

GEL  
GRUPO DE ESTUDOS LINGUÍSTICOS  
DO ESTADO DE SÃO PAULO

ESTUDOS LINGUÍSTICOS  
v. 42 n. 1

DESCRIÇÃO E ANÁLISE LINGUÍSTICA

ESTUDOS LINGUÍSTICOS, São Paulo, 42 (1): p. 1-575, jan-abr 2013

REVISTA ESTUDOS LINGÜÍSTICOS  
GRUPO DE ESTUDOS LINGÜÍSTICOS DO ESTADO DE SÃO PAULO GEL  
Universidade Federal de São Carlos  
Departamento de Letras  
Rodovia Washington Luiz, km 235  
CEP 13565-905 - São Carlos - SP - Brasil  
<http://www.gel.org.br/estudoslinguisticos/>  
estudoslinguisticos@gel.org.br

**Comissão Editorial**

Claudia Zavaglia  
Gladis Massini-Cagliari  
Juanito Ornelas de Avelar  
Luciani Ester Tenani  
Manoel Mourivaldo Santiago Almeida  
Marco Antônio Domingues Sant'Anna  
Maximina M. Freire  
Olga Ferreira Coelho  
Vanice Maria Oliveira Sargentini

**Editor responsável**

Marcelo Módolo

**Revisão e normatização**

Adélia Maria Mariano da S. Ferreira  
Maria de Fátima de Almeida Baia  
Daniel Peres de Oliveira

**Revisão de língua estrangeira**

Maria de Fátima de Almeida Baia (inglês)

**Diagramação**

Editora Paulistana

**Conselho Editorial**

Aldir Santos de Paula (UFAL), Alessandra Del Re (UNESP), Alvaro Luiz Hattner (UNESP), Ana Ruth Moresco Miranda (UFPEL), Angel H. Corbera Mori (UNICAMP), Angélica Rodrigues (UFU), Anna Flora Brunelli (UNESP), Aparecida Negri Isquerdo (UFMS), Ataliba Teixeira de Castilho (UNICAMP), Carola Rapp (UFBA), Claudia Regina Castellanos Pfeiffer (UNICAMP), Claudio Aquati (UNESP), Cláudia Nívia Roncarati de Souza (UFF), Cleudemar Alves Fernandes (UFU), Cristiane Carneiro Capristano (UEM), Cristina Carneiro Rodrigues (UNESP), Cristina dos Santos Carvalho (UNEB), Edvania Gomes da Silva (UESB), Edwiges Maria Morato (UNICAMP), Erica Reviglio Iliovitz (UFRPE), Erotilde Goreti Pezatti (UNESP), Fabiana Cristina Komesu (UNESP), Fernanda Mussalim (UFU), Francisco Alves Filho (UFPI), Gladis Maria de Barcellos Almeida (UFSCAR), Gladis Massini-Cagliari (UNESP), Ivã Carlos Lopes (USP), João Bôscio Cabral dos Santos (UFU), Júlio César Rosa de Araújo (UFC), Leda Verdiani Tfouni (USP), Lígia Negri (UFPR), Luciani Ester Tenani (UNESP), Luiz Carlos Cagliari (UNESP), Maria da Conceição Fonseca Silva (UESB), Maria Helena de Moura Neves (UNESP/UPM), Maria Margarida Martins Salomão (UFJF), Marisa Corrêa Silva (UEM), Marize Mattos Dall Aglio Hattner (UNESP), Mauricio Mendonça Cardozo (UFPR), Márcia Maria Cançado Lima (UFMG), Mário Eduardo Viaro (USP), Mirian Hisae Yaegashi Zappone (UEM), Mônica Magalhães Cavalcante (UFC), Neusa Salim Miranda (UFJF), Norma Discini (USP), Pedro Luis Navarro Barbosa (UEM), Raquel Salek Fiad (UNICAMP), Renata Ciampone Mancini (UFF), Renata Coelho Marchezan (UNESP), Roberta Pires de Oliveira (UFSC), Roberto Gomes Camacho (UNESP), Ronaldo Teixeira Martins (UNIVAS), Rosane de Andrade Berlinck (UNESP), Sanderléia Roberta Longhin Thomazi (UNESP), Sandra Denise Gasparini Bastos (UNESP), Sebastião Carlos Leite Gonçalves (UNESP), Seung Hwa Lee (UFMG), Sheila Elias de Oliveira (UNICENTRO), Sonia Maria Lazzarini Cyrino (UNICAMP), Vânia Cristina Cassee Galvão (UFG), Vânia Maria Lescano Guerra (UFMS)

Publicação quadrimestral

Estudos Lingüísticos / Organizado pelo Grupo de Estudos Linguísticos  
do Estado de São Paulo v. 1 (1978). Campinas, SP: [s.n.], 1978

Publicada em meio eletrônico (CDROM) a partir de 2001.

Publicada em meio eletrônico (<http://www.gel.org.br/>) a partir de 2005.

Quadrimestral

ISSN 14130939

1. Linguística. 2. Linguística Aplicada 3. Literatura I. Grupo de Estudos  
Linguísticos do Estado de São Paulo.

## Reconhecimento

O presente volume da Revista *Estudos Linguísticos* contou com a colaboração dos pareceristas abaixo listados. Alguns são membros do Conselho Editorial e os demais atuaram como pareceristas *ad hoc*. A Comissão Editorial agradece o empenho de todos no sentido de tornar melhor nossa publicação.

Adail Ubirajara Sobral, Adriana Zavaglia, Adriane Teresinha Sartori, Alcides Cardoso dos Santos, Aldir Santos de Paula, Alessandra Del Re, Alvaro Luiz Hattnher, Ana Maria Costa de Araujo Lima, Angel H. Corbera Mori, Anna Flora Brunelli, Aparecida Negri Isquerdo, Ariani Di Felippo, Ataliba Teixeira de Castilho, Augusto Buchweitz, Beatriz Protti Christino, Carla Alexandra Ferreira, Carlos Alexandre V. Gonçalves, Carlos Piovezani, Carmi Ferraz Santos, Carola Rapp, Cilaine Alves Cunha, Claudia Maria Xatara, Claudia Regina Castellanos Pfeiffer, Claudia Zavaglia, Claudio Aquati, Cláudia Regina Brescancini, Cleudemar Alves Fernandes, Cloris Porto Torquato, Cristiane Carneiro Capristano, Cristina dos Santos Carvalho, Cristine Gorski Severo, Dilson Ferreira da Cruz Júnior, Dirceu Cleber Conde, Eduardo Penhavel, Edvaldo A. Bergamo, Edvania Gomes da Silva, Edwiges Maria Morato, Elaine Cristina Cintra, Elaine Cristina de Oliveira, Elias Alves de Andrade, Elizabeth Harkot-De-La-Taille, Elzimar Goettenauer de Marins Costa, Emerson de Pietri, Erica Lima, Erica Reviglio Iliovitz, Erotilde Goreti Pezatti, Ester Mirian Scarpa, Fabiana Cristina Komesu, Fabio Akcelrud Durão, Fábio César Montanheiro, Fernanda Correa Silveira Galli, Flaviane Romani Fernandes Svartman, Flávia Bezerra de Menezes Hirata Vale, Francisco Alves Filho, Frantome Bezerra Pacheco, Giovana Ferreira Gonçalves, Gisela Collischonn, Gisele Cássia de Sousa, Gladis Maria de Barcellos Almeida, Gladis Massini-Cagliari, Gláucia Vieira Cândido, Graziela Zanin Kronka, Isadora Valencise Gregolin, Ivã Carlos Lopes, João Bôscio Cabral dos Santos, José Borges Neto, José Sueli de Magalhães, Juliano Desiderato Antonio, Júlio César Rosa de Araújo, Larissa Cristina Berti, Lauro José Siqueira Baldini, Lilian Cristine Scherer, Lígia Negri, Lucia Rottava, Luciana Pereira da Silva, Luciana Salazar Salgado, Luciane Correa Ferreira, Luciani Ester Tenani, Luiz Carlos Cagliari, Luiz Carlos da Silva Schwindt, Luzia Aparecida Oliva dos Santos, Luzmara Curcino Ferreira, Marcello Modesto dos Santos, Marcelo Módolo, Maria Aparecida Lino Pauliukonis, Maria Célia Cortez Passetti, Maria Cristina de Moraes Taffarello, Maria da Conceição Fonseca-Silva, Maria Eduarda Giering, Maria Ester Vieira de Sousa, Maria José Bocorny Finatto, Maria José Cardoso Lemos, Maria Margarida Martins Salomão, Mariângela de Araújo, Marília Blundi Onofre, Mario Luiz Frungillo, Marisa Corrêa Silva, Marize Mattos Dall'Aglio Hattnher, Mauricio Mendonça Cardozo, Mayumi Denise Senoi Ilari, Márcia Maria Cançado Lima, Monica Filomena Caron, Mônica Veloso Borges, Nelson Viana, Norma Discini, Pedro Luis Navarro Barbosa, Raquel Meister Ko Freitag, Raquel Salek Fiad, Renata Coelho Marchezan, Roberta Pires de Oliveira, Roberto Gomes Camacho, Roberto Leiser Baronas, Ronald Taveira da Cruz, Ronaldo Teixeira Martins, Rosana do Carmo Novaes Pinto, Rosana Mara Koerner, Rosane de Andrade Berlinck, Rosane Rocha Pessoa, Sanderléia Roberta Longhin-Thomazi, Sandra Denise Gasparini Bastos, Sebastião Carlos Leite Gonçalves, Sebastião Elias Milani, Sheila Elias de Oliveira, Simone Azevedo Floripi, Sonia Maria Lazzarini Cyrino, Soraya Maria Romano Pacífico, Sônia Bastos Borba Costa, Taísa Peres de Oliveira, Tony Berber Sardinha, Valéria Faria Cardoso, Vanice Maria Oliveira Sargentini, Vânia Cristina Casseb Galvão, Vânia Maria Lescano Guerra, Wilmar da Rocha D'Angelis, Wilton José Marques.

## SUMÁRIO

APRESENTAÇÃO .....	8
FONÉTICA	
A entoação em falas teatrais: uma análise da raiva e da fala neutra <i>Aline Mara de Oliveira Vassoler e Marcus Vinícius Moreira Martins</i> .....	9
Existem línguas de ritmo silábico? <i>Luiz Carlos Cagliari</i> .....	19
Análise comparativa da entoação do português brasileiro e do inglês norte-americano no filme <i>Shrek</i> <i>Maíra Sueco Maegava Córdula</i> .....	33
Análise entoacional: uma comparação entre o modelo de Halliday (1970) e Pierrehumbert (1980) <i>Mariane Carvalho</i> .....	45
FONOLOGIA	
Adaptação de empréstimos de itens lexicais no papiamentu moderno <i>Manuele Bandeira, Shirley Freitas e Ana Livia Agostinho</i> .....	61
Ausência de assimilação de vozeamento no português como L2 por anglófonos - uma análise via Teoria da Otimidade <i>Miley Antonia Almeida Guimarães</i> .....	72
GRAMÁTICA FUNCIONAL	
O papel da modalização no jogo discursivo: um estudo da construção <i>pode ser(...), mas...</i> <i>Cibele Naidhig de Souza Carrascossi</i> .....	87
Variação e mudança linguísticas e preservação da informação <i>Deize Crespim Pereira</i> .....	100
Clivagem e construções similares sob a perspectiva da Gramática Discursivo-Funcional <i>Erotilde Goretti Pezatti</i> .....	112
É o seguinte: limites entre elaboração e projeção <i>Fernanda Abreu e Silva Alencar</i> .....	127

Alguns meios de expressão linguística das relações implícitas que se estabelecem entre orações e entre partes do texto <i>Juliano Desiderato Antonio</i> .....	137
Gramaticalização de construções como tendência atual dos estudos funcionalistas <i>Mariangela Rios de Oliveira</i> .....	148
As expressões <i>aí está</i> e <i>lá vai</i> como instanciações da construção LocV <i>Rossana Alves Rocha</i> .....	163
<b>GRAMÁTICA GERATIVA</b>	
Paralelos estruturais entre a quantificação universal e as orações adverbiais em Karitiana <i>Luciana Racanello Storto</i> .....	174
Uma análise teórico-empírica de predicação secundária <i>Marcos Barbosa Carreira e Andrea Knöpfle</i> .....	182
<b>HISTORIOGRAFIA LINGUÍSTICA</b>	
A representação do objeto norma na obra <i>Sociolinguística, os níveis da fala</i> , de Dino Preti <i>Gil Negreiros</i> .....	195
<b>LEXICOGRAFIA E LEXICOLOGIA</b>	
Vocabulário Ortográfico Comum (VOC): constituição de uma base lexical para a língua portuguesa <i>Gladis Maria de Barcellos Almeida, José Pedro Ferreira, Margarita Correia e Gilvan Müller de Oliveira</i> .....	204
Estudo lexical dos nomes de sintomas e de doenças nos séculos XVII e XVIII: comparação entre o português e o francês <i>Mariana Giacomini Botta</i> .....	216
Sobre usos do comitativo em línguas africanas e a análise do ponto de referência <i>Paulo Jeferson Pilar Araújo e Dayane Cristina Pal</i> .....	230
As cores da fauna e da flora: um dicionário especial composto por cromônimos <i>Sabrina de Cássia Martins e Claudia Zavaglia</i> .....	245
<b>LÍNGUAS INDÍGENAS E AFRICANAS</b>	
Proposta de sistema fonológico para a língua dos Cayapó do Sul aldeados em São José de Mossâmedes <i>Eduardo Alves Vasconcelos</i> .....	257

Descrição e análise preliminar sobre a categoria verbal no dialeto indígena Pykobjê-Gavião (Timbira)	
<i>Talita Rodrigues da Silva</i> .....	270

## LINGUÍSTICA DE CORPUS

O léxico e seu uso em textos de alunos de língua estrangeira: um estudo descritivo pautado na Linguística de <i>Corpus</i> como subsídio para o ensino	
<i>Celso Fernandes Rocha</i> .....	284
A tradução de obras de Patrícia Melo, um estudo à luz da Linguística de <i>Corpus</i>	
<i>Elisangela Fernandes Martins</i> .....	301
Instrumentos e atividades realizadas para a coleta de um <i>corpus</i> de aprendizes em língua inglesa para integrar o Br-ICLE	
<i>Paula Tavares Pinto Paiva</i> .....	312

## LINGUÍSTICA HISTÓRICA

A (não) realização da vogal epentética em Português Arcaico à luz derivacional	
<i>Ana Carolina Freitas Gentil Almeida Cangemi</i> .....	323
Um estudo descritivo-comparativo sobre a colocação pronominal em jornais de São Paulo e de Rio Claro	
<i>Caroline Carnielli Biazolli</i> .....	338
O comportamento prosódico dos monossílabos no Português arcaico por meio de <i>Cantigas de Santa Maria</i>	
<i>Daniel Soares da Costa</i> .....	354
Percurso diacrônico de completivas introduzidas por <i>como</i> no português	
<i>Gisele Cássia de Sousa</i> .....	366
A realização do sujeito pela elite paulistana do início do séc. XX	
<i>Hélcio Batista Pereira</i> .....	376
Um enfoque no domínio da junção: a gramaticalização de <i>mesmo assim</i>	
<i>Lúcia Regiane Lopes-Damasio</i> .....	385
As fricativas sibilantes nas <i>Cantigas de Santa Maria</i>	
<i>Mariana Moretto Gementi</i> .....	401
Análise morfológica de elementos lexicais em textos jornalísticos do português brasileiro	
<i>Teresa Leal Gonçalves Pereira</i> .....	412
A constituição e desconstituição dos termos <i>linguagem, língua e fala</i> de Saussure	
<i>Thayanne Raísa Silva e Lima</i> .....	421

## MORFOLOGIA

- O processo de aquisição de verbos irregulares no português brasileiro  
*Aline Garcia Rodero Takahira* ..... 430
- Trajectoria da constelação sufixal: *-ismo, -ista, -ístico(a) e -ística*  
- desde a origem grega às línguas modernas  
*Nilsa Areán-Garcia* ..... 442

## SEMÂNTICA

- Anáforas associativas baseadas em relações eventivas  
*Alex de Britto Rodrigues* ..... 455
- A integração entre affordances e restrições gramaticais  
no processo de compreensão de sentenças  
*Paulo Henrique Duque* ..... 470

## SINTAXE

- Condições de opcionalidade do objeto direto no português do Brasil  
*Maria Madalena Loredó Neta* ..... 486
- Construções de *small clauses* complexas em Português Brasileiro  
*Patrícia Rodrigues e Maria José Foltran* ..... 497
- A subversão da norma e a produção de sentidos  
nos deslocamentos enunciativos  
*Stéfano Grizzo Onofre* ..... 512

## SOCIOLINGUÍSTICA E DIALETOLOGIA

- Crenças e atitudes linguísticas no Sudoeste do Paraná: tendências de  
reação frente às diferentes línguas e etnias  
*Clarice Cristina Corbari e Aparecida Feola Sella* ..... 526
- V-ra* no português: uma análise diacrônica  
*Kellen Cozine Martins e Maria da Conceição de Paiva* ..... 540
- Linguística 0800: uma ação de divulgação científica  
*Ricardo Joseh Lima* ..... 553
- O "r caipira" em Mato Grosso do Sul - estudo baseado em dados do *ALMS*,  
*Atlas Linguístico do Mato Grosso do Sul*  
*Vandersí Sant'Anna Castro* ..... 566

## **Apresentação do vol. 42 (2013)**

A presente edição da Revista Estudos Linguísticos consolida formulação proposta nos volumes precedentes, assegurando assim a continuidade do projeto de uma Revista ampla com grande qualidade de seus artigos. Para o presente volume, foram submetidos 163 artigos, dos quais 111 tiveram sua publicação aprovada pelos pareceristas. Todos os artigos são provenientes de comunicações apresentadas durante o 60º Seminário do GEL (2012), realizado na Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. Além disso, são publicados cinco artigos originados de conferências e intervenções em mesas redondas daquela edição do Seminário do GEL, completando assim um total de 116 trabalhos. Os artigos estão distribuídos nos três números que compõem o presente volume, que correspondem aos três eixos temáticos definidos nos volumes precedentes, a saber, “Descrição e Análise Linguística”; “Linguística: Interfaces” e “Análise do Texto e do Discurso”. A Comissão Editorial gostaria de manifestar seu agradecimento aos autores e aos pareceristas, que contribuíram para que esta publicação fosse possível. Nesse sentido, uma página de reconhecimento aos nossos pareceristas tem sido publicada, nomeando todos os que doaram seu tempo e esforço para que a avaliação dos artigos do presente volume fosse a mais criteriosa possível. Os trabalhos publicados refletem a grande diversidade das pesquisas produzidas nos domínios da linguagem, não somente no Estado de São Paulo, como em todo o território brasileiro.

Marcelo Módolo  
Presidente da Comissão Editorial

# A entoação em falas teatrais: uma análise da raiva e da fala neutra

(The intonation in play-acted vocal emotions: analysis of anger and neutral speech)

Aline Mara De Oliveira Vassoler<sup>1</sup>, Marcus Vinícius Moreira Martins<sup>2</sup>

<sup>1</sup>Departamento de Estudos Linguísticos (Oralidade e Letramento) – Universidade Estadual Paulista (Unesp – Rio Preto)

<sup>2</sup>Departamento de Letras Clássicas e Vernáculas da Faculdade de Filosofia, Letras e Ciências Humanas e Instituto de Matemática e Estatística – Universidade de São Paulo (USP)

alineoliveiravassoler@gmail.com, marcusmartins@usp.br

**Abstract:** In order to investigate emotions in speech, specifically anger, three professional actresses (20 years in the profession) participated in the survey. The play-acted speeches were captured in the laboratory of the Studio Multimedia Electronic Computer Center of the University of São Paulo (CCE-USP), using Sound Forge 9 audio software. The software used to obtain the curve intonational average of three repetitions of anger and neutral speech was the routine ExProsodia. The average of  $F_0$  mean obtained for rage was 197 Hz and for neutral speech was 156 Hz. Anger behaves in high frequency bands, while the  $F_0$  of neutral speech is located in medium frequency bands. However, the correlation was 0.88 for both conditions, showing that the linear range is nearly identical.

**Keywords:** emotion; intonation; experimental phonetics; acoustical analysis.

**Resumo:** Para investigar as emoções na fala, especificamente a raiva, três atrizes profissionais (20 anos de profissão) participaram da pesquisa. A captura das vozes foi realizada no laboratório Estúdio Multimeios do Centro de Computação Eletrônica da Universidade de São Paulo (CCE-USP), captadas pelo *software* de edição de som *Sound Forge 9*. O programa utilizado para obtermos a curva entoacional média das três repetições da raiva e da fala neutra foi a rotina *ExProsodia*. A média das médias de  $F_0$  obtida para a raiva foi de 197 Hz e para a fala neutra, 156 Hz. A raiva se comporta nas faixas de altas frequências, enquanto a  $F_0$  da fala neutra está localizada nas faixas de médias frequências. No entanto, a correlação foi de 0,88 para ambas as situações, mostrando que a variação linear é quase idêntica.

**Palavras-chave:** emoção; entoação; fonética experimental; análise acústica.

## Introdução

As emoções humanas são transmitidas por meio das expressões faciais, dos gestos corporais e pela fala do indivíduo. É evidente que a compreensão da fala depende diretamente de informações sintáticas e semânticas. No entanto, pistas prosódicas também carregam conteúdos fundamentais para a compreensão da mensagem. Essas informações permitem que o ouvinte, não só codifique e decodifique toda a mensagem falada, como também compreenda importantes funções linguísticas como a (i) personalidade do indivíduo, (ii) a idade do falante, (iii) o sexo/gênero e (iv) a condição física e emocional do falante (MOZZICONACCI, 2002; MOZZICONACCI; HERMES, 1999). Os elementos paralinguísticos e/ou não verbais, como as expressões faciais, os gestos, os sons (como os suspiros), também estão inerentes à compreensão da mensagem (MOZZICONACCI,

2001). Os primeiros estudos envolvendo as emoções (DARWIN, 1872), já consideravam as mudanças vocais como parâmetro de análise das emoções, no entanto, o foco principal desses estudos, eram as expressões faciais das diferentes emoções.

A partir do século XIX, as pesquisas sobre a fala se intensificaram e as relações entre as emoções e as mudanças vocais passaram a ser investigadas. Os estudos deste período, envolvendo o sinal acústico, comprovaram as primeiras inferências impressionísticas sobre as relações entre a voz e as emoções. Essas pesquisas demonstraram que as emoções geram mudanças no padrão respiratório, na fonação e na articulação do aparelho fonador do indivíduo, alterando, assim, os parâmetros do sinal acústico (SCHERER, 1989; WILLIAMS; STEVENS, 1972; BANSE; SCHERER, 1996). Dessa maneira, diversos experimentos têm confirmado a possibilidade de diferenciar as emoções por meio de parâmetros acústicos (MURRAY; ARNOTT, 1993, VASSOLER, 2011; OLIVEIRA; MEDEIROS, 2010).

O objetivo deste trabalho é contribuir com este corpo de pesquisa, analisando a emoção da raiva por meio de análises de falas interpretadas por atrizes e pela fala neutra (melhores definidas na seção a seguir). Pretendemos, pois, entender se há padrões entoacionais nas curvas entoadas com as emoções acima apresentada. Nossa proposta, é analisar as curvas entoacionais da fala emotiva da raiva comparando-a com a fala neutra, utilizando-se do modelo teórico metodológico desenvolvido por Ferreira Netto (2006, 2008), bem como realizar uma análise automática usando a rotina ExProsodia.

## **A Entoação**

O fenômeno da entoação passou a ser melhor compreendido em suas funções linguísticas e paralinguísticas a partir de meados de década de 70. Neste período surge um grande número de publicações em torno do tema, tais como as teses de Libermann (1975), Maeda (1976) e Pierrehumbert (1980) e os trabalhos do IPO publicados por t'Hart (1967), t'Hart e Cohen (1973), t'Hart (1981) e reunidos em Collier, t'Hart e Cohen (1990).

Esta geração de pesquisadores preocupou-se, principalmente, em definir e caracterizar linguisticamente o fenômeno da entoação. Para este fim, assumiram um pressuposto teórico-metodológico inicial: a entoação deve ser compreendida, apenas como a variação de  $F_0$ . Tal definição, embora operacional, incluía ainda os tons lexicais, como aqueles que ocorrem nas línguas da família *bantu* ou no chinês, por exemplo. Por esta razão, uma segunda definição teve que ser assumida: a entoação é pós-lexical, não sendo intrínseca aos itens do léxico. Ainda que se note, em alguns casos, tons cristalizados em itens lexicais, como no marcador conversacional “né”, ao qual se atribui uma curva ascendente. A terceira definição ainda nos diz que a entoação transmite informações complementares como: “tipo de sentença ou ato de fala, ou foco e estrutura da informação” (LADD, 1996, p. 7). Ladd (1996) ainda completa “por esta definição, entoação exclui características como acento tonal e tom, os quais são determinados pelo léxico e que servem para distinguir uma palavra da outra.”.

A terceira definição assume que pode haver dois tipos de informações veiculadas pela entoação, sendo uma linguística e outra paralinguística. A informação linguística, diz respeito a aspectos especificamente linguísticos, como a asserção ou a interrogação,

ao passo que a informação paralinguística transmitiria informações suplementares que contribuem para a compreensão da linguagem de modo geral (LADD, 1996). A hipótese central é a de que a primeira (informação linguística) pode ser estruturada em termos de uma teoria ou metodologia que se preocupe unicamente com questões associadas à entoação e prosódia, como proposto pelo programa ToBI (SILVERMAN et al., 1992; JUN, 2006). Por outro lado, a informação paralinguística não pode ser organizada ou explicada em termos de uma teoria linguística e demandaria outros modelos de análise.

Pike (1945) ao discutir a cita diferença, propôs que haveria dois canais de informação sobrepostos. Um primeiro propriamente linguístico e um segundo corpóreo, por assim dizer. O primeiro está diretamente associado ao código linguístico, isto é, aos segmentos, a fonologia, sintaxe, entre outros. Enquanto o segundo está ligado aos gestos corporais, a face do interlocutor e outros aspectos corporais. Desta maneira, é a associação entre os dois canais que fornece a informação paralinguística. Como apresentado em Pike (1945): um insulto falado por uma face risonha, pode ser compreendido como um elogio, por exemplo.

Ladd (1996) aponta que embora estes aspectos tenham influência no que é entendido pelo falante, eles não alteram a forma fonológica do que é dito. Nosso trabalho se sustenta nesta afirmação, pois é a partir dela que podemos comparar uma mesma sentença, a qual é entoada de diversas maneiras. A forma fonológica nos garante que teremos uma mesma representação mental, ou alvo tonal (XU; WANG, 2001), o qual será realizado pelo falante. Embora, nesta realização possamos ter valores de  $F_0$  diferentes, seja intra-sujeitos, seja entre-sujeitos.

## Prosódia e emoção

Vaissière (2004) e Hirst e DiCristo (1998) discutem, em seus respectivos trabalhos, o fato de não haver uma distinção clara entre entoação e prosódia. Por esta razão, os autores assumem duas definições de prosódia, uma *estreita* e outra *ampla*. Na abordagem estreita, os fenômenos de  $F_0$  são tomados como base para a análise, seja ela linguística ou não, ao passo que na abordagem ampla outros fenômenos como a intensidade, a qualidade vocal, o ritmo, acento, entre outros, podem ser analisados. Os estudos linguísticos tem se pautado por uma análise restrita, isto é, os demais aspectos prosódicos não são analisados em conjunto, mas, sim, como parte do processo de produção (PIERREHUMBERT, 1980; T'HART; COLLIER; COHEN, 1990; FERREIRA NETTO, 2006, 2008), enquanto os estudos sobre as emoções tem se voltado para uma abordagem ampla, como veremos a seguir.

Os principais sinais acústicos analisados para a expressão vocal das emoções são: (i) o contorno da frequência fundamental; (ii) a variação de  $F_0$  e (iii) intensidade vocal. As principais técnicas de análise utilizadas são: (i) o cálculo da distribuição de energia no espectro de frequência (especialmente a energia que envolve as regiões de alta e de baixa frequência e que afeta a percepção da qualidade de voz ou do timbre); (ii) a localização dos formantes (F1 e F2 relacionados com a percepção da articulação) e (iii) variedade temporal dos fenômenos, incluindo durações e pausas (BANSE; SCHERER, 1996; MURRAY; ARNOT, 1993).

No entanto, Scherer et al. (1984) e Murray e Arnott (1993) comprovaram que a frequência fundamental  $F_0$  parece ser o parâmetro acústico mais importante na identificação

das emoções na fala. A entonação da fala é representada acusticamente pela variação de  $F_0$  (LADEFOGED, 1974), por isso, pode ser utilizada na diferenciação das emoções (WILLIAMS; STEVENS, 1972). A maioria das pesquisas que estudam os efeitos vocais das emoções são adeptas da teoria discreta das emoções, que estabelece entre 9 e 14 emoções como as emoções básicas ou fundamentais. Essas emoções caracterizam-se por apresentarem padrões de respostas fisiológicas muito específicas, tanto na expressão facial, quanto na vocal. Assim, as emoções consideradas básicas são a alegria, a tristeza, a raiva, o medo e a surpresa (SCHERER, 2003; LAUKKA, 2004).

Na literatura não se encontra homogeneidade nos achados acústicos para a raiva, salvo a intensidade vocal que prevalece forte. Dessa forma, optou-se, nesse estudo, por analisar a entonação da fala com raiva comparada a fala neutra. Williams e Stevens (1972) encontraram, na raiva, contornos de  $F_0$  superior às demais emoções. Laukka (2004) e Stibbard (2001) também encontraram altos valores de  $F_0$  com grande variação a longo termo e contorno entonacional crescente. Em contraposição, alguns estudos encontraram frequência fundamental baixa (SOBIN; ALPERT, 1999), o que mostra incongruências entre os resultados. Como Mozziconacci (2002) sugere que as emoções sejam comparadas com a fala neutra do próprio sujeito, a fim de possibilitar comparações dentro e entre estudos. Nesse estudo, foi comparada a fala neutra com a fala com raiva. Assumimos a hipótese de que a voz neutra possui valores de  $F_0$  sempre mais baixos, ou em um registro mais baixo do que aquelas vozes entoadas com raiva.

## Metodologia

Participaram da pesquisa três atrizes profissionais com idades entre 40 e 50 anos, com média de 20 anos de profissão cada uma. A coleta dos dados se deu no laboratório do Estúdio Múltiplos do Centro de Computação Eletrônica da Universidade de São Paulo (CCE-USP). As atrizes foram instruídas a ler um texto informativo com cerca de 126 palavras, cujo tema era imunologia (VAZ, 1983), de modo que o léxico não viesse a interferir na interpretação. Cada uma das atrizes leu o texto de maneira atuada seguindo quatro emoções: raiva, alegria, medo e tristeza. Foi ainda previsto um grupo de controle chamado “fala neutra” em que as atrizes liam o texto sem inflexões emotivas. Dessa maneira, obteve-se um total de 75 gravações (4 emoções + 1 controle x 3 atrizes x 5 repetições). As gravações foram captadas com um mesa de som *Mackie CFX16*, com uma taxa de amostragem de 44KHz e uma resolução de 32 bit. Para o presente trabalho foi selecionada a emoção “raiva”, a qual será comparada com o grupo de controle “fala neutra”. Foi realizada a segmentação do texto em unidades menores (frases), selecionando uma frase para a análise. A segmentação linguística foi feita com base em Nespor e Vogel (1986) e descrita em pormenores em Vassoler (2011).

Foram usadas a primeira, a terceira e a quinta repetição de cada sujeito, obtendo-se um total de 18 tokens para análise (3 repetições x 2 emoções x 3 sujeitos). A unidade analisada foi:

- (1) *as células do sangue que fabricam anticorpos são individualizadas.*

A análise das variações de  $F_0$  foram feitas usando a rotina de análise da prosódia ExProsodia (FERREIRA NETTO, 2010). Nessa análise, foi usado o método de detecção automática das sílabas, chamado de método *c.p.s.* (candidato a pico silábico) pela média.

O método mapeia valores de frequência fundamental e intensidade a cada 5 milissegundos, a fim de determinar se há ou não uma unidade entoacional naquele ponto. Para se obter uma unidade entoacional os valores de frequência e intensidade tem de se manter dentro dos limites escolhidos pelo analista e maiores que zero por mais de 20 milissegundos. Sendo satisfeitos todos estes parâmetros o programa acusará que naquele ponto há uma unidade entoacional. A consistência dos valores de  $F_0$  é calculada usando-se testes de hipótese, como um *teste-z*, a fim de que não haja variações muito grandes no interior de uma unidade. Neste método, portanto, temos dois tipos de análise: a primeira mapeia os valores de  $F_0$  (a unidades entoacionais), em busca de valores válidos, ao passo que a segunda analisa a constância de tais dados. O parâmetro de constância é sempre a média, valores muito distantes da média são excluídos.

No momento em que houver inconstâncias ou a frequência ou a intensidade igualam-se à zero ou ultrapassam os valores escolhidos pelo analista, a rotina compreenderá como o fim de uma unidade entoacional e recomeçará novamente o processo de mapeamento. Cada uma das orações foi analisada pelo método acima exposto. Os valores de média das unidades entoacionais das gravações foram calculados por meio da seguinte fórmula (1),

Onde:

$\tau$  = tom médio;

$zx$  = unidade entoacional em mapeamento;

$zx + 1$  = unidade entoacional a ser mapeada.

A fórmula acima exposta nos diz que a média será obtida somando-se as unidades entoacionais válidas e tirando-se médias acumuladas até que a  $zx + 1 = zx$ , ou seja, haverá uma relação temporal intrínseca entre os momentos  $zx$  até que a última unidade mapeada pelo processo.

## Análise

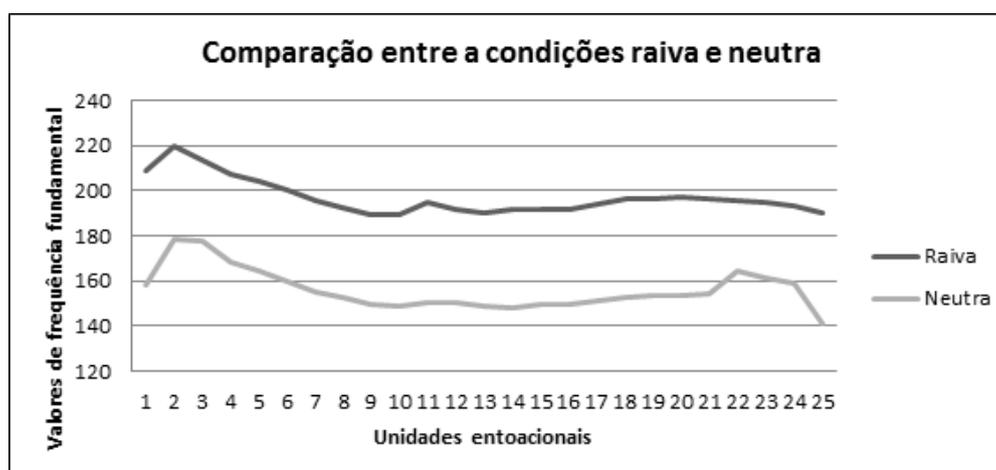
A partir dos valores obtidos foi feita uma média geral por unidade entoacional. Obtivemos, ao fim da análise, um total de 25 unidades para cada condição (raiva/fala neutra) em cada uma das repetições. Aquelas frases que possuíam menos do que 25 unidades entoacionais passaram por um processo de tratamento dos dados. Para este fim utilizamos um método de reposição de valores por amostragem, de modo que todos os vetores de dados pudessem ser equiparados, ou seja, ter a mesma quantidade de unidades entoacionais. Os valores obtidos em cada uma das unidades entoacionais foram somados aos seus pares (todas as primeiras unidade entoacionais, por exemplo, depois todas as segundas e assim por diante) e divididos por nove, o número de tokens por emoção. A partir desse procedimento, obteve-se uma média geral feita a partir da média das frequências das unidades entoacionais. Os valores apresentados no Quadro 1 resumem nossas estatísticas descritivas, as médias gerais foram de 197,54 Hz para a raiva, enquanto para a fala neutra obteve-se uma média de 156,13 Hz. Enquanto os valores de desvio-padrão geral foram de

7,46 para a primeira condição e 9,02 para a segunda. O coeficiente de variação intrínseco geral a cada condição foi de 3,78% e 5,78%, respectivamente. Os dados estão resumidos na tabela abaixo:

**Tabela 1: Resumo dos valores das estatísticas descritivas**

	<b>Raiva</b>	<b>Neutra</b>
<b>Média</b>	197,54	156,13
<b>Desvio-padrão</b>	7,46	9,02
<b>Coef. de Variação</b>	3,78%	5,78%

Os dois vetores de médias das unidades entoacionais de cada condição foram comparados por um teste de correlação de Pearson. O teste retornou um  $\rho=0,88$ , tal valor demonstra que ambas tem uma correlação muito forte, indicando que a despeito dos valores que as frequências fundamentais vieram a assumir em cada unidade entoacional a direção de ambos os vetores manteve-se a mesma. Foram também calculadas as correlações entre as condições entre sujeitos, tomando como base os valores de  $F_0$ , de modo que todas foram maiores que 0,8 ( $p>0,8$ ). Este valor indica que não se trata de uma idiossincrasia causada por conta das médias tomadas no processo de detecção automático. A Figura 1 mostra uma comparação entre as duas condições, os valores plotados no eixo y são as médias silábicas gerais, enquanto no eixo x se encontram as unidades entoacionais.



**Figura 1: Comparação entre a fala neutra e a raiva, sendo que os valores plotados no eixo y correspondem as médias silábicas gerais, enquanto no eixo x se encontram as unidades entoacionais**

Foram também calculados os coeficientes de variação para cada uma das condições. No caso da condição *raiva* obteve-se um coeficiente de variação de 3,78%, enquanto para a condição neutra houve um total de 5,78%. O que se pode notar é que a distribuição dos dados em ambos os casos é heterogênea, sugerindo que os dados estão alocados em intervalos bastante discriminados. Foi também executado um *teste-t* pareado, usando como fatores as condições raiva e neutra e como valores de análise as médias das unidades entoacionais, o primeiro teste foi executado usando-se um *bootstrapping*<sup>1</sup> com  $n=10000$  e

<sup>1</sup> A técnica de *bootstrapping* consiste em reorganizar aleatoriamente, com base em um método de amostragem aleatória simples, um certo conjunto de dados. Em tese, o conjunto re-amostrado mantém as mesmas características de distribuição do conjunto de dados iniciais, uma vez que eles são amostras aleatórias de uma dada população.

$\alpha=5\%$  e um segundo sem *bootstrapping*, com  $n=25$  e  $\alpha=5\%$ . Ambos os testes retornaram  $p<0,001$ , demonstrando que há, de fato, um efeito entre as duas amostras e que, provavelmente, ambas são diferentes.

## Conclusão

Como se pode observar a condição “raiva” se comporta nas faixas de altas frequências, enquanto a frequência fundamental da fala neutra está localizada nas faixas de médias frequências, corroborando a literatura pesquisada. No entanto, o coeficiente de variação da raiva está em torno de 3,7 e 5,7 para a fala neutra, o que mostra mais consistência nas repetições da raiva. O desvio padrão da fala neutra (9) também foi maior se comparado com a raiva (7,4). No entanto, a correlação foi de 0,88 para ambas as situações, mostrando que a variação linear é quase idêntica.

Estes números apontam que, salvo variações intrínsecas ao sujeito analisado, as emoções possuem padrões de contorno muito específicos e bastante regulares. A diferença entre elas está nos *pitch ranges*. As frases entoadas com raiva tendem a ter um registro mais agudo, ao passo que as frases entoadas de maneira neutra mantêm-se em um registro ligeiramente mais central. A hipótese explicativa, de um ponto de vista fisiológico, refere-se à ideia de que a raiva atuada exige do ator um nível de enrijecimento dos músculos intercostais e das cartilagens laringais diferentes do que daqueles da voz neutra. Este aumento de tônus causaria um aumento na pressão pulmonar e, por consequência, do fluxo de ar e pressão subglotal, alterando os valores de frequência fundamental (TITZE, I. R.; SCHMIDT, S. S.; TITZE, M. R., 1995). Contudo, sem as medidas fisiológicas, tais hipóteses são apenas explicativas, de modo que a verificação demandaria uma nova pesquisa.

Do ponto de vista linguístico, é possível afirmar que há uma forma entoacional padronizada para a sentença escolhida, de modo que é a implementação fonética desta forma o fator determinante de seu valor modal. A figura 1 e o valor de correlação de Pearson demonstram que há certa regularidade neste sentido. A forma entoacional segue os princípios de teorias como a de Ladd (1996), as quais hipotetizam tal forma como um objeto mental, sendo que a realização segue as características do sujeito, bem como as condições de implementação (estado emocional, contexto conversacional e etc.). Ressaltamos que não podemos determinar com exatidão qual tipo de elocução é entoada com raiva ou neutra apenas com os dados acústicos, uma vez que a implementação é o fator determinante, como acima citado. Por outro lado, podemos vislumbrar a hipótese de que frases iguais entoadas de maneiras variadas mantêm certo padrão nas suas variações de frequência fundamental.

## REFERÊNCIAS

- BANSE, R.; SCHERER, K. R. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, Genebra, v. 70, n. 3, p. 614-636, 1996.
- COLLIER, R.; ‘T HART, J. T.; COHEN, A. *A perceptual study of intonation*. Cambridge: Cambridge University Press, 1990.

DARWIN, C. R. *The expression of the emotions in man and animals*. London: John Murray, 1872.

FERREIRA NETTO, W. ExProsodia. *Revista da Propriedade Industrial – RPII*, Rio de Janeiro, v. 2038, n. 167, p. 167, 2010.

\_\_\_\_\_. Decomposição da entoação frasal em componentes estruturadoras e em componentes semântico-funcionais. In: *X Congresso Nacional de Fonética e Fonologia / IV Congresso Internacional de Fonética e Fonologia*. Niterói; UFF, 2008. p. 1-6.

\_\_\_\_\_. *Formação da prosódia da língua portuguesa*. 2006. Tese (Livre-Docência) – Universidade de São Paulo-USP, São Paulo.

HIRST, D.; DI CRISTO, A. A survey of intonation systems. In: HIRST, D.; DI CRISTO, A. (Ed.) *Intonation Systems*. Cambridge: Cambridge University Press, 1998.

JUN, S.-A. (Ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, 2006.

LADD, D. R. *Intonational phonology*. Cambridge: Cambridge University Press, 1996.

LADEFOGED, P. *Elements of acoustic phonetics*. Chicago & London: The University of Chicago Press, 1974.

LAUKKA, P. Vocal expression of emotion: Discrete-emotions and dimensional accounts. 2004. Tese (Doutorado). Uppsala, Sweden: Acta Universitatis Upsaliensis, 2004. v. 141. p. 1-80.

LIBERMANN, M. *The intonational system of English*. 1975. Tese (Doutorado em Linguística) - Massachusetts Institute of Technology, Cambridge, 1975.

MAEDA, S. *A Characterization of American English Intonation*. 1976. Tese (Doutorado em Engenharia Elétrica) – Massachusetts Institute of Technology – M.I.T., Cambridge.

MOZZICONACCI, S. J. L.; HERMES, D. J. Role of intonation patterns in conveying emotion in speech. *Proceedings of International Conference on Speech Prosody*. San Francisco: ISCA/SProSIG 1999. p. 2001-2004.

MOZZICONACCI, S. J. L. Prosody and Emotions. In: BEL, B.; MARLIEN, I. (Org.) *Proceedings of 1st International Conference on Speech Prosody*. Aix-en-Provence: ISCA/SProSIG, 2002. p. 1-9.

\_\_\_\_\_. Emotion and attitude conveyed in speech by means of prosody. In: *For the 2nd Workshop on Attitude, Personality and Emotions in User-Adapted Interaction*. Sonthofen, Germany, July, 2001, p. 1-10.

MURRAY, I. R.; ARNOTT, J. L.: Towards the Simulation of Emotion in Synthetic Speech: A Review of the Literature of Human Vocal Emotion. *Journal of Acoustic Society of America*, New York, v. 93, n. 2, p. 1097-1198, 1993.

- NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.
- OLIVEIRA, A. M. Aspectos acústicos da voz na representação teatral das emoções: um estudo preliminar. *Congresso Brasileiro de Fonoaudiologia*, 2008, Campos do Jordão, Sociedade Brasileira de Fonoaudiologia, p. 1312.
- OLIVEIRA, A. M.; MEDEIROS, B. R. Aspectos prosódicos de quatro emoções na voz falada. *VI Simpósio de Cognição e Artes Musicais*, Rio de Janeiro: ABCOGMUS, 2010. p. 292-300.
- PIKE, K. N. *The intonation of American English*. Michigan: Ann Arbor – University of Michigan Press, 1945.
- PIERREHUMBERT, J. *The Phonology of Phonetic of English Intonation*. 1980. Tese (Doutorado em Linguística) – Massachusetts Institute of Technology – M.I.T., Cambridge, 1980.
- SCHERER, K. R. Vocal communication of emotion: A review of research paradigms. *Speech Communication*, v. 40, n. 1-2, p. 227-256, 2003.
- \_\_\_\_\_. Vocal measurement of emotion. In: PLUTCHIK, R.; KELLERMAN, H. (Org.). *Emotion: Theory, research, and experience. The measurement of emotion*. v. 4. New York: Academic Press, 1989. p. 233-260.
- \_\_\_\_\_ et al. On the nature and function of emotion: a component process approach. In: SCHERER, K. R.; EKMAN, P. (Org.) *Approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum, 1984. p. 293-318.
- SILVERMAN, K. et al. ToBI: a standard for labeling English Prosody *Proceedings, Second International Conference on Spoken Language Processing*, Alberta: ISCA. 1992, 2, 867-870.
- SOBIN, C.; ALPERT, M. Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research*, New York, v. 28, n. 4, p. 347-365, 1999.
- STIBBARD, R. M. *Vocal Expression of Emotions in Non-laboratory Speech: An Investigation of the Reading/Leeds Emotion in Speech Project Annotation Data*. Unpublished PhD thesis. University of Reading, Reading, UK, 2001.
- T'HART, J. On the anatomy of intonation. *Língua*, Leiden, v. 19, n. 2, p. 177-19, 1967.
- \_\_\_\_\_. Differential sensitivity to pitch distance, particularly in speech. *Journal of Acoustical Society of the America*, New York, v. 69, n. 3, p. 811-821, 1981.
- T'HART, J.; COHEN, A. Intonation by rule: a perceptual quest. *Journal of Phonetics*, Bloomington, v. 1, n. 4. p. 309-327, 1973.
- TITZE, I. R.; SCHMIDT, S. S.; TITZE, M. R. Phonation threshold pressure in a physical model of the vocal fold muscosa. *Journal of Acoustical Society of the America*, New York, v. 97, n. 5, p. 3080-4, 1995.

VAISSIÈRE, J. Perception of intonation. In: PISONI, D. B.; REMEZ, R. E. (Ed.) *Handbook of Speech Perception*, Blackwell, 2004. p. 236-263.

VASSOLER, A. M. O. *O papel da frequência fundamental da fala na representação teatral das emoções*. 2011. Dissertação (Mestrado em Semiótica e Linguística Geral do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas) – Universidade de São Paulo – USP.

VAZ, N. M. Reconhecer a si próprio – Ideias para uma nova imunologia. *Ciência Hoje II*, v. 7, p. 32-8, 1983.

XU, Y.; WANG, Q. E. Pitch targets and their realization: Evidence from Mandarin *Chinese Speech Communication*, v. 33, n. 2, 319-337, 2001.

WILLIAMS, C. E.; STEVENS, K. N. Emotions and speech: Some acoustical correlates. *Journal of Acoustics Society of America*, New York, v. 52, p. 1238-1250, 1972.

# Existem línguas de ritmo silábico?

(Do syllable-timed languages exist?)

**Luiz Carlos Cagliari<sup>1</sup>**

Faculdade de Ciências e Letras – Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)

lccagliari@gmail.com

**Abstract:** The language typology of speech rhythm has been investigated by many researchers in different languages. However, they have not reached a unanimous conclusion. Part of the misunderstanding is due to the methodology they used. In addition, there is some theoretical confusion on the definition of a syllable timed language. This paper refers to the last question, but without discussing others people works and theories. Instead, it points out some essential questions that have not been taken into consideration in the discussion.

**Keywords:** Speech rhythm; typology; syllable-timed rhythm.

**Resumo:** A tipologia do ritmo das línguas tem sido investigada por muitos pesquisadores de várias línguas, porém, sem que eles tenham chegado a conclusões unânimes. Parte das divergências se deve às metodologias usadas. Há também muita confusão teórica a respeito do que seria uma língua de ritmo silábico. O presente trabalho contempla esse último aspecto, sem introduzir discussões com trabalhos e teorias desenvolvidas, apenas chamando a atenção para alguns pontos importantes, nem sempre levados em consideração nessa discussão.

**Palavras-chave:** Ritmo da fala; tipologia; língua de ritmo silábico.

## Introdução

Existe Língua de Ritmo Silábico? Nos últimos 30 anos, muitos foneticistas discutiram a questão do tipo de ritmo das línguas, a partir da proposta de Kenneth L. Pike (1945, p. 34-43) e de David Abercrombie (1967, p. 95-102; 1973). Esses autores propuseram que as línguas ou são do tipo acentual (Inglês) ou do tipo silábico (Francês). A grande polêmica começou com a seguinte afirmação de Abercrombie:

The rhythm of speech is a rhythm of these two systems of (chest) pulses: it is a product of the way they are combined in producing an air-stream for talking... Two different kinds of periodicity can result from their organization.... Either the stress-producing pulses or the syllable-producing pulses can be in isochronous sequence, and in the former case we have a stress-timed rhythm, in the latter a syllable-timed rhythm... The languages of the world fall into two classes, depending on whether the rhythm with which they are spoken is stress-timed or syllable-timed. (ABERCROMBIE, 1973, p. 7)<sup>2</sup>

1 Este trabalho foi financiado em parte pelo CNPq (Processo n. 303440/2009-0).

2 O ritmo da fala é o ritmo daqueles dois sistemas de pulsos (torácicos): é o produto do modo como eles são combinados, produzindo uma corrente de ar para a fala... Dois tipos diferentes de periodicidade podem resultar de sua organização... Ou os pulsos que produzem o acento ou os pulsos que produzem a sílaba podem estar em sequência isócrona e, no primeiro caso, temos um ritmo temporizado como acentual, no segundo caso, um ritmo temporizado como silábico... As línguas do mundo são de duas classes, dependendo de o ritmo com que elas são faladas ser temporizado pelo acento ou temporizado pela sílaba (ABERCROMBIE, 1973, p. 7) – A tradução é minha. Traduzi aqui *stress-timed* e *syllable-timed* como

Abercrombie (1967, p. 24, 96-99) definiu o ritmo a partir da corrente de ar da fala (processo aerodinâmico), mas as críticas à sua definição vieram de análises acústicas, medindo o tempo dos segmentos, como as taxas de elocução (*speech rate*) e outras unidades. Nos estudos acústicos, o que faltou, em geral, foi o parâmetro andamento, que é uma medida diferente da taxa de elocução, sem o qual nenhum ritmo pode ser medido corretamente. O grande volume de trabalhos acústicos, com medições físicas sobre enunciados de muitas línguas, chegou à conclusão de que

- 1) a dicotomia é precária ou mesmo falsa;
- 2) há tipos mistos de ritmo (acentual+silábico);
- 3) há outros padrões rítmicos (por exemplo, moras do Japonês);
- 4) critérios físicos de medição da fala, aplicados sobre a duração dos segmentos, controlados estatisticamente, geraram outros padrões rítmicos para as línguas.

Nessa vasta discussão, as tradicionais línguas de ritmo acentual se mantiveram mais próximas da caracterização antiga, ou seja, houve pouca contestação a respeito delas. Como grande parte dos trabalhos sobre ritmo vinham da poética em língua inglesa, o fato de o inglês ser de ritmo acentual não causava estranhamento.

Por outro lado, os linguistas começaram a levar em conta os processos fonológicos para justificar o fato de as línguas de ritmo acentual apresentarem um conjunto de fatos fonológicos não encontrados em outras línguas (CAGLIARI; ABAURRE, 1986). Essas outras línguas seriam, por oposição, línguas de ritmo silábico ou de outro tipo, a ser definido. Portanto, o grande problema sempre foi a interpretação do que seja uma língua de ritmo silábico. Enquanto para as línguas de ritmo acentual foram feitas muitas pesquisas e descrições prosódicas com alto grau de sofisticação (cf. HALLIDAY, 1973; PIERREHUMBERT, 1980), poucas investigações foram feitas sobre os demais tipos de língua, agrupados como sendo de ritmo silábico, em princípio, seguindo uma dicotomia, então, descaracterizada da proposta inicial de Abercrombie, citada acima (WENK; WIOLLAND, 1982).

Por outro lado, a mesma abordagem teórica metodológica e, sobretudo acústica, aplicada às línguas de ritmo acentual foi aplicada às línguas de ritmo silábico. Isso veio complicar ainda mais a dicotomia original. Alguns grupos de foneticistas ficaram com teorias que desenvolveram (O'DELL; NIEMINEN, 1999; BARBOSA, 2000; 2004; 2006), mesmo que seus resultados nem sempre satisfaçam a percepção e a intuição sistemática que os falantes têm de sua própria língua. Outros foneticistas abandonaram a discussão em si, procurando descrever o ritmo em função dos acentos (grade métrica), das pausas (concatenação da fala) e de unidades prosódicas fonológicas (SELKIRK, 1984; NESPOR; VOGEL, 1986), ou apenas das unidades entoacionais (PIERREHUMBERT, 1980). Falar de ritmo da fala, hoje, é um problema, devido aos muitos pressupostos teóricos envolvidos (REIS, 1995).

Diante do quadro apresentado acima, o objetivo do presente trabalho não é apresentar uma análise de uma ou mais línguas de ritmo silábico (não acentual), nem de criar uma teoria a mais, menos ainda, discutir os resultados obtidos nessas diferentes abordagens sobre

---

*temporizado pelo acento e temporizado pela sílaba* para inserir o valor semântico da palavra *timed*, que não é comumente traduzida.

o ritmo da fala, mas apenas voltar a velhas ideias que acabaram esquecidas no caminho. Em outras palavras, o objetivo do presente estudo é revisitar a questão das línguas de ritmo silábico, analisando suas características rítmicas e outros fenômenos prosódicos que interferem na definição do ritmo da fala, como a tonicidade, a duração, o andamento, a velocidade de fala e a concatenação. A relevância desta revisão é esclarecer alguns pontos que não ficaram muito claros, que foram ignorados ou mal interpretados, gerando parte da confusão que se estabeleceu na tipologia rítmica das línguas. Metodologicamente, apresenta-se um apanhado geral das principais afirmações sobre a natureza do ritmo da fala e de seus tipos.

A seguir, será apresentada uma caracterização do que se entende por língua de ritmo silábico no presente trabalho, incluindo os tipos de línguas de ritmo silábico, o mecanismo de produção do ritmo da fala, sua dependência do processo aerodinâmico, a percepção do ritmo na música e na fala, o andamento, as unidades rítmicas da fala, a estrutura do léxico, características acústicas e perceptuais do ritmo.<sup>3</sup>

## **Considerações gerais sobre o ritmo**

### **O que é o ritmo**

Antes de mais nada, é preciso definir o que é o ritmo em geral e o que é o ritmo da fala. Ritmo é uma medida de simetria, do ponto de vista físico. Do ponto de vista perceptual, o ritmo é uma expectativa de repetição: algo vai acontecer de novo, num certo tempo e espaço, com características vistas anteriormente. Em termos populares, ritmo é uma repetição, um *déjà vu*. Essa definição serve tanto para descrever fatos reais como serve de base para metáforas.

### **O ritmo da música**

O ritmo aparece de modo muito claro na música (COOPER; MEYER, 1960). Por exemplo, definimos o ritmo como bolero, samba, tango, valsa, etc. Contrapomos a música gregoriana à música compassada. As notas do canto gregoriano ou neumas (do latim *pneumatis* ou *neumatis*) procuram acompanhar a duração das sílabas do latim. O mesmo tipo de comportamento musical encontramos em outros tipos de cantochão. Nesse tipo de música, não há batidas rítmicas, mas o que comanda o ritmo é o andamento no desempenho das durações dos neumas. Esse tipo de música servia bem o latim e o grego antigos, que eram línguas com sílabas longas e breves. Com a perda da duração contrastiva do latim nas línguas românicas, e com o surgimento da saliência tônica em substituição à duração contrastiva, surgiu um novo tipo de música para o canto: a música compassada, ou seja, uma música que marca o compasso em medidas regulares de tempo. Dentro da medida de um compasso (2/4, 3/4, etc.), as notas cabíveis deveriam prever a duração estipulada pelo compasso. Por essa razão, as durações das notas foram definidas em termos do dobro da duração da nota de duração imediatamente inferior: a mínima vale 1 tempo; a semínima vale 1/2 tempo; a colcheia vale 1/4 tempo, a semibreve vale 2 tempos e a breve vale 4 tempos. Essa notação musical era conveniente para o latim e para as línguas românicas. Por um lado, contemplava as línguas que mantinham a oposição entre sílabas longas e breves e, por outro, introduziam um compasso com as batidas regularmente espaçadas para as línguas com oposição de

3 Neste trabalho não serão considerados os processos fonológicos ligados a processos rítmicos.

tonicidade.<sup>4</sup> Obviamente, regras existem e podem ser quebradas. Isso significa que nem toda música cantada precisa levar em conta essa relação de duração de modo absoluto. Porém, quando foge desse padrão, começam os estranhamentos. A música tocada tem sua vida própria, diferente da cantada (com relação à fala).

Portanto, o cantochão marca o ritmo pelo andamento e a música compassada, pelos compassos. Quando se muda o andamento, não se muda o ritmo necessariamente. Uma música pode ser tocada ou cantada mais velozmente ou mais vagarosamente, sem alterar a percepção do ritmo. Mas as notas ficam alteradas em suas durações acústicas. Por outro lado, quando se altera o ritmo (por exemplo, um samba vira rock ou tango...), a percepção do ritmo é logo percebida pelo ouvido e, muitas vezes, a música fica muito estranha para os ouvintes. Isso mostra que os ouvintes têm expectativas com relação ao ritmo das músicas que conhecem, ou seja, eles têm interiorizado em suas mentes representações rítmicas padronizadas. Fato semelhante ocorre com o ritmo da fala. O que diferencia uma nota musical das sílabas da fala é o fato de, na música, a altura melódica das notas se manter constante durante o período da realização da duração da nota no canto (ou emitida por instrumento musical), mas, na fala, não apenas as durações previstas para as sílabas costumam variar até um certo limite, como a própria melodia atribuída às sílabas varia durante a pronúncia da sílaba. Desse modo, a fala e o canto têm características muito diferentes de como a melodia e o ritmo são executados.

### **O mecanismo de produção da fala e o ritmo**

A fala tem um mecanismo próprio que envolve vários processos da ação da mente e do corpo (ABERCOMBRIE, 1967, p. 20-33; LAVER, 1994, p. 161-183; CAGLIARI, 2007a, p. 17-31). O ritmo da fala está diretamente ligado à duração das sílabas, assim como o ritmo da música está ligado às notas ou aos neumas. As sílabas são produzidas especificamente pelo processo aerodinâmico (CATFORD, 1977, p. 24-46), através dos pulsos torácicos (ABERCROMBIE, 1967, p. 24, 34-37). A fala usa a corrente de ar da respiração para gerar um som e modificá-lo através do processo de ressonância do aparelho fonador (FANT, 1968). A corrente de ar não só gera um som, mas gera também um ritmo. A corrente de ar é acionada pelo diafragma e pelos músculos intercostais (LAVER, 1994, p. 161-183). Na respiração normal tranquila, a variação de pressão da corrente de ar é do tipo senoidal, com períodos regulares de inspiração e de expiração. Todavia, esse modelo não serve para a fala, justamente por causa de como a fala processa o ritmo. Como o ritmo depende da sílaba, esta precisa ser produzida corretamente. Uma sílaba é uma unidade aerodinâmica de fala que apresenta uma duração e uma intensidade. Dependendo das palavras, que são formadas por sílabas, uma sílaba será mais longa ou mais breve, mais intensa ou menos intensa do ponto de vista aerodinâmico (MASSINI-CAGLIARI; CAGLIARI, 2000, p. 107-110). Essa variação tem como consequência acústica unidades (sílabas) de duração variável e de volume variável, além de carrear também uma variação melódica, gerada na glotes (cordas vocais) ou mesmo no processo de ressonância, quando a corrente de ar modificada para a fala atravessa o aparelho fonador.

Portanto, a corrente de ar gera dois padrões concomitantes de sílabas: sílabas longas ou breves e sílabas tônicas ou átonas. A tonicidade inicial, gerada pelo mecanismo

<sup>4</sup> Um fato semelhante: vamos encontrar nos poetas como Gonçalves Dias, que conseguia fazer versos isossilábicos e ao mesmo tempo com sílabas tônicas ocorrendo em intervalos regulares de duração: *Meu canto de morte / guerreiros ouvi / sou filho das selvas / nas selvas cresci*, etc.

aerodinâmico, costuma ser reforçada por outros fatores, como a altura melódica e o volume do som. Desse modo, todas as sílabas de todas as línguas terão uma duração e uma tonicidade inerentes, geradas pelo mecanismo aerodinâmico. Em outras palavras, qualquer sílaba de qualquer língua pode ser analisada como sendo longa ou breve e, ao mesmo tempo, sendo tônica ou átona. O fato de uma sílaba ser longa ou breve, tônica ou átona vai depender de uma comparação entre ela e as demais sílabas que ocorrem antes e depois. Ou seja, essas medidas são comparativas, relativas e contrastivas e não trazem valores fixos absolutos. Dependendo da velocidade de fala ou do andamento, uma sílaba pode ser longa em um contexto, mas do ponto de vista físico, sua duração pode ser breve, em outro contexto. Todavia, como a língua é um sistema que se mantém nos diferentes contextos de realização da fala, as durações relativas das sílabas no primeiro e no segundo contextos se mantêm constantes. Assim, a sílaba longa é vista como breve, no novo contexto, porque, nesse contexto, as sílabas breves do contexto anterior são vistas como muito mais breves, o que mantém nos dois contextos o mesmo valor proporcional da duração das sílabas e, conseqüentemente, do ritmo. O andamento não destrói o ritmo, apenas muda a velocidade com que o mesmo é realizado.

### Os tipos de ritmo da fala

As línguas evitam a monotonia da fala: não somos máquinas. Em princípio, uma língua poderia ter todas as sílabas com igual duração e igual tonicidade. Isso, porém, não acontece. A variação é uma característica essencial dos sistemas linguísticos, em todos os níveis de análise. Por essa razão, as línguas criam palavras com sílabas que variam de duração e de tonicidade. Compare as palavras apresentadas a seguir, nas quais colocamos a marca [l] para sílabas longas e tônicas e a marca [b] para sílabas breves e átonas:

- (1)    l   b   b   b   l   b   b   b   l   l   b   l   b   l   b  
lâm.pa.da,   ba.ta.ta,   u.ru.bu,   an.ti-in fla.cio.ná.rio

Sequências de sílabas desse tipo apenas transferem do léxico para a fala os padrões duracionais e tônicos das sílabas. Por outro lado, podemos manter tais padrões e mudar o andamento (ou velocidade da fala) de tal modo que as sílabas tônicas (em geral mais longas também) ocorram em intervalos regulares de tempo, ou seja, marcando “compassos” à moda da música. Esse tipo de composição das durações e da tonicidade, associadas à mudança de andamento, cria um novo padrão rítmico, pondo em saliência não mais apenas a oposição entre sílabas longas e breves, mas entre sílabas tônicas e átonas, alterando o andamento das sílabas dentro desses intervalos. Essa é a razão pela qual Abercrombie (1973, p. 7) afirma que o sistema linguístico pode ter apenas dois tipos de ritmo, caracterizando o que se convencionou chamar de línguas de ritmo silábico e línguas de ritmo acentual. Não há outra possibilidade de realização.

Porém, como dito anteriormente, os sistemas das línguas não só admitem, mas incentivam, a variação linguística, para não tornar a língua uma realização monótona. Como um fenômeno linguístico, o ritmo também pode variar, como sempre, dentro de certos limites, para não destruir o sistema da língua.

### Línguas de ritmo acentual

O rótulo de língua de ritmo acentual é antigo e, como dito anteriormente, vem da interpretação da métrica poética em língua inglesa (ABERCROMBIE, 1973). Nessa língua,

os poetas procuram fazer versos metrificados não com um mesmo número de sílabas (versos isossilábicos), mas com um mesmo número de acentos poéticos em cada verso. Esse tipo de metrificação é muito diferente do modelo italiano ou francês, nos quais os poetas fazem versos com igual número de sílabas, independentemente do número de acentos poéticos em cada verso. Mas os estudiosos antigos não chamaram essas outras línguas de línguas de ritmo silábico. Também não chamaram o grego antigo e o latim de línguas de ritmo silábico, embora fossem línguas que faziam versos de modo muito diferente do inglês, também não se preocupando com a ocorrência da tonicidade poética nos versos.

Os rótulos contrastivos de língua de ritmo acentual, por um lado, e de língua de ritmo silábico, para todas as outras, veio inicialmente com Pike (1945, p. 34-43), baseando-se numa ideia de Daniel Jones (1960[1917]), como ele mesmo reconhece. O ensinamento de Daniel Jones ficou com alguns linguistas britânicos e a ideia de Pike se espalhou pelo mundo (ROACH, 1982; DAUER, 1983; BERTINETTO, 1989). Depois de muitos trabalhos na área, o inglês foi a única língua que se manteve mais ou menos incólume, com o rótulo de língua de ritmo acentual, todas as demais línguas indicadas inicialmente como línguas de ritmo silábico passaram a ser questionadas e gerando novos tipos de línguas com relação ao ritmo.<sup>5</sup>

### **Línguas de ritmo silábico**

As metodologias usadas para a análise das línguas que não eram de ritmo acentual levaram a teorias contraditórias com relação à própria natureza do ritmo linguístico. Porque as línguas presumidamente de ritmo acentual podiam ser facilmente analisadas acusticamente, por exemplo, medindo os intervalos de tempo entre as sílabas tônicas poéticas, tal procedimento foi transferido para se estudar qualquer língua, não só nos versos poéticos, mas também na fala comum. Nesse momento, surgiram erros metodológicos e equívocos de interpretação do que seria, na verdade, uma língua de ritmo silábico.<sup>6</sup> Os pesquisadores esqueceram-se do processo aerodinâmico e ficaram apenas com as medidas acústicas e um tratamento estatístico dos dados acústicos. Isso levou o estudo do ritmo da fala ao caos.

Porém podemos ver as línguas de ritmo silábico com outros olhos. Em primeiro lugar, levando em conta o que foi dito anteriormente sobre o mecanismo de produção da fala, encontramos no léxico das línguas sílabas longas e breves, tônicas e átonas. Uma língua como o grego antigo e o latim tinha uma oposição fonológica entre sílabas longas e breves. Os tratadistas mostraram que uma sílaba longa valia por duas breves. O fato de uma sílaba ser longa ou ser breve dependia de sua natureza ser tal ou dependia do contexto, podendo uma sílaba breve ser longa, em certos casos, mas não o contrário, uma sílaba que era longa por natureza sempre ficava longa. Algumas regras indicavam contextualmente

<sup>5</sup> As referências bibliográficas com relação a esse problema são muito extensas e contraditórias e não cabe discuti-las no presente trabalho.

<sup>6</sup> O erro mais comum foi achar que as línguas de ritmo silábico teriam todas as sílabas com igual duração. Esse erro veio em decorrência da ideia de isocronia silábica atribuída por Abercrombie a línguas específicas, como um exemplo, e não se aplicava a todas as línguas consideradas de ritmo silábico. Ele conhecia bem como era o ritmo das línguas clássicas e jamais diria que, nessas línguas de ritmo silábico, todas as sílabas tinham a mesma duração. Abercrombie lecionava em Edimburgo (Escócia) e sabia muito bem que o inglês escocês era de ritmo diferente do inglês RP. Fui aluno de Abercrombie e era assim que ele ensinava. Lamentava que não houvesse estudos rítmicos sobre outras línguas de ritmo silábico, que não fossem o grego e o latim.

se uma sílaba era longa ou breve.<sup>7</sup> Quando as palavras vinham em um enunciado (poético ou não), havia uma sequência de sílabas longas e breves. Na poesia, havia um rearranjo de modo a se criar unidades poéticas chamadas pés métricos. Um verso tinha um padrão métrico definido, por exemplo, um hexâmetro, formado de 6 pés métricos, os quatro primeiros podendo ser dátilos (longa, breve, breve) ou espondeus (longa, longa), mas sendo o quinto pé métrico um dátilo e o sexto um espondeu. Pelas medidas de duração, um dátilo tinha a mesma duração de um espondeu. O que os diferenciava era que o dátilo ocorria em três sílabas e o espondeu, em duas.

Embora todas as línguas tenham palavras formadas de sílabas longas e breves, o grego antigo e o latim fizeram dessa realidade um modelo rítmico próprio, que era sentido na poesia, porque nesse contexto havia simetria e expectativa de repetição de unidades de valor duracional determinado. Na prosa, ocorriam as mesmas sílabas com suas respectivas durações, mas os enunciados não escolhiam as palavras para formar estruturas poéticas de versos.<sup>8</sup>

Quando surgiram as línguas românicas, derivadas do latim, houve uma mudança na estrutura lexical que envolveu muitos tipos de transformações (WILLIAMS, 1975). Entre elas, houve a perda da oposição fonológica entre sílabas longas e breves, exceto em italiano. Como não fazia mais sentido compor versos usando pés métricos poéticos, os poetas começaram a fazer versos com número determinado de sílabas, acrescentando as rimas em fim de verso. As rimas baseiam-se tipicamente na estrutura prosódica dos grupos tonais, marcando a sílaba tônica frasal (CAGLIARI, 1984). Criou-se então uma tradição de metrificação que foi estendida às línguas românicas. Tal modelo podia ser perfeitamente adaptado ao italiano, ao francês e ao espanhol, mas encontrou problemas sérios com relação ao português, como mostra Cagliari (1984; 2007a, p. 142-160).

Com esse novo modelo de ritmo, os estudiosos deixaram de lado a duração lexical das sílabas, e assumiram a tonicidade como saliência rítmica. Essa tonicidade veio do processo aerodinâmico, porém não atribuído às palavras, mas aos grupos tonais. Um grupo tonal pode ser, por exemplo, uma frase, como:

(2) Davi brinca com Artur no quintal.

Nesse enunciado (grupo tonal) ocorrem várias palavras, cada qual dita isoladamente carrega uma sílaba tônica (em itálico): *Davi brinca com Artur no quintal*. O enunciado poderia também ser dito com o seguinte esquema de tonicidade: Davi *brinca* com *Artur* no *quintal*, em que a palavra Davi ficou somente com sílabas átonas. Como o enunciado é uma frase declarativa afirmativa, a curva melódica é descendente, tendo um declínio maior na sílaba *-tal* da palavra *quintal*. Essa marca entoacional associada à tonicidade e ao fato de o enunciado ter uma estrutura sintática-semântica de tema/rema, dado/novo (HALLIDAY, 1970; HALLIDAY; GREAVES, 2008), faz com que a sílaba *-tal* seja o foco semântico-sintático, ou seja, o acento frasal ou foco prosódico do grupo tonal. Essa

7 Nas línguas de ritmo acentual, os linguistas mostram que há regras de atribuição de tonicidade em certos contextos.

8 Antônio Feliciano de Castilho (1908) estudou o ritmo do português através da notação de sílabas longas e breves, mostrando que a língua tem uma estrutura rítmica não apenas pela tonicidade, pela rima, mas também pela duração relativa das sílabas. Veja também o trabalho de Cavalcanti Proença (1955). Também é interessante ver como Jerônimo Soares Barbosa tratou o ritmo do português no final do século XVIII (CAGLIARI, 1985).

interpretação representa uma possível resposta para uma pergunta do tipo: *Onde Davi brinca com Artur?* Obviamente, a estrutura sintática-semântica poderia ser modificada, fazendo com que o foco prosódico caia, por exemplo, na sílaba *brin-* da palavra *brinca*. Neste caso, o foco lembra que o enunciado é uma possível resposta a uma pergunta do tipo: *O que Davi faz com Artur no quintal?*

Quando reduzimos um grupo tonal a uma palavra isolada, mantemos o mesmo esquema prosódico, porém já não temos mais o esquema sintático-semântico. Por essa razão, uma palavra como *lâmpada* tem o acento (saliência prosódica) na sílaba *lâm-*, dita com um tom médio (entoacional), o qual decresce até o final da palavra, indo pelas sílabas átonas. O mesmo ocorre com todas as palavras pronunciadas isoladamente, na mesma situação.

Por causa desse fenômeno prosódico, algumas línguas podem optar por ter no léxico sílabas de igual duração, portanto, nem longas nem breves, nem tônicas nem átonas, quando comparadas entre si. Quando ocorre um enunciado (uma ou mais palavras), surge a entoação e a tonicidade própria do esquema de grupos tonais, e a língua passa a ter uma marca de tonicidade para as palavras em contextos de foco do grupo tonal. Uma língua como o espanhol e o francês não têm marcas lexicais de tonicidade, esta só vem nos enunciados. O espanhol associa à tonicidade do foco, em geral, no final do enunciado, uma duração maior, o que reforça a saliência fônica nesse contexto. O francês tem um esquema lexical que define algumas sílabas como sendo breves (*chuá*) em certos contextos e costuma marcar o final de grupos tonais com uma saliência entoacional que varia de acordo com o fato de o grupo tonal ser seguido de outro (entoação suspensiva, mesmo em frases declarativas afirmativas) ou ser final de fala (entoação descendente).

Temos, então, três modos de implementar o ritmo: um no modelo greco-latino, um no modelo espanhol e outro no modelo francês. Podemos reduzir esses três tipos a dois: um modelo baseado na duração contrastiva das sílabas (greco-latino) e outro baseado na tonicidade do grupo tonal (espanhol e francês). O italiano e o alemão são línguas de ritmo semelhante ao grego e ao latim. Em italiano e em alemão, temos uma oposição fonológica com a duração relativa das sílabas longas e breves:

- (3) italiano oposição /t/ - /t/: *seta* (seda); *setta* (seita)
- (4) alemão: oposição ü – ü: *füllen* (encher); *fühlen* (sentir)

O português é tipicamente uma língua de ritmo acentual. Basta dizer uma lista de palavras ou contar números sem pausas intermediárias, que salta à vista o fato de o português ser uma língua de ritmo acentual. O que acontece com o português é inviável em italiano, em francês, em espanhol e em alemão. No exemplo abaixo, o asterisco marca a isocronia das sílabas tônicas em português:

- (5) \* \* \* \* \* (batidas rítmicas)  
vinte e **um**, vinte e **dois**, vinte e **três**, vinte e **quatro**, vinte e **cinco**, vinte e **seis**...

### O caso do japonês

Por motivos estranhos, surgiu entre os trabalhos de estudo do ritmo das línguas a ideia de que o japonês era uma língua que tinha um ritmo peculiar, chamado de ritmo moraico

(PORT et al., 1987). A ideia se espalhou e surgiram trabalhos baseados nessa ideia, por exemplo, de Beckman e Pierrehumbert (1986).

O japonês é uma língua que tem uma metrficação poética baseada na noção de duração silábica (ou moras) (WARNER; ARAI, 2001). Muito certamente, em tempos anteriores, o japonês devia ser uma língua com oposição fonológica de duração, tendo, portanto, um esquema rítmico do tipo do grego antigo e do latim. A língua se transformou e as durações silábicas acabaram por ser iguais como o espanhol. Mas, diferentemente dessa língua, o japonês tem um andamento diferente, o que motivou a observação de que o ritmo do japonês era do tipo *machine gun* (JAMES, 1940). Lloyd James era um engenheiro interessado em telegrafia. Portanto estudar o ritmo das línguas (e seus sistemas de escrita) ajudava a melhorar o sistema dos telégrafos. Línguas com oposição entre sílabas longas e breves eram boas para os telégrafos, assim como línguas de ritmo acentual, porque davam um ritmo bom à transmissão e ao mesmo tempo apresentavam variações de sílabas longas e breves. O japonês era um problema à parte, porque, além de ter sílabas iguais (o telégrafo trabalha com unidades longas e breves), ainda tinha um andamento peculiar que não coincidia com as fronteiras dos grupos tonais de outras línguas. Entretanto, o rótulo dado por Lloyd James para a língua japonesa ajudou a confundir mais os estudos sobre o ritmo das línguas na segunda metade do século XX.

Segundo a teoria exposta acima, o japonês é uma típica língua de ritmo silábico. Mais antigamente, era semelhante ao modelo greco-latino. Algumas dessas características ficaram marcadas na escrita que chegou até hoje. Além disso, a escrita é um elemento essencial da poética japonesa (cf. os *hai-kais*...). De acordo com a escrita japonesa, a algumas letras é atribuída uma mora (duração poética) e a outras, duas moras. Portanto a escrita japonesa faz lembrar uma língua de ritmo semelhante ao modelo greco-latino. Porém a fala segue outro padrão (do tipo francês, espanhol – com andamento diferenciado). Uma palavra como *nippon* tem duas sílabas fonéticas, de igual duração. Porém, a partir da escrita, a palavra vale diferentemente: a primeira “sílabas” *ni-* vale uma mora e a segunda, *ppon*, vale duas moras. Esse fato cultural tem repercussão no modo como as pessoas pensam a língua e a usam para fazer poesia e propagandas, por exemplo. Do ponto de vista da fonética, o que ocorre são duas sílabas de igual duração (LADEFOGED, 1975, p. 224).

### **Como o ritmo de uma língua pode variar**

Relembrando o que já foi dito anteriormente, todo sistema linguístico não só admite, como precisa de uma porcentagem de variação, para tornar o processo comunicativo melhor. Isso acontece em todos os níveis do sistema, incluindo o ritmo. Vejamos alguns casos.

O latim era uma língua de ritmo silábico e gerou línguas com ritmos diferentes, incluindo o português, que é uma língua de ritmo acentual. Na passagem do latim para o francês e o espanhol, as durações silábicas opositivas foram modificadas, conforme explicado acima. No caso do português, a tonicidade sobrepujou a duração silábica, marcando o ritmo à moda dos compassos musicais, ou seja, surgiu uma língua de ritmo acentual (MASSINI-CAGLIARI, 1994; CAGLIARI; MASSINI-CAGLIARI, 1998). As línguas românicas perderam o modelo dos pés métricos poéticos latinos. O italiano conservou uma oposição fonológica de duração entre as sílabas, mas também perdeu o modelo de pés métricos poéticos latinos. No caso de uma língua de ritmo silábico (latim) passar a

acentual (português), costuma ocorrer a perda da oposição fonológica de duração, mas há a introdução do fator andamento para gerar intervalos isocrônicos entre as sílabas tônicas. No caso de uma língua acentual passar a silábica (variações dialetais do inglês, do português), ocorre uma perda da isocronia acentual gerada pelo andamento e uma tendência a tornar todas as sílabas com duração igual, eventualmente, com alguma sílaba muito longa ou muito breve em determinados contextos. Um exemplo de como uma língua de ritmo acentual pode ser falada com um modelo de língua silábica é o inglês escocês e o inglês indiano, embora cada uma dessas variedades difira uma da outra. O inglês escocês tende a ter sílabas iguais (do tipo espanhol) e o inglês indiano tende a ter sílabas com durações alternadas longas e breves com um andamento peculiar. Outro exemplo é o português falado nas proximidades das fronteiras do Brasil com os países vizinhos, falantes de espanhol. Essa fala de gaúcho de fronteira caracteriza-se por pronunciar todas as sílabas com igual duração, como ocorre com o espanhol. Tal pronúncia soa para os demais falantes de português como se os gaúchos estivessem silabando, quando falam.

Portanto, uma língua pode ser falada com diferentes ritmos, dependendo do tipo de dialetos que ela tem. Mas o resultado será ou uma fala com ritmo acentual ou uma fala com ritmo silábico: não há outros tipos.

Entretanto, um mesmo falante pode, em determinadas circunstâncias, alterar o ritmo do que diz. Ninguém pede socorro usando expressões com ritmo acentual, mas pronunciando cada sílaba com duração igual. Eventualmente, haverá um destaque para alguma palavra. O mesmo costuma acontecer em certas circunstâncias, quando alguém quer chamar a atenção para uma palavra, destacando-a. Nesses casos, a pessoa acentua e aumenta mais a duração da sílaba tônica das palavras. Por razões emocionais, também pode ocorrer uma mudança de ritmo acentual para silábico. Todavia não se conhecem casos em que uma língua de ritmo silábico use de um modelo rítmico acentual nas mesmas circunstâncias apontadas acima, ou em outras de outro tipo. O que pode ocorrer é uma mudança de modelo rítmico silábico. Por exemplo, o falante pode optar por dizer todas as sílabas com igual duração, independentemente da regra que estabelece uma oposição fonológica de duração.

Por fim, há evidências segundo as quais todas as crianças, de todas as línguas, ao adquirirem a linguagem, começam falando com ritmo silábico (do tipo todas as sílabas iguais – a última mais longa) e só mais tarde passam a usar outros tipos de ritmo, próprios de suas línguas. Por outro lado, os bebês falantes variam muito a melodia carregada pelas sílabas (balbucio).

As observações acima mostram que não se pode definir *a priori* o ritmo de uma língua, sem saber qual é a variedade dialetal usada ou sem saber em que circunstâncias o falante usou a língua.

## **Conclusão**

As línguas produzem as sílabas através do processo aerodinâmico de produção da fala. Nessa etapa, as sílabas são pulsos torácicos que vão se definir como sílabas depois que passam pela glote. De acordo com o sistema linguístico interiorizado na mente dos falantes, as sílabas são categorizadas como sendo longas ou breves, por oposição de umas com as outras. Elas são também categorizadas em função da pressão da corrente de ar e

de sua consequência acústica de intensidade, gerando padrões de saliência que chamamos de sílabas tônicas ou átonas, também com valores relativos e contrastivos, porém não absolutos fisicamente.

Esses padrões são categorizados na mente dos falantes em função das palavras em que essas sílabas ocorrem. Portanto todas as línguas do mundo apresentam, em princípio, sílabas longas e breves, concomitantemente com o fato de elas serem também tônicas ou átonas (CAGLIARI, 1992, 2007b). Os itens lexicais são reestruturados em termos de tonicidade ou mesmo das moras (CAGLIARI, 2007a, p. 135-136) em função da estrutura prosódica dos grupos tonais. As ocorrências das sílabas longas ou breves irão depender da estrutura sintático-semântica ou pragmática, ligada à estruturação prosódica dos grupos tonais.

Em função desses parâmetros, as línguas ou são de ritmo acentual ou são de ritmo silábico, sem possibilidade de outro tipo. Essa é a categorização que existe na mente dos falantes e que eles usam para falar, ouvir e reconhecer os sons de suas línguas. Por causa de fatores fonéticos como o andamento, a velocidade de fala, a concatenação e de fatores sintáticos, semânticos e pragmáticos, a estrutura rítmica que está na mente dos falantes pode variar fisicamente. Por outro lado, há mudanças mais radicais na estrutura rítmica de uma língua, fazendo com que uma língua possa ser falada com um tipo de ritmo em um dialeto e com outro tipo de ritmo em outro dialeto. As medidas acústicas de duração e de tonicidade, obtidas através de análises instrumentais, não conseguem definir com precisão os padrões rítmicos que os falantes têm em suas mentes e que constituem o verdadeiro sistema linguístico rítmico das línguas. Em vez de destruir o sistema linguístico e os padrões mentais do ritmo da mente dos falantes, os foneticistas deveriam procurar os parâmetros que fazem a interface entre a mente e o mundo físico (articulatório, acústico e auditivo) da fala, do mesmo modo como se procurou fazer com as propriedades distintas fonológicas e seus correlatos articulatórios, acústicos e auditivos.

## REFERÊNCIAS

ABERCROMBIE, D. A phonetician's view of verse structure. In: JONES, W. E.; LAVER, J. (Ed.) *Phonetics in Linguistics: a book of readings*. London: Longman, 1973. p. 6-13.

\_\_\_\_\_. *Elements of general phonetics*. Edinburgh: Edinburgh University Press, 1967.

BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas: Pontes Editores, 2006.

\_\_\_\_\_. Elementos para uma tipologia do ritmo (linguístico) da fala. *Cognito - Cadernos românicos em ciências cognitivas*, Grenoble, França, v. 2, n. 1, p. 31-58, 2004.

\_\_\_\_\_. Syllable-timing in Brazilian Portuguese: uma crítica a Roy Major (Tempo-silábico em Português do Brasil: a critic to Roy Major). *DELTA*, São Paulo, v. 16, n. 2, p. 369-402, 2000.

BECKMAN, M.; PIERREHUMBERT, J. Intonational structure in Japanese and English. *Phonology Yearbook*, v. 3, p. 255-309, 1986.

BERTINETTO, P. Reflections on the dichotomy stress vs. syllable-timing. *Revue de phonétique appliquée*, p. 91-93; 99-130, 1989.

CAGLIARI, L. C. *Elementos de fonética do Português Brasileiro*. São Paulo: Paulistana, 2007a.

\_\_\_\_\_. Prosódia: ontem e hoje. In: FONSECA-SILVA, M. C.; PACHECO, V.; LESSA-DE-OLIVEIRA, A. S. C. (Org.). *Em torno da língua(gem): questões e análises*. Vitória da Conquista: Edições UESB, 2007b. p. 15-40.

\_\_\_\_\_. Da importância da prosódia na descrição de fatos gramaticais. In: ILARI, R. (Org.). *Gramática do Português Falado - Níveis de Análise Linguística*. Campinas: Editora da Unicamp, 1992. p. 39-64, v. II.

\_\_\_\_\_. O ritmo do Português na interpretação de Jerônimo Soares Barbosa. In: ENCONTRO NACIONAL DE FONÉTICA E FONOLOGIA – UFSC Working papers in linguistics, I, 1984. *Anais...* ANGENOT, J. P. (Ed.). Florianópolis: UFSC, v. 1, p. 27-38, 1985.

\_\_\_\_\_. Análise fonética do ritmo em poesia. *EPA: Estudos portugueses e africanos*, UNICAMP – IEL, Departamento de Teoria Literária, Campinas, n. 3, p. 67-96, 1984.

CAGLIARI, L. C.; ABAURRE, M. B.. Elementos para uma investigação instrumental das relações entre padrões rítmicos e processos fonológicos no Português brasileiro. *Cadernos de estudos lingüísticos*, Campinas, UNICAMP-IEL-DL, n. 10, p. 39-57, 1986.

CAGLIARI, L. C.; MASSINI-CAGLIARI, G. Quantidade e duração silábicas em Português do Brasil. *DELTA*, São Paulo, PUC-SP, v. 14, Número Especial, p. 47-59, 1998.

CASTILHO, A. F. *Tratado de metrificação portuguesa*. Lisboa: Empreza da História de Portugal, Sociedade Editora, 2 vol. 1850, ed. de 1908.

CATFORD, J. C. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press, 1977.

CAVALCANTI PROENÇA, M. *Ritmo e poesia*. Rio de Janeiro: Org. Simões, 1955.

COOPER, G.; MEYER, L. B. *The rhythmic structure of music*. Chicago: Un. of Chicago Press, 1960.

DAUER, R. M. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, Amsterdam, Elsevier, v. 11, p. 51-62, 1983.

FANT, G. Analysis and synthesis of speech processes. In: MALMBERG, B. (Ed.) *Manual of phonetics*. Amsterdam: North-Holland, 1968. p. 173-277.

HALLIDAY, M. A. K. The tones of English. In: JONES, W. E; LAVER, John (Ed.). *Phonetics in linguistics*. London: Longman, 1973. p. 103-126.

\_\_\_\_\_. *A course in spoken English: intonation*. London: Oxford University Press, 1970.

- HALLIDAY, M. A. K.; GREAVES, W. S. *Intonation in the grammar of English*. London: Equinox, 2008.
- JAMES, L. A. *Speech signals in telephony*. London: Pitman, 1940.
- JONES, D. *An outline of English phonetics*. Cambridge: Cambridge University Press, 1960. [1. ed. 1917]
- LADEFOGED, P. *A course in phonetics*. New York: Harcourt Brace Jovanovich, 1975.
- LAVIER, J. *Principles of phonetics*. Cambridge: Cambridge University Press, 1994.
- MASSINI-CAGLIARI, G. *Acento e ritmo*. São Paulo: Contexto, 1994.
- MASSINI-CAGLIARI, G; CAGLIARI, L. C. Fonética. In: MUSSALIM, F.; BENTES, A. C. (Org.). *Introdução à linguística: domínios e fronteiras*. São Paulo: Contexto, 2000. p. 105-146.
- NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986.
- O'DELL, M.; NIEMINEN, T. Coupled oscillator model of speech rhythm. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, XIV, San Francisco, 1999. *Proceedings...* OHALA, J. (Ed.). Los Angeles: University of California, 1999. v. 2. p. 1075-1078.
- PIERREHUMBERT, J. B. *The phonetics and phonology of English intonation*. Ph.D. Thesis. Cambridge Massachusset: M.I.T., 1980.
- PIKE, K. L. *The intonation of American English*. Ann Arbor: The University of Michigan Press, 1945.
- PORT, R. F.; DALBY, J.; O'DELL, M. Evidence for mora-timing in Japanese. *Journal of the acoustical society of America*, New York: ASA Publications, v. 81, n. 5, p. 1574-1585, 1987.
- REIS, C. A. C. *L'interaction entre l'accent, l'intonation et le rythme en portugais brésilien*. Thèse de doctorat. Aix-en-Provence: Université de Provence, Institut de Phonétique, 1995.
- ROACH, P. On the distinction between *stress-timed* and *syllable-timed* languages. In: CRYSTAL, D. (Ed.). *Linguistic controversies*. London: Edward Arnold, 1982. p. 73-79.
- SELKIRK, E. *Phonology and syntax: the relation between sound and structure*. Cambridge Massachusset: The M.I.T. Press, 1984.
- WARNER, N.; ARAI, T. The role of the mora in the timing of spontaneous Japanese speech. *Journal of the acoustical society of America*, New York, v. 109, n. 3, p. 1144-1156, 2001.

WENK, B.; WIOLLAND F. Is French really syllable-timed? *Journal of phonetics*, Amsterdam: Elsevier, n. 10, p. 193-216, 1982.

WILLIAMS, E. *Do latim ao português*. Rio de Janeiro: Tempo Brasileiro, 1975.

# Análise comparativa da entoação do português brasileiro e do inglês norte-americano no filme *Shrek*

(Comparative analysis of the intonation of Brazilian Portuguese and North-American English in the movie *Shrek*)

Maíra Sueco Maegava Córdoba<sup>1</sup>

<sup>1</sup> Universidade Federal do Triângulo Mineiro (UFTM), Universidade Estadual Paulista (Unesp)

mcordula@yahoo.com.br, mcordula@gmail.com

**Abstract:** The aim of this article is to compare the intonation of Brazilian Portuguese and North-American English through the analysis of a sequence of scenes from the movie *Shrek* (2001). The theoretical basis of the analysis is the systemic-functional approach to intonation (HALLIDAY, 1970; CAGLIARI, 2007). The factors discussed are: the number of tonal groups, the choice of tones, the syntactic choices and the choice of semantic and pragmatic meanings. The similar situational and linguistic context made it possible to observe the choice of tones as regards syntactic, semantic and pragmatic meanings, which are the same or different in the two versions of the same movie in analysis. It was also noted the special role played by the interpretation of the voice actors.

**Keywords:** English; Portuguese; intonation.

**Resumo:** O presente trabalho tem como objetivo apresentar alguns dados comparativos entre a entoação do português brasileiro e do inglês norte-americano, por meio da análise de uma sequência de cenas de um filme animado: *Shrek* (2001). A partir de uma visão sistêmico-funcional da entoação (HALLIDAY, 1970; CAGLIARI, 2007), foram considerados os seguintes fatores: número de grupos tonais, escolha de tons, escolhas sintáticas e escolhas de sentidos semânticos e pragmáticos. O contexto situacional e linguístico similar permitiu observar o funcionamento da escolha de determinados tons com relação à produção de sentidos sintáticos, semânticos e pragmáticos coincidentes ou divergentes entre as versões do filme em análise, com especial destaque para o papel da interpretação dos atores/dubladores.

**Palavras-chave:** língua inglesa; língua portuguesa; entoação.

## Introdução

O elemento prosódico da entoação tem papel significativo na produção de sentidos de uma língua. A partir desse pressuposto, o presente trabalho apresenta alguns resultados<sup>1</sup> de um estudo de caráter comparativo entre a variação melódica em português brasileiro, PB, e em inglês norte-americano, IA. O material analisado foi retirado de um filme animado, *Shrek* (2001), com o objetivo de controlar algumas variáveis de importância linguística significativa. A análise seguiu os pressupostos de uma abordagem sistêmico-funcional da entoação e foram selecionados os dados relativos à quantidade de grupos tonais (unidades de análise) em cada versão e a atribuição de tons distintivos a essas unidades.

---

<sup>1</sup> Os dados que aparecem nesse artigo fazem parte dos resultados de uma pesquisa de doutorado do Programa de Pós-graduação em Linguística e Língua Portuguesa da Universidade Estadual Paulista (Unesp), *campus* de Araraquara.

## Metodologia

Os dados deste trabalho foram obtidos a partir da análise do material de áudio de duas versões de um mesmo filme animado, *Shrek* (2001). A escolha por um *corpus* composto do áudio extraído de um filme animado se deve, entre outros fatores, à possibilidade de conhecimento do contexto de produção dos enunciados, à similaridade de contexto situacional e linguístico nas duas versões analisadas e à restrição dos atores/dubladores ao recurso artístico da voz.

O material de áudio foi extraído de um trecho de cinco minutos do filme animado selecionado, com a presença de quatro personagens de vozes masculinas: Shrek (o ogro, herói às avessas), Burro (um burro falante que se une ao ogro em suas aventuras), Lorde Farquaad (o vilão que governa um reino, mas que ainda não é rei) e o Guarda (guarda do reino do Lorde Farquaad). Nas cenas em questão, Shrek está acompanhado do Burro e se depara com o Lorde Farquaad, que lhe oferece uma troca de favores: o ogro salva uma princesa do dragão e, em troca, recebe seu pântano de volta (pântano este que foi invadido por criaturas de conto de fadas em refúgio à perseguição criada por Lorde Farquaad); após a consolidação do acordo, o herói do filme e seu acompanhante fazem uma longa discussão sobre a essência dos ogros e, finalmente, os dois chegam ao castelo onde a princesa que deve ser resgatada por eles está aprisionada.

As duas versões do material de áudio foram analisadas seguindo os pilares de uma abordagem sistêmico-funcional da entoação desenvolvida e aplicada por Halliday (1970) ao inglês e aplicada ao português brasileiro por Cagliari (2007),<sup>2</sup> línguas-alvo da análise comparativa deste trabalho. De acordo com esse pressuposto teórico, são três os aspectos fundamentais para a análise da variação melódica: a tonalidade, a tonicidade e os tons. Em resumo, a tonalidade refere-se ao recorte do material linguístico em grupos tonais, ou seja, as unidades de análise; a tonicidade está associada à posição da sílaba tônica de cada grupo tonal e os tons remetem às escolhas entre um conjunto de tipos de movimento da altura melódica significativos linguisticamente. Com relação à interface da fonologia com a semântica e a léxico-gramática, seguimos a proposta de Halliday e Greaves (2008) para o sistema de entoação, além das descrições (e glosas) apresentadas em Halliday (1970) e Cagliari (2007).

## Discussão dos resultados

A partir da descrição dos dados a partir dos pressupostos de análise de Halliday (1970) e Cagliari (2007), podemos partir para uma análise comparativa entre as versões dubladas em PB e em IA. Vale lembrar que esse estudo comparativo tem a vantagem de contar com enunciados produzidos em mesma situação contextual e com duração aproximada, já que o tempo de cena não pode ser excedido. Além disso, é importante destacar que as diferenças ou semelhanças advindas da comparação dos dados refletem escolhas específicas dos falantes estudados dentro do sistema linguístico disponível e dentro da norma considerada padrão para cada idioma, não podendo ainda servir para uma generalização do uso de determinados contornos melódicos em cada língua.

---

<sup>2</sup> A versão de 2007 é uma publicação da tese de livre docência do mesmo autor, defendida em 1982 (CAGLIARI, 1982), na Universidade Estadual de Campinas (Unicamp).

A análise comparativa entre os dados do PB e do IA a seguir considera os seguintes fatores no trecho do filme selecionado para a pesquisa: número de grupos tonais, escolha de tons, escolhas sintáticas e escolhas de significados semânticos e pragmáticos.

Em primeiro lugar, devemos observar o número de grupos tonais existentes nos trechos analisados. Em PB, foram destacados 142 grupos tonais enquanto na versão em IA, 150.

Como não há igualdade de números de GTs tanto na comparação entre cenas como na comparação entre personagens, é importante levantar os casos específicos de diferenças no número de GTs. Os casos em que há discrepância entre os números de GTs nos enunciados correspondentes em cada língua são apresentados nos quadros 1, 2 e 3: no primeiro quadro, são apresentados os enunciados que em PB são produzidos com um GT e que em IA são produzidos com dois ou mais GTs; no quadro seguinte, estão listados os enunciados produzidos com 2 GTs ou mais em PB; porém, com apenas um GT em IA; finalmente, no Quadro 3, são apresentados alguns enunciados que diferem com a presença de um vocábulo em uma das versões que não apresenta correspondente na outra versão analisada.

**Quadro 1 – Correspondência 1 GT em PB para 2 ou mais GTs em IA**

	PB	IA
(01)	//1 eu lhes /dou o /nosso /campe/ão//	//3 ^ I give /you// //1 ^ our /champion//
(02)	//1 É o /meu /pântano//	//1 Yeah// //1 my swamp//
(03)	//13 Certo/Ogro//	//3 All /right// //1 ogre//
(04)	//5 ^ e /eu lhe de/volvo o /seu /pântano//	//3 ^ and I'll /give you// //1 your /swamp back//
(05)	//1 Acho que /tem um /bom mo/tivo pros /burros não fa/larem//	//5 maybe there's a good /reason// //1 donkeys shouldn't /talk//
(06)	//1 ^ eu não en/tendo /Shrek//	//1 I don't get it// //2 Shrek//
(07)	//1 há /mais do que se ima/gina nos /ogros//	//3 there's a /lot more to /ogres// //1 ^ than people /think//
(08)	//13 ^ O/kay/ ãh //	//3 O/kay// //1 uhm//
(09)	//3 Já conheceu al/guém que você fa/lasse//	//1 Have you ever /met a /person// //3 you /say//
(10)	//3 “Ei vamos co/mer pa/vê”//	//1 Hey// //1 Let's get some par/fait//
(11)	//1 ^ eu /acho que eu pre/firo você can/tando//	//1 ^ I /think// //1 ^ I pre/ferred your /humming//
(12)	//53 é /bom você avisar os /outros antes de soltar /um//	//3 ^ you gotta /warn somebody// //1 ^ before you just crack one /off//
(13)	//1. Eu tava a/té de boca a/berta//	//3 ^ My /mouth was /open// //1 ^ and /everything//

(14)	//1 Tá le/gal, en/xofre//	//1 <u>Yeah</u> // //1 <u>right</u> // //1 <u>brimstone</u> //
(15)	//1 ^ Com cer/teza é /grande//	//1 <u>Sure</u> // //1 ^ it's big e/ <u>nough</u> //
(16)	//2 Lembra quando você /disse que os / monstros têm ca/ <u>mad</u> as//	//3 ah// //2 ^ re/ <u>member</u> // //3 ^ when you / <u>said</u> that// //1 ogres have / <u>layers</u> //

**Quadro 2 – Correspondência 2 GTs em PB para 1 GT em IA**

	<b>PB</b>	<b>IA</b>
(17)	//1 O/ <u>kay</u> // //3 Deixa eu en/tender / <u>bem</u> //	//1 ok let me /get this / <u>straight</u> //
(18)	//3 mo/er os / <u>ossos</u> // //1 e /pôr no / <u>pão</u> //	//53 grind his / <u>bones</u> to make your / <u>bread</u> //
(19)	//3 <u>Sabe</u> // //1 ^ serviço com/ <u>pleto</u> //	//1 ^ you /know the /whole / <u>ogre</u> trip//
(20)	//1 oh// //3 ^ enten/di//	//5 ^ I /know / <u>what</u> //
(21)	//1 <u>Uhm</u> //	(Sniffs)
(22)	//1 <u>Ah</u> // //3 deixa /eles no / <u>sol</u> //	//3 ^ You /leave them out in the / <u>sun</u> //
(23)	//3 <u>Ogros</u> // //3 não / <u>são</u> // //5 como / <u>bolos</u> //	//1 Ogres are / <u>not</u> like / <u>cakes</u> //
(24)	//1 ^ O pa/vê deve ser a /coisa mais delici/osa de/ <u>todo</u> o / <u>mundo</u> // //1 ^ Eu a/ <u>doro</u> pa/vê//	//1 ^ Par/faits may /be the most de/licious thing on the /whole damn / <u>planet</u> //
(25)	//1 <u>Sabe</u> // //1 tá a maior su/ <u>jeira</u> a/ <u>qui</u> //	//1 ^ Cause I'm /making a / <u>mess</u> //
(26)	//3 Só de fa/lar em pa/vê// //1 ^ eu me ba/ <u>bei</u> todo//	//1 Just the / <u>word</u> par/fait make me start / slobbering//
(27)	//1 Não / <u>vem</u> com essa his/tória de enxofre / <u>não</u> // //2 <u>tá</u> //	//3 ^ Don't be talking about it's the / <u>brimstone</u> //

**Quadro 3 – Alguns casos em que vocábulos presentes em apenas uma das versões analisadas não gera diferença no número de GTs correspondentes em PB e em IA**

	<b>PB</b>	<b>IA</b>
(28)	//4 não /acho legal / <u>não</u> //	//1 no, not /really, / <u>no</u> //
(29)	//3 <u>Shrek</u> // //2 cê tem um /lenço de pa/pel / <u>ai</u> //	//2 ^ Do you /have a / <u>tissue</u> // //1 ^ or / <u>something</u> //
(30)	//3 Pode acredi/ <u>tar</u> //	//1 ^ Be/ <u>lieve</u> me, / <u>Donkey</u> //

O número de GTs em cada língua analisada é muito próximo, porém a diferença em cada versão se deve a mais de um fator, que são listados a seguir: a) existência de alguns enunciados em apenas uma das versões; b) estrutura sintática diferente dos enunciados em cada uma das versões; c) diferente escolha dos falantes ao recortar os enunciados em GTs.

Em primeiro lugar, em PB, há 7 enunciados não presentes na versão em IA, que são: “Oh” (20), “Uhm” (21), “Ah” (22), “Eu adoro pavê” (24), “sabe” (25), “tá” (27), “Shrek” (29). Já, em IA, existem 6 enunciados que não estão presentes na versão em PB: “Yeah” (2), “and everything” (13), “ah” (16), “no” (28), “or something” (29), “Donkey” (30). Nota-se que nos enunciados destacados acima dos GTs (13) e (29) dos quadros 1 e 3, respectivamente, durante o processo de tradução do texto, houve uma preferência, em PB, pelo uso de advérbios em vez das locuções de final de sentença em IA, iniciadas com uma conjunção mais um composto de *some*: (29) “Do you have a tissue or something” foi traduzido por “Você tem um lenço de papel aí?” e (13) “my mouth was open and everything” por “Eu tava até de boca aberta”. Esses dois enunciados também podem ser considerados na categoria de escolhas sintáticas diferentes nas duas versões, como pode ser visto a seguir. Ainda vale notar que o acréscimo de enunciados em algumas ocasiões não gerou a produção de um novo GT, como é o caso dos itens (28) e (30) do Quadro 3<sup>3</sup> em que as palavras “no” e “Donkey” não aparecem na versão em PB.

Com relação à diferentes estruturas sintáticas, foi possível constatar a presença de 4 enunciados com essa característica nas duas versões, assim podendo apresentar um número de grupos tonais diferentes. Os seguintes enunciados são constituídos de um grupo tonal em IA e de dois em PB: a) (18) “grind his bones to make your bread” por “moer os ossos e por no pão” (adição de uma oração coordenada); b) (26) “Just the word parfait make me start slobbering” por “Só de falar em pavê, eu me babei todo” (introdução de uma oração subordinada) e c) (27) “Don’t be talking about it’s the brimstone” por “Não vem com essa história de enxofre não, tá?” (uso de declarativa interrogativa com final ‘tá’ – *question tag*). Já no seguinte enunciado: (7) “there’s a lot more to ogres than people think” por “há mais do que se imagina nos ogros”, ocorre o oposto, a mudança de foco para o final da sentença propicia a presença de apenas um grupo tonal em PB enquanto há dois em IA. Há ainda o caso em que há a presença de dois advérbios mais um substantivo funcionando como três grupos tonais em IA, enquanto em PB há dois grupos tonais: uma afirmação e uma oração de uma única palavra (substantivo): (14) “yeah, // right, //brimstone” e “tá legal, // enxofre”.

Outro fator constatado é a escolha de recorte dos grupos tonais. Nesse quesito, vale lembrar que a escolha de recorte de grupo tonal está diretamente relacionada ao foco das sentenças, em outras palavras, cada grupo tonal recorta a informação linguística entre o dado e o novo; se há mais de um grupo tonal, há mais de um foco.

Foram constatados 12 casos em que a versão em IA apresenta mais de um grupo tonal enquanto a versão em PB apresenta apenas um; dessa forma, há dois focos informacionais nos enunciados em IA, enquanto em PB há apenas um. Nos exemplos encontrados no *corpus*, é possível destacar alguns tipos de recorte: a) entre o verbo e seu objeto; b) entre a oração principal e vocativos ou interjeições e c) entre oração principal e oração subordinada.

---

3 A expressão em inglês “or something” do enunciado (27) é absorvida em apenas um GT no enunciado correspondente em PB, como foi explicitado no parágrafo desta nota; porém, como em PB, há também o acréscimo de um vocábulo não presente em IA, “Shrek”, os enunciados pertencentes ao mesmo trecho do *corpus* analisado: “Shrek, cê tem um lenço de papel aí” e “Do you have a tissue or something?”, possuem a mesma quantidade de GTs, 2. Optou-se, portanto, na exposição dos dados, por incluir o enunciado (27) no Quadro 1, no qual se destacam os casos em que o número de GTs se mantém apesar das diferenças de aspectos sintáticos ou semânticos entre os enunciados correspondentes.

Em primeiro lugar, nos dois enunciados de Lorde Farquaad: (1) “*I give you // our champion*” (“eu lhes dou o nosso campeão”) e (4) “*and I’ll give you // your swamp back*” (“e eu lhe devolvo o seu pântano”), o fato de o objeto direto estar separado do restante da oração indica que a ação do verbo “dar” é tão enfatizada quanto o que é “dado”; destaca-se, então, a ação de oferecimento do Lorde, o que contribui para um discurso construído pelo rei não legalizado, Lorde Farquaad, que o representa como um monarca bom e benevolente, mesmo diante da constatação de que foi esse personagem que mandou aprisionar as criaturas de contos de fadas na história do filme *Shrek* (2001), o que resultou no desalojamento do ogro ao qual o Lorde promete devolver o pântano.

Há 5 casos em que há o recorte entre a oração e expressões sintaticamente ligadas de forma mais solta: (3) “*All right // ogre*” (“Certo, ogro”); (6) “*I don’t get it // Shrek*” (“eu não entendo, Shrek”); (8) “*okay // uhm*” (“okay, ãh”); (10) “*Hey // Let’s get some parfait*” (“Ei, vamos comer pavê”) e, finalmente, (15) “*Sure // it’s big enough*” (“Com certeza é grande”). No primeiro caso o vocativo, Shrek, forma um GT à parte; no segundo exemplo, a interjeição “*hey*” está isolada em um GT e, no terceiro enunciado, o adjunto adverbial “*sure*” também está destacado compondo um GT. Há, então, a valorização desses elementos de interação entre os falantes na versão em IA: nos dois primeiros casos listados, há a atenção para chamar a atenção do outro falante e, no segundo, há a expressão da opinião do falante sobre o conteúdo do seu enunciado.

Ainda é preciso citar os exemplos em que as orações principais estão separadas das orações subordinadas por meio da entoação: (5) “*maybe there’s a good reason // donkeys shouldn’t talk*” (“Acho que tem um bom motivo pros burros não falarem”); (9) “*have you ever met a person // you say*” (“já conheceu alguém que você falasse”); (11) “*I think // I preferred your humming*” (“eu acho que eu prefiro você cantando”); (12) “*you gotta warn somebody // before you just crack one off*” (“é bom você avisar os outros antes de soltar um”) e (16) “*ah // remember // when you said that // ogres have layers*” (“lembra quando você disse que os monstros têm camadas”).

Há ainda três casos em que ocorre o oposto do visto acima, há mais de um grupo tonal na versão em PB enquanto há apenas um grupo tonal na versão em IA: (17) “*Okay, // deixa eu entender bem*” e “*ok, let me get this straight*”; (19) “*sabe, // serviço completo*” e “*you know, the whole ogre trip*”; (23) “*Ogros // não são // como bolos*” e “*Ogres are not like cakes*”. É possível dividir os exemplos em duas categorias: a) divisão entre oração principal e marcador conversacional e b) divisão interna entre elementos essenciais da oração, no caso, sujeito e verbo (predicado).

Com relação aos dois primeiros enunciados, vale destacar que as duas falas são do personagem Burro e seu intérprete em inglês (Eddie Murphy) tem uma velocidade de fala alta como característica artística. Nesses casos, a alta velocidade de fala influi no recorte dos enunciados em grupos tonais, nos quais as expressões indicam a interação com o interlocutor, funcionando como um elemento de manutenção do turno de fala.

Já no terceiro enunciado (23), em IA, a ênfase está na negativa (o elemento focalizado é o advérbio de negação *not*, que é a sílaba tônica saliente) e não nos elementos comparados que são considerados como Dado na unidade da informação textual.<sup>4</sup> O recorte da oração em PB em três GTs destaca a negação e os dois itens da comparação proposta no enunciado:

4 Durante a conversa do filme em análise, tanto ogros quanto bolos já tinham sido mencionados, apenas a negação da comparação é que não tinha sido feita nesse momento.

“ogros” e “bolos”, deixando claro que a comparação entre esses elementos não é aceitável pelo enunciador, Shrek. Vale destacar que, em IA, foi utilizado outro recurso prosódico além do contorno melódico nessa oração: há pausas após a palavra “*ogre*” e após a palavra “*not*”, além da presença de pausa em final de GT, ou seja, após a palavra “*cakes*”; dessa forma, o destaque para cada elemento da comparação foi realizado em IA por meio do elemento prosódico da pausa e a negação da comparação foi mais enfatizada pelo contorno melódico do enunciado com a sílaba tônica saliente no advérbio de negação.

Além da diferença no número de grupos tonais, também é interessante observar se os enunciados similares foram proferidos utilizando-se o mesmo tipo de tom ou não. Em PB, foram encontrados 142 grupos tonais no recorte selecionado, desses 77 são tom 1; 14, tom 2; 35, tom 3; 2, tom 4; 10, tom 5; 3, tom 13; e 1, tom 63.<sup>5</sup> Em IA, a classificação dos grupos tonais ficou como segue: 84 grupos tonais são tom 1; 19, tom 2; 32, tom 3; 5, tom 4; 8, tom 5; 1, tom 13; e 1, tom 53; totalizando 150 grupos tonais. Uma comparação entre a distribuição de tons em PB e em IA demonstra que os tons 1 e 3 respondem por grande parte das escolhas de tom dos GTs do *corpus*, cerca de 65% deles são caracterizados por um desses dois tons, ressaltando ainda a predominância do tom 1. A ordem de frequência em IA também é a mesma do PB: os tons 1 e 3 são os mais frequentes, seguidos do tom 2 e depois do tom 5. A frequência baixa do tom 4 e dos tons compostos também é característica das duas versões.

Dentre os enunciados similares em PB e IA do *corpus* que apresentam classificação diferente de tons, pode-se destacar a diferença entre a escolha dos falantes entre o tom 1 e o tom 3, os dois mais frequentes em ambos os recortes analisados. Tendo em mente as considerações sobre o sentido de completude e certeza expresso pelo tom 1, e o de incompletude e dependência expresso pelo tom 3, é possível fazer uma comparação entre um tom final, tom 1, e um tom suspensivo, tom 3 (cf. CAGLIARI, 2007, p. 173).

**Quadro 4 - Escolha entre tom 1 e tom 3**

PB	IA
//1 O/ <u>kay</u> // //3 Deixa eu en/ <u>tender</u> / <u>bem</u> //	//1 <i>ok let me /get this /straight//</i>
//3 só pro Farqua/ <u>ad</u> te devol/ <u>ver</u> o / <u>pântano</u> //	//1 <i>^just so /Farquaad will give you back your /swamp//</i>
//1 <u>Sabe</u> //	//3 <i>^You know /what//</i>
//1 <u>Sabe</u> //	//3 <i>^You /know//</i>
//3 <i>^Estrangu</i> / <u>lar</u> //	//1 <i>Throttle him//</i>
//3 <i>^pendu</i> / <u>rado</u> as ca/ <u>beças</u> //	//1 <i>^ and /put their /heads on a /plate//</i>
//3 <i>^pe</i> / <u>gado</u> uma / <u>faca</u> //	//1 <i>gotten a /knife//</i>
//3 Nós /somos /como ce/ <u>bolos</u> //	//1 <i>ogres /are like /onions//</i>
//3 “Ei vamos co/ <u>mer</u> pa/ <u>vê</u> ”//	//1 <i>Hey//</i> //1 <i>Let’s get some par/<u>fait</u>//</i>
//3 <i>Bye</i> / <u>bye</u> //	//1 <i>Bye-<u>bye</u>//</i>
//1 <u>Cara</u> //	//3 <i>Man//</i>
//1 Ah / <u>Shrek</u> //	//3 <i>Shrek//</i>

<sup>5</sup> Os tons compostos são notados com uma sequência de números para cada tom simples que compõe o tom composto; assim, o tom 13 é entendido como tom um três e o tom 63, tom seis três.

No Quadro 4, são citados todos os GTs correspondentes cuja função gramatical foi considerada a mesma no levantamento dos dados em que, em uma das versões, escolheu-se o tom 1; enquanto, na outra, a escolha tonal se deu pelo tom 3. Das 12 ocorrências, pode-se observar que, em 8 delas, o tom 1 foi utilizado em IA e o tom 3, em PB. Assim, é possível notar uma preferência, em PB, do uso do tom 3, que indica incompletude ou dependência, em relação ao uso do tom 1, em contextos possíveis. Ainda, é preciso considerar que usar um tom suspensivo, que indica a necessidade de se complementar o que está sendo dito, é uma estratégia conversacional para a manutenção do turno. O elevar da altura melódica indica a não finalidade e o interlocutor é informado que é bom aguardar, pois o locutor ainda deve completar sua fala.

Um dado interessante é que, em IA, o locutor só utilizou o tom 3 em casos em que o GT tem a função específica de manter ou tomar o turno, pelo uso de um marcador discursivo de introdução de conteúdo, “*You know what*”, “*you know*”, ou de vocativos: “*Man*”, “*Shrek*”. Essa constatação reforça a tendência em IA de se finalizar o discurso por meio do tom, enquanto em PB há a preferência por segurar a fala, como estratégia de manutenção do turno.

Ainda vale dizer que, em todos os casos apontados no Quadro 4, os GTs citados são acompanhados de mais enunciados produzidos pelo mesmo personagem. A única exceção é no caso de “Nós somos como cebolas”. Em PB, a proposta da metáfora criada pelo ogro a partir da pergunta do Burro, e expressa com o uso de tom 3, acentua um estado de dúvida do personagem sobre a escolha do melhor elemento comparativo. Shrek terá a oportunidade de confirmar ainda várias vezes após esse episódio que os ogros são como cebolas, portanto, a nuance de sentido adicionada ao enunciado em PB pela escolha do tom 3 é coerente com a produção de sentidos no texto, assim como a escolha pelo tom 1, tom de declarativa, para o mesmo enunciado em IA.

Outra comparação interessante a ser feita é com relação ao tom assertivo, tom 1, e o tom interrogativo, tom 2. No Quadro 5, são listados os três casos em que se atribuiu o tom 1 ao GT em PB, enquanto, em IA, foi utilizado o tom 2. Vale destacar que o padrão entoacional para interrogativas abertas para respostas sim ou não (interrogativas polares) é o tom 2, enquanto o uso de tom 1 é o padrão para perguntas com alternativas (interrogativas múltiplas) ou perguntas iniciadas com um pronome interrogativo (interrogativas não polares).

**Quadro 5 – Tom 1 em PB e Tom 2 em IA**

PB	IA
//1 <sup>^</sup> o /que que você /acha / <u>disso</u> //	//2 <sup>^</sup> Does that /sound good to / <u>you</u> //
//1 <u>fedem</u> //	//2 They / <u>stink</u> //
//1 <sup>^</sup> e /soltam aqueles cabe/ <u>linhos</u> //	//2 start /sproutin' little white / <u>hairs</u> //
//1 <u>Sim</u> //	//2 <u>aye</u> //

No primeiro caso apresentado no Quadro 5, a escolha pelo tom 1 em PB e pelo tom 2 em IA se deve ao fato de que cada versão apresenta um tipo sintático de interrogativa diferente, portanto um tom diferente também. Em PB, a pergunta foi formada com pronome interrogativo e, em IA, a sintaxe escolhida para a interrogativa é do tipo aberta para resposta sim ou não, logo, sem o uso de pronome interrogativo.

Na segunda linha do Quadro 5, a escolha diferenciada de tons – tom 1, em PB, e tom 2, em IA – pode ser explicada pela mesma razão apontada para o primeiro GT da lista presente no mesmo quadro: diferente forma sintática dos enunciados em cada uma das versões analisadas. Em PB, o GT “e soltam aqueles cabelinhos” está terminando uma lista de possibilidades: “Deixa eles no sol, eles ficam marrons, e soltam aqueles cabelinhos”, nesses casos, o último item listado apresenta um tom descendente, mesmo que esteja no modo interrogativo. Em IA, em que não se utiliza a conjunção “e” entre as orações, a noção de lista não é tão destacável quanto modo interrogativo em “*you leave them out in the sun, they get all brown, start sprouting little white hairs*”. Assim, pode-se dizer que a estrutura sintática foi um agente motivador na adoção do tom 1, em PB, e do tom 2, em IA, no terceiro enunciado apresentado no Quadro 2.

Os enunciados “fedem” e “*they stink*” são pronunciados pelo personagem Burro quando Shrek afirma que os ogros são como cebolas. Em IA, o personagem Burro, ao ouvir a afirmação de Shrek, cheira a cebola na mão do ogro e faz uma pergunta sobre a natureza da comparação entre os dois elementos; há, portanto, um questionamento, que pode ser glosado da seguinte maneira: “Não entendi bem por que os ogros são como cebolas, será que os ogros fedem como a cebola?”. Assim, em IA, há o uso do tom 2, o padrão para a função sintática do GT apresentado: declarativa-interrogativa.

Já em PB, o personagem Burro ouve a afirmação de Shrek de que os ogros são como cebolas e exclama um “Uhm” com o sentido de “Sim”, mostrando compreensão da comparação apresentada, logo após apresenta a próxima fala como uma proposição do tipo de semelhança entre ogros e cebolas a que a personagem principal do filme se refere nesse episódio. Após aceitar a comparação proposta o Burro, a sugestão feita pelo enunciado “fedem” pode ser considerada como: “Está certo, vocês fedem como a cebola, claro!”.

Assim, nota-se que a diferença nas escolhas de tons, nesse caso, é decorrente das diferentes escolhas de interpretação das falas da personagem Burro, de acordo com as possibilidades apresentados pelo contexto, como mostrado no Quadro 6.

**Quadro 6 – Contexto do diálogo entre Burro e Shrek sobre a comparação entre ogros e cebolas**

<b>Shrek</b>	//3 Nós /somos /como ce/ <u>b</u> olas//	//1 ogres /are like / <u>o</u> nions//
<b>Burro</b>	//1 Uhm//	(Sniffs)
	//1 <u>f</u> edem//	//2 They / <u>s</u> tink//
<b>Shrek</b>	//1 <u>S</u> im//	//1 <u>Y</u> es//
	//5 <u>N</u> ão//	//2 <u>N</u> o//
<b>Burro</b>	//1 Fazem vo/cê cho/ <u>r</u> ar//	//1 ^ or they /make you / <u>c</u> ry//
<b>Shrek</b>	//5 <u>N</u> ão//	//2 <u>N</u> o//
<b>Burro</b>	//1 <u>A</u> h//	//3 ^ You /leave them out in the / <u>s</u> un//
	//3 deixa /eles no / <u>s</u> ol//	//3 ^ they /get all / <u>b</u> rown//
	//3 ^ eles /ficam ma/ <u>r</u> rons//	//2 start /sproutin' little white / <u>h</u> airs//
<b>Shrek</b>	//5 <u>N</u> ão//	//1 <u>N</u> o//

Ao observar os GTs dessa passagem do filme, percebe-se que, em PB, há duas exclamações – “Uhm” e “Ah”, que indicam a concordância do personagem Burro com

as afirmações de Shrek. Após as falas de Shrek, o Burro procura demonstrar seu entendimento, a partir de propostas sobre o que seriam as bases da comparação entre ogros e cebolas. No primeiro momento em que ouve a comparação entre ogros e cebolas, o Burro demonstra concordância com a fala “Uhm” e sugere que os ogros e as cebolas fedem. Quando o ogro diz não ao Burro, este sugere que ambos (cebola e ogro) fazem as pessoas chorarem, mais uma vez Shrek descarta essa opção de comparação. O Burro, então, concorda mais uma vez por meio de uma interjeição e faz mais uma proposta, considerando a descrição do apodrecimento de uma cebola.

Por outro lado, em IA, não há a presença de marcadores discursivos de concordância, ao contrário, há enunciados que indicam a incompreensão da personagem Burro com relação à afirmação de Shrek de que ogros são como cebolas; assim, tem-se tom 2 no primeiro e no penúltimo enunciado da seleção. Vale notar que o quarto enunciado foi proferido com o contorno melódico de tom 1, tom padrão de uma alternativa final parte de uma interrogativa múltipla; dessa forma, a escolha pelo tom 1 no quarto enunciado ainda é coerente com a produção de sentido em que as falas do personagem Burro demonstram um sentido de questionamento diante da comparação proposta pelo ogro.

Com relação aos dados acima analisados, vale lembrar que há a interferência da tradução na diferença da entoação das duas versões, em outras palavras, a interpretação dos dubladores diante da situação apresentada foi divergente, assim suas escolhas de tons também não foram as mesmas. É importante destacar que o contexto situacional e linguístico em análise permitiu a interpretação diferente dos dubladores sem comprometimento da coerência do filme *Shrek* (2001).

Enfim, nos casos apresentados anteriormente, a escolha de tons foi responsável pela expressão dos sentidos nos enunciados. Além disso, esses enunciados foram proferidos com o tom padrão para seu tipo sintático dentro de cada sistema linguístico analisado.

No Quadro 7, no entanto, há um enunciado em que a escolha de tom não reflete o tipo sintático da oração. Esses dois enunciados seguem o padrão sintático para escolha de tom 2 para interrogativas, porém essa não é a escolha do dublador norte-americano; assim, pode-se concluir que o tom 1 é aceitável para se fazer perguntas em IA.

**Quadro 7 – Enunciado em IA em que o tom escolhido não segue o padrão sintático**

//2. enten  deu//	//1. You   get it//
-------------------	---------------------

Por um lado, vale notar que há apenas uma ocorrência desse caso no *corpus*, porém a simples presença de um caso mostra que essa escolha é uma opção para o falante de IA ao fazer interrogativas. Por outro lado, é importante considerar que, devido ao *corpus* englobar um número restrito de falantes, não é possível, a partir dos dados acima, traçar generalizações do não uso de tom 1 para interrogativas em PB só pelo fato de que isso não ocorre nos dados analisados.<sup>6</sup> Assim, pode-se apenas afirmar a possibilidade do uso de tom 1 para interrogativas em IA.

Ainda considerando o caso acima, deve-se ressaltar que essa pergunta está inserida no contexto em que a personagem que a profere está buscando tornar clara sua posição de que ogros são como cebolas e seu interlocutor insiste em não compreender a comparação,

<sup>6</sup> O tom 1 é classificado por Cagliari (2007, p. 181) para expressar a função de fala vigorosa, imponente em interrogativas sem palavras interrogativas.

então, o enunciado analisado também expressa um sentido de ordem; tem-se uma pergunta, pois é preciso ver se o interlocutor compreende sua fala, mas também implica o fim da discussão, por isso a escolha por um tom padrão de afirmativas, o tom 1. Em outras palavras, o enunciado “*you get it*” é uma interrogativa com significado de imposição de resposta afirmativa; corroborando com a classificação de Halliday (1970, p. 27) em que o tom 1 para interrogativa implica significado de pergunta vigorosa, imponente ou impaciente.<sup>7</sup>

Em resumo, pode-se afirmar que as diferenças de recortes em GTs do *corpus* nas duas línguas analisadas se devem, em alguns casos, à interferências de tradução, que possibilitaram a inclusão de vocábulos em uma versão e não na outra, ou à produção de enunciados cuja estrutura sintática é díspar nas duas versões. Houve também várias diferenças e semelhanças entre as escolhas de tom nos dados das versões do filme *Shrek* (2001) em PB e em IA. Com relação a isso, uma questão interessante foi a escolha entre tons para cada GT, destacando-se a importância da estrutura sintática dos enunciados e também algumas funções semânticas e pragmáticas exercidas pelos enunciados. Assim, escolhas por diferentes estruturas sintáticas ou por diferentes interpretações de um contexto situacional ou linguístico geram escolhas por diferentes tons.

### Considerações finais

Há várias diferenças e semelhanças entre as escolhas linguísticas da entoação nos dados das versões do filme *Shrek* (2001) em PB e em IA. Uma das diferenças está no número de grupos tonais em cada arquivo de áudio. Outra questão importante são os tons escolhidos para cada grupo tonal, destacando-se a importância dos tipos de sintaxe em que os enunciados são expressos e também os sentidos semânticos e pragmáticos a eles carregados. Há que se ressaltar duas ponderações importantes, todo texto de linguagem humana permite, por um lado, uma abertura a divergentes interpretações e, por outro lado, apresenta uma estrutura na qual os sentidos são construídos (sintaxe, semântica, pragmática). O sistema linguístico de padrões entoacionais não é diferente, há um conjunto de tons possíveis para determinado contexto situacional e linguístico; porém, o falante faz escolhas que enfatizam ou indicam determinados sentidos semânticos e pragmáticos a seus enunciados. No *corpus* em análise, essa natureza variável e estrutural da língua se destaca, pois o filme apresenta um contexto muito específico; porém, aberto a escolhas divergentes entre os atores/dubladores das duas versões, mesmo em um pequeno trecho, como o analisado.

### REFERÊNCIAS

CAGLIARI, L. C. *Elementos de fonética do Português Brasileiro*. 1981. Tese (Livre-docência) – Universidade Estadual de Campinas, Campinas, 1982.

\_\_\_\_\_. *Elementos de fonética do Português Brasileiro*. São Paulo: Paulistana, 2007.

HALLIDAY, M. A. K. *A course in spoken English: Intonation*. London: Oxford University Press, 1970.

<sup>7</sup> No original: *yes/no question: tone 2, neutral; tone 1, strong (forceful or impatient)* (HALLIDAY, 1970, p. 27).

HALLIDAY, M. A. K.; GREAVES W. S. *Intonation in the grammar of English*. London: Equinox Publishing Ltd., 2008.

SHREK. Direção de Andrew Adamson; Vicky Jenson. Produção de Aron Warner; John Williams; Jeffrey Katzenberg. Los Angeles: DreamWorks Picture, 2001. 1 DVD (93 min), son., color.

# **Análise entoacional: uma comparação entre o modelo de Halliday (1970) e Pierrehumbert (1980)**

(Intonation analysis: a comparison  
between Halliday's (1970) and Pierrehumbert's (1980) model)

**Mariane Carvalho<sup>1</sup>**

<sup>1</sup>Universidade Estadual Paulista "Júlio de Mesquita" (Unesp)

mazicarvalho@bol.com.br

**Abstract:** This paper investigates, according to the main methods of intonation description and from the standpoint of perceptual model of functional phonology (HALLIDAY, 1970; CAGLIARI, 1982; 2007) and autosegmental phonology (PIERREHUMBERT, 1980), the relation between the perception and the physical reality of a sound in relation to the phenomenon of intonation. Our main interest is to understand whether the perception corresponds to the acoustic analysis or whether it has its own characteristics. In order to conduct this study, we also want to demonstrate that the correlation between an acoustic analysis and an auditory perception interpretation provides a better explanation of the melodic variation of the language, as well as its relation with others levels of grammar (syntax, semantics and pragmatics).

**Keywords:** intonation; acoustics; auditory analysis.

**Resumo:** O presente trabalho investiga, por meio dos principais métodos de descrição da entoação, do ponto de vista perceptivo do modelo de fonologia funcional (HALLIDAY, 1970; CAGLIARI, 1982; 2007) e da fonologia autosegmental (PIERREHUMBERT, 1980), qual é a relação entre a percepção e a realidade física do som com relação ao fenômeno da entoação. Nosso principal interesse é entender se a percepção corresponde à análise acústica ou se ela apresenta características próprias. A partir disso, também queremos mostrar que a correlação entre uma análise acústica e uma interpretação auditiva perceptiva permite uma melhor explicação da variação melódica da língua, bem como sua relação com outros níveis da gramática (sintaxe, semântica e pragmática).

**Palavras-chave:** entoação; fonologia entoacional; Halliday; Pierrehumbert.

## **Introdução**

Os estudos que se baseiam em análises auditivas, atualmente, estão recebendo muitas críticas. Isso ocorre porque uma grande parte dos foneticistas não admitem a possibilidade de se estudar fonética do ponto de vista auditivo, acreditando apenas em dados estatísticos. No entanto, segundo Cagliari (2012, p. 3), "Os trabalhos mais antigos sobre a entoação foram feitos através de análises auditivas". A percepção da fala é o que motiva o uso da linguagem entre as pessoas, "[...] fundamental para o falante e para o ouvinte, e serve como controlador daquilo que se fala e do que se ouve" (CAGLIARI, 2012, p. 3), de modo que o ouvinte é capaz de perceber variações no tom de voz do falante, identificando padrões prosódicos com significado do tipo ofensivo, irônico, de decepção, que expressem alegria, entre outros. Esse fato mostra o valor dos estudos auditivos e perceptivos nos estudos linguísticos.

Decorre disso o nosso interesse em abordar esse aspecto tão criticado, mostrando que é possível e também viável desenvolver análises que tratem desse fenômeno. Por isso, optamos por fazer um estudo comparativo entre as análises acústica e auditiva, a fim de verificar se a análise acústica se distancia ou não do reconhecimento auditivo.

Nas transcrições auditivas, as análises são menos detalhadas, baseando-se em padrões interpretativos mais gerais. Em compensação, temos a precisão dos dados acústicos. Por isso, é da comparação entre ambos que podemos estabelecer as faixas de frequência que são importantes, pois pequenas variações da frequência fundamental (F0) não devem ser consideradas como relevantes para a análise linguística, uma vez que

Estabelecer esses limites é algo que não se pode fazer apenas estatisticamente, porque há outros fatores que interferem na avaliação da F0 como a tessitura. Tratamentos estatísticos, em geral, evidenciam apenas grandes contrastes ou se aplicam em condições de fala muito controladas. (CAGLIARI, 2010, p. 3)

## **Pressupostos teóricos**

Os estudos sobre a entoação podem ser divididos entre aqueles que foram desenvolvidos na primeira e segunda metade do século passado. No início do século XX, os estudos sobre a variação melódica da fala passaram a ser mais detalhados, a partir da preocupação dos linguistas em descrever as línguas tonais e os padrões entoacionais das demais línguas. O ensino de línguas estrangeiras ajudou a chamar a atenção para a entoação. Na linha americana, segundo Cagliari (2011), os trabalhos vêm desde Bloomfield (1933), Bloch e Trager (1942), Pike (1945) e Trager e Smith (1951). Já na tradição britânica, um dos linguistas mais representativos desse período foi Daniel Jones, seguindo a notação de Paul Passy para marcar a prosódia com curvas, linhas paralelas e bolinhas. Em sua obra *An outline of English Phonetics* (1917, descreveu a entoação do inglês britânico, apresentando dois tons básicos: os padrões de tons descendentes referentes à asserção e os padrões ascendentes das perguntas. Além da variação melódica dos enunciados, através de uma notação especial (tamanho das bolinhas) registrava também as ocorrências da tonicidade e do ritmo.

Posteriormente, Kenneth Lee Pike publica um estudo minucioso, intitulado *The intonation of American English* (1945), descrevendo a entoação do inglês americano. Seu objetivo era descrever o fenômeno e ajudar no ensino de inglês como língua estrangeira. Seu modelo descritivo não indicava a variação melódica de cada sílaba, mas somente onde havia um desvio da altura melódica para cima, para baixo, ou estável em um determinado nível. Nessa obra, Pike (1945) discute também questões teóricas e descritivas do ritmo das línguas, colocando claramente que as línguas ou são de ritmo acentual ou são de ritmo silábico. As línguas de ritmo acentual, como o inglês, marcam as batidas rítmicas nas sílabas tônicas, como se constata na prática poética de versos metrificados na Língua Inglesa. As línguas de ritmo silábico não fazem isso e o acento tônico dos enunciados não serve para compor versos metrificados.

De acordo com Cagliari (2011), foi a partir de Palmer (1922) e, posteriormente, com Kingdon (1958) que a notação entoacional passou a ser feita no próprio texto. Palmer (1922), segundo o autor (CAGLIARI, 2011), também introduziu as noções de *head*,

*nucleus e tail* e Kingdon (1958), as de *head, pré-head e body*. O'Connor e Arnold (1973) apresentaram estudos sobre a entoação em conversas. Para Cagliari (2011), o problema dos estudos de tradição britânica está no fato de eles se preocuparem mais com o ensino da língua inglesa para estrangeiros do que em desenvolver uma teoria mais elaborada do fenômeno e, ainda, pelo fato de terem investido muito pouco em análises acústicas. Uma teoria mais bem formulada aparece, segundo o autor (CAGLIARI, 2011), nos trabalhos de Crystal e Quirk (1964), juntamente com a teoria de Halliday (1967; 1970; 1973).

Entrando na segunda metade do século XX, há a incidência de muitos trabalhos e diferentes pontos de vista sobre os estudos entoacionais. Liberman e Michaels (1962) desenvolvem um estudo mais acústico da entoação, trabalhando, a partir da síntese da fala, a interpretação do ouvinte em relação às variações de atitudes acarretadas pela curva melódica, de acordo com Cagliari (2011). Em seus trabalhos, os autores reconheceram a importância da frequência fundamental (F0), mas viram que ela não é adequada para definir as emoções ocasionadas pela entoação.

Dentre os vários modelos apresentados, que foram propostos e usados tanto em análise auditiva quanto em análise acústica, é difícil avaliar todas as teorias e resultados que cada um deles produz. Como sabemos, as visões e as perspectivas frente ao objeto de estudo são distintas, cada uma com suas peculiaridades. Sendo assim, o trabalho conjunto, assim como revela Cagliari (2012), permite uma melhor compreensão e observação do fenômeno, mesmo diante de algumas aparentes contradições, pois um completa as falhas do outro.

Diante disso, este trabalho irá privilegiar, para a análise auditiva, o modelo teórico proposto por M. A. K. Halliday (1970) e adaptado para o Português do Brasil por Cagliari (2007). Para a análise acústica da entoação, o modelo teórico utilizado será o de Pierrehumbert (versão 1980), construído a partir da teoria métrica de Liberman (1975), Liberman e Prince (1977) e sobre a análise de Bruce (1977) para o acento de palavra do Sueco.<sup>1</sup> A escolha pelos dois tipos de abordagem deve-se ao fato de eles serem, atualmente, os modelos mais desenvolvidos no que diz respeito à descrição da entoação dentro do componente fonológico da gramática funcional e gerativa.

## **Os modelos teóricos: semelhanças e diferenças**

Como este artigo pretende mostrar, a partir de uma análise entoacional, a possibilidade de correlação entre as análises acústicas e auditivas, optamos por apresentar as duas abordagens utilizadas tendo em vista suas semelhanças e diferenças.

No modelo proposto por Halliday (1970) e por Pierrehumbert (1980), a entoação é vista como um fenômeno fonológico.<sup>2</sup> A primeira diferença está em como os autores

1 Vale destacar que há análises entoacionais recentes feitas para o Português do Brasil (PB) que tiveram como base o modelo de Pierrehumbert (1980), como por exemplo, as teses de L. Tenani (2002), F. Fernandes (2007) e M. Córdula. (2012); os artigos de L. Lucente e P. Barbosa (2009) e C. Gonçalves (2009), entre outros.

2 A fonologia gerativa, seguida por Pierrehumbert, parte de uma análise acústica (SPE: *Sound Pattern of English*) e relaciona fatos acústicos com fatos articulatórios e auditivos. Por outro lado, a fonologia funcional sistêmica de Halliday parte da análise auditiva e descreve os fatos fonológicos por oposição sistêmica, no sistema da língua. Quer o modelo de Pierrehumbert, quer o de Halliday têm como ponto mais abstrato e geral uma análise fonológica. Porém, o modo como os valores fonológicos são conseguidos e interpretados

enquadram a entoação. Pierrehumbert (1980) desenvolve seu modelo teórico a partir das ideias propostas pela Gramática Gerativista. O arcabouço de Halliday (1970), por outro lado, é o da gramática sistêmica funcionalista, que entende as funções linguísticas como uma relação entre forma e estrutura.

Para Halliday (1970), não é possível trabalhar só com a fonologia ou só com a fonética. É preciso juntar as duas, pois o trabalho individual leva a distorções, tanto em uma quanto em outra. Por isso, diferentemente de Pierrehumbert, o autor acredita não ser possível descrever a entoação exclusivamente com a análise fonética (auditiva ou acústica), bem como não é possível associar os padrões dos tons às categorias gramaticais diretamente. Isso porque essas descrições são independentes umas das outras. “Ocasionalmente, há uma relação entre categorias gramaticais de outro nível com o sistema entoacional, mas não há nenhuma categoria que exija um padrão entoacional único e exclusivo ou vice-versa” (CAGLIARI, 2011, p. 3). Contudo, conforme Cagliari, a característica abstrata e geral dos padrões entoacionais também é considerada pela autora.

Nesse sentido, Halliday (1970) acredita que a entoação, como um processo gramatical, depende de outros itens. É um processo fonológico que está inserido dentro de um sistema ligado à sintaxe, à pragmática e à semântica. Essa visão vai contra o que Pierrehumbert (1980) faz. Para ela, a entoação não é algo sistemático, mas fluido, enquanto, para ele, tudo na língua é sistemático, com limites controlados de variação. Ambos são sistemas fonológicos, e não um arranjo ocasional de dados. O próprio Halliday justifica seu ponto de vista, considerando que os contrastes fonológicos são tratados por um sistema de termos discretos, aplicando-se tanto para os sistemas fonológicos quanto para os contrastes gramaticais.

Halliday (1970) e Pierrehumbert (1980) começam a descrever seu modelo teórico a partir da caracterização dos dialetos com que vão trabalhar. Halliday, com a entoação do inglês britânico (RP) e Pierrehumbert, com a entoação do inglês americano (*standard*).

Assim como Halliday, Pierrehumbert também quer mostrar, como nota Cagliari (2011), que, em inglês, não há uma relação do contraste de variação melódica com o significado lexical, como ocorre, por exemplo, com o tailandês ou com o vietnamita, ambas línguas tonais. Na visão teórica dos autores, é possível descrever as línguas tonais e entoacionais seguindo o mesmo modelo.

Para a descrição dos padrões melódicos, Halliday e Pierrehumbert partem, inicialmente, da atribuição de valores aos pés rítmicos (para ele) e aos pés métricos (para ela). Em Halliday (1970), os pés estão organizados a partir de sua estrutura sintática. Em seu modelo, o enunciado é dividido em dois componentes, tema e rema, ambos com objetivos informativos. Esses componentes são chamados: componente tônico (CT) e componente pré-tônico (CPT). Pierrehumbert (1980) não fala em níveis gramaticais, mas também separa o enunciado em dois componentes, chamados por ela de componente nuclear e pré-nuclear.

No modelo funcional, a divisão entre componente tônico e componente pretônico é chamada de grupo tonal (GT). São os pés rítmicos que operam nessa separação. A autora, por sua vez, não está preocupada com essa nomenclatura e nem sequer a menciona

---

varia em uma e em outra abordagem.

em sua tese de doutoramento. No entanto, podemos comparar o GT de Halliday, em Pierrehumbert, ao domínio do contorno entoacional, ou seja, ao sintagma entoacional (*intonational phrase*).

No sistema entoacional de Halliday, os cinco níveis entoacionais são definidos pela configuração do contorno melódico do componente tônico. Já em Pierrehumbert, essa anotação é feita a partir da sílaba tônica frasal e do *final boundary stress*.

Quanto à atribuição de tonicidade, Halliday e Pierrehumbert classificam as sílabas entre tônicas e átonas. Ambos desconsideram as sílabas átonas na análise entoacional. Para saber onde está localizada a sílaba de maior proeminência, chamada por ela de *pitch accent*, a autora parte da estrutura métrica do enunciado. O autor, por sua vez, a nomeia de sílaba tônica saliente (STS), definindo-a a partir do foco entoacional do enunciado, ou onde ocorre a mudança mais notável da direção do contorno entoacional.

Para a marcação dos tons dos padrões entoacionais, Halliday propõe o uso de cinco tons melódicos (alto, meio-alto, médio, meio-baixo, baixo). Pierrehumbert trabalha somente com os tons altos, simbolizados por H (*high*) e baixos, simbolizados por L (*low*), que são gerados por um sistema fonológico, como parte da gramática da língua. Em outras palavras, há regras para a formação desses tons, chamadas de *mapping rules*.

O não reconhecimento de um tom intermediário, nesse caso, o médio, traz alguns problemas para o modelo teórico da autora pois, algumas vezes, ela acaba traduzindo um tom médio como um tom alto ou baixo, o que não ocorre na abordagem de Halliday.

Para o autor, o sentido gramatical do padrão entoacional é obtido por meio da curva melódica do componente tônico. Para a autora, esse padrão é dado pelo que ocorre entre o *pitch accent* e o *boundary tone*.

Quanto ao significado entoacional, Pierrehumbert faz a sua descrição da entoação a partir da observação das características da frequência fundamental (F0). A F0 presente entre os tons é caracterizada pela autora por meio de regras fonéticas, tendo em vista quais são esses tons e como eles se relacionam no tempo e na frequência. Feita essa observação, o passo seguinte seria comparar os usos do padrão de F0 que são fonologicamente distintivos. Porém, a autora não segue adiante, parando no primeiro método. Ela acredita que o sistema de investigação fonológica sobre o significado entoacional seja útil, mas não desenvolve essa questão em seu trabalho. Em alguns casos, ela faz uma descrição mais superficial do significado ou do uso de algum contorno específico. Mas isso, segundo Pierrehumbert, apenas para ajudar o leitor a entender o contorno que está em discussão. Para ela, não interessa se a frase é afirmativa ou interrogativa, mas sim o que o falante fez e não o que ele significou. Por isso é que ela observa a curva entoacional, fazendo uma relação entre os fatos acústicos e fonológicos. Tal atitude teórica é decorrente do fato de Pierrehumbert trabalhar no interior da gramática gerativa (CHOMSKY, 1965), a qual separa o componente fonológico do sintático e do semântico, cada qual operando com regras próprias.

Em contrapartida, Halliday (1970) faz uma relação entre os padrões entoacionais e as atitudes do falante. Diferentemente de Pierrehumbert, seu objetivo é descrever como o ouvinte percebe a entoação. Para ele, o significado de um tom está diretamente vinculado com outras escolhas gramaticais em um enunciado.

O modelo de Pierrehumbert (1980) exige uma marcação para o final do enunciado. Halliday também considera o contexto terminal do enunciado como sendo muito importante para a descrição entoacional. Para isso, através de símbolos visuais e linhas horizontais, marca o movimento da altura melódica para cada pé no final do enunciado. Assim como Halliday, Pierrehumbert também admite pausas entre as *intonationals phrase*.

Em Halliday, conforme Cagliari (2011), as pausas são incorporadas na descrição do ritmo e podem ocupar lugares silábicos nos pés ou, até mesmo, na duração de pés inteiros. As pausas ocorrem por uma questão sintática e semântica. Pierrehumbert, por sua vez, não leva esse ponto em consideração.

Em relação à tessitura, Halliday mostra que esse efeito iguala realizações fonéticas diferentes no sistema, “[...] mas que apresentam um contraste estabelecido como sendo o mesmo no nível fonológico” (CAGLIARI, 2011, p. 5). Nesse sentido, conforme Cagliari, uma sílaba pode apresentar uma variação da altura melódica de acordo com o valor da frequência fundamental, mas permanecer sendo alta ou baixa, por exemplo. Segundo o autor, isso ocorre porque, mesmo com a variação acústica, o que caracteriza um tom não é a realização fonética de um enunciado, mas o valor fonológico atribuído a ele.

O efeito da tessitura em Pierrehumbert pode ser comparado com a sua discussão sobre as regras de *downstep*. De acordo com Cagliari (2011), as variações no valor de H e L são um fato fonético e não fonológico, e a tessitura é um fato fonético e fonológico. Entretanto, “[...] sua função não é re-ordenar os valores da F0, mas indicar marcas discursivas, como enunciados parentéticos ou enfatizados (por exemplo, em altura melódica de falseto)” (CAGLIARI, 2011, p. 6).

Outro aspecto que coloca as duas metodologias em confronto é a questão de uma trabalhar com a questão acústica da entoação (Pierrehumbert) e a outra trabalhar com a questão auditiva da entoação (Halliday). Para a autora, sua opção em representar e mapear a entoação a partir da realização do contorno da frequência fundamental (F0) tem uma importância particular, devido às falhas apresentadas pelos outros métodos ao fazerem uma representação mais abstrata da entoação, embora a descrição auditiva não seja exatamente uma interpretação mais abstrata.

Resumindo, podemos dizer que as diferenças mais significativas entre as duas perspectivas estão relacionadas aos seguintes pontos: Pierrehumbert descreve a entoação a partir de uma análise fonética. O autor, por sua vez, trabalha com a união da fonética e da fonologia. O sistema entoacional de Halliday é formado por cinco níveis entoacionais que são definidos a partir da configuração do contorno melódico do componente tônico. Em Pierrehumbert há dois níveis, anotados de acordo com a sílaba tônica frasal e do final *boundary stress*. Para a anotação da sílaba tônica saliente, a autora recorre à estrutura métrica do enunciado e o autor, ao foco do enunciado. Outra diferença está no fato de Halliday fazer uma relação entre os padrões entoacionais e as atitudes do falante, e Pierrehumbert, não. Por fim, a autora trabalha em sua análise com a questão acústica da entoação e ele, com a questão auditiva.

Em relação aos pontos de encontro, os dois autores mostram que não existe uma relação entre o contraste da variação melódica com o significado lexical. Os dois modelos atribuem valores aos pés, dividindo-os em dois componentes. As sílabas são classificadas entre tônicas e átonas nas duas análises, sendo as sílabas átonas desconsideradas. A

marcação para o final do enunciado é importante para a descrição entoacional tanto em Halliday quanto em Pierrehumbert. Finalmente, as duas abordagens admitem pausas nos enunciados e levam em consideração o efeito da tessitura em sua descrição entoacional.

A comparação acima mostrou apenas alguns aspectos essenciais dessas metodologias. Mas, mesmo assim, é possível chegar à conclusão de que os dois modelos não são tão distantes assim pois, apesar das diferenças, as semelhanças são muitas. Como afirma Cagliari (2011, p. 17):

Embora os modelos sejam diferentes formalmente, quando comparados, eles são surpreendentemente semelhantes ou, pelos menos, apresentam uma ponte que permite passar uma análise realizada em um modelo para uma análise no outro modelo.

Portanto, mesmo com as diferenças e singularidades de cada modelo, é perfeitamente possível juntá-los e correlacioná-los em uma análise entoacional. Esse fato também mostra a viabilidade do nosso trabalho, ou seja, abre espaço para a interpretação da entoação através da correlação entre uma análise acústica e uma auditiva da entoação.

## **Materiais e métodos**

Os dados do presente estudo fazem parte da leitura de um pequeno trecho do livro *História sem fim*, de Michael Ende (s/d). A opção por analisar um fragmento de texto e não frases soltas está relacionada ao fato de que as últimas, às vezes, dão margem para uma intuição do falante diferente da esperada em experimentos programados com a investigação de fatos específicos, dificultando a interpretação semântica dos enunciados. Também optamos por não estudar a fala comum devido às suas dificuldades, como, por exemplo, a falta de funções comunicativas, interativas e emotivas, causadas por gravações de pesquisa. Em seguida, o material sonoro selecionado foi lido por um homem de meia idade, falante do dialeto paulista da região de Araraquara e com instrução superior.

Em relação à análise acústica, o texto foi segmentado em trechos correspondentes aos grupos tonais, para facilitar a análise acústica, que não comporta longos trechos gravados. Os dados tiveram uma análise acústica processada através do programa PRAAT, desenvolvido por Paul Boersma e David Weenink, do Institute of Phonetic Sciences (University of Amsterdam), disponível gratuitamente em <http://www.fon.hum.uva.nl/praat/>. A gravação do *corpus* foi processada em um ambiente acusticamente adequado, via microfone unidirecional, modelo Microsoft LifeChat LX-3000.

O texto foi segmentado em 15 enunciados correspondentes aos grupos tonais. Também foi feita a transcrição fonética<sup>3</sup> e ortográfica de cada um dos enunciados. Em seguida, as frases foram descompactadas e subdividas em três níveis: 1) segmentos, 2) sílabas e 3) palavras. Esse método foi feito a partir do *TextGrid* do PRAAT e por meio da observação das informações carreadas pelo espectrograma.

Para a análise do *pitch*, foi marcado o valor da frequência fundamental (F0) para cada sílaba, localizando-a no centro da duração das vogais, isto é, a partir do seu ponto medial. Para a realização dessa etapa, foi preciso trabalhar com dois métodos diferentes: 1) um referente às sílabas vozeadas e 2) outro para as sílabas desvozeadas. Para as sílabas vozeadas, utilizamos a função CTRL + 0 do PRAAT. Com esse comando, foi possível

---

3 A transcrição fonética seguiu o modelo do IPA.

analisar, com pequenas variações, o ponto medial, que selecionado vai pelo cruzamento de zero mais próximo do ponto médio do segmento. O ponto medial estabelecido depende de como o início e o fim do segmento foi determinado. Essa segmentação é feita através de dois procedimentos:

- (1) através de uma observação auditiva, levando em consideração as possibilidades articulatórias e as realizações fonéticas da língua;
- (2) através de uma análise visual, que interpreta parâmetros acústicos atribuíveis ao segmento, às suas transições e aos vizinhos (transições).

Para as sílabas desvozeadas, foi feita a seleção da sua parte vozeada e, em seguida, utilizamos a função: *Picth > get picth* do *software*. Não foi possível utilizar o mesmo método porque, muitas vezes, o ponto médio mais próximo do cruzamento de zero estava na vogal desvozeada.

É importante salientar a possibilidade de se mediar a frequência fundamental (F0) por meio de outros métodos, como por exemplo, através da análise da fronteira do segmento. Porém, entendemos que, nesse momento, há alguns fatores que influenciam e dificultam a segmentação, como por exemplo, a co-articulação. Segundo Cagliari (1980), o que ocorre nessa etapa são modificações fonéticas de vizinhança, devido à interferência das características articulatórias de um som sobre outro. Essas influências podem ocorrer tanto no início quanto no fim de um elemento, o que é natural da fala. Por isso, optamos por não fazer anotações pelo início ou fim de fronteira, uma vez que poderia haver uma interferência nos resultados da frequência fundamental (F0).

## Análise dos dados

### O padrão entoacional dos enunciados: uma comparação

A partir do fragmento de texto,<sup>4</sup> foram analisadas 15 frases. Seguindo o modelo de Halliday (1970), encontramos 22 grupos tonais, de tons primários e secundários. Os tons encontrados referem-se aos padrões entoacionais característicos, para esse falante em cada enunciado. Esses tons estão distribuídos no texto da seguinte forma:

1+	1	1	+1	1	1-	3	1-
1+	1	:1	1	1	3	1-	
3+	-1	3	1	+1	3	1	

Estabelecido o padrão entoacional característico de cada tom de acordo com o modelo funcional, convertemos a forma fonológica obtida para a de Pierrehumbert. Isso foi feito a fim de estabelecermos uma padronização entre os tons e verificar a aproximação ou não entre os dois modelos.

Assim, os padrões alcançados (segunda coluna) para esses tons foram:

---

4 O fragmento de texto analisado aparece em anexo neste artigo. A lista de tons apresentada é referente à descompactação do texto em 15 enunciados e separação dos grupos tonais (GTs). A análise das frases, com a respectiva divisão dos GTs não foi apresentada por motivos de espaço.

**Tabela 1: Conversão dos tons primários e secundários de acordo com o modelo de Pierrehumbert**

Tipos de tons	Forma fonológica no modelo de Pierrehumbert
Tom 1 (primário) <sup>5</sup>	H+L* L L%
Tom 3 (primário)	L+H* H H <sup>L</sup> %
Tom 1+ (secundário)	L+H* L L% ou H+L* L L%
Tom +1 (secundário)	H+L* L L% ou L+H* L L%
Tom 1- (secundário)	L+H* L L% ou L* L L%
Tom -1 (secundário)	L* L L%
Tom :1 (secundário)	L+H* L L%
Tom 3+ (secundário)	H* H H <sup>L</sup> %

A conversão mostrada acima foi feita a partir da análise das frases do texto, conforme os dois modelos teóricos. Primeiro, fizemos a transcrição da curva entoacional (obtida pelo PRAAT). Depois, foi feita a análise seguindo a metodologia proposta nas duas teorias. Para Pierrehumbert, primeiro, de acordo a análise da grade métrica e, em seguida, por meio da observação das características da F0, com respectiva anotação dos eventos tonais. Para o modelo de Halliday e Cagliari foi feita a divisão dos pés, dos GTs, da STS e anotação auditiva dos limites da altura melódica para a atribuição dos tons aos enunciados. Por fim, foi feita a correspondência entre os dois modelos. É importante salientar que os padrões fonológicos obtidos para o modelo de Pierrehumbert tiveram como ponto de partida para a sua transcrição o *nuclear pitch accent*. Isso porque, para a autora, o que vem antes do acento nuclear não é relevante para a análise. O que interessa é o alvo, isto é, o final do movimento a partir da sílaba mais proeminente, servindo como ponto de ancoragem. Também é necessário destacar que, para Tom 3, seja ele primário ou secundário, propomos para a análise um tom L sobrescrito ao lado do tom H de fronteira, já que é comum em PB que toda a curva melódica ascendente acabe com um valor de Hz um pouco mais baixo.

Como o procedimento realizado para se chegar aos resultados mostrados na tabela acima é muito extenso, apresentaremos aqui apenas o modo como a transposição foi feita para o tom 1.

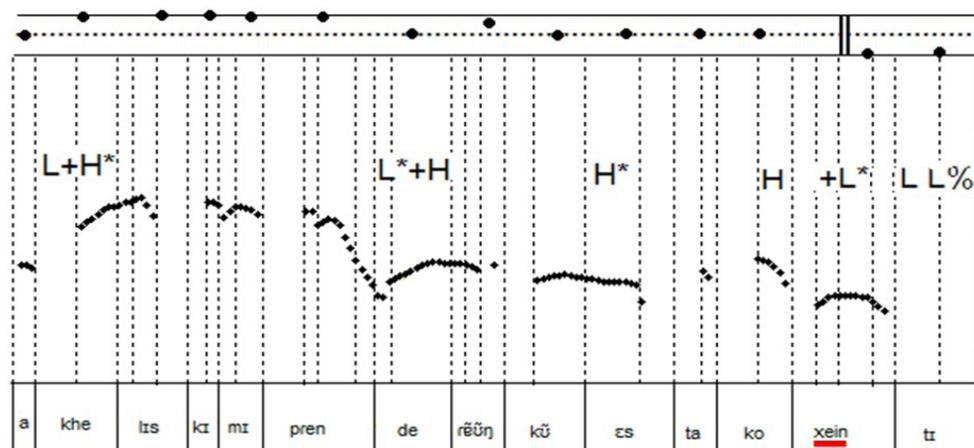
Para todos os grupos tonais identificados com o tom 1 primário, a curva melódica foi classificada com um padrão descendente em nossos dados. No modelo de Halliday, esse mesmo tom caracteriza-se por ter um movimento tônico descendente com tendência

5 Os diacríticos de + e – que aparecem em sobrescrito ao lado dos tons referem-se aos tons secundários propostos no modelo de Halliday (1970) e Cagliari (1982; 2007), variantes dos tons primários (que não levam essa marcação). Os tons secundários são caracterizados por trazerem uma conotação semântica mais forte do que o tom primário.

6 Em relação à notação de Pierrehumbert, o tom alto (H) ou baixo (L) que traz o símbolo de asterisco (\*) representa a sílaba tônica frasal, ou seja, o foco entoacional do enunciado. Já o símbolo de porcentagem (%) ao lado de cada unidade tonal (H% ou L%) representa os tons de fronteira e estão associados ao final da frase entoacional.

terminal baixa ou média baixa, bem como em Cagliari (2007). Ao fazermos a análise segundo o modelo de Pierrehumbert, o padrão mostrou uma F0 que se inicia mais alta e, em seguida, mantém-se em queda na posição final do enunciado, caracterizando também um padrão entoacional descendente.

Veja um exemplo da configuração da curva melódica no sentido da queda da Hz, para o modelo de Pierrehumbert<sup>7</sup> e de acordo com os pontos auditivos propostos para a anotação de Halliday.



**Figura 1. Exemplo de sentença com padrão entoacional descendente segundo os padrões de Halliday (1970) /Cagliari (1982; 2007) e Pierrehumbert no enunciado: “Aqueles que me prenderam com esta corrente”**

Fonte: Elaboração própria.

Na figura acima, está representado, no final do enunciado, o padrão obtido para as sentenças declarativas neutras em nosso *corpus*, tanto para a descrição de Pierrehumbert quanto para a de Halliday/Cagliari. O exemplo mostra a tendência à declinação da curva melódica para esse tipo de sentença nos dois modelos teóricos. Os outros oito grupos tonais, de padrão declarativo, também seguiram esse mesmo tipo de contorno entoacional para o componente tônico. Com os outros tipos de tons também foi possível chegar a uma correspondência entre as duas abordagens teóricas. Os resultados obtidos entre as formas fonológicas dos dois modelos entoacionais, mostram, portanto, a aproximação entre uma análise auditiva ao lado de uma análise acústica.

### Interpretação da pauta entoacional em valores acústicos

Também é possível mostrar que há uma relação entre as análises acústicas e auditivas, através da interpretação da variação melódica da F0 em função dos cinco tipos de tons da pauta entoacional usada pelo modelo de Halliday (1970) e Cagliari (1982; 2007).

<sup>7</sup> Em sua tese de doutorado, Pierrehumbert ilustra todas as possibilidades de combinação entre o *pitch accent*, *phrase accent* e *boundary tone* para o final do enunciado. A combinação proposta por ela resultou em um total de 22 diferentes padrões entoacionais e a forma fonológica obtida nesse trabalho está entre eles. Para conferir veja Pierrehumbert (1980, p. 399).

Nessa análise, os resultados obtidos mostraram que para esse informante, a tessitura variou de 80cps a 192cps, aproximadamente. Veja abaixo:

**Tabela 2: Intervalo de frequência que caracteriza os tons, dada a tessitura desse informante e a interpretação auditiva dos tons**

Baixo	Meio baixo	Médio	Meio alto	Alto
entre 79,9 e 110,0 cps	entre 110,2 e 124,3 cps	entre 124,3 e 141,7	entre 142,7 e 150,0	entre 150,7 e 191,9

Os valores descritos na Tabela 2 mostram que há uma correspondência bem clara de frequências acústicas para cada faixa de tom, tendo em vista a interpretação auditiva e, com isso, válida a utilização de análises auditivas. É claro que essa correspondência não foi unânime para todas as faixas de frequência. Por exemplo, nos 23 tons baixos encontrados em nossos dados, a variação foi de três tons, de modo que dois deles foram classificados no intervalo de frequência dos tons meio-baixo e o outro, na faixa de frequência do tom médio. Quanto aos outros tons, a variação foi de um tom para o meio-baixo e baixo. Para tom alto, por sua vez, não houve nenhum caso de variação. No entanto, essas diferenças são justificáveis e não indicam uma falha na análise auditiva ou sequer na análise acústica. Pois, auditivamente, esses resultados que fogem do padrão são irrelevantes no conjunto da frequência. O ouvido não considera as diferenças entre esses valores como sendo importantes, uma vez que eles ocorrem como variação da curva melódica que não é *stead state*, ou seja, não tem valores fixos, mas pequenos altos e baixos, que são homogeneizados pelo ouvido. Assim, esses resultados são interpretados em função do contexto em que ocorrem, considerando a tendência da curva melódica.

### **Análise semântica e sintática dos enunciados<sup>8</sup>**

Por meio do estudo semântico e sintático dos enunciados é possível observar a relação da entoação com a gramática da língua, uma vez que todo padrão entoacional ou contorno melódico liga-se a algum significado semântico e pragmático, a alguma estrutura sintática e, especialmente, aos atos de fala e às atitudes do falante.<sup>9</sup>

Para tanto, a análise foi feita a partir do ponto de vista auditivo (HALLIDAY, 1970; CAGLIARI, 1982; 2007) e acústico (PIERREHUMBERT, 1980). Nos enunciados (cf. as figuras 1 e 2 abaixo), a tonicidade foi atribuída a partir da análise auditiva da fala como um todo e não por meio da leitura individual de cada palavra. Os enunciados foram segmentados em sílabas, de modo que o intervalo entre uma sílaba tônica e outra constitua um pé rítmico (ABERCROMBIE, 1967, p. 96-98). A divisão em pés é marcada com barras inclinadas simples (/) e cada pé apresenta pelo menos uma sílaba acentuada (ou tônica) e um número relativo de sílabas não acentuadas. As figuras mostram também a análise prosódica. Os grupos tonais (GTs) são marcados com barras inclinadas duplas (//) em seu início e fim. Um GT caracteriza-se por apresentar um ou mais pés (P) e por ter uma sílaba tônica saliente (STS), que representa o foco. A STS está destacada por dois

8 A análise semântica e pragmática feita para esses enunciados não está relacionada aos elementos formais das frases, mas sim a uma interpretação dos padrões entoacionais que, são responsáveis, entre outras coisas, por estruturar o conteúdo dos enunciados e por ter um papel na realização semântica de atos de fala (CAGLIARI, 2007).

9 Como dissemos, a descompactação do texto resultou em 15 enunciados. Para esse artigo, trazemos somente a análise de dois deles, devido à extensão do trabalho e ao limite de páginas permitido.

traços que a sublinham e ocorre imediatamente após as barras verticais duplas (||). As sílabas tônicas silenciosas aparecem representadas pelo diacrítico (.) e ocorrem imediatamente após os números que indicam o tipo de tom do enunciado.

As três linhas horizontais representam os limites da altura melódica para a marcação dos tons, ou seja, o valor auditivo relativo da variação melódica que foi atribuído a cada sílaba. Essa representação auditiva é feita por meio de bolinhas cheias. Abaixo dessas três linhas, mostra-se o valor obtido para a F0 em números. Feita a anotação da F0, segue a análise acústica a partir do modelo autossegmental de Pierrehumbert (1980). Abaixo dessa transcrição destacamos o contorno da curva melódica de F0 para cada enunciado. Esse processo foi feito a partir do *praat picture*.

Logo abaixo da representação do contorno entoacional, aparece a transcrição fonética acompanhada da transcrição ortográfica. Os tons dos padrões entoacionais vêm assinalados com os respectivos números.

A escolha dos tons relacionou-se com os tipos sintáticos de sentença como: declarativas, interrogativas, exclamativas, entre outras; com os atos de fala, ou seja, se as frases indicam ordem, pedido, sugestão, etc.; e com as atitudes do falante, segundo seu valor semântico e pragmático, como: expressões de surpresa, polidez, indiferença, destaque, entre outras. O número que aparece do lado esquerdo de cada figura, acima das três linhas horizontais, indica o número do enunciado, obtido a partir da descompactação do texto. Veja a seguir os dois enunciados, com suas respectivas análises:

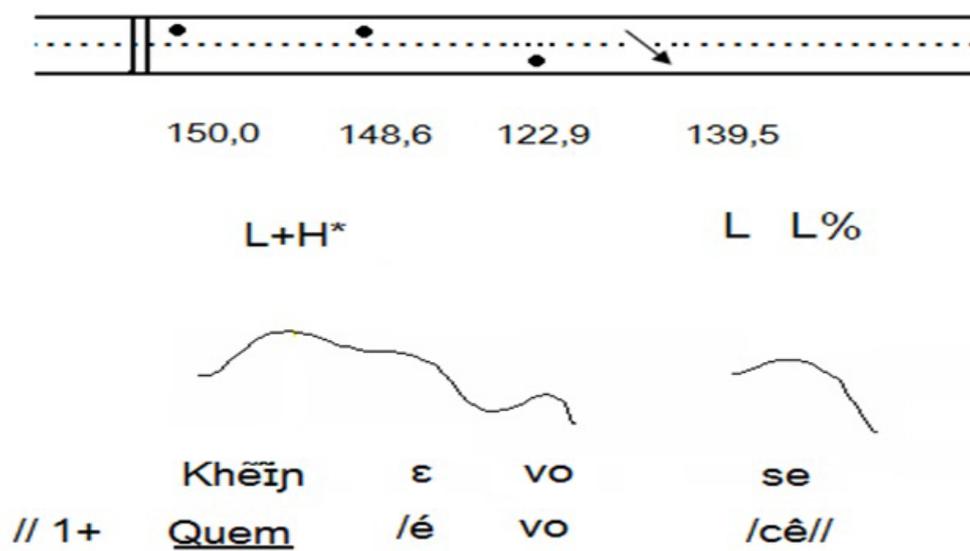


Figura 2. Análise semântica e pragmática do enunciado “Quem é você?”

No enunciado acima, o GT “Quem é você?” foi anotado com o tom 1+ secundário. Esse tom caracteriza-se por ter uma altura melódica (meio-alta) mais alta do que a do tom 1 primário, o qual se apresenta no nível médio constante. Nesse enunciado, a variação melódica começa em um nível meio-alto e, depois, segue com um padrão descendente até acabar em um nível baixo, com uma pequena ênfase na sílaba tônica (se), assim como indicado pela flecha. Não há variação do CPT, somente do CT.

Esse padrão entoacional caracteriza-se por ter uma estrutura sintática interrogativa de palavra interrogativa, *Quem* que, por sua vez, é o foco do enunciado. O fato de o foco estar no início do GT gera um valor de curiosidade à pergunta, o personagem está curioso e deseja saber quem é essa pessoa desconhecida e que surgiu do nada. É uma curiosidade que vem marcada também por uma precaução, por uma desconfiança. Por isso, uma tessitura inicial que se inicia mais alta e termina mais baixa.

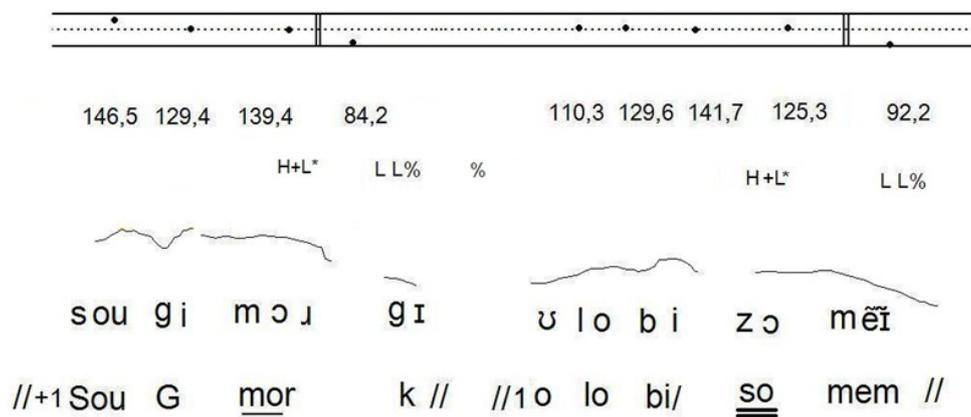


Figura 3. Análise semântica e pragmática do enunciado “Sou Gmork, o lobisomem”

E2, por sua vez, apresentou dois GTs, um de tom +1 secundário “Sou Gmork” e o outro de tom 1 simples, “o lobisomem”, os dois tons em sentenças declarativas. O tom +1 começa mais alto (meio-alto) do que o tom 1 simples do segundo GT, que se inicia em um nível médio. Esse tom também se caracteriza por ser uma variação do CT do seu tom primário correspondente. Nesse GT, houve um erro de leitura, mostrado na transcrição fonética. O leitor trocou a consoante (k) do final da palavra GmorK pela consoante (g), o que ocorre, às vezes, em leituras.

A tessitura começa com uma altura melódica mais alta no início do enunciado e depois começa a diminuir. Isso é próprio dos valores semânticos observado nos dois tons. No primeiro GT, a resposta é mais enfática. É como se o falante, interpretando a fala do personagem, quisesse sanar a curiosidade gerada pela pergunta anterior além de dar um valor mais irônico a sua resposta, isto é, o falante esperava que o seu interlocutor já soubesse quem era ele: “Como assim, quem sou eu? Sou Gmork, o lobisomem”. Depois disso, a expectativa do primeiro momento passa, caracterizando a tessitura mais baixa e a leitura mais neutra e pouco enfática do segundo GT.

Como se pode ver, a curva melódica acaba com uma altura melódica bem baixa no primeiro GT e, no segundo GT, começa em um nível médio. Acusticamente, o final do primeiro GT e o início do segundo são baixos, mas auditivamente tiveram classificações diferentes, um anotado como tom baixo, assim como os outros tons desse tipo e o outro como médio. Isso ocorreu no segundo GT porque há uma estrutura típica de platôs dando a sensação auditiva de que todos os tons estão na mesma altura, ou seja, auditivamente não houve uma queda na F0 ou na tessitura. Esse padrão também prova a neutralidade da fala nesse GT. Já o fato de a curva melódica não ter se mantido constante entre um GT e o outro se deve à ocorrência de uma pausa (que está representada pelo símbolo de %) entre

eles, o que também evidencia a mudança no padrão entoacional e, conseqüentemente, no diferente significado semântico das duas unidades.

## Conclusões

No decorrer do trabalho, constatamos a possibilidade de se estabelecer uma correlação entre as análises acústicas e auditivas, já que a análise acústica acoplada à análise auditiva ajuda a conciliar os detalhes da fala com as unidades fonológicas do sistema da língua.

A comparação entre o modelo auditivo e acústico mostrou que, mesmo sendo distantes teoricamente um do outro, é possível estabelecer um paralelo entre eles, o que também foi comprovado quando fizemos a correspondência entre os dois sistemas fonológicos. Além disso, a análise da tessitura também mostrou a possibilidade dessa correlação, uma vez que, estabelecidas as faixas de frequência, auditiva (alturas melódicas) e acústica (análise da F0), observamos a equivalência entre os dois modelos. As variações foram pouco expressivas e não foram significativas para inviabilizar os resultados. Elas apenas mostraram os ajustes que o ouvido faz durante a percepção da curva entoacional.

A breve análise semântica e sintática dos enunciados mostrou também a possibilidade em se definir linguisticamente os padrões entoacionais a partir da gramática da língua, revelando as intuições dos falantes e ouvintes.

## REFERÊNCIAS

- ABERCROMBIE, D. *Elements of general phonetics*. Edinburgh: Edinburgh University Press, 1967.
- BRUCE, G. *Swedish word accents in sentence perspective*. Travaux de L'Institut de Linguistique de Lund, CWK Gleerup: Mamö, 1977.
- BLOCH, B.; TRAGER, G. L. *Outline of linguistics analysis*. Baltimore, MD: Waverly Press, 1942.
- BLOOMFIELD, L. *Language*. New York: Holt, Rinehart & Winston, 1933.
- BOERSMA, P.; WEENINK, D. Praat: doing phonetics by computer (Version 5.1) [Computer Program]. 2009. Disponível em: <[http://: www.praat.org/](http://www.praat.org/)>. Acesso em: 10 jun. 2012.
- CAGLIARI, L. C. *A entoação no componente fonológico da gramática*: Halliday (1970) e Pierrehumbert. 2012. p. 1-18. ms.
- \_\_\_\_\_. *A entoação no componente fonológico da gramática*. 2011. p. 1-3. ms.
- \_\_\_\_\_. *Comparação entre interpretação auditiva e dados acústicos na análise entoacional*: estudo baseado na leitura de um trecho de uma história. 2010. p. 1-25. ms.

\_\_\_\_\_. *Elementos de fonética do português brasileiro*. São Paulo: Paulistana, 2007.

\_\_\_\_\_. *Elementos de fonética do português Brasileiro*. 1982. 192f. Tese (Livro de graduação em Linguística) – Instituto de Estudos Linguísticos, Universidade Estadual de Campinas, Campinas.

\_\_\_\_\_. A entoação do português brasileiro. *Estudos Linguísticos*, Araraquara, n. 3, p. 308-329, 1981.

CÓRDULA, M. S. M. *Análise fonético-fonológica dos padrões entoacionais do Português Brasileiro e do Inglês Norte-americano no filme Shrek (2001)*. 2012. 234f. Tese (Doutorado em Linguística) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara.

CHOMSKY, N. *Aspects of the Theory of Syntax*. Cambridge: The MIT Press, 1965.

CRYSTAL, D.; QUIRK, R. *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton, 1964.

ENDE, M. *História sem fim*. São Paulo: Martins Fontes, [s.d.].

FERNANDES, F. R. *Ordem, focalização e preenchimento em português: sintaxe e prosódia*. 2007. 445f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Unicamp, Campinas.

GONÇALVES, C. A. V. Uma abordagem autosegmental para a morfologia. *Cadernos de Letras da UFF – Dossiê: Difusão da língua portuguesa*, n. 39, p. 211-232, 2009.

HALLIDAY, M. A. K. *Intonation and Grammar in British English*. Janua Linguarum, Series Practica 48. The Hague: Mouton, 1967.

\_\_\_\_\_. *A course in spoken English: Intonation*. London: Oxford University Press, 1970.

\_\_\_\_\_. The Tones of English. In: JONES, W. E.; LAVER, J. (Ed.). *Phonetics in Linguistics*. London: Longman, 1973. p. 103-126.

JONES, D. *An outline of English phonetics*. Cambridge: Cambridge University Press, 1918.

KINGDON, R. *The roundwork of English intonation*. London: Longmans, 1958.

LIBERMAN, M. *The intonation system of English*. PhD thesis, MIT 1975. Distributed by IULC, 1978. Published by Garland Press: New York, 1985.

LIEBERMAN, P.; MICHAELS, M. Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *Journal of the Acoustical Society of America*, v. 34, n. 7, p. 922-927, 1962.

LIBERMAN, M.; PRINCE, A. *On stress and linguistic rhythm*. Linguistic Inquiry, n. 8, p. 249-336, 1977.

LUCENTE, L.; BARBOSA, P. A. The role of alignment and height in the perception of LH contours. *Speech Prosody*, 2010, Chicago: Proc. of the Speech Prosody 2010, p. 1-4.

O'CONNOR, J. D.; ARNOLD, G. F. *Intonation of colloquial English*. London: Longman, 1973.

PALMER, H. E. *English intonation, with systematic exercises*. Cambridge: Heffer, 1922.

PIERREHUMBERT, J. *The Phonology and phonetics of English intonation*. 1980. 402f. Tese (Doutorado em Linguística) – Department of linguistics and philosophy, Massachusetts Institute of technology, Indiana University Linguistics Club.

PIKE, K. L. *The intonation of American English*. Ann Arbor: The University of Michigan Press, 1945.

TENANI, L. E. *Domínios prosódicos no Português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos*. 2002. 331f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Unicamp, Instituto de Estudos da Linguagem, Campinas.

TRAGER, G. L.; SMITH, H. L. *An outline of English structure*. Washington: American Council of Learned Societies, 1951.

## ANEXO

Sou GmorK, o lobisomem

Por que está preso aqui? Esqueceram-se de mim quando se foram embora

Quem? Aqueles que me prenderam com esta corrente

E para onde foram? Gmork não respondeu. Olhou Atreiu com os olhos meio fechados.

Depois de um longo silêncio, disse: Você não é daqui, pequeno estrangeiro, não é desta cidade nem deste país. O que procura? Atreiu baixou a cabeça.

# Adaptação de empréstimos de itens lexicais no papiamentu moderno

(Loanwords adaptation of lexical items in modern papiamentu)

Manuele Bandeira<sup>1</sup>, Shirley Freitas<sup>2</sup>, Ana Livia Agostinho<sup>3</sup>

<sup>1,2,3</sup> Faculdade de Filosofia, Letras e Ciências Humanas – Universidade de São Paulo

manuelebandeira@usp.br, shirleyfreitas@usp.br, ana.agostinho@usp.br

**Abstract:** In this study, we investigate phonological processes of nativization in modern papiamentu economic lexicon. Papiamentu is an Iberian-based creole spoken in Aruba and Netherlands Antilles by about 200 million people. Corpus of this study was formed by words incorporated into papiamentu in the second half of the 20<sup>th</sup> century. We investigate phonological and morphological alterations in the lexical items of languages loaned to the target language (papiamentu) and we account segmental, syllabic and word levels. We conclude that the economic lexicon is adapted following papiamentu linguistic patterns, such as *atministrá* < *administrar* “to manage”, in which the consonant sonority loss in syllabic coda is triggered by a phonotactic restriction that prevents voiced obstruent codas.

**Keywords:** Papiamentu; phonological adaptation; economics.

**Resumo:** No presente estudo, foram investigados processos fonológicos de nativização/ adaptação do léxico do mundo da economia no papiamentu moderno. O papiamentu é um crioulo de base ibérica falado em Aruba e nas Antilhas holandesas por cerca de 200 mil pessoas. O corpus deste trabalho foi formado por palavras incorporadas ao papiamentu a partir da segunda metade do século XX. Foram investigadas alterações fonológicas e morfológicas nos itens lexicais das línguas emprestadoras para a língua-alvo (o papiamentu), considerando a adaptação no nível segmental, silábico e no nível da palavra. Constatou-se que o léxico do campo da economia é nativizado segundo o padrão linguístico do papiamentu, como se vê em *atministrá* < *administrar* “administrar”, em que a perda de sonoridade da consoante na coda silábica é engatilhada por uma restrição fonotática que impede codas obstruintes sonoras.

**Palavras-chave:** papiamentu; adaptação fonológica; economia.

## Introdução

Este estudo busca investigar os processos morfológicos e fonológicos na adaptação/ nativização de palavras originadas de outras línguas (sobretudo do holandês, inglês, português, espanhol<sup>1</sup>) no papiamentu moderno com os seguintes objetivos: (i) avaliar se os ‘empréstimos’ são nativizados segundo o padrão linguístico do papiamentu ou se há uma gramática especial para essas palavras (PARADIS; LABEL, 1994; PARADIS, 1986; KENSTOWICZ, 2001; KENSTOWICZ; SUCHATO, 2004); (ii) identificar os padrões fonológicos e morfológicos existentes na adaptação de empréstimo no papiamentu.

<sup>1</sup> Nem sempre é possível definir se uma palavra é de origem portuguesa ou espanhola no Papiamentu, assim preferiu-se denominar como de origem ibérica. Segundo Birmingham (1970 apud LIPSKI, 2008), cerca de sessenta por cento do léxico do Papiamentu tem origem no Espanhol ou no Português. Desta forma, *hasi* ‘fazer’, *bieu* ‘velho’, *mucho* ‘muito’, *traha* ‘trabalhar’ são de origem espanhola, ao passo que *bai* ‘vai’, *trese* ‘trazer’, *nasementu* ‘nascimento’ e *papia* ‘falar’ são de origem portuguesa. Contudo, há palavras como *frio* ‘frio’, *largu* ‘longo’ e *boka* ‘boca’, que podem ser originárias de qualquer uma dessas línguas ibéricas.

A partir de um *corpus* composto por itens incorporados a partir da segunda metade do século XX, analisar-se-ão palavras pertencentes ao campo lexical da Economia. Por Economia, as seguintes acepções foram aceitas como critérios avaliativos: “[...] ciência que estuda os fenômenos relacionados com a obtenção e a utilização dos recursos materiais necessários ao bem-estar; conjunto de disciplinas constituintes do curso de nível superior que forma economistas.” (HOUAISS; VILLAR, 2001; verbete *economia*). Serão investigadas alterações morfológicas e fonológicas nos itens lexicais das línguas emprestadoras para a língua-alvo (o papiamentu). Tal estudo é importante, pois a nativização de palavras de origem estrangeira é, como afirmam Calabrese e Wetzels (2009), um *locus* privilegiado para a observação do funcionamento do sistema linguístico da língua receptora. Ao investigar como o papiamentu adapta palavras emprestadas, é possível também analisar como a dada língua funciona, sob uma abordagem linguística sincrônica.

O *corpus* utilizado foi composto por palavras retiradas de fontes diversas, tais como: jornal impresso *Èxtra*, produzido e vendido em Curaçao; guias e livros sobre economia (HEILIGERS-HALABI, 1988; SINDIKATONAN, 1985; KOMERSIO, 1986; HOYER, 1944); *dicionário bilíngue papiamentu/inglês* (RATZLAFF-HENRIQUEZ, 2008). A metodologia para essa análise foi baseada em padrões de correspondência de som entre palavras do papiamentu e palavras de empréstimo de outras línguas. Nesse estudo, dois objetivos foram traçados: o primeiro foi estabelecer algumas regras que palavras emprestadas parecem estar sujeitas quando entram no léxico do papiamentu; o segundo, e não menos importante, tratou-se de identificar de que língua tal palavra foi emprestada para o papiamentu. Uma parcela dos dados e a discussão empreendida serão expostas nesse artigo.

Para análise, serão consideradas apenas palavras nativizadas em meados do século XX, isto é, serão observados apenas empréstimos recentes. Assim, o *corpus* conterà preferencialmente itens lexicais relacionados à Economia, por supor que a constituição desse campo lexical no papiamentu ocorreu, de modo geral, há pouco tempo. Além disso, será observado qual o papel da **Teoria de Restrições e Estratégias de Reparo (TCRS)** (PARADIS, 1988) – **Princípio da Preservação**, bem como será observada também a atuação das regras de conversão (ANDERSEN, 1974) sobre as adaptações fonológicas e morfológicas feitas nas palavras de empréstimo recente no papiamentu.

## **Papiamentu: características gerais**

O papiamentu é um crioulo de base ibérica, falado em Aruba e nas antigas Antilhas Holandesas por cerca de 200 mil pessoas. Ele é uma das línguas oficiais das ilhas de Aruba, de Bonaire e de Curaçao (ABC). O papiamentu também convive com o holandês e com outras línguas não oficiais, como o espanhol e o inglês. Para se ter uma ideia mais exata de tal afirmação, da população total de 208.000, aproximadamente um sexto é falante nativo de outras línguas, como o holandês, o inglês e o espanhol (HOLM, 2000). A proximidade com a América do Sul faz com que o léxico do papiamentu seja influenciado pelo espanhol americano e receba ainda a influência do português do Brasil (sobretudo graças à forte influência comercial brasileira, uma vez que Curaçao é um dos maiores entrepostos comerciais do Caribe e parceiro privilegiado do Brasil).

Além de ser falado nas ilhas ABC, o papiamentu também é falado nas ilhas de Saba, Saint Eustatius (Santo Eustáquio) e Saint Maarten (São Martinho). Em termos demográficos,

Curaçao é a ilha que possui a maior população (142.000<sup>2</sup>), seguida de Aruba (108.000) e, por último, Bonaire (13.000). Em Saba, Saint Eustatius e Saint Maarten, o inglês é a língua oficial (e a mais falada), ao lado do holandês e do papiamentu.

Basicamente, há três hipóteses no que diz respeito à gênese do papiamentu. A primeira afirma que ele pode ser resultado de uma relexificação de um proto-crioulo afro-português, como acreditam Lenz (1928) e Martinus (1996). A segunda hipótese supõe que o papiamentu era genuinamente um crioulo de base portuguesa, formado em Curaçao no final do século XVII, devido à chegada de judeus sefarditas, falantes do português, provenientes do nordeste brasileiro, acompanhados, por sua vez, de escravos (SMITH, 1999). Finalmente, a terceira hipótese defende que o papiamentu era originalmente um crioulo de base espanhola, cujos elementos portugueses foram trazidos pelos judeus sefarditas (Cf. MUNTEANU, 1996).

Para Mcwhorter (1995), não é possível ter havido crioulos de base espanhola no Atlântico. Segundo o estudioso, o papiamentu não pode possuir tal base, uma vez que os espanhóis, por estarem limitados pelo Tratado de Tordesilhas, não puderam criar condições sócio-históricas necessárias para o surgimento de línguas crioulas no Atlântico Ibérico. Já Lipsky (2008) argumenta que a influência atual da língua espanhola no papiamentu torna difícil investigar com exatidão a fonte dos vocábulos portugueses no papiamentu.

Como se pôde observar, há centros de desacordo com relação à origem do vocabulário inicial ibérico presente no papiamentu. Como muitos estudiosos não sabem ao certo se é um crioulo de base portuguesa ou espanhola, prefere-se denominar o papiamentu como um crioulo de base “ibérica”.

### **Adaptação de empréstimos no papiamentu**

Os métodos dessa análise são baseados em padrões de correspondência de som entre palavras do papiamentu e palavras de empréstimo de outras línguas. Assim, o objetivo primordial da análise é tentar estabelecer algumas regras que palavras emprestadas pareçam estar sujeitas quando entram no léxico do papiamentu. Ao analisar empréstimos, tal estudo também busca identificar de que língua tal palavra foi emprestada para o papiamentu. Tarefa de difícil execução, uma vez que a semelhança de um mesmo vocábulo em duas línguas diferentes nem sempre significa empréstimo como afirma Viaro (2011, p. 98): “Dadas duas línguas quaisquer, se um elemento de seu vocabulário é parecido ou idêntico, tanto no significante, quanto no significado, isso pode dever-se basicamente a três fatores distintos: coincidência, empréstimo ou origem comum.”

Assim, tendo em vista que os falantes do papiamentu conviveram e ainda convivem em um ambiente multilinguístico, estando as seguintes línguas muito presentes: holandês, espanhol, inglês e português, sendo a última língua mais recorrente no passado histórico da ilha, a análise buscou correspondências de som entre palavras dessa língua crioula e as línguas supracitadas. É importante salientar que nem sempre apenas esse procedimento é suficiente para afirmar com precisão a língua emprestadora de cada item lexical, assim o estudo admite, de antemão, possíveis falhas e equívocos quanto a essa questão.

Antes que os processos fonológicos e morfológicos comecem a ser analisados, é preciso que algumas definições sejam dadas. No presente trabalho, o termo ‘empréstimo’,

2 Dados fornecidos pelo Central Bureau of Statistics, Antilhas Holandesas, 2009.

como definido em (01), é aqui empregado como item(ns) lexical(is) introduzido(s) na língua alvo (L1) por falantes nativos que têm acesso à língua emprestadora (L2). A nativização ou a adaptação dos empréstimos de L2 em L1 são regidas por padrões fonológicos de L1, padrões impostos pelos falantes de L1 (PARADIS; LABEL, 1994).

- (01) **Empréstimo:** uma palavra simples ou composta, ou uma sentença oriunda de L2, incorporada ao discurso de L1.

Em relação aos empréstimos, Paradis (1996) afirma que os falantes de L1 tendem a interpretar a estrutura de L2 conforme a estrutura de L1. Além disso, os falantes de L1 frequentemente descartam das palavras incorporadas via empréstimo informações, contidas em L2, percebidas como redundantes do ponto de vista de L1. No percurso de L2 (espanhol, português, holandês e inglês) até L1 (papiamentu) ocorre uma série de modificações, uma vez que a estrutura linguística dessas línguas é diferente.

Para Calabrese e Weltzels (2009), a análise de empréstimos em uma língua nos oferece uma janela direta tanto para o estudo de como os signos acústicos são categorizados em termos de traços distintivos relevantes para o sistema fonológico de L1; como para o estudo da fonologia sincrônica verdadeira de L1 ao observar seu processo fonológico em ação.

Kenstowicz (2003) afirma que segmentos ou estruturas salientes tendem a ser preservados em processos de adaptação e, quando reparos são necessários, modificações serão evitadas o quanto possível. Paradis (1996) defende um princípio chamado **Princípio de Preservação**, em (02). Tal Princípio estabelece que a informação segmental não é arbitrariamente destruída (PARADIS, 1996):

- (02) **Princípio da Preservação:** A informação segmental é maximamente preservada, dentro do Princípio do Limiar (Threshold Principle).

O **Princípio do Limiar**, em (03), postula que todas as línguas impõem um limite para preservação do segmento, tal princípio não excede dois processos<sup>3</sup> dentro de um dado domínio de restrição.

- (03) **Princípio do Limiar**

- (a) Todas as línguas têm um limiar de tolerância para a preservação do segmento;  
(b) Esse limiar é estabelecido em dois passos (ou dois reparos) dentro de um dado domínio de restrição (PARADIS; LA CHARITÉ, 1997).

Além disso, há também a **Estratégia de Reparo**, definida em (04):

- (04) **Estratégia de Reparo:** Operação fonológica não contextual e universal que é engatilhada por uma violação de uma restrição fonológica, e que insere, apaga segmento ou estrutura para garantir a conformidade com a restrição (PARADIS, 1996).

O Princípio de Preservação (02) se aplica ao papiamentu, pois em palavras, fruto de empréstimo, que possuem estrutura silábica e fonemas semelhantes à língua crioula,

<sup>3</sup> O presente estudo defende que existam estratégias de reparo e o subsequente limiar para preservação de um segmento ou o seu apagamento, contudo, tal limiar não seria limitado a dois estágios obrigatoriamente, mas pode demandar um número superior ao citado. Rose (1999) afirma que falantes de lama, por exemplo, fazem uso de estratégias de reparo, contudo, ele é feito em três estágios, diferente do previsto por Paradis e La Charité (1997).

nota-se que o vocábulo se manteve bastante similar ou mesmo idêntico, como se pode ver nos exemplos:

- (05) *Agrikultor* (papiamentu) < *Agricultor* (“ibérico”) “Agricultor” – Alteração ortográfica (<k> <c>)
- (06) *Logístiko* (papiamentu) < *Logístico* (“ibérico”) “Logístico” – Alteração ortográfica (<k> <c>)
- (07) *Finansiamentu* (papiamentu) < *Financiamento* (“ibérico”) “Financiamento” – Alterações ortográficas (<s> <c>) (<u> <o>)
- (08) *Tarifa* (papiamentu) < *Tarifa* (“ibérico”) – Nenhuma alteração
- (09) *Aporte* (papiamentu) < *Aporte* (“ibérico”) – Nenhuma alteração

Nos itens lexicais analisados, o apagamento ocorreu com menor frequência, contudo, de forma bastante regular e com a mesma finalidade: evitar a violação de uma restrição na L1. Em todos os verbos no infinitivo e nomes formados pelo nominalizador deverbal **-dor**, todos de origem “ibérica”, houve o processo de apócope, subtração de um segmento sonoro final de palavra, nesse caso o **-r** final, como se observa a seguir:

- (10) *Ajustá* (papiamentu) < *Ajustar* (espanhol) “Ajustar”
- (11) *Desaroyá* (papiamentu) < *Desarollar* (espanhol) “Desenvolver”
- (12) *Atministrá*<sup>4</sup> (papiamentu) < *Administrar* (“ibérico”) “Administrar”
- (13) *Presupuestá* (papiamentu) > *Presupuestar* (espanhol) “Orçamentar”
- (14) *Probechá*<sup>5</sup> (papiamentu) > *Provechar* (espanhol) “Lucrar”
- (15) *Probechadó* (papiamentu) > *Probechador* (espanhol) “Beneficiador”

Outra restrição no papiamentu está relacionada à impossibilidade de haver obstruintes sonoras em posição de coda. Assim, diante dessa restrição, o papiamentu faz uso das obstruintes surdas como forma de adaptação, como se pode observar nos exemplos seguintes:

- (16) *Atministrá* (papiamentu) < *Administrar* (“ibérico”) “Administrar”
- (17) *Kontabilidat* (papiamentu) < *Contabilidad* (espanhol) “Contabilidade”
- (18) *Supsidio* (papiamentu) < *Subsídio* (“ibérico”) “Subsídio”

Um processo bastante recorrente na análise dos dados foi o de palatalização, sobretudo nos contextos em que o segmento fônico /s/ estava diante da vogal anterior /i/. Assim, por um processo de palatalização, **-si** torna-se **-sh** ([ʃ]). Como se pode observar nos exemplos seguintes:

- (19) *Privatisashon* (papiamentu) < *Privatización* (espanhol) “Privatização”
- (20) *Adishonal* (papiamentu) < *Adicional* (“ibérico”) “Adicional”
- (21) *Inflashon* (papiamentu) < *Inflación* (espanhol) “Inflação”

4 Além da apócope, nota-se a dessonorização do segmento consonantal *t* < *d*.

5 Nesse vocábulo, há também o processo de betacismo (*b* < *v*).

(22) *Inflashonista* (papiamentu) < *Inflacionista* (espanhol) “Inflacionária”

(23) *Negosiashon* (papiamentu) < *Negociación* (espanhol) “Negociação”

Em alguns vocábulos, foi possível observar que, além da palatalização ([ʃ] < [si]), houve também o apagamento do segmento no final de palavra (apócope). Tal apagamento foi acionado como mais uma regra de conversão que, segundo Andersen (1974), busca converter itens lexicais que não qualifiquem como “papiamentu nativo” em formas que qualifiquem. Assim, no papiamentu, não há ditongos formados por **-io** e **-ia**. Diante disso, o segundo segmento do ditongo na palavra da língua emprestadora sofre apagamento, como se pode observar a seguir:

(24) *NegoshiØ* (papiamentu) < *Negócio* (“ibérico”) “Negócio”

(25) *GanashiØ* (papiamentu) < *Ganancia* (“ibérico”) “Ganho, benefício”

Segundo Andersen (1974), haveria regras de conversão que adaptam itens lexicais que violem a restrição de mesma altura para formas que não violem. Por exemplo, a palavra do espanhol /su'dor/ tornou-se em papiamentu /so'dɔ/, nota-se que o papiamentu assimila a vogal para altura da vogal acentuada seguinte, contudo, os sons ficaram próximos em relação à altura, mas não idênticos (o ≠ ɔ). Segundo Viaro (2011, p. 179), na assimilação, “sons distintos aproximam seus pontos articulatórios ou acabam por tornar-se idênticos.” Na análise de palavras de adaptação recente, observa-se a aplicação dessa regra de conversão, em que a altura da vogal da sílaba acentuada é assimilada pela vogal que a antecede, como se pode ver em (26).

(26) *Fondo* (Papiamentu) < *Fundo* (“ibérico”) “Fundo”

## Processos segmentais regulares em palavras recentes

Embora a presente pesquisa não tenha analisado todos os itens pertencentes ao *corpus*, é possível depreender padrões fono e morfológicos de adaptação de empréstimos.

Nas subseções seguintes, serão vistas, primeiramente, as adaptações realizadas no **nível segmental**, tais como: *ensurdecimento*, *palatalização* e *betacismo*. Em seguida, serão vistos processos que ocorrem no **nível silábico**, como *aférese*. Posteriormente, analisar-se-ão os processos fonéticos que ocorrem no **nível da palavra**, tais como a *ressilabificação*.

### Ensurdecimento

A transformação de consoantes originalmente sonoras em surdas, ou seja, o ensurdecimento ou dessonorização, mostrou-se bastante regular nos itens analisados, uma vez que no papiamentu, há a restrição de obstruintes sonoras em posição de coda. Diante disso, ocorre a adaptação para um segmento surdo (ver exemplos).

	ESPAÑHOL	PAPIAMENTU	GLOSA
(27)	<i>Administración</i>	<i>Atministrashon</i>	“Administração”
(28)	<i>Administrativo</i>	<i>Atministrativo</i>	“Administrativo”

Na impossibilidade de haver obstruintes sonoras ([d], [b], [v]) em posição de coda, o papiamentu adapta tais segmentos de L2 através do uso de obstruintes surdas ([t], [p], [f]). É importante dizer que não é a sonoridade da consoante no onset seguinte que influencia o processo. A seguir, encontram-se exemplos dessa adaptação no papiamentu:

#### Coda Medial

	L2	PAPIAMENTU
(29)	<i>Subsídio</i> (“ibérico”)	<i>Sup.si.dio</i>
(30)	<i>Administrável</i> (português)	<i>At.mi.nis.tra.bel</i>

#### Coda final

	L2	PAPIAMENTU
(31)	<i>Contabilidad</i> (espanhol)	<i>Kon.ta.bi.li.dat</i>

Como esperado, a limitação se restringe apenas à sonoridade, pois não há violação de restrição em palavras emprestadas com obstruinte surda na coda, como se pode ver em **Kòrf.bal** que é proveniente do holandês *Korfbal* (“Corfebol, esporte coletivo praticado na Holanda”). Neste exemplo, percebe-se que a obstruinte surda ([f]) foi mantida na mesma posição.

### Palatalização

Um processo de transformação bastante recorrente na análise dos dados foi o de palatalização, sobretudo nos contextos em que o segmento fônico /s/ estava diante da vogal palatal anterior /i/. Assim, por um processo de palatalização, **-si** torna-se **-sh** ([ʃ]) nos seguintes exemplos:

	ESPAÑHOL	PAPIAMENTU	GLOSA
(32)	<i>Ac[si]ón</i>	<i>Ak[ʃ]on</i>	“Ação em bolsa de valores”
(33)	<i>Fluctua[si]ón</i>	<i>Fluktua[ʃ]on</i>	“Flutuação”
(34)	<i>Nego[si]ar</i>	<i>Nego[ʃ]á</i>	“Negociar”
(35)	<i>Privatiza[si]ón</i>	<i>Privatisa[ʃ]on</i>	“Privatização”

Apenas no vocábulo **finansiá** “financiar”, o processo de palatalização não ocorreu como em **negoshá** “negociar”. Embora, nesse item, tenha havido o contexto que favorece a palatalização (sibilante e vogal anterior alta [si]), não houve a mudança para a palatal [ʃ].

### Betacismo

O betacismo é um fenômeno conhecido como a alternância de pronúncias entre o [b] e [v]. Inicialmente, ao se observarem itens como **buelo** “voo” (< *vuelo*, espanhol), **desponibel** “disponível” (< *disponível*, “ibérico”), **nobo** “notícias” (< *novo*, “ibérico”), pode-se supor que tais palavras seriam exemplos dessa alternância de [b] e [v]. Entretanto, ao se ter acesso às pronúncias do espanhol venezuelano<sup>6</sup>, observa-se que todas as palavras

6 Considera-se o espanhol venezuelano, em detrimento das outras variedades do espanhol, devido à contiguidade e à constante presença de venezuelanos em Curaçao.

citadas possuem, onde se encontra o grafema <v>, a fricativa bilabial sonora ([β]). Assim, o que se poderia considerar como betacismo nada mais é que uma alternância grafemática (<b> < <v>).

	ESPAÑHOL	PAPIAMENTU	GLOSA
(36)	<i>Go[β]ernador</i>	<i>Go[b]ernador</i>	“Governador”
(37)	<i>Go[β]ernamental</i>	<i>Go[b]ernamental</i>	“Governamental”
(38)	<i>[β]uelto</i>	<i>[b]uèltu</i>	“Troco”
(39)	<i>[β]enda</i>	<i>[b]enta</i>	“Venda”

### Aférese

A aférese é um processo de subtração de um segmento fônico ou, às vezes, de sílabas inteiras no início dos vocábulos. Tal processo incide principalmente sobre vogais átonas, geralmente em sílabas sem coda (VIARO, 2011, p. 139). Ao se analisarem os itens seguintes, nota-se que as palavras adaptadas para o papiamentu que sofreram aférese possuem as características descritas por Viaro (2011): são palavras iniciadas por vogais átonas e núcleo de sílabas sem coda, com a exceção de **stropiá** de *estropiado*, fonte “ibérica”:

	ESPAÑHOL	PAPIAMENTU	GLOSA
(40)	[a]provechar	[Ø]probechá	“Lucrar”
(41)	[a]postar	[Ø]pustá	“Apostar”
(42)	[a]provechado	[Ø]probechadó	“Aquele que lucra”

### Ressilabificação

Segundo Andersen (1974), algumas regras de conversão no papiamentu parecem eliminar encontros consonantais não permitidos ou produzir encontros consonantais mais nativos. No caso da palavra **bankrut** adaptada do holandês *bankroet* “falência”, observou-se que o encontro consonantal [kr] foi desfeito. Embora tal encontro seja permitido no papiamentu, essa adaptação mostra que a estrutura escolhida parece ser mais usual para a língua recipiente.

	HOLANDÊS	PAPIAMENTU	GLOSA
(43)	ban[kr]oet	ban[k].[h]ut	“Falência”

Diante da análise dos dados, pode-se confirmar que a Teoria de Restrições e Estratégias de Reparo (TCRS), proposta por Paradis (1988), possui um importante papel no entendimento do processo de adaptação fonológica no papiamentu. Tal teoria, contudo, não é a única a explicar alguns processos fonológicos significativos na análise dos dados da referida língua. É possível notar a aplicação e funcionalidade das regras de conversão descritas por Andersen (1974), regras essas que visam adaptar itens lexicais não-nativos para nativos no papiamentu.

## Considerações finais

O papiamentu oferece uma visão privilegiada para linguistas que buscam investigar processos fonológicos e morfológicos em uma língua crioula, haja vista que a referida língua, diferentemente da maioria das línguas crioulas de base portuguesa do Atlântico goza de bastante prestígio social, sendo usada por todos os segmentos da sociedade em Curaçao. Tal prestígio não é encontrado, por exemplo, pelo santomé, língua crioula de base portuguesa de São Tomé e Príncipe, pontualmente evitada pela elite falante do português. Outro fator que faz do papiamentu uma língua com traços sociais peculiares, sem dúvida, é o fato de ela ser também língua oficial em seu país e de possuir um sistema de escrita padrão, veiculado em meios de comunicação e arte, tais como revistas, jornais e livros de literatura.

O processo de nativização está relacionado à adaptação de um item lexical na língua-alvo, L1, por indivíduos que estão expostos à língua emprestadora, L2, e é regido por restrições fonológicas de L1 conforme foi discutido ao longo do presente estudo. A partir da análise de palavras de empréstimos de línguas estrangeiras do campo da Economia no papiamentu, tal constatação foi reafirmada, uma vez que os vocábulos analisados também foram nativizados segundo o padrão linguístico do papiamentu.

Considerando a Teoria de Restrições e Estratégias de Reparo (TCRS), proposta por Paradis (1988), confirmou-se que a informação segmental tende a ser maximamente preservada. Portanto, o apagamento não foi o recurso mais usado, ocorrendo apenas quando era necessário evitar a violação de uma restrição na L1.

Além de analisar a adaptação fonológica sob a ótica da TCRS, é possível também considerar o papel das regras de conversão descritas por Andersen (1974) acerca dos processos linguísticos básicos que produziram o léxico do papiamentu nativo, sobretudo nos anos iniciais dessa língua (período de 1650 a 1750). Nota-se que tais regras ainda se aplicam nas palavras de adaptação recente no papiamentu.

Percebe-se também que a nativização de palavras de origem estrangeira é, como afirmam Calabrese e Wetzels (2009), um *locus* privilegiado para a observação do funcionamento do sistema linguístico da língua receptora. Ao estudar como o papiamentu adapta palavras emprestadas, é possível também analisar como a dada língua funciona em ação, sob uma abordagem linguística mais sincrônica.

Sobre os padrões de adaptação encontrados, observou-se que, em todos os verbos no infinitivo e nomes formados pelo nominalizador deverbal **-dor**, sobretudo de origem “ibérica”, houve o processo de apócope, subtração de um segmento sonoro final de palavra, nesse caso o **-r** final. Além disso, o papiamentu parece não permitir que fricativas surdas sejam sonorizadas (/v, z, ʒ/), tanto em posição de onset, como em coda. Assim, uma Estratégia de Reparo é engatilhada através da substituição da série sonora de L2 por um segmento surdo em L1 (v > f) como forma de adaptação.

Durante a análise, observou-se uma proibição de obstruintes sonoras em posição de coda (*administrar* > **atministrá**). Outro processo comum foi a palatalização nos contextos em que o segmento fônico /s/ ou /t/ estava diante da vogal anterior /i/ ([si] vira [ʃ]), [ti] torna-se [tʃ]).

Além da palatalização, ocorreu também o apagamento do segmento no final de palavra. Tal apagamento foi acionado, uma vez que, no papiamentu, não há ditongos formados por **-io** e **-ia**. Por essa razão, o segundo segmento do ditongo na palavra da língua emprestadora sofre apagamento (*negócio* > **negoshi**).

De todo modo, pode-se afirmar que a discussão teórico-metodológica desta análise não se esgota, de maneira alguma, aqui. Como próximo passo para este estudo, tem-se o objetivo de analisar mais vocábulos, fruto de empréstimo, em outros domínios lexicais, também formados recentemente. Serão analisados itens lexicais relacionados à tecnologia (incluindo a informática), ao esporte, entre outros. Dessa maneira, será possível verificar se os padrões de adaptação são semelhantes em todos os domínios lexicais ou se há variações de adaptações a depender do campo e/ou da língua fonte (L2).

## REFERÊNCIAS

ANDERSEN, Roger William. *Nativization and Hispanization in the Papiamentu of Curaçao, N.A.: a Sociolinguistic Study of Variation*. 1974. Tese (Doutorado em Filosofia) – Universidade do Texas, Texas, Austin.

CALABRESE, Andrea; WETZELS, Leo. *Loan Phonology*. Amsterdam: John Benjamins Publishing Company, 2009.

CENTRAL Bureau of statistics, *Antilhas Holandesas*. 2009. Disponível em: [http://www.cbs.an/population/population\\_b2.asp](http://www.cbs.an/population/population_b2.asp). Acesso em: 14 mai 2011.

HEILIGERS-HALABI, Bernadette. *Guia pa empresa chikí*. Kòrsou: Fundashon Empresa Chikí Kòrsou i Korpodeko, 1988. 55 p.

HOLM, John. *An Introduction to Pidgins and Creoles*. Cambridge: Cambridge University Press, 2000.

HOUAISS, Antônio; VILLAR, Mauro de Salles. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Editora Objetiva, 2001.

HOYER, John. *A Little Guide English – Papiamento*. Curaçao: Boekhandel Bethencourt Heerenstraat, 1944.

KENSTOWICZ, Michael. *Saliency and similarity in loanword adaptation: A case study from Fijian*. Unpublished manuscript, MIT. 2003. Retrieved August 11, 2005. Disponível em: <<http://web.mit.edu/linguistics/www/kenstowicz/kenstowicz-'03.pdf>>. Acesso em: 6 maio 2012.

\_\_\_\_\_. The role of perception in loanword phonology. *Studies in African Linguistics*, Portland, n. 3, p. 95-112, 2001.

KENSTOWICZ, Michael; SUCHATO, Atiwong. Issues in Loanword Adaptation: a Case Study from Thai. Manuscript. 2001. The role of perception in loanword phonology. *Studies in African Linguistics*, Portland, n. 32, p. 95-112, 2004.

- KOMERSIO, Bon. *Mas Trabou. Miho Kòrsou*. Kòrsou: Kámara di Komersio i Industria, 1986.
- LENZ, Rodolfo. *El Papiamentu: la lengua criolla de Curazao*. Santiago de Chile: Balcels & Cia, 1928.
- LIPSKI, John M. Spanish-Based Creoles in the Caribbean. In: KOUWENBERG, Silvia; SINGLER, John Victor (Ed.). *The Handbook of Pidgin and Creole Studies*. Oxford: Wiley-Blackwell, 2008. p. 543-564.
- MARTINUS, Frank. *The kiss of a slave: Papiamentu's West African connections*. Amsterdam: Universiteit van Amsterdam, 1996.
- MCWORTHER, John. The scarcity of Spanish-based creoles explained. *Language in society*, n. 24, p. 213-244, 1995.
- MUNTEANU, Dan. *El papiamentu, lengua criolla hispánica*. Madrid: Gredos, 1996.
- PARADIS, Carole. On constraints and repair strategies. *The Linguistic Review*, n. 6, p. 71-97, 1988.
- \_\_\_\_\_. The inadequacy of filters and faithfulness in Loanword Adaptation. In: DURAND, Jacques; LAS, Bernard (Org.). *Current trends in phonology*. Salford: University of Salford Publications, 1996.
- PARADIS, Carole; LABEL, Caroline. Contrasts from Segmental Parameter Settings in Loanwords: Core and Periphery in Quebec French. Proceedings of the MOT Conference on Contrasts in Phonology. *Toronto Working Papers in Linguistics*, n. 13, p. 75-94, 1994.
- PARADIS, Carole; LA CHARITÉ, Darlene. Preservation and minimality in loanword adaptation. *Journal of Linguistics*, n. 33, p. 379-430, 1997.
- RATZLAFF-HENRIQUEZ, Betty. *Dikshonario Papiamentu-Ingles/Ingles-Papiamentu*. Bonaire: Jong Bonaire, 2008.
- ROSE, Yvan. Perception, representation and correspondence relations in loanword phonology. In: CHANG, Steve S.; LIAW, Lily; RUPPENHOFER, Josef (Ed.). *Proceedings of the twenty-fifth annual meeting of the Berkeley Linguistic Society*. Berkeley, CA: Berkeley Linguistics Society, 1999. p. 383-394.
- SINDIKATONAN. *Na kaminda pa un era nobo: Alternativa sosio-ekonómiko di sindikatonan uni di Korsou*. Curaçao: SUK, 1985.
- SMITH, Norval. Pernambuco to Surinam 1654-65? The Jewish slave controversy. In: HUBER, Magnus; PARKVALL, Mikael. *Spreading the word: the issue of diffusion among Atlantic creoles*. London: University of Westminster Press, 1999. p. 251-298.
- VIARO, Mario Eduardo. *Etimologia*. São Paulo: Contexto, 2011.

# Ausência de assimilação de vozeamento no português como L2 por anglófonos – uma análise via Teoria da Otimidade

(Absence of voice assimilation in L2 Portuguese by native speakers of English – an analysis within Optimality Theory framework)

Miley Antonia Almeida Guimarães<sup>1</sup>

<sup>1</sup> Faculdade Filosofia, Letras e Ciências Humanas – Universidade de São Paulo (USP)

miley.guimaraes@gmail.com

**Abstract:** This paper aims to analyse within Optimality Theory framework (PRINCE; SMOLENSKY, 1993; MCCARTHY; PRINCE, 1995) the lack of regressive voicing assimilation in Portuguese as a second language by Anglophone speakers. It was demonstrated that the time of residence in Brazil was the most significant factor for the production of the voiced alveolar fricative – although variation persisted even within the data from more experienced learners, who had been living in the country for over ten years. Analysing the data within Optimality Theory constraint-based model, it was possible to attest the interaction between markedness effects and transfer of the English constraint ranking to the learners' interlanguage.

**Keywords:** Portuguese as a second language; segmental phonology; Optimality Theory.

**Resumo:** Este artigo trata da análise, via Teoria da Otimidade (PRINCE; SMOLENSKY, 1993; MCCARTHY; PRINCE, 1995), da ausência de assimilação regressiva de vozeamento no português como segunda língua por aprendizes anglófonos. Verificou-se que o tempo de residência no Brasil foi o fator mais significativo para a produção do vozeamento da fricativa alveolar, embora a variação tenha persistido mesmo nos dados de falantes mais experientes, com tempo de residência no país superior a dez anos. Pela análise por meio do modelo baseado em restrições proposto pela Teoria da Otimidade, foi possível constatar a interação entre processos de marcação e de transferência do ranqueamento de restrições do inglês para a interlíngua do aprendiz.

**Palavras-chave:** Português como segunda língua; fonologia segmental; Teoria da Otimidade.

## Introdução

É consenso na literatura em fonologia do português que a fricativa coronal em coda diante de segmento vozeado realize-se sistematicamente como vozeada, tanto em posição de interior como de fronteira de palavra, p. ex., 'de[z]de' e 'o[z] dois (ou 'de[ʒ] de' e 'o[ʒ] dois', a depender da variante regional) (BISOL, 2005; MATTOSO CÂMARA JR., 2007 [1970]). Esse processo de assimilação regressiva de vozeamento (ARV), em que o segmento vozeado espraia seu traço de sonoridade para a fricativa coronal final que se lhe antecede, consiste em um padrão fonológico do português, adquirido pela criança falante nativa nos primeiros estágios de aquisição do idioma.

No português como segunda língua (PL2) por aprendizes anglófonos, em contexto de fronteira de palavra, observa-se que o vozeamento da fricativa nem sempre ocorre, p. ex., 'o[s] dois'. Essa ausência do padrão de ARV na produção em PL2 por anglófonos será aqui tratada por meio do modelo baseado em restrições universais proposto pela Teoria da Otimidade (TO) (PRINCE; SMOLENSKY, 1993; MCCARTHY; PRINCE, 1995).

Originada como recurso formalizador da gramática das diferentes línguas, a TO tem sido também utilizada como modelo linguístico aplicado ao estudo em aquisição de segunda língua (L2).<sup>1</sup> Pesquisadores que se utilizam do modelo da TO geralmente assumem que o aprendiz utiliza-se da mesma hierarquia de restrições da L1 no início da aprendizagem da L2 (PATER, 1997; BROSELOW et al., 1998; HANCIN-BHATT; BHATT 1998; DAVIDSON et al., 2004; HANCIN-BHATT, 2008). Em inglês, p. ex., obstruintes vozeadas finais são permitidas, [gʊd] *good* ‘bom’, indicando que a restrição universal de marcação que impede obstruintes vozeadas nesse contexto está em posição baixa na hierarquia do idioma, sendo violada em favor da restrição de fidelidade que requer correspondência entre a forma subjacente, /gʊd/, e a forma de superfície, [gʊd]. Aprendizes de inglês como L2 cuja L1 não permite obstruintes vozeadas finais, o que demonstra que a restrição que milita contra obstruintes vozeadas finais ocupa posição prioritária na hierarquia, podem produzir [gɒt] para *good*, transferindo assim a hierarquia de restrições da L1 para o inglês como L2.

Objetiva-se, neste trabalho, formalizar a hierarquia de restrições inicial do PL2 por aprendizes anglófonos no que concerne ao padrão de vozeamento da fricativa final, bem como verificar, a partir de um *corpus* de fala espontânea proveniente de entrevistas com 11 aprendizes anglófonos de PL2 – seis americanos e cinco britânicos –, se o tempo de residência no Brasil correlaciona-se à realização do processo de ARV.

Este artigo segue dividido em três seções. Na seção “Teoria da Otimidade e aquisição da fonologia de L2”, os principais conceitos sobre a TO são expostos, assim como sua pertinência na área de aquisição de L2. A seção “Metodologia”, contém informações sobre os informantes e os procedimentos utilizados na coleta e análise dos dados. Na seção “Análise via Teoria da Otimidade”, é realizada a análise via TO do aspecto da fonologia do PL2 por anglófonos aqui proposto a ser estudado: a ausência de vozeamento da fricativa em posição final de palavra antecedendo segmento vozeado. Em seguida, encerra-se este trabalho com as considerações finais.

## Teoria da Otimidade e aquisição de L2

Introduzida por Prince e Smolensky (1993) e McCarthy e Prince (1993), a TO é um modelo linguístico que busca explicar o funcionamento da gramática das línguas naturais por meio de um *ranqueamento* (ou *hierarquia*) de restrições universais violáveis. Diferentes hierarquias, i.e. diferentes posicionamentos das restrições, responderiam pelas diferentes línguas existentes.

De acordo com o modelo, existem duas grandes famílias de restrições: a de restrições de fidelidade e a de restrições de marcação. Da família de restrições de fidelidade, fazem parte as restrições que requerem igualdade entre as formas de *input* e *output*, proibindo, p. ex., apagamento, epêntese ou modificação de um traço. Em contrapartida, as restrições de marcação privilegiam formas não marcadas, militando, p. ex., a favor de sílabas com *onset* e proibindo sílabas com coda, além de privilegiarem traços menos marcados em certos contextos linguísticos, como o não vozeamento de obstruintes finais. A relevância de determinada restrição na língua é dada pelo seu posicionamento na hierarquia.

---

<sup>1</sup> O termo *segunda língua* (L2) refere-se ao idioma aprendido após a aquisição da língua materna ou primeira língua (L1). Os termos *aprendizagem* e *aquisição* são aqui tratados como sinônimos.

O mecanismo pelo qual a TO seleciona o candidato a *output* que melhor obedece a princípios de boa formação linguística é constituído por 1) um conjunto de restrições de representações fonológicas, denominado CON (*constraints* ‘restrições’); 2) um criador de candidatos a *outputs*, denominado GEN (*generator* ‘gerador’); e, por fim, 3) um componente denominado EVAL (*evaluator* ‘avaliador’), que avalia os potenciais *outputs* e, considerando o ranqueamento de restrições de dada língua, seleciona o *output* ótimo para o *input* em questão.

A Figura 1 abaixo resume o funcionamento da gramática proposto pela TO.

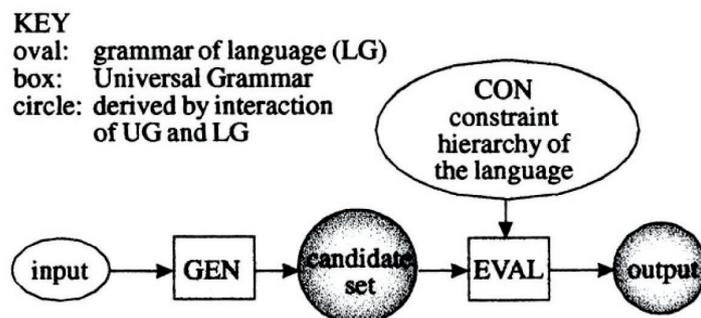


Figura 1. Esquema de funcionamento da gramática proposto pela TO (ARCHANGELI, 1999, p. 534)

As informações da gramática de uma língua particular (*input* e CON) interagem com a Gramática Universal, representada por GEN e EVAL. Para um dado *input*, GEN gera um conjunto de candidatos, os quais possuem correspondência com os elementos do *input*. EVAL então seleciona o candidato ótimo para aquele *input*, utilizando-se da hierarquia de restrições daquela língua (CON), ou seja, seleciona o candidato mais harmônico, que melhor obedece a restrições de boa formação linguística segundo a hierarquia daquela língua. O candidato escolhido por EVAL emerge então na fala, como *output* ótimo.

*Tableaux* são usados na aplicação do modelo de TO (Figura 2).

	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>
a. ☞ <i>candidate a</i>		*	*
b. <i>candidate b</i>		**!	
c. <i>candidate c</i>	*!		

Figura 2. Exemplo de *tableau* (KAGER, 1999, p. 25)

Na linha superior, as restrições (C) são dispostas de acordo com a hierarquia da língua. Os candidatos a *output* estão na primeira coluna. O asterisco (\*) marca uma violação em determinada restrição, a exclamação (!) indica que a violação é fatal, ou seja, por conta dessa violação, o candidato é eliminado, e o símbolo (☞) aponta para o *output* ótimo, o qual deve emergir na fala. A área sombreada indica as restrições que não são mais relevantes para a escolha do candidato.

Por meio da utilização de *tableaux*, é possível identificar o estágio inicial de aquisição da L2 pelo aprendiz. Segundo Hancin-Bhatt (2008), o processo de transferência da L1

para a L2, comum principalmente em níveis iniciais, torna-se explícito ao ser considerado um modelo que leva em conta hierarquias de restrições, sendo assim possível prever as formas pelas quais o aprendiz irá resolver estruturas fonológicas da L2 ainda não aprendidas. Torna-se também possível pelo modelo, de acordo com Broselow et al. (1998), explicitar a emergência de padrões não marcados, característica inerente ao sistema linguístico do aprendiz de L2 (cf. ECKMAN, 1987; MAJOR, 2001).

A TO, então, por abranger propriedades universais de marcação, pressuposto fundamental no modelo para a escolha da forma de *output*, por apresentar propriedades específicas das línguas, a partir do conceito de restrições violáveis e, finalmente, por poder se ocupar de diferentes níveis fonológicos, constitui-se como importante ferramenta para a explicação de processos linguísticos em L2.

## Metodologia

O *corpus* desta pesquisa constitui-se por trechos de fala espontânea coletados em entrevistas com 11 aprendizes anglófonos de PL2, seis americanos e cinco britânicos, todos residentes na cidade de São Paulo no momento da coleta de dados, realizada entre os meses de janeiro e fevereiro de 2011.

Serem falantes nativos de inglês que tivessem iniciado a aprendizagem de PL2 no Brasil e em idade adulta foi o critério estabelecido para a escolha dos informantes. Vale ressaltar que a ausência de ARV diante de segmento vozeado, p. ex., *loose* /lu:s/ ‘solto, folgado’ pronunciado como [lu:s] em *loose end* ‘ponta solta’ e em *loose dress* ‘vestido folgado’, é característica tanto das variedades do inglês americano, como do britânico: Smith (1997) e Myers (2010) para o *General American*; Gimson e Cruttenden (2008) para o *Received Pronunciation*; Wells (1982) para o inglês de Yorkshire.<sup>2</sup>

Foi realizada uma sessão por informante, com duração entre 15 a 20 minutos, gravadas em um aparelho da marca *Zoom H4*, a uma taxa de amostragem de 44,1 kHz. Temas como rotina, profissão, atividades de lazer e impressões sobre o Brasil e a língua portuguesa foram abordados durante as entrevistas.

Ocorrências de fricativas coronais alveolares<sup>3</sup> em coda final diante de palavra iniciada por segmento vozeado (obstruinte vozeada, consoante soante ou vogal) foram extraídas do *corpus* e transcritas foneticamente por meio de uma análise de oitiva, totalizando 411 ocorrências. Os segmentos cujo estatuto de sonoridade se deu por incerto, o que correspondeu a 7% dos casos, foram excluídos da análise.

Objetivando encontrar fatores linguísticos e extralinguísticos que pudessem se correlacionar com os resultados, fez-se uso do programa de análise estatística de dados linguísticos *Goldvarb* (SANKOFF et al., 2005), o qual selecionou apenas o *tempo de residência no Brasil* como fator relevante para a aplicação do processo de ARV (tabela 1). Os fatores linguísticos *contexto seguinte* (obstruinte vozeada, consoante soante, vogal) e *tonicidade da sílaba* (tônica, átona), assim como os fatores extralinguísticos *nacionalidade*

2 Considera-se, no entanto, haver, em inglês, um tipo de “assimilação regressiva de desvozeamento”, em que fricativas e africadas tornam-se desvozeadas diante de segmento desvozeado, p. ex., *lose* /lu:z/ torna-se [lu:s] em *lose ten-nil* ‘perder de dez a zero’ (ROCA; JOHNSON, 1999).

3 Não foram encontradas, na produção dos informantes desta pesquisa, ocorrências de fricativas coronais (pré-) palatais, [ʃ, ʒ], em coda final de palavra, apenas de fricativas alveolares, [s, z].

(americana, britânica), *conhecimento de espanhol* (sim, não) e *tempo de estudo formal de PL2* (nenhum, menos de 6 meses, 2 a 3 anos) não foram tidos como significativos para a aplicação do processo de ARV.

**Tabela 1. Pesos relativos de aplicação de ARV no PL2 por aprendizes anglófonos para o fator tempo de residência no Brasil.**

Tempo de residência no Brasil	Ocorrências de ARV em relação ao total de contextos relevantes		Peso relativo
	N	%	
3 meses	19/112	16	.22
2 a 3 anos	29/110	26	.36
9 a 11 anos	69/113	61	.72
20 a 31 anos	51/76	67	.77

Pela tabela 1, nota-se que falantes residentes no Brasil por um período superior a nove anos tenderam a produzir mais a contraparte vozeada da fricativa alveolar, [z], do que a desvozeada, [s]. Em contrapartida, os falantes residentes no país até três anos produziram majoritariamente fricativas alveolares desvozeadas nesse contexto. Nota-se ainda o valor crescente do peso relativo para a aplicação do processo: quanto maior o tempo de residência, maior a probabilidade de vozeamento.

Segue então a análise, via Teoria da Otimidade (PRINCE; SMOLENSKY, 1993; MCCARTHY; PRINCE, 1995), da ausência de ARV na fala dos aprendizes, característica, conforme consta da tabela 1, encontrada predominantemente nos estágios iniciais de aquisição (3 meses a 3 anos).

### Análise via Teoria da Otimidade

O desvozeamento de obstruintes em posição final é um aspecto não marcado característico de diversas línguas (p. ex., alemão, catalão, holandês, polonês), sendo aqui representado pela restrição (01) abaixo.

(01)  $]\_o/*\text{VOICE}$  – obstruintes vozeadas são proibidas em posição final de sílaba (MCCARTHY, 2002)

Em conflito com  $]\_o/*\text{VOICE}$ , está a restrição de fidelidade que milita a favor da manutenção no *output* do traço de vozeamento do *input*, apresentada em (02).

(02) IDENT(voice) – segmentos correspondentes no *input* e no *output* devem ter valores idênticos para o traço [sonoro] (MCCARTHY; PRINCE, 1995)

Em inglês, em posição final absoluta, a restrição de fidelidade IDENT(voice) domina a restrição de marcação  $]\_o/*\text{VOICE}$  (MCCARTHY, 2002), assim, o traço de vozeamento do *input* é preservado no *output*, p. ex., *lose* /luz/ [lu:z] ‘perder’ vs. *loose* /lu:s/ [lu:s] ‘folgado’ (tableau 1).

**Tableau 1.** Contraste de vozeamento em posição final (inglês)

<i>input</i> <i>evadão</i> <i>valovalor</i> <i>altomento</i> . <i>desvozeado</i> , <i>a</i> <i>tende pausa</i> <i>ou de segmento</i> <i>vozeado</i> /lu:z/	IDENT(voice)	] <sub>σ</sub> /*VOICE
☞ a. lu:z		*
b. lu:s	*!	
<i>input</i> /lu:s/		
☞ c. lu:s		
d. lu:z	*!	*

A violação de IDENT(voice) pelos candidatos (1b) e (1d) é fatal, resultando na eliminação desses candidatos. Os candidatos (1a) e (1c), por satisfazerem a restrição prioritária na hierarquia, IDENT(voice), são selecionados como *outputs* ótimos. Nota-se que os efeitos de ]<sub>σ</sub>/\*VOICE são invisíveis ou irrelevantes, sua violação pelo candidato (1a), por exemplo, não impede que esse candidato seja selecionado como *output* ótimo.

Em PB, presume-se que ocorra o ranqueamento oposto ao do inglês, ]<sub>σ</sub>/\*VOICE >> IDENT(voice). O fato de não haver obstruintes vozeadas em posição final absoluta pode indicar a dominância da restrição ]<sub>σ</sub>/\*VOICE na hierarquia do PB, conforme expresso no *tableau 2* para a palavra ‘mas’.

**Tableau 2.** Neutralização de vozeamento em posição final (PB)

<i>input</i> /mas/	] <sub>σ</sub> /*VOICE	IDENT(voice)
a. maz	*!	*
☞ b. mas		
<i>input</i> /maz/		
c. maz	*!	*
☞ d. mas		*

No *tableau 2*, o ranqueamento ]<sub>σ</sub>/\*VOICE >> IDENT(voice) leva à escolha de (2b) e (2d) como *outputs* ótimos, independentemente da forma que lhes serviram como *input*, /mas/ ou /maz/. Vale ressaltar que, considerando que o *input* tenha sido /maz/, o *output* [mas] justificaria o ranqueamento ]<sub>σ</sub>/\*VOICE >> IDENT(voice). Caso o *input* tenha sido /mas/, tem-se como argumento para a hierarquia o fato de que, na aquisição de L1, restrições de marcação dominam restrições de fidelidade (DEMUTH, 1995; SMOLENSKY, 1996; GNANADESIKAN, 2004), não havendo, em PB, *inputs* que comprovem a modificação desse ranqueamento no decorrer do processo de aquisição.

Pelo pressuposto da Riqueza de Base, não há restrições para formas de *input*, as diferentes hierarquias é que são responsáveis pelos diferentes inventários gramaticais das línguas (SMOLENSKY, 1996).<sup>4</sup> No entanto, como este trabalho trata de dados de aquisição, deve-se considerar um *input* que sirva como ponto de partida para as análises. Seguindo Mattoso Câmara Jr. (2007 [1970]), adota-se aqui a contraparte desvozeada da fricativa coronal, /s/, como forma de *input* para o PB.

4 No entanto, as formas de *input* devem obedecer a propriedades básicas e universalmente respeitadas – não se assumiria, em português, p. ex., /batata/ como *input* para a palavra *pera*, ou /txrt/ para *flor*.

Para que o ranqueamento proposto dê conta do fenômeno de ARV, deve-se considerar ainda a pertinência de outras restrições, acrescentando-as às hierarquias já estabelecidas.

A tendência não marcada das línguas para que duas oclusivas adjacentes compartilhem o mesmo traço de vozeamento pode ser representada pela restrição (03).

(03) AGREE – *clusters* de oclusivas devem concordar em vozeamento (MCCARTHY; PRINCE, 1995)

AGREE em posição elevada na hierarquia, acima de restrições de fidelidade, faz com que os elementos de um *cluster* compartilhem o mesmo valor para o traço de vozeamento, p. ex., de/sd/e > de[z d]e, o/s d/ois > o[z d]ois, em que o a fricativa passa a compartilhar do traço [+sonoro] da oclusiva seguinte.

Como a concordância de vozeamento pode também ser estendida, em algumas línguas, para o contexto que envolve fricativa diante de soante heterossilábica, p. ex., o[z m]eses, o[z o]utros, adota-se aqui a restrição em (04).

(04) AGREE FS – a sequência heterossilábica oclusiva contínua + soante deve concordar em vozeamento

A restrição AGREE FS leva em conta a aproximação que existe entre fricativas e soantes em relação a oclusivas e soantes. Em catalão, por exemplo, de acordo com Cebrian (1999), as fricativas finais podem se tornar vozeadas tanto pela influência da oclusiva vozeada seguinte como da vogal heterossilábica seguinte, diferente das oclusivas finais, que apenas são vozeadas regressivamente por outra oclusiva. Essa aproximação das fricativas com as vogais, segundo Cebrian, ocorre devido ao caráter de continuidade desses segmentos, os quais mereceriam uma restrição de concordância à parte: AGREE FV – a sequência heterossilábica oclusiva contínua + vogal deve concordar em vozeamento. Essa restrição proposta por Cebrian dá base, neste trabalho, ao estabelecimento de AGREE FS, a qual, ao abarcar as demais soantes, atende às especificidades do PB.

A restrição AGREE FS pode ser dominada por AGREE em línguas que não possuem ARV da fricativa diante de soantes heterossilábicas, ocorrendo compartilhamento de traço apenas entre *clusters* de oclusivas. Em holandês, por exemplo, não há ARV desencadeada por soante de início de palavra prosódica, o segmento é desvozeado na forma de *output* e assim permanece diante de soantes, p. ex., *die/v/+achtig* > *die[f].achtig* ‘gatuno’, podendo ser vozeado apenas diante de outra oclusiva, p. ex., *za/k/+d/oek* > *za[g.d]oek* ‘lenço’ (GRIJZENHOUT, 2000, p. 5).

Propõe-se aqui que as restrições AGREE e AGREE FS estejam em posição elevada na hierarquia do PB, respondendo pelo vozeamento da fricativa tanto diante de consoantes vozeadas como diante de soantes heterossilábicas (*tableau* 3).

**Tableau 3.** Assimilação regressiva de vozeamento (português)

/desde/	AGREE	AGREE FS	]σ/*VOICE	IDENT (voice)
a. de[sd]e	*!			
F b. de[zd]e			*	*
/os#dois/				
c. o[s d]ois	*!			
F d. o[z d]ois			*	*
/os#outros/				
e. o[s o]utros		*!		
F f. o[z o]utros			*	*

Para que ocorra o vozeamento da fricativa em coda diante de segmento vozeado, AGREE e AGREE FS devem dominar ]σ/\*VOICE, caso contrário, a fricativa permaneceria desvozeada. Os candidatos (3a), (3c) e (3e), por violarem as restrições prioritárias, AGREE e AGREE FS, são eliminados, restando os candidatos ótimos (3b), (3d) e (3f), que violam restrições inferiores na hierarquia, ]σ/\*VOICE e IDENT (voice).

Para o inglês, segue-se aqui a hierarquia de restrições baseada em Grijzenhout (2000), a qual faz uso de uma restrição de fidelidade limitada ao domínio do radical, IDENTSTEM (voice), expressa em (06).

- (06) IDENTSTEM (voice) – no que concerne ao domínio do radical, um segmento que é [α voice] no *input* é [α voice] no *output* (GRIJZENHOUT, 2000)

O *output* ‘ca[dz]’ para *cats* ca/t+z/ ‘gatos’, por exemplo, violaria a restrição IDENTSTEM (voice), pela mudança do traço de vozeamento do segmento pertencente ao radical: de /t/, no *input*, para [d], no *output*. Já o *output* ‘ca[ts]’ não violaria essa restrição, uma vez que mantém o traço de sonoridade pertencente ao domínio morfológico do radical.

Grijzenhout (2000) propõe então um ranqueamento que serve tanto para a assimilação progressiva de vozeamento que ocorre em inglês – nos casos de plural dos nomes, verbos na 3ª pessoa do singular e formas de genitivo –, como para a assimilação regressiva de desvozeamento que ocorre em palavras como *fifth* ‘quinto’, fi/v/<sub>root</sub>+/θ/ > fi[fθ], e *width* ‘largura’, wide + th > [witθ], em que AGREE é relevante. Além disso, sua hierarquia responde também pelo padrão de vozeamento de *clusters* em posição de fronteira de palavra.

Segue abaixo o ranqueamento proposto por Grijzenhout (2000) para o inglês, com o acréscimo, aqui estabelecido, da restrição AGREE FS.

**Tableau 4.** Padrão de vozeamento baseado em Grijzenhout (2000) (inglês)

fi/v/ <sub>root</sub> + /θ/	IDENTSTEM (voice)	AGREE	IDENT (voice)	]σ/*VOICE	AGREE FS
a. fi[vθ]		*!			
b. fi[vð]			*	*!	
☞ c. fi[fθ]			*		
ca/t/ <sub>stem</sub> + /z/					
d. ca[tz]		*!		*	
e. ca[dz]	*!		*	*	
☞ f. ca[ts]			*		
do/g/ <sub>stem</sub> + /f/ight <sub>stem</sub>					
☞ g. do[g.f]ight		*		*	
h. do[k.f]ight	*!		*		
thi/s d/og					
☞ i. thi[s.d]og		*			
j. thi[z.d]og	*!		*	*	
thi/s a/nimal					
☞ k. thi[s.æ]nimal					*
l. thi[z.æ]nimal	*!		*	*	

No *tableau* acima são prontamente eliminados os candidatos a *output* ótimo (4e), (4h), (4j) e (4l), por não serem fiéis ao traço de vozeamento do *input* no que concerne ao domínio do radical. São também eliminados os candidatos (4a) e (4d), uma vez que violam AGREE, restrição importante no domínio prosódico em questão. Candidatos que violam a restrição IDENT(voice), em favor da concordância entre *clusters* de obstruintes tautosilábicas, são permitidos, caso não violem IDENTSTEM(voice), como (4c) e (4f).

Na produção dos informantes, a qualidade do segmento seguinte (vogal, consoante soante ou obstruinte vozeada), assim como a tonicidade da sílaba em que se encontra o segmento (átona, tônica), não foram características selecionadas como relevantes pelo programa *Goldvarb* para a produção de ARV. Apenas *tempo de residência no Brasil* foi selecionado como fator significativo.

O grupo de falantes residentes no país por três meses frequentemente produziu, diante de segmento vozeado, a contraparte desvozeada da fricativa alveolar. Foram 75% ocorrências de [s] em contextos relevantes para a aplicação do vozeamento, p. ex., ‘amigo[s] no Brasil’, ‘dua[s] horas’, ‘todo[s] o[s] dias’, ‘ma[s] depoi[s] um pouco’, ‘ma[s] aqui,’ ‘meu[s] amigo[s] brasileiros’.

Assumindo-se que os aprendizes iniciantes estejam se utilizando da mesma hierarquia do inglês em sua produção em PL2, foi, primeiramente, estabelecida a seguinte hierarquia para esses falantes: IDENTSTEM (voice) >> AGREE >> IDENT (voice) >> ]σ/\*VOICE >> AGREE FS, conforme pode-se verificar no *tableau 5*.

**Tableau 5.** Ranqueamento ilustrativo da transferência da L1 (PL2)

ma/s/ <sub>stem</sub> + /d/epois	IDENTSTEM (voice)	AGREE	IDENT (voice)	]σ/*VOICE	AGREE FS
☞ a. ma[s d]epois		*			
b. ma[z d]epois	*!		*	*	
meu/s/ + /a/migos					
☞ c. meu[s a]migos					*
d. meu[z a]migos			*!	*	
ma/z/ <sub>stem</sub> + /d/epois					
e. ma[s d]epois	*!	*	*		
⊗ f. ma[z d]epois				*	
meu/z/ + /a/migos					
g. meu[s a]migos			*!		*
⊗ h. meu[z a]migos				*	

Pelo ranqueamento acima, deve haver, prioritariamente, correspondência entre o traço de vozeamento do *input* e do *output* em relação à obstruinte que ocupa o radical. Pode não haver concordância de vozeamento entre obstruintes de um *cluster*, i.e. AGREE pode ser violada, desde que a restrição IDENTSTEM (voice) seja satisfeita, tal como ocorre com (5a).

Caso /s/ tenha sido a forma subjacente adotada pelos informantes, pode-se assumir que eles, ao produzir [s] como *output*, estejam se utilizando da hierarquia do inglês no início de sua aprendizagem do PL2, ao satisfazerem, de modo prioritário, IDENTSTEM(voice), ou ainda que estejam rranqueando ]σ/\*VOICE para uma posição superior na hierarquia, como se verá adiante. Alternativamente, pode se considerar que os informantes tenham a fricativa alveolar vozeada, /z/, como forma subjacente. Apesar de ouvirem [s] em posição final de palavra em PB, podem ter notado que, diante de segmento vozeado, a fricativa se realiza como [z], assim como também é realizada como [z] no plural de palavras com término em [s], p. ex., ‘me[s]’, ‘me[z]es’. Desse modo, não se deve descartar a possibilidade de que /z/ esteja como *input* na produção dos informantes.

Pressupondo-se que /z/ seja então a forma subjacente, o *output* ótimo realizado com [s] revelaria que as restrições de fidelidade IDENTSTEM(voice) e IDENT(voice) não têm prioridade na hierarquia do aprendiz: ‘ma[s d]epois’ viola IDENTSTEM(voice), pela mudança no traço de vozeamento da obstruinte final do radical, portanto, violando também IDENT(voice), e ‘meu[s a]migos’ viola IDENT(voice), como se pode constatar no *tableau 5*, em que “L” indica que não foi escolhido o *output* ótimo correto pela hierarquia estabelecida.

Uma das propostas que se faz aqui é a de que a restrição ]σ/\*VOICE tenha sido rranqueada para a posição dominante nesse primeiro estágio de aquisição do PL2, respondendo pelo não vozeamento da fricativa final. Assim, os *outputs* ótimos ‘ma[s d]epois’ e ‘meu[s a]migos’ seriam selecionados, não havendo erros no ranqueamento (*tableau 6*).

**Tableau 6.** Ranqueamento com a restrição ]σ/\*VOICE em dominância (PL2)

ma/s/ <sub>stem</sub> + /d/epois	]σ/*VOICE	IDENTSTEM (voice)	AGREE	IDENT (voice)	AGREE FS
F a. ma[s d]epois			*		
b. ma[z d]epois	*!	*		*	
meu/s/ + /a/migos					
F a. meu[s a]migos					*
b. meu[z a]migos	*!			*	
ma/z/ <sub>stem</sub> + /d/epois					
F a. ma[s d]epois		*	*	*	
b. ma[z d]epois	*!				
meu/z/ + /a/migos					
F a. meu[s a]migos				*	*
b. meu[z a]migos	*!				

A produção do *output* em [s], tendo /z/ por *input*, justificaria a posição dominante de ]σ/\*VOICE na hierarquia. Por outro lado, considerando /s/ como *input*, um argumento que justificaria [s] como resultante da dominância de ]σ/\*VOICE, e não de IDENTSTEM(voice), teria por base o fato de que, em dados de aquisição de L1, restrições de marcação dominam restrições de fidelidade, conforme visto, o que poderia estar também definindo a hierarquia de aprendizes anglófonos de PL2 em início de aprendizagem.

Apesar de não ser possível determinar o valor para o traço de vozeamento do *input* pertencente à produção dos informantes, conseqüentemente não sendo possível estabelecer se, caso o *input* seja com [s], IDENTSTEM(voice) ou ]σ/\*VOICE dominam, pode-se afirmar que as restrições que militam a favor da concordância de vozeamento entre segmentos adjacentes, AGREE e AGREE FS, são dominadas em sua hierarquia, independentemente do *input*. Isso permite, tal como em inglês como L1, a formação de *cluster* de obstruinte final desvozeada e consoante vozeada, p. ex., ‘ma[s d]epois’ (PL2) e *this dog* ‘esse cão’ > /ðis/ thi[s d]og (inglês), e a produção da sequência fricativa final desvozeada + soante, p. ex., ‘meu[s a]migos’ (PL2) e *less is more* ‘menos é mais’ > /les/ le[s i]s more (inglês).

Ao se considerar AGREE e AGREE FS em posição baixa na hierarquia do PL2 por anglófonos, sendo dominadas por ]σ/\*VOICE, ou mesmo por restrições de fidelidade ao vozeamento subjacente – uma das possibilidades quando se considera /s/ como *input* –, o ranqueamento desses aprendizes iniciantes de PL2 difere substancialmente do ranqueamento do PB, em que as restrições AGREE e AGREE FS vêm em posição dominante, devendo ser obrigatoriamente satisfeitas.

Em (07), segue o ranqueamento proposto para o PL2 por anglófonos, em que as restrições AGREE e AGREE FS estão em posição inferior na hierarquia, podendo ser dominadas ou pela restrição de marcação ]σ/\*VOICE ou pelas restrições de fidelidade IDENTSTEM(voice) e IDENT(voice).

- (07) Ranqueamento em que AGREE e AGREE FS são dominadas (PL2)  
 {]σ/\*VOICE, IDENTSTEM(voice), IDENT(voice)} >> AGREE, AGREE FS

Assim, a hierarquia dos informantes difere-se da hierarquia do PB, assemelhando-se então à do inglês, no que concerne à posição dominada de AGREE e AGREE FS. Como não

foi determinado se o *input* dos informantes é com /z/ ou com /s/, não é possível precisar se eles, ao produzirem [s] no *output*, estariam se utilizando de uma hierarquia que prioriza restrições de fidelidade ao vozeamento do *input*, o que iria se caracterizar como transferência da L1, ou que prioriza a restrição de marcação que milita a favor do não vozeamento de obstruintes finais, ]σ/\*VOICE, característica que se constituiria como um exemplo da *emergência do não marcado*. No que concerne a padrões de vozeamento de *clusters* e de obstruintes diante de soante heterossilábica, ]σ/\*VOICE não é uma restrição prioritária nem em inglês, nem em português. No primeiro, a restrição de fidelidade ao vozeamento do *input*, IDENTSTEM(voice), vem em posição prioritária; no segundo, AGREE e AGREE FS são as restrições dominantes na hierarquia. A posição dominante de ]σ/\*VOICE no ranqueamento dos informantes refletiria um dado da emergência do não marcado, em que, conforme Broselow et al. (1998), uma restrição obscurecida na L1 do falante passaria a atuar, pelo rerranqueamento, em sua produção em L2.

Apesar de não ser comprovada a emergência do não marcado na produção dos informantes, uma vez que não se pode determinar o valor para o traço de vozeamento do *input* – somente o *input* em /z/ poderia assegurar a emergência do não marcado –, o modelo de TO é eficiente ao explicitar a transferência da L1 no que diz respeito à posição dominada de AGREE e AGREE FS no ranqueamento dos informantes.

### Considerações finais

Este estudo tratou da descrição e análise, via Teoria da Otimidade (PRINCE; SMOLENSKY, 1993; MCCARTHY; PRINCE, 1995), de um aspecto saliente da fonologia do PL2 por aprendizes anglófonos: a ausência de vozeamento da fricativa alveolar final diante de segmento vozeado.

Os dados descritivos revelaram que o processo de ARV correlacionou-se com o tempo de residência do aprendiz no Brasil: quanto maior o tempo de residência, maior a frequência de vozeamento. No entanto, o fenômeno não se mostrou plenamente adquirido, uma vez que a variação persistiu mesmo na produção de aprendizes mais experientes, com tempo de residência no país superior a nove anos.

A partir da análise via Teoria da Otimidade de dados de fala de aprendizes iniciantes, com até três meses de residência no país, demonstrou-se que a ausência de ARV pode resultar tanto da posição dominante de uma restrição de fidelidade ao traço de vozeamento do *input*, caso a forma subjacente seja /s/, como da posição dominante de uma restrição de marcação que proíbe o vozeamento de obstruintes finais. O primeiro caso indicaria transferência da hierarquia da L1, já que a restrição de fidelidade que exige correspondência entre o traço de vozeamento do *input* e *output*, IDENTSTEM(voice), dominante na hierarquia do inglês, passaria a ser dominante também na hierarquia do PL2. No segundo caso, os dados de fala dos aprendizes exemplificariam a emergência do não marcado, devido ao rerranqueamento de ]σ/\*VOICE para a posição dominante. Em inglês, a atuação de ]σ/\*VOICE seria obscurecida por restrições de fidelidade mais altas, e, em português, seria dominada pelas restrições de concordância de vozeamento AGREE e AGREE FS.

Ainda que não se tenha constatado, na hierarquia inicial do aprendiz, a dominância da restrição IDENTSTEM(voice) ou de ]σ/\*VOICE, provou-se que, assim como ocorre em inglês, AGREE e AGREE FS são restrições dominadas. Portanto, há transferência do inglês

na hierarquia inicial do aprendiz, validando a tese de que, em início de aprendizagem, segue-se a hierarquia da L1, ainda que parcialmente (HANCIN-BHATT, 2008).

A discussão, neste trabalho, do envolvimento entre transferência de L1 e processos de marcação e a formalização do ranqueamento inicial que governa o padrão de vozeamento na produção do aprendiz anglófono de PL2 podem servir de encaminhamento para pesquisas futuras, em que, a partir de um *corpus* mais robusto de análise, leve-se em conta a variação e o rerranqueamento de restrições no decorrer do processo de aprendizagem – tarefa que pode ser executada por meio de algoritmos de aprendizagem (BOERSMA, 1997; TESAR; SMOLENSKY, 2000; BOERSMA; HAYES, 2001).

## REFERÊNCIAS

ARCHANGELI, D. Introducing Optimality Theory. *Annual Review of Anthropology*, v. 28, p. 531-552, 1999. Disponível em: <<http://www.jstor.org/stable/223405>>. Acesso em: 16 abr. 2011.

BISOL, L. (Org.). *Introdução aos estudos de fonologia do português brasileiro*. 4. ed. Porto Alegre: EDIPUCRS, 2005.

BOERSMA, P. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences*, v. 21, p. 43–58. University of Amsterdam, 1997.

BOERSMA, P; HAYES, B. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, v. 32, n. 1, p. 45-86, 2001.

BROSELOW, E.; CHEN, S.; WANG, C. The Emergence of the Unmarked in Second Language Phonology. *Studies in Second Language Acquisition*, v. 20, p. 261-280, 1998.

CEBRIAN, J. Catalan obstruents and opacity: A case for representational ambisyllabicity. In: SMALLWOOD, C.; BÉJAR, S. (Ed.) *Toronto Working Papers in Linguistics*, v. 17. Toronto: University of Toronto, 1999. p. 58-80.

DAVIDSON, L.; JUSCZYK, P.; SMOLENSKY, P. The initial and final states: theoretical implications and experimental explorations of Richness of the Base. In: KAGER, R.; PATER, J.; ZONNEVELD, W. (Org.) *Constraints in Phonological Acquisition*. Cambridge University Press, 2004. p. 321-368.

DEMUTH, K. Markedness and the development of prosodic structure. *NELS*, v. 25, p. 13-25, 1995.

ECKMAN, F. Markedness and the Contrastive Analysis Hypothesis. In: IOUP, G; WEINBERGER, S. H. (Org.) *Interlanguage Phonology: The acquisition of a second language sound system*. Cambridge: Newbury House Publishers, 1987. p. 55-69.

GIMSON, A. C.; CRUTTENDEN, A. *Gimson's Pronunciation of English*. 7th edition. London: Hodder Education, 2008.

GNANADESIKAN, A. Markedness and faithfulness constraints in child phonology. In: KAGER, R.; PATER, J.; ZONNEVELD, W. (Org.) *Constraints in Phonology Acquisition*. Cambridge University Press, 2004. p. 109-157.

GRIJZENHOUT. Voicing and devoicing in English, German, and Dutch: evidence for domain-specific identity constraints. Two papers on constraint domains. *SFB 282 working paper no. 116*. Düsseldorf: HHU, 2000.

HANCIN-BHATT, B. Second Language Phonology in Optimality Theory. In: HANSEN EDWARDS, J. G.; ZAMPINI, M. L. (Ed.). *Phonology and Second Language Acquisition*. Amsterdam, Philadelphia: Benjamins, 2008. p. 117-148.

HANCIN-BHATT, B.; BHATT, R. Optimal L2 syllables – interactions of transfer and developmental effects. *Studies in Second Language Acquisition*, v. 19, p. 331-378, 1998.

KAGER, R. *Optimality Theory*. Cambridge: Cambridge University Press, 1999.

MAJOR, R. *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates, 2001.

MATTOSO CÂMARA JR., J. *Estrutura da língua portuguesa*. 40. ed. Petrópolis: Vozes, 2007 [1970].

MCCARTHY, J. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press, 2002.

MCCARTHY, J.; PRINCE, A. Prosodic Morphology I: Constraint Interaction and Satisfaction. *Report no. RuCCS-TR-3*. New Brunswick, NJ: Rutgers University Center for Cognitive Science, 1993.

\_\_\_\_\_. A. Faithfulness and Reduplicative Identity. In: BECKMAN; WALSH DICKEY; URBANCZYK (Orgs.) *Papers in Optimality Theory*. Amherst, MA: Graduate Linguistic Student Association, 1995. p. 249-384.

MYERS, S. Regressive voicing assimilation: Production and perception studies. *Journal of the International Phonetic Association*, v. 40, n. 2. The University of Texas at Austin, p. 163-179, 2010.

PATER, J. Minimal violation and phonological development. *Language Acquisition*, v. 6, p. 201-253, 1997.

PRINCE, A.; SMOLENSKY, P. *Optimality Theory: constraint interaction in generative grammar*. Baltimore: The Johns Hopkins University, 1993.

ROCA, I.; JOHNSON, W. *A Course in Phonology*. Oxford e Malden, MA: Blackwell Publishers, 1999.

SANKOFF, G.; TAGLIAMONTE, S.; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Toronto, CA: Department of Linguistics, University of Toronto, 2005. Disponível em <<http://individual.utoronto.ca/tagliamonte/goldvarb.htm>>. Acesso em: 10 out. 2011.

SMOLENSKY, P. The initial state and 'richness of the base' in Optimality Theory. *Technical Report JHU-CogSci-96-4*. Baltimore: Cognitive Science Department, Johns Hopkins University, MD, 1996.

SMITH, C. The devoicing of /z/ in American English: effects of local and prosodic context. *Journal of Phonetics*, v. 25, p. 471-500, 1997.

TESAR; SMOLENSKY. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press, 2000.

WELLS, C. *Accents of English 2: The British Isles*. Cambridge: Cambridge University Press, 1982.

# O papel da modalização no jogo discursivo: um estudo da construção *pode ser (...), mas ...*

(The role of modalization on discourse game:  
an analysis of construction *pode ser (...), mas ...*)

Cibele Naidhig de Souza Carrascossi<sup>1</sup>

<sup>1</sup>Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Araraquara

cibelecarrascossi@hotmail.com, cibelecarrascossi@yahoo.com.br

**Abstract:** The goal of this paper is to analyze occurrences of expression *pode ser* (as a predicate of an embedded proposition or as an independent structure), followed by an objection introduced by *mas* (*pode ser (...), mas*) in Portuguese. Based on the functionalist theory, we study the modal value and the communicative and argumentative strategy which are expressed in these contexts. The analysis reveals the modality as an essential category for the management of different points of view in discourse which also involves the defense and protection of face.

**Keywords:** functionalism; modal verbs; *pode ser*; adversative relation

**Resumo:** O objetivo deste trabalho é analisar, na língua portuguesa, ocorrências da expressão *pode ser* (como predicado encaixador de proposição ou como construção independente), seguida de objeção introduzida por *mas* (*pode ser (...), mas*). Com base na teoria funcionalista, estudam-se o valor modal e as estratégias comunicativas e argumentativas que se manifestam nesses contextos. O exame revela a modalidade como uma categoria fundamental para o trabalho de defesa e proteção de faces, para o gerenciamento de diferentes perspectivas no discurso.

**Palavras-chave:** funcionalismo; verbos modais; *pode ser*; relação adversativa.

## Introdução<sup>1</sup>

Embora muito já se tenha escrito sobre modalização linguística, pouco esclarecido está, ainda, o papel da categoria como orientadora da interação comunicativa, função que se salienta ao exame de determinadas construções modalizadoras, como *pode ser*.

Neste trabalho investigam-se ocorrências de *pode ser* em dois contextos sintáticos distintos: como predicado encaixador de proposição, exemplificado por (1); como construção independente, exemplificado por (2).

- (1) Acomodou o corpanzil como pôde no espaço que nos cabia, olhou para mim de um jeito conhecido e disse:  
— Quer dizer que estão comentando o fato de você ter deixado o pássaro voar?  
Não havia maldade, ironia ou censura na frase. Pelo que eu conhecia de Aurélio, a observação significava: “Cuidado, amigo, podem te ferrar”.

---

<sup>1</sup> As considerações apresentadas neste artigo são parte revista da tese de doutorado desenvolvida sob orientação da Profa. Dra. Maria Helena de Moura Neves, com bolsa do CNPq (processo n. 141343/2008-7), na Unesp, Faculdade de Ciências e Letras, Araraquara (CARRASCOSSI, 2011).

- O pássaro que deixei voar, Aurélio, não consegue ultrapassar nem o muro do quintal. **Pode ser que** seja menos inofensivo do que imaginei, **mas** não acredito que seja um assassino. Trata-se de um tímido. Falta a ele o desassombro do matador. (OSC)<sup>2</sup>
- (2) — Deixa de besteira, Luís. Ela deu o colar a tia Matilde, em Poços.
- **Pode ser... Mas** abre logo essa mão e passa minha parte. (COT)

No primeiro caso, *pode ser* constitui um predicado encaixador de proposição, introduzindo um conteúdo oracional, analisado como sujeito oracional. Em ocorrências como (2), por outro lado, *pode ser* é uma fórmula avulsa.

Os contextos em que a expressão *pode ser* é seguida de uma objeção introduzida por *mas*, como (1) e (2), são muito recorrentes e constituem fórmula extremamente produtiva para o gerenciamento de diferentes perspectivas no discurso, conforme revelarão as análises. Usos discursivos semelhantes ao de *pode ser*, nesses enunciados, são encontrados também em outras línguas, tais como o inglês (COATES, 1983 apud SWEETSER, 1990; SWEETSER, 1990; NUYTS, 2001), o holandês e o alemão (NUYTS, 2001). Essas considerações despertaram a atenção por uma investigação mais cuidada desses dados, na língua portuguesa.

Interessa, então, a este trabalho estudar a construção *pode ser* (predicado encaixador de proposição ou construção independente.), *mas...* (*pode ser(...)*, *mas...*) analisando-se o valor modal e as estratégias comunicativas, argumentativas que aí se manifestam. Será demonstrado que a investigação dessa construção revela o papel da modalidade como organizadora e reguladora da interação comunicativa, função que é negligenciada pelas conceituações mais tradicionais.

Como suporte teórico-metodológico, adota-se a teoria funcionalista da linguagem, mais especificamente nos termos de Dik (1997a, 1997b), em que se privilegia, acima de tudo, a funcionalidade das estruturas linguísticas em usos efetivos. O estudo da modalização, que se realiza nas situações de interação entre os usuários da língua, requer uma base teórica que considere a língua em uso, que inclua na análise, além da estrutura linguística, a situação comunicativa, o propósito do evento da fala, seus participantes e seu conteúdo discursivo.

Cabe esclarecer, ainda, que o *corpus* da pesquisa é constituído por textos escritos e falados. Analisaram-se textos do banco de dados do Laboratório de Lexicografia da UNESP de Araraquara, que possui mais de 200 milhões de ocorrências em textos escritos diversificados pelos tipos dramático, oratório, de propaganda, técnico, romanesco e jornalístico. Selecionaram-se, para exame, os textos do período compreendido entre 1950 e 2000. Os textos falados investigados são provenientes do Projeto NURC (Norma Urbana Culta). Examinaram-se 60 inquéritos distribuídos igualmente pelas cinco capitais (Porto Alegre, Rio de Janeiro, São Paulo, Recife e Salvador) e pelos três tipos (D2 – diálogo entre dois informantes; DID – diálogo entre informante e documentador; EF – elocução formal) abarcados pelo projeto.

Além da introdução e das considerações finais, o artigo está dividido em três seções. Primeiramente, discute-se a questão da modalidade e assenta-se a proposta de estudo da

<sup>2</sup> As letras entre parênteses são uma abreviatura do nome da obra, tal como foi dada no banco de dados do Laboratório de Lexicografia da Unesp, FCL, Araraquara, de onde essas ocorrências foram extraídas. Adiante será esclarecida a constituição do corpus da pesquisa.

modalidade a partir de sua multifuncionalidade nos enunciados. Em seguida, analisa-se a fórmula *pode ser(...), mas...* como peça do jogo discursivo, como estratégia de gerenciamento de pontos de vista divergentes, e busca-se, ainda, investigar o valor adversativo nessas construções.

## A modalidade

As modalidades constituem categoria linguística de grande complexidade e, por isso, difícil de ser apreendida, conforme atesta o grande volume de material que já se produziu sobre o tema, comparado à abundância de dissensos e de questões polêmicas.

Não cabe nos limites deste artigo explorar a instigante complexidade das modalidades, o que implicaria problematizar diversas questões que ao analista se impõem.<sup>3</sup> Para este artigo, importa demonstrar, apenas, que as noções mais convencionais não dão conta da funcionalidade das modalidades no gerenciamento da interação verbal.

As primeiras reflexões sobre modalidade têm suas origens nos estudos aristotélicos, em que se propõem as noções modais de possibilidade e de necessidade, e a ideia de que esses conceitos são interdefiníveis a partir da negação. No entanto, parece ponto pacífico, nos estudos linguísticos, que a análise lógica não se mantém, dado que nela a modalidade fica definida em termos de relações de verdade, independentemente do enunciador. Apesar disso, como observa Neves (2006), os estudos linguísticos são muito influenciados pela visão lógica e, então, os conceitos de possibilidade e de necessidade dificilmente não estão presentes na definição de modalidade.

Outra forma de delimitar a modalidade relaciona-se com a noção de *factualidade* (LYONS, 1977; NARROG, 2005) também referida pelos termos *real/irreal* (PALMER, 1986) ou *validade* (KIEFER, 1987). De modo bastante geral, a modalidade seria marcada pela não factualidade, ou seja, pela indeterminação em relação ao estatuto factual, o que significa que um enunciado modalizado não é nem positivamente, nem negativamente factual. Porém, como diz Palmer (1986, p.18), seria um erro limitar o estudo da modalidade à não factualidade, porque as declarações factuais, como as sentenças que marcam opinião ou julgamento do enunciador (não factuais), são subjetivas, representam pontos de vista do falante.

A concepção de modalidade baseada na noção de subjetividade é, em geral, bastante aceita, e encontra-se diferentemente expressa nas definições de teóricos e estudiosos:

[...] um meio usado pelo falante para expressar sua opinião ou atitude acerca de uma proposição que a sentença expressa ou a situação que a proposição descreve (LYONS, 1977, p. 436);<sup>4</sup>

[...] o modo pelo qual o significado de uma frase é qualificado de forma a refletir o julgamento do falante sobre a probabilidade de ser verdadeira a proposição por ele expressa (QUIRK et al., 1985, p. 219);

[...] a relação que se estabelece entre o sujeito da enunciação e seu enunciado (MAINGUENEAU, 1990, p. 180);

3 Para uma reflexão mais profunda, sugere-se a leitura do capítulo 2 de Carrascossi (2011), em que se problematiza a categoria modalidade, discutindo-se, principalmente, sua conceituação e sua classificação.

4 As citações em língua estrangeira foram traduzidas para o português, tradução sobre a qual a autora deste artigo responsabiliza-se.

[...] a expressão da subjetividade de um enunciador que assume com maior ou menor força o que enuncia, ora comprometendo-se, ora afastando-se, seguindo normas determinadas pela comunidade em que se insere. (CORACINI, 1991, p. 113)

Embora esse tipo de conceituação seja o mais comumente encontrado nos estudos de modalidade, ele não prevê a atuação da categoria em determinadas situações de interlocução. A modalidade marca a participação do enunciador na comunicação, que é realizada sob certas pressões interacionais e com o intuito de alcançar determinados objetivos comunicativos ou determinados efeitos discursivos. É preciso, então, atentar-se para o recurso pragmático que a modalidade representa no discurso. Especialmente na conversação, a modalidade pode assumir a função de reguladora e organizadora da interação verbal, e esse aspecto é negligenciado pelos trabalhos que a investigam com base em definições como as apresentadas.

Mais especificamente, o entendimento da modalidade como opinião ou atitude do falante em relação ao conteúdo proposicional por ele expresso tem alcance restrito, afinal, a modalidade:

- a) não é só atitude/opinião do falante em relação ao seu enunciado, pode ser atitude/opinião em relação ao enunciado do interlocutor também;
- b) não incide necessariamente sobre um conteúdo proposicional;
- c) pode estar fortemente ligada ao domínio discursivo, em determinados enunciados, e, por isso, nesses casos, sua principal função não é expressar um posicionamento de quem fala, e sim gerenciar a interação verbal.

Tais pontos serão explorados na seção seguinte, com a análise dos dados. Por ora, cabe ressaltar que se entende, neste estudo, que um tratamento adequado da modalidade deve considerar seu papel no processamento textual, na interação, ou “tensão”, comunicativa que se estabelece entre falante e ouvinte. Algumas reflexões direcionadas para a discussão da modalidade como orientadora da interação são encontradas em Hoffnagel (1997) e em Miranda (2000).

Analisando a modalidade epistêmica, Hoffnagel (1997) chama a atenção para o caráter intersubjetivo da modalização indicando que “além de explicar as atitudes do falante com respeito às proposições propostas, os modalizadores epistêmicos são empregados, também, para comunicar atitudes do falante, em relação a seus interlocutores” (HOFFNAGEL, 1997, p. 150).

Miranda (2000), em uma perspectiva sociocognitivista, argumenta que a modalidade deve ser considerada como operador sobre domínios dentro da cena comunicativa. A modalidade, segundo a autora, sinaliza sempre o trabalho com a face, regula a entrada em cena dos sujeitos participantes e o jogo de forças que se instaura nessa encenação.

Com base nas considerações apresentadas, propõe-se, neste trabalho, estudar a modalidade a partir de duas dimensões, pelas quais a categoria se realiza nos enunciados. Essas direções de modalidade, embora distintas, não são dicotômicas:

- a) a **orientada para o falante** – refere-se àquilo que o falante tem em mente (atitudes, crenças, julgamentos, etc.), serve para codificar experiência, especificar atitude em relação à validade do conteúdo modalizado no nível do conteúdo;

- b) a **orientada para o ouvinte** – marca estratégia comunicativa do falante, serve para regular a interação verbal, influir sobre os interlocutores.

No primeiro caso, a modalidade corresponderia ao entendimento mais convencional que a define como atitude ou opinião do sujeito enunciador em relação ao seu enunciado (LYONS, 1977; QUIRK et al., 1985; MAINGUENEAU, 1990; CORACINI, 1991; entre outros) e marcaria subjetividade.

Orientada para o ouvinte, a modalidade assume a função destacada por Hoffnagel (1997) e Miranda (2000), ou seja, é reguladora da situação comunicativa, sinalizando o trabalho de defesa e de proteção de faces (BROWN; LEVINSON, 1987).

Pode-se postular que essas duas direções da modalidade correspondem, respectivamente, às funções ideacional e interpessoal, como definidas por Halliday (1970, 1985), que não são excludentes, mas complementares. As sentenças são construídas simultaneamente como mensagem e como evento de interação (HALLIDAY, 1970; DIK, 1997a, 1997b). Como observa Halliday, “há uma região semântica em que as duas funções, a ideacional e a interpessoal, se sobrepõem, aquela do comentário do falante sobre o conteúdo” (HALLIDAY, 1970, p. 349).

Na **função ideacional**, a linguagem serve para expressar conteúdo, relacionado às experiências do falante, do mundo real ou mental (HALLIDAY, 1970, p. 325). Neves (1997, p. 12) esclarece que, no componente ideacional, as experiências do mundo mental podem envolver reações, cognições, percepções, assim como atos linguísticos de falar e de entender.

Na **função interpessoal**, por outro lado, a linguagem é utilizada como um meio de participação na situação de fala, ou seja, a linguagem relaciona-se com o estabelecimento de papéis sociais na comunicação, é um meio de estabelecer relações entre falante e ouvinte (HALLIDAY, 1970, p. 325). O elemento interpessoal serve, “num contexto mais amplo, ao estabelecimento e à manutenção dos papéis sociais, que, afinal, são inerentes à linguagem” (NEVES, 1997, p. 13).

As duas funções são entendidas como pontos extremos de um *continuum*, dentro do qual não há fronteiras rígidas entre o domínio ideacional e o interpessoal. Ou seja, não há relação dicotômica entre as duas funções, embora uma delas possa ser preponderante em determinado contexto.

A construção *pode ser (...), mas...* revela claramente essa multifuncionalidade da modalidade e, principalmente, a indica como categoria de gerenciamento e organização da interação comunicativa, conforme a análise que se segue.

### ***Pode ser (...), mas, estratégia de gerenciamento de pontos de vista divergentes***

Das 357 ocorrências da expressão *pode ser* (predicado encaixador de proposição ou construção sintaticamente independente) localizadas nas amostras de fala e de escrita, 84 (ou seja, 23%) são seguidas de uma objeção introduzida por *mas* (*pode ser (...), mas...*), como nas ocorrências (1) e (2), apresentadas na introdução.

A estratégia do enunciador, nessas sentenças, consiste em atribuir validade a uma possível objeção a seu raciocínio, mas considerá-la irrelevante. Em outras palavras, o enunciador parte de um pensamento oposto ao seu e sobre ele assenta sua declaração.

Considere-se a ocorrência (3):

- (3) Tendo ouvido falar que eu, quando menino, fui escoteiro, uma jovem leitora de São Paulo, Célia Maria, me escreve para dizer que, como bandeirante, tem curiosidade em saber o que significou o escotismo para mim.

Antes de mais nada: ainda existem escoteiros? A bem dizer, há muito tempo que não vejo um. A não ser que hajam passado a andar à paisana, como os padres e os militares mais discretos.

Segundo aquela definição lusitana, o escotismo vem a ser um bando de miúdos vestidos de parvos, comandados por um parvo, vestido de miúdo. **Pode ser** que hoje em dia assim seja, **mas** para mim foi um pouco mais que isso. (FE-CR)

O enunciador apresenta uma breve avaliação sobre o que representa o escotismo nos dias atuais, para, em seguida, introduzir, em sentido oposto, a sua impressão. Desse modo, o produtor do texto prevê um ponto de vista contrário ao seu e se antecipa a ele, admitindo-o e atribuindo-lhe alguma relevância (*pode ser que hoje em dia assim seja*). A concordância não é, entretanto, registrada sem resistência, não é absoluta. O fato admitido vem modalizado, situado no campo da incerteza, da possibilidade. O argumento definitivo (*para mim foi um pouco mais que isso*) surge em sentido contrário à admissão, tirando-lhe, assim, a validade absoluta.

Essa operação constitui um recurso na tarefa do *fazer crer*, na medida em que diminui a assertividade da declaração, ou, em outros termos, “prepara o terreno” para que a argumentação seja bem recebida pelo interlocutor. Um discurso sem concessões, autoritário, tem menos força argumentativa que um fundamentado na liberalidade, ainda que aparente.

A objeção assumida nesses enunciados pode ser advinda de um raciocínio do próprio falante que pressupõe posicionamentos contrários, como na ocorrência (3), mas também pode ser a retomada da fala do interlocutor. Esses últimos casos, naturalmente, ocorrem apenas em textos dialógicos:

- (4) Luiz Raul - Sempre tive a impressão que você tem um profundo desprezo pela gente... Você sempre age como se fosse melhor que todos nós. .. Tão superior, tão grave, tão coerente...

Léo - Isso não é verdade.

Luiz Raul - **Pode ser** que não seja, **mas** é a minha impressão... (RE-D)

O que se observa em ocorrências como (4) é que o enunciador finge admitir a validade da objeção. O falante apropria-se do enunciado do interlocutor e, modalizando-o como possível/eventual, diminui seu valor de verdade. Em seguida, ele reitera a validade de seu pensamento (*mas é a minha impressão*), o que significa que a admissão é apenas retórica.

Concordar com o outro, nesses enunciados, pode ser entendido como uma forma de o falante evitar um “confronto” direto, protegendo-se, assim, de uma possível contra-argumentação.

Essa admissão de objeção feita por meio da retomada de um argumento do interlocutor não é muito frequente com *pode ser* – predicado encaixador de proposição, como

em (4). Nesses casos, o mais comum é *pode ser* – construção independente. Note-se que a ocorrência (4) é perfeitamente parafraseável por “pode ser, mas é minha impressão”. Dado que esses são usos fortemente ligados ao domínio discursivo, na maioria dos casos, o enunciador opta pela forma mais reduzida, que, em geral, tem âmbito de incidência bastante claro, como na ocorrência (5):

- (5) — Quem? O Habib? Ah, ele é turco mas é ótima pessoa. Um grande amigo.  
— **Pode ser, mas** é muito esquisito. Eu ouvi o Martinho da venda contando que esse turco come carne crua. Não dessas assim, mal passadas. Crua mesmo. (ID)

Em (5), a concordância do enunciador é em relação ao enunciado anterior, e isso é facilmente recuperável no discurso: *eu concordo/ eu admito que ele seja ótima pessoa, um grande amigo*. Como nas ocorrências analisadas anteriormente, há a mesma operação de “avanço e recuo” em relação ao argumento alheio: o falante concorda com seu ouvinte, porém apresenta um argumento contrário (*é muito esquisito*), o qual prevalece.

Usos discursivos semelhantes ao de *pode ser* nesses enunciados são encontrados também em outras línguas. No inglês, Coates (1983, p. 134 apud NUYTS, 2001, p.224), Sweetser (1990) e Nuyts (2001), entre outros, identificaram usos discursivos semelhantes para o modal *may*. Nuyts (2001) observa que os modais *kunnen*, no holandês, e *können*, no alemão, também se apresentam nesse tipo de uso. Nesses enunciados, os modais indicam que, embora o enunciador não queira ou não possa rejeitar uma sugestão do interlocutor, ele também não quer aprová-la (NUYTS, 2001, p. 225).

Com base em sentenças com o verbo modal *may*, em construções muito próximas a *pode ser (...), mas...*, Sweetser (1990) postula que verbos modais se estendem do mundo epistêmico para o mundo conversacional (atos de fala). Analisem-se os enunciados (6) e (7), retirados de Sweetser (1990, p.70):

- (6) He may be a university professor, but he sure is dumb.  
(Ele pode ser um professor universitário, mas certamente é ignorante.)
- (7) He may be a university professor, but I doubt it because he’s so dumb.  
(Ele pode ser um professor universitário, mas eu duvido que seja, porque ele é muito ignorante)

A interpretação de sentenças como (6) pressupõe a verdade do primeiro segmento (*ele é um professor universitário*), que é uma admissão do enunciador situada no domínio discursivo. Sweetser (1990) propõe que esse enunciado pode ser entendido em termos de “no nosso mundo conversacional, **eu admito que** ele seja professor universitário, **apesar disso, eu insisto que** ele é ignorante”. Sweetser (1990) denomina esse tipo de modalização como “modalidade no nível dos atos de fala”.

Enunciados como (7), por outro lado, têm leitura epistêmica normal, pois envolvem uma conjectura do enunciador. Em (7), “ele pode ser professor universitário” é entendida como uma inferência do falante. Sweetser (1990) indica que uma paráfrase apropriada para essa sentença seria “**É possível que** ele seja professor universitário”.

De modo semelhante, no caso de *pode ser (...), mas*, algumas ocorrências estão mais ligadas ao domínio epistêmico e outras estão mais relacionados ao domínio discursivo. Comparem-se os enunciados (4) e (5), repetidos por conveniência, com o enunciado (8):

- (4) Luiz Raul - Sempre tive a impressão que você tem um profundo desprezo pela gente... Você sempre age como se fosse melhor que todos nós. .. Tão superior, tão grave, tão coerente...  
Léo - Isso não é verdade.  
Luiz Raul - **Pode ser** que não seja, **mas** é a minha impressão... (RE-D)
- (5) — Quem? O Habib? Ah, ele é turco mas é ótima pessoa. Um grande amigo.  
— **Pode ser, mas** é muito esquisito. Eu ouvi o Martinho da venda contando que esse turco come carne crua. Não dessas assim, mal passadas. Crua mesmo. (ID)
- (8) Deve ser coisa importante, pois ouvi a campainha tocar várias vezes, uma a caminho da porta e pelo menos três dentro do sonho. Vou regulando a vista, e começo a achar que conheço aquele rosto de um tempo distante e confuso. Ou senão cheguei dormindo ao olho mágico, e conheço aquele rosto de quando ele ainda pertencia ao sonho. Tem a barba. **Pode ser** que eu já tenha visto aquele rosto sem barba, **mas** a barba é tão sólida e rigorosa que parece anterior ao rosto. (EST)

Nas ocorrências (4) e (5), *pode ser* marca uma admissão situada no mundo conversacional. Por isso, a concordância em (4) e (5) é apenas aparente, serve como atenuadora da declaração do enunciador, conforme discutido, e a modalidade, nesse caso, é orientada para o ouvinte. Diferentemente, em (8), o enunciador expressa uma crença sobre a possibilidade de determinada proposição ser verdadeira e se tem, então, a modalidade orientada para o falante. Em (8), *pode ser* sinaliza uma avaliação do enunciador de que é possível ele já ter visto aquele rosto sem barba.

O que se verifica, no *corpus* da pesquisa, é que os casos em que a *pode ser* (predicado encaixador de proposição ou construção independente) se segue objeção introduzida por *mas*, a modalidade é em poucos casos como (8), orientada para o falante, sendo predominantemente orientada para o ouvinte, como (4) e (5).

Representando uma admissão falsa, assentada no mundo discursivo, a construção *pode ser (...)*, *mas* sinaliza, portanto, um jogo em que o enunciador finge concordar com uma opinião divergente a fim de diminuir o impacto negativo que sua declaração possa ter. Quando o enunciador utiliza esse mecanismo, revela consciência de que não está transitando por um território pacífico, ausente de controvérsias. O efeito que se obtém é a atenuação da força do contra-argumento, que, afinal, o enunciador introduz. Como tal, a operação *pode ser* pode ser entendida como um recurso de proteção e defesa de faces (BROWN; LEVINSON, 1987).

Esses usos podem ser considerados como marcadores de rejeição (*disclaimers*), na classificação dos marcadores de atenuação proposta por Rosa (1992) que é baseada em Fraser (1980). Os marcadores de rejeição precedem atos que são avaliados pelo enunciador como possivelmente ameaçadores da sua face ou da do interlocutor (críticas, proibições, ordens, etc.) e, então, têm a função de afastar, por antecipação, possíveis reações desfavoráveis do ouvinte.

A construção *pode ser (...)*, *mas* é um recurso que o enunciador utiliza para evitar um ato que possa ameaçar a face, tanto do falante, quanto do ouvinte. Ocorre que, ao mostrar-se atencioso com relação ao ponto de vista contrário ao seu, o enunciador busca minimizar a possibilidade de desaprovação, protegendo sua face. Simultaneamente, nesse movimento o enunciador preserva a face do interlocutor, pois não rejeita diretamente a sua opinião, evitando, assim, um ato que seja ameaçador à face do ouvinte.

## A relação adversativa em *pode ser (...) mas*

As construções adversativas caracterizam-se por apresentar sempre admissão e oposição. Nas palavras de Neves (1984, p. 23) reconhece-se uma entidade, “para posterior registro de sua desconsideração, negação, anulação, rejeição”. A autora propõe que a oposição vai de um mínimo, a condição de simples desigualdade, a um máximo, anulação, enquanto a admissão vai de um mínimo, o simples reconhecimento ou registro de existência, a um máximo, a concessão.

Conforme as análises apresentadas na seção anterior, a construção *pode ser*, quando seguida de *mas*, sinaliza uma admissão, em que fica clara uma concordância. Por isso, *pode ser* constitui um ambiente favorável à ocorrência de uma oração adversativa com *mas*. É o que se nota no enunciado (3), já apresentado:

- (3) Segundo aquela definição lusitana, o escotismo vem a ser um bando de miúdos vestidos de parvos, comandados por um parvo, vestido de miúdo. **Pode ser** que hoje em dia assim seja, **mas** para mim foi um pouco mais que isso. (FE-CR)

Admissão e oposição estão na base argumentativa desses enunciados. A primeira parte, concordância com uma possível objeção (*pode ser que hoje em dia assim seja*), aponta para uma conclusão que validará o fato admitido. No entanto, contrariando essa expectativa, a oração seguinte (*mas para mim foi um pouco mais que isso*) apresenta a opinião definitiva que encaminha a conclusão para uma direção oposta à da primeira oração.

A contrariedade a uma expectativa criada é uma invariante das construções adversativas e é, também, o ponto que as aproxima das construções concessivas. Tanto nas adversativas, quanto nas concessivas, o falante sente o peso da argumentação contrária e se antecipa, atribuindo alguma validade a essa objeção, porém, desconsiderando-a. Para Koch (1995), a diferença é de estratégia argumentativa:

Do ponto de vista semântico, os operadores do grupo *mas* e os do grupo *embora* têm funcionamento semelhante: eles opõem argumentos enunciados de perspectivas diferentes, que orientam, portanto, para conclusões contrárias. A diferença entre os dois grupos diz respeito à estratégia argumentativa utilizada pelo locutor: no caso de *mas*, ele emprega (segundo E. Guimarães) a “estratégia do suspense”, isto é, faz com que venha à mente do interlocutor a conclusão R, para depois introduzir o argumento (ou conjunto de argumentos) que irá levar à conclusão ~R; ao empregar o *embora*, o locutor utiliza a “estratégia de antecipação”, ou seja, anuncia, de antemão, que o argumento introduzido pelo *embora* vai ser anulado, “não vale”. (KOCH, 1995, p. 36-37)

Dada essa semelhança, as orações adversativas são parafraseáveis por orações concessivas, como preveem os esquemas de Neves (1999, p. 565-564), aqui adaptados, usando-se o exemplo (3) como padrão.

- Pode ser que hoje em dia assim seja, mas para mim foi um pouco mais que isso.

### **Esquema adversativo (admissão, assentimento)**

(a) **eu admito / eu concordo** que hoje em dia assim seja.

(b) **de qualquer modo / ainda assim**, para mim foi um pouco mais que isso.

- Embora hoje em dia assim seja, para mim foi um pouco mais que isso.

**Esquema concessivo (refutação de uma possível objeção)**

(a) **você/ alguém** pode me objetar que hoje em dia assim seja e eu **não desconheço isso**.

(b) **de qualquer modo / ainda assim**, para mim foi um pouco mais que isso.

Esses esquemas mostram que a adversativa salienta a admissão de uma objeção, enquanto a concessiva enfatiza a refutação de uma possível objeção. O ponto de contato entre as construções é o não atingimento da expectativa, marcada por uma relação contrastiva.

O valor opositivo do *mas* é estabelecido segundo diferentes escalas. Neves (1984) subdivide os enunciados com *mas* em dois grandes grupos, contraposição (as duas entidades relacionadas por *mas* são confrontadas, distinguidas) e eliminação (o segundo membro coordenado elimina o primeiro, de algum modo), os quais abrigam outros subagrupamentos. A contraposição pode estabelecer-se em direções opostas, na mesma direção ou em paralelo. A eliminação pode ocorrer ligada a uma relação temporal ou não, com ou sem re colocação.

Analisando a oposição que o *mas* estabelece nas construções do tipo *pode ser (...)*, *mas* verifica-se que o segmento introduzido por *mas* não anula totalmente a admissão apresentada pelo segmento anterior. O efeito de sentido dessas construções é que o argumento contrário possui alguma validade para o enunciador.

O *mas* que coocorre com *pode ser*, se visto de acordo com a proposta de Neves (1984), situa-se no primeiro grupo, a contraposição, já que relaciona argumentos em direções opostas ou em paralelo, como exemplificam as ocorrências seguintes:

- (9) **Pode ser que** o engenheiro de algumas escolas, de saída, não esteja com uma formação obsoleta, **mas** certamente pode ter conhecimentos totalmente insuficientes (PT-T)
- (2) — Deixa de besteira, Luís. Ela deu o colar a tia Matilde, em Poços.  
— **Pode ser... Mas** abre logo essa mão e passa minha parte. (COT)

Em ambos os casos, estabelece-se uma relação de desigualdade, que é a relação que Neves (1984) aponta como invariante no uso de *mas*, o que se liga à etimologia da palavra. Em (9), o segmento *mas certamente pode ter conhecimentos totalmente insuficientes* restringe a admissão anterior (*o engenheiro de algumas escolas, de saída, não está com uma formação obsoleta*), marcando uma oposição em direção oposta (NEVES, 1984). Em (2), o segmento *mas abre logo essa mão e passa minha parte* não marca uma oposição àquilo que foi admitido (*ela deu o colar para a Tia Matilde*), e sim acrescenta um novo foco à conversação. Desse modo, considera-se que haja, em (2), uma contraposição em direção paralela (NEVES, 1984).

A divergência entre os dois valores de *mas* em (9) e (2) deve ser vista em termos discursivos. A relação de desigualdade em (9) está assentada no campo do conhecimento, das ideias (mundo epistêmico), enquanto em enunciados como (2) o *mas* marca um desnível no campo discursivo (mundo conversacional, dos atos de fala). Note-se que, com *pode ser, mas...*, em (2), o enunciador muda o rumo da interação entre os interlocutores, indicando que não está interessado em continuar interagindo no plano das ideias, pois pretende obter a realização de uma ação.

Uma proposta de exame do item *mas* que vem ao encontro de tais considerações é a apresentada por Castilho (2010), que examina o valor do conectivo *mas* no sistema discursivo e no sistema gramatical, relacionando-o à sua gramaticalização. Como Neves (1984), o autor lembra que *mas* deriva do advérbio latino *magis*, cujo valor semântico básico era estabelecer comparações de quantidade e de qualidade e, secundariamente, servia para marcar a inclusão de indivíduos em um conjunto. O valor adversativo surgiu por meio de um processo metonímico, em que *mas* passa a atuar, no sistema da gramática, como uma conjunção de contração. Castilho mostra que, no sistema discursivo, como marcador discursivo e como conectivo textual, *mas* tem preservado o valor de inclusão, como se verifica no enunciado (2). Nesses casos, diz-se que o *mas* possui um valor de inclusão, pois introduz um argumento, uma ideia nova, mudando o rumo do discurso.

Nem sempre é fácil, porém, distinguir o valor de contração e de acréscimo em *mas*. Segundo Castilho (2010), esses valores coexistem em diferentes expressões sincrônicas da língua, podendo situar-se em um eixo que iria de *inclusão* para *contração*, sem muita nitidez separando um uso do outro.

A análise das ocorrências revela que, nos usos mais discursivos de *pode ser*, o *mas* acrescenta algo novo, como exemplificado em (2). Porém, mesmo nesses casos, o conectivo marca um desnível, o que corrobora a postulação de Neves (1984), segundo a qual a noção de desigualdade é básica no valor de *mas*.

## Considerações finais

O trabalho buscou enfatizar a modalidade como indexadora de uma estratégia pragmática do enunciador. A pesquisa revela a modalidade como uma categoria multifuncional que não se limita à expressão de um constructo do falante em relação a uma proposição, como muitas vezes está definida nos estudos, mas que marca, principalmente, a participação do enunciador na comunicação, realizada sob certas pressões interacionais e com o intuito de alcançar determinados objetivos comunicativos ou determinados efeitos discursivos.

O estudo da modalidade não pode desconsiderar tal multifuncionalidade da categoria. Desse modo, propôs-se, neste trabalho estudar a modalidade reconhecendo-se duas direções: orientada para o ouvinte e orientada para o falante. No primeiro caso, a modalidade corresponde a uma estratégia pragmática, serve para influir sobre os interlocutores, é gerenciadora da interação comunicativa. No último caso, a modalidade indexa uma experiência do enunciador, especifica atitude em relação à validade do conteúdo modalizado, no nível do conteúdo. As duas direções de modalidade são distintas, mas, de forma alguma, excludentes.

As análises indicam que a fórmula *pode ser (...), mas...* destaca grandemente o papel da modalidade como peça de gerenciamento discursivo, orientada para o ouvinte. Nesses contextos, com *pode ser*, o enunciador admite uma contra-argumentação, atribuindo-lhe alguma relevância. A objeção introduzida por *mas* traz o argumento definitivo. A admissão (*pode ser*) é apenas aparente, é uma forma de evitar um “confronto direto”.

Verificou-se, ainda, que os casos em que a construção não funciona como gerenciadora da “tensão” comunicativa, e sim como orientada para o falante, são pouco frequentes.

Por fim, o estudo da relação adversativa nessas ocorrências confirma a operação de admissão e oposição, de avanço e recuo e indicam que há usos mais ligados à interação

comunicativa em que o conectivo *mas* tende a marcar um desnível no campo discursivo e não no campo do conhecimento.

## REFERÊNCIAS

BROWN, P.; LEVINSON, S. C. *Politeness. Some universals in language usage*. Cambridge: Cambridge University Press, 1987.

CARRASCOSSI, C. N. S. *Gramaticalização e (inter)subjetivização na modalização em português: um estudo de pode ser*. 2011. 170 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Unesp, Araraquara.

CASTILHO, A. T. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2010.

COATES, J. *The Semantics of the Modal Auxiliaries*. London: Croom Helm, 1983.

CORACINI, M. J. R. F. E a questão da modalidade? In: \_\_\_\_\_. *Um fazer persuasivo. O discurso subjetivo da ciência*. São Paulo: Pontes/Educ, 1991. p. 112-132.

DIK, S. C. *The Theory of Functional Grammar*. Part 1: The Structure of the Clause. 2. ed. rev. Kees Hengeveld (Ed.). Berlin, New York: Mouton de Gruyter, 1997a.

\_\_\_\_\_. *The theory of Functional Grammar*. Part 2: Complex and derived constructions. 2. ed. Kees Hengeveld (Ed.). Berlin, New York: Mouton de Gruyter, 1997b.

FRASER, B. Conversational Mitigation. *Journal of Pragmatics*, v. 4, p. 341-350, 1980.

HALLIDAY, M. A. K. Functional Diversity in Language as Seen from a Consideration of Modality and Mood in English. *Foundations of Language*, v. 6, n. 3, p. 322-361, 1970.

\_\_\_\_\_. *An Introduction to Functional Grammar*. Baltimore: Edward Arnold, 1985.

HOFFNAGEL, J. C. A modalização epistêmica no processamento textual da fala. In: KOCH, I. G. V.; BARROS, K. S. M. (Org.) *Tópicos em linguística do texto e análise da conversação*. Natal: EDUFERN, 1997. p. 147-156.

KIEFER, F. On Defining Modality. *Folia linguística*, v. 21, n. 1, p. 67-93, 1987.

KOCH, I. G. V. A articulação entre orações no texto. *Cadernos de Estudos Linguísticos*, Campinas, n. 28, p. 9-18, jan./jun.1995.

LYONS, J. *Semantics 1*. Cambridge: Cambridge University Press, 1977.

MAINGUENEAU, D. *Éléments de linguistique pour texte littéraire*. Paris: Bordas, 1990.

MIRANDA, N. S. *A configuração das arenas comunicativas no discurso institucional: professores versus professores*. 2000. 196 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal de Minas Gerais, UFMG, Belo Horizonte.

NARROG, H. On defining modality again. *Language Sciences*, v. 27, p. 165-192, 2005.

NEVES, M. H. M. O coordenador interfrasal *mas* – invariância e variantes. *Alfa*, São Paulo, v. 28, p. 21-42, 1984.

\_\_\_\_\_. *A gramática funcional*. São Paulo: Martins Fontes, 1997.

\_\_\_\_\_. As construções concessivas. In: NEVES, M. H. M. (Org.) *Gramática do português falado: novos estudos*. v.7. Campinas: Ed. da Unicamp, 1999. p. 545-591.

\_\_\_\_\_. Imprimir marcas no enunciado. Ou: a modalização na linguagem. In: \_\_\_\_\_. *Texto e gramática*. São Paulo: Contexto, 2006. p. 151-166.

NUYTS, J. *Epistemic modality, Language, and Conceptualization: a cognitive-pragmatic perspective*. John Benjamins: Amsterdam/Philadelphia, 2001.

PALMER, F. R. *Mood and Modality*. Cambridge: Cambridge University Press, 1986.

QUIRK, R.; GREENBAUM, L.; LEECH, G.; SVARTVIK, J. *A Grammar of Contemporary English*. London: Longman, 1985.

ROSA, M. *Marcadores de atenuação*. São Paulo: Contexto, 1992.

SWEETSER, E. E. Modality. In: *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: University Press, 1990. p. 49-75.

# Variação e mudança linguísticas e preservação da informação

(Variation and linguistic change and preservation of information)

**Deize Crespim Pereira**

Faculdade de Filosofia, Letras e Ciências Humanas – Universidade de São Paulo (USP)

deize.pereira@usp.br

**Abstract:** The aim of this paper is to consider the influence of the linguistic factor informational status on two phenomena under linguistic variation and change in Popular Brazilian Portuguese: (i) explicitness x omission of reflexive pronouns, and (ii) application x non-application of the subject-verb agreement rule with the 1<sup>st</sup> plural person. The data under study consist of interviews from the project *Português Popular em São Paulo* (*Popular Portuguese spoken in São Paulo*) and the project *Filologia Bandeirante* (“*Bandeirante*” *Philology*). The theoretical and methodological tools are those from Functional Linguistics and Labovian Variationist Sociolinguistics. The results of data’s quantitative analysis show a tendency to preserve referential information in contexts in which it conveys new information. Thus, the subject-verb inflectional agreement as well as the reflexive pronouns tend to be explicit when their absence results in loss of information.

**Keywords:** Functional Linguistics; preservation of information; Popular Brazilian Portuguese

**Resumo:** O objetivo deste trabalho é refletir sobre a influência do fator linguístico *status* informacional em dois fenômenos de variação e mudança linguística no português popular brasileiro: (i) a realização x não-realização do pronome reflexivo e (ii) a aplicação x não-aplicação da concordância verbal de 1<sup>a</sup> pessoa do plural. O *corpus* se compõe de inquéritos do Projeto *Português Popular em São Paulo* e do Projeto *Filologia Bandeirante*. Os pressupostos teórico-metodológicos são retirados da Linguística Funcional e da Sociolinguística Variacionista Laboviana. Os resultados da análise quantitativa dos dados mostram que existe uma tendência de preservar a informação referencial em contextos nos quais esta veicula informação nova. Assim, tanto a concordância verbal quanto o pronome reflexivo tendem a ser realizados pelo falante, quando sua ausência implica comprometimento da informação.

**Palavras-chave:** Linguística Funcional; preservação da informação; Português Popular Brasileiro

## Introdução

O objetivo deste trabalho é refletir sobre a influência do fator linguístico *status* informacional em dois fenômenos de variação e mudança linguística no português popular brasileiro: (i) a realização x não-realização do pronome reflexivo e (ii) a aplicação x não-aplicação da concordância verbal de 1<sup>a</sup> pessoa do plural.

O *corpus* sob análise se compõe de inquéritos que integram dois projetos: *Português Popular em São Paulo* e *Filologia Bandeirante*.<sup>1</sup> O Projeto *Português Popular em São Paulo* se compõe de entrevistas com informantes adultos, de ambos os sexos, de baixa ou nula escolaridade, paulistanos e migrantes de vários estados brasileiros, que moram em diversas favelas da cidade de São Paulo. O Projeto *Filologia Bandeirante* é composto de inquéritos de informantes idosos de ambos os sexos, analfabetos ou de baixa escolaridade,

---

1 Foram utilizados 72 inquéritos do Projeto *Português Popular em São Paulo*, para analisar a realização x não-realização dos pronomes reflexivos, e 15 inquéritos do Projeto *Filologia Bandeirante*, para analisar a aplicação x não-aplicação da concordância verbal de 1<sup>a</sup> pessoa do plural.

nascidos e criados na zona rural dos estados de São Paulo e de Minas Gerais, na área correspondente às trilhas históricas das bandeiras paulistas.

Os pressupostos teórico-metodológicos são retirados da Linguística Funcional (HALLIDAY, 1987, 1994; DIK, 1989, 1997) e da Sociolinguística Variacionista Laboviana (LABOV, 1991, 1994, 2001). Para a análise estatística dos dados foi utilizado o programa computacional Goldvarb.

## **Pressupostos teóricos**

O ponto de partida do presente trabalho são as reflexões de Labov (1994), apresentadas em sua obra *Principles of Linguistic Change: Internal Factors*, mais especificamente, nos capítulos 19 e 20, respectivamente intitulados “The Overestimation of Functionalism” e “The Maintenance of Meaning”. Nestes capítulos, Labov faz uma crítica aos funcionalistas, argumentando que fatores de ordem funcional não dão conta de explicar a variação e a mudança linguísticas. Os fatores de ordem funcional são entendidos como aqueles que têm o efeito de preservar a informação. Vejamos, pois, um breve resumo da visão funcionalista da variação e mudança.

Os linguistas funcionalistas postulam que a principal função da língua é a comunicação. Assim, em conformidade com a maneira como veem a língua, os funcionalistas hipotetizam que a variação e a mudança linguísticas são governadas pela função comunicativa. Halliday (1987, 1994), por exemplo, argumenta que a língua muda justamente para satisfazer melhor as necessidades comunicativas dos falantes. O uso linguístico é governado pela necessidade de comunicar informação referencial, e há uma tendência geral de evitar a perda de informação.

Explicações funcionais da variação e mudança linguísticas pressupõem que os falantes, quando formulam as expressões, levam em conta a informação pragmática compartilhada com o ouvinte (conferir modelo de interação verbal de DIK, 1989) e, se houver a possibilidade de escolha entre duas formas alternativas (em termos sociolinguísticos, duas “variantes”), eles favorecem aquela que transmite o sentido da forma mais eficiente e efetiva.

Uma outra hipótese formulada pelos funcionalistas é que a mudança linguística pode ser barrada nos contextos em que implica comprometimento da informação. Para os funcionalistas, existe uma tendência de a informação semanticamente relevante ser mantida na estrutura superficial (KIPARSKY, 1982 apud LABOV, 1994). Assim, uma mudança fonética, por exemplo, pode não ocorrer em determinados contextos nos quais levaria à perda de distinções morfológicas.

Uma análise funcionalista da variação e mudança parte do pressuposto de que nenhuma informação é totalmente perdida, uma vez que sempre há informação redundante. Logo, a supressão de um determinado item linguístico seria condicionada pela quantidade de informação presente no entorno linguístico e no contexto situacional.

Labov (1994) nota que estudos variacionistas tentaram verificar a validade destas hipóteses funcionalistas, mas, segundo ele, tais estudos não as comprovaram. Há muitas mudanças fonéticas que apagam flexões, levando à perda de distinções morfológicas. Sua opinião, portanto, é que a necessidade de comunicar não governaria a variação e a mudança.

Estas podem, ainda segundo Labov, ser explicadas pela atuação de fatores estruturais e mecânicos, cegos para as necessidades comunicativas dos falantes.

Para exemplificar esses fatores mecânicos, Labov menciona vários fenômenos de variação sincrônica fonológica e morfológica, entre os quais: (i) a variação entre realização e não-realização de *-s* morfema de plural em sintagmas nominais, no espanhol porto-riquenho da Filadélfia (POPLACK, 1980 apud LABOV, 1994), a qual seria condicionada somente por fatores de natureza fonética; (ii) a variação da concordância verbal no português do Brasil (SCHERRE; NARO, 1993), a qual é fortemente influenciada pelo fator paralelismo – marcas levam a marcas, e zeros levam a zeros, isto é, verbos marcados tendem a ser seguidos por verbos igualmente marcados, enquanto verbos não-marcados tendem a ser seguidos por verbos igualmente não-marcados. Para Labov (1994), o paralelismo faz com que a flexão do verbo seja omitida quando necessária, e utilizada quando redundante.

O autor, contudo, não chega a ser tão radical a ponto de negar totalmente a influência de fatores funcionais. Segundo ele, estes não são capazes de explicar o *processo* de variação e mudança, mas podem aparecer como *produtos* da mudança linguística, quando esta finalmente se completa.

When language changes, its information carrying capacity is often threatened; but in the long run, most languages do preserve their means of conveying information, more or less, by one route or another. Though speakers may not behave wisely and thoughtfully as they choose one variant or the other, somehow the system does react. (LABOV, 1994, p. 568)

It is important to note that in the course of language evolution, change does go to completion, and variable rules become invariant. When this happens, there is inevitably some other structural change to compensate for the loss of information involved. (LABOV, 1991, p. 223)

Principle of structural compensation: when the rate of deletion of a meaningful feature of a language increases, the frequency of features that redundantly carry this meaning will increase. (LABOV, 1994, p. 604)

Para Labov, portanto, quando uma mudança linguística se completa, o sistema tende a se reajustar para preservar a informação que foi perdida. Esse reajuste do sistema é explanado a partir do conceito que Labov denomina “*probability matching*”: mecanismo de aprendizagem inconsciente, através do qual os indivíduos percebem as formas linguísticas mais frequentes, passando a empregá-las em sua fala. Um exemplo deste reajuste funcional é o aumento na frequência de preenchimento do sujeito por pronomes, para compensar a perda de flexão número-pessoal no paradigma verbal (HOCHBERG, 1986 apud LABOV, 1994).

Este aumento na frequência de uso de um determinado item linguístico, para compensar a perda de informação ocasionada pela supressão de outro, seria motivado pelo aumento nos casos de incompreensão gerados pela mudança linguística (cf. KROCH, 1989 apud LABOV, 1994). Em outras palavras, os falantes não conseguem se entender e inconscientemente passam a empregar formas que veiculam a informação suprimida pela mudança linguística.

Labov (1994) lança a hipótese de que os casos de incompreensão entre interlocutores sejam muito frequentes na comunicação diária. Contudo, passam despercebidos, e somente são notados quando o modo como o ouvinte entendeu a expressão não se encaixa com determinada situação pragmática.

Os contextos que geram desentendimentos na comunicação diária seriam os mesmos que levam os linguistas à análise e codificação equivocada dos dados. Estas são apontadas como uma das principais causas de se ter superestimado os fatores funcionais.

Para ilustrar como se dá esta codificação equivocada de ocorrências, Labov cita o estudo de Poplack (1980 apud LABOV, 1994). Esta autora constata que a supressão de *-s*, morfema de plural nos sintagmas nominais (*las cosas bonitas*), e de *-n*, morfema de 3ª pessoa de plural nos verbos (*mandan*), jamais ocorrem quando não há informação desambiguadora nos contextos linguístico ou situacional, que indique que se trata de uma forma de plural. Mas, segundo Labov (1994), este achado não pode ser tomado como argumento a favor da hipótese funcionalista de preservação da informação, porque casos totalmente não-marcados (i.e. em que não há marca formal de plural em nenhum elemento do sintagma nominal nem no verbo, tampouco informação contextual que indique que se trata de uma forma de plural) devem necessariamente ter passado despercebidos pelo linguista.

Vale ressaltar que uma posição funcionalista extrema não dá margem para sequer se admitir esta possibilidade. Uma visão funcionalista da língua assume que, se o falante quis expressar o plural, ele vai necessariamente explicitar esta informação de algum modo em seu discurso.

De fato, o postulado de que a função comunicativa da língua governa seu uso é caro para os funcionalistas. O próprio Labov (2001) reconhece que os falantes querem transmitir para o ouvinte o conteúdo proposicional do que eles estão dizendo, e ficam irritados e confusos quando não são bem sucedidos. Tais reflexões serviram de ponto de partida para o desenvolvimento deste artigo.

Em estudos anteriores, empreendi análise quantitativa da variação na realização da concordância verbal de 1ª pessoa do plural (PEREIRA, 2004) e da variação no uso dos pronomes reflexivos (PEREIRA, 2007) no português popular brasileiro, considerando vários fatores, tanto de natureza linguística como de natureza social. No presente trabalho, utilizo estes dados com o objetivo de refletir sobre o peso do fator *status* informacional na variação e mudança linguísticas.

Através deste fator, formulado com base em Halliday (1994) e Dik (1989, 1997), analiso o discurso em termos de informações dadas e novas. A informação dada é entendida como aquela que é dada/conhecida para o ouvinte, algo que já foi apresentado no contexto verbal ou situacional. O Dado é geralmente um elemento fórico. Por ser passível de ser recuperado, pode ser elíptico.

A informação nova, por outro lado, é aquela que é apresentada como nova/não-conhecida para o ouvinte. Tipicamente ela é nova porque não havia sido mencionada anteriormente, mas pode ser também algo inesperado ou importante, previamente mencionado ou não, algo que o falante apresenta como novo. A parte nova pode ser assinalada por proeminência tônica, o que caracteriza a parte mais importante, o Foco Informacional.

### **Variação no uso dos pronomes reflexivos**

A análise do *status* informacional consiste em verificar se o pronome reflexivo veicula informação dada ou nova. Do ponto de vista discursivo e semântico, o pronome reflexivo serve para transmitir a informação de que o sujeito e o objeto do verbo têm

o mesmo referente, são correferenciais. Vejamos o exemplo abaixo, retirado de Lima-Hernandes (2004).

Doc. Alaga a rua toda! O senhor não tem vontade de mudar um dia? / Inf. Como assim? Mudar o que? Mudar em que?

Neste contexto, a forma reflexiva, que foi omitida (*mudar-se*) constituía informação nova, e a sua omissão ocasionou comprometimento da informação, uma vez que o entrevistado não compreendeu a pergunta do documentador, isto é, não entendeu a identidade entre sujeito e objeto (*mudar-se ele mesmo* de lugar).

Há, por outro lado, ocorrências nas quais esta identidade claramente constitui informação dada: a omissão do pronome reflexivo não prejudica a informação, e a sua presença é redundante, porque o verbo normalmente é empregado com sentido reflexivo e, por assim dizer, já incorpora este sentido. Exemplos:

- (1) Quando eu lembro meu coração dói (I.p, p.11)
- (2) meu pai casô de novo, hoje tem mais cinco filho (I.e, p.12)
- (3) eu me arrependi tanto (I.r, p.3)

Mesmo entre os verbos que prototipicamente não se empregam reflexivamente (*entender, considerar*), ou que podem ou não ser usados na forma reflexiva (*divertir*), o entorno discursivo pode esclarecer o sentido reflexivo. Em outras palavras, a presença de um complemento ou predicativo (em itálico) pode tornar o sentido reflexivo informação dada. Exemplos:

- (4) eu ficava sozinha, é que entrava gente aqueles caipira, entrava muita gente saía, uns fazia baldeação, eu divertia ali, ( ) mais tinha aquelas coisa de (...) (I.z,p.17)
- (5) Eu sempre trabalhei desde que eu entendi por gente de idade de déis ano pra ( ) até uns seis anos atrás. Trabalhava sempre não tinha domingo nem feriado nem nada. (I.z,p.14)
- (6) mas nós sofremos demais aqui ... viu?... hoje em dia nós pode considerar rico/ Doc. é verdade/ Inf. é... tem água tem luz dentro de casa. (I.P<sup>2</sup>.p.16)

Tomando por base a perspectiva funcionalista, analisa-se o discurso como um processo e não como um produto. Logo, se o sentido reflexivo somente é esclarecido em orações posteriores (em itálico) à ocorrência, esta é considerada um contexto de informação nova, mesmo que o pronome não tenha sido realizado pelo falante. Exs.:

- (7) Inf. e a gente tá aqui esse tempo todo / Doc. E você acostumô aqui?/ Inf. Costumei mais tenho tanta vontade de saí daqui não por mim pelo meus filho sabe? Eu não quero que eles cria cê vê né? era pra gente vim ficá um ano dois ano vai fazê cinco ano e meio que a gente tá aqui/ Doc. Bastante tempo/ Inf. Então eu tive dois filho aqui. Ave Maria não quero que *eles cria aqui* não (I.y,p.29)
- (8) Inf. A fera é boa. Eu comprava num senhor ali vô largá né? e agora apertei fais vai fazê dois meis que *eu não dô dinheiro pra ele*. (I.z,p.12)
- (9) Eu que num ligo pra aparência não. Ligo é pro caráter das pessoa... então eu falo é... às veze a irmã fala que eu sou muito chata. Num é chata a pessoa tem que dá valô. *Num é qualqué carinha que mexe que deve ficá derreteno pra ele*. (I.9<sup>1</sup>,p.36)

O caso mais claro de informação nova é representado por instâncias em que a identidade, entre o referente do sujeito e outro referente da mesma oração, constitui informação totalmente inesperada, tomando o discurso imediatamente anterior. Na ocorrência a seguir, por exemplo, a ausência do pronome levaria a uma interpretação não-reflexiva: “esse [torcedor fanático] é capaz de matar [alguém] por aquilo ali”.

- (10) Inf. eu não não sô eu sô um eu sô o seguinte sô um torcedô, mais não sô um torcedô fanático. Porque ((grito da filha)) inxiste inxiste o torcedô fanático né? /Doc. É, realmente / Inf. O fanático é esse que sai daqui com chuva com sol ele vai lá discuti / Doc. e com radinho né? / Inf. É, esse é o fanático, então esse aí é capais de se matá por aquilo ali, mais eu não sô um fanático. (I.2,p.12)

Muitas das ocorrências de informação nova são gramaticalmente marcadas como Foco, isto é, informação saliente, através principalmente do uso da estrutura clivada, mas também de formas pronominais tônicas, como em (11-13).

- (11) Vô te falá, eu truxe esse troço tudo na cabeça minha filha, lá de baxo praqui e ele não queria nem sabê. Falava “Que nada cê que se lasque” (I.a,p.12)
- (12) eu fiquei internada duas vez com pneumonia. Aí eu falei “Sabe duma coisa? Eu vô cuidá é de mim” (I.f,p.9)
- (13) Era eu mesmo que tinha que me virá tinha que í lá nos cafundó do juda pedi caxão porque não podia comprá mesmo né? (I.f,p.16)

Os dados incluem, por fim, algumas ocorrências ambíguas (exemplos 14-16), com o intuito de capturar os zeros que estão faltando (“*zeroes missing*”) mencionados por Labov (1994), isto é, as ocorrências de zero que normalmente escapam do linguista, justamente pelo fato de constituírem informação nova e não serem marcadas. Segundo o autor, estes zeros, que passam despercebidos e não são recolhidos, foram a causa de se ter superestimado o peso de fatores de ordem funcional em fenômenos de variação e mudança.

Nas ocorrências a seguir, o sentido reflexivo constitui informação nova. Na verdade, elas contêm referentes em competição (em itálico>) e podem ser entendidas tanto como reflexivas, quanto como compreendendo dois referentes distintos.

- (14-15) Mais eu eu nessa hora eu nem sabia que (pode) que que eu fazia. Ela “acode mãe que a nenê *se machucô* a nenê caiu da cama”. Eu pensei na hora que eu peguei a menina que vi lavada de sangue eu pensei ( ) a menina que bateu *a cabeça* na perna da cama e tinha cortado. Que eu peguei ela a coitadinha tava até sem fala ((ri)) (...) aí elas danô a gritá “ô (marido) a nenê cortô olha pra nenê” e eu comecei também ficá desesperada de vê tanto sangue num nenenzinho (I.t,p.13)
- (16) Inf. não, esse aí não foi eu que fiz não mais a gente pinta isso aí / Doc. ahn ahn e e/ Inf. Fazia mais era era pintava mais pano de prato, fazia *cadás coisinha bonita* que as muié admirava “Ah mais óia, dona Geralda ( ) tá fazendo cada *trem bonito*, fazia aquelas florona bacana” ((ri)) *Elas se admirava* / Doc. Gostoso. E a senhora não fais pra vendê assim? (I.n,p.11)

**Tabela 1: frequência e peso relativo de realização do pronome reflexivo de acordo com o status informacional (*corpus* do Projeto *Português Popular em São Paulo*)**

Status Informacional	Frequência	Peso relativo
Dado	437/1401=31%	0.40
Novo	258/291=89%	0.88
<b>Total</b>	695/1692=41%	

A tabela 1 indica que a variação pronominal é fortemente influenciada pelo fator *status* informacional. O falante tende a realizar o pronome quando este veicula informação nova no discurso (frequência: 89%; peso relativo: 0.88). Na maior parte das ocorrências encontradas no *corpus*, porém, o pronome corresponde à informação dada, conhecida para o ouvinte. Neste contexto, em que a presença do pronome é redundante, predomina a não-realização (frequência: 31%; peso relativo: 0.40).

### **Variação da concordância verbal de 1ª pessoa do plural**

Do ponto de vista informacional, a concordância verbal serve para assinalar quem é o sujeito do verbo; ela transmite, portanto, a informação relativa ao referente do sujeito. Esta informação pode ser veiculada tanto pelo sujeito explícito, quanto pela flexão número-pessoal do verbo. Nas instâncias em que o referente constitui informação dada, tanto o sujeito quanto a flexão número-pessoal do verbo podem ser omitidos; já nas instâncias de referente Novo no discurso, ao menos uma das duas formas de explicitação deste referente terá de ser utilizada (i.e. o sujeito ou a flexão verbal).

Na análise do *status* informacional do sujeito, a questão foi colocada do seguinte modo: o referente de um determinado verbo – independentemente de o sujeito ser explícito ou zero, independentemente de o verbo estar ou não flexionado – constitui informação conhecida ou nova? Um teste para determinar o status informacional do sujeito consiste em extrair a flexão número-pessoal do verbo e o sujeito (se estes estiverem presentes). Caso seja possível, ainda assim, depreender o referente do verbo, este constitui informação dada.

Examinemos os exemplos a seguir.

(17) trabaiava aqui *eu i us dois fio meu* (I.13,p.192)

No exemplo (17), o sujeito explícito composto, posposto ao verbo, representa informação nova, já que é a primeira vez em que aparece no discurso. Tal sujeito é introduzido inclusive no lugar típico da informação nova, que geralmente ocupa uma posição mais ao final da sentença. O falante está, portanto, introduzindo este referente no discurso.

(18-21) *nói era* muito pobrezinho nu tinha nada...é...tinha nada só usava uma prantinha (I.2,p.19)

Em (18-21), o sujeito da primeira sentença é explícito, representado pelo pronome *nós*. Neste contexto também representa informação nova, porque está sendo introduzido no discurso. Nas demais instâncias em que a casa vazia do sujeito retoma “*nós*”, sendo o verbo não-marcado quanto à flexão número-pessoal, o sujeito é dado, uma vez que já sabemos que o seu referente é *nós*.

- (22) *são tiago* hoje é *uma cidade* boa pra *nóis* que somo aqui...*desse lugarzinho pobre* né? graças a deus temos bons professores boas escola...bons padre (I.1,p.15)

Em (22) há o uso do sujeito zero e o verbo flexionado na 1ª pessoa do plural. Neste contexto, a flexão representa informação nova. Isto ocorre porque o verbo tem um antecedente (*nóis*), mas a não-realização da flexão verbal implicaria ambiguidade na identificação do referente, o qual poderia ser interpretado como a cidade: “São Tiago tem bons professores, boas escolas, bons padres”.

- (23) aqui na estrada memo otu dia umas par de veis *nóis* matemo cascavel nessa estrada aqui vem do artu pa bebê água nu riberão né? lá de vorta ela no barro ela num sobi...ali memo fai uns doi meis *nói* matemo dua nu pé daqueli pau lá naquele naquele postu lá (I.4,p.57)

No exemplo (23), a ausência do sujeito pronominal ou da flexão dá margem para uma interpretação do referente do verbo como indeterminado. Se o falante dissesse *matou duas*, a casa vazia do sujeito poderia ser interpretada como de referência indeterminada. Logo, neste contexto, o referente do verbo constitui informação nova.

- (24) *foi foi nós* que fizemo (I.2,p.34)

- (25) *nóis* é que tinha que fazê o (nosso) separado (I.3,p.40)

- (26) fico comendo o churrasco que *nói* paguemo ... ((risadas)) *nói* que paguemo (I.5,p.72)

Os exemplos (24-26) ilustram casos em que o referente do sujeito é entonacionalmente e estruturalmente marcado como informação nova através do uso da clivada.

**Tabela 2: frequência e peso relativo de realização da regra de concordância verbal de 1ª pessoa do plural de acordo com o status informacional (corpus do Projeto Filologia Bandeirante)**

Status informacional	Frequência	Peso relativo
Dado	20/111 = 18%	.22
Novo	85/206 = 41%	.67
Total	105/317 = 33%	

Os resultados exibidos na tabela 2 evidenciam que a concordância é favorecida, quando o sujeito constitui informação nova (frequência: 41%; peso relativo: 0.67), e decisivamente desfavorecida, se o referente equivale a informação dada (frequência: 18%; peso relativo: 0.22).

Na medida em que a informação relativa ao referente do verbo pode ser transmitida tanto pela flexão número-pessoal como pelo sujeito explícito, a análise incluiu o cruzamento de fatores, de modo a examinar concomitantemente ambas as possibilidades de explicitação do referente. Os resultados são exibidos nas tabelas 3 e 4.

**Tabela 3: frequência de realização do sujeito conforme o status informacional (*corpus* do Projeto *Filologia Bandeirante*)**

Realização do sujeito	Status informacional	
	Sujeito Dado	Sujeito Novo
Sujeito explícito	34/111 = 31%	177/206 = 86%
Sujeito zero	77/111 = 69%	29/206 = 14%

**Tabela 4: frequência de realização da regra de concordância verbal de 1ª pessoa do plural de acordo com a presença/ausência do sujeito e o status informacional (*corpus* do Projeto *Filologia Bandeirante*)**

Realização do sujeito	Status informacional	
	Sujeito Dado	Sujeito Novo
Sujeito explícito	8/34 = 23%	57/177 = 32%
Sujeito zero	12/77 = 16%	28/29 = 97%

Vejam, primeiramente, a realização x não-realização do sujeito conforme seu *status* informacional. Uma vez que o índice geral de concordância verbal na 1ª pessoa do plural é extremamente baixo no português popular (no nosso corpus, apenas 33%), fato já constatado por outros estudos (cf. RODRIGUES, 1987), era de se esperar que o falante realizasse o sujeito com maior frequência em todas as instâncias, como um tipo de compensação funcional, estratégia sugerida por Labov. No entanto, a tabela 3 mostra que a realização do sujeito não se dá indiscriminadamente, mas segue um padrão. O falante do português popular tende a optar pelo sujeito explícito nos casos em que este veicula informação nova (177/206=86%), ao passo que nas instâncias de sujeito Dado, a opção preferencial é a sua omissão (77/111=69%).

Passemos à tabela 4, que mede a realização da concordância verbal conforme o *status* informacional e a realização x não-realização do sujeito. Do ponto de vista funcional, a concordância verbal não se faz necessária nos casos de sujeito explícito, uma vez que este já contém a informação relativa ao referente do verbo. E, de fato, não há uma diferença muito significativa nos índices de concordância com sujeito explícito, conforme este represente informação nova (32%) ou dada (23%).

Voltemos, pois, nossa atenção para as ocorrências de sujeito zero. A tabela 4 indica que o contexto de sujeito zero equivalente à informação dada inibe fortemente a realização da concordância verbal (16%). Os exemplos a seguir ilustram este contexto.

(27-31) tava chuvendu *nóis* tava batendu a enxada nu parava ia nu ranchu ranchu di capim trocava a ropa botava aquela enxugá lá nu fogu i ia trabaiá (I.13,p.192)

Os exemplos (27-31) mostram que o referente tende a ser explicitado na posição de sujeito da primeira oração e apagado nas demais; a concordância verbal igualmente não se realiza. Nestes casos, a ausência concomitante do sujeito e da flexão número-pessoal do verbo não resulta em comprometimento da informação, uma vez que o falante já apresentou ao ouvinte o referente (*nós*) no trecho do discurso imediatamente anterior, tornando tal referente um Tópico Dado (cf. também RODRIGUES, 1989).

O contexto de sujeito zero equivalente à informação nova, por outro lado, apresenta o padrão oposto, favorecendo a concordância verbal, praticamente de maneira categórica:

de um total de 29 ocorrências de sujeito zero Novo, em 28 a concordância se realiza (97%).<sup>2</sup> Este contexto engloba tanto casos em que não há um antecedente de 1ª pessoa do plural expresso no discurso anterior (como no exemplo 32), quanto aqueles em que este antecedente está presente, mas a flexão número-pessoal do verbo se faz necessária para evitar ambiguidade na interpretação do referente, como no exemplo 22 mencionado anteriormente e reproduzido a seguir.

(32) então *eu* fui nu casamentu da bevelina (...) aí fomu nu casamentu ...quando casô a gente foi...*u vosso avô...meu bisavô ...foi u grigório (...)* i foi *u bevelinu* (I.12,p.183)

(22) *são tiago* hoje é *uma cidade* boa pra *nóis* que somo aqui...*desse lugarzinho pobre* né? graças a deus temos bons professores boas escola...bons padre (I.1,p.15)

De acordo com as reflexões de Labov (1994) vistas no início, podemos inferir que se o autor tivesse acesso a estes resultados, argumentaria que casos totalmente não-marcados (de sujeito zero e não-realização da concordância) passaram despercebidos e não foram recolhidos.

Embora isto seja uma possibilidade, os resultados compilados neste estudo sobre a realização dos pronomes reflexivos e a realização do sujeito e da concordância verbal apontam para a outra conclusão possível. Nominalmente, existe realmente uma forte tendência de preservação da informação, quando esta representa informação nova no discurso.

### Considerações finais

Ao contrário da visão de Labov (1994), nossos dados demonstram que há uma tendência de preservar a informação referencial em contextos nos quais esta representa informação nova.

Assim, tanto a concordância verbal de 1ª pessoa do plural quanto o pronome reflexivo tendem a ser mais frequentemente realizados pelo falante, quando sua ausência implica comprometimento da informação.

Os resultados dos cálculos quantitativos evidenciam, além disso, que o falante tende a evitar a redundância. Em outras palavras, nos contextos em que a concordância verbal e o pronome reflexivo equivalem a informação dada, predomina a sua omissão.

Os dados da presente pesquisa apontam, portanto, para a relevância do fator *status* informacional como uma das forças que governam fenômenos de variação e mudança linguística no português popular brasileiro.

---

2 Há apenas uma ocorrência de verbo não-marcado: *quando nós viemo praqui é... é nós três né?* (I.11,p.165). Esta se insere num contexto discursivo em que o documentador pergunta à informante quantas pessoas moravam em sua casa. O sujeito constitui então a informação nova requisitada, sendo explicitado na função de Rema da oração subsequente. Interpretamos a casa vazia do sujeito como informação nova no sentido de não conhecida, na medida em que seu referente só é esclarecido posteriormente.

## REFERÊNCIAS

DIK, S. C. *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Dordrecht Holland/ Province RI:USA: Foris Publications, 1989. 433 p.

\_\_\_\_\_. *The Theory of Functional Grammar. Part 2: Complex and Derived Constructions*. Berlin/ New York: Mouton de Gruyter, 1997. 477 p.

HALLIDAY, M. A. K. *Language as social semiotic – the social interpretation of language and meaning*. Great Britain: Edward Arnold, 1987. 256 p.

\_\_\_\_\_. *An Introduction to functional grammar*. Great Britain: Edward Arnold, 1994. 434 p.

LABOV, W. *Sociolinguistic Patterns*. 11. ed. Philadelphia: University of Pennsylvania Press, 1991. 344 p.

\_\_\_\_\_. *Principles of Linguistic Change: Internal Factors*. Oxford–UK / Cambridge-USA: Blackwell Publishers, 1994. v. 1. 641 p.

\_\_\_\_\_. *Principles of Linguistic Change: External Factors*. Cambridge/ Philadelphia: Blackwell Publishers, 2001. v. 2. 572 p.

LIMA-HERNANDES, M. C. O processamento das funções de base comparativa: Proposta de análise funcional. In: *Seminários de Linguística Funcional*. São Paulo: USP, 2004.

PEREIRA, D. C. *Concordância verbal na língua falada nas trilhas das bandeiras paulistas*. 2004. 116 f. Dissertação (Mestrado em Letras. Área de Concentração: Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

\_\_\_\_\_. *Variação e mudança no uso dos pronomes reflexivos no português popular da capital paulista: Uma abordagem funcionalista e cognitivista*. 2007. 350 f. Tese (Doutorado em Letras. Área de Concentração: Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

PEREIRA, D.C.; RODRIGUES, A. C. S. O peso do fator discursivo Dado X Novo na variação da concordância verbal de 1ª pessoa do plural no Português Brasileiro. In: XV CONGRESSO INTERNACIONAL DA ALFAL, 2008, Montevideo. *Anais do XV Congresso Internacional da ALFAL*. Montevideo: Universidad de la República. Disponível em <[http://alfal.easy-planners.info/programa/buscar.php?id\\_tl=645#](http://alfal.easy-planners.info/programa/buscar.php?id_tl=645#)>. Acesso em 18 jul. 2012.

RODRIGUES, A. C. S. *A concordância verbal no português popular em São Paulo*. 1987. 259f. Tese (Doutorado em Letras. Área de Concentração: Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

\_\_\_\_\_. Concordância verbal e estratégias de pronominalização. *Estudos Linguísticos*, Lorena, v. 18, p. 546-554, 1989.

SCHERRE, M. M. P.; NARO, A. J. Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil. *Delta*, v. 9, n. 1, p. 1-14, 1993.

# Clivagem e construções similares sob a perspectiva da Gramática Discursivo-Funcional<sup>1</sup>

(Cleft Constructions and Similar Structures in the Functional Discourse Grammar)

Erotilde Goreti Pezatti<sup>1</sup>

<sup>1</sup>Instituto de Biociências, Letras e Ciências Exatas - Universidade Estadual Paulista (UNESP-SJRio Preto)

pezatti@ibilce.unesp.br

**Abstract:** This paper deals with so called cleft constructions and other similar structures in eight Portuguese varieties, aiming to show that the formal differences correspond to different functional strategies used by the Speaker to achieve his communicative intentions, in accordance with the theory of Functional Discourse Grammar (HENGEVELD; MACKENZIE, 2008). This study takes as evidence real spoken occurrences extracted from a *corpus* organized by the Linguistic Centre at the University of Lisbon. The results show the following distribution: *cleft construction* is a strategy for the pragmatic function *Contrast marking*; *Focus-ser constructions* are used to highlight a constituent and *(é) que* is used as an operator of emphasis which shows the speaker's wish to intensify certain pieces of information.

**Keywords:** lusophony; cleft constructions; focus; contrast; emphasis.

**Resumo:** Este estudo trata de estruturas denominadas clivagem ou construções *-(é) que*, nas variedades portuguesas, objetivando mostrar que as diferenças formais correspondem a diferentes estratégias utilizadas pelo Falante para atingir seus objetivos comunicativos, conforme preconiza a teoria da Gramática Discursivo-Funcional (HENGEVELD; MACKENZIE, 2008). Como material, utiliza ocorrências reais de uso extraídas do *corpus* oral organizado pelo Centro de Linguística da Universidade de Lisboa. A análise dos dados lusófonos revela que clivagem constitui uma estratégia de marcação da função pragmática *Contraste*; construções de *Foco-Ser*, por outro lado, são usadas para focalizar um constituinte, ou seja, marcam a função pragmática *Foco*; *(é) que*, por sua vez, funciona como um operador de *Ênfase*, já que assinala o desejo do Falante de intensificar uma parte de informação.

**Palavras-chave:** lusofonia; clivagem; foco; contraste; ênfase.

## Apresentação

O estudo da clivagem em português tem sido tratado sob várias perspectivas e de diferentes modos, como se pode observar nos trabalhos de Longhin (1999), Longhin e Ilari (2000), Travaglia (2006), Braga (2009) entre outros. Esses trabalhos, no entanto, mostram uma preocupação maior em se distinguir os vários tipos de estruturas e suas origens, relegando em segundo plano sua função, já que tais estruturas são tratadas, de modo geral, todas como “estratégias de realce ou de focalização de partes de informação”, de “marcação de relevo positivo” ou contraste. Nesse aspecto surge então a questão: se há tantas formas diferentes, há também diferentes funções? Responder a essa questão é o objetivo deste estudo, que toma como aparato teórico a Gramática Discursivo-Funcional (HENGEVELD; MACKENZIE 2008). Esse modelo funcionalista considera que devem ser explicados todos os aspectos do contexto que têm impacto sistemático sobre as escolhas gramaticais disponíveis para o Falante na formulação.

<sup>1</sup> Bolsista de Produtividade em Pesquisa CNPq (Proc. n. 301210/2009-8); Bolsista Capes – Pós-Doutorado no Exterior – Lisboa/Portugal (Proc. n. 5784/10-7).

Como material, foram utilizadas ocorrências reais de uso, extraídas do *corpus* oral organizado pelo Centro de Linguística da Universidade de Lisboa, em parceria com a Universidade de Toulouse-le-Mirail e a Universidade de Provença-Aix-Marselha. Tal amostragem recebe o nome de “Português oral” e desenvolveu-se no âmbito do Projeto “Português Falado: Variedades Geográficas e Sociais”, do qual resultou um *corpus* de amostragens de variedades do português falado. Para este estudo, selecionaram-se as amostragens referentes às variedades que constituem língua oficial do país, ou seja, a brasileira, a portuguesa, as africanas (de São Tomé e Príncipe, Angola, Cabo Verde, Guiné-Bissau e Moçambique) e a timorense.<sup>2</sup> Não se pretende, no entanto, assumir um tratamento variacionista e/ou sociolinguístico; pelo contrário, o uso desse material apenas nos permite despreziosamente observar como esse fenômeno ocorre no português.

Neste estudo nos concentraremos nas estruturas denominadas por Braga (2009) de clivagem, construções-*é que*, construções-*que* e construções *Foco-ser*, conforme exemplificam respectivamente (01), (02), (03) e (04).

- (01) já não *são* as senhoras *que* vão se sentar à frente (PT96:MeioPequeno:9)
- (02) por que *é que* não aceitaram depois? (TL99:IdentidadePovo:41)
- (03) nós nunca *que* iríamos vender a nossa liberdade e a nossa independência (TL99:IdentidadePovo:32)
- (04) mas ficava o pão melhor, *era*, com a batata? (PT94:AmassarCozer:93)

O texto encontra-se dividido em duas partes. Na primeira, a atenção é dedicada à interação, ou seja, às ações e estratégias adotadas pelo Falante para obter o seu propósito comunicativo, levando em conta o conhecimento de mundo que supõe ter seu interlocutor, conforme propõe a GDF. A segunda parte, por sua vez, apresenta três seções e trata especificamente das construções objeto de estudo. A seção 2.1 é reservada para Clivagem; 2.2 trata das construções de Foco-ser e, por fim, 2.3 enfoca as construções-(*é*) *que*. As duas partes são seguidas pelas Considerações Finais.

## Uma questão interacional

O modelo funcionalista aqui adotado, a Gramática Discursivo-Funcional (doravante GDF), assume o pressuposto de que as propriedades de enunciados linguísticos são adaptáveis aos objetivos comunicativos que o usuário de língua, na interação com outros usuários, procura alcançar ao usar tais enunciados. Assim a GDF (HENGEVELD 2004a, 2004b) é entendida como uma arquitetura modular, com uma organização descendente, do discurso para a forma das expressões linguísticas, pois considera que a construção de expressões linguísticas se inicia na codificação da intenção do falante e se desenvolve até a articulação. Essa direção é motivada pela concepção de que um modelo de gramática será mais eficaz quanto mais sua organização se assemelhar ao processamento linguístico no indivíduo, um processo descendente, conforme mostram estudos psicolinguísticos (cf. LEVELT, 1989).

A arquitetura geral da GDF é constituída de quatro componentes: o conceitual, o contextual, o gramatical e o de saída (cf. Figura 1 a seguir).

---

<sup>2</sup> Os materiais foram obtidos no endereço [http://www.clul.ul.pt/sectores/linguistica\\_de\\_corpus/projecto\\_portuguesfalado.php](http://www.clul.ul.pt/sectores/linguistica_de_corpus/projecto_portuguesfalado.php). Acesso em: 11 mar. 2009.

O Componente Conceitual é pré-linguístico. Nele, a intenção comunicativa (por ex.: aviso de perigo) e a representação mental correspondente (por ex.: o evento causador de perigo) são relevantes. Por meio da operação de Formulação essas representações conceituais são traduzidas em representações pragmáticas, no Nível Interpessoal, e semânticas, no Nível Representacional. As regras usadas na Formulação são específicas de cada língua.<sup>3</sup> Como resultado, representações conceituais similares podem receber representações pragmáticas e semânticas diferentes em diferentes línguas.

As regras de formulação fazem uso de um conjunto de primitivos que contém *moldes*, lexemas e operadores. As configurações, nos níveis Interpessoal e Representacional, são traduzidas em estruturas morfossintáticas na Codificação Morfossintática. As regras de Codificação Morfossintática caem em um conjunto de primitivos que contém *Padrões* Morfossintáticos, Morfemas Gramaticais e Operadores Morfossintáticos. Similarmente, as estruturas dos níveis Interpessoal, Representacional e Morfossintático são traduzidas em estrutura fonológica no Nível Fonológico. As regras de codificação fonológica deságuam em um conjunto de primitivos que contém *Padrões* fonológicos, formas supletivas e Operadores Fonológicos.

Ao organizar o Componente Gramatical desse modo, a GDF considera a abordagem funcional de língua em seu extremo lógico: dentro da organização descendente da gramática, a pragmática comanda a semântica, a pragmática e a semântica comandam a morfossintaxe e a pragmática, a semântica e a morfossintaxe comandam a fonologia.

O Nível Fonológico de representação é o *input* para a operação de Articulação, que, no caso de um Componente de Saída Acústico (diferentemente da escrita ou da língua de sinais), contém regras fonéticas necessárias para alcançar um enunciado adequado. A articulação ocorre fora da gramática propriamente dita.

Os vários níveis de representação dentro da gramática alimentam o Componente Contextual, capacitando subsequente referência a vários tipos de entidades relevantes para cada um desses níveis uma vez introduzidos no discurso. O Componente Contextual alimenta as operações de formulação e codificação, ao disponibilizar antecedentes, referentes visíveis e participantes do ato de fala que podem influenciar a composição do Ato Discursivo subsequente. Para criar uma especificação contextual, o Destinatário tem de reconstruir todos os níveis de representação com base no *output* real da gramática, i. e., o enunciado fonético. A GDF parte da perspectiva da produção linguística e concentra-se no Componente Gramatical.

Os quatro níveis de representação são hierarquicamente estruturados em camadas de vários tipos. A mais alta camada do Nível Interpessoal é o Movimento, que consiste de um ou mais Ato Discursivo; um Ato discursivo é, por sua vez, organizado sobre um esquema ilocucionário, que contém dois Participantes do Discurso, Falante e Destinatário, e o Conteúdo Comunicado, composto, por seu turno, de Subatos. Cada Subato é uma forma de ação comunicativa do Falante, que pode ser uma tentativa de evocar uma propriedade, denominado Subato de Atribuição, ou de evocar um referente, ou seja, um conjunto nulo, único ou múltiplo de entidades, denominado, por isso, de Subato de Referência.

---

3 A GDF não pressupõe noções semânticas e pragmáticas universais.

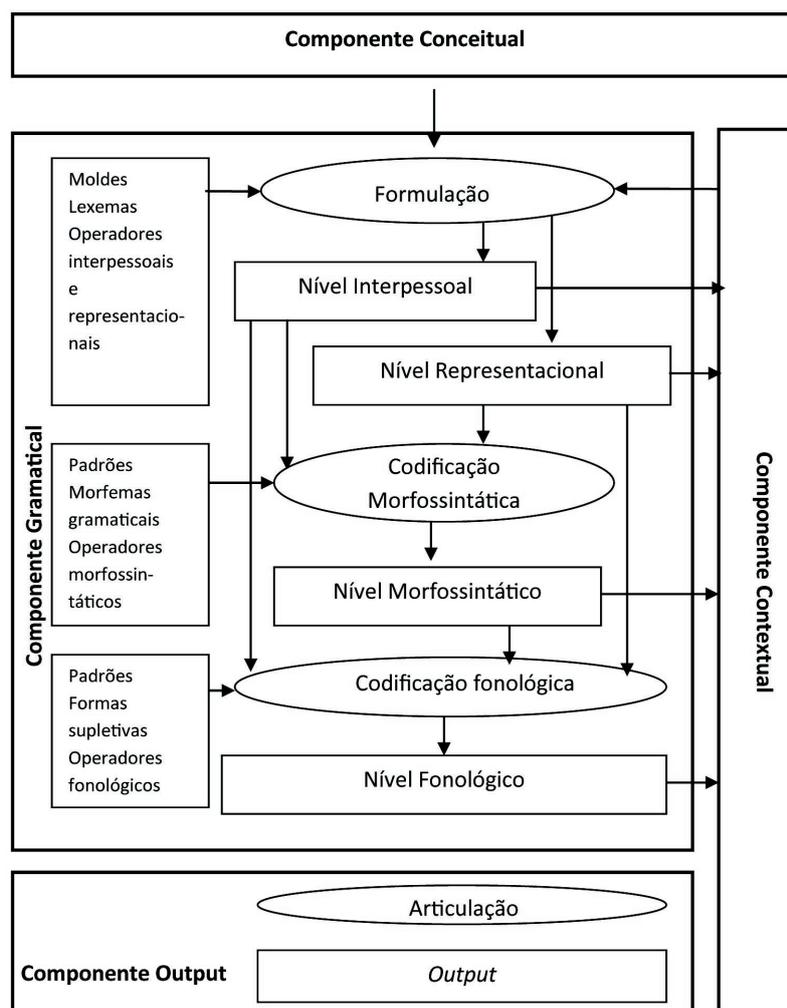


Figura 1. Esboço geral da GDF (adaptado de HENGEVELD; MACKENZIE, 2008, p. 13)

No Nível Representacional a unidade mais alta é o conteúdo proposicional (p), que contém um ou mais episódio, que, por sua vez, contém um ou mais estado-de-coisas, organizado em propriedade, indivíduo, lugar, tempo, modo, quantidade e razão.

O Nível Morfossintático é responsável pelas representações estruturais em termos de propriedades lineares das unidades linguísticas e, similarmente a outros níveis, é hierarquicamente organizado em expressão linguística, oração, sintagma e palavra.

Cada um dos níveis de representação dentro do Componente Gramatical é estruturado de modo próprio. O que todos eles têm em comum é uma organização hierarquicamente ordenada em camada, ou seja, são dispostos numa estrutura em camadas. Cada camada é composta de um Núcleo (obrigatório), que pode ser restringido por um Modificador (opcional), especificado por um Operador e ter uma Função. Núcleos e Modificadores representam estratégias lexicais, enquanto operadores e funções representam estratégias gramaticais. Função é sempre relacional e ocorre entre unidades da mesma camada, já o operador se aplica a uma unidade em si mesma.

A intenção do Falante não surge em um *vacuum*, mas sim em um multifacetado contexto comunicativo. Uma dessas facetas constitui o Componente Contextual, que contém dois tipos de informação:

(1) a informação imediata (de curto prazo) recebida do Componente Gramatical relativa a um enunciado particular que é relevante para a forma que os enunciados subsequentes assumem;

(2) informações de longo prazo sobre a interação corrente que é relevante para as distinções que são requeridas na língua em uso e que influenciam a formulação e a codificação nesta língua.

A GDF admite que fatores relativos a questões de gênero, registro, estilo etc. são aspectos do contexto de interação que podem ter impactos sobre as escolhas linguísticas do Falante. Assim, somente quando o contexto tem um impacto sistemático sobre as escolhas gramaticais disponíveis para o Falante na formulação é que esses aspectos merecem explicação.

Em português, informação vinda do contexto situacional, tal como a diferença de gênero (sexo) e a de relação social, é relevante para a codificação. Em (05), a escolha da forma ‘senhor’ em vez de ‘você’ reflete a formalidade da relação entre os interlocutores; já a escolha de *cansado* (e não *cansada*) e de *senhor* (e não de *senhora*) sinaliza o gênero (sexo) do Destinatário. Tanto a relação social quanto o gênero são especificações do Componente Contextual que são refletidos na gramática do português.

(05) O *senhor* parece cansado!

Tratada de uma perspectiva funcionalista, a clivagem e suas construções correlatas são sempre vistas como estratégias do Falante para realçar ou focalizar partes de informação avaliadas por ele como mais importantes no momento da interação. Isso significa que a clivagem constitui um aspecto formal da unidade linguística que reflete o seu (da unidade linguística) papel na interação entre Falante e Destinatário. Trata-se, portanto, das ações e estratégias adotadas pelo Falante para obter o seu propósito comunicativo, levando em conta o conhecimento de mundo que supõe ter seu interlocutor. São, portanto, aspectos contextuais que trazem consequências para a forma da expressão linguística.

São diversas as estratégias utilizadas pelo Falante com o intuito de atingir sua meta comunicativa. As estratégias de natureza intencional, na GDF, referem-se à *pragmática*. A pragmática constitui o modo como o Falante modela as suas mensagens em relação às expectativas que tem do estado mental do Destinatário. Isso determina as partes de uma unidade linguística que serão apresentadas como particularmente salientes, as que serão escolhidas como ponto de partida do Falante e as que serão consideradas compartilhadas pelo Falante e pelo Destinatário. A influência desses aspectos sobre a estrutura das unidades linguísticas recebe o nome de *função pragmática*, que se aplica ao conteúdo comunicado ou a um de seus constituintes. A GDF distingue três funções pragmáticas, Tópico, Foco e Contraste, que são linguisticamente marcadas, e têm como contraparte, que raramente é marcada nas línguas, respectivamente o fundo, o comentário e a sobreposição. Interessam-nos neste estudo apenas as funções pragmáticas Contraste e Foco.

Contraste, para Hengeveld e Mackenzie (2008), constitui uma função pragmática, ao lado de Foco e Tópico,<sup>4</sup> e assinala o desejo do Falante de realçar diferenças particulares entre dois ou mais conteúdos comunicados ou entre um conteúdo comunicado e informações

---

4 Hengeveld e Mackenzie (2008) consideram Tópico como a função pragmática atribuída a um constituinte para assinalar como o conteúdo comunicado se relaciona ao registro construído gradualmente no componente contextual, fornecendo um tipo específico de orientação para o estoque de informação nova a ser apresentada.

contextualmente disponíveis, no cotexto ou na situação discursiva. Esse aspecto será mais bem explicado na seção 2.1.

Foco, por outro lado, sinaliza a seleção estratégica do Falante de informação nova para preencher uma lacuna na informação do Destinatário, ou para corrigir uma informação do Destinatário. Essa função será mais detalhadamente explicitada na seção 2.2.

Outra categoria pragmática considerada pela GDF e relevante para este estudo é a Ênfase. A Ênfase consiste numa estratégia utilizada pelo Falante para intensificar, por meios lexicais ou gramaticais, um constituinte ou toda a expressão linguística, visando a atingir seus objetivos comunicativos, conforme se verá na seção 2.3.

## Clivagem e construções similares na GDF

### Clivagem

Ocorrências como (06) constitui um claro exemplo de construção clivada. É possível observar que a estrutura clivada – *são as senhoras que* – estabelece uma clara oposição entre *as senhoras* e *os jovens*, que aparece na oração seguinte, na continuação do texto: *já não são as senhoras que se vão sentar à frente da cadei[...], com a cadeira à frente do, do palanque, já é os jovens e, e pronto.*

(06) *já não são as senhoras que se vão sentar à frente da cadei[...], com a cadeira à frente do, do palanque, já é os jovens e, e pronto* (PT96:MeioPequeno:9)

Isso indica que as construções tradicionalmente denominadas clivadas, consideradas formas de marcação de Foco (DIK, 1997; LONGHIN, 1999; HENGEVELD; MACKENZIE, 2008, p. 91), constituem, em português, uma estratégia de marcação de Contraste.

Bolinger (1961, p. 83) define contraste como o fenômeno pelo qual dois ou mais itens são contrabalançados, indicando-se a preferência por um deles ou por vários membros do grupo. Ressalta, porém, que só há *contrastive accent* quando o número de candidatos a ocupar a posição potencialmente contrastiva é limitado.

O fenômeno da contrastividade, para Chafe (1976), envolve três fatores: o conhecimento pressuposto (*background knowledge*), entendido como dado; o número, geralmente limitado, de candidatos possíveis para exercer o papel em questão; e a asserção que contrasta o candidato correto com outros possíveis, ou seja, “eu *acredito que você acredita que algo aconteceu, que você tem um conjunto limitado de candidatos na mente e eu digo a você que esse alguém é alguém específico, em vez de outro*”. Em inglês, segundo o autor, pode-se exprimir a contrastividade por meio da entonação mais alta e acento mais forte no foco de contraste, mediante o uso de sentenças clivadas, de sentenças pseudo-clivadas e ainda por meio do tópico.

Taglicht (1984), por sua vez, considera a noção de contraste a partir das perspectivas entonacional e contextual. Segundo ele, um elemento é contrastivo quando apresentado como um de um par de opostos. Para ele, a oposição é um conceito pragmático-contextual, que pode ser representada por termos semanticamente opostos (ex.: frio vs. quente, agradável vs. chato, etc.) ou por estruturas sintaticamente paralelas (ex: João é agradável

e Paulo, um chato). O contraste também pode ser caracterizado como explícito (ambos os membros do par de opostos estão presentes no enunciado) ou implícito (apenas um elemento do par está presente). De acordo com o autor, é a forma do enunciado que transmite a implicação de que alguma coisa não dita teria contido o outro membro.

Dik (1989), por outro lado, considera contraste um tipo de foco, a que denomina Foco Contrastivo. São os casos em que a informação focalizada nem sempre é totalmente nova, mas é colocada em foco em virtude de algum contraste implícito ou explícito com alguma parte da informação. Assim, esse tipo de foco envolve sempre algum contraste entre o constituinte Foco e as partes alternativas de informação que podem ser apresentadas de forma explícita ou pressuposta. No primeiro caso, tem-se Foco Paralelo e os elementos focalizados desempenham a mesma função sintática. Nos outros tipos de Foco Contrastivo, a informação apresentada contrasta com outra que o Falante pressupõe estar armazenada na memória do destinatário. Implica sempre uma pressuposição, sendo por isso denominado de contrapressuposicional, que, por sua vez pode ser Substitutivo (o Falante pressupõe que o destinatário possui uma parte incorreta de informação X, que deve ser substituída por uma correta informação Y); Expansivo (o Falante pressupõe que o destinatário possui uma informação X, mas que X não está completa, havendo uma informação Y que é também importante que o destinatário conheça); Restritivo (o Falante presume que o destinatário dispõe de uma parte de informação correta e outra incorreta. Nesse caso, o Falante corrige a informação pragmática do destinatário ao restringir um conjunto de itens pressupostos àqueles que considera ter os valores para a posição envolvida); Seletivo (o Falante pressupõe que o destinatário acredita que a informação X ou Y está correta, mas não sabe qual).

A função pragmática Contraste, para Hengeveld e Mackenzie (2008), constitui uma função pragmática, tal como Foco e Tópico, que assinala o desejo do Falante de realçar diferenças particulares entre dois ou mais conteúdos comunicados ou entre um conteúdo comunicado e informações contextualmente disponíveis, no contexto ou na situação discursiva. É sob esse conceito que são aqui analisados os dados lusófonos.

O Contraste, em português, é geralmente marcado por operadores (elementos gramaticais), que normalmente antecedem o termo contrastado. Assim, operadores como *apenas* e *só* restringem um conjunto de itens pressupostos àquele que o Falante considera adequado, uma vez que presume que o Destinatário possui uma parte de informação que não é correta. Nesse caso, ele corrige a informação pragmática do Destinatário ao restringir um conjunto de itens pressupostos àqueles que considera ter os valores para a posição envolvida. Em (07), com o operador *só* o Falante opõe *sal* a outros tipos de temperos, para justificar que a carne muito temperada fica enjoativa.

(07) então você assa a carne só *no sal* (Bra80:ComerFalarBem:12)

Já operadores como *também* adicionam uma informação a outra pressuposta ou já mencionada, pois o Falante acredita que o Destinatário possui uma informação que não está completa. Assim, em (08), a intenção do Falante é acrescentar o surgimento da esterilidade às outras consequências do aborto clandestino, já mencionadas.

(08) *também* aparece a esterilidade (GB95:Aborto:56)

Por outro lado, operadores como *principalmente* e *sobretudo* indicam a crença do Falante de que o Destinatário acredita que mais de uma informação pode ser correta, por isso seleciona a informação mais adequada, marcando-a com um desses operadores, como se pode constatar em (09).

(09) eisso vai setraduzir numaprimeiraexplosãoescolarsobretudo no secundário (Ang97:EnsinoAngola:49)

Além do uso dos operadores acima mencionados, o Falante pode recorrer a outra estratégia para contrastar duas informações. Nesse caso, no entanto, o Falante indica ao Destinatário que substitua uma informação incorreta pela que ele (Falante) considera a correta. É exatamente isso que expressam as construções tradicionalmente denominadas clivadas, como em (10). Facilmente se percebe, nessa ocorrência, uma oposição entre *a classe dominante* expressa no texto e a classe menos favorecida, não claramente expressa mas pressuposta no componente contextual pelo cotexto anterior.

(10) e à cultura, o cara não tem direito a nada, meu, só vive que nem um animal.  
- - e nem de acesso ao material que ele produz  
- e é o cara, e é o cara que produz esse prédio aqui, em que a gente está morando e tudo o mais. *e então é a classe dominante que usufrui de, da, cultura, pô* (Bra87:Economia Sociedade:44)

Assim, tanto em (06) quanto em (10), a construção *é...que* constitui uma forma de codificar, considerando as informações compartilhadas pelos interlocutores, a intenção do Falante de solicitar ao Destinatário que substitua uma informação incorreta pela que considera correta para assim atingir seu propósito comunicativo.

Essa intenção, que ocorre no Nível Interpessoal, é formalmente, expressa no Nível Morfossintático, por meio da construção clivada, constituída do verbo *ser* mais *que*. A cópula tende a refletir o número e pessoa do nome contrastado, e o modo e tempo expresso no verbo da oração principal, como mostram *são*, em (06) que copia o tempo e modo verbal de *ir*, e *é*, em (10), que faz o mesmo com relação ao verbo *usufruir*.

O operador de Contraste *é...que* caracteriza-se pela descontinuidade, ou seja, posiciona-se parte (verbo *ser* flexionado) antes e parte (partícula *que*) após o núcleo que especifica. Isso diz respeito à ordenação de constituintes da oração que, na GDF, também pertence ao Nível Morfossintático, uma vez que a linearização de constituintes serve como um dos mecanismos pelos quais relações e funções vindas de níveis mais altos podem ser formalmente expressas. Assim, o operador de Contraste permanece no domínio do sintagma, uma vez que especifica um Subato, girando em torno de seu núcleo, permitindo assim que o sintagma, por sua vez, permaneça na sua posição canônica dentro da oração.

Além disso, como informação compartilhada, os sintagmas nominais, que representam os Subatos de Referência, são marcados pelo operador de Identificabilidade, também pertencente ao Nível Interpessoal. Identificabilidade, na GDF, refere-se à avaliação do Falante sobre a possibilidade de o referente de um Subato ser identificável, tanto para o Destinatário quanto para si próprio. Assim um referente pode ser identificável (+id) ou não (-id) para o Destinatário, e específico (+s) ou não (-s) para o próprio Falante.

Em construções clivadas o Subato Referencial é marcado pelo operador (+id), uma vez que o Falante pressupõe que o referente é identificável para o Destinatário, e pelo operador (+s), já que é específico para o próprio Falante. Essa identificabilidade vinda do

Nível Interpessoal é marcada em português pelo artigo definido *o*, obviamente flexionado em gênero e número, de acordo com o núcleo a que especifica, como fica claro nos exemplos (06) e (10) respectivamente com *as senhoras* e *a classe dominante*.

Como se vê, construções clivadas apresentam uma função bem específica, que é a de estabelecer uma oposição entre partes de informação, ou seja, constitui uma construção gramaticalizada para marcar a função pragmática Contraste.

### Construções de Foco-Ser

Segundo Dik (1989), a função Foco de uma expressão linguística é a informação relativamente mais importante ou saliente num dado contexto de interação verbal, e avaliada pelo Falante como essencial para ser integrada na informação pragmática do destinatário. A informação focal se refere então às mudanças que o Falante deseja provocar na informação pragmática do Destinatário. Dessa forma, deve haver sempre alguma diferença entre a informação pragmática do Falante e o quadro que ele faz da informação pragmática do Destinatário. A informação focal é apresentada como “nova” para o Destinatário.

A GDF, baseando-se em Dik (1989), considera que, na interação, a função Foco sinaliza a seleção estratégica do Falante de informação nova para preencher uma lacuna na informação do Destinatário, ou para corrigir uma informação do Destinatário. Em ilocuções Declarativas, a função Foco, como estratégia de informação nova, ocorre em construções existenciais. Nessas construções, a unidade semântica, que não é argumento de nenhum predicado, contém a informação nova, sendo, portanto o Foco, o que lhe vale a posição final da oração, conforme se verifica em (11), uma construção existencial com *ser*.

- (11) são *milhões de pessoas* que se deslocam e ficam concentradas durante meses no mesma região em que destroem completamente todo o coberto vegetal (Ang97:Guerra e Ambiente:32)

Essa função de sinalizar a seleção estratégica do Falante de informação nova para preencher uma lacuna na informação do Destinatário, como mostram (12) e (13), pode ser codificada por meio do que Braga (2009) denomina construção Foco-ser. Em (12), o sintagma *destas raízes* constitui a informação nova solicitada pelo documentador já no início da entrevista sobre a origem da palavra *morna*. Observe que, sem o operador de Foco *foi*, a oração *saiu destas raízes* seria informacionalmente neutra, e *destas raízes* não teria destaque.

- (12) - portanto, estava-me a dizer, eh, estávamos a falar sobre a origem da palavra “morna”.  
-> sim, eh, dizem que a palavra “morna” nasceu de, do vocábulo “mourni[...]”, “mourning”, eh, quer dizer que as pessoas que, que cantavam não cantavam, quer dizer, eh, sabe que às vezes quando no[...], eh, [...], como se diz, na, quando, eh, aconteceu que morreu alguma pessoa  
- sim  
-> etc., então o inglês d[...], o inglês diz “they are mourning”  
- hum, hum.  
-> estão no sentimento do morto, etc.  
- hum, hum.  
-> e então dizem que o senti[...], o senti[...], o sentimento, o canto da morna  
- hum, hum  
-> *saiu foi destas raízes* (CV95:AsMornas:13)

A ocorrência (13), por seu turno, apresenta no discurso um constituinte pela primeira vez, ou seja, fornece uma nova informação (*com aquarela*), considerada pelo Falante muito importante para o Destinatário acrescentar à sua informação pragmática.

- (13) a única coisa que tinha a fazer era só investigar assim sozinho através de, eh pá, material que eu tinha, não é, lápis de carvão e tal. mais tarde dediquei-me, comecei a dedicar-me na pintura *comecei a pintar; principalmente foi com aquarela* (Moç83:CantarPintar36)

Essa intenção de salientar um constituinte que carrega a informação nova, uma escolha efetuada no Nível Interpessoal, é codificada morfossintaticamente pelo operador *foi*, ou seja, o verbo *ser* flexionado no mesmo modo (indicativo) e tempo (pretérito perfeito) do verbo principal, *saiu* e *comecei*, respectivamente em (12) e (13).

Construção de Foco-ser pode ainda indicar a correção de uma informação que o Falante considera obscura, inadequada ou incompleta. É o que se percebe em (14), em que o Falante solicita do Destinatário uma correção em sua informação pragmática, ao responder a pergunta contida no Ato Interrogativo *ficava o pão melhor; era, com a batata?* O Foco, nesse caso, é marcado pelo verbo *ser* flexionado na terceira pessoa do singular, no mesmo modo (indicativo) e tempo (pretérito imperfeito) do verbo principal.

- (14) -> não havia. há batata doce mas está quase a quatrocentos escudos.  
- não, mas antigamente, não se costumava misturar?  
-> antigamente quem tinha de casa... misturava. amassava, cozia-se a batata, pela[...], descascava-se, e ao depois ama[...], amassava-se separado e ao depois é que se deitava no pão. limpava-se bem limpinho para não levar...  
- mas ficava o pão melhor, **era**, com a batata?  
-> ficava! é du[...], é, dura mais dias, eh, o pão mais fofo  
- hum, hum. (PT94:AmassarCozer:93)

Pelos exemplos, é possível observar que os referentes dos Subatos envolvidos não precisam necessariamente ser identificáveis para o Destinatário, uma vez que em (12) *raízes* é entendida como identificável, tanto para o Destinatário quanto para o Falante (+id, +s R), uma vez que pode ser inferida do cotexto, o que é codificado pelo uso do demonstrativo *esta*. O mesmo se observa em (14), com o sintagma *a batata*, claramente referida no contexto precedente. Já em (13), *aquarela* é referida como não identificável (-id) para o Destinatário e específico (+s) para o Falante, não sendo, portanto, marcada morfossintaticamente.

Construção de Foco-ser constitui, na verdade, uma estratégia para destacar elementos que, na ordenação canônica, já se posicionam no final da oração, por isso a necessidade do uso do operador de função pragmática Foco. Esse operador, por seu turno, sempre se colocará à esquerda do sintagma que especifica, indicando um movimento para frente. Em outros termos, construção de Foco-ser, diferentemente de construções clivadas, é usada para assinalar a função pragmática Foco.

### **Construções-(é) que**

Diferentemente de Braga (2009), não distinguimos construções-é *que* de construções-*que*, pois, da perspectiva da GDF, essas duas estruturas constituem uma mesma estratégia

do Falante. Construções-*é que* e sua variante *que* assinalam o desejo do Falante de intensificar partes de informação. Trata-se, portanto, de uma estratégia de Ênfase, marcada pelo operador (*é*) *que*.

Ênfase não tem sido tratada na literatura linguística como uma categoria à parte. Ela é sempre referida como uma forma de salientar constituintes, e muitas vezes confundida com Foco. Na verdade, Ênfase e Foco de modo geral são tratados indistintamente.

Na GDF, no entanto, é entendida como uma categoria pragmática que perpassa todas as camadas do Nível Interpessoal (Ato, Ilocução, Conteúdo Comunicado e Subatos). A Ênfase, na GDF, constitui mais uma estratégia utilizada pelo Falante para atingir seu propósito comunicativo. É, então, definida como a intensificação, por meios lexicais ou gramaticais, de um constituinte ou de toda a expressão linguística. Deve, no entanto, ser distinguida das funções pragmáticas Tópico e Foco<sup>5</sup>, conforme acima definidas, uma vez que é possível enfatizar um constituinte Tópico, como mostram (15) e (16), cuja intensificação é efetuada pelo operador *todos* e pelo modificador *mesmo* respectivamente; e Foco, como em (17), em que o constituinte focal da construção existencial – *dois aspectos que têm que se, que fazer* – é intensificado pelo operador *já*.

(15) e os filhos<sub>Top</sub> ficaram **todos**<sub>Enf</sub> ricos (PT97:DesportoDinheiro:84)

(16) eu<sub>Top</sub> **mesma**<sub>Enf</sub> levei um susto agora (Bra80:Fazenda:105)

(17) portanto há aqui **já**<sub>Enf</sub> *dois aspectos que têm que se, que fazer*<sub>Foco</sub> (Ang97:Guerra e Ambiente:106)

Como observado, essa categoria interacional pode ser aplicada a diferentes camadas do Nível Interpessoal. Um Ato Discursivo, por exemplo, pode ser intensificado por meio de um modificador como *caramba*, para expressar irritação ou raiva, conforme demonstra (18).

(18) sabe, eu brinco de lutar com meu, com meu filho, *caramba!* (Bra80:CriarFilhos:18)

Operador de Ênfase pode referir-se a todo o conteúdo comunicado, assumindo a posição inicial (P<sup>I</sup>) e especificando tudo o que vem à sua direita. Em (19), *aí* intensifica o conteúdo comunicado que ele introduz em relação aos conteúdos comunicados anteriormente expressos.

(19) é muito mais bonita porque é muito mais antiga. ela deve ser, está-se presumindo que essa casa já foi, nós encontrámos, quer dizer, nós não, meus tios encontraram com, quando compraram a fazenda já tinha essa casa. *aí mandaram construir essa que nós moramos* (Bra80:Fazenda:91)

A Ênfase pode recair sobre parte do conteúdo comunicado, como exemplificam (20) e (21), em que os operadores *ainda* e *já* denotam, em (20), a intenção do Falante de ressaltar o momento antecipado da indolência das ovelhas; e em (21), de salientar o momento de ocorrência do estado-de-coisas.

(20) *ainda* com sol [as ovelhas] já estão encostadas (PT97:SerPastor:88)

(21) *já* em sessenta e oito houve um surto de escolas, que foram abertas (CV95:RaparigasCV:07)

Construções-*é que* e sua variante *que* constituem, dessa perspectiva teórica, também uma estratégia para enfatizar uma informação. Como bem assinalam Longhin e Ilari 5 A função pragmática Contraste não permite Ênfase.

(2000), o operador de ênfase (*é*) *que* é comumente usado em ilocuções Interrogativas para enfatizar o constituinte-Q, de qualquer categoria semântica. Não há restrição também quanto ao tipo de Interrogativa, que pode ser direta, como em (22), (23) e (24), em que o operador Q representa respectivamente as categorias semânticas de modo, indivíduo e causa; ou indireta, como em (25) e (26), cuja interrogação incide, no primeiro caso, sobre a locação e, no segundo, sobre uma das unidades de uma construção de identificação.

(22) como *que* é o relacionamento com a sua irmã mais nova? (Bra95:MuitoIguaisDiferentes:83)

(23) e o que *é que* aconteceu? Bra87:EconomiaSociedade:101

(24) por que *é que* então não pode entrar nesse estudo? (Moç97:Maternidade53)

(25) o meu irmão mais velho, não me recordo onde *é que* estaria, (Moç86:Chuva:31)

(26) não estou a ver exactamente qual *é que* é a história(PT97:BoaPontaria02)

Esse operador de Ênfase não se restringe, no entanto, a ilocuções interrogativas, podendo marcar também constituintes de ilocução declarativa, como atestam as ocorrências (27), (28) e (29).

(27) nós nunca *que* iríamos vender a nossa liberdade e a nossa independência (TL99:IdentidadePovo32)

(28) - ah! um bom churrasco, você pega a carne, lava a carne direitinho, joga um sal grosso na carne e leva ela ao fogo, simplesmente. é o bom churrasco. é o churrasco gaúcho, é. porque a carne com um certo tempero, ela fica enjoativa, então você assa a carne só no sal e prepara o molho à parte: cebola, tomate picadinho, alho, vinagre, azeite, certo, sal, mais um pouquinho, uma pimentazinha de leve; então depois a pessoa que gosta bota no prato

- hum, hum.

- à parte, uma farofa! *isso é que* é o churrasco... ao molho... (Bra80:ComerFalarBem:16)

(29) - há um, há um ditado aqui em São Tomé que diz que... Deus fez o mar para todos os peixes, *mas o tubarão que tem a fama é que* se apossou do mar. tomou o mar só para ele.

-> só para ele.

- isso é, tem o sentimento que isto é verdade? é verdade que o tubarão é que reina no mar, aqui em São Tomé?

-> em São Tomé, o tubarão reina aí no mar.(To-Pr96:Pesca:86)

Em (27), não há dúvida de que o elemento enfatizado é o operador de tempo negativo *nunca*. (28) e (29), porém, podem sugerir tratar-se de contraste. Informações contextuais, no entanto, indicam, em (28), que a intenção do Falante naquele momento é ressaltar o que considera um verdadeiro churrasco, sem estabelecer contraste com outras partes de informação; e em (29), que o Falante deseja apenas assinalar que o tubarão é o dono do mar.

Braga (2009) observa que as Construções-*é que*, na fala, tendem a dispensar a correlação modo-temporal e a concordância número-pessoal. Apesar de não ter sido atestado em nosso *corpus*, é perfeitamente possível construções como (28a) e (29a).

(28a) *isso é que* era o churrasco

(29a) *mas* o tubarão que tem a fama *foi que* se apossou do mar

Segundo Braga (2009), a restrição à variação sugere que a expressão *é que* está constituindo um todo amalgamado, imune à correlação modo-temporal, à concordância número-pessoal e à interferência de material linguístico entre os dois itens que a integram. Isso indica que a expressão *é que* está se gramaticalizando como uma locução sinalizadora de foco e que os segmentos vinculados por ela não constituem uma estrutura bioracional. Para a autora, construções-*que*, para as quais as questões relacionadas à concordância número-pessoal e correlação modo-temporal não se colocam, visto que esta estratégia de focalização dispensa o verbo copular, o item *que* abandona sua propriedade conectora e passa a funcionar como um marcador gramatical de foco.

Concordamos com as afirmações da autora com relação à gramaticalização de (*é*) *que*, mas divergimos no tipo de categoria interpessoal que veicula, pois, como demonstrado, tanto *é que* quanto sua forma mais gramaticalizada *que* constituem, nos termos da GDF, um operador de Ênfase e não uma locução sinalizadora de função pragmática.

Os exemplos mostram que o operador de Ênfase (*é*)-*que* aplica-se a Subatos do Conteúdo Comunicado. Esses Subatos, no entanto, podem ser de Atribuição ou de Referência. Em (27), esse operador aplica-se ao Subato de Atribuição temporal, que se expressa morfossintaticamente por meio de *nunca*. Já em (28) e (29), especifica Subatos Referenciais, respectivamente *isso* e *o tubarão*. Nos dois casos trata-se de Subatos identificáveis (+id +s) para o Destinatário e para o Falante. A identificabilidade, em (28), é expressa pelo demonstrativo *isso*, que retoma todo contexto anterior; em (2), por outro lado, *o tubarão*, por ser ancorado pela oração adjetiva e constituir uma informação dada no discurso, é marcado morfossintaticamente pelo artigo definido *o*.

Outra marca formal desse operador de Ênfase é a posição que ocupa na linearização dos constituintes. Como se pode observar pelos exemplos, o operador (*é*)-*que* posiciona-se, dentro do sintagma a que pertence, sempre à direita do seu núcleo, indicando um movimento para trás, independentemente da categoria semanticamente expressa: modo, tempo, razão ou indivíduo.

Desse modo, construção-(*é*)*que*, diferentemente de construção clivada e construção de Foco-ser, que assinalam respectivamente a função pragmática Contraste e a função pragmática Foco, marca uma categoria do Nível Interpessoal, denominada Ênfase.

## Considerações finais

Assumindo, de acordo com o aparato teórico aqui adotado, que os aspectos do componente contextual que têm impacto sistemático sobre as escolhas gramaticais disponíveis para o Falante na formulação precisam ser explicados, a proposta deste artigo é mostrar que as diferentes formas de clivagem correspondem a estratégias discursivas distintas utilizadas pelo Falante para atingir seus objetivos comunicativos.

A GDF distingue funções pragmáticas (Foco, Tópico e Contraste) de outras categorias do Nível Interpessoal, como Ênfase, Mitigação, Identificabilidade, Exatidão e Aproximação. Cada uma dessas categorias pragmáticas aplica-se a camadas específicas do Nível Interpessoal, com exceção da Ênfase que é atribuída a todos os tipos de unidades acionais: Ato discursivo, Ilocução, Conteúdo Comunicado e Subatos de Atribuição e de Referência.

Como demonstramos, diferentes formas correspondem a diferentes estratégias discursivas de que se serve o Falante para conseguir seu propósito comunicativo. Assim, clivagem e construção de Foco-ser constituem formas de marcação de função pragmática, pois leva em conta o modo como o Falante modela as suas mensagens em relação às expectativas que tem do estado mental do Destinatário. Dessa forma, construção Foco-ser determina as partes de uma unidade linguística que são apresentadas como particularmente importante para o Destinatário acrescentar à sua informação pragmática ou corrigi-la; assinala, portanto, a função pragmática Foco. Por outro lado, ao se servir de uma construção clivada, o Falante, considerando as informações que pressupõe estar armazenada na memória do Destinatário, salienta uma delas com base em diferenças ou semelhanças entre elas; marca, dessa forma, a função pragmática Contraste. Construções-(é) *que*, por sua vez, permitem ao Falante intensificar um Subato dentro do Conteúdo Comunicado; desse modo, não introduz uma função pragmática, mas sim uma categoria interacional denominada Ênfase.

Como é possível notar, essas construções (clivagem, construções-é *que*, construções-*que* e construções Foco-ser) revelam ações e estratégias adotadas pelo Falante para obter propósitos comunicativos diferenciados (contrastar, salientar ou intensificar partes de informação), contrariando o que Braga propõe ao afirmar que “Tais fatos mostram que, no que diz respeito às construções clivadas, no português falado no Brasil, a hipótese de um isomorfismo entre forma e função deve ser rejeitada em favor de uma abordagem sensível ao caráter não discreto das categorias linguísticas” (BRAGA, 2009, p. 192).

## REFERÊNCIAS

- BOLINGER, Dwight. Contrastive accent and contrastive stress. *Language*, v. 37, n. 83-96, 1961.
- BRAGA, Maria Luiza. Construções clivadas no português do Brasil sob uma abordagem funcionalista. *Matraga*, 16, p.173-196, 2009.
- CHAFE, Wallace. Givenness, contrastiveness, definiteness, subjects and topics. In: LI, C. N. (Org.) *Subject and topic*. New York: Academic Press, 1976. p. 26-55.
- DIK, Simon. *The theory of functional grammar*. Part. 2. Complex and derived constructions, Berlin/New York: Mouton de Gruyter, 1997.
- \_\_\_\_\_. *The theory of functional grammar*. Part I: The structure of the clause. Dordrecht/ Providence RI: Foris Publications, 1989.
- HENGEVELD, Kees. The architecture of a functional discourse grammar. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. de los A. (Ed.) *A new architecture for functional grammar*. Berlin/New York, Mouton de Gruyter: 2004a. p. 01-21.
- \_\_\_\_\_. Epilogue, en J. L. Mackenzie e M. de los A. Gómez-González (Ed.), *A new architecture for functional grammar*. Berlin/New York: Mouton de Gruyter, 2004b. p. 365-378.

HENGEVELD, Kees; MACKENZIE, J. Lachlan. *Functional discourse grammar*, Oxford: Oxford University Press, 2008.

LEVELT, William. *Speaking*. Cambridge: Cambridge University Press, 1989.

LONGHIN, Sanderléia Roberta. *As construções clivadas: uma abordagem diacrônica*. 1999. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas. Inédita.

LONGHIN, Sanderléia Roberta; ILARI, Rodolfo. Uma leitura hallidayiana das sentenças clivadas do português. *Alfa*, v. 44, p. 193-213, 2000.

TAGLICHT, John. *Message and emphasis: on focus and scope in English*. London: Longman, 1984.

TRAVAGLIA, Luiz Carlos. O relevo no processamento da informação. In: JUBRAN, C. C. S.; KOCH, I. G. V. (Org.). *Gramática do português culto falado no Brasil: construção do texto falado*. v. 1. Campinas: Editora da Unicamp, 2006. p. 167-215.

# “É o seguinte”: limites entre elaboração e projeção

(“É o seguinte”: boundaries between Elaboration and Projection)

Fernanda Abreu e Silva Alencar<sup>1</sup>

<sup>1</sup>Programa de Pós-graduação em Linguística e Filologia, Faculdade de Letras – Universidade Federal do Rio de Janeiro (UFRJ)

alencar.nanda@gmail.com

**Abstract:** In this paper, we analyse the semantic relations of *elaboration* and *projection* in spoken contemporary Brazilian Portuguese. We limit the focus of our analysis to elaboration of the NP “o seguinte”, with the goal of identifying the various types of constructions that exist with this NP and the types of semantic relations they instantiate. Within this complex clauses, structures like “é o seguinte”, “X é o seguinte” and “outros verbos o seguinte” can be found. The structure “X é o seguinte” and “é o seguinte” point to an interpretation along the lines of elaboration. The occurrences of “outros verbos o seguinte”, may indicate more or less clearly the boundaries between elaboration and projection, which depends on the relationship between the transitivity of the verb in the first segment and the nature of the second segment.

**Keywords:** elaboration; projection; NP “o seguinte”.

**Resumo:** Neste artigo, analisamos as relações semânticas de elaboração e de projeção na modalidade falada do português brasileiro contemporâneo. Delimitamos nosso objeto de análise às construções que expandem o SN “o seguinte”, procurando identificar os tipos de construções com esse SN e o tipo de relação semântica que elas instanciam. Nos complexos oracionais em questão, podem ser encontradas estruturas do tipo “é o seguinte”, “X é o seguinte” e “outros verbos o seguinte”. As construções que apresentam as estruturas “X é o seguinte” e “é o seguinte” apontam para uma interpretação como elaboração. As ocorrências com “outros verbos o seguinte” no primeiro segmento, a depender da transitividade do verbo e da natureza do segundo segmento, podem indicar mais ou menos claramente os limites entre elaboração e projeção.

**Palavras-chave:** elaboração; projeção; SN “o seguinte”.

## Introdução

Neste artigo, analisamos as relações de elaboração e de projeção na modalidade de fala do português brasileiro contemporâneo. Tomamos como base a proposta de Halliday (2004), para quem a *elaboração* se realiza nos contextos em que uma parte do texto elabora ou especifica o conceito expresso pela parte que a antecede, e a *projeção*, nos contextos em que uma oração é constituída como a representação do “conteúdo” linguístico de outra, como mostram, respectivamente, os exemplos (1) e (2) abaixo:

- (1) *Cada argumento era fatal para o outro: ambos não poderiam ser verdadeiros.*<sup>1</sup> (HALLIDAY, 1994, p. 226)
- (2) *Eu disse “Estou ficando velho; vou ter que sentar-me à sombra”, e ela disse, “Graças a Deus, pelo menos nós podemos fazer isso”.*<sup>2</sup> (HALLIDAY, 2004, p. 445)

1 No original: *Each argument was fatal to the other: both could not be true.* (HALLIDAY, 1994, p. 226)

2 No original: *I said “I’m getting old; I’m going to have to sit down in the shade,” and she said, “Thank God, at least we can do it”.* (HALLIDAY, 2004, p. 445)

Embora o autor considere a Elaboração e a Projeção como relações distintas, em algumas ocorrências, os limites entre elas podem não estar bem definidos. Essas ocorrências compreendem o SN “o seguinte”,<sup>3</sup> como ilustrado abaixo:

(3) *Ele me disse o seguinte: que ele quer ir lá na piscina.* (Amostra Censo 80 – Falante 07)

Neste artigo, analisamos os complexos de orações integrados por dois segmentos cujo primeiro apresenta o SN “o seguinte”, com o objetivo de investigar (i) os tipos de construções com a expressão “o seguinte” (ii) o tipo de relação semântica que elas instanciam.

Para a consecução desse objetivo, utilizamos, como representativa da modalidade de fala, a amostra Censo 80, que integra o acervo do grupo PEUL (Programa de Estudos sobre o Uso da Língua). Realizamos o processamento estatístico dos resultados pelo uso do programa *Makecell*, que compõe o pacote computacional Goldvarb 2001, para a obtenção de frequências.

Este texto está organizado da seguinte forma: na primeira parte, apresentamos as relações de Projeção e Elaboração, na segunda, os tipos de construções com a expressão “o seguinte”. Na terceira, o tipo de relação semântica que elas instanciam. A quarta parte compreende as considerações finais e a quinta, as referências bibliográficas.

## Projeção e elaboração

Halliday (2004) investiga as orações complexas<sup>4</sup> segundo dois eixos: o eixo sintático e o eixo semântico. O eixo sintático diz respeito às relações de interdependência entre orações que se podem realizar por *parataxe*, *hipotaxe* e *encaixamento*. Na parataxe, sentenças de igual estatuto encontram-se justapostas; na hipotaxe, combinam-se sentenças de estatuto desigual numa relação de dependência, sem que uma seja parte da outra e no encaixamento<sup>5</sup> ou subordinação, sentenças de estatuto desigual estão em uma relação de dependência, uma funcionando como parte (um argumento) da outra.

O eixo semântico se refere às relações proposicionais existentes entre termos ou orações que se agrupam em dois tipos fundamentais: Projeção (*Projection*) e Expansão (*Expansion*). A *expansão*, por sua vez, se subdivide em Elaboração (*Elaboration*), Extensão (*Extension*) e Realce (*Enhancing*).

Na Projeção, uma oração se projeta por meio de outra que a apresenta como uma *locução* ou uma *ideia*. Segundo o autor, *projeção* é “a relação lógico-semântica entre orações que têm como função não uma representação direta da experiência (não-linguística), mas a representação de uma representação (linguística)”<sup>6</sup> (HALLIDAY, 2004, p. 441), como mostram os exemplos abaixo:

3 As orações complexas que apresentam o SN “o seguinte” foram consideradas por Alencar (2010) e Ferreira (2009) como uma instanciação da relação de Elaboração.

4 As orações complexas ocorrem quando uma oração núcleo está relacionada a outras orações que a modificam.

5 Neste trabalho, não investigaremos as orações encaixadas.

6 No original: [...] *the logical-semantic relationship whereby a clause comes to function not as a direct representation of (nonlinguistic) experience but as a representation of a (linguistic) representation.* (HALLIDAY, 1994, p. 250)

- (4) *Quando você soube que era um escritor? – Eu sempre soube. Não me lembro nem mesmo de ter pensado em fazer outra coisa depois de quinze ou dezesseis anos.*<sup>7</sup> (HALLIDAY, 2004, p. 441)
- (5) *Mãe, você sabe onde a tesoura está? Preciso desesperadamente, desesperadamente dela. – O que? A tesoura. – Sim, pendurei-a no lugar de costume.*<sup>8</sup> (HALLIDAY, 2004, p. 441)

A projeção pode ser uma representação do conteúdo de uma clausula mental, ideia, ou do conteúdo de uma clausula verbal, locução.

A Expansão pode instanciar-se em uma Elaboração, uma Extensão ou um Realce. Na Elaboração, uma oração reformula, especifica ou comenta o significado daquela com que se liga.

- (6) *Você está velho demais para esse jogo; você não poderia se curvar.*<sup>9</sup> (HALLIDAY, 2004, p. 398)

Para este estudo, interessa-nos a Elaboração vinculada a sintagmas nominais, que se divide em dois grandes grupos. O dos SNs, que apresentam conteúdo semântico, e o dos SNs, essencialmente fóricos (ALENCAR, 2010). Vejamos alguns exemplos.

No primeiro, encontram-se os pronomes substantivos, os pronomes indefinidos e os SNs que funcionam como rótulos.

#### *Pronomes substantivos*

- (7) *Poderia acrescentar às frases do Eliakim uma que li nos jornais e que me apresso em transcrever: “apertem os cintos, vamos decolar” (Bebeto de Freitas). E aos que pensam em ironizar o otimismo do presidente do Botafogo, cuidado! Bebeto (foto) tem pedigree, é sobrinho do Saldanha. E, além da valentia, é homem sério. Nem cara de cartola ele tem.* (EXTRA 02-01-04 apud ALENCAR, 2010 p. 68)

#### *Pronomes indefinidos*

- (8) *Mas algo não pode ser esquecido: o voto de confiança foi dado a Lula, e não tanto a seu partido, que foi derrotado para os governos estaduais dos principais Estados (São Paulo, Rio, Minas, Bahia, Rio Grande do Sul, Pernambuco e Ceará). Qualquer passo em falso será cobrado nas urnas daqui a quatro anos.* (JB 01-11-02 apud ALENCAR, 2010 p. 69)

*Rótulos que entram em uma categoria mais geral* (FRANCIS, 1994)

- (9) *Mas de uma coisa eu tenho certeza: não faltou empenho e por isso nenhuma daquelas palavras serviu para o grupo.* (JB 06-03-04 apud ALENCAR, 2010, p. 70)

*Indiciais* (ALENCAR, 2010) – nomes indicadores que se referem a resultados estatísticos, valores de referência, indicadores financeiros, etc.

- (10) *Os elitistas que me desculpem, mas o aspecto de feira, a cara de supermercado, as pessoas saindo com sacolas cheias, os filhos atormentando os pais com pedidos, tudo isso é fundamental para tentar melhorar mais um desses índices que envergonham a cidade: segundo o Data/Gois/ UniCarioca, quatro em cada dez cariocas leem apenas um livro por ano. Não é engano não – um por ano. E isso significa que há muita gente que nem isso.* (O GLOBO 14-05-05 apud ALENCAR, 2010, p. 70)

7 No original: *When did you know you were a writer? – I always knew. I can't remember even considering doing anything else after I was about fifteen or sixteen.* (HALLIDAY, 2004, p. 441)

8 No original: *Mum, do you know where the scissors are? I desperately, desperately need them. – What? The scissors. – Yes, I hung the scissors up, in their usual spot.* (HALLIDAY, 2004, p. 441)

9 No original: *You're too old for that game; you couldn't bend over.* (HALLIDAY, 2004, p. 398)

*Resultado de atividade física ou mental* (ALENCAR, 2010).

- (11) *Essa notícia não será dada por ser alarmante e ajudar a vender jornal - mas sim por ser verdadeira e sua difusão necessária. E nunca de forma a servir de chamariz para outros bárbaros, por acaso desatentos. Não é tarefa fácil: depende de incontáveis decisões subjetivas e está sujeita a erros humanos. Mas é nosso trabalho. Por favor, como já disse o outro, não fuzilem o carteiro.* (O GLOBO 03-06-03 apud ALENCAR, 2010 p. 70)

No segundo conjunto, denominado de SNs com função dêitica inerente, estão incluídos os pronomes demonstrativos e o SN “o seguinte”, SNs que funcionam apenas como um indicador de direcionalidade discursiva, como mostram, respectivamente, os exemplos abaixo (12)<sup>10</sup> e (13):

- (12) *Ó, o senhor quer saber de uma coisa, vamos fazer isso: olha, está na hora do meu padraço chegar, ele é um cão de homem! Se eles encontrar vocês aqui, ele vão querer saber, vão querer até meter a língua em vocês. Vamos fazer isso: vocês vão, me esperam num caramanchão lá nas barca, que eu vou dar o almoço a ele. Ele e o amigo. Assim que eles sair eu apanho a bicicleta, vou atrás de vocês.* (Amostra Censo 80 – Falante 12)
- (13) *E: Então, era uma profissão que começou com os aviões grandes.  
F: É, começou aumentando com o tamanho do avião. Não, porque é o seguinte: o DC, nos aviões anteriores de- inclusive no DC, em termos comerciais, era avião que necessitava apenas dois- duas pessoas: o piloto aliás, inicialmente, ele começou até com um piloto só, não é? Depois os aviões comerciais precisaram de um co-piloto.* (Amostra Censo 80 – Falante 45)

Acreditamos que há uma diferença de grau entre os SNs com maior ou menor carga semântica, que aponta para uma função mais textual em elementos como a palavra “coisa” e se torna mais evidente no SN “o seguinte”.

Ao investigar as construções que apresentam “o seguinte” no primeiro segmento, percebemos que os limites entre Elaboração e a Projeção podem não ser tão definidos, como considera Halliday (2004). Nas próximas seções, procuraremos identificar quais são os tipos de construções existentes com o SN “o seguinte” e os tipos de relação semântica que elas estabelecem.

### **Tipos de construções com a expressão “o seguinte” no primeiro segmento**

As orações complexas que envolvem o SN “o seguinte<sup>11</sup>” em nossa amostra são compostas por dois segmentos, no primeiro dos quais é possível encontrar estruturas do tipo “é o seguinte”, representada pelo exemplo (14); “X é o seguinte”, em que a posição X é ocupada por SNs com maior ou menor conteúdo semântico, tais como “negócio” e “economia”, como mostram os exemplos (15) e (16) e “outros verbos + o seguinte”, ilustrada pelo exemplo (17).

- (14) *E- E vem cá, como faz essa para comprar um jogador, como os clubes fazem, não é, a transação?  
F- É o seguinte: a gente, quanto assina um contrato, com um clube, existe duas cláusulas, não é? Ou uma cláusula com dois item. Ou você tem passe livre, ou então você é preso ao clube, o passe*

<sup>10</sup> Os grifos nos exemplos (12) a (30) são nossos.

<sup>11</sup> Risso (1999), como introdução para o estudo dos marcadores *bem, bom, olha e ah*, identifica o SN “o seguinte” como um marcador discursivo em estruturas “é o seguinte” e como uma formulação metadiscursiva em construções “pergunta a você o seguinte”.

*pertence ao clube, ou então você, quando faz, faz o contrato com passe livre, que não é interessante para o jogador não. O jogador fica meio desvalorizado. Quando ele está livre, aí livre ninguém quer, não é? Está livre é porque ninguém quer! Então, esses jogadores que têm o passe caro são justamente os mais procurado, porque realmente também têm qualidades, não é verdade? E, então é isso, o clube fica com esse direito do passe. Então você, quando quer se transferir, o clube estipula o preço do seu passe. De acordo com o seu futebol! Você é bom, no caso, aqui, vamos citar: hoje o bom é o Zico. Então, se o clube rejeita a ofertas aí de cem, cento e cinquenta milhões, não é? Então ele rejeita isso aí, porque acha que o jogador tem mais valor. Essa é que é a lei da transferência, a lei do passe. Que nós somos jogadores. Eu digo: nós, os jogadores, somos preso sob contrato. (Amostra Censo 80 Falante 14)*

- (15) *Vamos dar o valor. Ora se aquilo botar [um] um salário mínimo, aí já ajuda que mínimo não vem. O salário máximo. Isso é que é não é nada de mínimo. **O negócio é o seguinte**: você- trabalha, que você terá mais. Pronto, está certo? Fez menos, ganha menos; trabalhou mais, tem mais, não é isso? É não. Nada disso vamos trabalhar, e aí dá para os outro. Escuta, vamos abrir estradas num lugar que possa fazer trilho de trem, porque o com o combustível está caro, vamos abrir estrada de ferro então, pra trazer – de lá, não é? (Amostra Censo 80 Falante 32)*
- (16) *I: Mas é mais econômico mesmo o carro a álcool?  
F: Não, **a economia é o seguinte**: um carro a gasolina, por exemplo, um fusquinha mil e trezentos, ele reguladinho ele chega a fazer onze quilômetros com um litro, com gasolina, Com álcool, ele faz seis, seis e meio, sete, cinco, quatro. Então, se você botar na balança, realmente, vai ficar elas por elas, quer dizer, no meu ponto de vista, eu não sei. Eu estou falando por aquilo que eu sei e por aquilo que eu mexo e vejo, certo? (Amostra Censo 80 Falante 45)*
- (17) *Olha, João, você amanhã **faz o seguinte**: o dia que você vim se for um dia de domingo, um de sábado, você vim de manhã, você chega mais cedo e tira ele, vamos dizer, cinco e meia, meia hora antes. (Amostra Censo 80 Falante 46)*

A tabela abaixo mostra a distribuição dos tipos de estrutura com a expressão “o seguinte” encontradas em nosso *corpus*:

**Tabela 1: Tipo de estrutura com a expressão “o seguinte”**

Tipo de estrutura	N	%
É o seguinte	15	33
X é o seguinte	13	29
Outros verbos o seguinte	17	38

Como mostra a Tabela 1, há um predomínio de ocorrências com o verbo relacional (HALLIDAY 1994) “ser”. Nas estruturas “outros verbos o seguinte” observam-se os tipos de verbo *materiais, mentais, verbais e existenciais*.

Halliday (1994) concebe os processos em termos da conjugação de três componentes: o processo por si mesmo, seus participantes e as circunstâncias nele envolvidas. Dessa forma, vai ao encontro da posição de outros autores como Givón (2001) e Hooper e Thompson (1980), alargando o conceito de transitividade para além da fronteiras do núcleo verbal.

Os processos *relacionais* são representados por verbos que estabelecem atributos, identificação e posse entre duas entidades. Os processos **relacionais intensivos**, por exemplo, caracterizam-se pela presença do verbo “ser” (*be*) e de dois grupos nominais. No exemplo abaixo, uma qualidade é atribuída a uma entidade (carregador): “x é um membro da classe de a”:

- (18) *F: Frequento com o padre João. Eu estou há quinze ano na feira da ruas. Eu fui o fundador das feira da rua. Isso é, a feira da ruas é o seguinte: é, cada rua monta uma barraca e a renda da rua é para a igreja, entendeu? Ah! Essa festa dá muito dinheiro! (Amostra Censo 80, Falante 07)*

Os processos *materiais* são instanciados por verbos de ação<sup>12</sup> em que há, pelo menos, um participante, o autor da ação e, facultativamente, um segundo participante, que é o objetivo do processo. Vejamos um exemplo extraído de nosso *corpus*:

- (19) *O sargento, faz o seguinte: o senhor engrena a dodge, que eu vou dar uma sangria, porque o freio quando tranca, é as rodas que tranca, porque o burrinho não dá retorno, o burrinho de freio. (Amostra Censo 80 Falante 07)*

Os processos *mentais* são exemplificados por verbos que expressam sentimento, pensamento ou percepção de um participante humano, o experienciador, e se pode referir a alguma entidade (pessoa, criatura, objeto, instituição ou abstração), ou a algum processo (ação, evento, qualidade, estado ou relação), como podemos observar neste exemplo de nosso *corpus*:

- (20) *F: bom. Eu acredito o seguinte: para você dar uma resposta, assim, leva muita coisa junto. Mas eu acredito e se eu acho que isso aí é – é uma falta de politicamente, todo mundo não estar engajado. Essa é que é a verdade. (Amostra Censo 80 Falante 45)*

Os processos *verbais* são relacionados à fala e são instanciados por verbos *dicendi*: *dizer, perguntar, responder*, etc., como podemos observar no exemplo (21):

- (21) *Outro dia, eu estava conversando com um colega meu e a respeito de política. A respeito de quê? Agora você vai votar? Não, não vou votar mais é nesses camarada não fazem nada. Então, eu disse para ele o seguinte: que à medida que a gente cobrar do político que a gente votou, esse político vai ter que fazer mais alguma coisa para a gente. E a mesma coisa seria o governo e todas essas coisa de governo. (Amostra Censo 80 Falante 45)*

Os processos *existenciais* representam algo que existe ou acontece no mundo real e podem ser representados, no português, por verbos como *haver, existir*, como ilustra o trecho (22):

- (22) *F: Os jovem não estão se afastando. Sabe que que acontece? Acontece o seguinte: é que sabe como é que é: a rapaziada vai naquela ilusão, conforme eu falei sobre o - casamento, é a mesma coisa. O jovem se deixa levar, muita da vezes pela namorada. (Amostra Censo 80 Fal. 07)*

No que concerne à vinculação sintática entre os dois segmentos que são objeto de nossa análise, observamos que ela pode dar-se por parataxe, como em (23), ou por hipotaxe, como mostra o exemplo (24):

- (23) *F: A primeira recomendação que nós tínhamos: “tem uma mutuca, era uma mosca, sabe? Uma mosca tipo grande assim. Aquela mosca é o seguinte: ela morde, onde ela mordeu, se o sujeito tiver um preparado, por exemplo o alho, coisas preventivo que se levavam, nós, soldados, não, tinha medicamento tem que botar logo para queimar, porque onde ela mordeu abre uma chaga, abre uma ferida, e ferida essa que batalha para fechar, e não é fácil fechar. (Amostra Censo 80 Falante 46)*

- (24) *F: Ah! Eu acho que sim. Se todas fizessem assim, acho que todas aqui estavam morando em apartamento. Porque a tendência da favela é crescer. Como você vê: os barracos de tábuas estão*

<sup>12</sup> Os verbos de ação são geralmente aqueles que envolvem uma mudança num intervalo de tempo.

*sumindo! Você só vê- daqui a pouco isso aqui vira uma selva de pedra não é? Então, eu acho o seguinte: que muitas mulheres se acomodam, sabe? Por ter por morar sabe? Se acomoda aquilo ali: em ter um teto, em ter aquela comida e outras coisas, mas, naquela situação ali, tudo bem. Mas não – sabe? Não pensa um amanhã de: "ah, eu vou construir; eu vou ajudar meu marido, eu vou fazer e acontecer!" Sabe? Por quê? Porque "ele" mas, para isso, para isso elas não servem, mas, para ficar às vezes se admirando não é? Então, outras falam: "ih! Fulano, crescendo, fulano está roubando!" (Amostra Censo 80 – Falante 10)*

Ocorreram apenas quatro dados de hipotaxe, predominando, portanto, a ligação paratática (43/47 = 91%). Entretanto, uma análise mais localizada dos 4 casos de ligação hipotática permite lançar luzes sobre a escassez dessa forma de ligação.

A vinculação por hipotaxe nos dados analisados envolve, sobretudo, o uso do conector “que”,<sup>13</sup> na introdução do segundo segmento, como ilustra (25):

(25) *Ele me disse o seguinte: que ele quer ir lá na piscina. (Amostra Censo 80 Falante 07)*

Em dois dos quatro casos de construções vinculadas hipotaticamente, observa-se a presença de verbos do tipo *relacional*, associados à estrutura *X* “é o seguinte”, em que a posição *X* pode ser ocupada por um SN genérico, como mostra o exemplo (26). A presença do elemento que constituindo estruturas de foco *ser* indica uma possibilidade de que ele esteja associado à função focalizadora da oração onde se encontra o SN “o seguinte”.

(26) *F: Porque a única coisa que não vai bem é o seguinte: que nós temos aqui uma dificuldade muito grande de colocar a documentação do bar em dia, por ser local por ser essa povoação assim, entendeu? (Amostra Censo 80 – Falante 10)*

As outras possibilidades de tipos de processo no primeiro segmento das construções vinculadas hipotaticamente foram os verbos *dicendi*, exemplo (25), e *mental*, exemplo (24).

Na próxima seção, discutiremos como a elaboração e a projeção associam-se aos tipos de construções com a expressão “o seguinte”.

## Tipo de relação semântica

Conforme especificado, segundo Halliday (2004), na Elaboração, um segmento desenvolve ou especifica o conteúdo de outro segmento. Dessa forma, nas estruturas em que o primeiro segmento corresponde ao tipo *X é o seguinte*, o N que ocupa a posição *X* é desenvolvido pelo segundo segmento, caracterizando uma relação de elaboração, como mostra o exemplo abaixo:

(27) *I: E como é que faz, assim: vocês se reúnem, assim, de repente. Você estava contando, ontem, que jogam, não é? As mesmas pessoas, não é?*  
*F: Não, não é reunião, o caso é seguinte: essa pelada que eu joga lá no Acari, foi criado pela gente mesmo. Então era todo mundo da mesma rua. (Amostra Censo 80 Falante 25)*

No que diz respeito ao SN “o seguinte” nessas construções, como mencionado na seção 1, essa expressão demonstra estar perdendo sua função ideacional e assumindo uma função essencialmente textual, atuando apenas no plano da organização do texto. Esse

<sup>13</sup> O outro conector possível nesses contextos seria *se* que não foi encontrado na amostra analisada.

esvaziamento semântico se evidencia, principalmente, em construções desse tipo, em que “o seguinte” funciona como um sinalizador da relação de elaboração.

Estudos na área da Teoria da Estrutura Retórica mostram que a sinalização contribui para a indicação das relações semânticas no texto. Segundo Dale (1992), para a língua escrita, três formas de sinalização explícita são mais recorrentes: as lexicais, como as expressões que funcionam como pistas (*next, anyway, to get back*), as marcas prosódicas como vírgulas, ponto e vírgula, dois-pontos e marcas gráficas, como quebra de parágrafo.

Na modalidade falada, os traços prosódicos contribuem para a classificação das relações retóricas. No que diz respeito à relação de Elaboração, Murray, Taboada e Renals (2006) mostram que tanto o núcleo como o satélite tendem a se caracterizar por alta duração, esses dois segmentos são separados por uma pausa longa e o segmento satélite apresenta baixa velocidade de fala.

Em nossos dados, a ocorrência de uma pausa<sup>14</sup> não final entre os dois segmentos foi tomado como um índice de ligação entre eles.

Nos casos de “X é o seguinte”, no entanto, a ligação entre os dois segmentos pode ser explicitada também pelo elemento catafórico “o seguinte” presente no primeiro segmento, como mostrado no exemplo (27) acima.

Nas estruturas em que o primeiro segmento corresponde a “é o seguinte”, a recorrência da estrutura “X é o seguinte” indica a possibilidade de se subentender que, naquelas construções, havia um N que foi apagado na posição X, por isso optamos por considerá-la também como um caso de elaboração, entretanto, em construções com “*outros verbos o seguinte*”, a depender da transitividade do verbo (HALLIDAY, 1994) e da natureza do segundo segmento, a relação estabelecida pode indicar mais ou menos claramente os limites entre elaboração e projeção.

Nos casos de *verbo dicendi ou verbo mental + o seguinte*, considerando o tipo de verbo, a baixa carga semântica de “o seguinte” e o que explicita Halliday (2004) sobre a Projeção, “relação lógico semântica que têm como função não uma representação direta da experiência (não-linguística), mas a representação de uma representação (linguística)” (HALLIDAY, 2004, p. 441), a relação está em um limite entre elaboração e projeção, como mostram os exemplos abaixo:

(28) *Outro dia, eu estava conversando com um colega meu e a respeito de política. A respeito de quê? Agora você vai votar? Não, não vou votar mais é nesses camarada não fazem nada. Então, eu disse para ele o seguinte: que à medida que a gente cobrar do político que a gente votou, esse político vai ter que fazer mais alguma coisa para a gente. E a mesma coisa seria o governo e todas essas coisa de governo. (Amostra Censo 80 Falante 45)*

(29) *F: bom. Eu acredito o seguinte: para você dar uma resposta, assim, leva muita coisa junto. Mas eu acredito e se eu acho que isso aí é – é uma falta de politicamente, todo mundo não estar engajado. Essa é que é a verdade. (Amostra Censo 80 Falante 45)*

Como podemos observar nesses exemplos, o SN “o seguinte” evidencia uma função, principalmente, textual, simultaneamente, catafórica e segmentadora e o segundo segmento é uma representação de uma representação linguística.

<sup>14</sup> Neste trabalho foi controlada apenas a presença da pausa entre os dois segmentos. Um estudo mais aprofundado deste fator exigiria, no entanto, um controle da possível influência da duração da pausa.

Em estudos anteriores, Braga e Alencar (2008) e Alencar (2010) destacaram a existência de construções hipotáticas com verbo do tipo material relacionada à Elaboração de SNs. Isso pode ser considerado como um argumento a favor da relação de *projeção* nos casos em que o segundo segmento se refere à representação de uma representação linguística de um pensamento ou daquilo que é dito e em que há um maior esvaziamento semântico do SN, como podemos perceber nos exemplos (28) e (29) mostrados anteriormente.

Em construções *verbo material + o seguinte*, embora o SN “o seguinte” funcione como um indicador de direcionalidade discursiva, o verbo material, que está relacionado a ações no mundo físico, e a natureza do segundo segmento fazem com que ele agregue a função de núcleo da relação de elaboração, como se pode ver no exemplo (30):

- (30) *E- Sei. Por falar nisso, me falaram que tinha uma estátua também que desapareceu dali.*  
*F- “Mas também” aqui desapareceu, mas foi um tal de - agora, **ele faz o seguinte**: quando termina a festa, eles guardam eles levam. Só vem na época, você coloca ali outra vez, porque é de bronze. E sabe como é, bronze negro a pessoas procuram, não é? Infelizmente, o mal nosso do Brasil- se é os próprio ralos, aqueles de ferro, eles carregam. São já ouviu já viu quantos buraco aí nessas ruas aqui? Tiraram- eles tiram os ferro. E faço ideia uma estátua de branze, não é isso? E aqui, a primeira roubaram, não é? Então, fizeram outra, mas, quando termina a festa, eles procuram levar ela para guarda, compreendeu? É. (Amostra Censo 80 – Falante 32)*

No primeiro segmento, “ele” é o participante 1, aquele que realiza a ação; o verbo fazer (faz) é o Processo material e “o seguinte” é o participante 2, a Meta. No segundo, encontra-se uma especificação do SN “o seguinte”, configurando uma representação direta da experiência (não linguística).

## Considerações finais

A análise dos limites entre elaboração e projeção em construções que apresentam o SN “o seguinte” na fala revelou que esse SN demonstra estar perdendo sua função ideacional e assumindo uma função essencialmente textual. Os tipos de construções identificadas indicam que, no primeiro segmento dos complexos oracionais em questão, podem ser encontradas estruturas do tipo “é o seguinte”, “X é o seguinte” e “*outros verbos* o seguinte”.

As ocorrências que apresentam a estrutura “X é o seguinte” apontam para uma interpretação como elaboração do SN na posição X, sendo que “o seguinte” funcionaria como sinalizador de elaboração. A recorrência dos casos de “X é o seguinte” indica a possibilidade de se subentender que, nas construções com “é o seguinte”, havia um N que foi apagado na posição X, por isso optamos por considerá-las também como casos de elaboração.

Nas construções com “*outros verbos* o seguinte” no primeiro segmento, a depender da transitividade do verbo (HALLIDAY, 1994) e da natureza do segundo segmento, a relação estabelecida pode indicar mais ou menos claramente os limites entre elaboração e projeção. Os casos de “verbo material + o seguinte” podem ser considerados como elaboração, já os de “verbo dicendi + o seguinte” e “verbo mental + o seguinte”, como casos limítrofes entre elaboração e projeção.

## REFERÊNCIAS

- ALENCAR, F. A. S. *A relação semântica de elaboração na fala e na escrita: forma e função*. 2010. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2010.
- BRAGA, M. L.; ALENCAR, F. A. S. É o seguinte: ajudar os amigos. In: VOTRE, S.; RONCARATI, C. (Org.). *Antony Julius Naro e a linguística no Brasil*. Rio de Janeiro: FAPERJ, 2008. p. 295-305.
- DALE, R. *Exploring the Role of Punctuation in the Signalling of Discourse Structure*. In: WORKSHOP ON TEXT REPRESENTATION AND DOMAIN MODELLING, 1991, Berlin. *Proceedings...*: Technical University of Berlin, 1992. p. 110-120.
- FERREIRA, J. C. L. “*Vamos fazer o seguinte...*”: a construção apositiva na interface gramática e interação. 2009. 129 f. Tese (Mestrado em Linguística) – Universidade Federal de Juiz de Fora, Juiz de Fora, 2009.
- FRANCIS, G. Labelling discourse: an aspect of nominal group lexical cohesion. In: COULTHARD, M. (Ed.) *Advances in written text analysis*. Londres: Routledge, 1994. p. 83-101.
- GIVÓN, T. *Syntax: an introduction*. v. I e II. Amsterdam/Philadelphia: John Benjamins, 2001.
- HALLIDAY, M. A. K. *An introduction to functional grammar*. 3. ed. London: Arnold, 2004 [1985]. p. 689.
- HOPPER, P. J.; THOMPSON, S. A. Transitivity in Grammar and Discourse. *Language*, v. 56, n. 2, p. 251-299, 1980.
- MURRAY, G.; TABOADA, M.; RENALS, S. Prosodic Correlates of Rhetorical Relations. In: WORKSHOP “ANALYZING CONVERSATIONS IN TEXT AND SPEECH”, 2006, New York. *Proceedings...* New York: North American Association for Computational Linguistics, June 2006. p. 1-7.
- RISSO, M. S. Aspectos textuais-interativos dos marcadores discursivos de abertura *bom, bem, olha, ah*, no português culto falado. In: NEVES, M. H. M. (Org.) *Gramática do Português Falado*. v. VII: Novos Estudos. São Paulo: Fapesp; Unicamp, 1999. p. 256-296.

# Alguns meios de expressão linguística das relações implícitas que se estabelecem entre orações e entre partes do texto<sup>1</sup>

(Some means of linguistic expression of the implicit relations which are held between clauses and between text spans)

Juliano Desiderato Antonio<sup>1</sup>

<sup>1</sup> Departamento de Letras, Programa de Pós-Graduação em Letras – Universidade Estadual de Maringá (UEM)

jdantonio@uem.br

**Abstract:** One of the main principles of Rhetorical Structure Theory is that, besides the propositional content conveyed by the clauses of a text, there are implicit prepositions which arise from the combining of clauses or of bigger text spans. Relations may be linguistically marked by discourse markers or connectives, by conversational implicatures, by the meaning of the verb, by the type of clause combining, verb tense and verb mood. The aim of this paper is to discuss two means of expression of rhetorical relations in spoken language: positive regard towards the content of the nucleus that the speaker intends to create in his addressee; sentence mood and modality.

**Keywords:** rhetorical relations; means of linguistic expression; Rhetorical Structure Theory.

**Resumo:** Um dos principais pressupostos da Teoria da Estrutura Retórica do Texto é o de que, além do conteúdo proposicional explícito veiculado pelas orações de um texto, há proposições implícitas que surgem da combinação entre orações ou entre porções maiores de textos. As relações podem ser marcadas linguisticamente de várias maneiras, como por marcadores discursivos ou conectivos, por implicaturas conversacionais, pelo significado do verbo, pelo tipo de combinação entre as orações, pelo modo e tempo verbais etc. Neste trabalho, pretende-se abordar duas formas de expressão das relações retóricas na língua falada: conceito positivo em relação ao conteúdo do núcleo que o falante pretende criar em seu destinatário; modo da oração e modalidade.

**Palavras-chave:** relações retóricas; meios de expressão; Teoria da Estrutura Retórica.

## Introdução

A compreensão de textos depende, dentre outros fatores, do reconhecimento de relações implícitas que são estabelecidas entre as partes do texto. Essas relações, chamadas “proposições relacionais” (MANN; THOMPSON, 1983), “relações discursivas”, “relações de coerência” ou “relações retóricas” (TABOADA, 2009) permeiam todo o texto, desde as porções maiores até as relações estabelecidas entre duas orações e ajudam a dar coerência ao texto, conferindo unidade e permitindo que o produtor atinja seus propósitos com o texto que produziu.

Um tratamento adequado a essa questão das relações de coerência é oferecido pela RST (*Rhetorical Structure Theory* – Teoria da Estrutura Retórica do Texto), uma teoria descritiva que tem por objeto o estudo da organização dos textos, caracterizando as relações que se estabelecem entre as partes do texto (MANN; THOMPSON, 1988; MATTHIESSEN; THOMPSON, 1988; MANN; MATTHIESSEN; THOMPSON, 1992).

---

<sup>1</sup> Este trabalho apresenta resultados de projeto financiado pela Fundação Araucária-PR, por meio de bolsa de Produtividade em Pesquisa (Convênio 939/2012 – Fundação Araucária – UEM).

A RST parte do princípio de que as relações retóricas que se estabelecem no nível discursivo organizam desde a coerência dos textos até a combinação entre orações (MATTHIESSEN; THOMPSON, 1988).

Ao tratar das relações retóricas tanto no nível discursivo quanto no nível gramatical (combinação entre orações), a RST demonstra sua filiação à Linguística Funcional, um grupo de teorias que consideram essencial para o estudo da língua a função dos elementos linguísticos na comunicação (BUTLER, 2003; NEVES, 1997; NICHOLS, 1984). Mais especificamente, a RST foi desenvolvida no âmbito de outras duas teorias funcionalistas: a Gramática Sistemico-Funcional de Halliday e o Funcionalismo da Costa-Oeste dos Estados Unidos (ANTONIO, 2009).

De acordo com a RST, as proposições relacionais surgem no texto independentemente de sinais específicos de sua existência: não há necessidade de inclusão, no texto, de elementos linguísticos que tenham por função indicar as relações estabelecidas (MANN; THOMPSON, 1983). No entanto, pesquisas têm sido realizadas no sentido de identificar os meios linguísticos utilizados pelos falantes como “pistas” que permitam a identificação das relações retóricas por parte dos destinatários. De acordo com Gómez-González e Taboada (2005) e Taboada (2009), alguns dos meios mais utilizados pelos falantes para marcar as relações são os conectivos e os marcadores discursivos (doravante MDs), que funcionam como “cue words”, ou seja, são palavras que fornecem pistas para a identificação das relações estabelecidas. Outros meios de expressão sugeridos por essas autoras são tempo verbal, modo da oração, encaixamento sintático, significado do verbo, implicaturas conversacionais. Em sua pesquisa a respeito da relação de contraste em diálogos, Ford (2000) observou que o significado das porções de texto também pode ser útil na identificação das relações.

As relações que se estabelecem entre as orações podem ser descritas com base na intenção comunicativa do enunciador e na avaliação que o enunciador faz do enunciatário, e refletem as escolhas do enunciador para organizar e apresentar os conceitos. A identificação dessas relações pelo analista, por sua vez, se baseia em julgamentos funcionais e semânticos, que buscam identificar a função de cada porção de texto, e verificar como o texto produz o efeito desejado em seu possível receptor.

Esses julgamentos são de plausibilidade, pois o analista tem acesso ao texto, tem conhecimento do contexto em que o texto foi produzido e das convenções culturais do produtor do texto e de seus possíveis receptores, mas não tem acesso direto ao produtor do texto ou aos seus possíveis receptores, de forma que não pode afirmar com certeza que esta ou aquela análise é a correta, mas pode sugerir uma análise plausível (MANN; THOMPSON, 1988).

Assim, este trabalho tem como objetivo discutir dois meios de expressão linguística de relações retóricas encontradas no *corpus* de língua falada do Funcpar (Grupo de Pesquisas Funcionalistas do Norte/Noroeste do Paraná), formado por elocuições formais gravadas durante aulas e por entrevistas com professores pesquisadores. Os meios de expressão de que este trabalho irá tratar são os seguintes: conceito positivo em relação ao conteúdo do núcleo que o falante pretende criar em seu destinatário; modo da oração e modalidade.

## Fundamentação teórica

### RST

Uma lista de aproximadamente vinte e cinco relações foi estabelecida por Mann e Thompson (1988) após a análise de centenas de textos, por meio da RST. Essa lista não representa um rol fechado, mas um grupo de relações suficiente para descrever a maioria dos textos.<sup>2</sup>

As relações retóricas são definidas pela RST com base em quatro características: a) restrições sobre o núcleo; b) restrições sobre o satélite; c) restrições sobre a combinação entre o núcleo e o satélite; d) efeito. De acordo com Gómez-González e Taboada (2005), a RST tem um viés em relação ao criador do texto, de forma que a característica mais importante na definição das relações é o efeito que o produtor do texto deseja atingir em seu destinatário, ou seja, sua intenção.

No que diz respeito às funções globais, as relações da RST podem ser divididas em dois grupos (MATTHIESSEN; THOMPSON, 1988):

a) relações que dizem respeito ao assunto (*subject matter*), que têm como efeito levar o enunciatário a reconhecer a relação em questão: elaboração, circunstância, solução, causa, resultado, propósito, condição, interpretação, meio, avaliação, reafirmação, resumo, sequência, contraste;

b) relações que dizem respeito à apresentação da relação (*presentational*), que têm como efeito aumentar a inclinação do enunciatário a agir de acordo com o conteúdo do núcleo, concordar com o conteúdo do núcleo, acreditar no conteúdo do núcleo ou aceitar o conteúdo do núcleo: motivação, antítese, fundo, competência, evidência, justificativa, concessão, preparação.

No que diz respeito à organização, as relações podem ser de dois tipos (MANN; THOMPSON, 1988):

a) núcleo-satélite (hipotáticas), nas quais uma porção do texto (satélite – S) é ancilar da outra (núcleo – N), como na Figura 1 a seguir, em que um arco vai da porção que serve de subsídio para a porção que funciona como núcleo.



Figura 1 – Diagrama de relação núcleo-satélite

b) multinucleares (paratáticas), nas quais uma porção do texto não é ancilar da outra, sendo cada porção um núcleo distinto, como na Figura 2 a seguir.

<sup>2</sup> Uma lista com as relações e suas definições pode ser encontrada no *site* <http://www.sfu.ca/rst/07portuguese/definitions.html>.

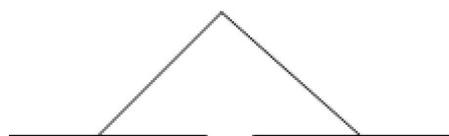


Figura 2 – Diagrama de relação multinuclear

## Análise das ocorrências

### Conceito positivo em relação ao conteúdo do núcleo que o falante pretende criar em seu destinatário

Na visão da RST, a noção de conceito positivo diz respeito a objetivos que o falante deseja alcançar com seu texto em relação ao destinatário (THOMPSON; MANN, 1987). Alguns objetivos mencionados por Thompson e Mann (1987) são persuadir, ou seja, levar o destinatário do texto a crer no conteúdo do núcleo, criar uma atitude de aprovação ou interesse em relação ao conteúdo do núcleo, criar um desejo, uma intenção de realizar o conteúdo do núcleo.

A partir da noção de conceito positivo, a RST realiza a distinção entre as relações de contraste (cf. Quadro 1) e de concessão (cf. Quadro 2). A última é considerada um recurso utilizado pelo falante para influenciar a crença do destinatário no conteúdo do núcleo, ou seja, é uma relação do tipo *presentational*, ao passo que a primeira é neutra, isto é, quando o falante utiliza a relação de contraste, não tem como intenção influenciar a crença do destinatário no conteúdo do núcleo.

Quadro 1. Definição da relação de contraste

Nome da relação	Restrições sobre cada par de núcleos	Intenção do falante
Contraste	Não mais do que dois núcleos; as situações nesses núcleos são (a) compreendidas como semelhantes em vários aspectos; (b) compreendidas como diferindo em alguns aspectos e (c) comparado com respeito a uma ou mais dessas diferenças.	O destinatário reconhece a comparabilidade e as diferenças levantadas pela comparação sendo feita.

Fonte: Mann e Taboada (2010).

Quadro 2. Definição da relação de concessão

Nome da relação	Restrições sobre N ou sobre S individualmente	Restrições sobre N + S	Intenção do falante/escritor
Concessão	Sobre N: o destinatário tem conceito positivo por N. Sobre S: o destinatário não está defendendo que S não tenha algo a ver.	O destinatário reconhece uma potencial ou aparente incompatibilidade entre N e S; reconhecer a compatibilidade entre N e S aumenta o conceito positivo do destinatário em relação a N.	O conceito positivo do destinatário em relação a N aumenta.

Fonte: Mann e Taboada (2010)

A observação dessa diferença entre as relações de contraste e de concessão pode ser observada nos exemplos das figuras 3 e 4 a seguir. Deve-se notar que, além da diferença

formal entre uma relação multinuclear (contraste) e uma relação núcleo-satélite (concessão), é critério decisivo para a RST a noção de conceito positivo na identificação dessas relações.

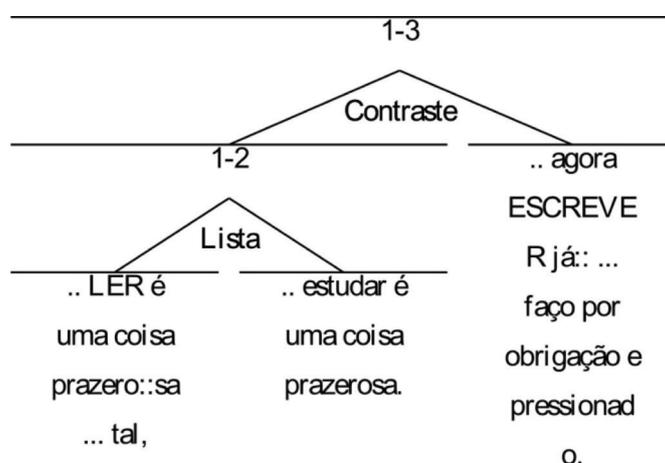


Figura 3 – Exemplo da relação de contraste

Fonte: Autoria própria.

No exemplo da Figura 3 (ANTONIO, 2012), o professor fala sobre suas atividades docentes. O gosto pela leitura e pelo estudo são apresentados como atividades que têm um mesmo estatuto (unidades 1 e 2). Essas atividades são, então, contrastadas com a obrigação profissional de escrever por meio do marcador discursivo *agora* iniciando a unidade 3.



Figura 4 – Exemplo da relação de concessão

Fonte: Autoria própria.

No exemplo da Figura 4 (ANTONIO, 2011), a oração concessiva posposta exerce função discursiva de adendo ou *afterthought* (DECAT, 2001). O fato de se utilizar, na transcrição, o ponto final após a unidade 1 indica uma entonação descendente, típica de final de frase, o que ajuda a comprovar prosodicamente essa análise. O professor, que vinha falando ao longo da aula a respeito da substituição das plantações de alimentos por plantações de cana-de-açúcar, afirma que prefere comida (arroz e feijão) ao combustível (álcool). Na oração concessiva, o falante faz uma ressalva, mas o vocábulo álcool, nessa oração, tem sentido de bebida alcoólica, e não de combustível, como na unidade anterior, provocando, dessa maneira, um efeito de humor durante a aula.

## Modo da oração e modalidade

O modo da oração e a modalidade também podem servir como pista para identificação de relações. Além das relações de solução e de preparação, identificadas por Antonio e Takahashi Barbosa (2012) como sendo estabelecidas por orações interrogativas, como nos exemplos das figuras 5 e 6, as relações de motivação e de competência podem ser identificadas quando realizadas por orações no imperativo (cf. figuras 7 e 9) ou por orações que contenham verbos na modalidade deôntica (cf. Figura 8).

Conforme apontado por Antonio e Takahashi Barbosa (2012), no exemplo da Figura 5, o aluno faz uma pergunta ao professor. Essa pergunta, que funciona como porção nuclear (unidades 596-597), traz um problema, que é solucionado pelas informações apresentadas pelo professor na porção que funciona como satélite (unidades 598-601).

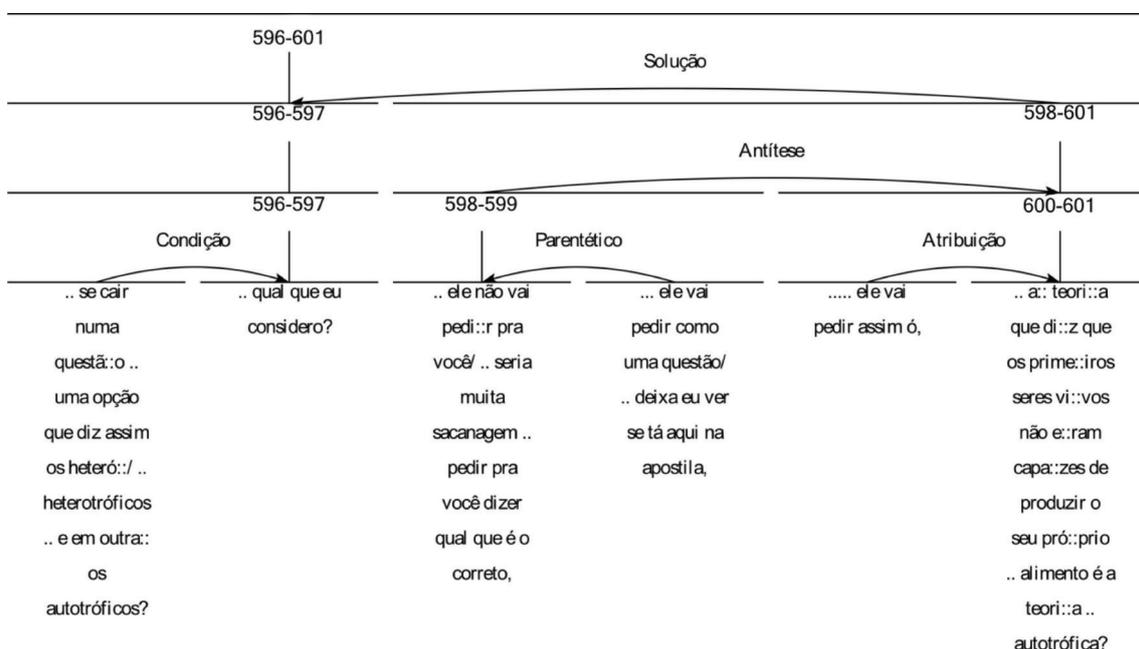


Figura 5 – Exemplo da relação de solução

Fonte: A autoria própria

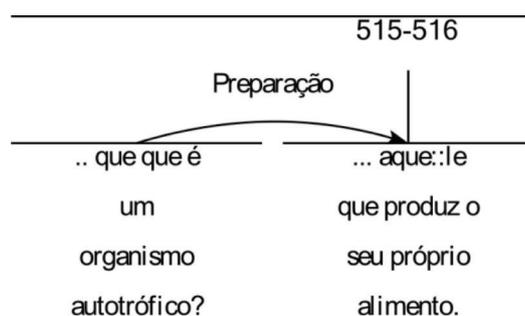
A definição da relação de solução é apresentada no Quadro 3.

Quadro 3. Definição da relação de solução

Nome da relação	Restrições sobre o núcleo ou sobre o satélite individualmente	Restrições sobre núcleo + satélite	Intenção do falante
Solução	Sobre N: N apresenta um problema.	S é uma solução para o problema apresentado em N.	O destinatário reconhece S como uma solução para o problema apresentado em N.

Fonte: Mann e Taboada (2010).

Conforme observado por Antonio e Takahashi Barbosa (2012), há casos em que o professor não espera que seus alunos respondam, apresentando ele mesmo a resposta. Trata-se das perguntas retóricas, utilizadas pelo professor ao longo da aula como forma de despertar o interesse de seus alunos no conteúdo que apresentará em seguida, como no exemplo da Figura 6.



**Figura 6 – Exemplo da relação de preparação**

Fonte: Aurtoria própria.

A relação que se estabelece entre a pergunta e a resposta é a relação de preparação (cf. Quadro 4).

**Quadro 4. Definição da relação de preparação**

Nome da relação	Restrições sobre o núcleo ou sobre o satélite individualmente	Restrições sobre núcleo + satélite	Intenção do falante
Preparação	Não há.	S precede N no texto; S tende a fazer com que o destinatário esteja mais preparado, interessado ou orientado para ler/ouvir N.	Tornar o destinatário mais preparado, interessado ou orientado para ler/ouvir N.

Fonte: Mann e Taboada (2010).

O modo imperativo e verbos modais deônticos são utilizados pelo professor para marcar as relações de competência e de motivação, definidas, respectivamente, nos quadros 5 e 6.

**Quadro 5. Definição da relação de competência**

Nome da relação	Restrições sobre o núcleo ou sobre o satélite individualmente	Restrições sobre núcleo + satélite	Intenção do falante
Competência	Sobre N: N apresenta uma ação que deverá ser realizada pelo destinatário.	A compreensão do conteúdo de S, por parte do destinatário, aumenta sua habilidade potencial para realizar a ação apresentada em N.	O destinatário tem aumentada sua habilidade potencial para realizar a ação apresentada em N.

Fonte: Mann e Thompson (1998).

**Quadro 6. Definição da relação de motivação**

Nome da relação	Restrições sobre N ou sobre S individualmente	Restrições sobre N + S	Intenção do falante/escritor
Motivação	Sobre N: N é uma ação não realizada com respeito ao contexto de N na qual o destinatário é o ator (incluindo aceitar uma oferta).	Compreender S aumenta o desejo do destinatário de realizar a ação em N.	O desejo do destinatário de realizar a ação em N é aumentado.

Fonte: Mann e Taboada (2010).

No caso da relação de competência, o imperativo é utilizado pelo professor para indicar aos alunos comportamentos e atitudes que eles devem ter para que possam realizar as ações apresentadas no núcleo. No exemplo da Figura 7, o professor de matemática sugere que os alunos “observem”, “olhem” a maneira como ele resolveu o exercício. Isso aumentará a capacidade de eles visualizarem os termos da equação.

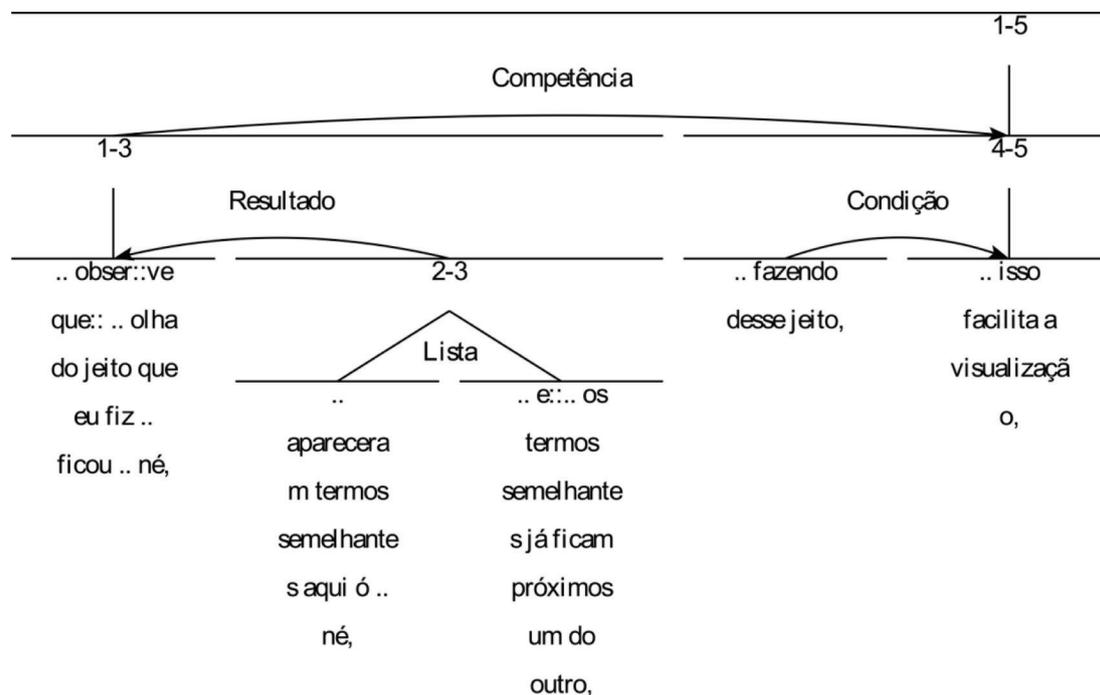
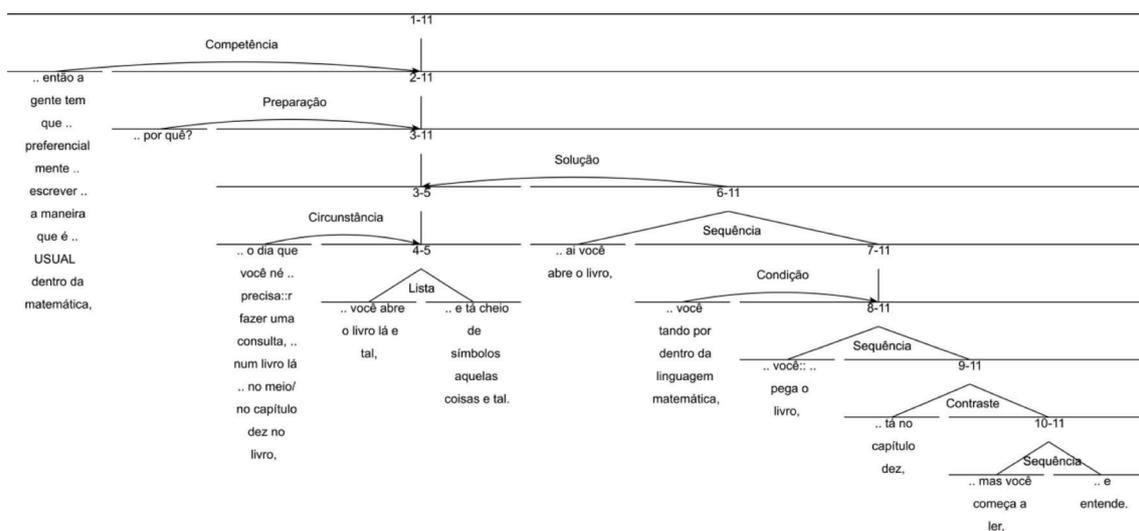


Figura 7 – Exemplo da relação de competência pelo modo imperativo

Fonte: Autoria própria.

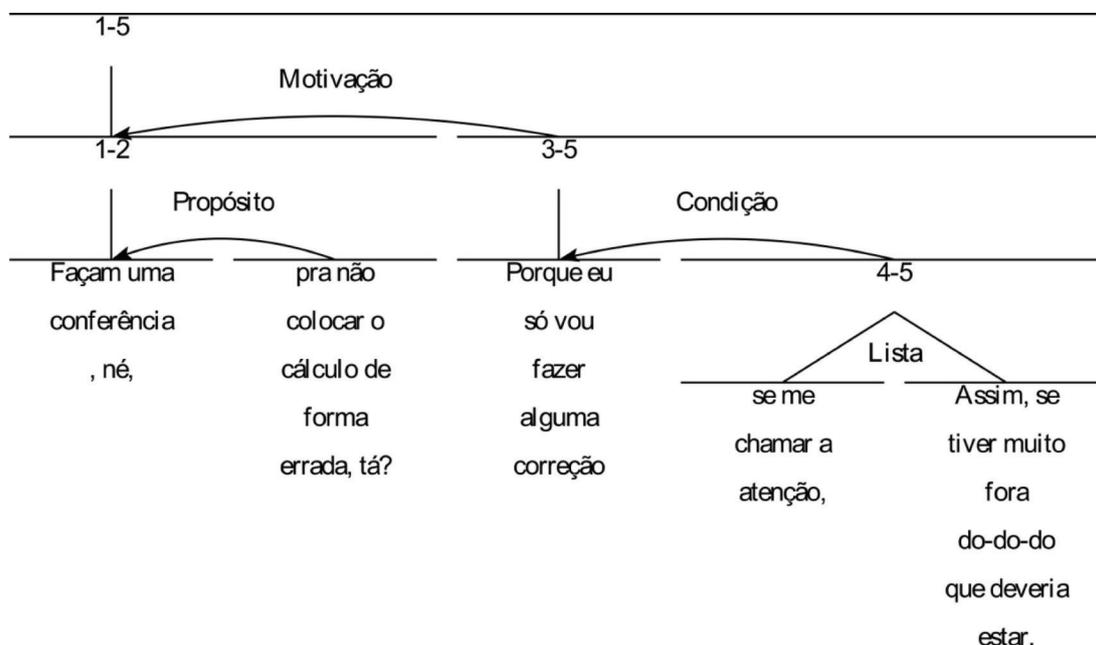
A modalidade deôntica, que atua no eixo da conduta (NEVES, 2006), também é utilizada pelos professores para aumentar a habilidade dos alunos para realizar o conteúdo do núcleo. No exemplo da Figura 8, o professor explica aos alunos, por meio da expressão modal deôntica “ter que”, a necessidade de utilizarem a maneira usual de se escrever na Matemática (unidade 1). Isso trará a eles habilidade para realizar o conteúdo do núcleo (unidades 2-11), a saber, compreender a linguagem matemática caso venham a consultar um livro dessa disciplina.



**Figura 8 – Exemplo da relação de competência marcada por verbo modal deontico**

Fonte: Autoria própria.

No caso da relação de motivação, a intenção do falante é aumentar o desejo do destinatário de realizar a ação apresentada no núcleo. Como observado por Antonio (2012), no exemplo da Figura 9, o conteúdo do satélite (unidades 3-5) traz uma motivação (o professor só irá corrigir os cálculos se realmente existir um erro que chame a atenção) para que os alunos realizem a ação da porção nuclear (unidades 1-2), que os alunos verifiquem seus cálculos. Caso os alunos não realizem o conteúdo do núcleo, os cálculos permanecerão errados.



**Figura 9 – Exemplo da relação de motivação**

Fonte: Autoria própria.

## Conclusão

Partindo do pressuposto de que há relações implícitas que são estabelecidas a partir da combinação entre orações e entre partes do texto, este trabalho teve como objetivo discutir dois meios de expressão linguística de relações implícitas encontradas em um *corpus* de língua falada formado por elocuições formais e por entrevistas. Os meios de expressão investigados foram o conceito positivo em relação ao conteúdo do núcleo que o falante pretende criar em seu destinatário; modo da oração e modalidade.

A noção de conceito positivo diz respeito a objetivos que o falante deseja alcançar com seu texto em relação ao destinatário. Essa noção foi utilizada para distinguir as relações de contraste e de concessão. A última é considerada um recurso utilizado pelo falante para influenciar a crença do destinatário no conteúdo do núcleo, ao passo que a primeira é neutra, isto é, não é utilizada pelo falante com intenção de influenciar a crença do destinatário no conteúdo do núcleo.

No caso de modo da oração e modalidade, orações no interrogativo podem marcar relação de solução se houver um pedido legítimo de informação na porção nuclear, considerado um problema, para o qual será apresentada uma solução na porção satélite. Orações no interrogativo também podem marcar relação de preparação quando o falante utiliza a pergunta apenas para criar na audiência o interesse pelo que será falado na sequência. Orações no imperativo ou com verbos modais deônticos são utilizados pelos professores informantes do *corpus* para indicar aos alunos comportamentos e atitudes que eles devem ter para que possam realizar as ações apresentadas no núcleo. Orações no imperativo também podem marcar relação de motivação, em que a intenção do falante é aumentar o desejo do destinatário de realizar a ação apresentada no núcleo.

## REFERÊNCIAS

ANTONIO, J. D. Expression of cause, evidence, justify and motivation rhetorical relations by causal hypotactic clauses in Brazilian Portuguese. *Acta Scientiarum Language and Culture*, Maringá, v. 34, n. 2, p. 253-268, 2012.

\_\_\_\_\_. Expressão da relação retórica de concessão em elocuições formais e entrevistas orais. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 19, p. 143-166, 2011.

\_\_\_\_\_. O texto como objeto de estudo na Linguística Funcional. In: ANTONIO, J. D.; NAVARRO, P. (Org.) *O texto como objeto de ensino, de descrição lingüística e de análise textual e discursiva*. Maringá: Eduem, 2009. p. 61-80.

ANTONIO, J. D.; TAKAHASHI BARBOSA, C. Relações retóricas estabelecidas por perguntas em elocuições formais. *Revista Todas as Letras*, São Paulo, v. 14, n. 2, p. 186-197, 2012.

BUTLER, C. S. *Structure and function: a guide to three major structural-functional theories*. Part 1: approaches to the simple clause. Amsterdam/Philadelphia: J. Benjamins, 2003. 570 p.

DECAT, M. B. A articulação hipotática adverbial no português em uso. In: DECAT, M. B. et al. (Org.) *Aspectos da gramática do português: uma abordagem funcionalista*. Campinas: Mercado de Letras, 2001. p. 103-166.

FORD, C. The treatment of contrasts in interaction. In: COUPER-KHULEN, E.; KORTMANN, B. (Ed.). *Cause – Condition – Concession – Contrast: Cognitive and Discourse Perspectives*. Berlin: Mouton de Gruyter, 2000. p. 283-310.

GÓMEZ-GONZÁLEZ, M. A.; TABOADA, M. Coherence Relations in Functional Discourse Grammar. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. A. (Ed.) *Studies in Functional Discourse Grammar*. Berne: Peter Lang, 2005. p. 227-259.

MANN, W. C.; THOMPSON, S. A. Rhetorical Structure Theory: toward a functional theory of text organization. *Text*, v. 8, n. 3, p. 243-281, 1988.

\_\_\_\_\_. *Relational propositions in Discourse*. ISI/RR-83-115, 1983.

MANN, W. C.; MATTHIESSEN, C.; THOMPSON, S. A. Rhetorical Structure Theory and text analysis. In: MANN, W.C.; THOMPSON, S. A. (Ed.) *Discourse description: diverse linguistic analyses of a fund-raising text*. Amsterdam/Philadelphia: J. Benjamins, 1992. p. 39-77.

MANN, W. C.; TABOADA, M. *RST Web Site*. 2010. Disponível em: <<http://www.sfu.ca/rst/01intro/definitions.html>>. Acesso em: 28 ago. 2012.

MATTHIESSEN, C.; THOMPSON, S. The structure of discourse and ‘subordination’. In: HAIMAN, J.; THOMPSON, S. (Ed.) *Clause Combining in Grammar and Discourse*. Amsterdam/Philadelphia: John Benjamins, 1988. p. 275-329.

NEVES, M. H. M. *Texto e gramática*. São Paulo: Contexto, 2006.

\_\_\_\_\_. *A Gramática Funcional*. São Paulo: Martins Fontes, 1997.

NICHOLS, J. Functional theories of grammar. *Annual review of Anthropology*, Palo Alto, v. 43, p. 97-117, 1984.

TABOADA, M. Implicit and explicit coherence relations. In: RENKEMA, J. (Ed.) *Discourse, of course*. Amsterdam: John Benjamins, 2009. p. 127-140.

THOMPSON, S. A.; MANN, W. C. A discourse view of concession. In: DELANCEY, S.; TOMLIN, R. (Ed.). *Proceedings of the Second Annual Pacific Linguistics Society*. Eugene: University of Oregon, 1987. p. 435-447.

# Gramaticalização de construções como tendência atual dos estudos funcionalistas

(Grammaticalization of constructions as current tendency of functionalist studies)

Mariangela Rios de Oliveira<sup>1</sup>

<sup>1</sup>Instituto de Letras – Universidade Federal Fluminense (UFF)/CNPq/Faperj

mariangela.rios@terra.com.br

**Abstract:** This paper discusses the process of the grammaticalization and the degrees of grammaticality of the linguistic usage in a unique view. This perspective links the functionalist and cognitivist theoretical orientation through a constructional approach. It is assumed that such association contributes positively to the more integrated research on patterns research, in terms of the correlation between function x form from which they result.

**Keywords:** Grammaticalization; constructions; Portuguese; Functionalism.

**Resumo:** O artigo trata do processo de gramaticalização e dos graus de gramaticalidade dos usos linguísticos numa perspectiva que une a orientação teórica funcionalista à cognitivista, por intermédio da abordagem construcional. Assume-se que tal associação contribui positivamente para a pesquisa mais integrada dos padrões de uso, em termos da correlação função x forma de que resultam.

**Palavras-chave:** Gramaticalização; construções; língua portuguesa; Funcionalismo.

## Introdução

Neste artigo, voltamo-nos para um dos rumos mais recentes e promissores da pesquisa de orientação funcionalista de vertente norte-americana. Estamos nos referindo a um tipo de investigação que, inspirado nas contribuições do Cognitívismo, mais especificamente em Croft e Cruse (2004) e Goldberg (2006), busca nos estudos sobre *construção* fundamentos teóricos para a análise interpretativa de trajetórias de gramaticalização, em perspectiva histórica, e de graus de gramaticalidade, em viés sincrônico, nos termos de Traugott (2008; no prelo) e Noël (2007), entre outros.

Assumimos que a investigação de padrões de uso linguístico, no Funcionalismo, aliada à abordagem construcional, no âmbito dos estudos cognitivistas, tende a conferir maior rigor analítico às pesquisas desenvolvidas nessa vertente. Tal rigor advém de dois aspectos que se salientam a partir de tal interface: a maior vinculação do binômio *função x forma*, que passa a ser considerado como de forte correspondência, e o maior controle na detecção, descrição e interpretação das propriedades atuantes na correlação de aspectos funcionais e formais.

Assim posto, nosso objetivo neste artigo é o de apresentar tal vertente de estudos, destacando sua contribuição para a pesquisa de padrões de uso do português e consequente abertura de uma vasta e instigante agenda de trabalho. Para tanto, na primeira seção do artigo, trazemos distintas concepções do termo *gramaticalização* no âmbito do Funcionalismo, desde as clássicas até as mais recentes; nessa parte, chamamos a atenção para a constante referência à dimensão contextual envolvida na mudança linguística. Na

segunda seção, dedicamo-nos de forma mais específica ao tratamento construcional nos estudos sobre processos de gramaticalização e graus de gramaticalidade, destacando as contribuições dessa vinculação para a pesquisa da língua em uso. A terceira seção é dedicada principalmente à exposição dos principais postulados de Bybee (2010) e Traugott (2008; no prelo), como fontes de alta relevância para a proposta do tratamento construcional na esfera dos estudos funcionalistas que aqui defendemos. Por fim, na quarta seção, de feição mais empírica face às demais, trazemos resultados de nossa pesquisa sobre padrões de uso de expressões verbais integradas por pronomes locativos, como *aí está* e *vamos lá*, a partir do aparato teórico aqui apresentado e defendido. Encerramos o artigo com a apresentação de possíveis rumos de pesquisa, na ratificação de que o tratamento construcional dos aspectos funcionais do uso linguístico abre uma extensa e promissora agenda de trabalho para os pesquisadores dessa área científica.

### Concepções de *gramaticalização* e o papel do *contexto*

Ao longo das pesquisas sobre mudança de forma e sentido em viés funcionalista, podemos observar a crescente valorização dos aspectos contextuais na descrição e na interpretação do uso linguístico. Desde a primeira referência ao termo *gramaticalização* feita por Meillet (1921 [1912]), que o definia como a atribuição de caráter gramatical a um elemento originalmente autônomo, de natureza lexical, passando pela definição de Kurylowicz (1965), para quem a gramaticalização consistia na passagem de um morfema de valor lexical para o estatuto de gramatical, ou ainda de uma categoria menos para outra mais gramatical, o conceito sobre o termo tem sofrido releituras e reformulações.

Tais reelaborações conceituais, longe de negarem as concepções iniciais sobre polissemia e mudança gramatical, têm contribuído para o refinamento do aparato teórico funcionalista, que, assim, acompanha o desenvolvimento das pesquisas nas distintas áreas da Linguística e das ciências afins, refinando e redimensionando seus pressupostos. Assim é que, ao longo dos anos, desde Meillet e Kurylowicz, registram-se várias referências que ampliam o escopo da análise funcionalista, na consideração das relações associativas ou contextuais na investigação dos fenômenos linguísticos.

À guisa de exemplificação, apresentamos a seguir cinco dessas referências, entre as muitas que poderiam aqui figurar:

- a. DeLancey (1993): estabelece que o ponto inicial do processo de gramaticalização é uma *construção produtiva*.
- b. Bybee (2003): define gramaticalização como *criação de novas construções*.
- c. Himmelmann (2004): destaca que elementos se gramaticalizam *em contextos sintáticos específicos*.
- d. Haspelmath (2004): conceitua gramaticalização como mudança diacrônica pela qual partes de um *esquema construcional* passam a desenvolver forte *dependência interna*.
- e. Traugott e Dasher (2005): defendem a relevância da metonimização para mudança gramatical.

Como podemos observar pelas cinco declarações referidas, preocupações de ordem contextual, atinentes a ambientes sintáticos mais amplos ou esquemas específicos de uso, passam a constituir foco de atenção e controle por parte dos estudiosos, principalmente a partir dos anos 90 do século XX. De todo modo, tal preocupação carece de maior rigor e controle, uma vez que, via de regra, termos como *construção produtiva*, *esquema construcional*, *contexto sintático específico*, entre outros, carecem de precisão conceitual.

Como bem salientam Noël (2007) e Traugott (2008; no prelo), essa incorporação da referência contextual ainda é feita de modo pouco sistemático e rigoroso. O *contexto* é tratado como entidade vaga, genérica, de contornos pouco ou nada definidos e, por isso mesmo, sua abordagem, tanto do ponto de vista teórico quanto metodológico, torna-se tarefa de difícil e complexa execução. Tal dificuldade coloca-se, ainda hoje, como verdadeiro desafio aos funcionalistas, uma vez que, ao assumir efetivamente as relações contextuais como motivadoras dos fenômenos linguísticos, é preciso dar conta dessa dimensão com maior rigor e precisão, sob pena de os resultados de pesquisa não serem capazes de testar hipóteses e atingir as metas estabelecidas inicialmente. Outro risco é, sob o rótulo *contexto*, serem abrigadas, de modo indiscriminado, motivações de natureza muito diversa, sem o devido embasamento teórico que justifique sua seleção e tratamento. Uma das alternativas teórico-metodológicas para a assunção do contexto de modo mais efetivo e rigoroso é justamente seu tratamento em perspectiva construcional.

O que as declarações de *a* e *e* destacam é que a investigação de itens isolados, como se fossem detentores de autonomia e percurso próprio, vai cada vez mais cedendo espaço no Funcionalismo à consideração de expressões maiores, de padrões de uso específico, de unidades pré-fabricadas, nos termos de Erman e Warren (2000), de tipo lexical, sintático ou pragmático. Nesse sentido, como demonstram Traugott e Dasher (2005), os processos metonímicos passam a ganhar destaque sobre os metafóricos, uma vez que derivações de sentido são interpretadas como resultantes de relações associativas, relacionais e indexais. Na verdade, levando em conta a relação motivada *função* > *forma*, tão cara à pesquisa funcionalista clássica, com o destaque para o eixo funcional, o que vemos agora é uma verdadeira guinada, que passa a destacar a dimensão estrutural na pesquisa dos usos linguísticos. Assim, os aspectos formais, que durante anos foram vistos como mera consequência de motivações semânticas ou pragmáticas, são reconsiderados e alçados a outro patamar, segundo o entendimento de que é preciso levá-los mais em conta, como já prenunciava Givón (1995).

A relevância dos aspectos estruturais e das relações associativas, com a seleção de objetos de pesquisa mais complexos e tomados de forma holística, acaba por aproximar, de modo efetivo, o Funcionalismo do Cognitivismo. Tal aproximação se consubstancia, de forma mais evidente, no viés dos estudos de *gramaticalização de construções*.

## **Gramaticalização como construcionalização**

Do conjunto dos postulados cognitivistas, têm interessado mais especificamente ao Funcionalismo os relativos à concepção construcional, uma vez que tais postulados têm se mostrado produtivos para a pesquisa da mudança gramatical e das relações contextuais aí envolvidas.

De acordo com Goldberg (1995; 2006), Croft (2001) e Croft e Cruse (2004), construções são entendidas como esquemas virtuais, altamente abstratos, com forte vinculação

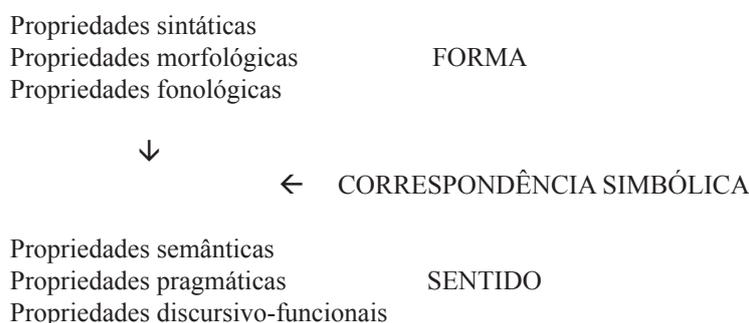
de forma e sentido. Essa vinculação se identifica por uma série de traços definidores, como a proposição de que o sentido de uma construção não corresponde à mera soma do sentido de cada um de seus constituintes e de que a alta vinculação semântico-sintática de uma construção tende a impossibilitar inversões de ordem interna entre seus constituintes e mesmo inserções de outros termos. Para os cognitivistas, a língua é concebida como um grande inventário de construções, de diversa extensão e tipologia, sem maior distinção entre os níveis do léxico e da gramática. Outra marca da abordagem construcional é sua desvinculação em relação à trajetória histórica das línguas, ou seja, não há propósito em investigar possíveis derivações de sentido e forma que motivariam mudanças de um padrão construcional, seja em menor seja em maior amplitude.

Assim posto, como salienta Noël (2007), é preciso selecionar e controlar quais os pressupostos construcionais que mais de perto interessam à pesquisa de cunho funcionalista, na consideração primeira das distinções que se manifestam entre as duas vertentes linguísticas. Em primeiro lugar, os funcionalistas entendem que há, ainda que de forma não categórica e discreta, o plano do léxico e o da gramática, considerando que existem sentidos, como os mais concretos e referenciais, e que há classes morfológicas, como os nomes e os verbos plenos, que prototipicamente se encontram no plano lexical, enquanto outros sentidos, como os procedurais e epistêmicos, por exemplo, e outras classes morfológicas, como as preposições e conjunções, se localizam de modo exemplar no plano da gramática. O tratamento das duas instâncias – léxico e gramática – separa as abordagens funcionalista e cognitivista, mas não inviabiliza o diálogo das pesquisas em ambas as áreas.

Outro ponto de distinção das duas abordagens é a consideração da perspectiva histórica no Funcionalismo, que se releva pelo foco nos estudos sobre polissemia e mudança gramatical. Mais uma vez, levando em conta tal distinção, é possível manter o diálogo entre as duas vertentes teóricas. Uma dessas possibilidades reside na consideração de que ao Cognitivism interessam os padrões construcionais fixados e estabilizados na língua, enquanto ao Funcionalismo cabe, além dessa tarefa, a pesquisa dos processos que conduziram a tal fixação, as mudanças de forma e sentido ocorridas na trajetória histórica das línguas. Conforme destaca Noël (2007), as construções de uma língua podem ser dispostas num *cline*, de modo que, desse conjunto, o Funcionalismo investigue aquelas que mais de perto lhe dizem respeito, como, por exemplo, as resultantes de mudança gramatical, motivadas por pragmatização.

Das contribuições cognitivistas, uma das mais relevantes para o Funcionalismo tem sido a proposta de pareamento *forma x sentido* defendida por Croft (2001, p. 18), assim exposta:

## CONSTRUÇÃO



De acordo com Croft (2001) e Croft e Cruse (2004), as dimensões da forma e do sentido, que compõem a construção, são constituídas, cada qual, por três propriedades básicas. As duas dimensões se encontram fortemente ligadas por correspondência simbólica. A frequência *type* é considerada determinante para a fixação do nível de entrenchamento das referidas dimensões enquanto esquemas, cabendo ao analista identificar e correlacionar as seis propriedades, de modo que, juntas e articuladas, sejam capazes de descrever a construção. De acordo com os autores, o modelo formulado é *baseado no uso*, ou seja, a virtualidade construcional é estabelecida na pesquisa pela coleta, descrição e análise de dados empíricos. Esses dados afetam e são afetados pelos modelos construcionais no momento da interação, que, por sua vez, tem forte apoio na cognição e na experiência.

Tal proposta de análise tem sido relevante e produtiva para as pesquisas funcionalistas que passam a investigar os usos linguísticos em perspectiva construcional. Em Bybee (2010) e mais notadamente em Traugott (2008; no prelo), encontramos referências e propostas de refinamento do modelo acima apresentado. Conforme a segunda autora, cabe ao Funcionalismo pesquisar de modo mais específico dois tipos de mudança pelas quais passam as construções: as que afetam subcomponentes de uma construção, as *mudanças construcionais*, e as que criam um novo e convencionalizado pareamento de forma e sentido, uma nova combinação, as *construcionalizações*. Ainda de acordo com Traugott (no prelo), as mudanças construcionais podem preceder ou acompanhar processos de construcionalização.

### Tendências atuais da pesquisa

Nesta seção, voltamo-nos mais especificamente para duas pesquisadoras funcionalistas que têm se dedicado à recente formulação teórico-metodológica na interface dos estudos funcionalistas e cognitivistas – Joan Bybee e Elizabeth Traugott. Trazemos aqui algumas de suas postulações que têm se mostrado produtivas e relevantes para a pesquisa na área.

Em Bybee (2010, p. 6-8), defende-se que os usos linguísticos são marcados, concomitantemente, pela sistematicidade e pela complexidade. Nesse sentido, variação, gradiência e mudança gramatical devem ser consideradas como manifestações que seguem determinadas tendências gerais e que afetam ambientes sintáticos mais amplos. Confor-

me a autora, em sua proposta também *baseada no uso*, são identificados cinco processos linguísticos de domínio geral:

- a) Categorização: propriedade cognitiva em reconhecer e associar entidades por semelhança, em representar e agrupar por meio de traços comuns.
- b) Encadeamento (*chunking*): vinculação de sentido e forma de constituintes, na formação de unidades complexas.
- c) Memória enriquecida: estocagem de detalhes da experiência com a língua, envolvendo conhecimento gramatical, semântico e pragmático dos usos linguísticos.
- d) Analogia: produção de enunciados com base em outros já produzidos, que servem de modelo para novos.
- e) Associação transmodal: experiências co-ocorrentes tendem a ser associadas em termos cognitivos.

Segundo Bybee (2010), construções são entendidas como padrões *efetivos* de uso. Em outras palavras, a autora se afasta um pouco mais do modelo cognitivista, ao tomar a construção como ocorrência ou dado empírico, e não como instância abstrata e virtual. De acordo com Bybee (2010), pelo processo de categorização, são definidos os representantes construcionais exemplares, que acabam também por fornecer modelos para mecanismos de analogização. A analogia, que na fase inicial dos estudos funcionalistas se encontrava relegada a plano secundário, passa a ganhar mais evidência e importância quando se destacam as relações metonímicas e associativas envolvidas nos usos linguísticos.

Em termos de vinculação de sentido e forma, a autora propõe o gradiente *encadeamento > idiomatismo / pré-fabricação > gramaticalização*. Assim posto, como Noël (2007), Bybee (2010) considera que, das esquematizações em geral, interessam ao Funcionalismo, na investigação da mudança gramatical, as mais convencionalizadas, que, via de regra, se constituem também como as mais entrincheiradas, as menos extensas e as mais recorrentes. Construções gramaticalizadas apresentam baixos níveis de composicionalidade (aferição do sentido a partir do sentido dos subcomponentes) e analisabilidade (identificação formal dos subcomponentes), o que lhes confere maior autonomia, generalização e frequência de uso.

Já em Traugott (2008) encontramos um tipo de abordagem que mais se coaduna com a perspectiva cognitivista. Ao definir a língua como sistema simbólico de pares de forma e sentido, inspirada principalmente por Croft (2001) e Croft e Cruse (2004), a autora parte de um conceito de construção distinto do de Bybee (2010). Para Traugott (2008), construções são consideradas como entidades teóricas e virtuais, projetadas para a captura de associações convencionais de forma e sentido. Com base nessa declaração, a autora defende que a análise funcionalista de vertente construcional precisa levar em conta e em conjunto as seis propriedades postuladas por Croft (2001) e Croft e Cruse (2004).

Assim posto, traços de natureza formal, como os atinentes aos níveis fonológico, morfológico e sintático, devem ser considerados em correlação direta com traços de sentido, desdobrados em semânticos, pragmáticos e discursivo-funcionais. Com base nessa declaração, as propriedades de forma deixam de ser vistas como consequência das de sentido, equilibrando, por assim dizer, a relação entre ambas as dimensões. Tal proposição tem implicações relevantes para o Funcionalismo, uma vez que a clássica derivação

*função* > *forma* cede lugar a uma abordagem que privilegia os traços de correspondência entre esses domínios, numa relação mais próxima à *função* = *forma*. O binômio *função* = *forma* deve orientar tanto a pesquisa funcionalista histórica, na detecção de trajetórias de gramaticalização, quanto a sincrônica, na identificação de graus de gramaticalidade.

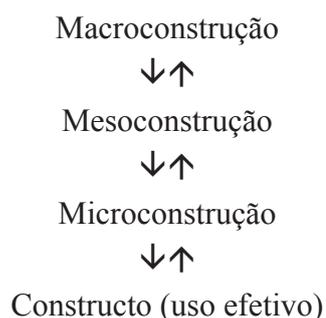
No destaque da correlação entre forma e sentido, Traugott (no prelo), tal como encontramos em Bybee (2010), enfatiza os mecanismos de analogização e reanálise como altamente relevantes para a fixação de padrões de uso gramatical. Conforme Traugott (no prelo), a analogização é concebida como um mecanismo de mudança que, partindo de determinado padrão fixo na língua, toma esse padrão para novos realinhamentos. Esses realinhamentos, que acabam por atribuir novos sentidos ou formas aos padrões já existentes, envolvem ainda o mecanismo de reanálise. Para a autora, seria mais adequado falar de *neoanálise* do que de *reanálise*, uma vez que o que ocorre, de fato, são novas interpretações ou associações contextuais, sem que, necessariamente, essas reelaborações sejam feitas com base numa pressuposta interpretação *padrão*, como algo estabelecido *a priori*. O destaque da analogização tem correspondência com proposta das representações exemplares, de Bybee (2010), e das relações de herança, de Goldberg (1995; 2006).

Seja como for, o fato é que, a partir do redimensionamento dos dois mecanismos – analogização e neoanálise – a investigação funcionalista não precisa buscar, compulsoriamente, a unidirecionalidade de todos os usos linguísticos, uma vez que grande parte desses usos tem na analogização e na neoanálise sua motivação maior. Dessa forma, via analogização e neoanálise, padrões de uso podem se tornar fixos e convencionais sem passar, necessariamente, pelas etapas de derivação de sentido e de forma já clássicas na pesquisa funcionalista, tais como, respectivamente, a trajetória metafórica proposta em Heine, Claudi e Hünnemeyer (1991), com o *cline* *pessoa* > *objeto* > *atividade* > *espaço* > *tempo* > *qualidade*, ou ainda a escala de Givón (1979) *discurso* > *sintaxe* > *morfologia* > *morfologia* > *zero*.

Conforme Traugott (2008), para que a pesquisa funcionalista adote a perspectiva construcional, é necessário rigor teórico-metodológico, uma vez que se está tratando da compatibilização de duas distintas vertentes de pesquisa linguística. Para tanto, a autora propõe quatro estágios de análise construcional, assim definidos:

- a) Macroconstrução: esquemas abstratos e virtuais;
- b) Mesoconstrução: conjunto de construções específicas, com função semelhante;
- c) Microconstrução: construção *type* individual;
- d) Constructo: *token* empiricamente comprovado, o *locus* da mudança gramatical.

Esses estágios, organizados hierarquicamente, são distribuídos e articulados de acordo com a seguinte disposição:



Conforme podemos observar, os quatro estágios são conectados por duas sequências de setas. As setas para baixo contemplam os mecanismos de analogização, uma vez que partem do esquema mais abstrato e virtual, que fornece o modelo para a fixação de novos alinhamentos; com base na macroconstrução e em seus desdobramentos mesoconstrucionais, ocorrem outros preenchimentos que vão, também, concorrer para a mudança linguística. As setas para cima se referem aos processos de construcionalização, em que a combinação de pares de sentido e forma, no nível do construto, via neoanálise, aumento de produtividade e esquematicidade, derivam em novos pareamentos de sentido e forma na língua, entendidos como convencionalizações inaugurais.

O olhar mais atento para os quatro estágios propostos por Traugott (2008) situa a pesquisa funcionalista no nível mais *baixo* dessa escala, correspondente ao construto. Referido como o de uso efetivo, o construto é o dado linguístico empiricamente constatado, a partir do qual são propostos ou formulados os demais níveis acima. Macro, meso e microconstruções são instâncias escalares e situadas em graus distintos de abstratização, sendo as macroconstruções as mais virtuais e esquemáticas. Assim, numa abordagem construcional, cabe ao Funcionalismo duas tarefas precípuas: a) captar relações entre pares de forma e de sentido, na tentativa de detecção de processos de construcionalização, ao longo da história dos usos linguísticos; b) investigar os mecanismos de analogização e reanálise, em perspectiva diacrônica e sincrônica, com base no levantamento de construtos e nos padrões micro, meso e macroconstrucionais.

Segundo Traugott (no prelo), mudanças se iniciam com pequenas inovações e em escalaridade, mas somente podem ser consideradas como tal no momento em que são convencionalizadas. A convencionalização passa por: a) aumento de esquematicidade e de associações paradigmáticas; b) incremento de produtividade e extensão para novos tipos (*host-class*); c) menor acessibilidade dos subcomponentes construcionais, em termos de sentido e de forma.

### **Padrões construcionais formados por pronomes locativos**

Da extensa pesquisa integrada sobre usos adverbiais do português, que tem ocupado nossa agenda de trabalho já há alguns anos no Grupo de Estudos *Discurso & Gramática*, conforme se constata em Oliveira e Cezario (2012), selecionamos, para esta seção, dois padrões construcionais de base verbal, formados por pronomes locativos – as macroconstruções locV e Vloc. Consideramos que esses dois esquemas virtuais se desdobram em construções específicas, a depender do tipo semântico do verbo e do pronome locativo. As mesoconstruções, por sua vez, se efetivam nos *types* individuais, nos exemplares específicos de cada subconjunto. Na pesquisa, examinamos os constructos, os *tokens* de uso, relacionando-os aos níveis superiores.

Assumimos como hipóteses que: a) ambos os padrões têm origem num tipo de construto de base verbal locativa, como parte de predicado oracional, que passou a ser reanalisado como expressão desvinculada do sujeito e de outro contexto oracional; b) a macroconstrução locV corresponde à retenção de um tipo de uso adverbial mais frequente na fase arcaica do português, em que era clara a tendência de posicionamento antes do verbo para os advérbios; c) a macroconstrução locV tem história mais recente na língua, reproduzindo o tipo de ordenação adverbial mais novo, referente ao posicionamento pós-

-verbal do advérbio; d) em termos de derivação semântico-funcional, ambos os padrões cumprem função conectiva em distintos níveis da língua – locV atua como operador, no nível sintático, enquanto Vloc funciona como marcador, no nível pragmático, estágio mais avançado de gramaticalização.

O primeiro padrão, locV, conforme se encontra em Rocha (2011) e Oliveira e Rocha (2011), se efetiva em torno de construtos exemplares como *daqui/daí vem* ou *aí/aqui está*. Trata-se de modelos de uso que servem de base para a fixação de outros, por analogização. Trazemos a seguir alguns dados de Rocha (2011), coletados do português contemporâneo em blogues da internet:

(1) *topico do mibr serve pra ajudar ser ajudado debater discutirpostar videos ensinar tutoriais etc... Po os Kra do site da mibr ficam o tempo todo fazendo peskisa ou tao trabaiaando la daí eles postam a noticia no site **daí vem** outro e copia? Pra q q ele vai faze intão...*

(2) *Para mim, muitos dos sintomas q atribuímos a menopausa tem outras fontes. E a menopausa (coitada) leva a culpa. Um dos principais sintomas e (O ninho vazio) e **daí vem** outras causas...*

Em (1) e (2), consideramos que se trata de usos nos quais *daí vem* ainda preserva traços de composicionalidade e analisabilidade. Nos dois fragmentos, podemos admitir que os termos *outro* e *outras causas*, respectivamente, atuam como sujeito da forma verbal *vem*, o que preservaria, em princípio, o estatuto mais lexical desta forma. A partir de tal interpretação, o sentido lógico e a função relacional caberiam primordialmente a *daí*. Contextos como esse, classificados como *atípicos* ou *críticos* por Diewald (2002), convidam a inferências, nos termos de Traugott e Dasher (2005), que permitem, de outra parte, leitura mais integrada de *daí vem*.

De acordo com Traugott e Dasher (2005), locutores lançam mão de termos mais referenciais, de natureza lexical, reelaborando-os para propósitos mais pragmáticos ou comunicativos. Nesse intento, convidam seus interlocutores a partilharem desses mesmos propósitos. Os usos mais integrados e convencionalizados de *daí vem* são típicos desses ambientes sugestivos. A maior integração semântico-sintática de *daí vem* se encontra, por exemplo, em fragmentos como o seguinte:

(3) *Há educadores tentando ensinar. **Daí vem** o interesse comercial e faz isso: “Eu rolo, tu rolas, ele \*\*\*\*\*”. Pode?*

Em (3), a possibilidade de atribuição de papel de sujeito a *o interesse comercial* é pequena, seja pela posposição em relação com elemento verbal, seja pela ausência de outros traços prototípicos dessa categoria, como animacidade, agentividade ou volição. Assim posto, *daí vem* é tomado, em termos de sentido e de forma, como um todo, que atua na conexão lógica entre as declarações anterior e posterior, na função de articulador de contraexpectativa.

Oliveira e Rocha (2011, p. 168) apresentam o seguinte levantamento, com a distribuição dos 48 usos de *daí vem* coletados a partir do *site* de busca Google, até a página sete dos *sites* sugeridos pelo buscador:

**Tabela 1: Usos de *daí vem* no português contemporâneo do Brasil**

Usos de <i>daí vem</i>	Número de dados
Expressão mais referencial	3
Articulador de contraexpectativa	13
Operador argumentativo	12
Sequenciador	3
Caso imbricado	5
Introdutor de tópico	12
Total	48

A Tabela 1 elenca seis funções para a expressão *daí vem*. Os tipos de uso mais frequentes, relativos aos papéis de articulador de contraexpectativa, operador argumentativo e introdutor de tópico, perfazem a quase totalidade dos dados e são também os mais convencionalizados. Registram-se ainda cinco dados considerados *imbricados*, ou seja, que se encontram em contextos críticos, nos quais prevalece a sobreposição funcional e a dificuldade de identificação efetiva do papel de *daí vem*.

A partir dos resultados obtidos na Tabela 1, Oliveira e Rocha (2011, p. 174) propõem o seguinte *cline* de integração semântico-sintática para *daqui vem*:

<i>menos gramatical</i>	$\longrightarrow$	<i>mais gramatical</i>
sequenciador + V >	sequenciador >	articulador de contraexpectativa operador argumentativo introdutor de tópico

No momento atual, a pesquisa da macroconstrução locV se desenvolve em nível de doutoramento. Para esta nova etapa, são elencadas as mesoconstruções específicas em que desdobra o esquema abstrato locV, a partir dos traços semântico-sintáticos do subcomponente verbal, e, na sequência, são levantados os *types* em uso no português contemporâneo, como *aí está*, *aqui concludo*, *lá vem*, entre outros.

A segunda macroconstrução de base verbal integrada por pronome locativo e por nós pesquisada é Vloc. Essa investigação, iniciada em Teixeira (2010) e hoje desenvolvida também como tese de doutoramento, parte do pressuposto de que o construto exemplar *vamos/vá lá* fornece a base analógica para padrões de uso como *quero lá*, *sei lá*, *vê lá*, entre outros.

Como destacado em Teixeira e Oliveira (2010, p. 74), defende-se o seguinte *cline* de gramaticalidade para *vamos lá*:

<i>menos gramatical</i>	$\longrightarrow$	<i>mais gramatical</i>
V + loc >	monitor textual >	monitor social

Os três fragmentos apresentados a seguir correspondem, respectivamente, aos três padrões de uso referidos na trajetória apresentada:

(4) *Por isso, não admito a indiferença. Eu jogo tranca com a mãe da diretora da loja. Ela sempre me diz: “Vamos lá que eu te apresento uma vendedora e você vai ser tratada feito rainha”. Mas é exatamente isso que me irrita: precisar que alguém saiba meu sobrenome para me tratar bem.*

- (5) I: sim mas ... o entrevistado sou eu ...  
E: é verdade ... **vamos lá** ... a parte do mar tá acabado?  
I: é ...  
E: você vai mexer agora só no céu?

(6) Já deu para entender o bastidor deste negócio de US\$ 700 bilhões, não? A questão complica na hora em que se define o papel de cada um. **Vamos lá**: quem decide quais papéis comprar? O Secretário. Quem decide quanto vai pagar? O Secretário. Quem escolhe os intermediários da compra? O Secretário. Para não deixar dúvidas, o texto reafirma que estes poderes podem ser exercidos “sem limitação.”

Em (4), coletado de reportagem da revista *Veja*, temos *vamos lá* como construto ainda no nível lexical, uma vez que *vamos* tem sentido de verbo pleno, referindo-se efetivamente a deslocamento físico, que se complementa por *lá*, representativo de espaço específico. Nesse fragmento, o locutor, de fato, convida o interlocutor a ir com ele a algum lugar. Identificamos nesse uso o estágio menos vinculado da expressão, que admitiria, inclusive, inserções de outros constituintes em seus subcomponentes, como, por exemplo *vamos agora lá* ou *vamos com certeza lá*.

Já em (5) trazemos um trecho de modalidade falada, retirado do *corpus* Discurso & Gramática da cidade do Natal (FURTADO DA CUNHA, 1998), que ilustra uso mais integrado e convencionalizado de *vamos lá*. Nesse trecho, os constituintes *vamos* e *lá* formam um todo de sentido e forma que atua no nível pragmático, uma vez que o locutor usa a expressão no convite ao interlocutor para que se pronuncie acerca de determinado tema; trata-se, conforme Traugott e Dasher (2005), da atuação da *inferência sugerida*, uma das motivações para a polissemia e a mudança linguística. O rótulo que elegemos, *monitor textual*, se justifica pelo tipo de funcionalidade desse uso, em que o locutor praticamente procura regular o turno do interlocutor, atuando no sentido de convidá-lo a se manifestar sobre determinado tema, num tipo de estratégia de intersubjetivização (TRAUGOTT; DASHER, 2005). O uso de *vamos lá* como monitor textual, por conta da especificidade funcional de que se reveste, é encontrado basicamente em textos da modalidade falada, em trechos de diálogo.

Por fim, em (6), ilustramos o estágio mais convencionalizado de articulação de *vamos lá* no português contemporâneo, levantado da revista *Veja on-line*. Na função de *monitor social*, a expressão é utilizada de modo mais abstratizado e rotinizado em relação a (5), uma vez que, em textos escritos, como (6), não temos a presença efetiva do interlocutor, portanto o convite se faz a um leitor hipotético, a uma entidade indiretamente referida. Assim usado, descolada em termos sintático-semântico do fragmento em que se insere, a expressão cumpre papel no nível pragmático-discursivo, concorrendo para a instauração do tom geral persuasivo de todo o fragmento. Em (6), a persuasão é instaurada, para além do uso de *vamos lá*, por intermédio da série de perguntas retóricas articuladas.

Teixeira (2010), com base no exame da correlação das propriedades de sentido e forma propostos por Croft (2001), identifica os seguintes traços para o uso mais convencionalizado de *vamos lá*, ilustrados aqui em (5) e (6), que foram nomeados genericamente como *marcadores de injunção*:

**Quadro 1: Propriedades do marcador de injunção *vamos lá***

	PROPRIDADES	MARCADOR DE INJUNÇÃO
FORMA	SINTÁTICAS	1) Independência de referência temporal: convencionalização da forma <i>vamos lá</i> ; 2) Posição restrita: início do enunciado; 3) Sujeito falante; 4) Verbo perde propriedade e não seleciona argumento; 5) Função clítica do locativo; 6) Marcado por pausa, codificado por vírgula, ponto, dois pontos.
	MORFOLÓGICAS	1) Não existe possibilidade de introdução de elementos entre os itens: amalgamento consolidado; 2) Cristalização na 1a.p.p. do presente do indicativo; não há contrações no <i>corpus</i> . * em pesquisas posteriores foi encontrada mudança morfológica como: <i>vamulá</i> , <i>vamula</i> , <i>vamlá</i> , <i>vamla</i> .
	FONOLÓGICAS	1) Há redução de material fônico; 2) Forma grupo de força V+LOC: em alguns contextos / <i>vamula</i> /; / <i>vamulá</i> /; <i>vamlá</i> /; / <i>vamla</i> /.
SENTIDO	SEMÂNTICAS	1) Perda do sentido original ligado ao <i>frame</i> em que não há movimento; 2) Sentido altamente intersubjetivo (codificação das perspectivas do falante apontando para o destinatário, convencionalizado na forma verbal: 1a. p.p. garante a intenção de compartilhamento de ideias e atitudes; 3) Há renovação de categorias já existentes: a de marcador discursivo; 4) Há nuance de sentido: deslocamento na intenção (expressa processos mentais); 5) É polissêmico; 6) Denota tendência de falantes recrutarem material para fins de fabricação de texto, ou seja, dando expressão simbólica para estratégias retóricas.
	PRAGMÁTICAS	1) Atuação de inferência sugerida: exortação = uso em contextos de conselho, pedido, sedução, desejo, bem como fatores intervenientes na interação: conduz o destinatário à percepção do ponto de vista e sua concordância; 2) Estratégia de produção do falante: 2.1) economia na produção = leva a reutilização de materiais antigos para novos meios (daí analogia) e rotinização, 2.2) maior explicitude: pressupõe falante tendo destinatário em mente; 3) Expansão de classe matriz: 3.1) locativo para clítico devido à ordenação de tal pronome, 3.2) favorecendo o recrutamento desse locativo para esse uso em razão de sua granularidade vasta.
	DISCURSIVO-FUNCIONAIS	1) Exclusividade de incidência em sequências injuntivas; 2) Comportamento totalmente abstrato e discursivo: exortação = nesse caso um dizer que devemos saber e fazer para alcançar um objetivo, o falante envolve o destinatário em sua opinião de forma que o faça aderir a ela.

Como podemos observar, em conformidade com Traugott (2008), Teixeira (2010) correlaciona propriedades formais e funcionais, na identificação e descrição do uso mais convencionalizado de *vamos lá*. Nesse levantamento, função e forma são tratadas como dimensões correspondentes, sem prevalência ou primazia de uma sobre a outra. Assim posto, numa perspectiva construcional, tal como a aqui defendida, a pesquisa funcionalista precisa dar conta de ambas as dimensões, numa abordagem mais holística dos usos linguísticos. Esse olhar mais amplo vai desde as propriedades fonológicas, no nível mais

elementar e específico, até as relativas ao ambiente discursivo-funcional, concernentes às sequências textuais e aos propósitos comunicativos.

### **Considerações finais**

O tratamento construcional hoje assumido pela pesquisa funcionalista abre uma vasta e instigante agenda de pesquisas. Trata-se de um campo de investigação promissor, que tem contribuído efetivamente para a descrição interpretativa dos padrões de uso linguístico. Uma das tarefas dos funcionalistas é, ao se debruçar sobre os dados, considerá-los representativos de *types*, como efetivações de modelos mais abstratos e virtuais.

Em conformidade com tal entendimento, cabe à pesquisa funcionalista, com base nos *tokens* levantados, identificar e estabelecer as micro, meso e macroconstruções que esses *tokens* realizam. Na pesquisa assim realizada, importa muito o rigor no estabelecimento da correlação função = forma, no esforço pela identificação das propriedades semânticas e estruturais que, combinadas, atuam nos padrões linguísticos em uso. Nesse viés, mecanismos de analogização e neanálise devem ser privilegiados.

Por fim, em perspectiva histórica, importa olhar para fases mais remotas da língua, na tentativa de se captarem dados atestadores de processos de construcionalização, com a criação de novos esquemas, e de processos de mudança construcional, com inovações dentre um dado esquema. Trata-se, nesse sentido, de vincular gramaticalização ao tratamento construcional de modo mais efetivo.

### **REFERÊNCIAS**

- BYBEE, J. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010. 252 p.
- \_\_\_\_\_. Mechanisms of change in grammaticalization: the role of frequency. In: JOSEPH, B. D.; JANDA, R. D. (Ed.). *The handbook of historical linguistics*. Oxford: Blackwell, 2003. p. 336-357.
- CROFT, W. *Radical Construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press, 2001. 448 p.
- CROFT, W.; CRUSE, D.A. *Cognitive Linguistics*. Cambridge: Cambridge University Press, 2004. 372 p.
- DELANCEY, S. Grammaticalization and linguistic theory. In: GOMEZ, J.; ROOD, D. (Ed.). *Proceeding of the 1993 Mid-America linguistics conference*. Boulder: Colorado, 1993. p. 1-22.
- DIEWALD, G. A model of relevant types of contexts in grammaticalization. In: WISCHER, I.; DIEWALD, G. (Ed.). *New reflections on grammaticalization*. Amsterdam/Philadelphia: John Benjamins, 2002. p. 103-120.

ERMAN, B.; WARREN, B. The idiom principle and the open choice principle. *Linguistic – an interdisciplinary journal of the language sciences*, Berlin/New York: Mouton de Gruyter, n. 2, p. 29-62, 2000.

FURTADO DA CUNHA, M. A. (Org.). *Corpus Discurso & Gramática – a língua falada e escrita na cidade do Natal*. Natal: Edufrn, 1998. 452 p.

HASPELMATH, M. On directionality in language change with particular reference to grammaticalization. In: FISCHER, O. et al. (Ed.). *Up an down the cline: the nature of grammaticalization*. Amsterdam: Benjamins, 2004. p. 17-44.

HEINE, B.; CLAUDI, U.; HÜNNEMEYER, F. *Grammaticalization: a conceptual framework*. Chicago: The University of Chicago Press, 1991. 328 p

HIMMELMANN, N. P. Lexicalization and grammaticalization: Opposite or orthogonal? In: BISANG, W. et al. (Ed.). *What makes grammaticalization?* Berlin: Mouton de Gruyter, 2004. p. 21-42.

KURYLOWICZ, J. The evolution of grammatical categories. *Esquisses linguistiques*, Munich, v. 2, p. 38-54, 1965.

GIVÓN, T. *Functionalism and Grammar*. Amsterdam; Philadelphia: John Benjamins, 1995. 503 p.

\_\_\_\_\_. *On Understanding Grammar*. New York: Academic Press, 1979. 379 p.

GOLDBERG, A. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press, 2006. 280 p.

\_\_\_\_\_. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press, 1995. 265 p.

MEILLET, A. L'évolution des formes grammaticales. In: CHAMPION, E. (Ed.). *Linguistique historique et linguistique générale*. Paris: Librairie Ancienne Honoré Champion, 1921 [1912]. p.130-148.

NOËL, D. Diachronic construction grammar and grammaticalization theory. *Functions of Language*, Amsterdam, n. 14, v. 2, p. 177-202, 2007.

OLIVEIRA, M. R.; CEZARIO, M. M. (Org.). *Adverbiais: aspectos gramaticais e pressões discursivas*. Niterói: EdUFF, 2012. 291 p.

OLIVEIRA, M. R.; ROCHA, R. A. As expressões “daqui vem” e “daí vem” como instanciações da construção Loc+SV no português contemporâneo. *Caligrama*, Belo Horizonte, v. 16, n. 2, p. 155-176, 2011.

ROCHA, R. *As construções “daqui vem” e “daí vem” no português do Brasil*. 2011. 98 p. Dissertação (Mestrado em Letras. Área de Concentração: Língua Portuguesa) – Instituto de Letras, Universidade Federal Fluminense, Niterói.

TEIXEIRA, A. C. M.; OLIVEIRA, M. R. Gramaticalização das construções “vá lá” e “vamos lá”. *Todas as Letras*, São Paulo, v. 16, p. 70-79, 2010.

TEIXEIRA, A. C. *Padrões de uso de “vá lá” e “vamos lá” na norma brasileira do português: microconstruções e gramaticalização*. 2010. 242 p. Dissertação (Mestrado em Letras. Área de Concentração: Língua Portuguesa) – Instituto de Letras, Universidade Federal Fluminense, Niterói, 2010.

TRAUGOTT, E. C. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In: ECKARDT, R.; JÄGER, G.; VEENSTRA, T. (Ed.) *Variation, Selection, Development - Probing the Evolutionary Model of Language Change*. Berlin/New York: Mouton de Gruyter, 2008. p. 219-250.

\_\_\_\_\_. “All that he endeavoured to prove was...”: On the emergence of grammatical constructions in dialogual and dialogic contexts. Forthcoming in KEMPSON, R.; COOPER, R. (Ed.). *Language in Flux: Variation, Change and Evolution*. Kings College London. (no prelo)

TRAUGOTT, E. C.; DASHER, R. *Regularity in semantic change*. Cambridge: Cambridge University Press, 2005. 341 p.

# As expressões *aí está* e *lá vai* como instâncias da construção LocV

(The *aí está* and *lá vai* expressions as instantiations of the LocV construction)

Rossana Alves Rocha

Universidade Federal Fluminense (UFF)

rossanalves@yahoo.com.br

**Abstract:** This paper investigates patterns of use of LocV construction in expressions such as “*aí está*” and “*lá vai*”. We intend to validate a possible joint which is more grammatical and conveyed by this constructional scheme. The research is qualitative, based on linguistic Functionalism, according to Traugott and Dasher (2005), Cunha et alli (2003), among others, from the study of grammaticalization and metaphorical and metonymic extension mechanisms as well as the cognitive approach of the construction grammar especially in Croft (2001), in the study of the symbolic structure construction.

**Keywords:** Functional Linguistics; grammaticalization; grammar construction.

**Resumo:** Este trabalho investiga os padrões de uso da construção LocV em expressões como “*aí está*” e “*lá vai*”. Tenciona-se validar possível articulação mais gramatical veiculada por esse esquema construcional. A pesquisa é de caráter qualitativo e fundamenta-se no funcionalismo linguístico, nos termos de Traugott e Dasher (2005), Cunha et alli (2003), entre outros, a partir do estudo da gramaticalização e do estudo dos mecanismos de extensão metafórica e metonímica como também na abordagem cognitiva da Gramática de Construção, mormente, nos termos de Croft (2001), no estudo da estrutura simbólica da construção.

**Palavras-chave:** linguística funcional; gramaticalização; gramática de construção.

## Introdução

Nosso ponto de partida é a constatação de que *aí está* e *lá vai* são expressões bastante produtivas nas interações dos falantes do português contemporâneo. Essas expressões são tratadas, neste trabalho, como instâncias da construção “locativo + verbo” (doravante LocV), um esquema abstrato que fica à disposição do falante/escritor para possível atualização.

Cumpramos ressaltar que o presente estudo tem sua origem na pesquisa intitulada *Pronomes locativos em construções nominais e verbais do português contemporâneo: ordenação, polissemia e gramaticalização* (OLIVEIRA, 2010), desenvolvida de modo mais específico pelo Grupo de Estudos *Discurso & Gramática* da UFF.

Partimos da hipótese de que construções verbais em torno de locativos, tais como as expressões *aí está* e *lá vai*, constituem unidade de sentido e forma que articula uma referência semântico-sintática e discursivo-pragmática distinta daquela advinda da soma de seus constituintes. Supomos, ainda, que pressões metonímicas atuam para a convencionalização de um tipo de unidade cumpridora de função gramatical, como elemento de conexão sintática ou textual. Consideramos a cristalização dos elementos nessas expressões fenômeno resultante de processo de gramaticalização.

Ressaltamos que a trajetória histórica das expressões estudadas ainda requer aprofundamento da análise e esse é o objetivo de pesquisa posterior que se pretende realizar. No entanto, o recorte sincrônico proposto por este artigo demonstra a mudança contínua de categorias e significados de *aí está* e *lá vai* sugerindo sua produtividade no português contemporâneo. Desse modo, pretendemos investigar padrões de uso do ponto de vista de modelos fluidos de uso linguístico de LocV, em expressões como *aí está* e *lá vai*, observando possível articulação mais gramatical veiculada por esse esquema construcional.

A análise, fundamentada no Funcionalismo Linguístico, nos termos de Bybee (2010), Heine e Kuteva (2007), Traugott e Dasher (2005), entre outros, aliada à linguística cognitiva, mormente em Croft (2001), leva em consideração fatores sintáticos, semânticos e discursivo-pragmáticos e entende que padrões fixos emergem da trajetória de gramaticalização dos elementos envolvidos a partir da frequência ou da regularidade com que são articulados na comunidade linguística.

Com base nessas premissas, tencionamos identificar as motivações para a mudança linguística, quais sejam cognitivas, subjetivas, discursivas e/ou pragmáticas, verificando os traços que diferenciam os distintos usos de *aí está* e *lá vai* como instanciações da construção LocV e, a partir de questões semânticas e formais, propor um *cline* de gradiência das expressões. Procuramos, também, a partir de ocorrências do banco de dados do *site Corpus* do Português, identificar os contextos preferenciais de uso dessas duas instâncias particulares.

## Funcionalismo Linguístico

A análise é fundamentada no funcionalismo linguístico norte-americano. Nesse sentido, consideramos, no estudo da língua, sua função de comunicar, a procedência do uso e a gradualidade da mudança linguística, principalmente, por meio do processo de gramaticalização. Segundo Noel (2006), esse processo é entendido por muitos teóricos como a criação de novas construções.

Nessa trajetória ascendente de gramaticalização, está envolvida alteração do estatuto categorial dos elementos, atribuindo funções gramaticais aos materiais lexicais, e, se já gramaticais, funções ainda mais gramaticais em contextos morfossintáticos e semântico-pragmáticos altamente restritos (TRAUGOTT, 2003, p. 645).

Na perspectiva do Funcionalismo Linguístico, a língua, conforme Bybee (2010), além de exibir aparente estrutura e regularidade de padrões, apresenta considerável variação em todos os níveis. A gradiência, assim, é essencial para o processamento linguístico e dá conta também de termos não gramaticalizados, sendo mais ampla que a gramaticalização.

A gramática, por sua vez, é definida por Bybee (2010, p. 8-9) como a formalização, o resultado da organização cognitiva de experiências com a língua. A autora considera a frequência um ponto alto no uso discursivo, pois, a partir da observação da repetição de padrões, evidencia-se a natureza da gramática baseada no uso.

Levamos em conta, ademais, mecanismos dominantes de mudança como a metáfora e a metonímia: esta relacionada às pressões estruturais, numa relação de superfície em que a contiguidade conceitual reflete associação e indexicalidade, atuando no nível sintagmático;

aquela, por sua vez, se refere à transferência de domínios, sendo a consequência, o resultado das relações metonímicas, atuando, assim, no nível paradigmático.

Traugott e Dasher (2005) referem-se a esses dois mecanismos, que operam na gramaticalização de construções linguísticas, e defendem que os aspectos metonímicos são mais básicos para o estudo da mudança linguística por destacarem a situação comunicativa. Os autores contrariam, dessa forma, o ponto de vista aceito por muitos linguistas, uma vez que redimensionam e valorizam o papel do contexto pragmático e das pressões cognitivas nos usos linguísticos.

Em conformidade com tal orientação, os autores (2005, p. 30) ressaltam o papel da subjetivação e a definem como um tipo de processo metonímico pelo qual *os emissores (falantes e escritores), ao longo do tempo, desenvolvem lexemas que codificam ou externalizam suas perspectivas ou atitudes elaboradas no ambiente comunicativo do evento de fala.* Através desse mecanismo, os falantes manifestam suas convicções e valores favorecendo a abstratização das expressões na medida em que sua criatividade licencia usos inovadores mais ligados a questões mais lógicas e, portanto, mais gramaticais.

Para estudar as motivações que levam a essa mudança, a abordagem funcionalista opera nas múltiplas instâncias do uso linguístico. Dessa forma, todas as ocorrências do objeto de estudo são consideradas. O uso e o impacto da experiência no sistema cognitivo, nesse entendimento, são tanto “ponto de chegada” quanto “ponto de partida” e, portanto, as construções linguísticas se desenham, sistematizam e se tornam disponíveis na comunidade linguística através dos processos de domínio geral (BYBEE, 2010).

Cada vez mais integrada aos pressupostos cognitivos, a abordagem funcionalista tem ampliado seus estudos. Bybee (2010, p. 6-8) pesquisa os processos subjacentes à estrutura da língua, que são aplicados a muitos domínios cognitivos. Tais processos, que a autora denomina de *domínio geral*, são:

- a categorização - ligação de identidade associada a representações estocadas;
- o encadeamento - união para a formação de unidades mais complexas;
- a memória enriquecida - estocagem mental de detalhes da experiência com a língua;
- a analogia - produção de enunciados baseada em outros enunciados;
- a associação transmodal - tendência de experiências co-ocorrentes serem cognitivamente associadas.

Assumindo que fatores de ordem pragmática e cognitiva interagem para a criação de *aí está* e *lá vai*, analisamos suas propriedades semântico-sintáticas e discursivo-pragmáticas, via modelo *top down*, quando o sistema linguístico é afetado descendentemente, promovendo assim o surgimento dessas expressões e, via *bottom up*, quando a fixação de seus usos afetam o sistema, ampliando sua representação. Em vista disso, nos fragmentos destacados na análise, observamos o grau de integração sintático-semântico, a função exercida em contextos específicos, a possibilidade de inserção de algum elemento e o tipo de referência feita.

## **Abordagem construcional**

A gramática de construções representa uma reação ao modelo componencial de organização da gramática. Esse novo tratamento foi motivado pelo fato de muitas estruturas gramaticais serem idiossincráticas com respeito às regras muito gerais de sintaxe, que são empregadas por esse modelo tradicional de estudo da língua. Nesse sentido, muitas expressões idiomáticas, como alguns dos usos analisados neste artigo, constituem um fenômeno problemático e, por isso, são um estímulo ao estudo e aplicabilidade dessa gramática que visa a dar conta de porções maiores de textos e seus processos cognitivos subjacentes.

Nessa perspectiva, levamos em consideração o aparato teórico da Gramática de Construção Radical (CROFT, 2001). Conforme Croft, a construção consiste em um pareamento de forma e sentido cujo significado não é dado somente pela soma dos sentidos dos membros da construção.

Nessa abordagem, Croft (2001, p. 18) destaca, nas construções, propriedades sintáticas, morfológicas e fonológicas, constituintes da forma, e propriedades semânticas, pragmáticas e discursivo-funcionais, referentes ao sentido convencional. Ao contrário do que determinam as análises componenciais, segundo o autor, o conjunto dessas propriedades é ligado internamente por um elo de correspondência simbólica (CROFT, 2001). Consideramos que essa conexão se deve ao entrincheiramento dessas duas faces, que passam a ter uma articulação sintático-semântica única convencionalizada em contextos pragmático-discursivos específicos.

Recentemente, as construções são vistas como um novo campo de pesquisa e são relacionadas por alguns funcionalistas aos estudos do processo de gramaticalização (NOEL, 2006). Tal processo vem atender a um desgaste natural da língua, pois a ritualização de um uso específico torna determinadas formas menos expressivas.

Por conta do tratamento das seis propriedades em conjunto, que tem como objetivo abranger as distintas dimensões envolvidas e suas ligações inter-relacionadas, Traugott (2008) defende que a Gramática de Construção Radical é a que melhor se ajusta ao estudo do processo multicamadas de gramaticalização que prevê uma série de mudanças correlatas. A autora procura dar aplicabilidade à teoria da interseção entre abordagem da gramática de construções e gramaticalização. Por tudo isso, os estudos dessa autora dão suporte ao estudo da construção LocV, alvo de análise deste artigo, e fornecem recursos para uma explicação mais consistente desse modelo virtual, cognitivo, instanciado por atualizações como *aí está* e *lá vai*.

### **Usos da expressão *aí está***

Nesta seção, levantamos alguns contextos em que *aí está* ocorre como uma instanciação da construção LocV.

#### ***Aí está* – Expressão menos entrincheirada**

Aos usos da expressão *aí está*, em que os elementos preservam sua autonomia sintático-semântica e discursivo-pragmática, denominamos expressão menos entrincheirada. Como se pode verificar, a posição dos elementos não determina um uso mais cristalizado.

No fragmento destacado em (1), podemos inverter a ordem de *aí* e *está* sem que se verifique alguma mudança expressiva de significado.

A sequência em que a expressão se encontra é predominantemente narrativa; no entanto, por as sequências não serem “puras”, *aí está* encontra-se em um contexto frasal que denota hipótese factual, confirmada pelo emprego dêitico de “agora”, “está” e “aí”. Pressupomos, portanto, que o fragmento permite leitura mais prototípica dos itens, sugerindo explicações, informações a respeito da situação comunicativa. Desse modo, o pronome locativo *aí* se refere a um espaço físico-concreto ou textual e o verbo *está* é usado no seu sentido referencial, com valor estativo ou copulativo.

- (1) Que diria uma cartomante se olhasse agora para as minhas mãos e para esta cidade que **aí está**? São tôdas umas putas mentirosas. A única que havia consultado há dois anos atrás, em Marabá, levara-lhe todo o dinheiro depois de dizer que ele seria muito rico e depois de praticar na cama de lona uma porção de sem-vergonhices. (<http://www.corpusdopotugues.com.br>)

Em (1), há um traço prototípico, anteposição do sujeito ao verbo, representado pelo pronome relativo *que*, o que motiva a leitura da expressão num uso mais concreto. O fragmento demonstra cabalmente uma referência à cidade, tendo como elemento de reforço o advérbio *aí*. Nessa articulação mais prototípica, a forma verbal *está* tem como argumentos o sujeito e o locativo, sendo este, nesse caso, necessário. Propomos que desse primeiro uso mais concreto de *aí está* são derivados outros mais abstratos, explanados a seguir.

### ***Aí está* – Expressão de projeção**

Como expressão de projeção, *aí está* migra para outro padrão de uso e passa a apresentar sentido menos concreto, mais construcional, com sentido próximo a uma relação lógico-semântica de projeção, pela qual um elemento projeta outro por meio de vários recursos léxico-gramaticais (HALLIDAY; MATTHIESSEN, 2004). Desse modo, a noção semântica da forma *está* e a foricidade de *aí* são modificados, num adiantado estágio de polissemia, para uma distinta atualização da construção LocV.

- (2) a Prefeitura prefere sacrificar o Festival impondo-lhe as precárias instalações do Teatro do Parque, inclusive para a abertura, porque não quer dividir os créditos. Olha, para as pessoas que têm dúvida quanto ao significado da palavra “provincianismo”, **aí está** um ótimo exemplo. (<http://www.corpusdopotugues.com.br>)

Considerando-se o curso da mudança linguística, em (2), os elementos *aí* e *está* tornam-se um todo composto de forma e significado (uso construcional), deixando de ser itens plenos e independentes. Portanto, se a ordem dos termos for invertida, ou se algum outro item for incluído, ocorrerá mudança no sentido da expressão. Há, nos termos de Croft (2001), um pareamento de forma e sentido, que constitui uma instanciação da construção LocV.

Nesse uso, *aí está* funciona como um articulador textual, ao estabelecer uma relação de sentido entre uma porção de texto anterior com uma porção subsequente. Posicionada entre as duas porções, o que remete à função prototípica do verbo de ligação, a expressão veicula a ideia de projeção ao apontar, indicar, uma ideia subsequente configurada por meio do SN abstrato “um ótimo exemplo”. A seleção de um verbo copulativo demonstra um resquício de ligação, o que nos remete ao princípio da persistência (HOPPER, 2003).

A sequência argumentativa corrobora para esse entendimento por se caracterizar pela progressão lógica de idéias. Trata-se de um contexto que comumente se constitui de elementos mais subjetivos ao se fazer a defesa de um ponto de vista, de uma ideia.

### ***Aí está* – Caso imbricado: expressão menos entrincheirada e expressão de projeção**

No fragmento (3) a seguir, apresentamos um dado de difícil análise e classificação de *aí está* por explicitar um contexto de hibridização. Esta ocorrência demonstra que as mudanças na língua são graduais, o que favorece a criação de ambientes com expressões ambíguas.

- (3) “ubi est caritas ubi est amor est Deus” qualquer coisa assim, “onde existe a caridade, onde existe o amor, **aí está** Deus”. O próprio fato dessa procura em latim talvez queira significar a dificuldade de cada um achar dentro de si essa caridade e esse amor ao próximo que são o endereço onde Deus mora e, aproveitando a abrangência do latim, em que “est” tanto significa # “está” como “existe”, são Deus. # Quando você achar a caridade, o amor ao próximo, você achará Deus, é o que se deduz, ou vice-versa. (<http://www.corpusdopotugues.com.br>)

Em (3), ilustrando a visão da gradiência exposta por Bybee (2010), temos, por conta de um contexto híbrido, uma etapa intermediária entre duas funções: a menos entrincheirada e a expressão de projeção. Temos, nesse caso, motivação para o processo de gramaticalização.

Podemos caracterizar a sequência como expositiva por apresentar explicações acerca do termo em latim, o que nos leva a uma leitura mais autônoma dos itens. Por sua vez, identificamos modalização feita pelo autor, numa estratégia de subjetivação, por meio de “talvez queira”, “quando você achar”, “é o que se deduz”, permitindo, também, uma leitura mais distante da prototípica.

O substantivo *Deus*, apesar de concreto numa visão formal, é semanticamente abstrato. Há, assim, o possível entendimento como expressão mais lexicalizada (referencial). O locativo *aí*, juntamente com o emprego de *onde*, igualmente de valor locativo, está retomando a localização virtual de Deus: “onde existe a caridade, onde existe o amor”, reforçando seu uso mais referencial. Por outro lado, a posposição do sujeito é um traço marcado, o que permite entendimento mais abstrato da expressão, com sentido próximo a uma relação lógico-semântica de projeção.

### ***Aí está* – Operador argumentativo**

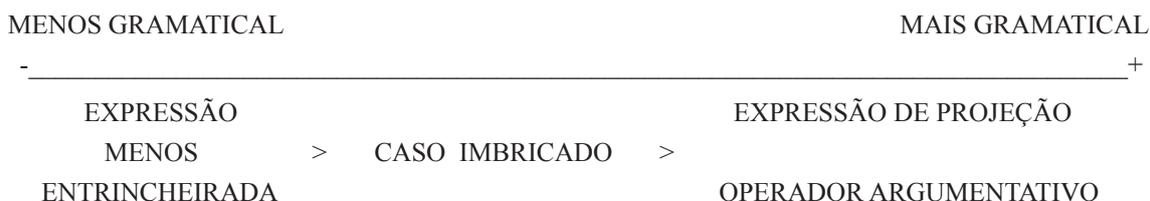
A mudança de sentido da expressão *aí está* para operador argumentativo, como no fragmento que segue, aponta para a crescente abstratização e pragmatização do significado da construção LocV. Há uma relação lógica de causa ou de consequência, pautada em elementos factuais. Tal uso indexicalizado encontra-se ancorado na situação comunicativa de argumentação e deriva dessas relações lógicas.

- (4) JC - Qual a diferença entre a sua gestão e a do ex-prefeito Jarbas Vasconcelos? ROBERTO MAGALHAES - No momento não há como se falar de diferenças, porque a tônica desses seis meses foi a continuidade. E **aí está** a quarta etapa do canal do Jordão concluída, faltando apenas as duas pistas de rodagens. **Aí está** o Hospital de Olhos, que não parou um dia sequer, e este ano ainda será inaugurado. (<http://www.corpusdopotugues.com.br>)

No fragmento acima, a expressão *aí está* estabelece noções semânticas de causa/consequência, reforçando um argumento relevante para a sustentação do ponto de vista

do falante. A posição da expressão a cada introdução de argumento promove progressão textual e ratifica a articulação dessas relações lógicas inerentes às sequências argumentativas. Tais sequências, por serem intrinsecamente subjetivas, representam um campo fértil para a gramaticalização, colaborando para a interpretação de *aí está* num outro padrão construcional. Aqui não apenas o dado, mas, sim, todo o ambiente é subjetivo, denotando pressão contextual típica do mecanismo de metonimização.

A partir dos usos da expressão *aí está*, propomos o *continuum* abaixo:



### Usos da expressão *lá vai*

Nesta seção, levantamos alguns contextos em que *lá vai* ocorre como uma instanciamento da construção LocV.

#### *Lá vai* – Expressão menos entrincheirada

No fragmento a seguir (5), *lá vai* possui sujeito anteposto, agentivo e de referência humana, o que configura uma função mais prototípica dessa expressão. O pronome locativo *lá* faz referência a um local determinado e o elemento *vai* indica deslocamento de um ponto a outro. Esses fatores contribuem para a leitura de *lá vai* como uma expressão menos cristalizada, mais referencial.

- (5) Bem feito! Mas o Miguel Reboredo, que peça! Felizmente para Viana ele só **lá vai** uma vez por semana. É uma vila grande, onde não há casa sem pára-raios e quando lá apanham o hospital cheio de doentes pobres botam-lhe fogo. (<http://www.corpusdoportugues.com.br>)

O fragmento evidencia, por conta do contexto metonímico pontuado por elementos que indicam localização, a independência do elemento *lá* como um pronome locativo, referindo-se ao local aonde vai “uma vez por semana”, e do elemento *vai* como um verbo de deslocamento, do lugar onde “ele” se encontra para a “vila grande”, seu destino. Os itens da expressão, gramaticalmente independentes, possuem os traços ditos estáveis dos dois elementos.

#### *Lá vai* – Expressão de apresentação

Nesse padrão construcional inovador, o grau de comprometimento do falante com a ação verbal é omitido. Aqui, a estratégia é tomar uma posição aparentemente neutra, atenuando as marcas da autoria da ação de deslocamento. Isso corrobora a mudança, pois, conforme Traugott e Dasher (2005), a mudança se dá do menos para o mais pessoal, por isso, num *continuum* de abstratização, esse uso de *lá vai* segue, possivelmente, a expressão mais referencial.

- (6) Ainda não decorrera metade do tempo marcado para o repouso do incansável espírito, recebi com a resposta este simples bilhete: “Não pude esperar. **Lá vai** a coisa. Se não servir, rasgue”. A coisa era esta extraordinária carta, cheia de mocidade e de fulgor. (<http://www.corpusdoportugues.com.br>)
- (7) Já tem data para voltar? Sua mãe não diz nada, mas vê-se que morre de saudades. **Lá vai** um abraço que bênção anda fora de moda. (<http://www.corpusdoportugues.com.br>)

Nos dois fragmentos (6) e (7) acima, a expressão *lá vai* posiciona-se no início do período, um local não prototípico, tornando a sequência menos concreta. Verificamos ainda que a autonomia do sintagma nominal posposto à expressão contribui para essa impessoalidade. Há uma estratégia intersubjetiva de desfocalização do agente.

O uso de *lá vai*, dessa forma, parece preservar a face do falante, que omite a ação de enviar, deslocar, como se “o abraço”, “a coisa”, por si sós, tivessem autonomia, agentividade. Essas inferências apoiadas nos emissores são pistas da mudança envolvendo subjetificação (TRAUGOTT; DASHER, 2005, p. 81).

### ***Lá vai* – Expressão de apresentação de tempo decorrido**

Como expressão de apresentação de tempo decorrido, entendemos o uso de *lá vai* no valor de verbo impessoal. Propriedades morfossintáticas, como a presença do advérbio de tempo *já* anteposto à expressão, e referências de tempo, como “um ano” e “quantos anos”, após *lá vai* concorrem para a sua leitura nesse valor construcional.

- (8) JORNAL DE NOTÍCIAS (JN) - Já **lá vai** um ano sem se conseguir eleger o Conselho de Arbitragem da FPF, uma situação que envolverá alguns perigos para o futebol português. (<http://www.corpusdoportugues.com.br>)
- (9) INF Era. Era dois. Era dois só. Dois. Eram dois moradores. INQ2 Isso já **lá vai** quantos anos, senhor, senhor Arquibaldo? INF Ah, isso já vai lá perto de duzentos anos. (<http://www.corpusdoportugues.com.br>)

Os dois fragmentos acima, como visto em Traugott (2008), mostram a importância do contexto morfossintático como fonte ou origem de gramaticalização. Assim, pressões metonímicas, como a presença do advérbio de tempo *já* antes da expressão *lá vai*, ilustram a força dessas pressões motivadoras da mudança linguística por gramaticalização.

### ***Lá vai* – Expressão de introdução intensificadora**

Configura-se, nos dois fragmentos a seguir, o uso da expressão *lá vai* como uma expressão de introdução evidenciando informação de valor quantitativo e/ou de intensidade. Por meio desses mecanismos, os falantes podem expressar sua opinião, sua posição diante de um determinado fato, ressaltar a importância desse fato para a compreensão de seu objetivo, transmitir seu estado emocional. Nesse contexto interacional de uso, o significado resulta de associações de ideias por meio da metáfora, um mecanismo inerente à faculdade humana de relacionar elementos análogos ou assemelhados.

- (10) Um homem desses acabar assim - que castigo! - dizia um. - É “cosa” feita! Foi inveja da “inteligência” dele! - dizia uma preta velha -. Gentes da nossa “cô” não pode “tê inteligência”! Chega logo os “marvado” e **lá vai** reza e “fêtiço”, “pa perdê” o homem - rematava a preta velha. (<http://www.corpusdoportugues.com.br>)

- (11) Quem sabe um dia, depois dele viúvo, a gente não arranje uma simpatia de Marianinha por ele? - Mas ele é muito mais velho do que ela. Ela só tem dezesseis anos e ele uns trinta e **lá vai** pedrada, meu amor. - E daí? (<http://www.corpusdoportugues.com.br>)

Nos fragmentos (10) e (11), a presença dos substantivos abstratos contáveis “reza” e “pedrada” e sua posposição à expressão, configurando um sujeito não agentivo, favorecem a abordagem construcional da expressão *lá vai* como uma expressão de introdução. Tal estratégia está diretamente ligada à intenção comunicativa do falante, à importância que ele concede a um determinado evento, a um sentimento, a um comentário ou à ideia que ele quer transmitir. A expressão, portanto, modificou sua natureza em direção a um sentido mais abstrato, por conta de sua relação com as formas adjacentes.

### ***Lá vai* – Expressão de introdução catafórica**

Como vimos em Bybee (2010, p. 5), há variação e gradiência em instâncias específicas de uma construção. Isso é comprovado pelo uso da expressão *lá vai* em outra funcionalidade, como a do fragmento a seguir:

- (12) Deve fazer umas quadrinhas novas... Porque não faz? - Fiz já. - Pode recitar? - Pois não. - Diga lá. - **Lá vai**: Ai, Filomena, Se eu fosse como tu, Punha uma máscara Na cara do Dudu. (<http://www.corpusdoportugues.com.br>)

No fragmento (12), a expressão teve seu uso estendido, sendo usada como um recurso para preparar o interlocutor para o que será anunciado. A presença de dois-pontos após a expressão e a sequência dialogal desempenham, em função da metonímia, um papel crucial para a leitura como expressão de preparação. Esse uso explicita o envolvimento de fatores cognitivos e pragmáticos na mudança pelo processo de gramaticalização.

Numa proposta de gradiência da expressão *lá vai*, temos, por hipótese, o *continuum* abaixo:

menos gramatical		mais gramatical
-		+
expressão menos entrincheirada	>	caso imbricado >
		expressão de apresentação expressão de introdução

### **Resultados preliminares**

De acordo com as análises feitas neste trabalho, as expressões *ai está* e *lá vai* confirmam, como visto em Bybee (2010), que a gradiência perpassa todos os níveis da língua. Tal variação possui relação com a competição de usos e com a unidirecionalidade, uma vez que nas duas expressões presumimos uma trajetória do mais para o menos concreto.

As expressões estudadas apresentam, como vimos, heterogeneidade, gradiência e variação, funcionando nas mais diversas situações e contextos. Há, assim, novos usos coexistindo com antigos, os quais podem ou não desaparecer (TRAUGOTT; DASHER, 2005). Esses processos dinâmicos são entendidos como resultados do processo de gramaticalização.

Nesse sentido, identificamos gradação do uso de *ai está* e *lá vai* com sentidos mais concretos, pela atuação como palavras autônomas, numa expressão menos entrincheirada,

para usos mais abstratizados, com as palavras em usos menos referenciais. Os usos das duas expressões são bastante diferentes em relação à função, refletindo a forte ligação entre linguagem, cognição, cultura e experiência.

Sabendo que a frequência das expressões *aí está* e *lá vai* determinam sua produtividade, verificamos que a expressão *lá vai* apresenta maior variação de sentido por ser mais frequente, menos marcada, evidenciando cinco usos. As experiências com a língua, o particular, como visto em Bybee (2010), afetam suas representações cognitivas, o geral.

Em conformidade com os pressupostos da linguística cognitiva e da linguística funcionalista defendemos que as funções das expressões estudadas desempenham um papel singular no processo de criação da gramática da língua portuguesa. A análise proposta necessita ainda se aprofundar nas observações e nos testes das trajetórias de abstratização. Trata-se de uma agenda de pesquisa promissora, que apenas se inicia.

## REFERÊNCIAS

BYBEE, J. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010.

CROFT, W. *Radical Construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press, 2001.

HALLIDAY, M.; MATTHIESSEN, C. *An introduction to functional grammar*. 3. ed. London: Arnold, 2004.

HEINE, B.; KUTEVA, T. *The genesis of grammar – a reconstruction*. Oxford: Oxford University Press, 2007.

HOPPER, P.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge: Cambridge University Press, 2003.

NOEL, D. *Diachronic construction grammar vs. grammaticalization theory*. 2006. Disponível em: <<http://hub.bku.hk/handle/123456789/38694>>. Acesso em: 24 jun. 2011.

OLIVEIRA, M. R. de. *Pronomes locativos em construções nominais e verbais do português contemporâneo: ordenação, polissemia e gramaticalização*. Rio de Janeiro: 2010. Projeto de pesquisa enviado ao CNPq.

OLIVEIRA, M. R.; TEIXEIRA, A. C. M. Gramaticalização das construções “vá lá” e “vamos lá”. *Todas as Letras*, v. 16, p. 70-79, 2010.

ROCHA, R. A. *As construções “daqui vem” e “daí vem” no português contemporâneo*. 98 p. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal Fluminense, Niterói, 2011.

TRAUGOTT, E. C. Constructions in grammaticalization. In: JOSEPH, B.; JANDA, R. (Ed.). *The handbook of historical linguistics*. Oxford: Blackwell Publishing, 2003. p. 624-647.

TRAUGOTT, E. C. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In: ECKARDT, R.; JÄGER, G.; VEENSTRA, T. (Ed.). *Variation, selection, development* –probing the evolutionary model of language change. Berlin/New York: Mouton de Gruyter, 2008. p. 219-250.

TRAUGOTT, E. C.; DASHER, R. *Regularity in semantic change*. Cambridge: Cambridge University Press, 2005.

# Paralelos estruturais entre a quantificação universal e as orações adverbiais em Karitiana

(Structural parallels between universal quantification and adverbial clauses in Karitiana)

Luciana R. Storto<sup>1</sup>

<sup>1</sup>Departamento de Linguística, Universidade de São Paulo

storto@usp.br

**Abstract:** This paper argues that universal quantification in Karitiana is obtained through the modification of the noun phrase by an adverbial clause. The structure of this clause includes a copular verb and a subordinator *tyym*, also used in temporal adverbial clauses. When the adverbial clause appears in sentence-initial position, it has, additionally, to include a pronoun to indicate co-reference with either a subject or a non-subject argument in the main clause. In this case, we consider that the adverbial clauses act as modifiers of the main clause.

**Keywords:** adverbial clauses; universal quantification; adverbial quantification; Karitiana; Tupi

**Resumo:** Este artigo argumenta que a quantificação universal em Karitiana é obtida via modificação de um sintagma nominal por uma oração adverbial. A estrutura desta oração inclui um verbo copular e o subordinador *tyym*, também usado em orações adverbiais temporais. Quando a oração adverbial ocorre no início da sentença, ela tem que incluir, adicionalmente, um pronome que indica correferência com o sujeito ou um não sujeito da oração principal. Neste caso, consideramos que as orações adverbiais atuam como modificadores de toda a oração principal.

**Palavras-chave:** orações adverbiais; quantificação universal; quantificação adverbial Karitiana, Tupi

## Objetivo

Fornecer evidências sintáticas e semânticas de que, em Karitiana (Tupi-Arikém, Rondônia, Brasil), sintagmas nominais traduzidos com leitura universal são, de fato, orações adverbiais formadas por um sintagma nominal, um verbo copular *aka* e o subordinador *tyym*.

- (01) [Taso aka-tyym]      ø-na-sokð̃ĩ-t      eremby  
homem cop-asp      3-decl-amarrar-nfut      rede  
‘[Todos os homens] amarraram a rede’  
Literalmente: [homens quando completos] amarraram a rede

## Evidências

A nova evidência a ser apresentada neste artigo a favor dessa hipótese é a de que é possível mover orações com leitura universal para o início da sentença da mesma maneira que é possível mover orações adverbiais temporais na língua (STORTO, 2007), e a correferência entre argumentos funciona da maneira idêntica em ambos os casos.

Além disso, argumentamos que como em Karitiana não há determinantes — artigos, demonstrativos — nem quantificadores no sintagma nominal (quantificação-D), a língua usa adjuntos adverbiais (quantificação-A) em forma de orações existenciais (nucleadas

por uma cópula) modificando o sintagma nominal no lugar de sintagmas nominais com quantificação universal.

### **Background: a ausência de determinantes e quantificadores no sintagma nominal**

Müller, Storto e Coutinho-Silva (2006) argumentaram que o Karitiana é uma língua sem artigos, demonstrativos, quantificadores nominais ou classificadores, na qual não há nenhuma marca de pluralidade no sintagma nominal. Neste artigo, nos referimos a línguas com esse perfil tipológico como línguas com nominais nus.

- (02) Maria Ø-naka-m-'a-t                      gooj  
 Maria 3-decl-caus-fazer-nfut      canoa  
 'Maria fez uma/a(s) canoa(s)'

Além disso, os autores sugerem que ambos os sintagmas demonstrativos e os quantificados universalmente são orações nucleadas por uma cópula nesta língua. A cópula *aka* está presente tanto nos sintagmas traduzidos como quantificação universal em (01), como em sentenças copulares (03) e orações demonstrativas (04):

- (03) Taso ø-na-aka-t                      i-se'a-t  
 Homem 3-decl-cop-nfut      participio-bom-conc.abs.cop.  
 'O homem é bonito'
- (04) Dibm ø-naka-tar-i                      [ony taso                      aka]  
 amanhã 3-decl-ir-fut                      [dêitico homem                      cop]  
 'Aquele homem irá amanhã'  
 Literalmente: [homem (que) está lá] irá amanhã

Duas dissertações de mestrado defendidas na Universidade de São Paulo forneceram evidências a favor da hipótese de que o Karitiana não tem determinantes nem quantificação-D: Sanchez-Mendes (2009) mostrou que todos os quantificadores são advérbios na língua (quantificação-A); Coutinho-Silva (2008) analisou as orações traduzidas com quantificação universal como relativas.

A diferença entre quantificadores-D e quantificadores-A é que os últimos quantificam sobre eventos:

- (05) ðonso                      ø-naka-ot-ø                      kandat                      ese  
 mulher                      3-decl-pegar-nfut                      muito                      água  
 'Uma/a(s) mulher(es) pegou/pegaram água muitas vezes'

Müller et al. (2006) e Sanchez-Mendes (2009) também observam que a distribuição de *kandat* é a de um advérbio. Storto (1999) mostrou que advérbios e sintagmas posposicionais (adjuntos) em declarativas SVO podem ocorrer no início e fim da sentença ou entre o verbo e o objeto, não entre S e V:

- (06)                      S                      V                      O  
 {kandat} ðonso ø-naka-ot-ø {kandat} ese {kandat}

Coutinho-Silva (2008) mostra que a quantificação com *akatyym* tem escopo local, quantificando sobre o sintagma nominal que precede a cópula, enquanto quantificadores adverbiais como *kandat* ‘muito’ funcionam como advérbios cuja quantificação tem escopo sobre toda a sentença (SANCHEZ-MENDES, 2009).

- (07) \*Taso    ø-na-sokõ’ĩ-t            aka-tyym            eremby  
 homem 3-decl-amarrar-nfut        cop-asp            rede

### A estrutura das orações com *akatyym*

Orações com *akatyym* devem ser precedidas pelo sintagma nominal que funciona como sujeito da cópula, nos casos abaixo *taso* ‘homem’ (em (08)) e *eremby* ‘rede’ em (09):

- (08) [Taso aka-tyym] ø-na-sokõ’ĩ-t            eremby  
 homem cop-asp 3-decl-amarrar-nfut        rede  
 ‘Todos os homens amarraram uma/a(s) rede(s)’
- (09) Taso    ø-na-sokõ’ĩ-t        [eremby        aka-tyym]  
 homem 3-decl-amarrar-nfut rede            cop-asp  
 ‘Um/o(s) homem(s) amarrou/-aram todas as redes’

Coutinho-Silva (2008) adota a proposta de Storto (1999) para a estrutura das orações subordinadas, na qual elas são VPs selecionadas por um núcleo aspectual (neste caso o subordinador aspectual *tyym*):

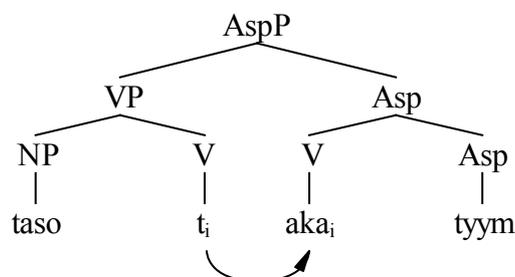


Figura 1: Estrutura Interna das subordinadas em Karitiana

Coutinho-Silva (2008) analisa as orações *akatyym* como relativas de grau (*degree relatives*), seguindo Grosu e Landman (1998). Essas relativas seriam responsáveis por uma leitura maximizadora dos indivíduos na posição de sujeito (ver STORTO; THOMAS, 2012, para maiores detalhes).

O conteúdo desses indivíduos é dado pelo sintagma nominal que precede *akatyym*, como em (01)-(02) acima ou pelo sintagma nominal coindexado com o pronome de terceira pessoa nas orações *akatyym* movidas para o início da sentença em (10)-(11):

- (10) oração “taakatyym” movida para o início da sentença  
 [Ta-aka-tyym] Ø-na-sokõ’ĩ-t            eremby taso  
 3anaf.suj-cop-asp 3-decl-amarrar-nfut        rede        homem  
 ‘Todos os homens amarraram a rede’

- (11) oração “iakatyym” movida para o início da sentença  
 [I-aka-tyym] Ø-na-sokō’ĩ-t eremby taso  
 3-cop-asp 3-decl-amarrar-nfut rede homem  
 ‘O homem amarrou todas as redes’

Coutinho-Silva (2008) considera que nestes exemplos *akatyym* moveu-se para a posição inicial da sentença e deixou para trás seu sujeito *taso* (10) ou objeto *eremby* (11). Os elementos de terceira pessoa na oração inicial da sentença (*ta* and *i*) não são pronomes para o autor, mas índices que têm a função de manter explícita a correferência - com um sujeito (*ta*) ou não sujeito (*i*).

Storto (2011) oferece argumentos contra a análise de Coutinho-Silva (2008), e sugere uma análise sintática para as orações de quantificação universal como orações adverbiais, que são introduzidas pelo mesmo núcleo subordinador *tyym* presente em outras orações adverbiais:

- (12) [A-yry tyym] yjxa Ø-na-oky-j  
 2-chegar asp nós 3-decl-matar-fut  
 ‘Quando você chegar, nós vamos matar (caça)’  
 Literalmente: [Quando sua chegada estiver completa] nós vamos matar

- (13) [Gok Maria amang tyym] Ø-na-oky-t him taso  
 mandioca Mari plantar asp 3-decl-matar-nfut caça homem  
 ‘Quando Maria plantou mandioca, o homem matou a caça’  
 Literalmente: [Quando o evento de Maria plantar mandioca estava completo], o homem matou a caça.

Um primeiro argumento apresentado por Storto (2011) contra a análise de que orações universais são relativas, proposta Coutinho-Silva (2009), é que as orações relativas *nunca* são introduzidas por subordinadores adverbiais em Karitiana (STORTO, 2012), enquanto orações adverbiais *devem invariavelmente* ocorrer com subordinadores como *tyym*.

- (14) [A-yry tyym] yjxa Ø-na-oky-j  
 2-chegar asp nós 3-decl-matar-fut  
 ‘Quando você chegar, nós vamos matar (caça)’  
 Literalmente: [Quando sua chegada estiver completa] nós vamos matar (caça)

Um segundo argumento contra a análise de relativas e a favor da análise de orações adverbiais (STORTO, 2011) é que a oração traduzida com quantificação universal pode ser movida para o início da sentença como são outras orações adverbiais. Quando isso ocorre, os pronomes *ta* ou *i* aparecem na oração e são correferenciais com o sujeito (*taso*) ou o não sujeito (*eremby*) da oração principal:

- (15) [Ta-aka-tyym] Ø-na-sokō’ĩ-t eremby taso  
 3anaf.suj-cop-asp 3-decl-amarrar-nfut rede homem  
 ‘Todos os homens amarraram a rede’
- (16) [I-aka-tyym] Ø-na-sokō’ĩ-t eremby taso  
 3-cop-asp 3-decl-amarrar-nfut rede homem  
 ‘O homem amarrou todas as redes’

## A correferência entre argumentos na oração principal e na subordinada adverbial

Note o mesmo padrão observado em sentenças com *taakatyym* e *iakatyym* nas orações adverbiais abaixo. Com o verbo intransitivo ‘ir’ na subordinada adverbial nucleada pela marca de aspecto *takit* ‘antes’ o pronome anafórico de terceira pessoa *ta* aparece quando o sujeito da subordinada é correferencial ao sujeito da oração principal. Caso contrário, aparece o pronome *i*, que se refere a uma terceira pessoa qualquer, desde que não seja o sujeito da sentença principal:

- (17) [Ta-tat takit] Ø-na-sombak-Ø João Porto Velho-ty  
 3.anaf.suj-ir antes 3-decl-ver-nfut João Porto Velho-obliquo  
 ‘Antes de ele<sub>i</sub> ir, João<sub>i</sub> visitou Porto Velho’  
 (ele = João, o sujeito da principal)
- (18) [I-tat takit] nasombak João Porto Velho-ty  
 ‘Antes de ele<sub>\*i</sub> ir, João<sub>i</sub> visitou Porto Velho’  
 (ele = outra pessoa, não o João)

Em todos os casos em que ocorre, *ta* é uma anáfora orientada para sujeito que permite correferência de longa distância; já *i* deve estar livre – sem correferência - no seu domínio local (na oração subordinada em que ocorre):

- (19) Taso Ø-na-oky-t ta-ota  
 homem 3-decl-machucar-nfut 3anaf.suj-amigo  
 ‘O homem<sub>i</sub> machucou seu<sub>i/\*j</sub> amigo’
- (20) Taso naokyt iota  
 ‘O homem<sub>i</sub> machucou seu<sub>\*i/j</sub> amigo’

Em (19) e (20) temos os mesmos pronomes de terceira pessoa usados como possuidores do nome *ota* ‘amigo’. Quando há correferência com o sujeito, *ta* é obrigatório.

Em (21), o verbo da oração principal é transitivo e o da subordinada adverbial é intransitivo, (ocorrendo com um objeto obliquo). Novamente, o pronome *ta* faz referência ao sujeito da oração principal:

- (21) Taso Ø-na-oky-t j̃onso [opok-ot ta-so’oot tykiri]  
 Homem 3-decl-machucar mulher inimigo-obl 3anaf.suj-ver quando  
 ‘O homem<sub>i</sub> machucou a mulher when ele/a<sub>i/\*j</sub> viu o inimigo’

Nos exemplos abaixo, temos pares de sentenças – (22)-(23), (24)-(25), (26)-(27) (28)-(29) – exemplificando o uso dos pronomes de terceira pessoa dentro de subordinadas adverbiais:

- (22) [Ta-tat takit] taso Ø-na-oky-t sojxa  
 3anaf.suj-ir antes homem 3-decl-matar-nfut porco  
 ‘Antes de ele<sub>i/\*j</sub> ir, o homem<sub>i</sub> matou o porco’  
 (ele = homem, não porco)

- (23) [I-tat takit] taso Ø-na-oky-t sojxa  
 3-ir antes homem 3-decl-matar-nfut porco  
 ‘Antes de ele<sub>j</sub> /<sub>\*i</sub> /<sub>k</sub> ir, o homem<sub>i</sub> matou o porco<sub>j</sub>’  
 (ele = porco, não homem; ou outra pessoa)
- (24) [Ta-hÿry] tykiri] Ø-na-terektereng-Ø i  
 3anaf.suj-cantar quando 3-decl-dançar-nfut ele  
 ‘Quando ele<sub>i</sub> cantou, ele<sub>i</sub> dançou’  
 (a mesma pessoa cantou e dançou)
- (25) [Ihÿry] tykiri] naterekteregng i  
 ‘Quando ele<sub>\*i</sub> cantou, ele<sub>i</sub> dançou’  
 (pessoas diferentes cantaram e dançaram)
- (26) Taso Ø-na-oky-t ãonso [opok-ot ta-so’oot tykiri]  
 homem 3-decl-matar-nfut mulher inimigo-obl 3anaf.suj-ver quando  
 ‘O homem<sub>i</sub> matou a mulher<sub>j</sub> quando ele<sub>i</sub> /<sub>\*j</sub> viu o inimigo’  
 (O mesmo homem viu o inimigo e matou a mulher)
- (27) Taso naokyt ãonso [opok-ot iso’oot tykiri]  
 ‘O homem<sub>i</sub> matou a mulher<sub>j</sub> quando ele/a<sub>\*i</sub> /<sub>j</sub> /<sub>k</sub> viu o inimigo’  
 (o homem matou a mulher e ela ou outra pessoa viu o inimigo)
- (28) [(João) Ta-mi tykiri] Ø-naka-hyryp-Ø ðwã  
 (João) 3anaf.suj-bater quando 3-decl-chorar-nfut criança  
 ‘Quando (João) bateu em ela<sub>i</sub>, a criança<sub>i</sub> chorou’  
 (a criança chorou quando ela mesma apanhou)
- (29) [(João) imi tykiri] nakahyryp ðwã  
 ‘Quando (João) bateu em ele/a<sub>\*i</sub>, a criança<sub>i</sub> chorou’  
 (a criança chorou quando outra pessoa apanhou)

A estrutura destas orações adverbiais em Karitiana é a seguinte: são orações subordinadas (AspPs) e podem adjungir-se à direita ou à esquerda do CP. Orações com leitura universal como *taakatyym* e *iakatyym* também podem aparecer depois da oração principal, assim como outras orações adverbiais:

- (30) ãonso Ø-na-amang-Ø [ony gok aka] [i-aka-tyym]  
 Mulher 3-decl-plantar-nfut dêitico mandioca cop 3-cop-asp  
 ‘A mulher plantou todas aquelas mandiocas’  
 Literalmente: A mulher plantou as mandiocas que estavam lá [quando estavam completas]

## Discussão

Storto e Thomas (2012) oferecem uma análise da proposta de Storto (2011) dentro da semântica formal. Eles assumem que há um pronome nulo na oração principal e que as orações [SN *akatyym*] dão uma interpretação a este pronome.

Esta análise unifica os usos de *tyym* em orações universais, orações adverbiais temporais, e seu uso como conjunção de VP. Ilustramos a ocorrência de *tyym* como conjunção entre VPs abaixo:

- (31) Yn Ø-naka'y-t asyryty, byyty **tyym**  
 Eu 3-decl-comer-nfut banana mamão também  
 'Eu comi banana e (comi) mamão também'
- (32) Otadnamyn-t yn Ø-naka-m->y-t yn gop yn  
 Quatro-obl eu 3-decl-caus-receber-nfut eu(tóp) vespa eu(tóp)  
 'Quanto a mim, eu recebi as vespas quatro vezes'
- (33) Gop sōwō-rã, gop miemo gop miemo sypom-p,  
 Vespa vermelha-qtfr vespa vespa dois-obl
- (34) Gop sowo-rã sypom-p **tyym**  
 Vespa vermelha-qtfr dois-obl também  
 'Vespas vermelhas, gop miemo, gop miemo duas vezes, vespas vermelhas duas vezes também'

No entanto, Storto e Thomas (2012) ainda tem um problema com as orações (*iakatyym* e *taakatyym*) movidas para o início da sentença, pois, em vez de um pronome nulo, há um sintagma nominal completo na oração principal (e pronomes ocorrem na oração adverbial).

Nossa proposta é analisarmos a oração formada pelo verbo copular e subordinador (*akatyym*) como um adjunto do sintagma nominal nos casos usuais em (8) e (9) - onde há uma unidade entre SN e *akatyym* - e como um adjunto da oração principal nos casos em que há movimento para o início da sentença de *taakatyym* ou *iakatyym* em (10) e (11). No primeiro caso, o sujeito da oração adjunto é um pronome nulo, e no último caso, os pronomes abertos *i* e *ta* estão presentes.

Podemos concluir que a quantificação em Karitiana é sempre quantificação-A, instanciada via advérbios (SANCHEZ-MENDES, 2009) ou orações adverbiais (STORTO, 2011). A oração adverbial é um adjunto da oração principal nos casos em que temos *taakatyym* ou *iakatyym*, por exemplo, em (10) ou (11), assim como são adjuntos os advérbios que funcionam como quantificadores (SANCHEZ-MENDES, 2009). Neste sentido, a quantificação universal é estruturalmente paralela a outros tipos de quantificação em Karitiana. A única diferença entre elas é que a quantificação universal tem a forma de uma oração adverbial e os outros tipos de quantificadores são advérbios.

Já quando uma oração adverbial aparece como adjunto do SN, ela funciona como modificador do nominal, de maneira semelhante a o que um adjetivo como “completo” faria em português, com a diferença de que a oração adverbial adjunto modifica o nominal ao introduzir um verbo existencial – a cópula - e o núcleo aspectual *tyym*, que denota completude. Uma paráfrase possível para este SN modificado pela oração adjunto seria “homem, quando (pro) está completo”.

## REFERÊNCIAS

- COUTINHO-SILVA, Thiago. *Aspectos dos sintagmas nominais em Karitiana: a quantificação universal*. 2008. Dissertação. (Mestrado em Linguística) - Universidade de São Paulo, São Paulo.
- GROSU, Alexander; LANDMAN, Fred. Strange Relatives of the Third Kind. *Natural Language Semantics*, Netherlands, v. 6, p. 125-170, 1998.
- MÜLLER, Ana; STORTO, Luciana; COUTINHO-SILVA, Thiago. Number and the count-mass distinction in Karitiana. In: FUJIMORI, Atsushi; SILVA, Maria Amélia Reis (Ed.) *University of British Columbia Working Papers in Linguistics*. v. 19. Proceedings of WSCLA XI. Vancouver: [s.n.], 2006. p. 122-135.
- SANCHEZ-MENDES, Luciana. *A quantificação adverbial em Karitiana*. 2009. Dissertação (Mestrado em Linguística) - Universidade de São Paulo, São Paulo.
- STORTO, Luciana. Subordination in Karitiana. *Amerindia*, Paris, v. 35, p. 219-237, 2012.
- \_\_\_\_\_. The clausal nature of demonstrative and universally quantified phrases in a Bare noun language. Paper presented in *Semantics of Under-represented Languages in the Americas 6*. Manchester, 2011. (ms)
- \_\_\_\_\_. Anáfora ta in Karitiana: um argumento de terceira pessoa não especificado para número. Trabalho apresentado na *ABRALIN V*. Belo Horizonte, 2007. (ms)
- \_\_\_\_\_. *Aspects of a Karitiana grammar*. 1999. Thesis (Ph.D) - MIT, Cambridge, 1999.
- STORTO, Luciana; THOMAS, Guillaume. The clausal nature of universally quantified phrases in Karitiana. In: BOGEL-ALLBITTEN, E. (Ed.) *Proceedings of the Sixth Conference on the Semantics of Under-represented Languages in the Americas and SULA-Bar*. Amherst: GLSA/University of Massachusetts at Amherst, 2012. p. 221-238.

# Uma análise teórico-empírica de predicação secundária

(An empirical-theoretical analysis of secondary predication)

Marcos Barbosa Carreira,<sup>1</sup> Andrea Knöpfle<sup>2</sup>

<sup>1,2</sup> Departamento de Linguística – Universidade Federal do Paraná (UFPR)

marcos.carreira@gmail.com, deaknoepfle@gmail.com

**Abstract:** This work compares the behavior of object oriented secondary predication constructions within Brazilian Portuguese and German, exemplified by the sentence João comeu a carne crua (João ate the meat raw x João ate the raw meat). As the English translation shows, this kind of data has a structural ambiguity regarding the attributive and predicative role of the adjective. We show that this ambiguity is not verified in German, a fact that might be relevant to a proper structural description (SD). Revising some standard analysis (STOWELL, 1983; ROTHSTEIN, 1995), we notice that they lead to a constituency paradox, already pointed out in Carreira (2008). We therefore propose a syntactic analysis based on Den Dikken's (2006) theory of predication, which might circumvent the alleged paradox.

**Keywords:** secondary predicate; small clauses; adjectives; structural ambiguity.

**Resumo:** O trabalho parte de uma análise comparativa entre o comportamento de predicados secundários orientados para o objeto entre português brasileiro e alemão, exemplificados na sentença João comeu a carne crua. A ambiguidade estrutural do dado se refere às possibilidades de leitura atributiva e predicativa do adjetivo. Mostramos que a ambiguidade estrutural apresentada na primeira língua não se verifica na segunda, fato de potencial relevância para a descrição estrutural (DE). Fazemos uma breve revisão bibliográfica (STOWELL, 1983; ROTHSTEIN, 1995), mostrando que as análises tradicionais levam a DE a um paradoxo de constituência, já colocado em Carreira (2008). Propomos uma possível análise para esse tipo de estrutura nos termos de Den Dikken (2006), que poderia dar conta do referido paradoxo.

**Palavras-chave:** predicado secundário; small clause; adjetivos; ambiguidade estrutural.

## Introdução

O objetivo deste trabalho é investigar e propor qual seria uma descrição estrutural adequada para dados de predicação secundária, translinguisticamente. Buscamos analisar em que medida algumas abordagens de predicação secundária apresentadas na literatura dentro da Gramática Gerativa Transformacional são descritivamente adequadas, considerando primeiramente como estudo de caso dados do português brasileiro (PB) e do alemão.<sup>1</sup> Assim, nosso trabalho parte de uma comparação entre o comportamento de predicados secundários orientados para o objeto nessas duas línguas.

Estamos chamando de 'Predicados Secundários' os constituintes que participam de uma predicação; nela, há a atribuição de uma propriedade a um nominal marcado tematicamente por um núcleo de um outro constituinte. Formalmente, tomemos a definição em Rothstein (2001) de predicado secundário, em que  $\alpha$  é um predicado secundário de  $\beta$

---

<sup>1</sup> Mais especificamente, referente ao quadro teórico, nos baseamos em Chomsky (1965, 1981, 1995 *inter alia*). Para dados e análise de predicação secundária, nossa base bibliográfica está em Stowell (1981, 1983), Rothstein (1983, 1995, 2001) e Den Dikken (2006).

sse  $\alpha$  é predicado de B, e  $\alpha$  e B se c-comandam e B é theta-marcado por um núcleo não contido em  $\alpha^2$ . Vejamos um exemplo em PB e seu correspondente em alemão.

(01) Maria bebe seu chá **frio**.

(02) Maria trinkt ihren Tee **kalt**.

A leitura relevante para o dado em (01) é a de que *O chá estava frio quando Maria o bebeu*, e não *Entre os chás disponíveis, Maria bebeu somente aquele que era frio*. O dado em (02) no alemão não permite a segunda interpretação, fato que descreveremos na segunda seção.

Ao investigar e propor qual seria uma descrição estrutural adequada para dados de predicação secundária orientados para objeto, tínhamos também em mente uma questão morfológica, observada na comparação dos dados entre PB e alemão. Tal questão, no entanto, não se mostrou relevante sintaticamente, com relação aos objetivos desse trabalho. Em (01), notamos que o adjetivo flexiona necessariamente em masculino e singular; já o adjetivo em alemão não apresenta quaisquer marcas de flexão.

O trabalho está assim dividido: na próxima seção, descrevemos e comparamos o comportamento dos adjetivos em PB e alemão, mostrando a ambiguidade estrutural em PB e sua não verificação em alemão. Na terceira seção, colocamos a questão morfológica. Em seguida, na quarta seção, trazemos algumas análises canônicas, suas vantagens e o paradoxo de constituência. Na quinta seção apresentamos a teoria de predicação secundária em Den Dikken (2006), e como ela poderia fornecer uma descrição estrutural (DE) que não gerasse o paradoxo. Concluimos com algumas considerações gerais e questões em aberto.

## Descrição dos dados

Nesta seção, vamos descrever e comparar o comportamento dos dados de predicação secundária em PB e alemão. Começando pelo PB, é bem conhecida na literatura a ambiguidade estrutural apresentada por determinados dados; ambiguidade essa gerada por duas descrições estruturais diferentes. Tomemos um exemplo canônico.

(03) João comeu a carne crua.

(04) *Foi a carne crua que João comeu.*

(05) *Foi crua que João comeu a carne.*

O dado em (03), na verdade, representa duas sentenças diferentes, geradas cada uma por uma DE distinta. Em uma delas, representada pela paráfrase em (04), [a carne] forma um DP com [crua]. Em outra DE, representada pela paráfrase em (05), [a carne] não forma um DP com crua.

Já em alemão, essa ambiguidade do PB não se coloca, pois cada uma das *strings de terminais* (no sentido de CHOMSKY, 1957) geradas pelas descrições estruturais apresenta marcas morfológicas e ordem diferente entre si. Vejamos as sentenças:

---

2 Não estamos dizendo que vamos assumir essa configuração para o predicado secundário. Vamos discutir melhor a questão na quarta seção.

- (06) João ass das Fleisch roh.  
 João comeu a carne cru  
 ‘João comeu a carne crua’ *no sentido de que*: “Foi crua que João comeu a carne”.
- (07) João ass das rohe Fleisch.  
 João comeu a crua carne  
 ‘João comeu a carne crua’ *no sentido de que*: “Foi a carne crua que João comeu”.

Em (06), [a carne] não forma um DP com [cru]; em (07), [a carne] forma um DP com [crua]. Crucialmente, notamos que, em (06), o adjetivo está posposto ao nome e não concorda com ele, diferentemente de (07), em que o adjetivo se antepõe ao nome e flexiona em gênero, número e caso de acordo com esse nome.

A gramática tradicional do alemão descreve os adjetivos como adjetivos em uso predicativo e adjetivos em uso atributivo. No primeiro caso, o adjetivo nunca é anteposto ao nome e não pode flexionar. No segundo caso, o adjetivo sempre é anteposto ao nome e concorda com ele – é o que estamos chamando de *adjetivo não formar um DP com o N* e *adjetivo formar um DP com N*, respectivamente.

Vamos colocar mais alguns dados do alemão para exemplificar melhor a diferença entre uso atributivo e predicativo, lembrando que as características colocadas para o adjetivo em uso predicativo em alemão não são restritas a predicados orientados para o objeto. Nos dados que seguem, destacamos as marcas morfológicas dos adjetivos, com as respectivas glossas.

- (08) Der Mann sang das Lied fröhlich.  
 O homem cantou a música feliz  
 ‘O homem cantou a música feliz’ *no sentido de que*: “O homem estava feliz quando cantou a música”.
- (09) Der Mann sang das fröhliche Lied.  
 O homem cantou a feliz<sub>neutro-sing-Acc</sub> música  
 ‘O homem cantou a música feliz’ *no sentido de que*: “Foi a música feliz que o homem cantou”.
- (10) Ich fand den Mann besoffen.  
 Eu encontrei o homem bêbado  
 ‘Eu encontrei o homem bêbado’ *nos sentidos de que*: “O homem estava bêbado quando eu o encontrei”./ “Eu estava bêbado quando encontrei o homem”.
- (11) Ich fand den besoffenen Mann.  
 Eu encontrei o bêbado<sub>masc-sing-Acc</sub> homem.  
 ‘Eu encontrei o homem bêbado’. *no sentido de que*: “Foi o homem bêbado que eu encontrei”.
- (12) Ein besoffener Mann ist gekommen.  
 Um bêbado<sub>masc-sing-Nom</sub> homem é chegado  
 ‘Um homem bêbado chegou.’ *no sentido de que*: “Foi um homem bêbado que chegou”.

Relevante para nosso trabalho é ter clara a assunção de que o adjetivo predicativo não forma um DP com o nome de quem é predicado, e que tal adjetivo não apresenta quaisquer marcas de flexão em alemão (diferentemente do PB).<sup>3</sup> Essa característica do

3 Vamos discutir a questão da constituência na quarta seção. Apesar dessa diferença importante, a análise aqui proposta não tem por objetivo explicá-la, pelo menos não nesta fase da pesquisa.

adjetivo predicativo no alemão é fato conhecido na literatura, e Kratzer (2005) nota que o adjetivo de resultativas adjetivais também é do tipo predicativo, a quem a autora chama de *bare* (nu), justamente por nunca apresentar quaisquer marcas de flexão.

(13) Er trank die Teekannen leer.  
Ele bebeu as chaleiras vazio<sub>bare</sub>  
'Ele esvaziou as chaleiras, bebendo todo o seu conteúdo.

(14) \*Er trank die Teekannen leeren.  
Ele bebeu as chaleiras vazias<sub>plur-Acc</sub>  
*Sentido pretendido:* 'Ele esvaziou as chaleiras, bebendo todo o seu conteúdo.

O contraste acima mostra que o adjetivo da resultativa adjetival precisa ser *bare* (13), do contrário a construção é mal formada (14). A autora não investiga mais a fundo o sentido de *bare*, mas observamos em Knöpfle (2011b) que *bare* seria a ausência de quaisquer marcas de flexão de caso, gênero e número no adjetivo das resultativas adjetivais, assim como os demais adjetivos predicativos do alemão – o que nos leva à questão morfológica, assunto da próxima seção.

### A questão morfológica

Kratzer (2005) especula que a variação paramétrica entre a ocorrência de resultativas (em línguas como alemão e inglês) ou a não ocorrência (em línguas românicas) estaria condicionada à morfologia. Pudemos observar, da sentença em (13), acima, que a tradução para o PB não segue a mesma estrutura que o alemão.<sup>4</sup> Segundo a autora, os adjetivos em alemão têm a propriedade de poderem apresentar a forma *bare*. Já tal propriedade não estaria disponível nas línguas românicas. Nessas, os adjetivos seriam flexionados desde o início da derivação sintática (não havendo estágio em que apareçam sem flexão).

Parece razoável afirmar que os adjetivos em PB estão sempre flexionados – ou, pelo menos, apresentam diferenças consideráveis em relação ao comportamento flexional dos adjetivos em alemão. Camara Jr. (1970) hipotetiza que os adjetivos em PB são flexionados para feminino com o sufixo flexional ou desinência '–a', e para o plural com sufixo flexional ou desinência '–s'.

O masculino e o singular são caracterizados pela ausência de marcação feminina e plural. Na teoria adotada por esse autor, isso quer dizer que existe um morfema zero, marcador de masculino, e outro, marcador de singular. Se essa análise estiver correta, então, mesmo sem marcação evidente de flexão, o adjetivo está flexionado, pois possui os morfemas zero marcadores de masculino e singular. Existindo sempre minimamente morfemas zero desse tipo, então de fato a flexão está sempre presente nos adjetivos do PB. Observamos, em PB, que o adjetivo precisa necessariamente flexionar em gênero e número com o NP, independentemente de o adjetivo formar ou não um DP com o NP.

(15) \*João comeu a carne cru.

4 Não vamos entrar aqui na discussão sobre o estatuto das resultativas em PB. Para tanto, remetemos o leitor a Foltran (1999), Lobato (2004), Barbosa (2008) e Knöpfle (2011a).

Assim, a primeira pergunta que colocamos é: [cru] é masculino ou bare? Uma segunda questão que levantamos é a seguinte: por que a concordância é obrigatória do PB e inexistente no alemão? Nesse sentido, trata-se de uma questão paramétrica? Não perseguiremos essas duas perguntas aqui, por questões de escopo e objetivo do trabalho. A terceira pergunta - que devemos ter em mente ao propor uma DE para dados de predicação secundária como os que estamos analisando – é se a concordância é relevante para definir a estrutura sintática desses predicados.

De acordo com o que vimos acima, é fato que, em PB, o adjetivo predicativo flexiona necessariamente. Também é fato que, em alemão, o adjetivo predicativo é *bare*, i.e. não pode apresentar quaisquer marcas de flexão de gênero/número/caso. Além, disso, como veremos, o alemão e o PB parecem ter a mesma DE para dados de predicação secundária. Portanto, o que nos parece, pelo menos à primeira vista, é que questões relacionadas à concordância/flexão de gênero/número/caso não são relevantes para definir constituição – pelo menos não para adjetivos enquanto predicados secundários. No entanto, talvez elas sejam sim relevantes para definir a posição do predicado secundário na DE, ou mesmo sua estrutura interna.<sup>5</sup>

### Análises canônicas

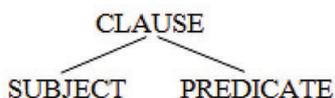
Nós encontramos na literatura sobre esses dados pelos menos duas teorias principais, que se tornaram mais conhecidas. Vamos chamá-las de análises canônicas.

A primeira delas, a Teoria das *Small Clauses*, afirma, conforme Stowell (1995), que as SC são estruturas de predicação em que um sujeito e um predicado estão em uma relação de irmandade, e entre esse sujeito e esse predicado há uma relação de predicação que é a mesma que há entre o sujeito de uma oração plena, exceto pelo fato de não haver nas SC um verbo auxiliar e nem uma morfologia de tempo e aspecto, apesar de o adjetivo da SC poder apresentar alguma marca de concordância. Assim, de acordo com Stowell,

[...] a teoria das SCs é baseada na convicção de que essa relação semântica [a predicação] é refletida uniformemente na estrutura de constituinte, no sentido de que a relação sujeito/predicado é sempre codificada em termos de um par de constituintes irmãos. (1995, p. 272)<sup>6</sup>

Veja a representação do autor repetida abaixo:

(16)



<sup>5</sup> Uma vez que a concordância não se mostra relevante para definir constituição, não esperamos que nossa análise sintática com base em Den Dikken (2006) se proponha a dar conta da variação quanto à morfologia (flexão gênero, número e caso) dos adjetivos predicativos entre PB e alemão. Agradecemos a observação do revisor anônimo da Revista Estudos Linguísticos por nos questionar sobre não termos nada a dizer a respeito dessa variação. Essa é realmente uma questão curiosa, para a qual ainda não temos uma resposta; também por isso nos restringimos a analisar a proposta de Den Dikken (2006) com relação à constituição. No entanto, insistimos em registrar a diferença morfológica, pois acreditamos que essa é uma questão pendente e precisa ser colocada e atacada em algum momento.

<sup>6</sup> Tradução nossa.

Dessa forma, nosso AP predicado secundário é tratado por essa teoria como um caso de SC-adjunta, formada a partir de AP. Trata-se de uma predicação que tem uma estrutura semelhante à apresentada acima, em (16). Com isso, esse AP terá em seu *Spec* um PRO (se assumirmos, como Stowell, que A' é o nível (predicado) relevante, levando em conta a forma da predicação proposta em (16)). Esse PRO é coindexado com o DP que é complemento do verbo e, com isso, a concordância teria também alguma análise possível, que deixamos de lado nesse trabalho.

A estrutura abaixo, na Figura (1), em que a SC é adjunta de VP, seria uma DE possível para nossos dados, considerando a proposta de Stowell (1995):

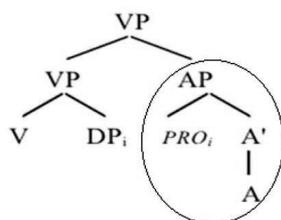


Figura 1. Small Clause adjunta a VP

A outra teoria, que chamamos aqui de Teoria de Predicação Secundária (ROTHSTEIN, 1983, 1995, 2001), trata esses predicados como adjuntos, mas sem a formação de uma SC-adjunta. Nesse caso, a predicação se dá via algo como m-comando e o AP predica de uma DP que já foi *theta*-marcado por um outro núcleo. Essa teoria não postula, portanto, um PRO na posição de *Spec* do AP, já que a predicação se dá de outra forma. Sendo assim, somente o AP é adjungido a alguma categoria como VP ou IP. A estrutura na figura 2 seria outra DE possível para nossos dados:

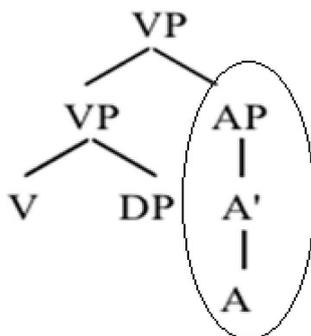


Figura 2. Predicado Secundário

O que é importante para nosso trabalho é a propriedade compartilhada por ambas as teorias que diz que o AP e o DP não formam juntos um constituinte. Assim, as análises canônicas parecem fazer previsões corretas a respeito da distribuição dessas categorias, com relação a testes de constituência como: pronominalização, clivagem etc.

As análises canônicas fazem a seguinte previsão a respeito do dado abaixo: uma vez que o DP e o AP não formam um constituinte, a sequência ‘o bife cru’ grifada em (17) não pode ser movida, ou estar entre ‘ser’ e ‘que’ em uma clivada (cf. 18):

(17) A Duda comeu o bife cru

(18) \*Foi o bife cru que a Duda comeu \_\_

A sequência em destaque (grifada) em (17) não pode ser fronteada (cf. 18), uma vez que não forma um constituinte (considerando a leitura relevante, a predicativa). Essa é uma previsão compatível com ambas as teorias discutidas. Observe a árvore abaixo, na Figura (3):

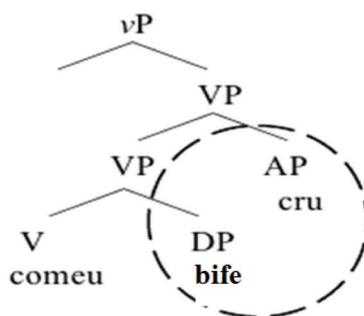


Figura 3. Predicado Secundário e o DP não formam um constituinte movível

A figura (3) demonstra essa previsão, pois nota-se que o DP e o AP não formam um constituinte sozinhos; entretanto, como veremos na próxima seção há pelo menos um dado que a desafia.

### Paradoxo de constituência

Na seção anterior, vimos que o DP mais o predicado secundário, isto é, a sequência ‘o bife cru’, não pode ser fronteada e a impossibilidade desse movimento está de acordo com as previsões que tanto a Teoria das Small Clauses quanto a Teoria de Predicação Secundária fazem. No entanto, o dado abaixo nos mostra um contexto em que esse deslocamento é possível:

(19) Esse bife mal passado eu até comeria \_\_, mas ele cru eu não como \_\_ de jeito nenhum.

Com isso chegamos a um Paradoxo de Constituição, que já foi descrito em Carreira (2008), conforme segue abaixo:<sup>7</sup>

Um fato interessante chama a atenção. Os exemplos de SC adjuntas com verbos do tipo de ‘contratar’ não pareciam permitir fronteamento e isso também era o esperado, pois as propostas de estrutura para esse verbo tem sido algo como [...] em Stowell (1981) com as SC adjuntas, ou [...] como a adotada por Rothstein (1983) e Foltran (1999) para os Predicados Secundários. Em ambos os casos, não há a previsão desse fronteamento, porque essas predicções NP+XP não seriam um constituinte formado de DP e XP, portanto ambos não seriam movíveis. (CARREIRA, 2008, p. 80)

A caracterização dada por Carreira (2008) é bem mais completa e complexa, uma vez que o paradoxo surge na verdade do cruzamento de resultados de diferentes testes, além de a classe do verbo ‘comer’ ter sido comparada a outras quatro classes de verbos. O que o autor constatou foi que o resultado de um teste, como teste de pronominalização, não nos permite pronominalizar as predicções – o que faria a previsão de que não há um

<sup>7</sup> Assuma, ao ler a citação, que o verbo ‘contratar’ pertence à mesma classe de ‘comer’; e entenda XP como uma variável para categoria, que no nosso caso é um AP.

constituente – veja os exemplos (20) e (21) abaixo – mas o resultado do teste de fronteamto apresentado em (19) faz a previsão de que há sim um constituinte.

- (20) a. A Duda viu a Sheron Stone grávida  
b. A Duda viu algo (pronominalização possível)
- (21) a. O Felipe comeu a carne crua  
b. #O Felipe comeu algo (não pronominalizável)

Por algum motivo, com o verbo ‘ver’, o pronome ‘algo’ pode pronominalizar (estar no lugar de, fazer referência a mesma coisa que) à sequência ‘a Sheron Stone grávida’; porém, com o verbo ‘comer’, esse tipo de pronominalização não é possível.

Mesmo assim, podemos caracterizar esse paradoxo de uma forma mais simples que é: as teorias apresentadas nesse artigo fazem a previsão de que esse movimento (movimento da expressão predicativa ‘esse bife mal passado’) não exista, mas o dado em (19) faz a previsão de que ele exista, ou que algo a mais esteja em jogo.

Por isso, se esse deslocamento é possível no contexto sintático adequado, qualquer teoria sintática precisa fazer essa previsão. A busca de uma solução para o paradoxo passou por várias possibilidades que foram apresentadas em Carreira (2008). Entre elas estavam a possibilidade de fronteamto do VP (via *remnant movement*), movimento lateral e formação de um novo constituinte (à la TAKANO, 2002)<sup>8</sup> e ainda vários movimentos (possibilidade não tratada em Carreira 2008). Para esse artigo, no entanto, vamos considerar apenas a possibilidade de se pensar numa estrutura diferente na base, usando a teoria de Den Dikken (2006).

### Uma proposta: Den Dikken (2006)

Vimos na seção anterior que as teorias canônicas (STOWELL, 1981; 1983; 1995 *inter alia*; e ROTHSTEIN, 1983; 1995; 2001) levam-nos a um paradoxo de constituência. Nessa seção vamos apresentar uma possibilidade bastante promissora e interessante principalmente por propor uma acomodação da predicação na gramática e pelo esforço de levar sua tese às últimas consequências, de tal forma a promover um olhar bastante mais fino sobre a estrutura da gramática e seu funcionamento. Trata-se da Teoria dos *Relators* e *Linkers* de Den Dikken (2006).

Den Dikken propõe que todas as relações de predicação<sup>9</sup> sejam mediadas por um núcleo funcional abstrato (o RELATOR), responsável por estabelecer essa relação entre o predicado e seu sujeito na estrutura sintática.

O *RELATOR* – núcleo de RP (relator phrase) – é um *placeholder*, que acomoda em sua projeção o *sujeito* e o *predicado* e estabelece entre ambos uma relação de predicação conforme apresentado abaixo:

8 Takano (2002) discute dados de objeto duplo do Japonês que envolvem movimentos de argumentos que não formam um constituinte.

9 Den Dikken assume que o *predicado* é um *constituente sintático* que expressa uma propriedade atribuída ao *sujeito* (Tradução nossa). O autor prefere essa definição mais restrita a uma definição de predicado ‘n-ary’ (baseada em Frege), expressada em termos de grade temática. Acreditamos que seja assim, por essa ser uma forma de excluir a complementação como uma forma de predicação.

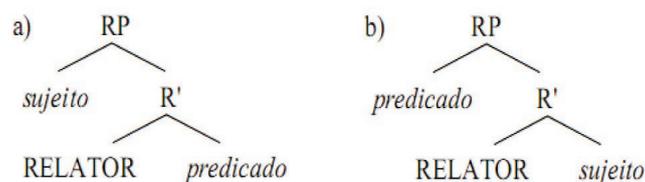


Figura 4. O Sintagma Relator como local da predicação

Observe, nas representações (a) e (b) da figura (4), que o sujeito e o predicado ocupam ou *spec* de RP ou sua posição *compl*. O autor chama (a) de estrutura *predicado-complemento* e (b) de estrutura *predicado-especificador*. O que resta de mais importante nesse ponto é que a relação de predicação deve ocorrer dentro do sintagma RP. Essa imposição dá conta da localidade da predicação, isto é, a predicação deve acontecer no domínio mínimo<sup>10</sup> do RELATOR e ser mediada por ele.<sup>11</sup>

É importante notar que a ideia de Den Dikken, apesar de semelhante à proposta de Bowers (1993), no sentido de que a predicação seja assimétrica e mediada por um núcleo funcional, difere desta, porque a categoria que Den Dikken propõe é necessariamente abstrata e pode se realizar de diferentes formas na gramática – por exemplo, como a cópula *be* (*ser*), como *as*, *for* etc. Por sua vez, a de Bowers propõe uma categoria funcional independente, de rótulo Pr, responsável pela predicação.

O autor propõe ainda que todas as relações entre sujeito-predicado sejam hierarquicamente assimétricas e não direcionais, mediadas pelo RELATOR. E, com isso, *todas* as construções com predicados secundários, por envolverem uma predicação, são, na verdade, um RP; e também as construções de SC de Stowell e Rothstein são na verdade RPs e, portanto, a relação entre o sujeito da SC e do predicado é mediada por um RELATOR. Observe que, com esse movimento, o autor está mais na linha de Stowell excluindo a predicação secundária Rothstein (1983, 1995, 2001) como uma possibilidade de instanciar a predicação.

Ao assumirmos um RELATOR como mediador da predicação e que as SC também são casos de predicação, não há qualquer possibilidade, nesse quadro, de se pensar a SC como uma categoria sem projeção funcional. E, como consequência, a SC é uma projeção do RELATOR, um constituinte. A única diferença das SC para os outros casos de predicação (por exemplo, os da oração plena – *full clause*) vem do fato de que as SC são “*a subject predicate structure lacking tense*” (DEN DIKKEN, 2006, p. 60).<sup>12</sup>

Nesse sentido, a Teoria dos *Relators* e *Linkers* se insere na tradição da Teoria das *Small Clauses*, pois o RP é na verdade uma SC nucleada por uma categoria funcional. As duas teorias se diferenciam na medida em que o núcleo da SC para Stowell é o predicado (um núcleo lexical e, portanto, a SC é uma projeção deste núcleo) e para Den Dikken é um núcleo funcional (o núcleo R, e por isso RP). Então, ambas as teorias estão juntas sobre a localidade da predicação, e também quanto à SC ser um constituinte, mas se diferenciam (entre outros aspectos) quanto à forma da predicação.

10 Da mesma forma, para dar conta da localidade, Rothstein (1995) trata da predicação como m-command. Stowell assume c-comando mútuo entre o sujeito e o predicado, como vimos em (16). Outros autores acabam também assumindo algum tipo de localidade para a predicação.

11 Não vamos tratar dos LINKERS nesse texto. Remetemos o leitor à Den Dikken (2006).

12 Uma estrutura sujeito predicado sem tempo (tradução nossa).

Há, no entanto, um caminho escolhido pela Teoria dos *Relators e Linkers* que nos chama a atenção e nos interessa pela relação que esse caminho tem com os nossos dados. O autor apresenta os exemplos abaixo (p. 22) para falar do que ele chama de *Construções de Predicação Secundária* (apesar da coincidência do nome, não se trata somente dos Predicados Secundários de Rothstein).

- (22) a. [<sub>VP</sub> considers [<sub>RP</sub> [Imogen] [R' RELATOR [<sub>AP</sub> smart]]]] (Small Clause complemento)  
 b. [<sub>VP</sub> hung [<sub>RP</sub> [his shirt] [R' RELATOR [<sub>PP</sub> on the line]]]] (construção locativa)  
 c. [<sub>VP</sub> hammered [<sub>RP</sub> [the metal] [R' RELATOR [<sub>AP</sub> flat]]]] (Resultativa)  
 d. [<sub>VP</sub> ran [<sub>RP</sub> [the pavement] [R' RELATOR [<sub>AP</sub> thin]]]] (Resultativa)  
 e. [<sub>VP</sub> paint [<sub>RP</sub> [the the book] [R' RELATOR [<sub>AP</sub> yellow]]]] (Resultativa)

Assim, pelas análises acima, vemos que Den Dikken trata as SC, os predicados secundários, as resultativas e os complementos duplos<sup>13</sup> como instâncias de construção de predicação secundária em que V seleciona um RP sem tempo (o RP é complemento em todos os casos de (22)). Esse recorte fica ainda mais claro a partir desta citação: “*I will assume here that all secondary predication constructions involve a small-clause structure*” (DEN DIKKEN, 2006, p. 60).<sup>14</sup>

A opção por uma DE que apresenta o RP como complemento de V fica mais clara pela escolha que o autor faz ao citar na literatura as três teorias concorrentes abaixo:

- (23) a. [<sub>VP</sub> V DP Pred ] → (WILLIAMS, 1980)  
 b. [<sub>VP</sub> V Pred ] DP ] → (NEELEMAN, 1994)  
 c. [<sub>VP</sub> V [ DP Pred ] → (KAYNE, 1984; HOEKSTRA, 1988)

O autor assume que somente (23c) apresente uma estrutura em que DP e Pred formam uma unidade que permita propor que o Relator seja um núcleo dessas unidades (sintagmas). Em (23a), o DP e o Pred não formam um constituinte sozinhos, além de (23a) ter uma estrutura ternária que foi excluída por outros motivos na teoria. E, em (23b), o *Pred* (predicado) não forma com seu DP sujeito um constituinte. Portanto, o autor assume (23c) como a estrutura adequada para os casos de Construções de Predicação Secundária.

Fica claro que, com essa escolha, a Teoria do *Relators e Linkers* nos fornece uma possibilidade de solucionar o paradoxo, porque propõe que a Construção de Predicação Secundária forme um sintagma, irmão de V, que pode ser movido.

Vejamos novamente o exemplo (19) repetido abaixo como (24) e considere as figuras em (05):

- (24) Esse bife mal passado eu até comeria \_\_, mas ele cru eu não como \_\_ de jeito nenhum.

Observe que a sequência ‘esse bife mal passado’ e ‘ele cru’ são expressões movidas. Seus lugares de origem estão indicados pelos traços baixos. Além disso, a sequência

13 As construções de objeto duplo são discutidas em outra parte do livro, na qual o autor assume, citando Hoekstra (1988), que esses dados são na verdade casos de SC complemento; portanto, RP complemento. Para isso, ele trata os verbos como ‘dar’ e ‘por’ não como verbos de três argumentos, mas de dois argumentos; i.e., eles selecionam um argumento externo DP e um argumento interno RP.

14 “Vou assumir aqui que toda construções de predicação secundária envolve uma estrutura de *small clause*” (tradução nossa).

‘ele cru’ indica claramente que se trata de uma predicação, por causa do pronome ‘ele’. A proposta do Den Dikken fornece uma maneira de codificar essa predicação como um constituinte e, nós argumentamos, ainda de permitir seu movimento para a periferia esquerda da sentença. Observe a figura abaixo que compara as DEs das Teoria do Predicação Secundária da Rothstein com a proposta do Den Dikken.

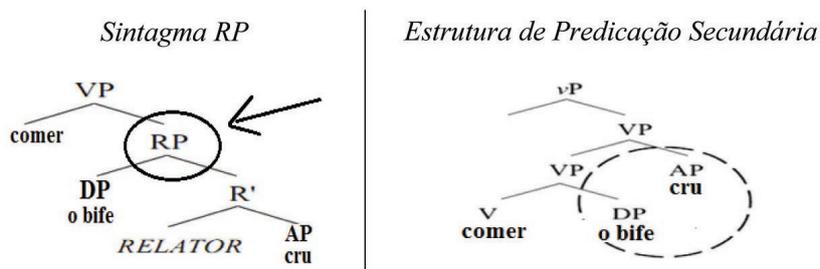


Figura 5. Sintagma Relator versus Predicado Secundário

O que vemos é que o nó RP instancia um constituinte e que é, portanto, passível de movimento. Essa possibilidade não está disponível em ambas as propostas canônicas anteriores.

Parece-nos, dessa forma, que temos uma solução para o Paradoxo de Constituição apresentado na seção anterior, apesar de essa proposta ainda precisar de uma discussão mais completa, colocando-a frente ou testando-a com os dados de um paradigma maior, como o apresentado em Carreira (2008).

### Considerações finais

Neste trabalho, descrevemos o comportamento dos predicados secundários em PB e alemão, sobretudo no que se refere à ambiguidade estrutural apresentada na primeira língua e ausente na segunda. Observando a morfologia dos adjetivos em predicado secundário (PB e alemão), notamos que a concordância não parece ser relevante para definir constituição. Fica para análise futura a questão da relevância para definir a posição do predicado secundário na SD, ou mesmo sua estrutura interna. Fizemos uma breve revisão bibliográfica de algumas das análises canônicas e mostramos que elas levam a um paradoxo de constituição.

Já uma análise com base em Den Dikken (2006) parece dar conta da constituição, sem esbarrar no paradoxo. No entanto, uma análise desse tipo levanta outras questões, que precisam ser colocadas minimamente. Considere os dados abaixo:

(24) \*John foolish was considered *ti* (ROTHSTEIN, 2001, p. 120)

(25) \*John foolish is what you considered *ti* (ROTHSTEIN, 2001, p. 120)

Porque o verbo ‘consider’ (considerar), que é o verbo mais citado como aquele que seleciona uma SC-complemento (portanto um RP), não permite o fronteamento?

Outra questão: por que o dado em (13b), repetido abaixo, resiste à leitura prediativa? Não estamos agora fazendo a previsão de que o movimento da sequência grifada seja possível, já que ela seria um RP?

- (13) a. A Duda comeu o bife cru  
b. \*Foi o bife cru que a Duda comeu \_\_

Apesar disso, acreditamos que o fato de esse movimento – como descrito em (13b) – não existir, pode ser resultado de uma motivação independente que ainda precisa ser demonstrada.

## REFERÊNCIAS

BARBOSA, J. W. C. *A estrutura sintática das chamadas “construções resultativas em PB”*. 2008. 134 f. Dissertação (Mestrado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

CÂMARA JR., J. M. *Estrutura da Língua Portuguesa*. 2. ed. Petrópolis R/J: Vozes. Limitada, 1970.

CARREIRA, M. B. *Diagnósticos de Constituição para Construções Predicativas Adjetivais*. 2008. 128 f. Dissertação (Mestrado em Letras) – UFPR, Curitiba.

CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.

\_\_\_\_\_. *The Minimalist Program*. Cambridge, Mass.: The MIT Press, 1995.

DIKKEN, M. den. *Relators and Linkers: the Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, Mass.: The MIT Press, 2006.

FOLTRAN, M. J. G. D. *As construções de predicação secundária no português do Brasil: aspectos sintáticos e semânticos*. 1999. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo.

HOEKSTRA, T. Small clause results. *Lingua*, v. 74, p.101-139, 1988.

KAYNE, R. *Connectedness and Binary Branching*. Dordrecht: Foris, 1984.

KNÖPFLE, A. Sintaxe das resultativas adjetivais no alemão: ‘contra exemplos’ e os dados do PB. *Estudos Linguísticos*, São Paulo, v. 40, n. 1, p. 420-433, jan./abr. 2011a. Disponível em: <[http://www.gel.org.br/estudoslinguisticos/volumes/40/el\\_vol.40\\_n.1\\_Integra.pdf](http://www.gel.org.br/estudoslinguisticos/volumes/40/el_vol.40_n.1_Integra.pdf)>. Acesso em: 03 mar. 2013.

\_\_\_\_\_. Resultativas adjetivais e o estatuto nu do adjetivo. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 19, n. 1, p. 115-142, 2011b.

KRATZER, A. Building resultatives. In: MAIENBAUM, C.; WÖLLSTEIN-LEISEN, A. (Ed.). *Event arguments in syntax, semantics, and discourse*. Tübingen: Niemeyer, 2005. p. 177-212.

LOBATO, L. Afinal, existe a construção resultativa em português? In: NEGRI, L.; FOLTRAN; M. J.; PIRES DE OLIVEIRA, R. (Org.). *Sentido e Significação*. São Paulo: Contexto, 2004. p. 142-180.

NEELEMAN, A. *Complex Predicates*. University of Utrecht, 1994. (Unpublished doctoral dissertation.)

ROTHSTEIN, S. *Predicates and their subjects*. Dordrecht: Kluwer, 2001.

\_\_\_\_\_. Small clauses and copular constructions. In: CARDINALETTI, A.; GUASTI, M. T. (Ed.). *Syntax and Semantics*. London & New York: Academic Press, 1995. p. 27-48.

\_\_\_\_\_. *The Syntactic Forms of Predication*. Tese (Ph.D. Dissertation) — MIT, 1983. Disponível em: <<http://hdl.handle.net/1721.1/15509>>. Acesso em: 03 mar. 2013.

STOWELL, T. Remarks on clause structure. *Syntax and Semantics*, San Diego, v. 28, p. 271-286, 1995.

\_\_\_\_\_. Small clause restructuring. In: FREIDIN, R. (Ed.). *Principle and parameters in comparative grammar*. Cambridge: MIT Press, 1991. p. 182-218.

\_\_\_\_\_. Subject across categories. *The Linguistic Review*, v. 2, p. 285-312, 1983.

\_\_\_\_\_. *Origins of phrase structure*. Tese. MIT, 1981. (Unpublished doctoral dissertation)

TAKANO, Y. Surprising constituents. *Journal of East Asian Linguistics*, v. 11, p. 243-301, 2002.

WILLIAMS, E. Predication. *Linguistic Inquiry*, v. 11, p. 203-238, 1980.

# A representação do objeto *norma* na obra *Sociolinguística, os níveis da fala*, de Dino Preti

(The representation of the object *norm* in the book *Sociolinguística, os níveis da fala* by Dino Preti)

Gil Negreiros<sup>1</sup>

<sup>1</sup>Universidade Federal de Santa Maria (UFSM)

gil.negreiros@yahoo.com.br

**Abstract:** The theme of this paper refers to the linguistic ideas that mark the academic production of Dino Preti. Our goal is to show how was the design of the object *norm* in the book *Sociolinguistics, levels of speech*. The book is theoretically based on the assumptions of Historiography Epistemology of Aurox (1992). This theoretical line considers language as a social fact, and it is based on the explanation and interpretation of philosophical-scientific human knowledge. There is an interest, in this theoretical framework, in the causal connections between the investigated object and others related to it, which is based on five parameters (system objects, temporal parameter, space parameter, external system parameterization that connects the objects to its context and system of interpretants). We follow, in this research, the Phenomenological Method from its four methodological moments, namely, immersion, discrimination, assignment and synthesis.

**Keywords:** History of Linguistic Ideas; Linguistics in Brazil; Dino Preti.

**Resumo:** O tema deste trabalho refere-se às ideias linguísticas que marcam a produção acadêmica de Dino Preti. Nosso objetivo é mostrar como se deu o delineamento do objeto *norma* na obra *Sociolinguística, os níveis da fala*. O trabalho se apoia, teoricamente, nos pressupostos da Historiografia Epistemológica de Aurox (1992). Essa linha teórica considera a língua como fato social e se baseia na explicação e na interpretação filosófico-científica do conhecimento humano. Há um interesse, a partir desse quadro teórico, nas ligações causais do objeto investigado com outros que lhe são relacionados, a partir de cinco parâmetros (sistema de objetos, parâmetro temporal, parâmetro espacial, sistema de parametrização externo que liga o sistema de objetos ao seu contexto e sistema de interpretantes). Seguimos, no processo de pesquisa, o Método Fenomenológico, a partir de seus quatro momentos metodológicos, a saber: imersão, discriminação, atribuição e síntese.

**Palavras-chave:** História das Ideias Linguísticas; Linguística do Brasil; Dino Preti.

## Introdução

A história das ciências é, atualmente, uma das áreas mais profícuas na investigação científica. Conhecer o passado, pensar a prospecção a fim de entender o presente: é essa a ideia que rege os estudos de historiografia. Contudo, mesmo diante dessa importância, os trabalhos de historiografia das ciências ocorrem de forma desequilibrada. Enquanto há uma quantidade satisfatória de discussões a respeito da história do pensamento matemático e da realidade natural, as ciências da linguagem não alcançam o mesmo nível de interesse por parte dos estudiosos (AUROUX, 2006, p. 106).

No caso específico da historiografia brasileira, dedicada ao estudo das ciências da linguagem, podemos afirmar que há muito a investigar. Como exemplo, basta remetermos aos poucos trabalhos de historiografia a respeito do pensamento e da obra de linguistas de destaque.

Também é exemplo dessas lacunas a falta de desenvolvimento de pesquisas que busquem interpretar o percurso de certos temas, como os relativos à Análise da Conversação e à Sociolinguística, no âmbito do pensamento linguístico brasileiro, embora aspectos orais e sociolinguísticos sejam temas relevantes e quase sempre presentes nas discussões acadêmico-linguísticas das últimas quatro décadas.

Nesse período, são significativas as produções acadêmicas que enfatizam a importância de um trabalho sério e científico com a oralidade e com temas ligados à Sociolinguística como, por exemplo, a variação linguística na formação escolar da sociedade.

Em consonância com essa proposta acadêmica, alguns “instrumentos linguísticos”, como livros didáticos e materiais bibliográficos de apoio ao ensino de linguagem, começaram a apresentar, em seu conteúdo, mesmo de forma pouco enfática, aspectos relativos à importância de trabalhos voltados para as questões orais e sociolinguísticas, em resposta positiva ao que estava sendo buscado pelas ciências da linguagem que, por meio de outros “instrumentos linguísticos”,<sup>1</sup> delineavam a oralidade e os usos da língua na sociedade como objeto de pesquisa.

Um dos pesquisadores brasileiros mais significativos de todo esse percurso é Dino Preti, autor preocupado em estudar, pesquisar e divulgar a Análise da Conversação e a Sociolinguística nos contextos de pesquisa e de ensino. Em quatro décadas, a obra de Preti sempre foi marcada pelo trabalho com esses temas, o que, possivelmente, fez com que seu pensamento fosse quase sempre reconstruído e redirecionado, de acordo com os aspectos históricos do momento de produção.

Diante disso, objetiva-se, neste trabalho, que pertence a uma pesquisa maior,<sup>2</sup> apresentar alguns resultados da investigação que propomos das ideias linguísticas e das causalidades históricas que nortearam a primeira obra de destaque de Dino Preti, que foi a adaptação, em livro, de seu trabalho de doutoramento, publicado sob o título de *Sociolinguística, os níveis da fala* (denominado, neste trabalho, pela sigla SNF). Mais especificamente, pretende-se mostrar, nesta comunicação, como se deu, a nosso ver, o delineamento do objeto *norma* na obra em xequê.

O *corpus* escolhido é a segunda edição da obra, publicada em 1974 pela Editora Nacional. No âmbito deste trabalho, serão feitas considerações sobre os três primeiros capítulos de SNF, referentes ao recorte teórico feito pelo autor: o primeiro, intitulado “A Sociolinguística e o fenômeno da *diversidade* na língua de um grupo social”; o segundo, cujo título é “A *norma* e os fatores de unificação linguística, na comunidade” e o terceiro, “A representação escrita dos *níveis de fala*”

O trabalho se apoia, teoricamente, nos pressupostos da Historiografia Epistemológica, na modalidade adotada por Auroux (1998, 2006 e 2008) e Puech (2006). Essa linha teórica considera a língua como fato social e se baseia na explicação e na interpretação filosófico-científica do conhecimento humano. Há um interesse, a partir desse quadro teórico, nas ligações causais do objeto investigado com outros que lhe são relacionados, a partir de cinco parâmetros (sistema de objetos, parâmetro temporal, parâmetro espacial, sistema

---

1 Para mais detalhes do conceito de “instrumentos linguísticos”, consultar Auroux (1992).

2 Este trabalho é parte da pesquisa *Ideias Linguísticas na obra de Dino Preti*, desenvolvida na Universidade de São Paulo sob a supervisão da Prof.<sup>a</sup> Dr.<sup>a</sup> Marli Quadros Leite. A pesquisa teve o financiamento do CNPq, órgão ao qual agradecemos.

de parametragem externo que liga o sistema de objetos ao seu contexto e sistema de interpretantes).

Torna-se oportuno, também, afirmar que, ao desenvolver uma pesquisa sobre a história científica da obra de Dino Preti, aqui representada pela análise do objeto “norma” na obra SNF, estamos no âmbito da “História do Presente”. Dessa maneira, não só tentaremos compreender o passado próximo, mas também o momento que vivemos atualmente. Sobre isso, Chaveau e Tétard (1999, p. 15) afirmam que a história pode ser, com um menor recuo e métodos particulares, não apenas um estudo do passado, mas também um estudo do presente.

Os mesmos autores também postulam que:

[...] o estudo historiográfico, metodológico e epistemológico dos tempos atuais está apenas desbravado. [...] A história do imediato e a do presente demandam uma definição mais precisa, em seu próprio funcionamento, a fim de ser melhor percebidas [sic], individualmente, e uma em relação à outra. (CHAVEAU; TÉTARD, 1999, p. 21)

Por fim, cabe lembrar que adotamos o Método Fenomenológico, que se enquadra no campo da análise qualitativa de dados, por meio de seus quatro momentos, que são:

1) imersão: leitura de todo material, com a finalidade de suspender as crenças e as opiniões, com aceitação acrítica de tudo quanto é proposto nos relatos;

2) discriminação: separação e organização dos dados a serem analisados, de acordo com o critério definido teoricamente;

3) atribuição de sentido: interpretação dos dados de acordo com suas referências teóricas e subjetivas. Trata-se de um momento de análise dissertativa e não de mera descrição;

4) síntese: mapeamento das atribuições de sentido obtidas por meio da análise dos dados, com o intuito de alcançar uma compreensão geral e superior do fenômeno pesquisado (APPOLINÁRIO, 2006, p. 167).

Filosoficamente, o Método Fenomenológico se refere ao exame da relação entre o mundo e os sentidos humanos que o experimentam. Segundo Husserl, criador do termo “fenomenologia” e um dos principais nomes desse método, o conhecimento é obtido por meio da intuição. Segundo o pensador, para experimentarmos determinada realidade, torna-se necessário destacar e separar todos os nossos pressupostos sobre ela, para que se possa detectar a “essência” dos fenômenos. Desse modo, ao suspender nosso julgamento sobre as coisas, temos acesso ao dado verdadeiro e puro e, assim, poderemos descrevê-lo como ele realmente se apresenta (APPOLINÁRIO, 2006, p. 165).

O método fenomenológico, como pertencente ao campo das análises qualitativas de dados, apresenta a possibilidade de a análise poder se iniciar até mesmo ao longo da coleta de dados. Além disso, não se buscam, por meio desse método, generalizações a respeito do tema. O que se busca é compreender um fenômeno no seu sentido mais intenso.

## Conceitos de *língua* e *fala* na representação do objeto *norma* em SNF

Preti, em seu SNF, aproveitando-se das ideias de Saussure sobre “a unidade” e as “características conservadoras” da *langue*,<sup>3</sup> chama a atenção para uma aparente contradição presente em fatos como esses: a) em uma mesma época, duas pessoas não falam da mesma maneira a língua; b) em uma mesma época, uma única pessoa não fala, em todos os momentos, da mesma maneira. Mesmo assim, com esse comportamento variante explícito, é possível afirmar que exista, no seio de uma comunidade linguística, um comportamento de manutenção de unidade:

[...] sabemos que, a partir do instante em que a comunidade aceita uma língua como meio primordial de comunicação, toda e qualquer variação lhe será prejudicial, motivo pelo qual a tendência é manter sua *unidade*, colaborando todos, consciente e inconscientemente, no sentido de sua nivelção, pois dessa maneira a compreensão será mais fácil, e a própria integração do indivíduo na cultura comum se dará com mais facilidade. (PRETI, 1974, p. 27)

Assim, ao mesmo tempo em que afirma existir uma variação “nos atos de fala”, e, conseqüentemente, que essas variações são “autorizadas” pelo sistema (pela *langue*), Preti, ao propor tal discussão, considera a “língua” como um “meio primordial de comunicação”, atribuindo-lhe a manutenção da comunidade.

Obviamente, visto sob o prisma do momento presente, poderíamos supor a existência de uma discrepância na questão, se considerarmos os conceitos saussurianos de *langue* e *parole*. É aceitável, hoje, que há diferenças entre uma língua (e os processos de mudança que nela ocorrem) e os conceitos saussurianos de sistema e de mudança linguística.<sup>4</sup> Contudo, essa postura teleológica não cabe aqui. O que importa, na verdade, é detectar aspectos característicos do momento de produção da obra SNF, marcado por influências teóricas muitas vezes heterogêneas.

Preti, ao discutir a questão dos usos frente a um “sistema” ideal, postula que a força conservadora, que age a favor da uniformidade linguística, é garantida pelos *usos* linguísticos de uma comunidade, que estão a serviço da necessidade de os falantes serem compreendidos. Todos esses usos, apesar de possuírem marcas heterogêneas, são regidos por uma regra geral, coletiva. Existe, assim, uma entidade linguística ideal, imposta no inconsciente das pessoas. Essas considerações de Preti advêm do pensamento de Sapir, Bally e de Vendryès, principalmente no que se refere ao uso comunitário de uma língua. Parece-nos que esses três autores, além de serem citados na bibliografia de SNF, embasam o discurso “pretiano”, em extenso processo de heterogeneidade constitutiva, conforme pretendemos demonstrar nas próximas linhas.

A questão de uma “força conservadora tácita” pode ser considerada como uma marca enunciativa do discurso de Sapir no discurso de Preti. É naquele que este se embasa para explicar a variedade na uniformidade.

Sapir, ao discutir as diferenças entre o inglês nova-iorquino e o inglês falado pela classe média de Londres (duas formas consideradas por Sapir como dois *dialetos*), postula que “as variações individuais mínguem e desaparecem diante de certas concordâncias

3 Ver Saussure, s/d, p. 88.

4 A mudança linguística, para Saussure, é vista a partir de uma analogia entre o sistema e um “jogo de xadrez”.

maiores – como, por exemplo, na pronúncia e no vocabulário – que ressaltam com vigor quando a língua do grupo em conjunto é comparada à do outro grupo” (SAPIR, 1980, p. 119).

Dessa forma, é possível observar que Sapir defende algo maior que as liberdades individuais dos falantes, que impõe aos falantes certos limites:

Isso mostra que há uma como que entidade linguística ideal, a dominar a fala habitual dos membros de cada grupo; que o sentimento de liberdade quase irrestrita, que tem cada indivíduo ao usar a sua língua é contido por uma tácita norma diretriz. Certo indivíduo joga com essa norma de um modo que lhe é peculiar; outro atém-se ao que existe naquilo em que o primeiro mais caracteristicamente inova, mas por sua vez diverge da média geral em certos pontos que lhe são próprios; e assim por diante. (1980, p. 120)

Bally, por sua vez, que se diz fiel à distinção saussuriana entre língua e fala, acrescenta ao campo da língua um item fundamental no entendimento de sua estilística: a língua falada focalizada no seu conteúdo afetivo e subjetivo. É assim que, para Bally, a estilística se enquadra na linguística geral (1913, p. 158-9).

Um dos itens de seu conteúdo subjetivo que merece destaque é, segundo o autor, a questão do valor subjetivo que se dá aos usos da comunidade:

Uma expressão é chamada usual: 1) quando designa uma coisa ou uma ideia indissolúvelmente ligada com a vida, considerada em suas características sociais comuns a todo um grupo linguístico; 2) quando esta designação dá imediatamente a qualquer membro deste grupo linguístico, não importa qual, a impressão de que “isto se diz assim, deve-se dizer assim, foi sempre dito assim, e se dirá sempre assim”. Esta crença é, sem dúvida, pura ilusão, pois que é desmentida pela evolução constante da linguagem; mas, pelo seu valor subjetivo, é uma realidade absoluta, sem a qual a consciência de um estado da língua não seria possível. O conjunto de meios de expressão que possuem este caráter uniforme no espírito de todos os indivíduos, eis o que para nós é a língua usual. (BALLY, s/d, p. 209)

Vendryès, por seu turno, afirma que o uso linguístico não é arbitrário, mas completamente a favor dos interesses da comunidade, ideia que vem ao encontro das proposições de Bally:

Ele [o uso] é determinado sempre pelo interesse da comunidade, que, neste caso, é a necessidade de ser compreendido. Cada um se opõe, por conseguinte, sem sabê-lo e por instinto, à introdução do arbitrário no *uso*. Quando por parte de um indivíduo isolado se produz uma infração, é corrigida imediatamente: o ridículo castiga suficientemente o culpado, para tirar-lhe todo desejo de repeti-la. Para que uma infração ganhe força de lei, é preciso que todos os membros da comunidade estejam igualmente dispostos a cometê-la, isto é, que seja recebida como regra e, por conseguinte, deixe de ser infração. (VENDRYÈS, 1968, p. 268)

Ao se apoiar nas ideias de Bally<sup>5</sup> e de Vendryès sobre o “uso linguístico em uma comunidade” para definir o fator de conservação de uma língua, Preti traz à discussão as ideias sobre *norma* de Coseriu (1967), pensador que atuou, nos dizeres de Ilari (2004), no limite entre o Estruturalismo e o “Pós-estruturalismo”:

5 Importante frisar aqui a presença das ideias de Bally no horizonte de retrospectiva de Preti. Talvez pelo envolvimento desse último com questões relacionadas ao aproveitamento dos usos linguísticos na literatura de língua portuguesa, o fato é que quase sempre é explícita, na produção de Preti, a influência do “Pai da Estilística”.

Para Coseriu, a possibilidade de delimitar uma sincronia é, até certo, uma ficção, pois a todo momento, em qualquer língua, convivem mecanismos gramaticais e recursos lexicais que são fruto de diferentes momentos da história. O velho convive com o novo, e é essa convivência de fragmentos de velhos sistemas com fragmentos de novos sistemas que caracteriza um estado de língua dado. Por isso, diz Coseriu, o linguista estará lidando o tempo todo com *pancronias*. Coseriu revisou também a oposição *língua (sistema linguístico) X fala*: entre a fala e o sistema, sugere que se considere uma instância intermediária, muito mais operacional e psicologicamente mais real que a própria língua: a norma. (ILARI, 2004, p. 81)

É justamente Coseriu o autor de uma das críticas mais construtivas a respeito da dicotomia saussuriana *langue / parole*. Coseriu afirma que “existía ‘uma identificación inicial’ entre lo que era ‘individual’ y lo que era ‘concreto’ (‘parole’), por una parte, y entre lo que era ‘social’ y lo que era ‘formal’ o ‘funcional’, por outra” (COSERIU, 1967, p. 53). A partir desse impasse, Coseriu propõe, segundo Matthews (1998), que não é o sistema (a *langue*) que se impõe ao indivíduo nos processos de mudança linguística a escolher determinado caminho, mas sim a adequação a uma norma, que são sistemas de obrigações obrigadas, de imposições sociais e culturais. Sobre isso, Matthews postula:

La solución que propone Coseriu a este conflicto consiste en reconocer dos niveles sucesivos de abstracción, en lugar de uno solo. Em el nivel superior cabe postular la existencia de un sistema. Sin embargo, no se trataría de otra cosa que de “um sistema de posibilidades”, es decir, de “coordenadas” que especifican lo que lês está permitido a lós hablantes y lo que les está vedado (“que indican caminos abiertos y caminos cerrados”, por citar literalmente a Coseriu). (MATTHEWS, 1998, p. 83)

Parece-nos que a mesma ideia é apresentada por Preti, que postula que:

A norma é o ponto de chegada no processo de uniformização e nivelamento da língua de uma comunidade. É o momento em que o uso é fixado em lei linguística. A própria sociedade se encarrega de preservar a *norma*, que ela própria estabeleceu. (PRETI, 1974, p. 30-1).

É nítida a influência de Coseriu, de formação estruturalista, no pensamento de Preti, o que se dá também a partir de heterogeneidades constitutivas, como a que se observa no trecho acima, bem como em referências diretas, como em citações do autor, como no trecho que se segue, retirado da nota de rodapé 54: “*Norma* para Coseriu, em seu famoso estudo, seria ‘o que se disse e tradicionalmente se diz numa comunidade’, e compreenderia ‘só o que na fala concreta é repetição de modelos anteriores’” (COSERIU apud PRETI, 1974, p. 30).

É no pensamento de Coseriu que a ideia pretiana de *norma* ganha mais precisão, justamente no fato de o linguista romeno falar de diversas *normas*, de vários níveis e a partir de vários critérios, ideia muito próxima do ideal da Sociolinguística e daquilo que Preti necessitava no momento, ou seja, uma teoria que sustentasse a audácia de se analisar a língua falada no diálogo literário. A norma, assim, é

[...] um sistema de realizações obrigatórias, de imposições sociais e culturais e varia segundo a comunidade. Dentro da mesma comunidade linguística nacional e dentro do

mesmo sistema funcional podem-se comprovar várias normas (linguagem familiar, linguagem popular, linguagem literária, linguagem elevada, linguagem vulgar etc.) distintas, sobretudo no que concerne ao vocabulário, mas a miúdo também nas formas gramaticais e na pronúncia. (COSERIU, 1967, p. 104)

No Quadro 1, é possível demonstrar como Preti constrói seu discurso sobre *norma*, ponto fundamental para a análise dos “níveis de fala” no discurso literário brasileiro:

**Quadro 1. As representações dos objetos em SNF**

<b>Objetos</b>	<b>Características constitutivas</b>
Sistema	<ul style="list-style-type: none"> <li>- Discurso baseado em Saussure.</li> <li>- Domínio geral; Instituição social. Sistema abstrato de signos.</li> <li>- Acervo linguístico.</li> </ul>
Norma	<ul style="list-style-type: none"> <li>- Discurso baseado em Bally, Sapir, Vendryès e, principalmente, Coseriu.</li> <li>- Usos específicos autorizados pela comunidade de fala.</li> <li>- Representação do sistema na fala, pode ser “cultura” ou “regionais”, dependendo da aceitabilidade social.</li> <li>- É a realização do sistema.</li> </ul>
Fala	<ul style="list-style-type: none"> <li>- Discurso baseado em Saussure.</li> <li>- Considerada como “ato de fala”, é o uso feito a partir dos conhecimentos da norma adotada socialmente (e, indiretamente, a partir do conhecimento do sistema abstrato).</li> </ul>
Níveis sociolinguísticos	<ul style="list-style-type: none"> <li>- Discurso baseado nos autores da Sociolinguística Variacionista.</li> <li>- Ocorre no plano interindividual e são moldados a partir de elementos externos sociais, culturais, etnográficos e naturais.</li> </ul>

A *norma* (cultura ou não, de caráter geográfico ou não), considerada como a realização do *sistema*, obviamente não se opõe ao conceito de variação dos usos linguísticos. Pelo contrário, é possível afirmar, então, em variedades dentro de uma uniformidade. A tendência unificadora normativa fragmenta-se em tendências unificadoras menores, que se diferenciam em um ou outro grau, o que cria os níveis sociolinguísticos.

Assim, fica claro, no pensamento de Preti, outro índice de seu horizonte de retrospectiva: a questão das *normas*. Esse horizonte pode ser assim descrito, tendo em vista a construção das imagens dos objetos *sistema*, *norma (s)*, *fala* e *níveis de fala*.

**Quadro 2. Relação entre objetos históricos e índices do horizonte de retrospecção de Preti**

<b>Horizonte de retrospecção</b>	<b>Interpretantes</b>	<b>Objetos</b>	<b>Objetivos de Preti no momento histórico</b>
Estruturalismo saussuriano	Saussure	Sistema e fala	Analisar a influência da língua falada nos diálogos literários
Estilística “linguística” francesa	Bally	Norma	
Estruturalismo linguístico norte-americano	Sapir		
“Linguística da fala” (influência de Meillet)	Vendryès		
Estruturalismo “crítico”	Coseriu		
Sociolinguística variacionista	Labov e seu grupo	Níveis de Fala	

No Quadro 2, notamos os índices, já abordados, de base estruturalista, que formam o horizonte de retrospecção de Preti discutidos neste item do trabalho. Contudo, essa heterogeneidade que constitui o horizonte de retrospecção pode ser demonstrada, também, pelo fato de que, em conjunto com essa força epistemológica estruturalista, de influência saussuriana, há uma influência do campo teórico da Sociolinguística, área de estudos vinculada ao funcionalismo.

### **Considerações finais**

Neste trabalho, verificamos como o objeto *norma* é moldado na obra SNF. Na verdade, o conceito de *norma* resolve um problema epistemológico enfrentado por Preti no momento da junção entre a ideia de *língua* saussuriana, vista como sistema abstrato (conceito muito forte e em evidência na época de produção da obra em xeque) e a necessidade de estudo dos usos linguísticos, área de estudos de uma área científica nova para a época: a Sociolinguística variacionista laboviana.

Os teóricos que compõem o sistema de interpretantes “pretiano” também devem ser colocados em evidência. A nosso ver, a junção de teóricos de diversas escolas, como o uso de teóricos do Estruturalismo de base saussuriana, do Estruturalismo linguístico norte-americano, da Estilística francesa de caráter linguístico, da Sociolinguística Variacionista, pode ser um indício do ainda nebuloso momento histórico da linguística no contexto acadêmico brasileiro.

Esse emprego heterogêneo também se justifica pela tentativa de acertar a questão entre conceitos como *sistema*, *norma*, *fala* e *níveis de fala*, objetos às vezes tratados separadamente por teóricos de escolas diferentes.

A proposta de Preti em SNF é analisar diálogos literários sob a luz da Linguística. A nosso ver, tal proposta, no início da década de 1970, é desafiadora e dá importância à obra. Além disso, o livro também pode ser considerado como um “ponto de encontro” de diversas tendências teóricas, todas pertencentes à mesma época e que moldaram os horizontes de retrospecção dos trabalhos da Linguística do Brasil na época.

Torna-se evidente, também, como se dão o resgate e a reconstrução de certas ideias linguísticas em diferentes tempos e espaços na Linguística, aspecto nuclear nos horizontes de retrospectiva dos sujeitos-pesquisadores. Resgatar o passado das ciências, assim, é uma forma de entendê-las não só no ontem, mas principalmente no hoje e promovê-las na prospecção. Comungamos das ideias de Auroux (2006, p. 16), que, ao falar sobre essas tarefas da historiografia das ciências, afirma que uma ciência sem passado torna-se uma ciência morta. E os mortos não possuem futuro.

## REFERÊNCIAS

- APPOLINÁRIO, F. *Metodologia da Ciência: filosofia e prática da pesquisa*. São Paulo: Thomson Pioneira, 2006.
- AUROUX, S. *A revolução tecnológica da gramatização*. Campinas: Unicamp, 1992.
- \_\_\_\_\_. *A filosofia da linguagem*. Campinas: Unicamp, 1998.
- \_\_\_\_\_. Les méthodes d'historicisation. In: *Histoire, Epistemologie, Langage*. Université Paris VII, Tome XXVIII, Fasc. 1, 2006. p. 105-106.
- \_\_\_\_\_. *A questão da origem das línguas*. Campinas: RG, 2008.
- BALLY, C. *Le langage et la vie*. Genebra: Droz, 1913. [*Em lenguaje y la vida*. Buenos Aires: Losada]
- \_\_\_\_\_. *Traité de Stylistique Française*. Genebra: Librairie George & Cie., s/d.
- CASTILHO, A. T. Apresentação. In: PRETI, D. *Sociolinguística, os níveis da fala*. São Paulo: Nacional, 1974.
- CHAVEAU, A.; TÉTART, P. (Org.). *Questões para a história do presente*. Bauru: Edusc, 1999.
- COSERIU, E. *Teoría del lenguaje y Lingüística General*. Madrid: Gredos, 1967.
- ILARI, R. O estruturalismo linguístico: alguns caminhos. In: MUSSALIM, F.; BENTES, A. C. *Introdução à linguística: fundamentos epistemológicos*. São Paulo: Cortez, 2004.
- MATTHEWS, P. *Breve história de la linguistic structural*. Madrid: Akal, 2009.
- PRETI, D. *Sociolinguística, os níveis da fala*. São Paulo: Nacional, 1974.
- PUECH, C. Pour une histoire de la linguistique dans l'histoire de la Linguistique? In: *Histoire, Epistemologie, Langage*. Université Paris VII, Tome XXVIII, Fasc. 1, 2006. p. 9-24.
- SAPIR, E. *A linguagem*. São Paulo: Perspectiva, 1980.
- SAUSSURE, F. *Curso de Linguística Geral*. São Paulo: Cultrix, s/d.
- VENDRYÈS, J. *Le langage*. Paris: Albin Michel, 1968.

# Vocabulário Ortográfico Comum (VOC): constituição de uma base lexical para a língua portuguesa

(Common Orthographic Vocabulary (VOC): development of a lexical database for the Portuguese language)

**Gladis Maria de Barcellos Almeida<sup>1</sup>, José Pedro Ferreira<sup>2</sup>, Margarita Correia<sup>3</sup>,  
Gilvan Müller de Oliveira<sup>4</sup>**

<sup>1</sup> Departamento de Letras – Universidade Federal de São Carlos (UFSCar)

<sup>2</sup> Instituto de Linguística Teórica e Computacional (ILTEC)

<sup>3</sup> Instituto de Linguística Teórica e Computacional (ILTEC) e Universidade de Lisboa (UL)

<sup>4</sup> Instituto Internacional da Língua Portuguesa (IILP)

gladis@ufscar.br, jpf@iltec.pt, mcf@iltec.pt, gimioliz@gmail.com

**Abstract:** The International Institute for the Portuguese Language, under the auspices of the Community of Portuguese Speaking Countries, is coordinating the development of the Common Orthographic Vocabulary (VOC). VOC will be a large on-line lexical database which will take into account the varieties of the eight Portuguese speaking countries (Angola, Brazil, Cape Verde, Guinea-Bissau, Mozambique, Portugal, São Tome and Príncipe, and East Timor). The project comprises two inter-related phases: 1) the merging of existing Portuguese and Brazilian vocabularies as an evidence of the lexicographic tradition; 2) the compilation of corpora for the development of national vocabularies for the remaining countries. This paper describes in detail the methodology of the task being pursued in the context of that project.

**Key-words:** lexical database; corpus; lexical resource; lexicographic tradition; orthographic vocabulary

**Resumo:** O Instituto Internacional da Língua Portuguesa (IILP), sob os auspícios da Comunidade dos Países de Língua Portuguesa, lidera a organização do Vocabulário Ortográfico Comum da Língua Portuguesa (VOC). O VOC constituir-se-á numa grande base lexical *on-line*, que contemplará as variedades dos oito países lusófonos (Angola, Brasil, Cabo Verde, Guiné-Bissau, Moçambique, Portugal, São Tomé e Príncipe e Timor-Leste). Esse projeto está dividido em duas fases: 1) a junção de vocabulários já existentes em Portugal e Brasil, evidenciando a tradição lexicográfica portuguesa; 2) a elaboração de *corpora* para a constituição dos vocabulários nacionais dos demais países. Assim, nesta oportunidade, pretende-se detalhar os aspectos metodológicos que subjazem às tarefas envolvidas no referido projeto.

**Palavras-chave:** base lexical; *corpus*; recurso lexical; tradição lexicográfica; vocabulário ortográfico

## Introdução

A constituição do Vocabulário Ortográfico Comum da Língua Portuguesa (VOC) é um desiderato da Comunidade dos Países de Língua Portuguesa<sup>1</sup> (CPLP) e uma determinação do Acordo Ortográfico da Língua Portuguesa de 1990<sup>2</sup> (AO90). O projeto

<sup>1</sup> Mais informações em <<http://www.cplp.org/>>.

<sup>2</sup> O Acordo Ortográfico da Língua Portuguesa foi assinado em Lisboa, em 16 de dezembro de 1990, por todos os países membros à época. A constituição de um Vocabulário Ortográfico Comum foi um requisito do Acordo: “Os Estados signatários tomarão, através das instituições e órgãos competentes, as providências necessárias com vista à elaboração, até 1 de Janeiro de 1993, de um vocabulário ortográfico comum da língua portuguesa, tão completo quanto desejável e tão normalizador quanto possível, no que se refere às terminologias científicas e técnicas.” (Artigo 2.º do texto do AO90).

encontra-se atualmente em curso, sob a coordenação do Instituto Internacional de Língua Portuguesa<sup>3</sup> (IILP). Tentativas anteriores de elaborar um vocabulário ortográfico comum esbarraram nas diferentes interpretações dadas ao texto das reformas ortográficas que pretendiam aplicar, as quais se refletiram no desenvolvimento de vocabulários ortográficos nacionais com divergências (Brasil e Portugal). Nesse contexto, o desenvolvimento do VOC nas condições descritas neste trabalho constitui não só a base da aplicação do AO90 em todos os países da CPLP, como um avanço sem precedentes na lexicografia de língua portuguesa.

A mais recente reforma ortográfica da língua portuguesa, em curso desde 2009,<sup>4</sup> que visa à unificação da ortografia nos vários países da CPLP, apenas unificou as regras de escrita, que preveem a coexistência de variantes nacionais em determinados contextos, e não exige a fusão das tradições lexicográfica e ortográfica. Isto é, continuarão a existir casos de divergência que dependem não da aplicação de regras, mas da tradição, como *úmido* no Brasil e *húmido* nos restantes países. Outro dos problemas que se colocam deve-se ao fato de os diferentes países se encontrarem em momentos diferentes da solidificação e normalização das suas variedades nacionais.

Por essas razões, a solução adotada no VOC é uma abordagem mista, na qual a tradição lexicográfica e o uso real em contextos escritos da variedade padrão são levados em conta para obter uma fatia representativa do léxico da língua. Por um lado, a tradição desempenha um papel fundamental na determinação das formas existentes e na das que não são alteradas pela reforma ortográfica; por outro lado, para que o VOC seja representativo do português como um todo, língua pluricêntrica de oito países, deve registrar, pelo menos, as formas mais frequentes em uso em todos os países, especialmente aqueles com menor ou inexistente tradição lexicográfica.

Assim, o presente artigo pretende detalhar todos os procedimentos envolvidos na elaboração do VOC, desde a junção dos vocabulários já existentes em Portugal e Brasil, pondo em evidência a tradição lexicográfica portuguesa e os aspectos computacionais atinentes à tarefa, como também a constituição dos *corpora* nos demais países membros, retratando a sua importância em projetos envolvendo o léxico.

## A junção dos Vocabulários de Portugal e do Brasil

O português dispõe já de uma longa tradição lexicográfica, construída, sobretudo, nos últimos dois séculos em Portugal e no Brasil, que deu origem a um acervo extenso, ainda que não comparável ao de outras línguas de relevância mundial. Por isso, o VOC será constituído por entradas com duas origens, que representarão duas partes distintas da obra: uma, correspondente à *memória lexicográfica do português*, composta pelas entradas dos vocabulários ortográficos oficiais já existentes; outra, correspondente ao léxico obtido através da mineração de *corpora*, representativa do léxico efetivamente em uso em todos os países.

---

3 Instituição vinculada à CPLP, que tem como objetivos “a promoção, a defesa, o enriquecimento e a difusão da língua portuguesa como veículo de cultura, educação, informação e acesso ao conhecimento científico, tecnológico e de utilização oficial em fóruns internacionais” (mais informações em <<http://www.iilp.org.cv/>>).

4 No Brasil, o AO90 foi promulgado pelo decreto n. 6.583, de 29/9/2008.

A memória lexicográfica do português será constituída por um subconjunto equilibrado das entradas do *Vocabulário Ortográfico do Português*<sup>5</sup> (VOP), sob a responsabilidade de Margarita Correia, produzido pelo Instituto de Linguística Teórica e Computacional<sup>6</sup> (ILTEC) e oficial em Portugal, e do *Vocabulário Ortográfico da Língua Portuguesa*<sup>7</sup> (VOLP), 5ª ed., coordenada por Evanildo Bechara, da Academia Brasileira de Letras. Ambas as obras são baseadas nas nomenclaturas de dicionários existentes,<sup>8</sup> ainda que os *corpora* venham a desempenhar um papel importante na integração de ambos no VOC, como se verá adiante.

Em qualquer obra dessa natureza, que se proponha a integrar informação provida de várias fontes lexicográficas e com características distintas, uma das principais tarefas passa pela homogeneização das entradas e pelo controle da representatividade de cada uma delas, equilibrando o resultado final. Um primeiro problema que a execução do VOC levanta deve-se ao fato de que as obras que integram a memória lexicográfica do português têm características diferentes quanto aos critérios lexicográficos, de inclusão e de identidade lexical, além de diferentes dimensões e propriedades (o VOLP tem uma nomenclatura muito mais extensa que o VOP, mas inclui muito menos informação para cada entrada). Esse fato impossibilita, portanto, a integração direta das duas obras lexicográficas.

Em face desses problemas, o VOC foi implementado numa plataforma centralizada de edição lexicográfica, que permite a homogeneização das obras, recorrendo a *corpora* para selecionar subconjuntos de dados. Foi escolhida pelo IILP a plataforma usada pelo ILTEC na construção do VOC, o OSLIN<sup>9</sup> (JANSSEN, 2005). Nesse sentido, o VOLP tem de ser inteiramente adaptado a essa plataforma.

O OSLIN (sigla de *Open Source Lexical Information Network*) é uma plataforma digital com base em serviços alojados na web que permite a rápida criação de recursos lexicais, a sua gestão e a permanente inserção de dados. A plataforma contém ferramentas dedicadas a diferentes funções,<sup>10</sup> pensadas para serem usadas por lexicógrafos, que facilitam a gestão e a criação, de forma integrada, a partir de um par *lema – classe gramatical*, de entradas lexicais com outra informação formal associada: paradigma flexional, todas as formas flexionadas, divisão silábica (para fins de translineação), acentuação, remissões para entradas relacionadas (entre elas as variantes diatópicas decorrentes da aplicação do AO90, muito relevantes no caso desta obra), relações funcionais com outras entradas (por ex., todos os nomes deverbais eventivos, nomes de qualidade e advérbios deadjetivais são associados explicitamente) e, dentro em breve, informação sobre a morfologia e a fonética (ASHBY; FERREIRA, 2010).

A integração do VOLP nessa plataforma está sendo feita por uma equipe coordenada por Gladis M. de Barcellos Almeida, na Universidade Federal de São Carlos (UFSCar).

---

5 Disponível em <[www.portaldalinguaportuguesa.org](http://www.portaldalinguaportuguesa.org)>.

6 Conferir em <[www.iltec.pt](http://www.iltec.pt)>.

7 Conferir em <[www.academia.org.br](http://www.academia.org.br)>.

8 Para uma descrição do VOP, consulte-se Correia; Ferreira [no prelo]. As fontes do VOLP não são referidas explicitamente na obra, embora, em comunicações públicas dos seus autores, seja ponto assente que teve como base obras lexicográficas do português.

9 Mais informações em <[www.oslin.org](http://www.oslin.org)>.

10 Para uma descrição mais detalhada do sistema multivalente de edição e manutenção dos recursos lexicais do OSLIN, consulte-se Ferreira et al. (2008).

A nomenclatura daquela obra foi primeiro importada para o sistema OSLIN, seguindo operações de equivalência, desdobramento e sobreposição de classes gramaticais, tornando-as (e às entradas que as contêm) compatíveis com as do VOC. Cada entrada do VOLP, agora já adaptada aos critérios de identidade lexical e ao *tagset* do OSLIN, foi depois marcada como estando ou não presente também no VOP, sendo a sua inclusão no sistema validada automaticamente, nos casos em que isso se verificava. O restante do material lexical do VOLP está sendo sistematicamente cruzado e integrado no VOC com base em interseções entre a obra e *corpora* existentes, de modo a finalizar essa fase e a validar os dados.

Em primeiro lugar, o VOLP foi cruzado com um léxico<sup>11</sup> computacional construído a partir do *corpus* do Núcleo Interinstitucional de Linguística Computacional<sup>12</sup> (NILC) (PINHEIRO; ALUÍSIO, 2003). Esse léxico foi primeiramente, tal como o VOLP, adaptado à estrutura lexical e categorial do VOC, depois de conformado com o Acordo Ortográfico, tarefa executada primeiramente utilizando-se o conversor *Lince*<sup>13</sup> (FERREIRA et al., 2012) e, depois, valendo-se de verificações manuais por padrão através de expressões regulares correspondentes aos contextos em que no português do Brasil existem potenciais mudanças.

O VOLP também está sendo cruzado com o *Corpus Brasileiro*<sup>14</sup> versão 1 (CEPRIL/PUCSP/FAPESP) (BERBER SARDINHA et al., 2009), que, além das operações indicadas acima, é adicionalmente necessário proceder à lematização dos *tokens* constantes do *corpus*, à sua etiquetagem morfossintática e à sua contagem e ordenação por frequência (o léxico do NILC já tinha sido alvo dessas operações, razão pela qual não foram, nesse caso, necessárias).

Depois dessas operações preparatórias, que permitiram a criação de bases de dados lexicais MySQL independentes para cada fonte, foi executado para cada uma delas o mesmo pré-processamento computacional para os pares *lema – classe gramatical* que foi feito para o VOLP, usando ferramentas disponíveis no sistema de administração do OSLIN, de modo a tornar os dados compatíveis com os das outras bases de dados lexicais. A base de dados contendo o VOLP já de acordo com o formato OSLIN é então fornecida, através de cruzamento, com informação de frequência obtida a partir dos *corpora*, de modo a obter um subconjunto de mais de 200 000 entradas do VOLP que estejam atestadas com suficiente frequência nos *corpora* de referência usados. As entradas a serem inseridas no VOC são, pois, sempre atestadas na fonte lexicográfica de referência, o VOLP, e em *corpora*.

## O uso de *corpora* para constituição de novos recursos

Para as restantes variedades do português, no entanto, não é possível adotar o procedimento acima descrito, dado que não existem ainda recursos lexicográficos de referência que as representem.

11 Esse léxico é o que dá suporte ao corretor ortográfico do *Microsoft Word*, versão brasileira.

12 Conferir em <[www.nilc.icmc.usp.br](http://www.nilc.icmc.usp.br)>.

13 O *Lince* é gratuito, de distribuição livre e pode ser obtido no *site* do Portal da Língua Portuguesa <[www.portaldalinguaportuguesa.org](http://portaldalinguaportuguesa.org)>.

14 Conferir em <<http://corpusbrasileiro.pucsp.br>>.

Seguindo as decisões no seio do IILP, dos representantes de todos os países da CPLP, optou-se pela elaboração de *corpora*. Cada Estado Membro da CPLP assegurou a existência de uma equipe nacional com, no mínimo, três indivíduos que executarão, em coordenação com uma equipe central do projeto, as tarefas necessárias ao prosseguimento do objetivo final.

Um primeiro problema a enfrentar para que isso fosse possível relacionava-se com a inexistência sequer de *corpora* de referência representativos dessas variedades. O pouco trabalho feito nesse sentido refere-se a variedades sociais que não se coadunam com o caráter dos recursos que devem ser criados, que se pretendem representativos de variedades diafasicamente relacionadas com a escrita em contexto formal e com a norma de cada país.

Uma primeira e essencial tarefa do projeto era, por isso, a constituição de *corpora* de referência para cada país da CPLP que deles não dispõe. Como se justifica no decorrer deste artigo, isso deverá ser feito atendendo à representatividade, ao balanceamento, à diversidade e ao tamanho, levando ainda em consideração a exequibilidade e os custos, essenciais para permitir a rápida execução do projeto. Assim, foi elaborado e aprovado pelos representantes técnicos dos vários países um conjunto de metas quanto a tipos textuais, sua proveniência e o peso no resultado final.

Cada *corpus* nacional terá no mínimo 30 milhões de palavras, distribuídas por sub-*corpora* com texto literário (20%), jornalístico (25%), legislativo e de sessões parlamentares (25%), técnico (saúde, educação, ambiente, pescas e agricultura: 25%) e de proveniência variada (5%). Essa distribuição assegura que todos os países possam ter as mesmas fontes, todas passíveis de serem obtidas em formato digital, de modo a reduzir o tempo de execução e o custo com recursos humanos que seria necessário para processar texto noutros suportes. Além disso, garantiu-se que o material necessário existe em todos os países, o que determinou os domínios de especialidade do sub-*corpus* técnico indicados acima.

Os *corpora* são obtidos por cada equipe nacional em formato digital, renomeados de acordo com o tipo e a origem da fonte e colocados numa estrutura de pastas hierarquizadas num servidor comum. Cada arquivo é depois processado recursivamente pela equipe central, que o converte para texto simples e sugere elementos não textuais para limpeza, e pela equipe nacional, que procede à validação desse primeiro processamento. Posteriormente, cada texto é dividido em extratos, que são anotados quanto à proveniência e tipo, *tokenizado*, lematizado, etiquetado e finalmente o seu material lexical é convertido num léxico de frequência contendo lema, classe de palavra e índice de frequência.<sup>15</sup>

Cada um desses léxicos de frequência é depois inserido numa base de dados própria no sistema OSLIN, tornando-se uma lista de candidatos à inserção no VOC, e, depois de definida uma linha de corte com base na frequência e distribuição por sub-*corpus*, as entradas são inseridas finalmente no VOC por cada equipe nacional usando a plataforma de gestão do sistema. Para isso, foi integrado na plataforma de gestão do OSLIN um módulo de inserção em massa de entradas lexicais. Esse módulo permite gerar automaticamente todas as propriedades formais de cada palavra a partir do lema e da classe gramatical de cada candidato. Os lexicógrafos da equipe nacional de cada país, tal como ocorreram com as entradas obtidas a partir do VOP e do VOLP, verificam manualmente a legitimidade de toda e

---

15 Para uma descrição detalhada do processo seguido, consultar Almeida e Ferreira (2012).

qualquer entrada e a correção das informações geradas automaticamente pelo sistema, definindo apenas o paradigma flexional de cada entrada, que não pode ser corretamente aduzido pelo sistema.<sup>16</sup>

Cada palavra constante do VOC é marcada explicitamente quanto à sua proveniência e quanto às fontes e países em que se encontra atestada e, nos casos em que esteja disponível, do seu índice de frequência e das variantes já atestadas no Vocabulário.

Está em desenvolvimento um módulo adicional do OSLIN que permite a identificação semiautomática de variantes, que facilitará a difícil tarefa de registrar todas as remissões correspondentes à variação nacional existente entre países, quer a resultante da aplicação das regras do Acordo Ortográfico, nos casos em que permite variação, quer a que não decorre da sua aplicação, dado que o Acordo Ortográfico não unifica a grafia do português, mas apenas a enunciação das suas regras.

### **O papel dos *corpora* para a avaliação e o estabelecimento de um padrão**

A utilização de *corpora* ou grandes quantidades de textos armazenados e selecionados de acordo com determinados critérios para servir de fonte de coleta e/ou análise de itens lexicais não é prerrogativa do mundo moderno. Há muito que se constituem *corpora* para a validação ou definição de uma norma lexical que, na tradição lexicográfica portuguesa e inglesa – para exemplificar com duas grandes línguas pluricêntricas –, viria a ser organizada e publicada em forma de grandes dicionários de língua.

Na língua portuguesa, pode-se citar o *Vocabulário Portuguez e Latino*, elaborado pelo Padre Rafael Bluteau e publicado em oito volumes, entre 1712-1728. O Vocabulário foi o primeiro dicionário para o qual foi fixado um *corpus* (MURAKAWA, 2001). Esse *corpus*, contendo 406 obras, aproximadamente, com autores dos séculos XV a XVII, foi utilizado como exemplário de uso linguístico para as palavras que constavam da nomenclatura do dicionário (MURAKAWA, 2001; 2006). Outro exemplo já no século XIX é o *Diccionario da Lingua Portuguesa*, 2ª edição, de António de Morais Silva, publicado em 1813, o qual também se valeu de um *corpus* (MURAKAWA, 2006) como fonte para a recolha de exemplos.

Na língua inglesa, pode-se referir *A Dictionary of the English Language*, publicado em 1755, para o qual Samuel Johnson reuniu um *corpus* de textos para poder observar as palavras em seu uso autêntico. Esse *corpus* lhe serviu também como fonte de exemplos para os verbetes. Outra importante obra é *The Oxford English Dictionary* (OED), cujo projeto lexicográfico, capitaneado por James Murray, também se valeu de uma grande quantidade de textos, obtidos a partir de uma carta de apelo enviada pelo próprio Murray em 1879, pedindo colaborações. Como resposta a essa carta, mais de 800 voluntários leitores passaram a enviar ao editor tiras de papel com citações que deveriam conter as palavras a eles atribuídas (BIBER; CONRAD; REPPEN, 1998). A reunião desses milhares de tiras de papel constituiu o primeiro *corpus* do OED.

Além de os *corpora* terem um importante papel em grandes projetos lexicográficos orientados por uma metodologia empírica, considera-se que a palavra seja a principal unidade de análise dos estudos realizados em Linguística de *Corpus*, dada a facilidade de sua

---

<sup>16</sup> Para uma descrição detalhada deste módulo, consultar Janssen (2011).

identificação pelas ferramentas computacionais (VIANA, 2011). Em projetos envolvendo o léxico, a palavra é praticamente a porta de entrada para a análise de *corpus*, ou, como afirma Calzolari (1996, p. 3), “all Language Engineering applications require knowledge about words”. Não é sem razão, pois, que projetos e/ou atividades de pesquisa situados no âmbito do léxico tenham no *corpus* seu maior aliado.

O que mudou daqueles projetos lexicográficos fundadores para cá é a concepção de *corpus*, entendido hoje como “a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research” (SINCLAIR, 2005, p. 10). Esse formato eletrônico, promovido pelo advento do computador, interferiu diretamente não só na concepção que se tem de *corpus* como também na sua forma de armazenamento e exploração, já que os recursos oferecidos pela máquina permitiram que grandes quantidades de textos pudessem ser processadas em questão de segundos, fazendo com que muitas hipóteses sobre determinados fenômenos linguísticos pudessem ser testadas rápida e eficientemente.

A moderna noção de *corpus* também carrega consigo requisitos que devem ser fortemente considerados num projeto de *corpus*. São eles: representatividade, balanceamento, diversidade e tamanho (McENERY; WILSON, 1996; KENNEDY, 1998; BIBER et al., 1998; RENOUF, 1998; SINCLAIR, 2005).

Dentre todos esses requisitos, a representatividade é crucial, haja vista que um *corpus* representativo tende a ser bem balanceado, ter boa diversidade e tamanho adequado aos objetivos da pesquisa. Assim, para a construção do VOC, especial atenção foi dada à representatividade, já que o vocabulário ortográfico deverá servir como importante instrumento de normalização lexical da língua portuguesa.

Quanto ao caso específico das variedades menos representadas do português, há que se levar em conta a discrepância entre as condições linguísticas da língua no conjunto dos oito países membros da CPLP: Angola, Brasil, Cabo Verde, Guiné-Bissau, Moçambique, Portugal, São Tomé e Príncipe e Timor-Leste.

Com exceção de Portugal e Brasil, que já têm *corpora*, grandes dicionários tradicionais (instrumentos linguísticos que permitem a construção e são muitas vezes a base de vocabulários ortográficos) e são países onde o português é majoritariamente a língua mais usada, os demais países (Angola, Cabo Verde, Guiné-Bissau, Moçambique, São Tomé e Príncipe e Timor-Leste) vivem uma situação linguística bastante diferente, uma vez que grande parte dos seus cidadãos não tem o português como primeira língua.

Observe-se, no Quadro 1, a quantidade de línguas que convive com o português nos Países Africanos de Língua Oficial Portuguesa (PALOP) e em Timor-Leste, numa lista necessariamente incompleta dada a inexistência de consenso em relação a cada país e, em alguns casos, à falta de dados e de estudos recentes. Apresenta-se, também, no mesmo Quadro, a população aproximada de cada um e o ano de sua independência, para se ter uma noção um pouco mais ampla do cenário linguístico, demográfico e político no qual se insere este projeto do VOC.

**Quadro 1: Línguas faladas nos PALOPs e em Timor-Leste, além do português<sup>17</sup>**

Estados-membros	Línguas faladas (além do português)	População	Ano da independência
Angola	<i>Kikongo, kimbundo, tchokwe, umbundo, mbunda, kwanyama, nhaneca, fiote, nganguela, entre outras</i>	18.498.000	1975
Cabo Verde	Crioulo de Cabo-Verde (ou caboverdiano)	499.796	1975
Guiné-Bissau	Crioulo da Guiné-Bissau (ou <i>kriol</i> ), balanta, mandjaco, fula, entre outras	1.449.230	1974
Moçambique	<i>Cicopi, cinyanja, cinyungwe, cisenga, cishona, ciyao, echuwabo, ekoti, elomwe, gitonga, maconde (ou shimakonde), kimwani, macua (ou emakhuwa), memane, suaíli (ou kiswahili), suazi (ou swazi), xichangana, xironga, xitswa e zulu</i>	20.366.795	1975
São Tomé e Príncipe	Forro, principense (ou <i>lunguyé</i> ) e angolar.	187.356	1975
Timor-Leste	Tétum, entre muitas outras línguas austronésias e algumas papuanas	1.066.582	1975 <sup>18</sup> e 2002 <sup>19</sup>

Tendo em vista, pois, esse cenário, antes de se pensar em constituir vocabulários nacionais de cada um desses países, é preciso primeiro constituir os *corpora* que representem, de fato, a norma lexical de cada um deles; *corpora* que incluam tanto palavras comuns a todos ou a alguns países como aquelas que são específicas de apenas um deles.

Além desse aspecto, foi preciso ter em conta que os *corpora* compilados fossem comparáveis em tamanho, em distribuição por gêneros e em relação ao peso de cada um dos gêneros no conjunto final (ALMEIDA; FERREIRA, 2012), de modo que o VOC, depois de pronto, evidenciasse uma gestão política de língua descentralizada, despolarizada (saindo do eixo Portugal e Brasil) e com participação ativa de todos os países-membros.

Crítérios iniciais que orientaram o projeto dos *corpora* foram: i) textos escritos e em contexto formal; ii) prioridade para textos já em formato digital. A partir daí, iniciou-se um trabalho coletivo, contando com a cooperação das equipes nacionais de cada país, de forma a estabelecer as características dos *corpora*. A partir dessas primeiras decisões, foram-se desenhando os *corpora*, de modo a atender o principal requisito, qual seja, a representatividade.

Ser **representativo** significa que um *corpus* deve ser elaborado de forma a espelhar determinadas características linguísticas da comunidade cuja língua está sob análise (SINCLAIR, 2005). Daí a importância de se fazerem escolhas adequadas para que o *corpus* possa realmente refletir comportamentos linguísticos. Questões que devem ser feitas durante a seleção dos textos são: quais documentos? Quais tipos de textos? Quais gêneros textuais? Enfim, o que de fato representa os usos linguísticos de uma comunidade? No

17 As informações foram obtidas no portal da CPLP (<http://www.cplp.org/>); nos portais dos respectivos governos (Angola - <http://www.governo.gov.ao/>; Cabo Verde - <http://www.governo.cv/>; Guiné-Bissau - <http://www.anpguinebissau.org/>; Moçambique - <http://www.portaldogoverno.gov.mz/>; São Tomé e Príncipe - <http://www.presidencia.st/>; Timor-Leste <http://timor-leste.gov.tl/>), na Wikipédia, especificamente para conferir os dados populacionais e em Batoréo e Casadinho (2009).

18 Independência de Portugal.

19 Fim da ocupação indonésia.

caso dos *corpora* para o VOC, a pergunta teve de ser ainda mais específica: quando se escreve em português num contexto formal nesses países, quais palavras são empregadas?

Diretamente associado à representatividade, está o conceito de **balanceamento**. Embora seja um conceito vago, de acordo com Sinclair (2005), é preciso ter em mente que o *corpus* deve ter um equilíbrio entre gêneros discursivos (informativo, científico, etc.), tipos de textos incluídos (artigo, editorial, entrevista, dissertação, etc.), temas (pesca, agricultura, saúde, educação, etc.), ou até mesmo títulos, ou autores. O ideal seria que se conseguisse levar em conta todas essas categorias, mas sempre atendendo às demandas da pesquisa que se pretende realizar.

Se o *corpus* for cuidadosamente balanceado, ele terá uma boa **diversidade**. A propósito disso, Biber et al. (1998, p. 248) assinalam que “there are important differences in the use of lexical, grammatical and discourse features across different varieties of language”. O que lhe dá argumentos para afirmar que o conceito de “língua geral” é hipotético, dado que cada gênero discursivo tem seus próprios padrões de uso. Nesse sentido, um *corpus* representativo deve conter uma diversidade de gêneros, tipos de textos e assuntos, pois a frequência de muitas palavras pode variar de acordo com o assunto (BIBER et al., 1998).

Um *corpus* pensado para ser representativo, balanceado e diversificado tem, em geral, um **tamanho** adequado ao tipo de pesquisa que se vai realizar e à metodologia a ser adotada na pesquisa (SINCLAIR, 2005). Quando se fala em tamanho de *corpus*, não se trata somente do número total de ocorrências (*tokens*) e de palavras diferentes (*types*), mas com quantas categorias (gêneros discursivos, tipos de textos, assuntos, títulos, autores, etc.) um *corpus* deve contar, quantas amostras de cada categoria e quantas palavras existem dentro de cada amostra (KENNEDY, 1998).

Assim, os *corpora* que estão sendo compilados nos PALOPs e em Timor-Leste atendem a esses requisitos de representatividade, balanceamento, diversidade e tamanho, acima especificados.

Todavia, para a gestão descentralizada de uma língua pluricêntrica, tão importante quanto esses requisitos é o envolvimento de todas as equipes nacionais no projeto; afinal, quem mais sabe responder às questões que devem ser feitas durante a seleção de textos para compor um *corpus* são os falantes de cada um desses países.

## **O uso de *corpora* no VOC**

Tendo em conta o que foi dito, os *corpora* assumem necessariamente um papel central na criação de novos recursos lexicográficos para a língua portuguesa quando tomada como um todo, como é o caso do VOC.

Por um lado, pela sua natureza, herança e princípios metodológicos, os dicionários existentes para o português são mais um perpetuar da tradição lexicográfica, ainda que em alguns casos parcialmente atualizada, do que um verdadeiro repositório atualizado das palavras de fato em uso nos países de língua portuguesa, o que afeta a sua representatividade. Por isso, uma primeira função dos *corpora* no VOC é a de validação dos dados constantes da tradição lexicográfica portuguesa.

Além disso, dado que a reforma ortográfica não torna homogênea a escrita de muitas palavras com formas ortográficas já antes divergentes (*úmido* no Brasil, mas *húmido* nos

restantes países, como já mencionamos no início deste artigo), qualquer repositório lexical multinacional para o português deve identificar, marcar e relacionar por meio de remissões as diversas variantes, o que só é possível executar num curto espaço de tempo recorrendo à análise semiautomática das divergências entre *corpora* representativos de cada variedade.

Por último, atendendo à falta de cobertura lexicográfica atual para muitas das variedades nacionais do português e à inexistência de *corpora* representativos dessas variedades, a criação e o uso de tais recursos de base são essenciais para captar a realidade da língua escrita em contexto formal em cada país e proceder à sua integração num recurso que a represente, fixando a ortografia do seu léxico. Nesse sentido, os *corpora* servem, pois, de fonte primária para as entradas da obra em criação.

### Considerações finais

Apesar de todo o trabalho desenvolvido, sobretudo no último século e meio, e da grande evolução trazida pelas publicações mais recentes de dicionários, a lexicografia de referência do português tem ainda claras lacunas face às outras línguas de relevância mundial. São particularmente prementes três problemas: a falta de recursos lexicais normalizadores disponíveis que permitam o processamento computacional da língua portuguesa; a falta de recursos feitos com base ou tendo em conta a informação obtida a partir de *corpora*; a falta de recursos representativos da diversidade do português, que possam agir nacionalmente como normalizadores em países que não dispõem neste momento de recursos próprios.

Além dessas questões puramente lexicográficas, o português carece, além disso, de um vocabulário ortográfico comum, prerrogativa do Acordo Ortográfico firmado por todos os países e representante da vontade política manifestada de manter a unidade da língua portuguesa.

O VOC pretende responder a esses problemas, criando um novo recurso multivalente que reaproveita os recursos já disponíveis e cria, quando necessário, novos recursos. O projeto, ainda em fase de desenvolvimento, dará os seus primeiros resultados em 2014. Os resultados serão, a partir dessa data, disponibilizados gratuitamente a partir de uma interface *on-line* de acesso livre.

### REFERÊNCIAS

ALMEIDA, G. M. B.; FERREIRA, J. P. *Manual para a elaboração de corpora*. Com vista à organização dos Vocabulários Ortográficos Nacionais dos países integrantes da CPLP. Lisboa (Portugal): ILTEC; São Carlos (SP, Brasil): NILC, 2012.

ASHBY, S.; FERREIRA, J. P. The Role of Morphology in Generating High-Quality Pronunciation Lexica for Regional Variants of Portuguese. In: BRANCO, A.; KLAUTAU, A.; VIEIRA, R.; LIMA, V. L. S. de; PARDO, T. A. S. (Ed.). *Computational Processing of the Portuguese Language – Lecture Notes in Artificial Intelligence*, v. 6001. Berlin Heidelberg: Springer-Verlag, 2010. p. 162-165.

BATORÉO, H.; CASADINHO, M. O português, uma língua pluricêntrica: o caso de Timor-Leste. *Revista Portuguesa de Humanidades, Estudos Linguísticos*, Braga: Universidade Católica Portuguesa de Braga, 13-1, p. 63-79, 2009. Disponível em: <[http://www.catedraportugues.ueem.mz/lib/docs/bib\\_timor/Batoreo\\_Casadinho\\_2009.pdf](http://www.catedraportugues.ueem.mz/lib/docs/bib_timor/Batoreo_Casadinho_2009.pdf)>. Acesso em: 12 ago. 2012.

BERBER SARDINHA, T.; MOREIRA FILHO, J. L.; ALAMBERT, E. The Brazilian Corpus: A one-billion word online resource. In: MAHLBERG, M.; GONZÁLEZ-DÍAZ, V.; SMITH, C. (Ed.). *Proceedings of the Fifth Corpus Linguistics Conference, CL2009*. Liverpool: University of Liverpool, UK, 20-23 July 2009.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus Linguistics – Investigating Language Structure and Use*. Cambridge, U.K.: Cambridge University Press, 1998.

CALZOLARI, N. Lexicon and Corpus: a multi-faceted interaction. In: GELLERSTAM, M.; JARBORG, J.; MALMGREN, S.-G.; NOREN, K.; ROGSTROM, L.; PAPMEHL, C. R. (Ed.). *Euralex '96 Proceedings*. Goteborg: Göteborgs Universitet, 1996. p. 3-16.

CORREIA, M.; FERREIRA, J. P. Vocabulário Ortográfico do Português Correia, Margarita e Ferreira, José Pedro. In: *Actas del V Congreso Internacional de Lexicografía Hispánica*, Madrid, 25-27 jun. 2012 (no prelo)

FERREIRA, J. P.; BARBOSA, S.; JANSSEN, M. Mordebe Admin: A Lexical Management System. In: *Proceedings of the 13th Euralex International Congress, Barcelona, 2008*. Barcelona: Institut Universitari de Lingüística Aplicada - Universitat Pompeu Fabra; Documenta Universitaria, 2008. p. 351-357.

FERREIRA, J. P.; LOURINHO, A.; CORREIA, M. Lince, an End User Tool for the Implementation of the Spelling Reform of Portuguese. In: CASELI, H. M.; VILLAVICENCIO, A.; TEIXEIRA, A. J. S.; Perdigão, F. (Ed.). *Computational Processing of the Portuguese Language - 10th International Conference, PROPOR 2012 Proceedings*. Lecture Notes in Computer Science 7243. Berlin, Heidelberg: Springer, 2012. p. 46-55.

JANSSEN, M. Open Source Lexical Information Network. In: BOUILLON, P.; KANZAKI, K. (Ed.). *Proceedings of the Third International Workshop on Generative Approaches to the Lexicon*, May 19-21 2005. Genebra: École de Traduction et d'Interpretation – Université de Genève, 2005. p. 79-106.

\_\_\_\_\_. Computer-Aided Inflection for Lexicography Controlled Lexica. In: KOSEM, I.; KOSEM, K. *Electronic Lexicography in the 21<sup>st</sup> Century New Applications for New Users – Proceedings of eLex 2011*. Liubliana: Trojina, Institute for Applied Slovene Studies, 2011. p. 96-105.

KENNEDY, G. *An Introduction to Corpus Linguistics*. London/NY: Longman, 1998.

McENERY, T.; WILSON, A. *Corpus linguistics*. Edinburgh: Edinburgh University Press, 1996.

MURAKAWA, C. A. A. Tradição lexicográfica em língua portuguesa. In: OLIVEIRA, A. M. P.; ISQUERDO, A. N. (Ed.). *As ciências do léxico: lexicologia, lexicografia e terminologia*. 2. ed. Campo Grande: Ed. UFMS, 2001. p. 153-159.

\_\_\_\_\_. *Antônio de Moraes Silva: lexicógrafo da língua portuguesa*. Araraquara: Laboratório Editorial FCL/UNESP; São Paulo: Cultura Acadêmica Editora, 2006. 228p.

PINHEIRO, G. M.; ALUÍSIO, S. M. *Cópus Nilc: descrição e análise crítica com vistas ao projeto Lacio-Web*. NILC-TR-03-03, 2003. Disponível em: <<http://www.nilc.icmc.usp.br/lacioweb/downloads/NILC-TR-03-03.zip>>. Acesso em: 8 set. 2012.

RENOUF, A. (Ed.). *Explorations in Corpus Linguistics*. Amsterdam: Rodopi, 1998.

SINCLAIR, J. *Corpus and Text – Basic Principles*. In: Wynne, M. (Ed.). *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow: p. 1-16, 2005. Disponível em: <<http://ota.ahds.ac.uk/documents/creating/dlc/chapter1.htm>>. Acesso em: 12 ago. 2012.

VIANA, V. Linguística de *corpus*: conceitos, técnicas & análises. In: VIANA, V.; TAGNIN, S. E. O. (Ed.) *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2011. p. 22-92.

# Estudo lexical dos nomes de sintomas e de doenças nos séculos XVII e XVIII: comparação entre o português e o francês

(Étude lexicale des noms de symptômes et de maladies aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles: comparaison entre le portugais et le français)

**Mariana Giacomini Botta<sup>1</sup>**

<sup>1</sup>GMHP – Grupo de Morfologia Histórica do Português (DLCV – FFLCH/USP)  
Faculdade de Ciências e Letras de Araraquara (FCL/CAr – Unesp)  
Université Paris 3 Sorbonne Nouvelle

marianabotta@gmail.com

**Résumé :** Le but de cette étude est la vérification et l'analyse de quelques unités lexicales appartenant à la terminologie de la médecine aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles. Même si à l'époque la médecine s'élevait encore en science, certains termes décrivent les symptômes, les maladies, les diagnostics, les traitements et les médicaments s'utilisaient déjà couramment. Pour étudier ces emplois, nous nous basons sur un corpus de textes de la *Gazeta de Lisboa*, entre 1715 et 1800, et de la *Gazette de France*, entre 1631 et 1786. Étant donné que l'échange d'informations entre les pays et que la traduction des nouvelles étaient des pratiques quotidiennes dans les journaux de l'époque, la présente analyse de la transition des termes de la médecine entre deux langues de pays différents, dans un contexte où les études sur santé se constituaient en science, semble intéressante et utile aussi bien en ce qui concerne la terminologie, que la traductologie et la linguistique historique.

**Mot-clés:** Linguistique historique; terminologie; médecine; gazettes; XVII<sup>e</sup> et XVIII<sup>e</sup> siècles.

**Resumo:** A proposta deste trabalho é a verificação e a análise de algumas das unidades lexicais que integram a terminologia da medicina nos séculos XVII e XVIII. Embora naquela época a medicina ainda estivesse se estabelecendo como ciência, termos que designam sintomas, doenças, diagnósticos, tratamentos e medicamentos podem ser observados em uso na linguagem corrente. Para se estudar esses usos, trabalha-se com unidades extraídas de um *corpus* formado por textos da *Gazeta de Lisboa*, entre 1715 e 1800, e da *Gazette de France*, entre 1631 e 1786. Considerando-se que a troca de informações entre países e que a tradução de notícias eram práticas cotidianas das gazetas daquela época, o estudo sobre a transição de termos da medicina entre as línguas de duas diferentes nações, em um período em que os estudos da saúde se constituíam como ciência, mostra-se interessante e útil tanto para a terminologia, quanto para a tradutologia e para a linguística histórica.

**Palavras-chave:** Linguística histórica; terminologia; medicina; gazetas; séculos XVII e XVIII.

## Introdução

A medicina é uma das mais antigas atividades do homem, e estima-se que tenha surgido entre 460 e 377 a.C., com Hipócrates, que é considerado o pai da medicina. Na Idade Média, a arte de curar ficava a cargo de religiosos e de barbeiros, estes últimos por saberem lidar com a navalha e, assim, poderem drenar abscessos. Foi somente a partir da Era Moderna que surgiram pessoas que se ocupavam apenas da saúde humana.

Nos séculos XVII e XVIII, não apenas a medicina, mas também a química e a farmácia ainda não estavam estabelecidas como ciência e baseavam-se em conhecimentos empíricos e supersticiosos, mesclando-se à alquimia. O surgimento de estudos sobre as drogas medicinais

que eram levadas da Índia e da América e também sobre a importância da higiene na prevenção de epidemias contribuiu para a mudança desse panorama.

Com o surgimento das primeiras gazetas impressas, a partir da segunda metade do século XVI, notícias de conteúdo médico começaram a ser publicadas, e uma parcela da população passou a ter acesso a termos pouco usados anteriormente. Isso pode ser verificado nas páginas de jornais impressos dos séculos XVII e XVIII, como o francês *Gazette de France*, e o português *Gazeta de Lisboa*.

Naquele tempo, em que os regimes monárquicos dominavam o cenário político da Europa, a vida pessoal de reis, imperadores e religiosos era de interesse geral. Considerados como fatos políticos, os relatos sobre a saúde dos monarcas são recorrentes nestas publicações, e termos como febre, médico, enfermo, doença/enfermidade, parto/nascimento, entre outros, integram seus textos.

Com o objetivo de fazer um levantamento de quais eram esses termos e estudar seus usos, propõe-se nesse trabalho uma análise da terminologia da medicina neste período<sup>1</sup>, usada em textos informativos publicados nos primeiros jornais impressos da França e Portugal: a *Gazette de France*, que surgiu em 1631, criada por Théophraste Renaudot, e a *Gazeta de Lisboa*, publicada a partir de 1715 sob a direção de José Freire de Monterroyo Mascarenhas. Os dois veículos de comunicação informavam sobre acontecimentos de toda a Europa e também da Ásia, África e América. As notícias eram produzidas por uma rede de correspondentes espalhados por diversas localidades e também eram traduzidas de gazetas de outros reinos.

Conjunto das unidades lexicais disponíveis de uma língua, o léxico é um inventário aberto, que se amplia na medida em que as comunidades linguísticas aperfeiçoam seus conhecimentos sobre a realidade e criam novas técnicas e ciências. Desta maneira, o léxico de uma comunidade documenta e acumula o conhecimento humano partilhado por uma sociedade.

A sofisticação do conhecimento humano e o desenvolvimento científico têm como resultado o surgimento de sistemas léxicos complexos, que são chamados de terminologias. Estas são uma necessidade do desenvolvimento das ciências, que precisam de novos termos para nomear os conceitos que surgem, o que resulta também na ampliação do repertório de signos lexicais de uma língua. Assim, entende-se a terminologia como um conjunto de termos rigorosamente definidos, que são específicos de uma ciência, de uma técnica, de um domínio particular da atividade humana.

Para evitar ambiguidades e mal-entendidos, toda ciência necessita de um conjunto de termos, definidos com rigor, por meio dos quais ela procura designar as noções que lhe são úteis. Em linguística, a terminologia é a disciplina que estuda sistematicamente a rotulação e a designação de conceitos particulares a um ou vários assuntos ou campos de atividade humana, por meio de pesquisa e análise dos termos em contexto, com a finalidade de documentar e promover seu uso correto. Portanto, a terminologia tem por objeto teórico as denominações dos objetos ou conceitos utilizados pelos diferentes domínios do saber,

---

<sup>1</sup> Este artigo, que traz o tema tratado em simpósio durante o 60º Seminário do GEL, é uma versão resumida e revisada do trabalho apresentado no evento “Traduction, terminologie, rédaction technique : des ponts entre le français et le portugais. En hommage au Professeur Armelle Le Bars, nos dias 13 e 14 de janeiro de 2011, em Paris, França.

o funcionamento na língua das unidades terminológicas, os problemas de tradução, de classificação e de documentação dessas unidades. Consiste ainda no estudo da escolha e dos usos dos termos que fazem parte dos vocabulários especializados. Em terminologia, entende-se um termo como a combinação indissociável de uma denominação e um conceito.

Durante todo o processo de desenvolvimento e de estabelecimento da medicina como ciência, uma linguagem própria foi desenvolvida, e esta até hoje é muitas vezes mal compreendida pelos leigos. Isso se deve ao fato de os termos médicos, além de serem formados a partir de radicais, prefixos e sufixos gregos e latinos (na sua maioria), ainda provêm (em menor número) de elementos vernáculos ou procedentes de outros idiomas. Considerando-se que a tradução era prática corrente nas gazetas dos séculos XVII e XVIII, acredita-se que um estudo sobre a transição de termos da medicina entre duas línguas de dois diferentes países, em um período em que os estudos da saúde se constituíam como ciência, seja interessante e útil tanto para a terminologia, quanto para a tradutologia e para a linguística histórica. Trabalha-se com uma terminologia de orientação descritiva, fundamentada pelos princípios da linguística, como a desenvolvida por Cabré (1999) e na perspectiva da socio-terminologia, definida, entre outros, por Gaudin (2003).

### **Composição do *corpus*, recolha das unidades e metodologia de análise**

As unidades em análise foram extraídas de um *corpus* formado por textos 12 edições do jornal *Gazeta de Lisboa* (entre 1715 e 1800) e 14 edições da *Gazette de France* (de 1631 a 1786), ou seja, abrangendo os anos finais do século XVII e todo o século XVIII. Foram encontradas mais de 80 unidades que designam noções que fazem parte da linguagem médica da época, entre sintomas, tratamentos, procedimentos e nomes de doenças. Destas, foram selecionadas 12 unidades, dez delas designando doenças, escolhidas devido à sua proximidade semântica e à existência de correspondência entre as línguas portuguesa e francesa.

Cada uma das unidades selecionadas foi analisada separadamente e o seu emprego nos contextos dos quais foram retiradas foi comparado com as informações de dicionários da época. Foram consultados o *Dictionnaire de la langue française*, de Émile Littré (1863, 1ª edição), o *Thresor de la langue françoise, tant ancienne que moderne*, de Jean Nicot (TLF, 1606), o *Dictionnaire de l'Académie française* (1694; 1762; 1798 – 1ª, 4ª e 5ª edições), o *Vocabulario portuguez & latino*, de Raphael Bluteau (1712 – 1728) e o *Diccionario da lingua portugueza*, de Antonio Moraes Silva (1813).

### **O estudo da terminologia a partir de um *corpus* histórico de língua corrente**

A Terminologia é uma disciplina inter e transdisciplinar, que se encarrega da tarefa de analisar a expressão lexical formal da organização dos conhecimentos científicos, técnicos e tecnológicos, ou seja, os termos técnico-científicos.

O termo, ou unidade terminológica, é um componente constitutivo da produção do saber que possui características que favorecem a precisão vocabular da comunicação técnico-científica. Qualquer unidade lexical pode se tornar um termo quando se adapta aos propósitos de uma determinada área do saber. Além disso, deve-se considerar que os termos sofrem os mesmos processos e transformações típicos do sistema linguístico, equiparando-se, assim, a uma unidade do léxico comum.

Embora os estudos terminológicos geralmente focalizem o funcionamento dos termos no momento presente, estudar o léxico de uma língua, e de uma língua de especialidade, em um determinado período pode permitir a captação de parte de sua história, que é alterada a todo momento pela dinâmica da renovação lexical. Para Baudet (1988, p. 856), a perspectiva diacrônica é fundamental aos estudos terminológicos, pois só é possível se estudar as línguas das ciências estudando o seu desenvolvimento. Por se ocupar com o estudo dos termos, a Terminologia trabalha com vocábulos que veiculam significações socialmente regulamentadas e inseridas em práticas institucionais.

Desta forma, segundo Gaudin (2003, p. 11), “o estudo diacrônico dos termos diz respeito à história das ciências, das técnicas, dos discursos socialmente regulamentados e à história das ideias”. Assim, ainda segundo este autor, os termos são ao mesmo tempo denominações e instrumentos de trabalho e de conceptualização, e “são tomados em momentos discursivos e históricos que alteram seus sentidos e os inscrevem na gênese e na transformação dos conhecimentos e do mundo” (GAUDIN, 2003, p. 205).

## Análise dos dados

Por meio da leitura dos textos do *corpus* e da delimitação do objeto de estudo às unidades usadas na denominação de doenças, constatou-se que duas noções são fundamentais para se entender qual era o conhecimento da época sobre o funcionamento do corpo humano e sobre a ocorrência de doenças. Apesar de não designarem propriamente doenças, as unidades **humor** e **vapor** faziam parte da terminologia da medicina da época e estavam na base das crenças sobre como e porque as doenças afetavam os corpos.

A unidade **humor** possuía sentido distante do atual e, em contextos com temática sobre a saúde, referia-se, ainda seguindo a teoria hipocrática, aos líquidos e fluidos presentes no corpo humano, como o sangue, a saliva, a lágrima, o leite e o esperma, entre outros, como pode ser visto na ocorrência:

- (1) Ex. e. Rev. Arcebispo Bispo do Algarve lançou do braço copia de **humor** sanguineo. (GDL,<sup>2</sup> 3 de abril de 1755)

Esse significado era o mesmo tanto em português quanto em francês, como apontam os dicionários da época:

[...] líquida substância nas plantas ou nos corpos dos animais [...] pelo humor não se entende apenas os quatro humores do corpo, como é o sangue, fleuma, colera e melancolia, mas todas as mais humanidades, como o leite e o esperma e ainda os humores recrementicios como é a saliva, as lagrimas, o soro do sangue, os quais ou por copia ou por vicio ofendem a saude. (BLUTEAU, 1712-1728)

Substance liquide sécrétée par un organisme vivant; Liquide normalement présent dans un organisme vivant, généralement humain ou animal (sang, salive, p. ex.). (TLF, 1606)

A partir desse conhecimento, pensava-se que a maioria das doenças era causada pelo excesso ou pela degeneração dessas substâncias nos corpos, como afirma Bluteau:

---

<sup>2</sup> A sigla GDL é usada neste artigo para indicar os exemplos extraídos do *corpus* da *Gazeta de Lisboa*.

[...] que por copia ou por vicio ofendem a saude. Por vicio quando o sangue é demasiado, sobejando nas veias, a fleima obstruindo os poros, o soro redundando por todas as veias. Por vicio quando o sangue apodrece, a fleima é salgada, a colera e a melancolia é podre ou requeimada, o esperma corrupto, o soro viciado, o leite grumoso, a saliva salgada, a urina acre, as lagrimas mordazes. (BLUTEAU, 1712-1728)

Pelas ocorrências do *corpus*, percebe-se que, ao contrário, a unidade **vapor** era usada para designar um tipo de sintoma:

- (2) Hum agradável, e eficaz remedio, entre todos os mayores, para defender o coração de todos os **vapores**, melancolicas, ancias, desmayos, palpitaçoens, e de toda outra malignidade de que for acometido [...]. (GDL, 5 de julho de 1725)

No dicionário de Bluteau (1712-1728), no verbete **vapor** encontra-se a seguinte informação: “nos animais se levanta do estômago e dos intestinos vapores, que ofendem o cérebro”, o que levaria a crer se tratar de gases / flatulência. Mas o Dicionário da Academia Francesa (1694; 1762; 1798) mostra que não se conhecia ao certo o que eram os tais vapores: “Exhalaison supposée monter du sang et des autres humeurs jusqu’au cerveau (ex. vapeur maligne); Étourdissements, vertiges, migraines, malaises divers”.

Além dessas duas unidades, foram selecionadas outras dez unidades, designando nomes de doenças, a maioria, acreditava-se, era causada pela desordem dos **humores** no corpo ou pela ocorrência dos **vapores**. A lista desses termos pode ser verificada a seguir:

**Tabela 1: Nomes de doenças que ocorrem nas gazetas dos séculos XVII e XVIII<sup>3</sup>**

Português	Francês
fluxo / defluxo	fluxion
Hemorragia	hémorragie
Gota	goutte / podagre
Apoplexia	apoplexie
Gonorreia	gonorrhé(e) / verole
Bexigas	petite verole / pourpre
Tisica	phtisie
sezonismo / febre terçã	fièvre tierce
pedra nos rins	pierre dans les reins / calcul rénal
Aborto	avortement / fausse couche / couche avant terme

Sabe-se que a maioria dessas doenças e sintomas já eram conhecidos e eram descritos desde a antiguidade. O interesse deste trabalho está em observar como este conhecimento era exposto nos primeiros jornais impressos e nos dicionários dos séculos XVII e XVIII. As análises das unidades são apresentadas nas seções seguintes.

<sup>3</sup> Em alguns casos, mesmo sem ter sido encontrada a mesma ocorrência nas publicações dos dois países, foi estabelecida a correspondência na outra língua de acordo com os dados encontrados nos dicionários da época já citados.

## Fluxos e fluxões

Por **fluxos** e **fluxões** designavam-se diferentes tipos de doenças e sintomas, que vão desde inchaços localizados, inflamações e tumores a diarreias, como pode ser visto pelos excertos extraídos do *corpus*:

- (3) [...] aquelles, que servem para curar zunidos dos ouvidos, dor de dentes, gonorrhœas, morreimas, **fluxos** de sangue, preservar de aborto, vertigens, dores de cabeça, confortar os nervos, desfazer tumores, confortar, e fazer alvos os dentes, rebater os vapores do utero, provocar os mezes, matar lambrigas, etc. (GDL, 5 de julho de 1725)
- (4) Il est entièrement delivré de la fièvre, & la **fluxion** sur la main droite est fort diminuée, de sorte qu'il pourra estre en estat de signer bien tost. (GDF, 02 de janeiro de 1700)
- (5) [...] mais la nuit du Dimanche il se trouva incommode d'un **flux** de ventre, qui luy causa une fièvre assez forte. (GDF, 04 de setembro de 1700)

Segundo os dicionários consultados, **fluxão** era a “transmissão de humor de uma parte a outra ou descarga de humores, que redundando nos vasos, caem em alguma parte do corpo”, ou ainda, “fluxo de humores supérfulos que a natureza descarrega por câmaras ou diarreia” (BLUTEAU, 1712-1728); “correnteza ou corrente de liquido ou humor que corre para alguma parte do corpo” (MORAES SILVA, 1813); “Écoulement d'humeurs malignes sur quelque partie du corps” (DAF, 1762); “Afflux de sang ou d'autres liquides sur certains tissus qui se tuméfient. Inflammation de la paroi thoracique, de la plèvre et du poumon; Gonflement inflammatoire des gencives ou des joues provoqué par une infection dentaire” (TLF, 1606).

## Hemorragia

Nos dicionários consultados, **hemorragia** se refere à perda de sangue, principalmente pelo nariz ou gengivas. No *corpus*, foi encontrada a seguinte ocorrência:

- (6) Escrevem de Halifax que em Harley Green, perto daquela cidade, se descubrio ha pouco huma nascente de agua mineral, de que se espera resulte grande utilidade: he ella muito ferrea, e parece sera efficaz contra as relaxações da fibra, **hemorragias**, fiatos. (GDL, 16 de fevereiro de 1790)

É interessante ressaltar que os dicionários já trazem a informação de se tratar de um termo da medicina. Neles encontram-se as seguintes informações: “perda de sangue por qualquer parte do corpo” (BLUTEAU, 1712-1728); “fluxo de sangue” (MORAES SILVA, 1813); “Écoulement de sang à la suite de la rupture d'un vaisseau sanguin” (TLF, 1606); “Perte de sang par le nez, par une plaie” (DAF; 1762).

## Gota

A **gota** é uma doença muito antiga, descrita desde o Império Romano e, no século V a.C., Hipócrates utilizava o nome podagra para designar este mal<sup>4</sup>. A partir do século IX, a **gota** passa a ser a designação de qualquer humor nocivo destilado nas juntas e nos órgãos, e apenas a partir do século XV corresponde ao que era conhecido por podagra, substituindo esse termo durante o Renascimento em quase todas as línguas europeias. No

4 Informação extraída de: Bouret, Treves e Marc (2009).

século XVII, o francês Guillaume de Baillou<sup>5</sup> (1583-1616) distingue a **gota** do reumatismo e da artrite e, em 1683, o médico inglês Thomas Sydenham (1624-1689) descreve a doença em seu *Tractatus de Podagra et Hydrope*.

As ocorrências encontradas no *corpus* provam se tratar de uma doença conhecida na época, pois não são fornecidas maiores informações sobre ela nos textos:

- (7) Il commença néanmoins des le 19, a estre considerablement soulagé, & présentement on assure qu'il est entièrement sans fièvre & comme la fluxion a commencé à se jeter sur la main & sur les genoux, on espere que ce nouvel accident pourra finir par une attaque de **goute**. (GDF, 16 de janeiro de 1700)
- (8) Sua mag. chr. ha muytos dias que se acha molestado de **gota**, & de dores de siatica com tanta força, que lhe interromperão o sono algumas noytes, & o obrigaraõ a estar na cama [...]. (GDL, 28 de setembro de 1715)
- (9) O Emperador teve a 8 deste mez hum ataque de **gota**, mas tam ligero, que no dia seguinte se achou em estado de poder ir a caça. (GDL, 3 de março de 1740)
- (10) Ces jours derniers, le Souverain Ponifice, a eu une légère attaque de **goutte**, mais ils ne s'en ressent plus actuellement. (GDF, 28 de março de 1768)

Nos dicionários da época também se encontram informações sobre esta enfermidade: “doença causada de acrimonia do humor que cai nas juntas e faz muita dor [...]; todo mal da gota procede das gotas do humor que corre e se embebe nas juntas” (BLUTEAU, 1712-1728); “doença que consiste em fixar-se nas articulações das mãos ou pés o humor grosso e cru que a natureza arroja às extremidades do corpo” (MORAES SILVA, 1813); “La goutte és pieds, Podagra” (TLF, 1606); “Fluxion acre & douloureuse qui tombe ordinairement sur les jointures” (DAF; 1762).

## Apoplexia

**Apoplexia** é o termo em desuso para o que hoje se conhece por acidente vascular cerebral (AVC), e pelo número de ocorrências no *corpus*, percebe-se que era um mal bastante comum nos séculos XVII e XVIII.

- (11) Faleceu em 19 do propio mez nesta Cidade, de huma **apoplexia**, depois de 21 dias de doença, em idade de 75 annos, o Dezembargador Eleuteiro Collares de Carvalho [...]. (GDL, 3 de março de 1740)
- (12) A Rainha D. Marianna de Neuburgo, viuva delRey Carlos II se acha já livre do accidente de **apoplexia**, que padeceo, e dava cuidado. (GDL, 4 de janeiro de 1725)
- (13) [...] noticia de haver alli falecido a 2 mons. Pociey, gram general de Lithuania de hum accidente de **apoplexia** [...]. (GDL, 9 de março de 1730)
- (14) [...] o Cardeal Joam Bautista Altieri havendo tido segundo accidente **apopletico** falecera no mesmo dia 12 [...]. (GDL, 12 de mayo de 1740)
- (15) [...] est mort hier en cette Ville, des suites de l'**apoplexie** dont il avoit été atteint il y a quelque temps. (GDJ, 25 de janeiro de 1768)

---

5 Informação obtida no site *Arthrolink*, na seção *Osteoarthritis Encyclopaedia*. Disponível em: <<http://www.arthrolink.com/eng/actualites/dossier-la-une/tous-les-dossiers/tous-les-dossiers-la-une/arthrose-dans-l-histoire-parti-4>>.

- (16) Não deixa de ser huma cousa curiosa o saber as diversas especies de enfermidades, que os causarão: 914 pessoas morrerão de bexigas, 462 de dentação, 612 de convulsões, 122 de catarraes, 42 mulheres sobre parto, 142 de febre quente, 119 de febres podres, 28 de hemorragias, 492 de molestias de peito, 1054 de tísica, 53 de hydropisia, 144 de asfyxia, 412 d'apoplexia, 25 de cancos, 15 de melancolia, 242 de velhice, &c. (GDL, 19 de fevereiro de 1790)

Pelas definições encontradas nos dicionários também é possível perceber se tratar de uma doença que era frequente e conhecida: “é um mal que, como um raio, fere e derruba subitamente. é uma obstrução dos ventriculos do cerebro, que tapando as artérias do rete mirable, impede as vias dos espiritos que sobe ao coração e tira de repente o movimento” (BLUTEAU, 1712-1728); “ataque do cerebro que se priva logo da sensibilidade e movimento com ronquido e dificuldade de respirar” (MORAES SILVA, 1813); “Maladie qui attaque le cerveau & oste tout à coup le mouvement & le sentiment.” (DAF, 1ª ed.); “Maladie qui attaque le cerveau, et qui ôte subitement la faculté des mouvemens volontaires” (DAF, 5ª ed.).

## Gonorreia

Durante muitos anos, a **gonorreia** foi confundida com a sífilis, portanto não se pode precisar se as ocorrências encontradas no *corpus* se referem a uma ou a outra doença. Entretanto, é curioso o fato de as duas doenças receberem diferentes nomes, alguns deles variando de acordo com o país, sobretudo no caso da sífilis, que também era conhecida naquela época como **vérole** ou **grosse vérole**.

Além destes nomes, popularmente, ela era conhecida pelos franceses como mal vénitien, de Naples ou napolitan, pelos italianos, como mal francês, e pelos portugueses, como mal espanhol. Segundo os dicionários portugueses da época, a **gonorreia** também era conhecida como gálico, mal francês ou mal venéreo e, ainda, popularmente, como chaude-pisse e esquentamento, estes últimos obviamente por causa de seus sintomas. Vê-se a seguinte ocorrência no *corpus*:

- (17) [...] que servem para curar zunidos dos ouvidos, dor de dentes, **gonorrhéas**, morreimas, fluxos de sangue, preservar de aborto, vertigens, dores de cabeça, confortar os nervos, desfazer tumores, confortar, e fazer alvos os dentes, rebater os vapores do utero, provocar os mezes, matar lambrigas, etc. (GDL, 5 de julho de 1725)
- (18) Ne voit-on pas , dit-il , tous les jours des hommes qui vous consultent pour une maladie vénérienne et qui cependant n'ont eu qu'une seule **gonorrhé**, qu'ils font remonter à 20 ou 30 ans.<sup>6</sup>

Nos dicionários, encontram-se as seguintes definições: “é fluxo de materia e emanação continua pelo cano da ourina de humores, acres, mordazes e corruptos do contagio Gallico, impresso nos vasos seminarios e parastratas ou glândulas” (BLUTEAU, 1712-1728); “esquentamento em que ardor na urina e purgação pela uretra” (MORAES SILVA, 1813); “Flux involontaire de semence” (DAF, 1762); “Écoulement par le canal de l'urètre, qui est dû le plus souvent à une affection vénérienne” (DAF, 1798).

---

6 Ocorrência extraída da *Gazette Médicale de Paris*, de 21 de maio de 1831, Tomo 2, n. 21. Disponível em : <<http://goo.gl/Epzw9>>. Utiliza-se essa ocorrência apenas como ilustração, pois, apesar de os dicionários franceses consultados apresentarem a unidade *gonorrhé* em sua nomenclatura, não foram encontradas ocorrências no jornal *Gazette de France*, muito provavelmente por conta do tabu cultural.

## Bexigas

O **mal de bexigas** aparece com bastante frequência nos textos da *Gazeta de Lisboa*, como se pode ver nos contextos seguintes:

- (19) [...] havendo falecido em Luneville de huma violenta doença de **bexigas**, que lhe acabou a vida dentro de 4 dias em 29 do passado. (GDL, 14 de setembro de 1715)
- (20) Faleceo em Namur de **bexigas** a 23 de Outubro, a Condessa Isabel Guilhelmina de Nassau d'Idyck, mulher do Conde Mauricio Luiz de Nassau Beverwert [...]. (GDL, 4 de janeiro de 1725)
- (21) O principe Manil de Saboya, que havia adoecido de **bexigas**, como se avizou a semana passada, faleceu na manhã de 28 de dezembro em idade de 42 annos. (GDL, 9 de fevereiro de 1730)
- (22) O mal de **bexigas** reina fortemente nesta Cidade, e tem levado grande numero de pessoas. (GDL, 18 de fevereiro de 1740)
- (23) Recebeu a Corte hum Expresso de Roma a 25 do passado com a noticia, de ser falecido de **bexigas** naquella Corte o Conde de Harrach [...].(GDL, 18 de fevereiro de 1740)
- (24) [...] deram noticia da morte do Conde de Lavague, filho do Principe de Masserano, morto em Novi de **bexigas** no dia antecedente. (GDL, 5 de outubro de 1745)
- (25) O principe de Condé vay bem na sua doença de **bexigas**. (GDL, 5 de fevereiro de 1750).
- (26) De Turin se escreve ser ali tam violento o mal de **bexigas**, que desde 15 de outubro passado tem perecido mais de 4U pessoas de toda a idade em ambos os sexos. (GDL, 19 de fevereiro de 1750).

O dicionário de Bluteau traz muitas informações sobre este mal:

[...] doença conhecida que cobre o couro de bostilas. Procede de um sangue viciado que causa esta efervescencia na massa sanguinaria e do sangue reconcentrado nas bostilas se gerãem uns pequenos abcessos com impressões corrosivas na pele, que nema deixam cicatrizes [...], é mal contagioso [...], ha bexigas negras, bexigas de pelo de lixa e bexigas doudas. (BLUTEAU, 1712-1728)

Pelas definições, percebe-se que se utilizava esse termo para designar diferentes tipos de doenças. Em Moraes Silva (1813), encontra-se: “espécie de empola que se ergue sobre a cutis, cheia de um humor acre e corrosivo”.

Pela designação usada em português, não foi possível encontrar a correspondência em lingua francesa. Desta forma, usou-se como base para a pesquisa o termo atual, pois o que era antigamente conhecido pelo nome de **bexigas** corresponde hoje a duas doenças: a **varíola** e a **varicela** (ou catapora). A partir desses dados e da verificação no *corpus*, encontram-se os correspondentes em francês: **petite verole** (mais frequente) ou **pourpre**.

- (27) Le fils unique du Comte de Guiscard, Ambassadeur de France en Suède, mourut, le 22 en cette Ville, de la **petite verole**. (GDF, 16 de janeiro de 1700)
- (28) Le même jour, l'Archiduc Charles se trouva indisposé, avec une grosse fièvre. On a reconnu depuis, que c'étoit la **petite verole**, la fièvre est diminuée considerablement, & on le croit hors de danger. (GDF, 16 de janeiro de 1700)
- (29) Le Duc de Glocester fut attaqué à Windsor le 5 de ce mois, d'une fièvre très violente mefiée de **pourpre**, dont il mourut le 10. (GDF, 21 de agosto de 1700)

Pelas acepções listadas nos dicionários franceses, percebe-se a correspondência: “Sorte de maladie maligne qui paroist au dehors par de petites taches rouges qui viennent sur la peau” (DAF, 1694) ; e no TLFi (1606): “Toute maladie caractérisée par une éruption de boutons, de plaques rouges sur la peau (rubéole, rougeole, scarlatine, variole, etc.)”.

## Tísica

Conhecida hoje como tuberculose, a **tísica** era bastante frequente nos séculos XVII e XVIII. Como a causa real dessa doença só foi conhecida no final do século XIX, outros males que provocavam os mesmos sintomas eram chamados de **tísica** (**phtisie** em francês). No *corpus*, encontra-se a seguinte ocorrência:

- (30) [...] 914 pessoas morrerão de bexigas, 462 de dentação, 612 de convulsões, 122 de catarraes, 42 mulheres sobre parto, 142 de febre quente, 119 de febres podres, 28 de hemorragias, 492 de molestias de peito, 1054 de **tísica**, 53 de hydropisia, 144 de asfyxia, 412 d’apoplexia, 25 de caneros, 15 de melancolia, 242 de velhice, &c. (GDL, 19 de fevereiro de 1790)

Em seu dicionário, Bluteau (1712-1728) diz que esta doença se caracteriza por “febre lenta, tosse, escarros de sangue e materia saniosa ou purulenta, que pouco a pouco atenuam o corpo”, e explica se tratar de:

[...] qualquer atenuação ou emaciação do corpo, mais particularmente aquela que procede de chagas no bofe [...], é uma exulceração do bofe, que insensivelmente derrete todo o corpo [...], causada de um humor acre e corrosivo.

Na primeira edição do Dicionário da Academia Francesa (1694), encontra-se : “Sorte de maladie qui desseche tout le corps, & qui est accompagnée d’une grande langueur”.

## Sezonismo ou febre terçã

Esta doença é conhecida hoje como paludismo ou malária, mas na época em estudo era designada por **febre terçã** (**fièvre tierce**). Era uma doença frequente na Europa e muitos estudiosos afirmam que foi levada às Américas pelos colonizadores. Sabe-se que, por exemplo, Cristovão Colombo era portador dessa doença, e que os jesuítas espanhóis encontravam um certo conforto para esse mal em técnicas que eram usadas pelos Incas para tratar de febres<sup>7</sup> mais brandas.

Em 1717, o médico italiano Giovanni Maria Lancisi, na publicação *De noxiis paludum effluviis eorumque remediis*, provou que a doença era transmitida por uma mosca, que era mais frequente em regiões de manguezais, e lhe deu o nome de malária (maus ares). As ocorrências são da *Gazette de France*:

- (31) Le Prince de Monaco Ambassadeur de France est malade depuis quelques jours, de la **fièvre tierce**, pour laqu’elle il a esté saigné deux fois [...]. (GDF, 02 de janeiro de 1700)
- (32) Le Cardinal Nerli est depuis quatre jours, incommodé d’unes **fièvre double tierce**: & le Cardinal Durazzo est aussi malade à Florence d’une **fièvre tierce**. (GDF 25 de setembro de 1700)

---

<sup>7</sup> Informação obtida em Kaufman e Ruveda (2005, p. 854-85).

Embora os dicionários de língua portuguesa cite a **febre terçãa**, não foram encontradas ocorrências na *Gazeta de Lisboa*, que traz apenas febres malignas, contagiosas, quentes, podres e febre amarela. Segundo esses dicionários, a febre terçãa é a “que dura dois dias consecutivos, procede de colera, que apodrece em dous lugares fora dos vasos maiores” (BLUTEAU, 1712-1728) e é um tipo de febre periódica ou errática, com repetição, “que torna a acometer dentro de dias certos” (MORAES SILVA, 1813).

Nos dicionários de francês, encontra-se **tierce** entre os tipos de febres descritos no verbete **fièvre**. No TLFi (1606) encontra-se na entrada **tièrce**: “Fièvre tierce - Variété de fièvre intermittente dont les crises reviennent le troisième jour, laissant entre elles un jour d’intervalle”, “Fièvre tierce doublée - Fièvre intermittente où il y a deux accès tous les deux jours avec un jour d’apyrexie”; “Fièvre double tierce - Fièvre où le malade a tous les jours des accès alternativement semblables, de sorte que le premier répond au troisième, le deuxième au quatrième, ainsi de suite”.

## Pedra nos rins

O problema de cálculos renais já era conhecido no século XVIII, como pode ser visto na seguinte ocorrência da *Gazeta de Lisboa*:

- (33) ElRey, que esteve alguns dias muy doente do seu **achaque de pedra**, se achava já convalecido [...]. (GDL, 6 de julho de 1745)

Nos dicionários, encontramos as seguintes definições: “doença, calculo que dos humores crassos e viscosos e endurecidos com o calor natural se forma nos rins ou na bexiga e impede as vias de ourina” (BLUTEAU, 1712-1728). “Une maniere de maladie appelée La pierre ou Gravelle, Calculus, Lithiasis ; La pierre engendrée au corps de l’homme, Humanus calculus” (TLF, 1606); “Pierre, se dit encore De l’amas de sable et de gravier qui se forme en pierre dans les reins, dans la vessie ou dans quelque autre partie du corps” (DAF, 1762).

Fato interessante é que os mesmos dicionários, na entrada cálculo, dizem se tratar de um uso antiquado para tratar das pedras nos rins: “Calcul, signifie aussi, la maladie de la pierre, soit dans les reins, soit dans la vessie. Il a le calcul. Il est vieux & hors d’usage”.

## Aborto (avortement / fausse couche / couche avant terme)

Nas gazetas também foram encontrados termos que tratam da finalização involuntária da gestação, como pode ser visto nos exemplos:

- (34) [...] servem para curar zunidos dos ouvidos, dor de dentes, gonorrheas, morreimas, fluxos de sangue, preservar de **aborto**, vertigens, dores de cabeça, confortar os nervos, desfazer tumores, confortar, e fazer alvos os dentes, rebater os vapores do utero, provocar os mezes, matar lambrigas, etc. (GDL, 5 de julho de 1725)
- (35) L’Amnassadrice de Venise est dangereusement malade dune **fausse couche** qu’elle a faite. (GDF, 25 de setembro de 1700)
- (36) L’Ambassadrice de Venise, qui avoit esté malade à l’extrémité, pour estre **accouchée avant terme**, est entièrement guérie. (GDF, 02 de outubro de 1700)

Na 1ª edição do *Dicionário da Academia Francesa* (1694), na entrada **avortement** é dada como definição “accouchement avant terme”, mas os exemplos tratam de um aborto provocado. No *Dictionnaire Critique de la Langue Française* (FERAUD, 1787-1788), encontra-se a diferenciação de uso:

L'avortement est l'accouchement avant terme. Avorter, c'est donc accoucher avant le terme ordinaire; mais en parlant des femmes, il ne se dit que d'un avortement volontaire et criminel, ou causé par un accident [...] ; Avorter se dit ordinairement des femelles des animaux.

O mesmo é explicado em outra edição do dicionário da Academia Francesa (1798):

En parlant Des femmes, on ne le dit guère que d'un accouchement avant terme provoqué par des moyens criminels. Procurer un avortement à une femme, au moyen de quelque breuvage. Lorsque l'accouchement avant terme a lieu par quelque accident, ou par l'effet d'une mauvaise constitution, on l'appelle fausse couche.

No francês essa diferenciação permanece até hoje, enquanto no português ela não existe. Em português, fala-se de aborto espontâneo ou natural e de aborto provocado. No dicionário de Bluteau (1712-1728), aborto é definido como “impre feita emissão do feto ou ação de parir ante tempo. [...] também se diz da criatura que morre no ventre da mãe, ainda que não saísse à luz do mundo [...]”.

## Considerações finais

O estudo dos termos da medicina usados nas gazetas da França e de Portugal nos séculos XVII e XVIII proporciona a ampliação do conhecimento não apenas dessa ciência, que estava se consolidando naquela época, mas também da maneira de pensar daquelas sociedades. Ele também confirma uma correspondência entre a maioria dos termos, o que reafirma a ideia de que as informações publicadas nestes jornais circulavam em diversos países.

A principal característica da terminologia da medicina daquela época, pelo que se pôde verificar nesta sucinta pesquisa, é a sua relação com a teoria hipocrática dos humores, ou teoria dos quatro humores, que foi a principal explicação racional sobre a saúde e a doença entre o século 4 a.C. e o século XVII. Segundo essa teoria, a vida seria mantida pelo equilíbrio entre quatro humores: sangue, fleuma, bÍlis amarela e bÍlis negra, relacionados, respectivamente, ao coração, sistema respiratório, fÍgado e baço. As doenças se deveriam a um desequilÍbrio entre os humores.

Percebe-se que as unidades encontradas no *corpus* eram de uso corrente e que os leitores dos dois jornais tinham algum conhecimento sobre elas. Pelas definições encontradas nos dicionários, confirma-se a filiação do conhecimento médico da época à teoria de Hipócrates, pois nas definições das doenças sempre há menção ao desequilÍbrio dos humores como causa.

## REFERÊNCIAS

- ARTHROLINK. *Osteoarthritis Encyclopaedia*. Disponível em: <<http://www.arthrolink.com/eng/actualites/dossier-la-une/tous-les-dossiers/tous-les-dossiers-la-une/arthrose-dans-l-histoire-parti-4>>. Acesso em: 03 jan. 2011.
- BAUDET, J. C. Histoire du vocabulaire de spécialité, outil de travail pour l'historien des sciences et des techniques. In: VV. AA. *Terminologie diachronique*. Actes du Colloque, Bruxelles, CILF, 1988. p. 56-67.
- BLUTEAU, R. *Vocabulario portuguez & latino: aulico, anatomico, architectonico...* Coimbra: Collegio das Artes da Companhia de Jesus, 1712-1728. 8 v.
- BOURET, B.; TREVES, R.; MARC, J-F. Affection rhumatismales dans la peinture artistique. In: CNIT – *Congrès français de rhumatologie*. Paris, 2009. Disponível em: <[http://www.publication3d.com/kiosque/cimpppa/\\_files/PDF-V1071-Lite\\_DVL\\_00/index\\_voix.html](http://www.publication3d.com/kiosque/cimpppa/_files/PDF-V1071-Lite_DVL_00/index_voix.html)>. Acesso em: 05 jan. 2011.
- CABRÉ, M. T. *La terminologia: representación y comunicacion*. Barcelona: Institut Universitari de Linguística Aplicada, 1999.
- DICTIONNAIRE de *l'Académie française*. 1. ed. Paris: Jean Baptiste Coignard, 1694. Disponível em: <<http://artfl.atilf.fr/dictionnaires/ACADEMIE/PREMIERE/premiere.fr.html>>. Acesso em: 04 jan. 2011.
- DICTIONNAIRE de *l'Académie française*. 4. ed. Paris: Vve B. Brunet, 1762. Disponível em: <<http://artfl.atilf.fr/dictionnaires/ACADEMIE/QUATRIEME/quatrieme.fr.html>>. Acesso em: 04 jan. 2011.
- DICTIONNAIRE de *l'Académie française*. 5. ed. Paris: J. J. Smits et Ce., Imp.-Lib, 1798. Disponível em: <<http://artfl.atilf.fr/dictionnaires/ACADEMIE/CINQUIEME/cinquieme.fr.html>>. Acesso em: 04 jan. 2011.
- FÉRAUD, J-F. *Dictionnaire critique de la langue française*. Marseille: Mossy, 1787-1788. Disponível em: <<http://dictionnaires.atilf.fr/dictionnaires/FERAUD/index.html>>. Acesso em: 04 jan. 2011.
- GAUDIN, F. *Socioterminologie: une approche sociolinguistique de la terminologie*. Bruxelles: De Boeck & Larcier, 2003.
- GAZETTE *médicale de Paris*: Journal de médecine et des sciences accessoires. Paris, 1831. Disponível em: <<http://books.google.ca/books?id=PHNEAAAACAAJ&hl=pt-BR&pg=PP1#v=onepage&q&f=false>>. Acesso em: 16 out. 2012.
- KAUFMAN, T.; RUVEDA, E. The quest for quinine: those who won the battles and those who won the war. *Angewandte Chemie International Edition*, v. 44, p. 854–885, jan. 2005.
- LITTRE, E. *Dictionnaire de la langue française*. Paris, 1863. Disponível em: <<http://francois.gannaz.free.fr/Littre/accueil.php>>. Acesso em: 07 jan. 2011.

MORAES SILVA, A. *Diccionario da lingua portugueza* - recopilado dos vocabularios impressos ate agora, e nesta segunda edição novamente emendado e muito acrescentado. Lisboa: Typographia Lacerdina, 1813. Disponível em: <<http://www.brasiliana.usp.br/diccionario/edicao/2>>. Acesso em: 09 jan. 2011.

NICOT, J. *Le Thresor de la langue françoise, tant ancienne que moderne*. Paris: David Douceur, 1606. Disponível em : <<http://artfl.atilf.fr/dictionnaires/TLF-NICOT/index.htm>>. Acesso em: 07 jan. 2011.

# Sobre usos do comitativo em línguas africanas e a análise do ponto de referência

(On Comitative uses in African languages and the reference point analysis)

Paulo Jeferson Pilar Araújo<sup>1</sup>, Dayane Cristina Pal<sup>2</sup>

<sup>1,2</sup>Departamento de Linguística – Universidade de São Paulo (USP)

<sup>1</sup>Bayreuth International Graduate School of African Studies – Universidade de Bayreuth (BIGSAS)

jefersonpilar@gmail.com, dayabrik@uol.com.br

**Abstract:** This article focuses on the syncretic nature of the comitative category and on the search for a conceptualist characterization of some constructions related to it. Some grammatical constructions are presented in African languages by observing the particularity of those languages in order to propose that the Reference Point Analysis, in the Cognitive Grammar framework, would be the common conceptualist base of all constructions related to the conceptual domain of comitative.

**Keywords:** Comitativo; african languages; reference points; cognitive grammar; conceptual domains.

**Resumo:** Este artigo focaliza a natureza sincrética da categoria Comitativo e busca uma caracterização conceptualista para as diversas construções relacionadas a essa categoria. Faz-se uma apresentação de construções gramaticais em línguas africanas observando-se a particularidade dessas línguas para então se propor que a Análise do Ponto de Referência, nos termos da Gramática Cognitiva, seja a base conceptualista comum a todas as construções relacionadas ao domínio conceitual do Comitativo.

**Palavras-chave:** Comitativo; línguas africanas; pontos de referência; gramática cognitiva; domínios conceituais.

## Introdução<sup>1</sup>

O domínio conceitual do Comitativo tem recebido uma maior atenção nos últimos anos, contando com um trabalho tipológico de fôlego (STOLZ; STROH; URDEZ, 2006), e aparecendo como uma categoria bem produtiva em diversas construções como as de posse predicativa (STASSEN, 2009; HEINE, 1997), coordenação (HASPELMATH, 2004; STASSEN, 2000), construções seriais (LORD, 1993), existenciais (ABDOULAYE, 2006), dentre outras. O Comitativo também já mereceu uma tentativa de delimitação com Arkhipov (2009), no entanto, o estatuto dessa categoria continua bastante controverso

---

<sup>1</sup> Agradecemos o apoio institucional do Fundo de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão-Fapema pela bolsa de estudos concedida entre 2009 e 2010 e 2013 e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Capes, pela bolsa sanduíche concedida durante o ano de 2011. Parte deste trabalho foi realizado enquanto o primeiro autor esteve na *Bayreuth International Graduate School of African Studies-BIGSAS*, Alemanha, na condição de *Associated Junior Fellow*. Agradecemos também os comentários da professora Margarida Petter, coordenadora do Grupo de Estudos de Línguas Africanas-Gela, vinculado ao Departamento de Linguística da USP, e à colaboração de colegas falantes de baulê e quizombo que contribuíram com a solução de dúvidas no uso dos exemplos de suas línguas maternas. Os erros e problemas remanescentes são inteiramente nossos.

em sua relação com outras categorias com as quais ela sincretiza.<sup>2</sup> Em nossos trabalhos com línguas africanas temos verificado o papel significativo do Comitativo em diversas construções não tão comuns em línguas indo-europeias, razão de acharmos oportuno dedicar uma maior atenção a essa categoria, sem deixar de atentar para as particularidades tipológicas das línguas africanas.

Neste trabalho nos ocupamos de construções em línguas africanas relacionadas ao Comitativo, para então verificar a possibilidade de analisar essas construções a partir de uma base conceitual comum, os pontos de referência cognitivos ou nos termos da Gramática Cognitiva (LANGAKER, 2009, 2000), a “Análise do Ponto de Referência”.<sup>3</sup>

Este artigo está organizado da seguinte forma: iniciamos com a relação do domínio conceitual do Comitativo com outros domínios e uma proposta de delimitação em termos semânticos. Em seguida, apresentamos as construções que devem ser consideradas para uma tipologia do Comitativo em línguas africanas. O passo seguinte é o de verificar de que forma essas construções podem ser consideradas como Construções de Pontos de Referência, nos termos de Langacker (2000). Por fim, traçamos as considerações finais.

### **A relação do Comitativo com outros domínios conceituais**

Heine e Kuteva (2002, p. 79-90) enumeram pelo menos dez possíveis processos de gramaticalização tendo as marcas de comitativo como origem, dentre eles o uso de partículas de valor comitativo para introduzir agentes de construções passivas; coordenação entre SNs ou entre sentenças; comitativos como marcas de progressivo; em construções existenciais; expressão do papel semântico de instrumento; expressão de posse predicativa, etc.<sup>4</sup>

O trabalho tipológico de Stolz, Stroh e Urdez (2006) tenta elaborar um mapa semântico dos sincretismos existentes entre o Comitativo e as funções gramaticais comumente associadas a ele, principalmente para as construções de agentividade, de coordenação, relações espaciais, possessivos e instrumento. Talvez o sincretismo mais conhecido seja o do Comitativo com Instrumento, conforme atestado pela Metáfora da Companhia de Lakoff e Johnson (1980), segundo a qual o uso do comitativo para a expressão de instrumento seria universal nas línguas do mundo.

Enquanto os trabalhos supracitados enfatizam a natureza polissêmica das partículas de comitativo nas línguas, Arkhipov (2009) tenta distinguir o Comitativo de outras categorias, algumas extremamente difíceis para uma distinção, como as de Acompanhamento e Associativo. Vejamos mais detalhadamente a proposta de Arkhipov (2009) para delimitação do domínio conceitual do Comitativo.

---

2 Sincretismo entre funções gramaticais é aqui entendido conforme Stolz (2001) para o caso em que uma mesma partícula tem usos para mais de uma função, a exemplo de comitativo e instrumento em português: “estou **com** uma faca na mão” e “eu corto o pão **com** uma faca”.

3 A terminologia é bastante variada para os processos cognitivos de relações de ponto de referência. Decidimos utilizar Análise de Pontos de Referência como um termo metateórico, enquadrando o que Langacker considera construções de pontos de referência.

4 Conferir o trabalho tipológico de Stolz, Stroh e Urdez (2006) para os possíveis sincretismos de funções gramaticais em um conjunto de mais de 400 línguas.

## Delimitando o domínio conceitual do Comitativo

Arkhipov (2009) apresenta a seguinte definição para o Comitativo:

In a nutshell, comitative is defined as a particular construction type used to ‘pluralize’ a participant – that is, to predicate the same state of affairs of two individual participants, such that the main predicate itself is not repeated and the two participants are not equal in their syntactic status. (ARKHIPOV, 2009, p. 223)

Um exemplo em português seria:

- (1) João saiu **com** a Maria (Eles (João e Maria) saíram = João e Maria saíram).

Seguindo a definição acima, Arkhipov acredita ter encontrado uma forma de delimitar uma Construção Comitativa (genuína) com o objetivo de distinguir a categoria Comitativo de outras categorias gramaticais, de forma que um trabalho tipológico, a exemplo dos de Stassen (2000 e 2009), possa ser realizado. As partículas de comitativo são altamente polissêmicas em praticamente todas as línguas do mundo e, conforme notado por Haspelmath (2010), decidir quais os melhores “conceitos comparativos” (usados em trabalhos comparativistas e tipológicos) para tratar de “categorias descritivas” (particulares de língua para língua) é uma tarefa extremamente difícil devido ao que Haspelmath chama de PARTICULARISMO CATEGORIAL (CATEGORIAL PARTICULARISM). Segundo esse conceito, cada língua tem uma forma diferente de atribuição de traços para a delimitação de uma categoria gramatical. Não adentraremos de forma aprofundada no cerne desta questão levantada por Haspelmath, mas fazemos menção à proposta do autor por acreditarmos que a forma de definir o Comitativo em Arkhipov (2009) é um bom exemplo de como conceitos usados para fins comparativos em linguística desconsideram a forma particular de conceptualização em cada língua. Na proposta de Arkhipov, haveria uma “Construção Comitativa” (genuína) (=ComC), enquanto todas as outras categorias como a de Instrumento, Companhia, Associativo, etc. seriam “construções quase-comitativas”. O desafio em definir categorias gramaticais de uma forma que possam ser caracterizadas em termos semânticos e conceituais ao mesmo tempo em que as particularidades de conceptualização das línguas sejam respeitadas é um problema de solução controversa.

Acreditamos, no entanto, que a proposta conceptualista da Gramática Cognitiva de Langacker (2000, 2009), que advoga a natureza simbólica da gramática, categorias como o Comitativo podem ser caracterizadas a partir de processos cognitivos básicos, respeitando a particularidade de cada língua por ser uma teoria baseada no uso e pautar-se em esquemas construcionais ao invés de regras e restrições. Seguindo o espírito Boasiano de Haspelmath, nas próximas sessões apresentamos construções em línguas africanas nas quais o comitativo exerce papel fundamental, e sob a lente de Arkhipov todas deveriam ser consideradas “quase-comitativas”. Buscaremos mostrar a particularidade de cada construção para em seguida propormos que a definição de Comitativo de Arkhipov pode ser caracterizada melhor se atentarmos para uma possível base conceptual comum às construções chamadas pelo autor de genuínas e quase-comitativas.

## Domínios conceituais do Comitativo em línguas africanas

Para este trabalho buscamos fazer, na medida do possível, uso de dados linguísticos acessíveis na literatura africanista ou específica para cada tópico discutido devido ao fato de não termos como preocupação fundamental a descrição detalhada de cada construção apresentada, mas tão somente exemplificar os principais traços tipológicos em que o Comitativo desempenha um papel importante nas línguas africanas e que de certa forma se diferenciam dos usos mais conhecidos nas línguas indo-europeias. As línguas selecionadas são as comumente citadas nos trabalhos voltados para os temas de cada subseção.

### Comitativo e coordenação entre SNs

Sem dúvida o protótipo de um comitativo é a expressão de companhia juntamente com o associativo (STOLZ; STROH; URDEZ, 2006). Um conjunto de construções que abarcam esses dois principais sentidos e usos do comitativo é o da coordenação entre SNs ou “NP-conjunction” na terminologia tipológica de Stassen (2000). Na sua tipologia, Stassen sugere que as línguas do mundo fazem uso principalmente de duas alternativas para marcar a coordenação entre dois participantes, da seguinte forma: línguas que fazem uso de duas marcas diferentes para construções como em “A **and** B” (A e B) e “A **with** B” (A **com** B), como é o caso de boa parte das línguas indo-europeias, são denominadas como *AND-languages*. Para as línguas que fazem uso da mesma marca para as construções coordenativas e comitativas, como em “A **with/and** B” (A e/**com** B), o autor denomina como *WITH-languages*, este é o caso para a maioria das línguas bantas.<sup>5</sup> Um dos achados tipológicos de Stassen (2000) é o que o autor chama de “AND-drift” (E-deriva), ou seja, a ideia de que as *WITH-languages* passariam por diferentes processos de gramaticalização para a formação de uma partícula coordenativa de NPs diferente da partícula comitativa. O continente africano é apontado pelo autor como a área linguística onde esses processos são bastante comuns (STASSEN, 2000, p. 27).

Um exemplo é a língua iraqw da Tanzânia. Nela, a adposição *nee* é usada tanto para a coordenação entre SNs (2a) como para o sentido de comitativo (2b):<sup>6</sup>

---

5 Seguimos a proposta de Fiorin e Petter (2008) e damos preferência às formas aportuguesadas das línguas africanas. Para não entrarmos no mérito da questão sobre a adequação ou não do aportuguesamento dos nomes das línguas, conforme comentários de um dos nossos pareceristas anônimos, remetemos o leitor à proposta supracitada e observamos que para o caso das línguas africanas, diferentemente das línguas indígenas brasileiras, muitas delas já contam com formas aportuguesadas e dicionarizadas, como é o caso de hauçá, quimbundo, suaíli, etc, inclusive a forma para plural e feminino “bantas” (HOUAISS, 2009).

6 As abreviaturas utilizadas são: ABL = ablativo; APL = aplicativo; COM = comitativo; CONJ = partícula com valor conjuncional; COP = cópula; DET = determinante; EXIST = existência; IMP = Imperativo; IMPF = imperfectivo; IMPS – impessoal; INST = instrumento; LOC = locativo; NAR = narrativo; NEG = negação; PAS = passado; PRS = presente; PRM = passado remoto; PSR = passado recente; PERF = perfeito; POSS = possessivo; VF = vogal final; 1-3SG = 1<sup>a</sup>-3<sup>a</sup> do singular; 1-3PL = 1<sup>a</sup>-3<sup>a</sup> do plural. Para indicar o sincretismo entre funções utilizamos a notação X/Y, como exemplo para o comitativo e instrumento: COM/INST. Para fins de padronização das glosas, os exemplos foram simplificados, de forma que foram omitidos os números indicadores das classes nominiais como também a indicação das zonas das línguas bantas, como é feito na tradição bantuísta. Sugerimos que as fontes originais sejam consultadas, seja para contrastar com a simplificação feita neste artigo como também para possíveis citações.

- (2) Iraqw (Afro-asiática, Cuchítica; MOUS, 2004, p. 112-113)
- a. *muu-dá nee dama-r-ín ta-ri waráahh*  
 Pessoas-DEM **COM/CONJ** panturrilha-3PL-POSS IMPS-NAR passar.PAS  
 ‘As pessoas e suas panturrilhas passaram’ (As pessoas com suas panturrilhas passaram)
- b. *ino’ín ta qaró waráahh nee hikwa-ín*  
 3PL IMP já passar.PAS **COM/CONJ** gado-POSS  
 ‘Os outros já passaram com o seu gado’

Em quizombo, dialeto do quicongo, há a possibilidade de algumas ambiguidades entre os sentidos de companhia e conjunção entre NPs, como atestado no exemplo (3a) abaixo:<sup>7</sup>

- (3) Quizombo (Nigero-congolês, Banto central; TECA; ARAÚJO, em preparação)
- a. *\*Yándì didì lòsò yé màdèsò*  
 3SG Comer.PSR arroz **COM/CONJ** feijão  
*yé mwànà*  
**COM/CONJ** criança  
 ‘Ele comeu arroz e feijão e o filho’ (Sentido de que ele comeu o filho juntamente com arroz e feijão)
- b. *yándì yé mwànà á-didì lòsò*  
 3SG **COM/CONJ** criança 3PL-comer.PSR arroz  
*yé màdèsò*  
**COM/CONJ** feijão  
 ‘Ele e o filho comeram arroz e feijão’

## Comitativo e possessivo

Nos dois principais trabalhos tipológicos sobre posse predicativa (STASSEN, 2009; HEINE, 1997), dentre as estratégias mais recorrentes para a expressão de posse nas línguas do mundo, encontramos o comitativo em construções que seguem uma fórmula “X is with Y” (X está com Y” no sentido de “X tem Y” (STOLZ, 2001).<sup>8</sup> Como grandes representantes dessa estratégia estão as línguas bantas. O uso de verbos comitativos para a expressão de posse é também encontrado em línguas dos continentes australiano e americano. Abaixo alguns exemplos de línguas bantas:

- (4)<sup>9</sup> Quimbundo (Nigero-congolês, Banto central; XAVIER, 2010, p. 131)
- ù-énd-à nì mù-zúmbù kì-ù-á-ǰímbà-él-él-è*  
 3SG-anda-IMPF **COM/POSS** LOC-lábios NEG-PRM-3SG-esquecer-PERF-APL-VF  
 ‘ele anda com beijos, ele não (se) perdeu’ (Ele tem boca, não se perde)
- (5) Tsuana (Nigero-congolesa, banto do Sudoeste; CREISSELS, [s.d.], p. 3)<sup>10</sup>

7 A ambiguidade em 3(a) pode ser desfeita com o acréscimo da palavra “kùmósi” antes de “yé” mwànà: (Ele comeu arroz e feijão junto com o filho).

8 As outras estratégias utilizadas pelas línguas para a expressão de posse são o uso de cópula ou verbos com sentido existenciais em construções locativas e existenciais ou o uso de um verbo equivalente a “to have” do inglês.

9 Provérbio equivalente a “Quem tem boca vai a Roma”.

10 Nestes dois exemplos, Creissels observa que em sentenças afirmativas o morfema “na” é usado junta-

- a. *ke na le madi.*  
 1SG COP COM/POSS dinheiro  
 ‘eu tenho dinheiro’ (Eu estou com dinheiro)
- b. *Ga ke na madi*  
 NEG 1SG COM/POSS dinheiro  
 ‘eu não tenho dinheiro’ (Eu não estou com dinheiro)

(6) Leembama (Nigero-congolesa, banto da floresta; OKOUDOWA, 2010, p. 182)

- wε na mw-ánáńkéri ka tuvr-áá ngúrúbalaxa ní*  
 2SG COM/POSS sobrinhoNEG insultar-HAB tio NEG  
 ‘Você com/tem sobrinho, não insulte seu tio’ (Se você tem sobrinho, não insulte seu tio)

Apesar do uso da estratégia comitativa ser um traço forte nas línguas bantas, Stassen (2009, p. 416, 418) apresenta exemplos de outros troncos linguísticos como o nilo-saariano, em (7) e o afro-asiático em (8):<sup>11</sup>

(7) Kukú (Nilo-saariano, nilótico oriental; COHEN, 2000, p. 133)

- ǰ kɔ píli*  
 1SG COM/POSS pilili  
 ‘eu tenho um pilili’

(8) Hauçá (Afro-asiático, chádico; NEWMAN, 2000, p. 161)

- Ta-na dà sabuwar munduwa.*  
 3SG-PROG COM/POSS novo bracelete  
 ‘Ela tem um novo bracelete’

## Comitativo e instrumento

Sabemos agora, com os estudos de Stolz, Stroh e Urdez (2006), que o sincretismo entre marcas de comitativo com instrumento não é tão universal como pretendiam Lakoff e Johnson (1980) com a Metáfora da Companhia. Boa parte das línguas do mundo usa uma estratégia em que o Comitativo e o Instrumento são geralmente distintos, como é o caso da língua iraqw (MOUS, 2004, p. 112), no entanto, em algumas línguas africanas, como o baulê, o eve, o fon, o quicongo, etc. essas funções são sincretizadas, traço tipológico bastante comum entre as línguas indo-europeias também. Vejamos o caso do baulê.

---

mente com “le”, carregando o sentido de “estar com”, sendo glosado como COP no primeiro exemplo, porém, em sentenças negativas, o “na” ocorre sozinho como um verbo comitativo, sendo assim glosado como COM no Segundo exemplo. Com isso, o autor faz algumas considerações sobre o estatuto do morfema “na”: “Historically, the variant *na-* (which has a very limited distribution and is never obligatory) is a reflex of the Proto-Bantu comitative preposition \**na*, whereas the productive variant *le* has obvious cognates in the other Southern Bantu languages only, and its etymology remains a mystery.” Conferir Nurse (2008) para uma discussão sobre a distribuição de *na* em línguas bantas.

<sup>11</sup> As línguas africanas são classificadas em quatro grandes troncos ou filios linguísticos: Nigero-congolês, Afro-asiático, Nilo-saariano e Coissan (Cf. BONVINI, 2008).

O baulê possui o morfema *nì* cuja função básica é a de exprimir relações de coordenação entre dois termos, pronomes e nomes. Esse morfema pode introduzir um nome com valor de acompanhamento. Em (9a), estabelece-se a coordenação de dois termos que preenchem a função sintática de argumento externo do verbo. A cena descrita nessa construção, porém, é a de alguém que trouxe consigo frutas. O mesmo morfema coordenativo *nì* é empregado para introduzir um instrumento (9b):

(9) Baulê (Nigero-congolês, cuá; CREISSELS; KOUADIO, 1977, p. 348, 344)

a. ɔ#      **nī**                      Wākāmă      bà-lī  
 3Ss      **CONJ/COM**      fruta              vir-PERF  
 ‘Ele com frutas veio’ (Ele trouxe frutas)

b. ɔ#      **kpé-li**                      kpa’ún              **Nī**              la’liē  
 3Ss      **cortar-PERF**              pão              **COM/INST**      faca  
 ‘Ele cortou o pão com uma faca’

### Comitativo e Existencial/Locativo

O comitativo também participa na expressão de existenciais, como nos exemplos do hauçá e quissuaili abaixo:

(10) Hauçá (Afro-asiático, chádico; ABDOULAYE, 2006, p. 1124)

*Dà                      mutàanee      màasu      shâ-n      wannàn      taabàa.*  
**COM/EXIST**      pessoas      donos.de      beber-de      este      tabaco  
 ‘Há pessoas que fumam esses cigarros.’

(11) Quissuaili (Nigero-congolês, Banto central; HEINE; KUTEVA, 2002, p. 84)

*Ku-      **na**                      asali      nyingi.*  
 LOC      **COM/EXIST**      mel      muito  
 ‘Há muito mel.’

Stolz, Stroh e Urdez (2006) atestam a pouca produtividade de sincretismo entre as marcas de comitativo com expressões de espaço nas línguas do mundo, porém, em quizombo encontramos o uso de *yé* (marca do comitativo) em construções de mudança de localização:

(12) Quizombo (Nigero-congolês, banto central; TECA; ARAÚJO, em preparação)

*túkà      kù-nzó                      yé                      kùnà      kókò 15              minùtì      mù-màlù*  
 ABL      LOC-casa                      **COM(até/para)**      LOC      rio      15              minuto      LOC-pé  
 ‘De casa para o rio são 15 minutos a pé’ (Da casa com o rio são 15 min. no pé.)

## Comitativo e construções seriais

Em seu estudo sobre gramaticalização em construções seriais, Lord (1993, p. 47-64) sugere que preposições de valor comitativo como *kpèlu* do iorubá, o *kplé* do eve e o *kpôdô* do fon surgiram de verbos em construções seriais e que essas preposições podem ainda assumir a mesma posição de verbo em configurações similares às de uma construção serial, indicando assim sua origem verbal (Comparar exemplos em (9a) e (13)).

Ainda em uma construção serial é possível depreender o sentido de comitativo mesmo sem um morfema específico para o valor de comitativo, como no exemplo abaixo do baulê:

(13) Baulê (Nigero-congolês, cuá; KOUADIO, 2000, p. 83)

ɔ            *fā-li*                      *dwō*                      *bà-lī*

3SG    pegar-PERF            inhame    vir-PERF

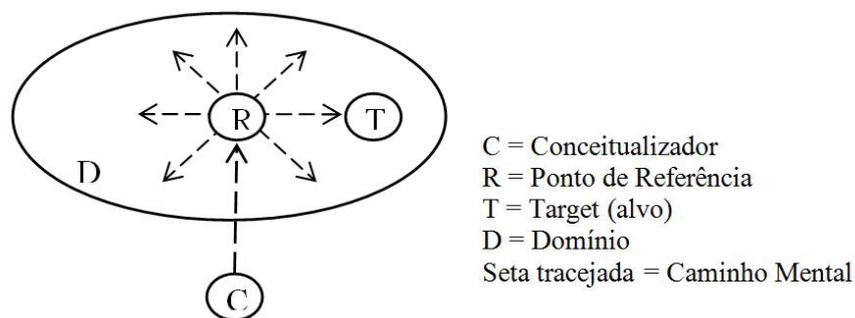
‘ele trouxe o inhame’ (Ele pegar inhame vir/Ele com inhame veio)

O verbo *fa* em seu sentido mais prototípico remete a traços de volição e posse e estabelece uma relação de posse/companhia entre o participante da cena descrita e o objeto nela relacionado. Por extensão metafórica, esses traços são também identificados em construções seriais. No caso do exemplo (13) acima, por meio desse verbo depreende-se que o participante da cena descrita carrega, traz consigo o inhame. A cena completa seria a de vir e trazer o inhame ou vir COM o inhame.

## A base conceptual para o Comitativo: a Análise do Ponto de Referência

Para todas as construções apresentadas nas subseções anteriores, buscamos atentar para a sugestão de Haspelmath (2010) quanto ao PARTICULARISMO CATEGORIAL das línguas africanas. Em cada construção apresentada, a definição de Construção Comitativa (genuína) de Arkhipov (2009) torna extremamente restritiva a relação do Comitativo com outras construções com as quais as fronteiras entre domínios conceituais são não discretas, desde construções mais prototípicas, como as de coordenação de NPs, para construções nem tão prototípicas como em construções seriais. Por outro lado, se tomarmos uma posição conceptualista de gramática como proposta por Langacker (2000) vemos que é possível identificar as operações cognitivas que apontem para uma base conceptual comum e esquemática suficiente para abarcar todas as construções em que o comitativo exerce algum papel.

Como apontado por Tribushinina (2008), a noção de pontos de referência cognitivos foi introduzido por Rosch (1975) em seus trabalhos sobre protótipos na categorização de cores. Os pontos de referência foram bastante estudados em Psicologia Cognitiva. Em Linguística temos com Langacker (2009; 2000) diversas aplicações da noção de Pontos de Referências Cognitivos para construções gramaticais de difícil caracterização, notadamente as construções possessivas (LANGACKER, 2003). A Análise do Ponto de Referência é também considerada como a base conceptual para diversas construções, além das construções possessivas já mencionadas, construções de tópico, dêiticos, metonímia (Langacker, 2000) e anáfora pronominal (VAN HOEK, 2007). A figura abaixo representa a operação de ponto de referência como utilizada em Gramática Cognitiva:



**Figura 1. Análise do Ponto de Referência**

Nas palavras do autor:

[...] the reference point ability is our capacity to invoke one conceived entity as a means of establishing mental contact with another, i.e. mentally accessing one entity **via** another. The conceptualizer (C) first directs attention to the entity serving as **reference point** (R). Attending to R evokes a set of associated entities, collectively called its **dominion** (D), one of which is the **target** (T). A reference point relationship is thus a matter of sequenced mental access, where directing attention to R makes it possible to then direct attention to T. (LANGACKER, 2009, p. 82)

Essa é a caracterização esquemática da Análise do Ponto de Referência. Para que se dê conta de todas as outras instanciações que tenham a habilidade do ponto de referência como base conceitual, Langacker demonstra que outros processos cognitivos como a relação assimétrica entre figura e fundo, nos termos técnicos de Langacker, o alinhamento entre Trajetor (TR) e Marco (MR) e o perfilamento decorrente do *construal*<sup>12</sup> de uma cena serão também determinantes. Na impossibilidade de apresentar todos os construtos usados em Gramática Cognitiva para representar em diagramas os padrões de uma construção, reservaremos algum espaço para exemplificar a forma como Langacker trabalha com sua notação, de forma que não fique pesadamente técnica a análise, mas que possa ser possível seguir intuitivamente o modelo de análise proposto em Gramática Cognitiva.

O alinhamento entre TR e MR especificam os elementos de maior e menor proeminência, respectivamente. Numa pergunta como “onde está o livro?” para uma resposta “o livro está em cima, na prateleira”, aponta para o “livro” como o elemento de maior proeminência, portanto, considerado o TR, enquanto que a “prateleira”, mesmo que esteja presente na cena, tem menor proeminência que o “livro”, sendo assim o MR. Nas figuras abaixo, representamos as duas situações em que o livro está localizado na parte de cima de uma prateleira e inserido nela. Os círculos e linhas espessos significam que a atenção do Conceitualizador (C), falante ou ouvinte, está tomando os elementos Ponto de Referência (R) e o alvo (T) como figura e o Domínio (D) como fundo na Figura 2(a). Na Figura 2(b) o Ponto de Referência se confunde com o Domínio, dando assim o sentido de inclusão da preposição *em* do português:

<sup>12</sup> Nem todos os termos da Gramática Cognitiva possuem uma tradução equivalente em português ou já aceita pelos pares brasileiros como o caso de perfilamento para *profiling* ou Trajetor e Marco para *Trajector* e *Landmark*. Decidimos manter o termo *construal*.

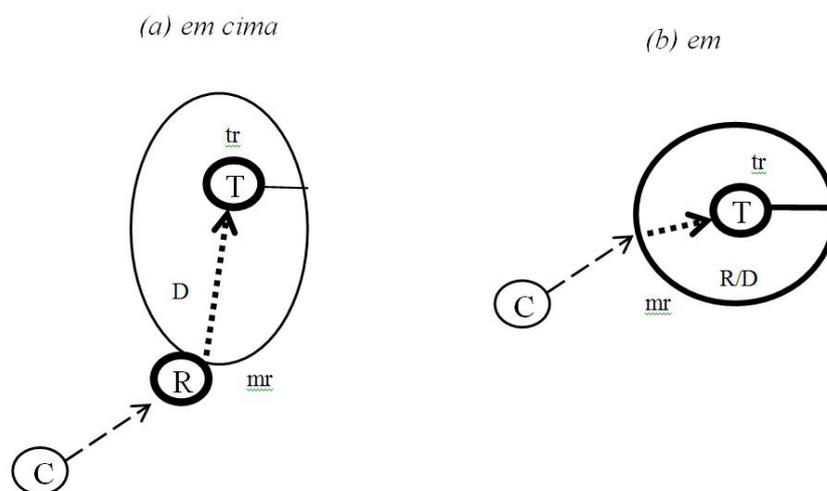
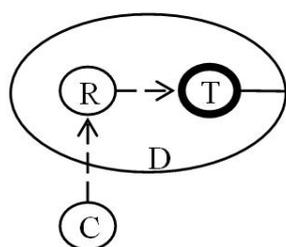
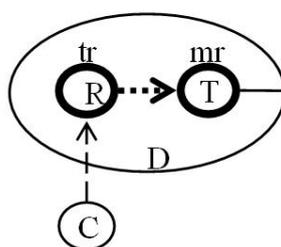
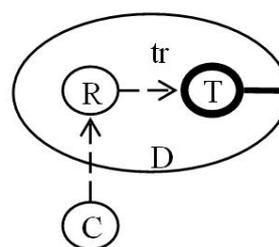


Figura 2. Orientação espacial na análise do ponto de referência

Como já mencionado, a Análise do Ponto de Referência tem sido usada principalmente para a descrição de construções de posse (LANGACKER, 2000, 2003), de forma que para adentrarmos na questão do Comitativo apresentamos primeiramente como Langacker trata das construções possessivas para então delineararmos nossa proposta de como a mesma análise pode ser usada em nosso estudo.

Langacker constata inicialmente que o domínio de posse abrange diversas situações relacionais entre duas entidades, um possuidor (PR) e um possuído (PD) e que tanto nas construções de posse adnominal como de posse predicativa uma entidade mais saliente é evocada com o propósito de estabelecer contato mental com uma entidade alvo menos saliente. Nesse caso, a habilidade do ponto de referência seria a operação cognitiva por trás das relações possessivas e que demonstraria de forma mais intuitiva o tipo de relação assimétrica que acontece entre as entidades PR e PD.

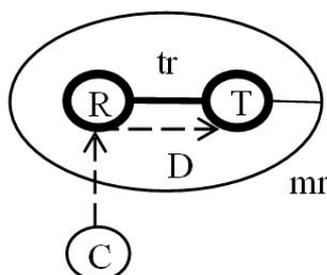
A Análise do Ponto de Referência explicaria melhor a natureza da relação entre PR e PD por indicar a assimetria da relação entre as duas entidades nas relações de posse, como exemplo: “o pescoço da menina” e a situação assimétrica: “a menina do pescoço”; “o assassinato do presidente” de “o presidente do assassinato”, etc. Para Langacker (2003, p. 21): as estratégias de expressão de posse predicativa são alternativas funcionais comparáveis. A equivalência entre uma língua que faz uso de verbos do tipo “to have” (*Have-possessive languages*) e línguas que fazem uso de cópula (*Be-possessive languages*) para expressar posse predicativa se reflete no nível estrutural composto (*composite structure level*), nível decorrente do emparelhamento de estruturas no nível de componentes (*component structures level*). As diferenças se dariam pelos processos de perspectivização, como já citados, de perfilamento, alinhamento entre TR e MR, etc. Nos gráficos abaixo, Langacker (2009, p. 89) demonstra que as operações de alinhamento entre TR e MR permitem visualizar a alternância entre um possessivo no nível nominal, na Figura 3(a). para o nível oracional, Langacker usa os termos *Be-possessivos* e *have-possessivos* para línguas que fazem uso de cópula ou verbos equivalentes a “ter”, respectivamente:

(a) *Possessivo Nominal*(b) *HAVE Possessivo*(c) *BE Possessivo*

**Figura 3. Distinção entre possessivos a partir da Análise do Ponto de Referência**

Na Figura 3(a), um possessivo nominal se caracteriza por perfilar uma *coisa*, enquanto um possessivo oracional perfila um *processo*, indicado pelos círculos e a seta entre R e T, mais espessa. A seta tracejada entre R e T na Figura 3(b) indica o controle de R sobre T, característica semântica de um possessivo. Uma entidade é possuída quando está sob controle, físico, experiencial ou social por outra entidade. Na Figura 2(c), as línguas que fazem uso de cópula ou verbos comitativos para a expressão de posse têm o seu ponto de referência com um papel mais passivo. O Alvo T continua sendo acessado via R, mas este exerce um papel mais tênue, sendo assim T é o elemento perfilado assumindo o estatuto de TR. Com isso, a Análise do Ponto de Referência seria a representação esquemática de um possessivo arquetípico, enquanto todas as outras formas de possessivo como propriedade (posse legal sobre algo), parte-todo (partes do corpo), parentesco (relações sociais), etc. seriam instanciações mais específicas de um esquema relacional de posse. De uma representação altamente esquemática como a da Figura 1 até as diferentes instanciações de possessivos das figuras 2(a-c) decorrem outros processos cognitivos como arraigamento (entrenchment) e subjetivização da experiência (LANGACKER, 2009, p. 2-3, 85).<sup>13</sup>

Para o Comitativo, propomos que seja o mesmo caso. Arkhipov considera o Comitativo como uma construção para “pluralizar” um conjunto numa situação assimétrica de forma que os participantes compartilhem da mesma asserção no evento descrito. Fica relativamente fácil de observar a proximidade dessa definição feita por Arkhipov com o que temos visto até o momento sobre os pontos de referência. Porém, como seria na notação da Análise do Ponto de Referência tal relação? Decidimos que uma possível convenção para o Comitativo seria esta apresentada abaixo:



**Figura 4. Comitativo e Análise do Ponto de Referência**

<sup>13</sup> Para arraigamento “entrenchment” (lit. entrincheiramento) entende-se como o estabelecimento de unidades linguísticas que se tornam padrões cognitivos ou rotineiros. Subjetivização é um processo similar ao da gramaticalização (Cf. LANGACKER, 2000).

A linha espessa entre R e T daria a ideia de que duas entidades podem ser conceptualizadas como um único participante, tendo a relação de pontos de referência ainda como a representação esquemática. O Conceptualizador perfila as duas entidades como um único TR, enquanto o Domínio serve como o MR. À primeira vista, a Análise do Ponto de Referência se adequa bem ao que Arkhipov chama de Construção Comitativa (genuína). Quanto às construções quase-comitativas, para o caso das cinco construções descritas acima, sugerimos que são as diversas operações de perspectivização, isto é, as diferentes possibilidades de alinhamento entre TR e MR é que caracterizariam semanticamente as diferentes instanciações de um comitativo.

A ambiguidade observada em *quizombo* no exemplo 3(a) se daria pelo fato de em *quizombo* a preposição *yé* não permitir alinhar o filho como TR juntamente com o pai, precisando fazer uso da palavra “*kumosi*”. Em 3(b), porém, pelo o filho estar juntamente com o pai numa posição de sujeito, ambos são perfilados como TR na sentença, enquanto que o MR, arroz e feijão, são menos proeminentes.

O uso de verbos com valor comitativo para posse já foi explicado nos parágrafos acima, seguindo a explicação de Langacker, a de que um PR é sempre um ponto de referência para o PD. Poderíamos considerar que uma relação prototípica para um comitativo, em que A está perto e acompanhado por B, conforme Figura 4, guarda a mesma relação daquela dos possessivos. *Grosso modo*, uma entidade a ser acompanhada serve como um ponto de referência R para que se acesse a entidade que será a companhia, neste caso, o alvo T da relação de companhia. Relação similar à de possessivo.

Para o uso de instrumento, como em 9(b), o instrumento é perfilado como TR, o primeiro elemento de proeminência, enquanto o pão cortado é o ponto de referência para a faca. A atenção do Conceptualizador passa pelo pão cortado para chegar à faca.

Talvez os casos menos prototípicos para comitativo sejam os das construções existenciais/locativas e seriais. Nelas as partículas de comitativo parecem perder basicamente o sentido prototípico de Companhia para serem usados como marcas de relações espaciais ou de estruturação de evento. Novamente afirmamos que todas elas são construções que tem como base pontos de referência. A habilidade de ponto de referência se mantém latente nessas construções, mesmo que elas sejam instanciações menos prototípicas. Podemos ainda perceber o caráter de ponto de referência de construções existenciais como em (10) e (11) se tomarmos os elementos existentes, pessoas e mel, como o alvo (T) que é perfilado. Em outras palavras, pessoas e mel são perfilados no Domínio (D) de existir, conforme Figura 3(c).

Considerar uma construção serial como em (13) com um sentido comitativo seria o mesmo que tomar o primeiro evento “pegar o inhame” como o ponto de referência para o qual o evento de vir é o alvo (T). Na conceptualização da construção serial como um único evento, os traços de volição e posse do verbo *fa* remetem ao de companhia do objeto de posse, permitindo uma leitura comitativa.

Em todas as construções, a proeminência dada a um ou outro elemento numa relação de ponto de referência é determinante para a forma como os itens lexicais podem integrar distintas construções gramaticais de acordo com o a opção de perfilamento feita pelo Conceptualizador, falante e/ou ouvinte.

## Considerações finais

Acreditamos que a Análise do Ponto de Referência constitui-se como um construto teórico promissor para a compreensão da variedade de casos em que partículas de comitativo atuam em diferentes construções, de modo particular em determinadas línguas. Alguns traços tipológicos de línguas africanas, como o uso de partículas de comitativo para posse predicativa, em expressões espaciais e construções seriais, apontam para essa variedade de casos. Frente a isso, caracterizar uma categoria como o Comitativo a partir de operações cognitivas básicas, como a capacidade de direcionar a atenção a uma entidade por via de outra entidade, ou seja, fazer uso de pontos de referência para a conceptualização de diferentes situações permite que se dê conta do alto grau de sincretismo que uma categoria gramatical pode apresentar. Neste ponto, a análise proposta por Langacker (2000) coaduna-se com a tentativa de Arkhipov (2009) de encontrar uma definição bastante esquemática para a categoria do Comitativo, dando conta das várias instanciações dessa categoria, desde as mais prototípicas até as menos prototípicas.

Ao mesmo tempo em que uma base conceptual é eleita como básica para boa parte dos usos de uma categoria, atentar para o fato de que cada língua tem uma forma específica de lidar com certas categorias, ou seja, tem diferentes formas para a conceptualização e categorização, respeita de certo modo o PARTICULARISMO CATEGORIAL defendido por Haspelmath (2010). Concordamos, no entanto, que a brevidade de como o construto teórico da Análise do Ponto de Referência foi apresentado neste artigo e o leque de construções consideradas deixam a desejar um maior aprofundamento da proposta. Este é nosso intento após esta primeira proposta cognitiva para o Comitativo.

## REFERÊNCIAS

- ABDOULAYE, M. Existential and possessive predication in Hausa. *Linguistics*, v. 44, n. 6, p. 1121-1164, 2006.
- ARKHIPOV, A. Comitativo as a cross-linguistically valid category. In: EPPS, P.; ARKHIPOV, A. *New challenges in typology – transcending the borders and refining the distinctions*. Berlim: Mouton de Gruyter, 2009.
- BONVINI, E. Línguas africanas e português falado no Brasil. In: FIORIN, J. L.; PETTER, M. *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto, 2008.
- COHEN, K. B. *Aspects of the grammar of Kukú*. Munique: LINCOM Europa, 2000.
- CREISSELS, D. Control and the evolution of possessive and existential constructions. [s.d.]. Disponível em: <<http://www.deniscreissels.fr/public/Creissels-ev.poss.exist.constr.pdf>>. Acesso em: maio 2011.
- CREISSELS, D.; KOUADIO, N. *Description phonologique et grammaticale d'un parler baoulé*. Institut de Linguistique Appliquée, Université Nationale de Côte-d'Ivoire, 1977.
- FIORIN, J. L.; PETTER, M. *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto, 2008.

HASPELMATH, M. Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, v. 86, n. 3, p.663-687, 2010.

\_\_\_\_\_. *Coordinating constructions*. (Typological Studies in Language 58) Amsterdam: Benjamins, 2004.

HEINE, B. *Possession*. Cognitive Sources, Forces, and Grammaticalization. Cambridge: Cambridge University Press, 1997.

HEINE, B.; KUTEVA, T. *World Lexicon of Grammaticalization*. Cambridge: CUP, 2002.

HOUAISS, A. *Dicionário eletrônico Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva. Versão 3.0 [CD-ROM], 2009.

KOUADIO, J. Les séries verbales em baoulé: questions de morphosyntaxe et de sémantique. *Studies in African Linguistics*, v. 29, n.1, spring 2000.

LAKOFF, G.; JOHNSON, M. *Metaphors We Live By*. Chicago; London: The University of Chicago Press, 1980.

LANGACKER, R. *Investigations in Cognitive Grammar*. Berlim: Mouton de Gruyter, 2009.

\_\_\_\_\_. Strategies of clausal possession. *International Journal of English Studies*, v. 3, n. 2, p. 1-24, 2003.

\_\_\_\_\_. *Grammar and Conceptualization*. Berlim: Mouton de Gruyter, 2000.

LORD, C. *Historical change in serial verb constructions*. Amsterdam: John Benjamins, 1993.

MOUS, M. The grammar of conjunctive and disjunctive coordination in Iraqw. In: HASPELMATH, M. *Coordinating constructions*. (Typological Studies in Language 58) Amsterdam: Benjamins, 2004.

NEWMAN, P. *The Hausa language: an Encyclopedic Reference Grammar*. New Haven: Yale University Press, 2000.

NURSE, D. *Tense and Aspect in Bantu*. New York: Oxford University Press, 2008.

OKOUDOWA, B. *Morfologia verbal do lembaama*. 2010. 170f. Tese (Doutorado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

ROSCH, E. Cognitive reference points. *Cognitive Psychology*, v. 7, p. 532-547, 1975.

STASSEN, L. *Predicative Possession*. Oxford: Oxford University Press, 2009.

\_\_\_\_\_. AND-languages and WITH-languages. *Linguistic Typology*, v. 4, n. 1, p. 1-4, 2000.

STOLZ, T.; STROH, C.; URDZE, A. *On Comitatives and related categories: a typological study with special focus on the languages of Europe*. Berlin; New York: Mouton de Gruyter, 2006.

STOLZ, T. To be with X is to have X: comitatives, instrumentals, locative, and predicative possession. *Linguistics, An Interdisciplinary Journal of the Language Sciences*, v. 39, n. 2, p. 321-350, May 2001.

TECA, A.; ARAÚJO, P. J. P. *A grammatical sketch of kizombo (H. 16h), dialect of Kikongo (Angola)*. (Em preparação).

TRIBUSHININA, E. *Cognitive reference points: semantics beyond the prototypes in adjectives of space and colour*. The Netherlands: LOT, 2008.

VAN HOEK, K. Pronominal anaphora. In: GEEAERTS, D.; CHYCKENS, H. (Ed.). *The Oxford Handbook of Cognitive Linguistics*. New York: Oxford University Press: 2007.

XAVIER, Francisco da Silva. *Fonologia Segmental e Supra-Segmental do Quimbundo – Variedades de Luanda, Bengo, Quanza Norte e Malange*. 2010. 158f. Tese (Doutorado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

# As cores da fauna e da flora: um dicionário especial composto por cromônimos

(The colors of fauna and flora: a special dictionary composed of chromonyms)

Sabrina de Cássia Martins<sup>1</sup>, Claudia Zavaglia<sup>2</sup>

Universidade Estadual Paulista (UNESP/IBILCE)

[martins\\_sabrina@ig.com.br](mailto:martins_sabrina@ig.com.br), [zavaglia@ibilce.unesp.br](mailto:zavaglia@ibilce.unesp.br)

**Resumo:** O presente trabalho tem como objetivo tratar de algumas considerações surgidas no processo de elaboração de um dicionário especial composto por expressões cromáticas, encontradas dentro dos domínios da Botânica e da Zoologia, bem como apresentar o modelo de macro e micro-estrutura formulados. Nossos esforços estão concentrados na busca por uma estrutura simples, completa e que seja a mais adequada para o público-alvo, isto é, os especialistas e estudantes ligados à área do meio ambiente.

Palavras-chave: dicionário especial, léxico da Biologia, expressões cromáticas.

**Abstract:** This work aims to talk about some considerations arisen from the process of elaboration of a special dictionary composed of chromatic phrases found in Botanic and Zoology domains, as well as to present its macrostructure and microstructure. Our efforts are concentrated on creating a simple and complete structure, which may be the most adequate for the audience, i.e., for specialists and students connected to the Environment area.

**Key-words:** special dictionary, lexicon from Biology, chromatic phrases.

## Introdução

O léxico é o conjunto dos signos linguísticos que representa toda a experiência cultural de uma sociedade, sendo por isso considerado o saber vocabular partilhado que está sujeito a todas as influências sócio-históricas desse povo. É um sistema aberto, estando, portanto, em constante expansão. Entendido como um sistema linguístico, pode ser subdividido em diversos outros microssistemas que caracterizam e nomeiam realidades de uma determinada comunidade. Um desses microssistemas é composto pelos nomes de cores.

Abrangendo combinações fixas das mais variadas naturezas, as expressões cromáticas (doravante EC) estão presentes tanto na nossa linguagem comum, compartilhada pelos falantes da língua, denominando fatos do nosso dia a dia, sensações e emoções, como também nas linguagens de especialidade, nomeando acontecimentos sócio-históricos, processos, compostos químicos, seres orgânicos e inorgânicos, entre tantos outros.

Nos subdomínios da fauna e da flora, por exemplo, as cores são presença constante na denominação das espécies, designando funções diversas e refletindo a variação intra-específica ou fatores ambientais.

---

1 Mestranda pelo Programa de Pós-graduação em Estudos Linguísticos pela Universidade Estadual Paulista, Câmpus de São José do Rio Preto/São Paulo.

2 Livre-Docente em Lexicografia e Lexicologia pela Universidade Estadual Paulista, Câmpus de São José do Rio Preto/São Paulo.

A grande frequência de tais expressões nos subdomínios acima mencionados, aliada à importância dada ao meio ambiente no cenário mundial, nos incentivou a formular um dicionário especial composto apenas por essa fatia especial do léxico.

## **Lexicografia especializada e dicionários especiais**

O dicionário é uma herança cultural transmitida entre as gerações e ampliada, à medida que novas palavras são criadas pela comunidade (BIDERMAN, 2001). Mais que um livro de consulta, é um ponto de referência, correspondendo às exigências de informação e de comunicação dos falantes. Além disso, é visto como instrumento pedagógico, pois tem como objetivo educar, respondendo a questões sobre o emprego das palavras, sua aceitabilidade e denotações/conotações (ZAVAGLIA, 2009).

Embora exista uma ampla tradição lexicográfica, foi só na segunda metade do século XX que suas reflexões críticas e metodológicas tomaram proporção. Até esse momento, a lexicografia permanecia à margem das reflexões e avanços da linguística moderna, sendo considerada por esta uma prática artesanal, cujo caráter científico era constantemente contestado.

Nesse mesmo período, surge o interesse na compilação de obras temáticas e especializadas. Isso se deve a fatores como a dificuldade de se abranger todo o léxico de uma língua numa única obra e a complexidade de se manter uma equipe de profissionais qualificados durante o tempo necessário de elaboração da mesma. Aliado a esses fatores, temos o advento da ciência e da tecnologia que, com a globalização, proporcionou a criação de diversos vocabulários, ligados às mais diversas áreas do saber científico, que necessitam de rápida difusão.

No que diz respeito à composição dos dicionários especiais, Cabré (1993) aventa que, comparado aos dicionários gerais que têm algumas características fixas, a presença de qualquer outra característica que se distancie do padrão dá ao dicionário o caráter de especial ou específico. Bergholtz e Tarp (1995) salientam que a Lexicografia Especializada, assim como a Terminologia, também trabalha com o léxico especializado, abordando algumas características especiais, por exemplo, a) um modelo de macroestrutura sistemático, que organiza as entradas de acordo com os campos semânticos aos quais pertencem, b) equilibra-se entre o caráter descritivo e prescritivo, c) visa tanto à decodificação quanto à codificação, sendo ótimas fontes de consulta para uma área específica. Para os autores,

Como uma parte especial da lexicografia em geral, a Lexicografia Especializada trabalha com as terminologias; a Lexicografia Especializada trabalha tanto com macroestruturas sistemáticas quanto alfabéticas, decidindo de acordo com cada caso qual é a mais apropriada; a Lexicografia Especializada deve necessariamente num grau maior ou menor ser tanto descritiva quanto prescritiva; a Lexicografia Especializada destina-se ao público em geral e aos especialistas; a Lexicografia Especializada prepara dicionários com fins de codificação como de decodificação. (BERGENHOLTZ; TARP, 1995, p. 11, tradução nossa)

É dentre as obras lexicográficas especializadas que se enquadra este trabalho, visto que

- 1) Aborda o vocabulário pertencente a um determinado campo de estudo, a saber, da Biologia, em particular a Botânica e Zoologia;

- 2) É composto por ECs, microsistema lexical característico da linguagem comum, cada vez mais presente nas línguas de especialidade;
- 3) É dedicado a um público específico, os especialistas e estudantes da área do meio ambiente;
- 4) Seu modelo de macroestrutura é onomasiológico, estando em concordância com os pressupostos taxonômicos;
- 5) As informações contidas na microestrutura são específicas da área e a definição transita entre os modelos lexicográfico, terminológico e enciclopédico.

No tocante ao processo de elaboração de obras lexicográficas especiais, este exige do lexicógrafo o planejamento detalhado da obra, de forma a atender às expectativas e necessidades do público-alvo. Sendo assim, é preciso ponderar a metodologia mais adequada e segui-la rigorosamente. A seguir, descreveremos a metodologia adotada para a elaboração do *Dicionário onomasiológico de expressões cromáticas da fauna e flora*.

## **Dicionário Onomasiológico de Expressões Cromáticas da Fauna e Flora**

### **A Macroestrutura: composição da nomenclatura**

Trata-se de um dicionário cuja macroestrutura é composta por expressões cromáticas, isto é, um sintagma nominal que contém em seu interior um nome de cor que pertença a um dos seguintes subdomínios cromáticos: *vermelho, verde, azul, amarelo, preto, branco, cinza, marrom, rosa, alaranjado, roxo, violeta e anil*. Além disso, a expressão deve estar localizada dentro dos domínios da Zoologia, especificamente dos *vertebrados* (anfíbios, répteis, peixes, aves e mamíferos) e da Botânica, especialmente das *angiospermas* (dicotiledôneas).

Para a coleta das ECs, partimos, primeiramente, de duas obras lexicográficas monolíngues brasileiras: o *Dicionário eletrônico Houaiss da língua portuguesa* (HOUAISS, 2009) e o *Novo dicionário eletrônico Aurélio* (FERREIRA, 2010). Aliado a esses dois dicionários, fizemos uso também do *Dicionário brasileiro de botânica* (PEREIRA; PUTZKE, 2010), do *Diccionario das plantas uteis do Brasil e das exóticas cultivadas* (CORRÊA, 1926), do *Dicionário das plantas úteis do Brasil* (CRUZ, 1979), do *Dicionário zoológico* (TIERNO, 1954), e do *Dicionário dos mamíferos do Brasil* (CARVALHO, 1979).

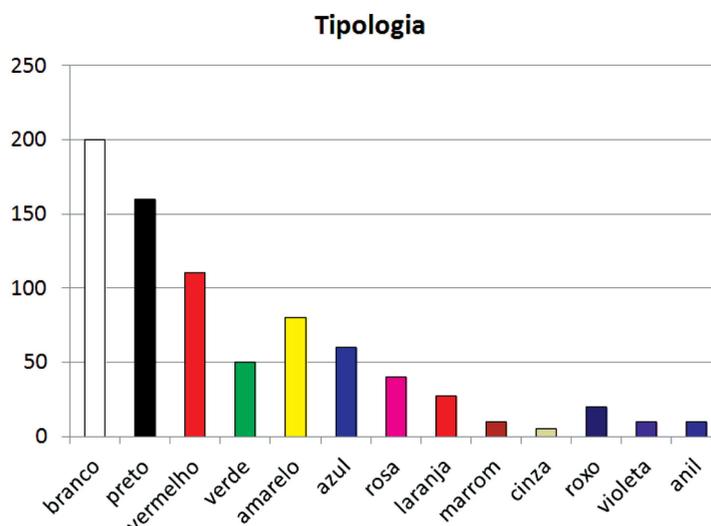
Tal levantamento tem sido aprimorado à medida que seguimos com a pesquisa por meio de consultas em *sites* da web, tais como revistas eletrônicas especializadas, *sites* do governo relacionados ao meio ambiente, *sites* sobre as propriedades medicinais das plantas e de espécies em extinção, etc., bem como consultas a dicionários e glossários *on-line*, tais como a *Avibase*, a *Biota neotropica* e a *Flora brasiliensis*.

A título de exemplificação, listamos a seguir alguns dos itens encontrados até o momento:

**Quadro 1: Exemplos de expressões cromáticas pertencentes à fauna e à flora**

Expressões cromáticas	Reino Planta Angiospermas	Reino Animal Vertebrados
<b>Branco</b>	abeto-branco, abóbora-branca, açai-branco, amoreira-branca, loiro-branco;	anambé-branco, andorinha-de-rabo-branco, anu-branco, beija-flor-de-papo-branco, lavadeira-de-cabeça-branca, rolinha-branca;
<b>Preto</b>	caapiá-preto, cabeça-de-negro, coração-de-negro, feijão-preto, jurema-preta;	badejo-preto, boca-preta, coruja-preta, garoupa-preta, urso-negro;
<b>Vermelho</b>	aroeira-vermelha, bicuíba-vermelha, bredo-vermelho, cacho-vermelho, caeté-vermelho;	bodião-vermelho, boto-vermelho, garça-vermelha;
<b>Amarelo</b>	ameixa-amarela, batatão-amarelo, camarú-amarelo, juá-amarelo;	bagre-amarelo, moreia-amarela, periquito-de-testa-amarela;
<b>Verde</b>	caroba-de-flor -verde, coração-verde, feijão-verde, heléboro-verde, ingá-verde;	beija-flor-de-garganta-verde, japu-verde, Martim-pescador-verde, moreia-verde;
<b>Azul</b>	loto-azul, maracujá-azul, tento-azul, tremoço-de-flor-azul;	donzela-azul, frango-d'água-azul, juriti-azul, saí-açu-azul;
<b>Anil</b>	anil-bastardo, anil-bravo, anil-trepador;	----
<b>Roxo</b>	angico-roxo, cambará-roxo, espinho-roxo, gervão-roxo;	anambé-roxo, bico-roxo;
<b>Violeta</b>	cipó-violeta, pau-violeta, violeta-africana;	beija-flor-de-orelha-violeta;
<b>Cinza</b>	----	tubarão-cinza-da-Groenlândia, tartaranhão-cinza, boto-cinza;
<b>Marrom</b>	----	donzelinha-marrom;
<b>Rosa</b>	cedro-rosa, dedal-de-rosa, ipê-rosa, jumbo-rosa, loiro-rosa, malva-rosa;	pescada-rosa, boto-cor-de-rosa, camarão-rosa, congro-rosa;

**Gráfico 1: Ocorrências dos subdomínios cromáticos**



Como pode ser observado a partir do quadro e do gráfico acima, a maior parte das expressões insere-se no subdomínio *branco*, seguido por *preto*, *vermelho*, *amarelo*, *azul* e *verde*. Os subdomínios *anil* e *marrom* foram encontrados, até o momento, apenas no domínio da Flora e em pequena quantidade. Também nesse domínio, é comum encontramos expressões cromáticas de diferentes campos como sinônimos da mesma espécie, por exemplo, o vermelho e o rosa (araçá-vermelho e araçá-rosa).

### A Macroestrutura: o modelo onomasiológico

O modelo de macroestrutura é o onomasiológico, composto por um esquema simples e completo que está de acordo com os princípios da Sistemática, ramo da Biologia que também se serve dos princípios da Onomasiologia para categorização das espécies. Segundo Amabis e Martho (2001), a proposta desse ramo é nomear os seres de forma que seus nomes os agrupem de acordo com as características em comum, identificando-os e proporcionando seu estudo. Desse modo, foi criada uma classificação decrescente que vai do *Reino* (mais genérico) até a *espécie* (mais específico). O sistema taxonômico permite, então, o agrupamento dos seres vivos conforme o grau de parentesco e possibilita ainda compreender a sua evolução. O modelo de dicionário onomasiológico, por sua vez, possibilita a permanência desse agrupamento, facilitando a pesquisa.

Tal estrutura foi configurada com o auxílio de um profissional da área de ciências biológicas e gerou um modelo de macroestrutura ordenado de acordo com os grupos hierárquicos, a saber:

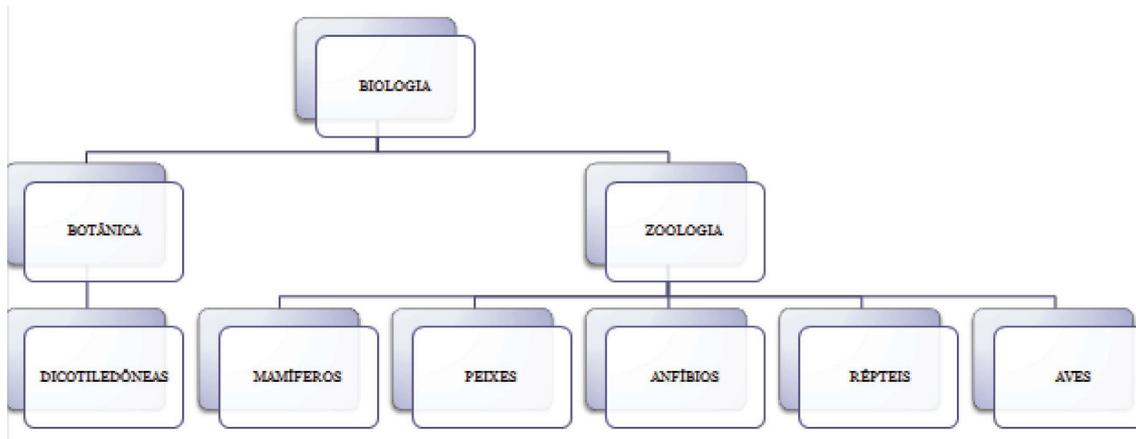
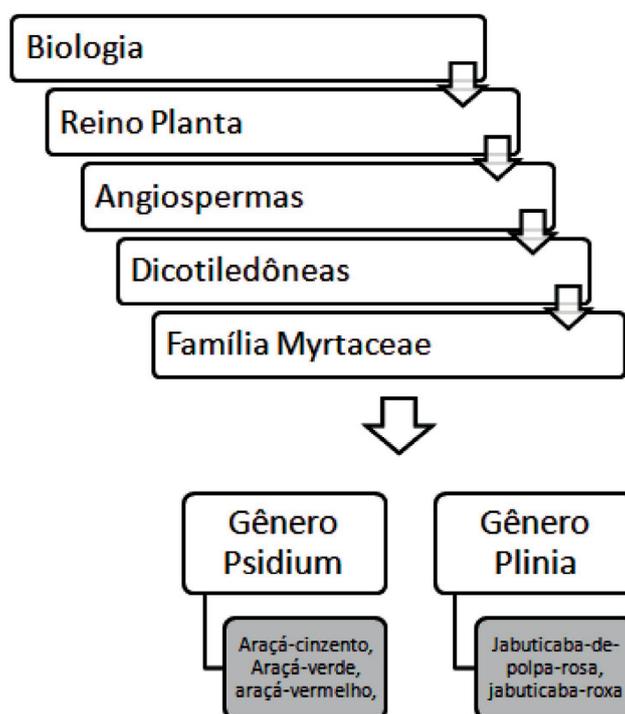


Figura 1: Modelo onomasiológico de macroestrutura

Como exemplo, citamos a classificação do item “araçá”, dentro do dicionário:



**Figura 2: Classificação das expressões cromáticas dentro da macroestrutura**

As famílias são separadas de acordo com a classe às quais pertencem e conforme suas características fundamentais. Dentro de cada família, as entradas são ordenadas alfabeticamente.

### A Microestrutura

A microestrutura é a base do dicionário, por isso pretendemos formular um modelo que atenda às expectativas do público-alvo, que seja de fácil compreensão e que reúna o máximo de informações possíveis em uma definição redigida de forma clara, além de exemplos de uso e sinônimos. Dessa forma, propomos que o verbete seja composto:

- pela palavra-entrada, no caso a expressão cromática;
- por informações linguísticas sobre o cromônimo, tais como classificação gramatical e plural irregular;
- nome científico da espécie;
- definição do cromônimo. No que diz respeito à definição, destacamos que ela é uma mistura de definição lexicográfica, visto que parte do gênero ou hiperônimo para as características específicas, terminológica, pois aborda as informações necessárias para a compreensão de um conceito, e enciclopédica, já que descreve as relações do objeto com o mundo real. Como o dicionário é destinado diretamente a especialistas das ciências biológicas, além de outras áreas afins, a definição é composta pela terminologia da área, reunindo informações específicas que diferenciem a EC de outras espécies da mesma família, porém sem trazer informações muito gerais.

- Contextualização real do cromônimo. Para a extração dos exemplos, fizemos uso do *corpus* web. Considerada uma revolução dos meios de comunicação, a web proporciona não só a conexão entre as pessoas, como também o compartilhamento de informações. Para Berber Sardinha (2003, p. 192), “o conjunto das interações e informações disponibilizadas na rede, continuamente, dia a dia, em qualquer parte do mundo em que a infraestrutura esteja presente, resulta num gigantesco *corpus* dessa interação”. A fim de apontar para o uso das expressões cromáticas tanto na linguagem comum quanto na especializada, destacando o seu funcionamento em ambas as formas de comunicação, concluímos que a web seria o *corpus* ideal, pois contém uma grande variedade linguística que vai desde publicações acadêmicas até blogs e fóruns de discussão.
- Sinônimos da EC, quando encontrados.

Assim, teremos o seguinte modelo de verbete:

**EXPRESSÃO CROMÁTICA** [inf. linguísticas] *nome científico*; ● definição  
 → contexto em linguagem especializada ⇔ contexto em linguagem comum  
**SINÔNIMOS**

Figura 3: modelo de verbete

Tal modelo dará origem a verbetes como:

REINO PLANTAE; DIVISÃO: MAGNOLIOPHYTA  
 ORDEM: MAGNOLIALES; FAMÍLIA: MYRISTICACEAE

**BICUÍBA-BRANCA** [s.f.; pl. bicuíbas-brancas; var. bicuíba branca] *Virola bicuhyba*, *Virola officinalis* ● Planta nativa do Brasil que pode ser encontrada nas regiões Sul, Sudeste e Nordeste, geralmente em florestas pluviais. É uma árvore de porte médio, podendo chegar a 25 m de altura. Apresenta folhas simples, inteiras, alado-pecioladas, ovado-oblongas; flores apétalas dispostas em racimos axilares; fruto baga drupácea, globosa e pequena. Sua madeira é de coloração pardo-clara, passando a vermelho-pardacenta com o decorrer do tempo; é moderadamente pesada, resistente e durável, sendo fácil de trabalhar. → Sua área de ocorrência é na floresta pluvial que se estende de Pernambuco ao Rio de Janeiro, ocorrendo no Parque Florestal do Rio Doce (MG) alguns exemplares de 50 m de altura e 0,72 m de diâmetro e, na mesma área, ocorre uma espécie afim, denominada **Bicuíba branca** (*Virola officinalis*). ⇒ O gênero *Virola* é bem representado por outras espécies, nas demais regiões do país, principalmente na zona litorânea, desde o Estado do Espírito Santo até o Estado de Santa Catarina, conhecidas por **BICUIBA-BRANCA**, **BICUÍBA-ROSA**, **BUCUVA** e **BUCUVUÇU** e, eventualmente, por **VIROLA**. **SIN. BICUÍBA-REDONDA, BICUÍBA, BUCUVA, BUCUVUÇU, VIROLA, BICUIBAÇU**

<http://ipr.dnit.gov.br/manuais/Manual%20de%20Vegetacao%20Rodoviaria%20-%20VOLUME%201.pdf>  
<http://www.refrima.com/pag/ucuuba-branca.html>

**UCUUBA-BRANCA** [s.f.; pl. ucuubas-brancas; var. ucuuba branca] *Virola surinamensis* ● Planta nativa do Brasil que pode ser encontrada nas margens alagáveis dos rios e igapós, como também em várzeas, principalmente nos estados do Acre, Amazonas, Rondônia, Roraima, Amapá, Pará, Tocantins, Maranhão e Ceará e nos países da Amazônia Central. É uma árvore de grande porte, medindo aproximadamente 30 a 35 m de altura e 60 a 100 cm de diâmetro, com troncos retos, horizontais. As folhas são alternas, dísticas, curto-pecioladas e simples; a lâmina possui formato oblongo, oblongo-lanceolado ou lanceolado, consistência coriácea, de 10-25 cm

de comprimento e 2-5 cm de largura; as flores são de coloração amarelo-pálido. Sua madeira possui textura média; alburno de coloração creme-claro; e cerne mais escuro, variando do levemente rosado ao castanho-vermelho intenso. É empregada em construção de interiores, carpintaria, marcenaria e na fabricação de caixas, palitos de fósforo, laminados, compensados, celulose e papel. → O resultado mostrou que na análise sensorial feita com as espécies não houve diferença estatística ( $P > 0,05$ ) entre as madeiras angelim-rosa (*Platycamus regenellii* Benth, 1974), muiracatiara (*Astronium lecointei* Ducke, 1988) e virola ou ucuúba-branca (*Virola surinamensis* (Rol.) Warb, 1988), porém a defumação feita com mandioqueira (*Ruizterania albiflora* Marcano Bert, 1988) foi diferente apenas para a sardinha. ⇒ *Virola Surinamensis* que eu encontrei na praia. (**Ucuúba Branca**). **SIN. UCUÚBA, UCUÚBA-CHEIROSA, UCUÚBA-DE-IGAPÓ, UCUÚBA-DA-VÁRZEA, UCUÚBA-VERDADEIRA, ÁRVORE-DE-SEBO, VIROLA**

[http://www.pubvet.com.br/artigos\\_det.asp?artigo=172dc381.4shared.com/doc/ADqR\\_6kR/preview.html](http://www.pubvet.com.br/artigos_det.asp?artigo=172dc381.4shared.com/doc/ADqR_6kR/preview.html)

**UCUUBA-VERMELHA** [s.f.; pl. ucuubas-vermelhas; var. ucuuba vermelha] *Virola sebifera* ● Árvore nativa do Norte da América do Sul. Apresenta casca com seiva de coloração avermelhada folhas coriáceas; flores em panículas; e frutos ovoides, cujo endocarpo contem uma espécie de gordura de cor amarela clara, consistente, muito usada na fabricação de velas e sabões. → Também apresenta semelhança com a *V. sebifera*, conhecida como ucuuba-vermelha devido à seiva avermelhada da casca, da qual se diferencia por possuir seiva incolor. ⇒ Informações fidedignas bem colocadas, a casca da ucuuba vermelha tem também um valor olfativo muito bom. **SIN. ÁRVORE-DE-SEBO, BICUÍBA, BUCUVA, CANANGA, UCUUBA, UCUUBEIRA, UXICURUÁ, UCUÚBA-DO-CERRADO, GORDURA-DE-VIROLA**

<http://www.iepa.ap.gov.br/arquivopdf/itucubabranca.pdf>

<http://blog.multivegetal.com/ucuuba-fonte-de-beleza-e-saude-para-a-pele/>

REINO: ANIMALIA; FILO: CHORDATA; CLASSE: REPTILIA ORDEM: SQUAMATA; FAMÍLIA: BOIDAE (SUBORDEM: OPHIDIA)
---

**JIBOIA-VERDE** [s.f.; pl. jiboias-verdes; var. jiboia verde] *Corallus caninus* ● Serpente amazônica de aproximadamente 2,2 m de comprimento, que apresenta corpo de coloração verde na região dorsal com listras transversais branco-amareladas, e região ventral amarela. É uma espécie constritora, não peçonhenta e de hábitos arborícolas e noturnos. Alimenta-se de roedores, pequenas aves e répteis. → Pode-se também encontrar o jacaré-tinga (*Caiman crocodilus*), o jacaré-açu (*Melanosuchus niger*), as cobras surucucu (*Lachesis muta*) e jibóia-verde (*Corallus caninus*), além de cinco variedades de rãs. ⇒ Cobra papagaio, periquitambóia, jibóia verde ou *Corallus caninus* exibindo seu look para a copa do mundo. **SIN. COBRA-PAPAGAIO, JIBÓIA-DA-AMAZÔNIA, ARABÓIA, ARARAMBÓIA, ARAUEMBÓIA, BOA, BOA-ARBORÍCOLA-ESMERALDA, PERIQUITAMBÓIA, PÍTON-VERDE-DA-ÁRVORE**

[http://www.livrosgratis.com.br/arquivos\\_livros/mre000041.pdf](http://www.livrosgratis.com.br/arquivos_livros/mre000041.pdf)

<http://cunhabio.blogspot.com.br/2010/05/cobra-papagaio-periquitamboa-jiboia.html>

**JIBOIA-VERMELHA** [s.f.; pl. jiboias-vermelhas; var. jiboia vermelha] *Epicrates cenchria cenchria* ● Serpente encontrada na Venezuela, Guianas, Peru e na região amazônica do Brasil. De aproximadamente 2,2 m de comprimento, apresenta dorso pardo-avermelhado com manchas negras e ventre amarelado. Alimenta-se de roedores, lagartos e pequenas aves. Não é peçonhenta e é muito utilizada como animal de estimação devido a sua beleza. → Moradores do Povoado Rajada, próximo ao município de Carnaúba dos Dantas, encontraram na manhã desta terça-feira (4), uma cobra do tipo salamanta, conhecida também como jiboia vermelha, medindo 2.20m. ⇒ Ela também é conhecida como Uaçubói e Jiboia Vermelha. Ela tem várias Cores pelo Corpo, como Vermelho, Preto e Branco. **SIN. SURUCUCU-DE-FOGO, SALAMANTA, SUAÇU, UAÇUBÓI**

[http://www.dnonline.com.br/app/noticia/cotidiano/2011/10/04/interna\\_cotidiano,82420/cobra-com-mais-de-dois-metros-e-encontrada-no-interior-do-rn.shtml](http://www.dnonline.com.br/app/noticia/cotidiano/2011/10/04/interna_cotidiano,82420/cobra-com-mais-de-dois-metros-e-encontrada-no-interior-do-rn.shtml)

<http://brunochavesanimais.blogspot.com/2011/10/salamanta.html>

## A questão da univocidade na Biologia: a homonímia e a polissemia

O princípio de homogeneidade caracterizou por muito tempo as unidades lexicais especializadas que eram vistas como uniformes, atemporais, associadas e neutras ideologicamente. De fato, as terminologias sempre foram particularizadas pela precisão conceitual tão requerida para o intercâmbio comunicacional entre especialistas de diversas áreas ao redor do globo.

Entretanto, como bem assinala Krieger (2001), os itens lexicais especializados não são constructos ideais e homogêneos restritos à comunicação dos especialistas. Ao contrário, elas são elementos constitutivos das linguagens naturais, estando presentes no nosso dia a dia e, por isso, são passíveis de toda forma de variação.

No campo da Biologia, é comum encontrarmos uma mesma EC que denomine diferentes espécies, pertencentes tanto à mesma família, porém a gêneros diferentes, quanto a famílias diversas. Segundo Quicke (1996), isso ocorre quando pesquisadores utilizam o mesmo nome científico para denominar diferentes táxons, o que se dá pelo fato de eles tentarem expressar as características físicas dos animais por meio do vocabulário latino ou grego, o que torna provável a utilização da mesma palavra.

Dessa forma, no decorrer dessa pesquisa, consideramos como caso de polissemia, isto é, “formas que apresentam ao menos um sema em comum” (ZAVAGLIA, C.; ZAVAGLIA, A., 2002, p. 241), a expressão *CAMBUÍ-ROXO*, que faz referência tanto à espécie *Eugenia candolleana* quanto à espécie *Myrtus rubra*, ambas pertencentes à família Myrtaceae, e que, portanto, apresentam algumas características em comum.

Tais casos são registrados no dicionário na mesma entrada, porém aparecerão dentro do verbete separados pelo gênero ao qual pertencem, numerados e seguidos pela definição e pelo contexto. Optamos pela presença do contexto em apenas uma das acepções, devido à dificuldade de se encontrar um exemplo adequado em que figurasse tanto o cromônimo quanto o nome científico, informação esta essencial para a distinção da espécie.

Como exemplo, citamos:

ORDEM: LAURALES; FAMÍLIA: LAURACEAE

**CANELA-AMARELA** [s.f.; pl. canelas-amarelas; var. canela amarela] **1.** *Nectandra lanceolata* (Nees & Mart.), *Nectandra leucantha*, *Nectandra oreadum*, *Nectandra pichurim*, *Nectandra puberula*, *Nectandra weddellii*, *Ocotea pichurim*, *Cryptocarya moschata* ● Árvore endêmica da Mata Atlântica, de médio porte que não ultrapassa os 25 m de altura e 90 cm de diâmetro. Sua madeira foi muito usada durante décadas no Brasil na confecção de aberturas, (portas e janelas) por ser uma madeira muito resistente e de fácil manejo. Por consequência, hoje quase não se encontra espécie adulta desta árvore, já que leva em torno de 25 a 35 anos para atingir seu auge. → *Caso do pirapitinga, peixe que se alimenta das sementes da canela amarela, árvore que na região de Visconde de Mauá ocorre em 90% da mata ciliar.* ⇒ *A canela amarela e a canela preta fizeram parte da vida de muitos brasileiros que se quer tiveram conhecimento, serviram de assoalhos e tetos durante anos.* **SIN. CANELA-DA-VÁRZEA,**

**CANELA-FEDORENTA, CANELA-LOURO, CANELA-VERMELHA, ESPORA-DE-GALO 2.** *Nectandra ambigua* ● Árvore nativa do Brasil, podendo ser encontrada em todo o país. De aproximadamente 10 m, apresenta madeira de coloração amarelada; folhas coriáceas; flores aromáticas de coloração branco-amareladas; bagas ovoides; raiz de casca amarga e tônica. Sua madeira é muito utilizada na marcenaria e carpintaria. **SIN. CANELA-SECA 3.** *Ocotea diospyrifolia* ● Árvore nativa da Argentina, Bolívia, Paraguai e Brasil, ocorrendo nas regiões centro-oeste, sudeste e sul do país. Pode ser encontrada em florestas estacionais semidecíduais, florestas ciliares e em várzeas. Tal espécie é típica de formações abertas, principalmente em matas de galeria e savanas úmidas. De aproximadamente 10 m de altura, apresenta casca tanífera; suas folhas são coriáceas; as flores são suavemente perfumadas e em panículas piramidais; as bagas são globosas. Sua madeira é de boa qualidade, sendo muito utilizada para postes e tábuas de assoalho. **SIN. CANELA-BARAÚNA, BATALHA, CANELA, LOURO-AMARELO, CANELA-LOURO, CANELÃO, CANELÃO-DE-MÓVEIS, CANELEIRO 4.** *Ocotea aciphylla* (Nees) Mez, *Aydendron Tenellum* ● Árvore nativa do Brasil, ocorrendo principalmente nos estados do Amazonas, Minas Gerais e São Paulo, de aproximadamente 20 m. Apresenta folhas esparsas, flores brancas em panículas terminais e casca aromática. **SIN. LOURO-AMARELO, CANELA-AMARELA-DE-CHEIRO, CANELA-BRANCA, CANELA-POCA, CANELA-PORCA, CANELA-PORCA, LOURO-AMARELO-DE-CHEIRO, LOURO-INAMUÍ-DA-TERRA-FIRME**

<http://www.viscondeesconde.com.br/historias01.asp?pp=101>

<http://canelasecanelas.blogspot.com.br/>

Do mesmo modo, consideramos como caso de homonímia, isto é, “formas que possuem mais de um significado para um mesmo significante” e que possuem ao menos um traço semântico distintivo (ZAVAGLIA, C.; ZAVAGLIA, A., 2002, p. 241), a expressão *BRANCA-URSINA*, que faz referência à espécie *Acanthus mollis* da família Acanthaceae e à espécie *Heracleum sphondylium* da família Apiaceae. Uma vez que não pertencem à mesma família, e por isso não compartilham de traços distintivos comuns, são consideradas como formas homônimas.

Tais casos serão registrados em entradas separadas e incluídas dentro das respectivas famílias, constando do final do verbete uma remissiva que fará referência à outra espécie, como pode ser observado a partir do exemplo a seguir:

ORDEM: APIALES; FAMÍLIA: APIACEAE

**BRANCA-URSINA** [s.f.; pl. brancas-ursinas; branca ursina] *Heracleum sphondylium* ● Planta nativa da Europa que pode chegar a 1,5m de altura. Apresenta caule longo e rígido; folhas verde-acinzentadas grandes e recortadas; flores brancas agrupadas em umbela. → Há uma grande confusão a respeito do nome “*Branca ursina*” usado popularmente para designar duas plantas diferentes: *Heracleum sphondylium* e *Acanthus mollis*. ⇒ Em Portugal a espécie da mesma família é a *Heracleum sphondylium* L., também conhecida como *Branca-ursina*. **SIN CANABRÁS, ESFONDILIO V. BRANCA-URSINA, família Acanthaceae**

[http://www.cesaho.com.br/biblioteca\\_virtual/arquivos/arquivo\\_212\\_cesaho.pdf](http://www.cesaho.com.br/biblioteca_virtual/arquivos/arquivo_212_cesaho.pdf)

[http://www.proz.com/kudoz/english.../2277115-giant\\_hogweed.html](http://www.proz.com/kudoz/english.../2277115-giant_hogweed.html)

**BRANCA-URSINA** [s.f.; pl. brancas-ursinas; var. branca-ursina] *Acanthus mollis* ● Planta herbácea originária da região mediterrânea. De aproximadamente 2 m de altura, apresenta grandes folhas de cor verde escuro e flores tubulares de cor branca, lilás ou rosa com espinhos verdes ou arroxeados, surgindo acima da folhagem. Floresce no final da primavera ou início do verão. Ela cresce em áreas secas, e é tolerante à seca e à sombra. Suas folhas, flores e raízes apresentam propriedades medicinais, como: adstringente, analgésico, antidiarreico, anti-inflamatório, expectorante, laxante e cicatrizante. → *Os sintomas seguintes erradamente foram incluídos na versão anterior do Radar como pertencentes a esta planta devido ao o nome popular Branca ursina ser usado para duas espécies diferentes.* ⇒ Popularmente é conhecida por Acanto, erva gigante e branca-ursina. É uma herbácea vivaz e robusta, originária da região mediterrânea. **SIN. ACANTO, ACANTO-NEGRO, ERVA-GIGANTE, PÉ-DE-URSO, GIGANTE, ACANTO-MANSO** v. **BRANCA-URSINA**, Família *Apiaceae*

[http://www.cesaho.com.br/biblioteca\\_virtual/arquivos/arquivo\\_212\\_cesaho.pdf](http://www.cesaho.com.br/biblioteca_virtual/arquivos/arquivo_212_cesaho.pdf)

<http://mariapuresa.blogspot.com/2011/10/acanto-ou-branca-ursina.html>

### Considerações finais

Como dito no início deste texto, o dicionário, além de ser o acervo da cultura de um povo, também é uma ferramenta fundamental para a transmissão do conhecimento, seja ele linguístico ou relacionado à realidade em que estamos inseridos. Engana-se, porém, aquele que tem a ilusão de que a composição de tais obras seja uma tarefa fácil. Antes, é um trabalho árduo que requer do lexicógrafo disciplina, persistência e foco. Procuramos descrever a metodologia adotada para o *Dicionário onomasiológico de expressões cromáticas da fauna e flora*, projeto de doutorado já concluído que teve como objetivo complementar o mercado das obras de referência, registrando essa fatia especial do léxico e, além de tudo, examinando como as ECs colaboram para a ampliação do léxico da língua portuguesa.

### REFERÊNCIAS

- AMABIS, J. M.; MARTHO, G. R. *Conceitos de Biologia*. São Paulo: Moderna, 2001.
- BERBER SARDINHA, T. Que tipo de *corpus* é a Web? *Revista da ANPOLL*, v.1, n.15, p. 191-220, 2003.
- BERGENHOLTZ, H.; TARP, S. *Manual of Specialised Lexicography: The Preparation of Specialised Dictionaries*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1995.
- BIDERMAN, M.T.C. *Teoria Linguística*. 2.ed. São Paulo: Martins Fontes, 2001.
- CABRÉ, M. T. *La terminología*. Teoría, metodología, aplicaciones. Barcelona: Editorial Antártida/Empúries, 1993.

KRIEGER, M. G. Terminologia revisitada. In: KRIEGER, Maria da Graça, MACIEL, Anna Maria Becker. (Org.) *Temas de Terminologia*. Porto Alegre/São Paulo: Ed. Universidade/UFRGS/Humanitas/USP, 2001, p. 47-60.

QUICKE, D. L. J. *Principles and Techniques of Contemporary Taxonomy*. 2. ed. London: Blakie Academic Professional, 1996.

ZAVAGLIA, C.; ZAVAGLIA, A. A elaboração de um dicionário trilingüe temático de cromônimos italiano-português-francês/francês-português-italiano: reflexões e considerações. *Linguística*, São Paulo, v. 12, p. 235-247, 2002.

ZAVAGLIA, C. *Sistematização crítica em Lexicografia e Lexicologia*. São José do Rio Preto, 2009. 92f. Tese (Livre-docência em Lexicologia e Lexicografia) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto.

## DICIONÁRIOS

CARVALHO, C. T. de. *Dicionário dos mamíferos do Brasil*. 2. ed. rev. São Paulo: Nobel, 1979.

CORRÊA, M. P. *Diccionario das plantas uteis do Brasil e das exóticas cultivadas*. Rio de Janeiro: Imprensa Nacional, 1926.

CRUZ, G. L. *Dicionário das plantas úteis do Brasil*. Rio de Janeiro: Civilização Brasileira, 1979.

FERREIRA, A. B. de H. *Novo dicionário eletrônico Aurélio versão 7.0*. 5. ed. Curitiba: Positivo Informática Ltda., 2010.

HOUAISS, A. *Dicionário eletrônico Houaiss da língua portuguesa*. Versão 1.0. Objetiva, 2009.

PEREIRA, A. B.; PUTZKE, J. *Dicionário brasileiro de botânica*. Curitiba: CRV, 2010.

TIERNO, J. C. *Dicionário zoológico: contendo, por ordem inversa, todos os termos registrados nos dicionários mais correntes da língua portuguesa*. Lisboa: Tertulia Edípica, 1954.

# Proposta de sistema fonológico para a língua dos Cayapó do Sul aldeados em São José de Mossâmedes

(Proposal for a phonological system for the language spoken by the South Cayapó speakers settled in São José de Mossâmedes)

Eduardo Alves Vasconcelos<sup>1</sup>

<sup>1</sup>Instituto de Estudos da Linguagem – Universidade Estadual de Campinas (Unicamp)

dudualves@gmail.com

**Abstract:** The existing linguistic information about the language spoken by South Cayapós was restricted to seven lists of words. Three lists were collected among those villagers gathered near Vila Boa (GO); two were recorded in 1819; one was done by Emmanuel Pohl (1772-1834) and another by Auguste de Saint-Hilaire (1779-1853). The third record is the oldest and possibly the one that was noted with less accuracy, it is a list of names in a baptism record, dated 1782. In this article, I present a proposal of graphematic analysis for the last record and I hold a discussion about the phonological system of South Cayapó language: its oppositions, processes and representation.

**Keywords:** South Cayapó, word list, Jê languages, phonology.

**Resumo:** As informações linguísticas existentes sobre a língua que foi falada pelos Cayapós do Sul restringem-se a sete listas de palavras, das quais, três foram coligidas entre aqueles aldeados nas proximidades de Vila Boa (GO). Destas listas, duas foram anotadas no ano de 1819, um por Emmanuel Pohl (1772-1834) e outro por Auguste de Saint-Hilaire (1779-1853). O terceiro registro é o mais antigo e possivelmente o que foi anotado com menos acuidade, trata-se de uma lista de nomes contida em um registro de batismo datado de 1782. Neste artigo, é apresentada uma proposta de análise grafemática para este último registro e realizada uma discussão sobre o sistema fonológico do Cayapó do Sul: suas oposições, processos e representação.

**Palavras-chave:** Cayapó do Sul, lista de palavras, línguas Jê, fonologia.

## Introdução

As listas de palavras coletadas por Emmanuel Pohl (1782-1834) e Auguste de Saint-Hilaire (1779-1853) entre os Cayapó do Sul<sup>1</sup> aldeados em São José de Mossâmedes (Goiás) já foram tematizadas em outro texto (VASCONCELOS, 2009), em que faço uma análise grafemática destas listas, mas não proponho análise fonológica. Neste presente estudo, volto à discussão das interpretações possíveis a estas listas, lançando mão de outras evidências sobre as características dos registros e buscando informação em uma terceira fonte para os Cayapó do Sul da região de Mossâmedes: nomes próprios presentes em um registro de Batismo de 1782, referente ao primeiro grupo Cayapó do Sul aldeado em Maria I.<sup>2</sup>

Quanto à interpretação de transcrições não fonéticas de línguas indígenas, adotarei os procedimentos utilizados em estudos como de Grannier Rodrigues (1990)<sup>3</sup> para o Guaraní

1 Povo indígena que foi contatado no século XVIII, quando os primeiros bandeirantes adentravam ao território que compõe hoje parte do estado de Goiás.

2 “Esta aldeia foi construída em 1781, nas margens do rio Fartura, a doze léguas de Vila Boa, no Governo de Luís da Cunha de Menezes.” (ATAÍDES, 1998, p. 79)

3 A dissertação foi defendida em 1974. Aqui utilizo a versão publicada em livro de 1990.

Antigo,<sup>4</sup> no qual, a partir do material produzido pelo Padre Ruiz de Montoya, ela busca identificar e analisar a fonologia daquela língua, partindo da seguinte metodologia:

(a) indicação explícita da natureza do som por Ruiz de Montoya; (b) o valor que tinham no Espanhol do século 17 as letras utilizadas na escrita do Guaraní; (c) particularidades da escrita do Espanhol de Ruiz de Montoya; (d) a ordenação alfabética do Tesoro; (e) a provável coerência do sistema fonológico do Guaraní; (f) alternâncias morfofonológicas do Guaraní; e (g) a situação correspondente em dialetos guaranis atuais. (GRANNIER RODRIGUES, 1990, p. 15)

Tal metodologia é aplicada neste estudo, com as devidas adaptações, a saber: (1) indicação explícita da natureza do som pelos anotadores; (2) o valor que tinham as letras utilizadas pelos anotadores na grafia adotada; (3) particularidades do registro; (4) alternâncias morfofonológicas que o registro permite depreender; (5) correspondências em outros registros do Cayapó do Sul e (6) correspondências em línguas da família Jê. A análise fonológica guia-se pelos princípios da Fonologia de Praga (TRUBETZKOY, 1939; JAKOBSON; FANT; HALLE, 1952) e pelos desenvolvimentos da Fonologia Autossegmental, na vertente da Geometria de Traços, principalmente aqueles levado a cabo por D'Angelis (1998, 2002).

### **Registro de Batismo de Vila Boa (1782)**

Luís da Cunha Menezes (Governador da Província de Goiás de 1778-1783) ao noticiar a coroa portuguesa sobre o aldeamento dos Cayapó do Sul envia, em anexo ao seu comunicado, uma lista das crianças que foram batizadas no ano de 1782, na Matriz de Vila Boa, em cerimônia realizada pelo Vigário João Antunes de Noronha. Este documento ganha importância quanto ao registro da língua que fora falada pelos Cayapó do Sul por constar nele os nomes indígenas dos pais das crianças: são exatamente 60 nomes próprios, anotados em ortografia de língua portuguesa, e, possivelmente, pelo responsável pela cerimônia.<sup>5</sup> Trata-se do registro mais antigo da língua que foi falada por este povo e apesar da falta de acuidade e preocupação com os significados dos termos, é uma peça relevante na tentativa de propor o funcionamento e a organização do sistema fonológico do Cayapó do Sul.

A seguir, apresento a lista dos nomes próprios do registro de Vila Boa e, na sequência, uma hipótese de interpretação grafemática.

---

4 Araújo (1992), ao realizar análise do *Wörterbuch der botokudensprache* de Bruno Rudolph, esclarece: “na determinação dos procedimentos a serem adotados na análise [...], de grande auxílio foi o trabalho de D. [Grannier] Rodrigues (1974), que apresenta uma análise fonológica do Guaraní do século XVII, documentado pelo Padre Ruiz de Montoya” (ARAÚJO, 1992, p. 20).

5 A cópia enviada à coroa portuguesa não está assinada, o que levanta dúvidas inclusive sobre quem foi o responsável pela cópia do documento paroquial.

Lista de nomes próprios anotada em Vila Boa (GO) em 1782:

1. Angrayocha	13. Ungaptuai	25. Apluace	37. Banequere	49. Unxepaan
2. Chiunequa	14. Combono	26. Xecrâ	38. Xuxiê	50. Inqueciare
3. Chaquenonan	15. Xuanampiae	27. Enconâ	39. Caxique	51. Cananquete
4. Cequaquai	16. Carampea	28. Tuuntu	40. Incunan	52. Poimre
5. Queampia	17. Pitubâ	29. Amixon	41. Canampuaxi	53. Mixiela
6. Yamaroi	18. Cambriopixom	30. Tanquerê	42. Toyoto	54. Pamquaque
7. Caçacabe	19. Canapixo	31. Emponi	43. Chaponhece	55. Tonito
8. Coimpa	20. Incapuim	32. Xampeu	44. Xampea	56. Conapicici
9. Pembaque	21. Quipanto	33. Poquea	45. Xataqueare	57. Carampea
10. Romexi	22. Potuaré	34. Bazeque	46. Pupuare	58. Cocri
11. Xaquean	23. Potinii	35. Taxiú	47. Caceterê	59. Tuinta
12. Uncrixiu	24. Acotamacê	36. Xapamapixixi	48. Ecotpaen	60. Canampui

Neste registro [p, b, t] estariam representados por *p*, *b* e *t*, enquanto [k] seria representado tanto por *c* quanto por *qu*, aquele quando diante de *a*, *o*, *u* e *qu* para *i*, *e*; ainda [g] representado por *g*, porém, somente precedido por *n* (cf. 1 e 13), sugerindo que essa sequência poderia ser uma representação para a nasal velar [ŋ], ou mesmo uma pré-nasalizada [ŋ̃g]. Consoantes pré-nasalizadas como [ŋ̃g], seriam representadas, nesta lista, pela sequência grafema nasal (*m*, *n*) seguido do grafema das obstruintes. Outra obstruinte descontínua sonora encontrada no registro é [b], que também pode ser associada às pré-nasalizadas: [mb] ~ [b]. Uma possibilidade de interpretação para estas ocorrências é que sejam consoantes pré-nasalizadas que se realizam com contorno oral diante de uma vogal oral [m̃p, m̃b, ñt, ŋ̃k, ŋ̃g], mas diante de vogal nasal realizar-se-iam como nasal [m, n, ŋ].

O grafema nasal, tal como nas demais listas Cayapó do Sul, indicaria: (i) consoante nasal, (ii) marca da nasalidade da vogal ou, ainda, (iii) nasalidade intrínseca à obstruinte seguinte. Nos registros em que há ocorrência das obstruintes sonoras elas apresentam variação com a contraparte surda ou com a contraparte nasal. Curiosamente, nesta lista, as obstruintes sonoras registradas são aquelas que compartilham o traço grave, lançando mão aqui da proposta de traços acústicos de Jakobson, Fant & Halle (1952).

As obstruintes contínuas [s], [z] e [ʃ] estão representadas, neste registro, por: *c*, *ç* para [s], *z* para [z] e *ch*, *x* para [ʃ]. Não há indícios para as africadas [ts̃] e [tʃ̃], encontradas tanto na lista de Pohl (1832) quanto na de Saint-Hilaire (1848). A sonora [z] está restrita a uma única ocorrência: 34. Bazeque.

Nas três listas são anotadas duas líquidas, a lateral [l] e o *tepe* [r], representados por *l* e *r*, respectivamente. A proposta assumida aqui é de que o fonema seja justamente o *tepe*. Já as soantes [j] e [w] estão representadas pelas vogais *i* e *u*, como em 4. *Cequaquai* e 32. *Xampeu*; já a palatal também é representada por *y*, quando em *onset* em 1. *Angrayocha*, 6. *Yamaroi* e 42. *Toyoto*. O /w/ também poderia estar ocorrendo em *clusters* e *codas* como em 13. *Ungaptuai*, 41. *Canampuaxi* e 35. *Taxiú*.

No quadro 1, seguinte, é apresentada a proposta dos fones das consoantes do registro de Vila Boa de 1782, os segmentos marcados com uma interrogação são aqueles cuja ocorrência não é possível comprovar, mas também não é possível descartar.

**Quadro 1: Lista de Vila Boa (GO) – Consoantes**

p	t		k
	s ~ z (?)	f	
m̄b ~ b			ŋg ?
m	n	ɲ	ŋ ?
w	r ~ l	j	

Para as vogais, é possível propor o quadro seguinte, em que interpreto o *â* como a representação para a vogal [ə] e as vogais nasais são representadas pela sequência *v+ç*.

**Quadro 2: Lista de Vila Boa (GO) – Vogais**

i, ĩ	ə	u, ũ
e, ê	a, â	o, õ

## Sistema fonológico: uma proposta

### Fricativas e africadas nas listas de Mossâmedes

Os dois quadros a seguir são reproduções da proposta de interpretação grafemática para as listas de Pohl e Saint-Hilaire em Vasconcelos (2009):

**Quadro 3: Pohl (1832)**

Consoantes				
p b	t	ts̄	tʃ̄	k
	z		f	
m	n			
w	r l		j	
Vogais				
i, ĩ				u, ũ
e, ê		a, â		o, õ

**Quadro 4: Saint-Hilaire (1848)**

Consoantes				
p b	t d	ts̄	tʃ̄	k
	s		f ʒ	
m	n		ɲ	
w	r l		j	
Vogais				
i, ĩ				u, ũ
e, ê		a, â		o, õ

A interpretação dada ao grafema *s*, da lista de Pohl, em itens como *unisi* ‘mãe’, *usúm* ‘pai’, *schuninsi* ‘galinha’, *schuninsischumá* ‘galo’, *sucomu* ‘serra’ e *kitesi* ‘ferro’, se tratava da sonora [z], seguindo assim o padrão da ortografia alemã em que *s* só tem valor de [s] em *coda silábica*.<sup>6</sup> Na lista de Vila Boa há somente uma ocorrência de fricativa sonora e em Saint-Hilaire há dois itens em que ele anota *j* (= [ʒ]). No diário de Pohl (1837, p. 33-34) consta uma lista de palavras dos Xavante da Aldeia do Carretão de Pedro III,<sup>7</sup> que fora anotada em 1819. Nesta lista também encontramos o *s*, porém, ao comparar os itens daquela lista com registros recentes da Língua Xavante (MCLEOD, 1974; PICKERING, 2010) e também da língua Xerente (SOUZA, 2008), observamos que corresponde a [tʃ̄]<sup>8</sup>

6 Dentre as listas Cayapó do Sul, três são anotadas por alemães: Pohl (1832), Kupfer (1870) e Nehring (1894), além da lista de Pohl (1832) o grafema *s* só ocorre em único item na lista de Kupfer (1870, p. 254): *sacoa* ‘boca’.

7 Em homenagem ao Rei de Portugal e Algarves que governou, ao lado de Maria I, entre 1777 e 1786.

8 Mantém-se o símbolo utilizado pela pesquisadora.

em McLeod (1974), a [s] em Pickering (2010) e ainda [s] no Xerente (SOUZA, 2008). As seguintes correspondências são encontradas com as demais listas Cayapó do Sul.<sup>9</sup>

‘galinha’	<i>schuninsi</i> (P)	<i>zurinzi</i> [tsurĩtsi] (K)	<i>chuninxi</i> [ʃũnĩʃi] (L)	<i>xinunxi</i> [ʃĩnũʃi] (B)
‘pai’	<i>usúm</i> (P)	<i>uxum</i> [uʃũ] (L)	<i>uçúm, vóçúm</i> [usũ] (B)	
‘mãe’	<i>unisi</i> (P)	<i>kuinzi</i> [kuĩtsi] (K) <sup>10</sup>		

A fricativa vozeada [ʒ] de Saint-Hilaire pode ser interpretada como variação com a contraparte desvozeada, contudo, as ocorrências das africadas [ts̄] e [tʃ̄] e das demais fricativas [s] e [ʃ] podem ser explicadas como casos de variação livre. Considerando os segmentos vocálicos que seguem as consoantes fricativas e africadas, nas três listas, temos em Vila Boa:

[s] i, e            [z] e            [ʃ] i, e, a, o, u

Já em Pohl, as realizações estão assim distribuídas:

[s] i, u            [ʃ] i, a, o, ó, u  
[ts̄] e, a, o        [tʃ̄] é, un<sup>11</sup>

Por fim, em Saint-Hilaire:

[s] i            [ʃ, ʒ] i, e, o, ó, u  
[ts̄] e, a, o        [tʃ̄] u

A proposta a ser levantada, segundo a qualidade do registro, é que há ao menos dois fonemas nesse conjunto de realizações. A questão é saber como se dá a oposição: pelo traço [± anterior] diferenciando [ts̄] e [tʃ̄]; ou pelo traço [± contínuo], opondo [ts̄] a [ʃ]. É preciso considerar que dos poucos itens comuns às listas de Pohl (1832) e Saint-Hilaire (1848) é regular a correspondência do traço [contínuo], ou seja, se em Pohl temos [tʃ̄] (*tsch*, não contínuo) em Saint-Hilaire teríamos [ts̄] (*ts*, não contínuo), a única exceção a esta regularidade é o termo galinha-pássaro:

‘arco’	<i>itsché</i>	<i>itse</i>
‘flecha’	<i>caschoné</i>	<i>cajone</i>
‘estrela’	<i>amschiti</i>	<i>amsiti</i>
‘burro’	<i>kitaschá</i>	<i>iquitachó</i> (‘cavalo’)
‘galinha’	<i>schuninsi</i>	<i>itschune</i> (‘pássaro’)

A hipótese de diferenciação pelo traço de continuidade tem como resultado uma série das obstruintes com quatro elementos /p, t, [ts̄], k/ e uma série das contínuas com apenas um representante: /ʃ/.

9 Aqui e em demais pontos: Vila Boa (VB), Pohl (P), Saint-Hilaire (SH), Kupfer (K), Lemos da Silva (L), Nehring (N), Barbosa (B).

10 As representações entre colchetes são hipotéticas.

11 Optei por manter a grafia do anotador.

## Obstruintes sonoras, nasais ou pré-nasalizadas?

O grafema nasal é o principal recurso utilizado para marcar nasalidade nas listas Cayapó do Sul: vogais nasais e consoantes nasais (em *coda* ou *onset*). Quando estes grafemas antecedem obstruintes é preciso considerar que tais ocorrências estejam nos dando indícios sobre a constituição destas consoantes. Para discussão do *status* destes grafemas antecedendo obstruintes, observemos os seguintes itens, retirados das listas de Pohl (1832) e Saint-Hilaire (1848):

(P) <i>impôti</i> , (SH) <i>impo</i>	‘veado’
(SH) <i>impute</i>	‘sol’
(SH) <i>impudé</i>	‘pescoço, nuca’
(P) <i>intá</i>	‘chuva’
(SH) <i>intó</i>	‘olho’
(SH) <i>intiera</i>	‘mulher’
(SH) <i>incó</i>	‘água’
(P) <i>incója</i>	‘aguardente’
(P) <i>pintue</i>	‘filho, filha’
(SH) <i>intompéiparé</i>	‘bonito, bonita’

Para estes itens, as seguintes interpretações são possíveis: (i) uma vogal nasal [ĩ], representada por *im/in*; (ii) uma vogal nasal [ĩ̃], porém, por variar entre *m/n*, representaria o espalhamento de nasalidade da vogal para obstruinte subsequente; (iii) a nasalidade é da consoante, neste caso, uma consoante nasal que estaria ganhando um contorno oral à direita, pela adjacência à vogal oral subsequente;<sup>12</sup> ou (iv) trata-se de consoantes subjacentemente pré-nasalizadas e a vogal precedente poderia ser simplesmente oral, ou ainda, erro de percepção dos anotadores.

Os demais registros do Cayapó do Sul pode nos ajudar a compreender melhor qual o status desta consoante, por exemplo, na lista de Lemos da Silva (1882) encontramos o item *pantó* ‘olho’, em que *pa-* é interpretado com um morfema que designa humano (ou ‘gente’) – comum a outros itens da lista como: *paquim* ‘cabelo’, *pacré* ‘orelhas’ – e o tema que corresponderia a ‘olho’ seria *-ntó*. Enquanto em Barbosa (1918), temos as seguintes ocorrências a partir do tema para ‘água’ *inkô*, *nkô*: *tinkó* ‘molhar’, *tinkônípín* ‘afogar’, *çankou* ~ *çankô* ‘saliva’ e *cinnankô* ‘diarreia’; e variações como: *inkô* ~ *nkô* ‘água’, *intó* ~ *ntó* ‘olho’ e *impó* ~ *mpó* ‘veado’. Tais ocorrências corroborariam uma interpretação que assume a hipótese em (iv).

Nos registros de Pohl (1832) e Saint-Hilaire (1848), as seqüências grafema nasal/grafema de obstruinte desvozeada não estariam restritas a início de palavras, como podemos observar nos dois últimos itens. Seguindo a interpretação em (iv), a hipótese é que a língua apresentaria uma série de obstruintes *p*, *t*, *ts*, *k* – (SH) *ípa* ‘braço’, (SH) *ité* ‘perna’, (P) *itsché* ‘arco’, (P) *cupá* ‘terra’ – e uma série de soantes pré-nasalizadas *m̃p*, *ñt*, *ñs*, *ŋ̃k* – (SH)

12 D’Angelis (1998, p. 20-21, p. 235-236) propõe que nas línguas Jê e Macro-Jê que apresentam esse processo, o contorno em questão é dessoantizado, o que significa que, além de oral (ou desnasalizado), perde também o vozeamento. Espectrogramas de alguns enunciados em Kaingang mostraram efetivamente isso. No entanto, diferenças de sincronização entre o vozeamento da vogal seguinte e o término da consoante em *onset* (que na literatura fonética costuma ser referido por VOT) produzem resultados distintos nas diferentes línguas: em algumas, percepção da dessoantização parcial é de *mp*, *nt*, etc.; e em outras, *mb*, *nd*, etc.

*impo* ‘veado’, (P) *intá* ‘chuva’, (SH) *amsiti* ‘estrela’,<sup>13</sup> *incó* ‘água’. Nesta hipótese, qual o status das nasais plenas? Ou são vistas como uma (terceira) série consonantal própria ou, alternativamente, devem se considerar que as nasais plenas são somente realizações fonéticas das pré-nasalizadas quando adjacentes a vogais nasais? Os itens a seguir ilustram a ocorrência dos grafemas nasais indicando consoante nasal em *onset* silábico:

Vila Boa	Pohl		Saint-Hilaire	
<i>Yamaroi</i>	<i>muschiú</i>	‘milho’	<i>macácá</i>	‘branco’
<i>Romexi</i>	<i>inromú</i>	‘mata’	<i>intomarca</i> <sup>14</sup>	‘feio’
<i>Combono</i>	<i>penatá</i>	‘farinha de milho’	<i>panariá</i>	‘um índio’
<i>Tonito</i>	<i>caschoné</i>	‘flecha’	<i>cajone</i>	‘flecha’
<i>Chaponhece</i>	<i>tapanió</i>	‘negro’	<i>tapanho</i>	‘um negro’

Como pode se observar nas três listas, aparentemente, as nasais plenas estariam ocorrendo diante de vogais orais, se é este o caso, excluimos das listas de hipóteses aquela que as pré-nasalizadas sejam resultado de processo de oralização causado pela vogal núcleo da sílaba, neste caso, é a vogal à esquerda que desencadearia o processo. Lembro que ocorrência como *pantó*, como exposto, na lista de Lemos da Silva (1882), seria um argumento contra esta hipótese. Interpretação alternativa é que a nasalidade da vogal não tenha sido registrada e justamente a ocorrência da consoante nasal seria indício da nasalidade da vogal, mantendo a hipótese em (iv).

D’Angelis (2002, p. 86), sob a perspectiva da geometria de traços, propõe que “quando Nasal for um traço fonológico (privativo) em uma língua determinada, ele estará alocado sob um nó articulador *Soft Palate* (SP), quer em vogais, quer em consoantes (i.e., onde quer que a oposição *nasal x oral* seja fonológica)” e esclarece que “a presença do traço fonético [nasal] não implica necessariamente a presença do traço fonológico correspondente”. Segundo o autor, línguas que fazem a oposição soante x obstruente realizam “soanticidade pelo abaixamento do véu palatino. Nesse caso, a nasalidade é apenas recurso para implementação do traço *Spontaneous* (ou *Sonorant*) *Voicing* (SV), ou seja, uma condição de implementação fonética” (D’ANGELIS, 2002, p. 86). Para a variedade do Kaingang do oeste catarinense (língua Jê meridional), D’Angelis (2002, p. 87) assume que “há uma oposição *nasal-oral* nas vogais, mas oposição *soante-obstruente* nas consoantes” e sugere “que o processo significativo que opera na língua é o *espalhamento* do nó SP aos segmentos adjacentes que não o portam subjacentemente”. Assim, segundo D’Angelis (1998), as consoantes nasais são soantes especificadas subjacentemente pelo traço SV, mas não para (SP); já as vogais são especificadas para SP (com especificação do traço nasal para as vogais nasais) e também para SV. As soantes nasais realizam-se plenamente diante de vogais nasais, e são pré-nasalizadas ou pós-oralizadas diante de vogais orais, pelo espalhamento do nó SP: /m̃/ [‘m̃] ‘jabuticaba’ x /mɔ/ [‘mbo] ‘espiga’; /mũ/ ‘mandi (peixe)’ x /muŋ/ [mbugŋ] ‘imbu’. Em Kaingang, o *output* é uma pós-oralizada vozeada, resultado da aplicação do Princípio de Contorno Obrigatório (OCP) que motivaria a fusão dos nós SV adjacentes.<sup>15</sup>

13 Davis (1966, p. 20) ‘estrela’ \*kan<sup>ye</sup>, no mesmo artigo Suyá, k<sup>h</sup>ane(ti).

14 Nas demais listas Cayapó do Sul: *tamancáre* (K), *tamancare* (L), *tómanká* (B). Este item é o único, não só na lista de Saint-Hilaire como também nas demais listas Cayapó do Sul, em que há a ocorrência de *r* em coda não final. Não se descarta, neste caso, algum problema na transposição do original manuscrito para a versão publicada, em que talvez a forma anotada por Saint-Hilaire se aproxime daquela encontrada em Barbosa.

15 Ver nota 12.

Em Cayapó do Sul teríamos processo semelhante, porém, as consoantes subjacentes seriam as pré-nasalizadas, segundo a hipótese em (iv). O que permitiria a realização da nasal plena seria justamente o espalhamento do nó SP das vogais nasais (especificadas para SV e para SP nasal) para as soantes nasais adjacentes (especificadas para SV e não especificadas para SP). Já o espalhamento do nó SP das vogais orais (especificadas para SV e SP) para as soantes adjacentes mantém a fase dessoantizada da pré-nasalizada.<sup>16</sup>

Por fim, a ocorrência de <mb> e <ng>, registrada na lista de Vila Boa, e as possíveis obstruintes vozeadas comuns a três listas estariam, em parte, relacionadas às pré-nasalizadas, ainda segundo D'Angelis (1998, p. 20-21), como um fenômeno de gradiência, resultado da passagem da fase soante para a fase dessoantizada.

Os itens a seguir ilustram a aplicação desta hipótese ao Cayapó do Sul:

/mpo/	<i>impo</i> (SH)	‘veado’	/mp̃u'ju/	<i>muschiú</i> (P)	‘milho’
/nta/	<i>intá</i> (P)	‘chuva’	/jamp̃ãroj/	<i>Yamaroi</i> (VB)	
/nto/	<i>intó</i> (SH)	‘olho’	/pentã'ta/	<i>penatá</i> (P)	‘farinha de milho’
/ŋko/	<i>incó</i> (SH)	‘água’	/tapa'nsõ/	<i>tapanió</i> (P), <i>tapanho</i> (SH)	‘negro’

Uma série de consoantes pré-nasalizada para o Cayapó do Sul não seria uma novidade e nem uma excentricidade às línguas Jê: Camâra Jr. (1959) propõe, para o Proto-Jê, radicais reconstruídos pré-nasalizados e Alves (2007) argumenta a favor de uma série de pré-nasalizadas para o Apãniekrá (língua Jê Setentrional).

### As soantes contínuas

As consoantes pré-nasalizadas formam a série das soantes descontínuas do Cayapó do Sul, enquanto /w, r, j/ formariam a série das contínuas. Nas listas de Vila Boa (1782), Pohl (1832) e Saint-Hilaire (1848) a soante /r/ ora é representada por *r*, ora por *l*. Assumo que a realização subjacente seja do *tepe*, porém, não se descarta que a alternância encontrada nos registros possam dar indícios de uma realização *flap* ou retroflexa (cf. sessão seguinte). Pode-se atestar nas listas o *tepe* em *onset* silábico – *Romexi* (VB), *robú* (P) ‘cão’, *poré* (SH) ‘pedaço de pau’ – e formando *clusters* com obstruinte – *Xecrá* (VB), *lenkré* (P) ‘dedo’, *icreti* (SH) ‘anta’. Já para /w, j/, ora são representados: (i) por *i* e *u*, *piápa* (P) ‘cesto pequeno’, *chuí* ‘dente’; (ii) por *y*, *Yamaroi* (VB), *Angrayocha* (VB); (iii) ou ainda em Pohl, por *j*: *pujanca* ‘deus’.<sup>17</sup>

### Clusters

Os *clusters*, nas línguas Jê, seguem restrições que, em Kaingang (D'AANGELIS, 1998), não permite dois segmentos com o mesmo articulador e, em Panará (VASCONCELOS, prelo), em que a oposição grave vs agudo é mais preponderante, não permite dois segmentos graves ou dois agudos. Nos registros de Mossâmedes é possível fazer o seguinte levantamento de possíveis *clusters*: pw (*Pupuaré* VB), pj (*piápa* ‘cesto pequeno’ P), pr (*itpé-pri* ‘menino’ P),

16 Em D'Angelis (1998, p. 235): “Segue-se que, numa consoante soante parcialmente desnasalizada, deveríamos esperar o surgimento de um contorno surdo, como por exemplo: [mp], [nt], [ŋk], etc.”

17 O termo para ‘deus’ em Saint-Hilaire (1848) é *puhancá*, sem nenhuma explicação, a versão brasileira da Livraria Itatiaia, de 1975, está *punhança*. A posição do *h* neste item é discrepante da opção de Saint-Hilaire em usar a ortografia de língua portuguesa como a base do registro, esperar-se-ia o uso de *-rr-*. Em Lemos da Silva (1882) o termo para ‘deus’ é *puancá*.

tw (*Potuaré* VB), tj (*itpentié* ‘moça’ P), jw (*chuí* ‘dente’ SH), fj (*Uncrixiu* VB), kw (*uncuá* ‘casa’ P), kj (*piankákianká* ‘papel’ P), kr (*chiccré* ‘orelha’ SH) e rj (*panariá* SH). Se o Cayapó do Sul mantém as mesmas restrições de línguas como o Kaingang e o Panará, o número de ocorrências seria mais reduzido do que o apresentado; no entanto, tal redução só é possível se também for possível identificar os traços que estão em jogo nessa posição silábica.

Se os traços envolvidos estão relacionados a articulador, pw e kw não podem co-ocorrer; sua realização estaria diretamente relacionada a como o /w/ é interpretado na língua: se tem caráter fonologicamente labial, o *cluster* possível é kw; se /w/ é fonologicamente velar, então o *cluster* possível é pw. Se a distinção for, antes, entre grave x agudo, nem pw nem kw formam sequências permitidas nessa posição, bem como, tj e fj não figuram entre os *clusters* licenciados. A sílaba final em itens como *Uncrixiu* e *Taxiú* seria uma sílaba CV, com *i* ou *u* como núcleo vocálico, ou uma sílaba CVC na qual *u* seria uma *coda* [w]; ou ainda, uma sequência de duas sílabas, CV.V: *xi.u*. Já para *-tié*, em Pohl, *-tie*, em Saint-Hilaire, uma possível interpretação é que seja *ti.é*, o que levaria a interpretar *intiera* como *i.nti.e.ra* ‘mulher’ (SH). Em Pohl há ainda sequências como *itschiú* ‘fogo’, que em Vasconcelos (2009) é interpretado como [tʃu], tal interpretação é estendida para o termo *muschiú* ‘milho’ [ʃu] e estaria relacionado à *tetaschú* ‘feijão’ e a *antoaáschú*<sup>18</sup> ‘munição’, indicando um morfema comum associado a grão.<sup>19</sup>

Dos itens em *clusters*, a sequência *rj* sai do padrão obstruinte + soante: (a) *panariá*, (b) *chiccria* ‘mão’, (c) *icria* ‘coxa’, (d) *icrian* ‘cabeça’ e ainda (e) *Apluace* (em que temos algo como /prwa/). Considerando a correspondências na lista de Barbosa (1918): (a) *panará*, (b) *cykiá*, (c) *inkré* e (d) *kián*; já em Panará (DOURADO 1990, 2001) (a) /panã’ra/, (b) *tsikia* ~ *sikia*, (c) /ĩ’krə/ e (d) /ĩ’kjã ~ /iki.ã. Em Apãniekrá (ALVES, 2004) (b) -ũkra, (c) *i-ke* e (d) *krã*. Segundo D’Angelis (comunicação pessoal), uma divisão silábica para (a) *panariá* seria *pa.na.ri. ’a*, ou mesmo, *pa.na.ri. ’?a*, enquanto para (b) *chiccria*, (c) *icria* e (d) *icrian*, o *a* final poderia ser um morfema, apresentando o mesma divisão silábica de (a), por exemplo, *i.cri.a* ‘mão’. Ou ainda, o *i* presente nestes itens é “uma percepção do anotador, para certa dessincronização de gestos na passagem do *cluster* para a vogal, talvez porque esse *r* não seja um simples *tepe*, talvez um *flap* ou um retroflexo. Efeito semelhante ocorreria em *Apluace* (aqui pela labialização da primeira consoante)”.

Sequências tw são licenciadas, porém, devido à precariedade dos dados, esse *cluster* não se sustentaria, pois, em Saint-Hilaire, *pintué* pode ser resultado de processo de juntura de morfemas envolvendo *tó* ‘dançar’ (BARBOSA, 1918), e na lista de Vila Boa, *Ungaptuai* e *Potuaré* podem ser interpretados como *Un.gap.tu.ai* e *Po.tu.a.ré*.

Ainda na lista de Vila Boa também há uma ocorrência *mre* (*Poimre*), sequência que, tal como *rj*, fugiria ao padrão obstruinte-soante. Uma análise possível é que o grafema *m* esteja indicando somente a nasalidade da vogal que o precede e possível espalhamento para a soante tautossilábica: [põj.re]; mas também, pode ser indicio de que a restrição no *cluster* se daria por contínuo vs não contínuo e não por obstruinte vs soante; ou ainda, que

18 Em Pohl, ‘espingarda’ *atoná*, que em Barbosa é *atôme* com a seguinte nota: “pronuncia-se a tó me”. Interpreta-se aqui o item para ‘munição’ como [ãtwa:’ʃu], considerando a correspondência *o:w* do registro do Xavante do séc. XIX com registros atuais.

19 Já para ‘arroz’ o termo é *tunishin*, para o qual é possível propor um possível equívoco na transposição do manuscrito para a forma impressa, e a forma anotada por Pohl seria *tunischíu*.

os segmentos nasais possam ser interpretados como obstruintes, neste caso, são segmentos especificados para SP [nasal] e não para SV, formando assim, um série de obstruintes nasais. É importante considerar que este é o único caso entre os registros disponíveis do Cayapó do Sul.

## Vogais

Em Vasconcelos (2009), foi proposta, para as listas de Pohl e Saint-Hilaire, cinco vogais orais: i, e, a, o, u; e cinco nasais: ã, ê, ã, õ, ù; enquanto para lista de Vila Boa é possível identificar seis vogais orais (i, e, ə, a, o, u) e cinco nasais (ĩ, ê, ã, õ, ù), pois em itens como 17. *Pitubá*, 26. *Xecrá* e 27. *Enconá* há indícios da realização de [ə], no entanto, a sua ocorrência restrita a final de palavra gera suspeita quanto ao seu status de fonema, pois pode tratar-se de uma variação ou mesmo equívoco do anotador.

É possível levantar a hipótese de uma vogal alta posterior não arredondada [ɨ]. Tal hipótese é resultado da comparação dos itens a seguir encontrados nos registros de Pohl e Saint-Hilaire com a lista de Barbosa (1918) e com registros recentes das línguas Panará (DOURADO 1990, 2001) e Apãniekrá (ALVES, 2004, 2007):

<i>itschiú</i> (P)	<i>icy</i> (B) <sup>20</sup>	[ɨ'tsi] ~ [ɨ'tʃi] (Pa)	ku'xi, ku'hi (Ap)	'fogo'
<i>muschiú</i> (P)	<i>môcê, môcy</i> (B)	/mõ'si/ (Pa)	põ-hi (Ap)	'milho'
<i>cupá</i> (P), <i>cúpa</i> (SH)	<i>kýpa</i>	/'kyipa/, 'kipa (Pa)		'terra'
<i>itú</i> (P)	<i>iútú</i> (B)	/y'i'ti/ (Pa)	tik (Ap)	'morrer'
<i>kiúti</i> (P)	<i>kir, ikih</i> (B)	kyi (Pa)	h-akri (Ap)	'(estar) frio'
<i>itpúti</i> (P), <i>imputé</i> (SH)	<i>iútât</i> <sup>21</sup>	piti, ãpiti (Pa)	pit (Ap)	'sol'

Em todos estes casos, a vogal [ɨ] estaria representada por *u/ú*, em que permaneceria tanto a altura como a posição da vogal. Contudo, nem todo item destes registros grafado com *u/ú* representam [ɨ], como podemos observar nos itens a seguir:

<i>putkuá</i> (P)	<i>pukuá</i> (B)	/puku'a/ (Pa)	'céu'
<i>cubu papa</i> (P) 'caçar'	<i>kbú, kubú</i> (B)		'passear'
<i>pujanka</i> (P), <i>puhancá</i> (SH), <i>puancá</i>			'deus'

Tal vogal também fica como ocorrência possível, mas não confirmada se nos restringimos ao registro de Mossâmedes, pois, é possível, que diferente do esperado para línguas da família Jê, a vogal [ɨ] corresponda a [u], reduzindo o quadro de protofonemas vocálicos proposto por Davis (1966). No quadro 5 é apresentada uma síntese da hipótese para as vogais de Mossâmedes a partir das listas de Vila Boa, Pohl e Saint-Hilaire.

**Quadro 5: Hipótese para as vogais de São José de Mossâmedes**

i, ã	ɨ?	u, ù
e, ê	ə?	o, õ
	a, ã	

20 Em Barbosa (1918, p. 40): "o y soa quase como o u francez"; o que interpreto como tentativa de representar [ɨ], som incomum para o a língua materna do anotador, o Português.

21 Em todas as listas Cayapó do Sul é encontrado um termos (ou mais de um) para 'sol': *itputi* (P), *imputé* (SH), *hiutote* (K), *puti* (L), *impütě* (N), *iútât, iútôt, iúktôt, iútáici* (B).

## Considerações finais

A proposta apresentada neste estudo ainda deixa pontos em aberto sobre o que tenha sido o sistema fonológico do Cayapó do Sul; não trato dos segmentos licenciados em *coda*, tema que ainda merece mais investigação. No entanto, a partir desta discussão, aponto que a oposição básica se dá entre obstruintes /p t ts k ʃ/ e soantes /mp̄ nt̄ ns̄ ŋk̄ w r̄ j̄/, porém, permanece em suspenso qual a relevância do traço [contínuo] para as obstruintes. Ocorrência como *Poimre* (VB) e a falta de definição sobre a *coda* não permitem descartar a hipótese de que as nasais sejam mais uma série de consoantes, possivelmente, de obstruintes nasais, indicando que oposição básica entre as consoantes seja, tal como nas vogais, entre oral vs nasal.

Os registros de Mossâmedes dão indícios de que a língua falada pelos Cayapó do Sul é uma variedade diferente daquela falada nas aldeias próximas a Sant'Anna do Paranaíba, às margens do rio Paraná, ponto mais meridional do seu território. Duas características indicariam essa distinção: (i) impossibilidade de reduzir os segmentos fricativos e africados a um único fonema, tal como ocorre nas demais listas, ou seja, aqui teríamos ts̄ e ʃ, enquanto nos demais há ou só s ou ʃ; e (ii) os itens para 'cabeça' *icrian* e 'mão' *chicria*, que se aproxima mais dos termos encontrados em Apãniekrá (ALVES, 2004), *krã* e *-ũkrã*, do que das demais listas Cayapó do Sul, *kián* e *cykiá* (BARBOSA, 1918).

## Listas de Palavras: fontes

(i) Registro de Batismo de Vila Boa. *Relaçãõ dos Indios da Naçaõ Cayapó q'se Baptizaraõ na Igreja de Villaboa de Goyáz no dia de hoje 12 de Outubro de 1782*. In: OFÍCIO do [governador e Capitão-Geral de Goiás] Luís da Cunha Meneses, ao [secretário de estado da Marinha e Ultramar], Martinho de Melo e Castro, sobre a Civilização dos índios Caiapó; o sucesso do aldeamento da dita nação; a criação da Aldeia Maria I, e enviando plantas e estampas da dita aldeia. 18 dez. 1782. Manuscrito. CD-ROM. *Projeto Resgate de Documentação Histórica "Barão de Rio Branco"*. Disponível no Arquivo Edgar Leuenroth, Universidade Estadual de Campinas (AEL/Unicamp).

(ii) Lista de palavras anotada por Emmanuel Pohl. *Sprachproben der Cayapós in der Aldeya S. José de Mossamedes*. In: POHL, J. E. *Reise im Innern von Brasilien: Auf allerhöchsten befehl seiner majestat des kaisers von osterreich, franz des ersten*. Wien: A Strauss's Sel Witwe & J B Wallishausser, vol.1, 1832. p. 447-448.

(iii) Lista de palavras anotada por Auguste de Saint-Hilaire. SAINT-HILAIRE, A. de. *Voyage aux sources du Rio de S. Francisco et dans la province de Goyaz II*. Paris: a. Bertrand, 1848. p. 108-109.

## REFERÊNCIAS

ALVES, F. C. *O Timbira falado pelos Canela Apãniekrá: uma contribuição aos estudos da morfossintaxe de uma língua Jê*. 2004. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.

\_\_\_\_\_. Sistema fonológico do Timbira Apãniekrá (Fonemas, sílaba e acento). In: RODRIGUES, A. D.; Cabral, A. S. A. C. (Org.). *Línguas e Culturas Macro-Jê*. Brasília: Editora UnB, 2007. p. 45-55.

ARAÚJO, B. A. C. *Análise do Wörterbuch der Botokudensprache*. 1992. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas.

ATAÍDES, J. M. *Sob o signo da violência: colonizadores e Kayapó do Sul no Brasil Central*. Goiânia: Editora UCG, 1998.

BARBOSA, A. S. *Cayapó e panará*. Manuscrito disponível no Arquivo do IHGB, 1918.

CÂMARA JR., J. M. Alguns Radicais Jê. *Publicações Avulsas do Museu Nacional*, 28, 1959.

D'ANGELIS, W. R. *Traços de modo e modos de traçar geometrias: línguas Macro-Jê & teoria fonológica*. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1998. 2 vols.

\_\_\_\_\_. Nasalidade e soanticidade em línguas Jê: o Kaingang paulista e o Mëbengokre. *Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da ANPOLL*, Belém, Pará, Brasil, p. 86-95, 2002.

DAVIS, I. Comparative Jê phonology. *Estudos Linguísticos*, v. 1, n. 2, p. 10-24, 1966.

DOURADO, L. *Estudo preliminar da fonêmica Panará*. 1990. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade de Brasília, Brasília.

\_\_\_\_\_. *Aspectos Morfosintáticos da Língua Panará (Jê)*. 2001. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.

GRANNIER RODRIGUES, D. M. *Fonologia do Guaraní Antigo*. Campinas: Editora da Unicamp, 1990. 86 p.

HALL, J.; MCLEOD R. A.; MITCHELL V. *Pequeno dicionário Xavante-Português*. 1987. Cuiabá: SIL. Disponível em: <<http://www.sil.org/americas/brasil/publens/dictgram/XVDict.pdf>>. Acesso em: 02 mar. 2012.

JAKOBSON, R.; FANT, C. G. M.; HALLE, M. *Preliminaries to Speech Analysis. The distinctive features and their correlates*. Cambridge, MA: The MIT Press, 1952.

KUPFER. Die Cayapo-Indianer in der Provinz Matto-Grosso. *Zeitschrift für der Gesellschaft für Erdkunde zu Berlin*, n. 5, p. 244-254, 1870.

LEMOS DA SILVA, J. *Os índios Cayapós*. Manuscrito disponível no Arquivo do IHGB, 1882.

MCLEOD, R. Fonemas Xavante. *Série Linguística*, n. 3, p. 131-152, 1974. Disponível em: <<http://www.sil.org/americas/brasil/SILapub.html>>. Acesso em: 02 mar. 2012.

- NEHRING, C. Sud-Cayapo: Wörterlisten. In: EHRENREICH, P. Materialien Zur Sprachekunde Brasiliens. *Zeitschrift für Ethnologie*, n. 26, p. 136-137, 1894.
- PICKERING, W. A. *A Fonologia Xavante: uma revisão*. 2010 Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- POHL, J. E. *Reise im Innern von Brasilien: Auf allerhöchsten befehl seiner majestat des kaisers von osterreich, franz des ersten*. Wien: A Strauss's Sel Witwe & J B Wallishausser, 1832-1837. 2v.
- POHL, J. E. *Viagem no Interior do Brasil: empreendida nos anos de 1817 a 1821 e publicada por ordem de Sua Majestade o Imperador da Austria Francisco Primeiro*. Tradução de Teodoro Cabral. São Paulo: Instituto Nacional do Livro, 1951, Parte I.
- SAINT-HILAIRE, A. de. *Voyage aux sources du Rio de S. Francisco et dans la province de Goyaz II*. Paris: Arthus Bertrand, 1848.
- \_\_\_\_\_. *Viagem à Província de Goiás*. Tradução de Regina R. Junqueira. Belo Horizonte: Itatiaia; São Paulo: Edusp, 1975.
- SOUZA, S. L. de. *Descrição fonético-fonológica da língua Akwe-Xerente*. 2008 Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade de Brasília, Brasília.
- TRUBETZKOY, N. *Princípios de fonologia*. Tradução de Wilmar R. D'Angelis. No prelo. Título original: Grundzüge der Phonologie, 1939.
- VASCONCELOS, E. A. Sobre as listas de palavras Cayapó do Sul de São José de Mossâmedes. *Revista Sínteses*, v. 14, p. 405-423, 2009.
- \_\_\_\_\_. Repensando la Fonologia de la Lengua Panará (Jê). In: *II Jornadas de Fonética e Fonologia*, Córdoba: UNC (prelo).

# Descrição e análise preliminar sobre a categoria verbal no dialeto indígena Pykobjê-Gavião (Timbira)

(Preliminary description and analysis on the verbal category in the indigenous dialect Pykobjê-Gavião (Timbira))

Talita Rodrigues da Silva<sup>1</sup>

<sup>1</sup>Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (FFLCH/USP)

talita.rodrigues.silva@usp.br

**Abstract:** In this paper we aim to describe and analyze in a very preliminary way some common verbal structures and uses in basic sentences of the indigenous dialect Pykobjê-Gavião (Macro-Jê Branch, Jê Family, Timbira Group). To develop this research, we take mainly the following references: Simon C. Dik (1997), Thomas E. Payne (1997) and Paul Schachter (2007). Following Payne (1997, p. 47), we will observe the distributional and structural properties of verbs, which can occur as head of a verbal phrase (VP) and as head of an adjective-verbal predicate. Then, we will discuss some types of verbs: active intransitive verb, state intransitive verb, process intransitive verb, extended intransitive verb, simple transitive verb and extended transitive verb. Finally, we will observe and analyze some tense and aspectual verbal operators.

**Keywords:** Brazilian indigenous language; Timbira dialect; morphosyntax; verbal phrase.

**Resumo:** Neste artigo buscamos descrever e analisar de forma bastante preliminar algumas estruturas e usos verbais comuns em sentenças básicas do dialeto indígena Pykobjê-Gavião (Tronco Macro-Jê, Família Jê, Complexo Timbira). Para desenvolver essa pesquisa, tomamos principalmente as seguintes referências: Simon C. Dik (1997), Thomas E. Payne (1997) e Paul Schachter (2007). De acordo com Payne (1997, p. 47), observaremos as propriedades distribucionais e estruturais de verbos, que podem se apresentar tanto como núcleo de um sintagma verbal (SV) quanto como núcleo de um predicado verbo-adjetival. Então, discutiremos sobre alguns tipos de verbos: verbo intransitivo ativo, verbo intransitivo estado, verbo intransitivo processo, verbo intransitivo estendido, verbo transitivo simples e verbo transitivo estendido. Finalmente, veremos e analisaremos alguns operadores verbais de tempo e aspecto.

**Palavras-chave:** língua indígena brasileira; dialeto Timbira; morfossintaxe; sintagma verbal.

## Introdução

Os Pykobjê-Gavião são, segundo Rodrigues (2002 [1986]), um dos sete povos membros do grande grupo étnico Timbira.<sup>1</sup> Dentre as sete variantes do dialeto Timbira, quatro dispõem de trabalhos de descrição e análise linguística com orientação de métodos científicos. São elas: Krahô (cf. SOUZA, 1997), Parkatejê-Gavião (cf. FERREIRA, 2003), Apãniekrá-Canela (ALVES, 2004) e o Pykobjê-Gavião, que já fora estudado com ênfase nos aspectos

---

<sup>1</sup> A denominação Timbira, segundo Nimuendajú (1946), poderia significar “os amarrados” (do tupi: *tin* (amarrar) + *pi'ra* (passivo)), em referência às fitas de palha ou faixas trançadas em algodão que esses povos usam em partes do corpo como braços, testa e tornozelos. Timbira é o nome que os não índios atribuíram a esses povos. Entre eles, contudo, o tratamento é por *mehê*, que significa “minha gente”, “minha carne”. Os sete povos Timbira são: Ramkokramekrá-Canela, Apãniekrá-Canela, Krenjê, Krahô, Krinkati-Gavião, Parkatejê-Gavião e Pykobjê-Gavião. Se houver interesse em conhecer melhor cada um dos povos, sugerimos a leitura de Melatti (1972; 1978).

fonéticos e fonológicos, por Sá (1999), morfofonológicos, por Amado (2004), e, morfosintáticos, por Silva (2012).<sup>2</sup>

Historicamente, o Pykobjê-Gavião é um povo que, segundo Nimuendajú (1946), manteve-se voluntariamente afastado dos não indígenas até meados do século XIX, devido ao seu *ethos* belicoso. Atualmente, vive a sudoeste do estado do Maranhão, a cerca de 18 km do centro da cidade de Amarante, na microrregião de Imperatriz.

Os Pykobjê-Gavião ocupam uma mesma terra indígena (com extensão média de 42 mil hectares) que fora delimitada em 1977 pelo convênio Funai/Radam, após intensas disputas com fazendeiros locais. Encontram-se divididos em quatro aldeias: Governador, a mais antiga delas, Rubiácea, Riachinho e Aldeia Nova, esta última organizada em meados de 2010. Ao todo, contabilizam cerca de 600 indivíduos, com crescimento populacional médio de 10% ao ano, segundo dados recentes do Centro de Trabalho Indigenista (CTI).<sup>3</sup>

Apesar de haver disputas internas e eventuais desentendimentos, as aldeias partilham da mesma unidade cultural e há intenso convívio entre seus membros, sobretudo em decorrência das festividades tradicionais e laços de parentesco.

Os Pykobjê-Gavião também são conhecidos na literatura como “Gavião do Maranhão” ou “Gavião do Leste” e denominam-se como *Pycop catëë jë*, que significa povo de *Pycop*. *Pycop* é uma divindade tradicional da cosmologia Timbira.

Ainda que não se trate de um povo com contato recente ou cujas relações iniciais tenham sido pacíficas com os não indígenas, chamados por eles de *cöpë*, a língua indígena tem se mantido como fator de afirmação cultural. Ou seja, o Pykobjê-Gavião continua a atuar como língua de comunicação interpessoal dentro das aldeias e é a língua materna de 100% desse povo.

Em tal situação, o Português do Brasil, doravante PB, costuma ser ensinado apenas em idade escolar, isto é, por volta dos sete anos. Contudo, sabemos haver letramento anterior por meio de veículos midiáticos disponíveis em algumas casas das aldeias. Dentre os principais difusores destacamos rádio e televisor.

A alfabetização tende a ocorrer nas escolas indígenas das aldeias, com atuação majoritária de professores não indígenas e missionários americanos. Alguns poucos indígenas também passaram a contribuir com este ensino mais formal, após receberem a oportunidade de cursarem módulos de licenciatura intercultural oferecidos por instituições públicas de ensino superior deste país, tais como a Universidade Federal de Goiás (UFG).<sup>4</sup> Contudo, a dificuldade de usar a escrita ainda é um tema de discussão comum entre esse povo, principalmente nas faixas etárias mais elevadas.<sup>5</sup>

2 Existe entre os estudiosos a controvérsia quanto ao fato de as sete línguas membros do Complexo Timbira serem línguas aparentadas ou se tratarem, de fato, de dialetos de uma única língua Timbira, que se afastou diacronicamente. Uma evidência a favor dos que apostam na ideia de que os Timbira formaram um grupo único no passado, com apenas uma língua em uso, o Proto-Timbira, é o fato de que todos os indivíduos que falam a língua indígena de seus "avós" (aceitando as variações diacrônicas existentes) conseguem se comunicar entre si através de linguagem.

3 Centro de Trabalho Indigenista (CTI), trata-se de uma organização não governamental atuante há mais de duas décadas entre os povos indígenas tratados.

4 Para entender melhor como funciona a educação intercultural entre povos indígenas brasileiros, sugerimos a leitura de Silva e Rocha (2006).

5 Essa temática deve ser relativizada levando-se em consideração a influência da Tradição oral no aprendizado e transmissão/recriação de conhecimentos por parte de povos indígenas brasileiros.

De todo modo, muitos Pykobjê-Gavião falam o PB com relativa proficiência, basicamente devido à necessidade de contato com representantes de órgãos de proteção aos indígenas, vendedores e demais pessoas, que movimentam serviços nas cidades próximas. Alguns poucos apresentam boa capacidade para se comunicarem em PB por meio da modalidade escrita, bem como o fazem em Timbira, e estes tendem a ficarem responsáveis por reivindicações junto aos órgãos de assistência.

Com relação ao letramento voltado às demandas desse povo indígena (cf. MATTOS E SILVA, 2004), apesar de a escrita uniformizada Timbira ter sido aprovada em dezembro de 2003, e de, na ocasião, contar com a anuência de representantes de seis dos sete povos Timbira, o que incluía membros do grupo Pykobjê-Gavião, a grafia adotada ainda é a missionária (Missão Novas Tribos do Brasil).<sup>6</sup>

Embora os educadores, especificamente os indígenas, reconheçam que a grafia missionária apresenta “problemas”, uma vez que não estabelece uma relação constante entre som e grafema, ela é a vigente por duas razões fundamentais: (i) faltam materiais didáticos que ensinem e apliquem a grafia uniformizada; (ii) os replicadores da escrita indígena (indígenas ou não indígenas), praticamente, desconhecem a existência da mesma. Tal situação evidencia que, embora a grafia Timbira tenha sido elaborada com embasamento teórico (assessoria das linguistas Flávia de Castro Alves e Rosane de Sá Amado) e sob uma égide democrática, houve falhas na transmissão desse conhecimento às partes interessadas.

Apesar das dificuldades expostas acima, optamos por apresentar os dados coletados em viagem de campo realizada em fevereiro e março de 2010 por Silva (2012) com a grafia uniformizada, por acreditarmos ser este um passo importante para seu estabelecimento no meio indigenista.

Tendo a ambição de atuar como uma exposição preliminar, que entende sua limitação quanto ao conteúdo verbal semântico da língua em estudo, o presente artigo buscará apresentar a distribuição e as estruturas que marcam os sintagmas verbais do Pykobjê-Gavião. Inicialmente, apresentaremos os tipos de sintagmas verbais, de acordo com a quantidade de argumentos e as funções sintáticas do sujeito que cada tipo verbal requer. Em seguida, apresentaremos alguns operadores aspecto-temporais.

### **Características morfosintáticas e morfossemânticas dos verbos em Pykobjê-Gavião: uma análise preliminar**

Segundo Dik (1997, p. 49-52), a estrutura subjacente da frase (*clause/speech act*) requer, em primeiro lugar, a existência de um predicado (*predicate*), que remeta a propriedade/relação (*property/relation*).<sup>7</sup> A esse predicado será aplicado um número adequado de termos (*terms*), que funcionarão, por sua vez, como argumentos (*arguments*) do predicado. Distintos dos argumentos são os satélites, uma vez que estes não são exigidos pelo predicado. Ou seja, enquanto os argumentos são obrigatórios para a interpretação do predicado, os satélites trazem informações adicionais. Dik (1997, p. 87) diz, ainda, que o mais natural é que o satélite ocupe posição periférica na frase e não central, como costuma ocorrer com os argumentos.

6 A grafia uniformizada Timbira já estava em discussão desde 1994, segundo Amado (2004, p. 6-7).

7 A organização de níveis de uma frase, segundo Dik (1997, p. 50), se organiza da seguinte forma: Nível 1: predicado(r) e termos (*core predication*); Nível 2: predicção (*extended predication*); Nível 3: proposição (*proposition*); Nível 4: frase ou cláusula (*clause*).

Contudo, o autor reconhece que, em muitos predicados, a noção de argumento se confunde com a de satélite, por isso as análises não se dão aprioristicamente.

Em Pykobjê-Gavião, distinguimos dois tipos de predicados: predicados nominais (N) e predicados verbo-adjetivais (VA). Os predicados VA, que são mais profícuos que os predicados N, podem ser analisados de acordo com o número (predicado de um lugar, de dois lugares ou três lugares) e a natureza dos termos que exige (funções semânticas do primeiro argumento (A1) e funções semânticas dos demais argumentos (A2 e/ou A3)).<sup>8, 9</sup> A predicção como um todo designa o denominado *State of Affairs* (SoA).<sup>10</sup>

Quando o verbo (predicado verbal) requer apenas um argumento para ter sentido completo, trata-se de um verbo intransitivo, capaz de gerar uma oração intransitiva simples (ordem canônica: S-V). Os verbos intransitivos variam de acordo com a natureza do único argumento, que tem função sintática de sujeito.<sup>11</sup> Dessa forma, distinguimos três tipos de verbos intransitivos simples em Pykobjê-Gavião:

**Verbo intransitivo Ativo:** A1 tem a função semântica de Agente, doravante Sa:<sup>12</sup>

1. jôm craa hyycrôn<sup>13</sup>  
QE criança dançarINTR<sup>14</sup>  
'O ser criança dança'

8 Dik (1997, p. 117) afirma, com tradução nossa: O quadro dos predicados verbais básicos pode ser de um, dois ou três lugares, para designar propriedades, respectivamente, de relações binárias ou ternárias. Texto original: "Basic verbal predicate frames can be one-place, two-place, or three-place, designating properties, binary or ternary relations, respectively."

9 Payne (1997, p. 47) lembra-nos que "argumento" é um termo usado para se referir em um predicado verbal aos participantes e seus papéis semânticos (*semantic roles*). Afirma, ainda, que os papéis semânticos expressam relações conceituais existentes no "mundo", que, embora influenciem a morfossintaxe das frases, não correspondem, de fato, a categorias deste nível. Dixon (1994) propôs as categorias S, A e O para identificar a função sintática dos argumentos verbais nas línguas naturais, de modo a estabelecer um método para a identificação dos sistemas de marcação de caso.

10 Uma tradução que tem sido utilizada para designar *State of Affairs* é 'Estado de Coisas'. Dik (1997, p. 124) considera, ainda, que o SoA diz respeito não apenas sobre **o que** é dito, mas também observa **como** é dito.

11 Payne (1997, p. 145) sugere que pode haver duas classes morfossintaticamente distintas de sujeitos intransitivos: argumentos Sa e argumentos Sp. Segundo ele, os argumentos Sa seriam gramaticalmente tratados como argumentos A de verbos transitivos, ao passo que os argumentos Sp seriam tratados como argumentos P de verbos transitivos.

12 Dik (1997, p. 118) afirma, com tradução nossa: Agente – a entidade que controla uma Ação. Texto original: "Agent – the entity controlling an Action".

13 Julgamos necessário manter a glosa 'INTR' para diferenciar verbos com sentido semelhante, mas com teia argumental distinta, como é o caso de *hyycron*, que significa 'dançar sozinho'. Já *peequên* remete à dança realizada em parelha.

14 Seguem abaixo, por ordem alfabética, as glosas das siglas utilizadas ao longo deste artigo, as quais foram baseadas nas *Leipzig Glossing Rules*, desenvolvidas em conjunto pelo Departamento de Linguística do Instituto Max Planck e pelo Departamento de Linguística da Universidade de Leipzig:

**ADV** – classe adverbial; **ASC** – aspecto semi-completo; **AUM** – aumentativo; **BEN** – benefactivo; **DAT** – Caso dativo; **DEM** – demonstrativo; **DIM** – diminutivo; **DIR** – diretivo; **DUR** – aspecto durativo; **EPI** – pronome independente enfático; **ERG** – Caso ergativo; **FUT/LEX** – futuro lexicalizado; **IMP** – partícula de imperativo; **INTR** – verbo intransitivo; **IRR** – modo *irrealis*; **LOC** – locativo; **NPr** – nome próprio; **PASS/LEX** – passado lexicalizado; **PAS/REM** – passado remoto; **PD**: pronome dependente; **PF** – partícula fonte; **PI** – pronome independente; **PL** – marca de plural; **PT** – posposição de tempo; **QE** – quantificador existencial; **QI** – quantificador interrogativo; **WH** – pró-forma interrogativa.

2. Bernardet xwa  
NPr banharINTR  
'Bernardete se banha'<sup>15</sup>
3. cormy wa cato  
ASC 1PI partirINTR  
'Acabei de partir'

**Verbo intransitivo Estado:** A1 tem a função semântica Zero ( $\emptyset$ ), doravante So (ex. 4), ou apresenta semântica relacionada à descrição (exs. 5 e 6):<sup>16</sup>

4. ry'my' ãj - p̃ym  
DUR 1PD-estar.caído/caído  
'Estou caindo'
5. weewee cate'te  
borboleta ser.grande/grande-AUM  
'A borboleta é grande'
6. cahỹj cãn cate  
mulher ser.mau/mau muitoADV  
'A mulher é muito má'

**Verbo intransitivo Processo:** A1 tem a função semântica Processado/Transformado, doravante Sio.<sup>17</sup> O sujeito desse tipo verbal sempre aparecerá marcado com a partícula {mỹ}, indicando ser ele o afetado pelo processo:<sup>18</sup>

7. hõmre - mỹ ã' - cre'cret  
homem-DAT 3PD-ter.medo/medroso  
'O homem tem medo/está medroso'
8. ãj - mỹ cõrcree-re<sup>19</sup>  
1PD-DAT ter.sede/sedento-DIM  
'Eu estou super sedento'
9. jõm craa - mỹ prỹm  
QE criança-DAT ter.fome/faminto  
'O ser criança está com fome'

15 No PB dialeto maranhense o uso é 'banhar' e não 'tomar banho', como no Sudeste (variante televisiva), por isso os indígenas optam por 'banhar' quando precisam usar o Português.

16 Dik (1997, p. 118) afirma, com tradução nossa: Zero ( $\emptyset$ ) – a entidade primariamente envolvida em um Estado. Texto original: "Zero ( $\emptyset$ ) – the entity primarily involved in a State".

17 Dik (1997, p. 118) afirma, com tradução nossa: Processado/Transformado – a entidade que sofre um Processo. Texto original: "Processed – the entity that undergoes a Process".

18 Para saber mais sobre os usos das partículas marcadoras de Caso em Pykobjê-Gavião, tal como {mỹ}, que indica Caso dativo, e {te}, que indica Caso ergativo, sugerimos a leitura de Silva (2011).

19 Observemos que, nesse caso, a marca de diminutivo atua como um intensificador. O mesmo ocorre com a marca de aumentativo em usos como:

Jõm craa tete'te 'O ser criança está limpíssimo'  
QE criança ser.limpo/limpo-AUM

**Verbo intransitivo Estendido:** pressupõem a existência de um constituinte oblíquo para ter sentido pleno (ordem canônica – Sa-V-Obl). Atestamos, por ora, que verbos intransitivos estendidos ocorrem apenas na subclasse dos intransitivos ativos, conforme segue:<sup>20</sup>

10. cē    te    amne  
3PI ir/vir para cá  
'Ele vem para cá'
11. wa    ha    Rōbehax    wyr    te<sup>21</sup>  
1PI IRR Rubiácea DIR ir/vir  
'Eu vou a Rubiácea'
12. ropre hōmre awjahē a'cēt cŷm  
gato macho caçar mata LOC  
'O gato caça na mata'

Além dos verbos intransitivos, o Pykobjê-Gavião também dispõe de verbos transitivos. Os transitivos apresentam, pelo menos, dois argumentos básicos, de acordo com Dik (1997): A (sujeito de verbo transitivo) e P (objeto de verbo transitivo).<sup>22 23</sup> Se os verbos transitivos forem bivalentes, chamá-los-emos de “transitivos simples”. Se forem trivalentes, tratá-los-emos por “transitivos estendidos”, uma vez que A3 será um constituinte oblíquo.

**Verbo transitivo Simples (ordem canônica: A -P -V):** se o tempo/aspecto for passado (modo *realis*), a marca de Caso ergativo será posposta ao sujeito (exs. 13 e 14), a menos que A1 tenha a função semântica [+processado/experienciador], pois, nesta situação, a marcação é de Caso dativo, igualmente posposto ao sujeito (exs. 15 e 16):

13. aa – te    cwyr    cor  
2PD-ERG mandioca comer  
'Você comeu mandioca'
14. rop – te    xoo    pro  
onça-ERG cão pegar  
'A onça pegou o cão'
15. aa – mŷ    ě' – prŷm – te<sup>24</sup>  
2PD-DAT 3PD-querer-AUM  
'Você o desejava'

20 Ferreira (2003, p. 86), assim como nós, divide os verbos intransitivos do Parkatêjê em “intransitivos simples” e “intransitivos estendidos” e os verbos transitivos em “transitivos simples” e “transitivos estendidos”.

21 Observemos que a diferença entre ‘ir’ e ‘vir’, {te}, é determinada pela natureza semântica do constituinte oblíquo que acompanha esse verbo.

22 Payne (1997, p. 133) diz que A é definido como o argumento de uma frase transitiva que mais se assemelha a Agente.

23 Payne (1997, p. 134) diz que P é definido como o argumento de uma frase transitiva que mais se assemelha a Paciente.

24 O verbo *prŷm*, em Pykobjê-Gavião, tem o sentido de ‘querer’, ‘precisar’. Pode requerer apenas um argumento e, neste caso, será usado com a ideia de ‘ter fome/estar faminto’. Quando requer dois argumentos, seu sentido mantém-se como ‘querer’. A forma *prŷmte* está lexicalizada no estado sincrônico da língua com o sentido de ‘desejar muito algo’. Tanto *prŷm* quanto *prŷmte*, em uso transitivo ou intransitivo, exigem que o sujeito (A1) tenha a função semântica Processado/Transformado (SiO), sendo marcado, portanto, com Caso dativo.

16. mam cahỹj – mỹ xoo enta xen  
 PAS/REM mulher-DAT cão DEM gostar  
 ‘Antigamente, a mulher gostava daquele cão’

Se o tempo/aspecto for presente (modo *realis*), apenas a marca de Caso dativo aparecerá posposta ao sujeito (ex. 17). Comparando (16) a (17), observamos que, nestas situações, a diferenciação entre tempo/aspecto passado e tempo/aspecto presente se fará, basicamente, por meio de operadores aspecto-temporais (ex. 16, uso de *mam*), dos quais trataremos a seguir.

17. cahỹj – mỹ xoo enta xen  
 mulher-DAT cão DEM gostar  
 ‘A mulher gosta daquele cão’

Se tempo/aspecto for futuro (modo *irrealis*), não haverá marca de Caso ergativo (exs. 18 e 19) ou de Caso dativo (exs. 20 e 21). Parece que duas partículas não podem se suceder. Então, neste contexto, a partícula verbal terá prioridade sobre a nominal:

18. ca ha cwyr cor  
 2PI IRR mandioca comer  
 ‘Você comerá mandioca’
19. rop ha xoo pro  
 onça IRR cão pegar  
 ‘A onça pegará o cão’
20. ca ha ẽ – prỹm – te  
 2PI IRR 3PD-querer-AUM  
 ‘Você o desejará’
21. cahỹj ha xoo enta xen  
 mulher IRR cão DEM gostar  
 ‘A mulher gostará daquele cão’

Com base nos dados vistos acima, observamos que há uma diferenciação pronominal. No tempo/aspecto passado (modo *realis*), os pronomes estão sempre prefixados a uma partícula, que pode ser {te} (ex. 13) ou {mỹ} (ex. 15), por isso podem ser chamados de pronomes prefixais, pronomes de forma presa ou pronomes dependentes. Quando o tempo/aspecto é presente (modo *realis*), a menos que o sujeito seja [+processado/experienciador], o pronome será independente ou forma livre (ex. 22). Já no tempo/aspecto futuro (modo *irrealis*), o pronome será sempre independente ou forma livre (exs. 18 e 20).<sup>25</sup>

22. ry'my' ca cwyr cor  
 DUR 2PI mandioca comer  
 ‘Você está comendo mandioca’

**Verbo transitivo Estendido (ordem canônica: A-Obl-P-V):** as mesmas discussões de Caso aplicadas a frases transitivas simples são verificadas em frases transitivas

25 Para entender melhor a diferenciação entre “pronome dependente” e “pronome independente”, em Pykobjê-Gavião, sugerimos a leitura de Silva (2012).

estendidas.<sup>26</sup> No entanto, neste contexto, A2 só terá função semântica meta (*goal*), ao passo que A3 tratar-se-á de um constituinte oblíquo, que contemple uma das seguintes funções semânticas (cf. DIK, 1997, p. 120): recipiente (*recipient*), localização (*location*), direção (*direction*), fonte/origem (*source*) ou referência (*reference*). Se o oblíquo for fonte/origem, haverá uso da partícula benefactiva/malefactiva {mỹ} (ex. 23), que, por sua vez, é homófona à marca de Caso dativo. Para as demais funções semânticas, o Pykobjê-Gavião apresenta algumas opções de núcleos posposicionais, dentre eles, citamos os que parecem ser mais profícuos:

23. aa - te co - mỹ cö jōor  
 2-ERG 3-BEN água dar  
 ‘Você deu água para ele’
24. cahỹj ha co – pē jōm craa hōtyr<sup>27</sup>  
 mulher IRR 3PD-PF QE criança salvar  
 ‘A mulher salvará a criança dele’  
 Sentido: ‘Criança será salva pela mulher do mal a ser infligido por ele’
25. ěj – te co wyr cö pyr  
 1PD-ERG 3PD DIR água trazer  
 ‘Eu trouxe água para ele’
26. wa ha cö cỹm tep cwyr  
 1PI IRR água LOC peixe pegar  
 ‘Pegarei peixe no rio’

Até o momento, vimos subtipos possíveis de predicados verbo-adjetivais (VA), de acordo com o número de argumentos exigidos. Em Pykobjê-Gavião, atestamos: verbo intransitivo simples (ativo (Sa), estado (So), processo (Sio)), verbo intransitivo estendido (apenas ativo (Sa)), verbo transitivo simples e verbo transitivo estendido.

A seguir, analisaremos alguns tipos de operadores verbais (tempo/aspecto) verificados.

## Operador aspecto-temporal em Pykobjê-Gavião

Conforme explicado por Givón (2001, p. 50-52), os verbos têm baixa estabilidade temporal, por isso, em Pykobjê-Gavião, as marcas de **tempo/aspecto** e **modo** estão diretamente relacionadas aos verbos.<sup>28</sup> Como veremos melhor adiante, a principal forma para indicar

26 Payne (1997, p. 171) alerta para a distinção existente entre “transitividade semântica” e “transitividade gramatical/sintática” quando se analisam frases como aquelas a que chamamos de “transitivas estendidas”. Sobre suas características, o autor explica, tradução nossa: A valência semântica de um verbo, V, refere-se ao número de *participantes* necessários na cena expressa por V. A valência sintática, então, indica o número de *argumentos* verbais em uma frase na qual V é o predador principal. Texto original: “[...] the semantic valence of a verb, V, refers to the number of necessary *participants* in the scene expressed by V. Syntactic valence, then, is the number of verbal *arguments* in a clause in which V is the main predicator.”

27 O sufixo –pē é a forma contraída do núcleo adverbial com função semântica fonte pēn.

28 Como pode ocorrer em outras línguas naturais (cf. GIVÓN, 1984, p. 65), no Pykobjê-Gavião, as noções de tempo e aspecto parecem estar condensadas em uma única marca morfológica, ao passo em que modo apareceria diferenciado. Autores afirmam que o mesmo se dá nas demais variantes Timbira já descritas (cf. SOUZA, 1997; FERREIRA, 2003; ALVES, 2004; FREITAS, 2008).

tempo/aspecto neste dialeto Timbira é através de itens lexicais com funções dêiticas, tais como ‘ontem’, ‘amanhã’ e ‘agora’.<sup>29</sup> As marcas de modo, em geral, são indicadas por meio de categorias gramaticais (partículas verbais).<sup>30</sup> Em uma mesma frase com predicado verbal, podem coexistir operadores de tempo/aspecto e operadores de modo, conforme vemos nos exemplos dados abaixo:<sup>31</sup>

27. **awca'te** wa **ha** cwyr cor  
 FUT/LEX IPI IRR mandioca comer  
 ‘Amanhã comerei mandioca’
28. **awca'te** ny cahỹj **ha** cwyr cor  
 FUT/LEX PT mulher IRR mandioca comer  
 ‘Amanhã a mulher comerá mandioca’

Comparando (27) a (28), vemos que em (28) o item aspecto-temporal lexicalizado, *awca'te*, vem acompanhado por uma posposição temporal, *ny*. O uso desta partícula é facultativo quando um advérbio temporal está sendo utilizado como operador aspecto-temporal. O mesmo não se observa nos casos em que a função do advérbio temporal é de satélite, conforme uso em (29). Observemos que, neste caso, a omissão de *ny* implicaria em estranhamento, como indicado em (30).

29. empo co ha me **awca'te** **ny** coco?<sup>32</sup>  
 WH 1EPI IRR PL amanhã PT comer  
 ‘O que nós vamos comer amanhã?’
30. \*empo co ha me **awca'te** coco?  
 WH 1EPI IRR PL amanhã comer  
 ‘O que nós vamos comer amanhã?’

Os operadores lexicais aspecto-temporais mais profícuos em Pykobjê-Gavião são:

- Para marcar tempo/aspecto **futuro imediato**, *awca'te*, que na forma adverbial significa ‘amanhã’ (exs. 27 e 28);
- Para marcar tempo/aspecto **futuro distante**, *an cỹm*, que, literalmente significa ‘no período de anos’, como dado abaixo:

29 Fiorin (2004, p. 162) dá as seguintes explicações acerca do uso dêitico: “Todo enunciado é realizado numa situação definida pelos participantes da comunicação (eu/tu), pelo momento da enunciação (agora) e pelo lugar em que o enunciado é produzido (aqui). As referências a essa situação constituem a dêixis e os elementos linguísticos que servem para situar o enunciado são os dêiticos.”

30 Alves (2004, p. 67) considera que, no Apãniekrá-Canela (Timbira), “a partícula **ha** tem mais um uso modal (*realis* vs. *irrealis*) que não requer referência ao tempo futuro”. No Pykobjê-Gavião, acreditamos, igualmente, que a partícula verbal {ha} atua mais como um operador modal do que como um operador aspecto-temporal.

31 Comrie (1985) apresenta as noções para codificar tempo, aspecto e/ou número de uma língua natural em três classes possíveis: 1) categorias gramaticais (flexionais ou livres); 2) itens lexicais com funções semelhantes a dos advérbios; 3) expressões compostas (classe aberta) que a língua pode fornecer.

32 O verbo ‘comer’ pode ser dito *co(r)* ou *coco(r)*. *Coco(r)* significa ‘alimentar-se’ e não apenas ‘comer o alimento x’, como parece ser o uso mais comum para *co(r)*.

31. **an cȳm wa ha cato**<sup>33</sup>  
 ano LOC 1PI IRR partirINTR  
 ‘Dentro de alguns anos irei embora’

- Para marcar tempo/aspecto **passado recente** (Momento da Fala é posterior ao Momento da Referência), *e’no’ny*, que pode ser usado adverbialmente com o sentido de ‘ontem’. Observemos, a partir dos exemplos dados abaixo, que as mesmas considerações feitas sobre o uso de *ny* para *awca te* se aplicam a *e’no’ny* (comparemos (32) a (33) e (34) a (35)):

32. **e’no’ny hōmre ngōr pex**  
 PAS/LEX homem dormirINTR bemADV  
 ‘Ontem o homem dormiu bem’

33. **e’no’ny ny hōmre ngōr pex**  
 PAS/LEX PT homem dormirINTR bemADV  
 ‘Ontem o homem dormiu bem’

34. **empo aa - te e’no’ny ny coco?**  
 WH 2PD-ERG ontem PT comer  
 ‘O que você comeu ontem?’

35. **\*empo aa - te e’no’ny coco?**  
 WH 2PD-ERG ontem comer  
 ‘O que você comeu ontem?’

- Para indicar ‘anteontem’ ou ‘**há poucos dias atrás**’, usa-se *e’no’cȳm*, que não deve estar acompanhado de *ny*.<sup>34</sup>

33 *Cȳm* ocupa a posição nuclear, que poderia ser de *ny*, por isso o uso de *ny* seria altamente redundante.

34 Acreditamos que *e’no’ny* e *e’no’cȳm* partilhem da mesma raiz primitiva, *e’no*. No estado sincrônico do Pykobjê-Gavião, *e’no* pode ser usado como quantificador nominal indefinido com o sentido de ‘algum pedaço de’, como vemos na frase abaixo:

ẽj – mȳ **e’no** cwyr j-acree ‘Vou escolher alguma mandioca para mim’  
 1PD-DAT QI mandioca PR-escolher

Literalmente: ‘Para mim uma parte das mandiocas eu escolho’

Contexto: Há um monte de mandiocas, dentre elas, escolherei a(s) que mais me apetece(m).

*E’no* também pode ser usado como verbo no sentido de ‘dar uma parcela, um pouco’, como veremos abaixo:

ẽj – mȳ cwyr to **e’no!** ‘Dá um pouco de mandioca para mim!’  
 1PD-DAT mandioca IMP dar (parte do todo)

Literalmente: ‘Para mim dê um pedaço de mandioca!’

Contexto: Alguém está comendo uma mandioca e o falante deseja comer um pouco daquele prato.

36. **e'no'c̣ym** aa – te a'tor coran<sup>35</sup>  
 pouco tempo 2PD-ERG lambu matar  
 'Há poucos dias você matou lambu'

- Para marcar tempo/aspecto **passado remoto** usa-se *mam*, que significa 'antigamente', podendo indicar desde algumas décadas atrás até períodos de um tempo mítico. Não é adequado usar qualquer marca adverbial depois de *mam*. Sua posição canônica é em início de frase.

37. **mam** co - m̃ cō xen  
 PAS/REM 3PD-DAT água gostar  
 'Antigamente ele gostava de água'

- Para marcar tempo/aspecto **presente durativo** (presente, Momento da Fala coincide com o Momento da Referência), o mais comum é usar *ry'my'*. Não atestamos quaisquer usos adverbiais para esse operador. Assim como ocorre com demais operadores aspecto-temporais, sua posição canônica é em início de frase:

38. **ry'my'** ējcre enta xēt  
 DUR casa DEM queimarINTR  
 'Essa/esta casa está queimando'

39. **ry'my'** cē cō ē' – pēn  
 DUR 3PI água 3PD-beber  
 'Ele está bebendo água'

- O operador *ry'hy* funciona apenas como marcador de **aspecto incompleto no presente**, mas tem a peculiaridade de exigir que um verbo deôntico (-'re) esteja unido ao verbo principal, conferindo, desse modo, ideia de urgência na completude, conforme aparece no exemplo abaixo:

40. cē **ry'hy** cwyr cor - 're  
 3PI ANC mandioca comer-DEO  
 'Você já devia ter comido a mandioca'

35 Assim como ocorre com outros verbos (exs. *co(r)* = comer, *cato(r)* = sair/partir, etc.), *cora(n)* é um tipo verbal que alterna formas longas e formas breves. Amado (2004, p. 99 em diante) analisa a variação que os verbos apresentam em sua raiz a depender do tempo que expressam. Quando se trata do tempo passado, que, segundo a autora, é o não-marcado em Pykobjê, os verbos apresentam aquilo que foi denominado "forma longa". Já se o tempo for o presente ou o futuro, temos a "forma breve". Sobre a diferenciação pronominal verificada, no tempo não-marcado, o pronome utilizado é dependente, ao passo que nos demais é independente. Vejamos, a seguir, os exemplos retirados de Amado (2004, p. 100, grifos nossos, glosa original), para entendermos melhor como se processa essa diferenciação, de nível fonológico, de acordo com o tempo/aspecto apresentado:

<b>ej</b> – te aʔjē <b>kor</b>	'eu comi a carne'
1 ERG 'carne' 'comer'	
awakaʔte <b>wa</b> ha aʔjē <b>ko</b>	'amanhã eu comerei a carne'
'amanhã' 1 FUT 'carne' 'comer'	
<b>wa</b> aʔjē <b>ko</b>	'eu estou comendo a carne'
1 'carne' 'comer'	

- O operador *cahör* também marca aspecto incompleto, com escopo restrito à classe semântica dos verbos de sucção (*chupar, sugar, comer, beber, etc.*), como vemos no exemplo que segue:

41. jôm craa - re **cahör** pryte' cy mak  
 QE criança-DIM ANC gado leite sugar  
 'O ser criança ainda suga o leite de vaca'

Por fim, julgamos relevante apresentar o operador utilizado para marcar uma ação que está a ponto de ser completada ou que foi completada a pouquíssimo tempo. Sua função, portanto, seria similar a do *just*, em Inglês (ex. 89). Chamaremos tal operador de marca de **aspecto semi-completo**.

42. Inglês:  
 I **just** meet her  
 'Acabei de conhecê-la'

43. wa **cormy** te cöte wyr  
 1PI ASC ir riacho DIR  
 Interpretação (a) 'Eu estou indo ao riacho'. Contexto possível: Caminhando rumo ao riacho, ainda na estrada.

44. tẽ ca **cormy** te cöte wyr  
 INT 2PI ASC ir riacho DIR  
 'Você já foi ao riacho?'

45. wa **cormy** te cöte wyr  
 1PI ASC ir riacho DIR  
 'Acabei de ir ao riacho'  
 Interpretação (b) 'Acabei de ir ao riacho'. Contexto possível: O falante chegou há pouco tempo do riacho em questão.

Com base nos dados vistos acima, podemos concluir que o tempo/aspecto não-marcado do Pykobjê-Gavião é o **presente completo**. Todos os outros tendem a ser ressaltados por meio de um item lexical com função de operador aspecto-temporal. Nesta subseção, vimos os seguintes operadores aspecto-temporais: futuro imediato (*awca'te*), futuro distante (*an cÿm*), passado recente (*e'no'ny*), passado 'há poucos dias' (*e'no'cÿm*), passado remoto (*mam*), presente durativo (*ry'my'*), presente incompleto (*ry'hy, cahör*) e presente semi-completo (*cormy*).

### Algumas considerações

O intuito central desse artigo foi descrever e analisar muito preliminarmente algumas das características mais proeminentes da categoria verbal em Pykobjê-Gavião. Buscamos tratar sobre características distribucionais e estruturais do verbo, enquanto núcleo de um sintagma verbal e/ou de um predicado verbo-adjetival (VA).

Os tipos verbais depreendidos foram: verbo intransitivo simples (ativo (Sa), estado (So), processo (Sio)), verbo intransitivo estendido (apenas ativo (Sa)), verbo transitivo simples e verbo transitivo estendido.

Na subsecção que tratou de alguns tipos de operadores verbais, apresentamos as seguintes marcas, muitas delas dêiticas: futuro imediato (*awca'te*), futuro distante (*an cÿm*), passado recente (*e'no'ny*), passado 'há poucos dias' (*e'no'cÿm*), passado remoto (*mam*), presente durativo (*ry'my'*), presente incompleto (*ry'hy, cahör*) e presente semi-completo (*cormy*). E, por fim, levantamos a hipótese de que o tempo/aspecto não-marcado do Pykobjê-Gavião poderia ser o presente completado.

Possuindo como base esta modesta contribuição aos estudos da categoria verbal, relegamos a estudos futuros uma discussão mais detida sobre a marcação da pluralidade verbal em Pykobjê-Gavião. Tomando como ponto de partida o tema dos operadores, brevemente visto aqui, em intersecção com a discussão da categorial adverbial, poderíamos analisar como esta língua marca mais de um evento, a depender de questões tais como tempo, espaço e/ou variação de participantes.

## REFERÊNCIAS

- ALVES, F. C. *O Timbira falado pelos Canela-Apãniekra: uma contribuição aos estudos da morfossintaxe de uma língua Jê*. 2004. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- AMADO, R. S. Aspectos morfofonológicos do Gavião-Pykobjê. 2004. Tese (Doutorado em Semiótica e Linguística Geral) – Universidade de São Paulo, São Paulo.
- COMRIE, B. *Tense*. Cambridge: Cambridge University Press, 1985.
- DIK, S. C. *The theory of functional grammar* – Part 1: The structure of the clause. 2. ed. Berlin; New York: Mouton de Gruyter, 1997.
- DIXON, R. M. W. *Ergativity*. England/Cambridge: Cambridge University Press, 1994.
- FERREIRA, M. N. O. *Estudo morfossintático da língua Parkatejê*. 2003. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- FIORIN, J. L. Pragmática. In: FIORIN, J. L. (Org.) *Introdução à linguística II. Princípios de análise*. São Paulo: Contexto, 2004. p. 161-186.
- FREITAS, M. F. P. de. *Revisitando os verbos em Parkatêjê: questões relevantes para um estudo morfossintático*. 2008. Dissertação (Mestrado em Letras, Estudos Linguísticos) – Faculdade de Letras, Universidade Federal do Pará, Belém.
- GIVÓN, T. *Syntax, an introduction*. v. 1. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2001.

MATTOS E SILVA, R. V. *O português são dois...* Novas Fronteiras, velhos problemas. São Paulo: Parábola Editorial, 2004.

MELATTI, J. C. *Ritos de uma tribo Timbira*. São Paulo: Ática, 1978.

\_\_\_\_\_. *O messianismo Krahô*. São Paulo: Herder, 1972.

NIMUENDAJÚ, C. *The Eastern Timbira*. Berkeley and Los Angeles: University of California Publications in American Archeology and Ethnology, 1946. v. 41.

PAYNE, T. E. *Describing morphosyntax – a guide for field linguists*. New York: Cambridge University Press, 1997.

RODRIGUES, A. D. *Línguas Brasileiras: para o conhecimento das línguas indígenas*. 4. ed. São Paulo: Edições Loyola, 2002 [1986].

SÁ, R. M. *Análise fonológica preliminar do Pykobyê*. 1999. Dissertação (Mestrado em Semiótica e Linguística Geral) – Universidade de São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo.

SCHACHTER, P. Parts-of-speech systems. In: SHOPEN, T. (Ed.) *Language Typology and Syntactic Description*. v. 1. 2. ed. Cambridge: Cambridge University Press, 2007. p. 1-60.

SILVA, M. S. P.; ROCHA, L. M. Educação bilíngue intercultural entre povos indígenas brasileiros. *Revista da UFG*, 2006. Disponível em: [http://www.proec.ufg.br/revista\\_ufg/dezembro2006/textos/educacao\\_bilingue.pdf](http://www.proec.ufg.br/revista_ufg/dezembro2006/textos/educacao_bilingue.pdf). Acesso em: 30 jun. 2013.

SILVA, T. R. O estatuto das partículas {te} e {mỹ} da língua indígena Pykobjê-Gavião. *Revista Estudos Linguísticos*, São Paulo, v. 40, n. 1, p. 351-363, jan./abr. 2011.

\_\_\_\_\_. *Descrição e análise morfossintática do nome e do verbo em Pykobjê-Gavião (Timbira)*. Amtsgricht München: Lincom Europa Academic Publications, 2012. [v. 66 da série LINCOM Studies in Native American Linguistics].

SOUZA, S. M. *A sintaxe de uma língua de verbo no final: Krahô*. 1997. Tese (Doutorado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

# O léxico e seu uso em textos de alunos de língua estrangeira: um estudo descritivo pautado na Linguística de *Corpus* como subsídio para o ensino

(Lexicon and its use by foreign language students: a descriptive study  
based on *Corpus* Linguistics to assist students with learning)

Celso Fernando Rocha<sup>1</sup>

<sup>1</sup> Faculdade de Ciências e Letras de São José do Rio Preto, UNESP - Universidade Estadual Paulista  
celsotrad@yahoo.com.br

**Abstract:** The aim of this article is to present part of the results of a study concerning the lexicon employed by students learning Spanish language in a Languages degree (Undergraduate Education program). We describe the use and the context in which two verbal forms occur. We made use of *Corpus* Linguistics theory in order to compile two *corpora* of descriptive and argumentative compositions and observe the use of two Spanish verbs (*haber* and *tener* – third person singular). 250 compositions from first and second year students were collected. The *WordSmith Tools* software was applied to generate the list of words and the list of concordance. The verbal forms *hay* and *tiene* were the most used, and in some cases they were applied inappropriately when compared to the traditional Spanish grammar and to an electronic *corpus*. The results were discussed in class and were important to raise consciousness in relation to the students' textual production.

**Keywords:** Applied Linguistics; *Corpus* Linguistics; Spanish learners' *corpora*; lexicon and teaching.

**Resumo:** O objetivo deste artigo é apresentar os resultados parciais de um levantamento lexical de redações de aprendizes de língua espanhola de um curso de licenciatura em Letras. Procuramos descrever o uso e o contexto nos quais duas formas verbais ocorrem. Mais especificamente, recorreremos à Linguística de *Corpus* para compilar dois *corpora* compostos de textos descritivos e argumentativos e observar o emprego dos verbos *haber* e *tener*, conjugados na terceira pessoa do singular. Com relação aos passos metodológicos adotados, foram coletadas 250 redações de alunos do primeiro e segundo anos do referido curso. O programa *WordSmith Tools* foi utilizado para extração das listas de palavras e concordância. Verificou-se que as formas verbais *hay* e *tiene* foram as mais empregadas, e, em alguns casos, seus usos não corresponderam ao prescrito pela gramática tradicional espanhola e ao apresentando em um *corpus* de espanhol. Cabe mencionar que os resultados desse levantamento foram levados à sala de aula e serviram de subsídio para a reflexão sobre a produção textual dos discentes.

**Palavras-chave:** Linguística Aplicada; Linguística de *Corpus*; *corpus* de aprendizes de espanhol; léxico e ensino.

## Introdução

Neste artigo, temos como escopo apresentar os procedimentos teórico-metodológicos e os dados parciais de um estudo maior de investigação do léxico em textos de aprendizes de espanhol como língua estrangeira (E/LE), obtidos por meio de um trabalho de coleta de 250 redações, escritas por dois grupos de alunos do curso de licenciatura em Letras com habilitação em língua espanhola (aproximadamente 15 alunos por grupo, falantes nativos de língua portuguesa). Objetivamos, mais especificamente, identificar as

duas formas verbais mais frequentes nos dois grupos e verificar os cotextos e contextos nos quais foram empregadas, contrastando-as com a norma padrão e com um *corpus* eletrônico de E/LE. Dessa forma, fazemos uso do instrumental disponibilizado pela Linguística de *Corpus* (LC) com intuito de, em um segundo momento, elevar o grau de conscientização do aprendiz, por meio da inserção em sala de aula dos dados extraídos dos *corpora*. Tal etapa de utilização dos dados para o ensino já começou a ser levada a cabo e, portanto, apresentaremos algumas reflexões iniciais sobre o procedimento adotado e seus resultados iniciais.

No que diz respeito à LC, pode-se dizer que vem estabelecendo relações com a Linguística Aplicada: ensino de língua estrangeira há algum tempo. Notamos, no ambiente pedagógico, a presença da LC em, pelo menos, quatro áreas de concentração, a saber: a) descrição de língua nativa, b) descrição de linguagem de aprendiz, c) ensino da metodologia de pesquisa em LC para alunos, e d) desenvolvimento de materiais didáticos.

Na primeira área, congregam-se o estudo da escrita e da linguagem oral de falantes nativos. Os desdobramentos desse tipo de pesquisa apresentam reflexos na produção de livros didáticos e de dicionários que são utilizados por professores e alunos. Desse modo, os levantamentos, análises e descrições linguísticas gerados pela LC estão indiretamente presentes em sala de aula por meio do material utilizado. O professor, nesse contexto, não faz uso direto da LC.

A segunda área de investigação possibilita diversos tipos de pesquisas, bem como de aplicação, em sala de aula, do conhecimento gerado. Para o docente, surge a possibilidade de coletar, armazenar em computador e descrever (qualitativa e quantitativamente) a linguagem produzida por aprendizes. Esses conjuntos de dados podem servir de referência para a produção de material de apoio, para o reforço de algum tópico desenvolvido em aula e para o planejamento de lições, com base nas dificuldades linguísticas mais frequentes dos alunos.

A terceira área estabelece relações com ensino explícito da metodologia da LC para alunos de língua estrangeira (LE). O instrumental analítico da LC auxilia os alunos (professor ou tradutor em formação) na aprendizagem/aquisição de léxico com o qual não tenha tido contato prévio; na compilação de *corpora* (comparáveis ou paralelos) de textos de determinadas áreas de especialidade; na elaboração de glossários bilíngues (espanhol e português); na preparação de aulas e de exercícios com base no léxico mais frequente, por exemplo.

A última área mencionada compreende a criação de abordagens ou metodologias de ensino que tomam como base conceitos advindos da LC. Na literatura da área, destacam-se, pelo menos, três modelos de aproveitamento do conhecimento gerado pela LC: o Currículo Lexical (WILLIS, 1990), a Abordagem Lexical (LEWIS, 1993) e o Ensino Movido por Dados (JOHNS, 1994).

A importância de pesquisas dessa natureza, levando em consideração os aspectos descritos anteriormente, reside no fato de que o mapeamento do conjunto lexical do *corpus* de aprendizes e a identificação de dificuldades mais comuns enfrentadas por graduandos de E/LE, ao empregarem o léxico e a gramática em LE nos seus textos, oferece subsídios para o direcionamento mais adequado do foco instrucional. Em outras palavras, aulas

com maior atenção dada às subáreas nas quais os aprendizes apresentem maiores dificuldades, levando-se em conta suas necessidades reais.

### **Corpus: algumas definições introdutórias**

Para a LC, o termo “*corpus*” possui diversas definições: originalmente significava qualquer coleção de texto em meio digital ou não, organizado de acordo com determinados padrões (BAKER, 1995). A definição sofreu mudanças e passou a significar uma coleção de textos em formato eletrônico, passíveis de serem analisados automaticamente ou semiautomaticamente. Além disso, *corpus* deixou também de significar apenas textos de língua escrita para incluir textos orais e, por isso, é comum encontrarmos *corpora* tanto da fala quanto da escrita de iniciantes em línguas estrangeiras, por exemplo.

Atualmente, a presença de *corpora* torna possíveis análises que não teriam grandes possibilidades de serem executadas há algumas décadas, dadas as dificuldades advindas dos contextos teórico, metodológico e tecnológico. Na década de 1960, as palavras eram transferidas manualmente para cartões perfurados, para serem lidas por meios eletrônicos. Em tal momento, o estudo de determinadas estruturas linguísticas e posterior levantamento de frequências e contrastes apresentavam restrições de execução, devido à abrangência e à constituição do objeto de estudo.

Para Berber Sardinha (1999), uma definição mais abrangente, por incorporar as características principais para a construção de *corpora* em formato eletrônico, é a de Sánchez:

[...] um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise. (SÁNCHEZ, 1995, p. 8-9 apud BERBER SARDINHA, 1999, p. 12)

Ainda, com relação ao *corpus*, sua natureza e seu tamanho permanecem variáveis, adaptando-se às situações e às perguntas de pesquisa, já que se podem ter *corpora* das mais diversas áreas e com variados propósitos. Cabe mencionar que sua compilação exige a observação de pré-requisitos operacionais, como: extensão, representatividade, especificidade e adequação.

Ao tratar do tamanho dos *corpora*, Berber Sardinha (2004, p. 26) argumenta que, apesar de ser um critério importante para sua compilação, há poucas pesquisas que abordam tal característica. O autor menciona três abordagens vigentes para determinar o número adequado de palavras em um *corpus*. A primeira, apresentada como “impressionística”, está baseada na observação e nas afirmações feitas por especialistas. A segunda, “histórica”, leva em consideração a monitoração dos *corpora* efetivamente utilizados pela comunidade ao longo de um período específico. A última tem como fundamento a “estatística”, para definir a quantidade necessária de palavras, a fim de retratar determinados aspectos ou características da língua.

A extensão e a representatividade estão interligadas. Entretanto, não dispomos de critérios objetivos para determinar o quão representativo é um *corpus*, pois, em última

instância, “o ônus de demonstrar a representatividade da amostra e de ser cuidadoso em relação à generalização dos seus achados para uma população inteira é dos usuários” (BERBER SARDINHA, 2004, p. 25).

A continuação, apresentaremos algumas conceituações em relação ao *corpus* de aprendizes.

### **Corpus de aprendizes e algumas pesquisas na área**

Há muito tempo os estudos baseados em *corpora* influenciam o ensino de LE, em especial, de inglês. No século XX (até meados da década de 1980), seis nomes destacaram-se na aplicação de princípios da LC na produção de listas de frequências, feitas manualmente, para servirem de suporte à criação de material didático: Edward Thorndike, Herold Palmer, A. S. Hornby, Irving Lorge e Michael West (cf. BERBER SARDINHA, 2004, p. 252). Edward Thorndike, por exemplo, foi responsável pela publicação do *Teacher's workbook* (1921), levantamento compilado a partir de 4,5 milhões de palavras. O material propagou o ensino da língua materna e estrangeira com base no controle do vocabulário, ou seja, os alunos entrariam em contato com o léxico mais frequentemente empregado. Palmer e Hornby trabalharam na produção de uma *Interim Report on Vocabulary Selection*, que, posteriormente (1953), foi assumido por West (incorporando os trabalhos de Thorndike, Lorge, Palmer e Hornby). O pesquisador foi responsável pela *General Service of English Words* (duas mil palavras mais empregadas do inglês e comparadas com as definições do *Oxford English Dictionary*).

A empreitada levada a cabo pelos pesquisadores mencionados merece destaque, uma vez que a coleta e análise ocorreram em uma época em que não havia computadores. Mesmo na atualidade, coletar milhões de palavras e proceder a análises, com o auxílio do computador, não é tarefa simples.

De acordo com Granger (2002, p. 7), a compilação de *corpora* de aprendizes (*learner corpora* ou CCL, abreviatura para *Computer Learner Corpora*) é uma tarefa que exige muito esforço e tempo inestimável por parte dos pesquisadores: por isso, é importante ter um plano específico de coleta, armazenagem e extração de dados. Cabe salientar que os resultados advindos desse tipo de pesquisa são relevantes, pois podem auxiliar no desenvolvimento de materiais e na elaboração de currículos mais adequados às realidades culturais dos alunos. Segundo Berber Sardinha (2004), o desenvolvimento de materiais didáticos a partir da perspectiva de falantes nativos pode produzir uma visão idealizada da aprendizagem de LE, deixando de explorar as particularidades inerentes ao aprendizado de um grupo específico de alunos. O autor também menciona que, ao compilar *corpora* de aprendizes e explorá-lo por meio de ferramentas computacionais, o professor pode tomar decisões mais pontuais no que concerne à identificação e ao encaminhamento das dificuldades mais comuns enfrentadas pelos alunos.

Neste sentido, estudar as características da língua utilizada pelo aprendiz de LE sob a perspectiva da LC e por meio do *corpus* de aprendiz propicia um ponto de observação privilegiado. Segundo Lavid,

La disponibilidad actual de grandes colecciones de textos en formato digital y de herramientas informáticas ha hecho posible la realización de estudios sobre las regularidades con las

que los hablantes utilizan los recursos gramaticales de una lengua, investigando la distribución de frecuencias de diferentes construcciones, y las relaciones entre estructuras gramaticales y otros factores lingüísticos y contextuales, por un lado, y los factores que afectan a la selección entre las posibles variantes estructurales, por otro. (LAVID, 2005, p. 336)

Nos últimos anos, estudos que utilizam *corpus* passaram a disponibilizar descrições mais precisas de diferentes registros (linguagem oral, discurso formal, discurso acadêmico, diferentes dialetos, etc.). Estudos em língua inglesa demonstraram que existem padrões de uso de língua diferentes entre nativos e não-nativos. Petch-Tyson (1998), por exemplo, estudou a visibilidade do autor em quatro subcorpora (holandês, finlandês, francês e sueco), observando o emprego do pronome de primeira pessoa I e sua co-ocorrência com outros itens linguísticos. A autora descobriu que o uso de expressões como *I believe that*, *In my opinion* e *I am not of the same opinion* mostrou-se mais frequente em textos de não-nativos. Verificou-se, por conseguinte, uma transposição de expressões típicas da fala para o meio escrito pelos aprendizes de língua inglesa. Outros estudos também apontam diferenças do conjunto lexical utilizados por falantes não-nativos de inglês, dando especial atenção à interlíngua.

Por sua vez, em língua espanhola, podemos mencionar o trabalho pioneiro de Soto Balbás (2003), que faz uso de um arcabouço teórico-metodológico da LC para identificar, em um *corpus* de 25 mil palavras, os erros mais frequentes em 216 redações de aprendizes E/LE, nível básico. Também temos conhecimento de Cintrão (2009), que compilou um *corpus* de aprendizes E/LE, com foco na descrição de competência tradutória, e o projeto levado a cabo na Universidade Autônoma de Barcelona, intitulado *Corpus Escrito del Español L2* (CEDEL2), por meio do qual são descritos aspectos de aprendizagem de E/LE por falantes de língua inglesa.

Abordaremos, a continuação, os aspectos metodológicos adotados na confecção do estudo.

### **Passos empregados para a compilação dos *corpora* e extração dos dados**

Com o intuito de analisar o uso dos verbos “*haber*” e “*ter*” (conjugados na terceira pessoa do presente do indicativo; uma vez que foram as formas mais frequentes na lista de palavras geradas pelo *WordSmith Tools* (doravante WST)) empregados por alunos de E/LE de um curso de licenciatura em Letras (primeiro e segundo anos – aproximadamente 15 alunos por turma) em 250 redações, procederemos à compilação de um *corpus* de aprendiz, composto por textos escritos no período de junho a dezembro de 2011. O quadro, a seguir, apresenta o número de redações, o prazo de entrega, os temas e o número de palavras por *corpus* coletado:

**Quadro 1. Número, temas, tipos de redações coletadas e quantidade de palavras**

Número da redação/ mês de entrega	Temas <i>Corpus I</i>	Temas <i>Corpus II</i>
01/ junho de 2011	Como eu sou	Como eu sou
02/ junho de 2011	Como meu melhor amigo é	Meio ambiente
03/ agosto de 2011	A organização pessoal	Descrição sobre a Internet
04/agosto de 2011	Descrever uma foto	Descrição sobre os sentidos
05/setembro de 2011	A casa dos sonhos	Enredo de novela
06/setembro de 2011	A vida em marte	Uma cidade
07/outubro de 2011	O zoológico	O que é a liberdade?
08/outubro de 2011	Uma cidade	Uma foto (feita em lab.)
09/novembro de 2011	O que é a liberdade?	Organização pessoal
10/novembro de 2011	Quais são suas lembranças?	
11/dezembro de 2011	Uma cidade (feita em lab.)	
Números	11 redações: 43.474 palavras	9 redações: 50.306 palavras

Quanto à extensão dos textos, pedimos para que os alunos escrevessem redações com no mínimo 300 e no máximo 500 palavras e as salvassem em formato .txt (textos sem formatação). Quando alguma redação não atingia o número mínimo, entrávamos em contato com o aluno e discutíamos as possibilidades de ampliá-la, ensinando-o a organizar suas ideias e a explorar melhor o tema. Por outro lado, quando o texto excedia o número máximo de palavras, era incorporado ao *corpus* sem nenhum corte. Dos 11 temas de redação do *corpus I*, um foi feito no laboratório de informática (uma cidade); do *corpus II* a redação que descrevia uma foto também foi escrita em laboratório, sem pré-levantamento de informações, consultas às obras de referências ou preparação de rascunhos. A intenção foi contribuir para o aumento da heterogeneidade na construção dos *corpora*. Pretendemos, futuramente, estudar a influência das condições de produção na qualidade do texto produzido.

Para as outras redações, o aluno recebeu o tema com um mês de antecedência e pôde realizar rascunhos em casa e enviar o texto por correio eletrônico. Foi permitido pesquisar em livros, dicionários e sítios da internet, bem como consultar o professor e os colegas. Houve vídeos e momentos de discussão sobre os temas em sala de aula, além de explicitação dos componentes essenciais de um texto descritivo ou argumentativo. Cabe salientar que a maioria dos textos são descritivos e que nosso *corpora* com textos argumentativos está sendo compilado.

A extração dos dados foi feita por meio do WST, um dos programas mais utilizados para pesquisas em LC. Esse *software*, criado por Michael Scott, professor da Universidade de Liverpool, possui três ferramentas: *WordList*, *Keywords* e *Concord*. No desenvolvimento deste trabalho, empregamos as ferramentas *WordList* e *Concord*. A ferramenta *WordList* permitiu criar listas de palavras por ordem de frequência e por ordem alfabética e, por isso, foi possível observar o léxico mais empregado no *corpus*.

Além das listas de palavras, a segunda ferramenta, a *Concord*, gerou listagens das ocorrências de itens específicos ou nódulos (verbos *haber* e *tener*, conjugados na terceira pessoa), acompanhados dos seus respectivos cotextos (texto ao redor da palavra), facilitando, assim, a identificação de sequências maiores nos textos (colocações). Essa ferramenta

também disponibilizou o utilitário *collocates*, responsável pela quantificação dos nódulos e identificação quantitativa de suas co-ocorrências. Dado o escopo da presente investigação, levantamos os primeiros dados referentes às coligações de *hay* e *tiene*.

Cabe acrescentar que as redações foram impressas, os “erros” (problemas de expressão escrita, inadequação lexical, regência verbal, concordância nominal, por exemplo) foram sublinhados manualmente, e os textos entregues aos alunos, sem maiores explicações. Caberia ao aluno, em sala, tentar corrigir seu texto. Caso não conseguisse, poderia perguntar ao colega e, posteriormente, ao professor ou ao grupo. Além dos trechos destacados nas redações impressas, levamos linhas de concordâncias geradas pelo WST para discussão em grupo. Desse modo, empregamos dois instrumentos para a extração dos dados: o WST e as discussões em sala. Também recorremos ao *corpus de referencia del español actual* (CREA, REAL ACADEMIA ESPAÑOLA, 2012), para constatar/ilustrar a possibilidade de alguns usos.

Apresentaremos, a seguir, os resultados obtidos por meio dos levantamentos efetuados.

### Os verbos mais frequentes nos *corpora* analisados

Como mencionado anteriormente, apresentaremos os dados referentes às análises feitas a partir dos *corpora* selecionados. Para efetuar o estudo, as redações foram convertidas em arquivos .txt e inseridas no programa WST. A Tabela 1 apresenta esses verbos – em suas respectivas formas mais frequentes – nos textos dos *corpora* I e II.

Tabela 1. Verbos de maior ocorrência nos *corpora*

Verbo (forma mais frequente)	Ocorrência total	Corpus I	Corpus II
<i>Hay</i>	327	229	98
<i>Tiene</i>	286	187	99
<i>Era</i>	222	123	99
<i>Hacer</i>	207	89	118
<i>Tengo</i>	196	73	123
<i>Soy</i>	180	75	105

Por meio da Tabela 1, é possível observar o uso mais recorrente do verbo *haber* na terceira pessoa do singular do presente do indicativo (*hay* 229 ocorrências), no *corpus* I, e seu menor emprego (106) nas redações dos alunos do segundo ano de licenciatura em letras. O verbo *tener* (*tiene*) e o verbo *ser* (*era*), por sua vez, também na terceira pessoa do singular, registraram menor número de emprego no *corpus* II. Por outro lado, as formas *hacer*, *tengo* e *soy* foram mais utilizadas no *corpus* II, registrando, respectivamente, 118, 123 e 105 ocorrências. Observaremos, a seguir, os usos dos verbos *haber* e *tener* em nossos *corpora*.

### Cálculo dos colocados mais frequentes nos *corpora* I e II com *hay*

Com o objetivo de analisar com maior acuidade as combinações que orbitam os verbos *haber* e *tener* (em suas formas conjugadas) nos dois *corpora*, procedemos ao le-

vantamento, por meio da ferramenta *concord* e de seu utilitário *collocates*, das construções prototípicas dos *corpora*. Em um primeiro momento, extraímos as colocações dos *corpora* e, em etapas futuras de pesquisa, interessa-nos (após a ampliação do *corpus*) estudar se há diferenças de emprego dessas combinatórias nas redações dos diferentes grupos de alunos.

Por ora, apresentamos a Figura 1, na qual se podem examinar os colocados mais frequentes no *corpus* I:

N	Word	With relation	Total	1st Left	2nd Left	L3	L4	L5	L6	L7	Centre	R1	R2	R3	R4	R5
1	HAY	hay 0,000	245	0	0	4	1	3	0	0	229	0	0	3	1	4
2	DE	hay 0,000	79	36	44	3	1	22	0	0	0	1	0	12	0	13
3	Y	hay 0,000	73	22	51	4	4	7	3	4	0	0	5	22	13	11
4	LA	hay 0,000	67	49	18	10	4	6	29	0	0	2	1	1	5	0
5	UNA	hay 0,000	63	2	61	0	0	2	0	0	0	41	7	1	0	3
6	UN	hay 0,000	58	0	49	1	5	2	1	0	0	40	2	4	2	1
7	EN	hay 0,000	52	30	22	7	4	16	3	0	0	2	4	0	5	5
8	NO	hay 0,000	43	31	12	0	4	1	0	0	26	0	3	2	3	4
9	QUE	hay 0,000	41	19	20	2	1	1	3	6	0	1	4	14	6	3
10	EL	hay 0,000	39	24	15	4	5	4	11	0	0	4	3	3	4	1
11	CON	hay 0,000	27	5	22	0	3	2	0	0	0	0	0	0	11	3
12	DONDE	hay 0,000	26	14	12	1	1	0	4	0	0	0	1	4	6	1
13	TAMBIÉN	hay 0,000	24	6	18	0	0	0	0	0	0	16	0	1	0	1
14	MUCHOS	hay 0,000	16	0	16	0	0	0	0	0	0	16	0	0	0	1
15	SCORE	hay 0,000	16	14	2	1	1	0	3	0	0	0	0	1	1	0

Figura 1. Tela com o cálculo dos 15 primeiros colocados mais frequentes com *hay* no *corpus* I

O verbo *hay* apresenta-se com as seguintes combinações mais frequentes à direita: “[x] *no hay* [x]” (26 ocorrências), “[x] *donde hay* [x]” (8) e “[x] *también hay* [x]” (6). Já à esquerda, observa-se o maior número de colocados repetidos: “[x] *hay una* [x]” (41 ocorrências), “[x] *hay un* [x]” (40), “[x] *hay también* [x]” (15) e “[x] *hay muchos* [x]” (15).

Por sua vez, na Figura 2, podemos destacar o uso do mesmo verbo no *corpus* II:

N	Word	With relation	Total	1st Left	2nd Left	L3	L4	L5	L6	L7	Centre	R1	R2	R3	R4	R5
1	HAY	hay 0,000	100	1	1	0	1	0	0	0	91	0	0	0	1	0
2	QUE	hay 0,000	66	14	42	1	0	0	7	6	0	5	10	14	0	5
3	NO	hay 0,000	36	30	6	3	1	2	0	24	0	0	0	3	3	0
4	DE	hay 0,000	33	10	23	4	3	0	3	0	0	0	4	7	7	5
5	EN	hay 0,000	18	8	10	2	2	2	2	0	0	4	3	2	1	0
6	Y	hay 0,000	18	6	12	0	2	1	1	2	0	0	3	3	3	3
7	LA	hay 0,000	15	3	12	2	1	0	0	0	0	2	3	1	1	5
8	EL	hay 0,000	12	4	8	2	2	0	0	0	0	4	1	1	1	1
9	UN	hay 0,000	11	0	11	0	0	0	0	0	0	0	2	1	0	0
10	PERSONAS	hay 0,000	10	2	8	0	1	0	1	0	0	6	1	2	0	0
11	LO	hay 0,000	10	3	7	1	1	0	1	0	0	1	2	0	4	0
12	UNA	hay 0,000	9	1	8	0	0	1	0	0	0	6	2	0	0	1
13	LOS	hay 0,000	8	1	7	0	1	0	0	0	0	0	0	1	4	2
14	PERO	hay 0,000	8	7	1	0	0	0	0	7	0	0	0	0	1	0
15	POR	hay 0,000	7	1	6	0	0	0	1	0	0	1	1	3	0	1

Figura 2. Tela com o cálculo dos 15 primeiros colocados mais frequentes com *hay* no *corpus* II

De modo semelhante, a combinação “[x] *no hay* [x]” (24 ocorrências) foi a mais frequente, seguida de “[x] *pero hay* [x]” (7). À direita, observa-se o uso de “[x] *hay un* [x]” (8) e “[x] *hay una* [x]” (5).

Um aspecto importante a ser observado é o número de uso absoluto dos itens que orbitam a forma *hay*, apresentado na quinta coluna das figuras 1 e 2. Pode-se constatar

que, por exemplo, há emprego maior no *corpus* II de combinações que orbitam o relativo *que* e a partícula negativa *no*. Por outro lado, tais combinações ocorrem em frequência menor no *corpus* I, uma vez que as construções mais recorrentes são aquelas em que aparecem os artigos indeterminados. De modo semelhante, a forma *tiene* (ver figuras 3 e 4) é usada com maior frequência no *corpus* II em perífrases que expressam obrigações e com maior frequência no *corpus* I para expressar posse.

O fator seriação das turmas e a existência de alguns textos diferentes nos *corpora* podem explicar o uso apresentado e, também, indicar ao professor traços de sobreuso ou subuso de determinados padrões, além de contribuir para o desenvolvimento e a inserção de atividades específicas, atentando, por conseguinte, para as necessidades de aprendizagem em determinado grupo.

### Cálculo dos colocados mais frequentes nos *corpora* I e II com *tiene*

A Figura 3, abaixo, ilustra as combinações mais usuais no *corpus* composto por redações dos alunos do primeiro ano:

N	Word	Verb	Total	Left	Right	L1	L2	L3	L4	L5	Center	R1	R2	R3	R4	R5
1	TU	tiene	109	1	1	0	1	0	0	0	107	0	0	0	1	0
2	Y	tiene	76	23	54	4	2	4	0	0	0	2	3	14	21	14
3	LA	tiene	54	33	21	0	3	0	16	0	0	30	0	2	3	0
4	DE	tiene	53	23	30	7	3	11	2	0	0	0	3	14	7	0
5	QUE	tiene	50	31	19	0	3	4	0	13	0	7	0	4	2	0
6	UNA	tiene	50	9	41	1	2	0	0	0	0	30	2	5	2	2
7	UN	tiene	36	6	30	2	4	0	0	0	0	15	3	7	1	4
8	EL	tiene	36	13	23	1	3	4	0	0	0	18	0	2	0	1
9	ES	tiene	34	23	11	7	0	0	0	0	0	0	1	0	1	4
10	LOS	tiene	22	5	17	1	0	2	2	0	0	7	1	1	3	0
11	EN	tiene	19	0	19	1	3	3	1	0	0	2	0	0	3	1
12	TAMBIÉN	tiene	17	0	17	0	0	0	1	7	0	0	0	0	3	1
13	NO	tiene	17	14	3	0	2	1	0	13	0	0	1	0	0	2
14	MI	tiene	16	16	0	4	0	2	10	0	0	0	0	0	0	0
15	POR	tiene	16	9	7	2	3	1	3	0	0	0	0	3	0	4

Figura 3. Tela com o cálculo dos 15 primeiros colocados mais frequentes com *tiene* no *corpus* I

Por meio da Figura 1, constatamos que as combinações mais frequentes no *corpus* foram: “[x] *tiene una* [x]” (30 ocorrências), “[x] *tiene un* [x]” (15), “[x] *tiene el* [x]” (14), “[x] *tiene la* [x]” (10), “[x] *tiene los* [x]” (7); no caso de *también* e *que*, verificam-se combinatorias à esquerda e à direita de forma alternada: “[x] *también tiene* [x]”/ “[x] *tiene también* [x]” (7+6) e “[x] *que tiene* [x]”/ “[x] *tiene que*[x]” (13+7).

No *corpus* II, chama-nos a atenção o uso pouco frequente de perífrases que expressam negação ou obrigação. Podemos constatar tal fato por meio da observação do resultado do cálculo de colocados na Figura 4:

N	Word	With	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	TIENE	tiene	99	0	0	0	0	0	0	0	99	0	0	0	0	0
2	QUE	tiene	47	32	15	1	4	10	6	11	0	3	0	2	5	5
3	DE	tiene	46	16	30	6	4	4	2	0	0	1	4	13	4	8
4	LA	tiene	40	19	21	2	3	5	9	0	0	12	0	1	3	5
5	Y	tiene	31	10	21	2	2	0	3	3	0	0	2	7	9	3
6	EL	tiene	29	14	15	3	2	5	4	0	0	7	3	1	4	0
7	EN	tiene	15	6	9	1	2	3	0	0	0	1	2	0	3	3
8	ES	tiene	15	11	4	1	5	1	4	0	0	0	0	0	2	2
9	NO	tiene	14	13	1	0	0	1	0	12	0	0	0	0	0	1
10	LOS	tiene	14	3	11	1	1	0	1	0	0	2	1	1	4	3
11	UN	tiene	13	6	7	2	1	1	1	1	0	5	1	0	1	0
12	UNA	tiene	12	3	9	1	0	1	0	1	0	6	1	0	0	2
13	LAS	tiene	10	3	7	3	0	0	0	0	0	3	0	1	1	2
14	MUY	tiene	9	4	5	1	0	2	1	0	0	0	1	4	0	0
15	PERO	tiene	9	6	3	0	2	1	1	2	0	0	0	1	2	0

Figura 4. Tela com o cálculo dos 15 primeiros colocados mais frequentes com *tiene* no *corpus* II

No que concerne às combinações com o uso de “que”, observa-se o emprego desse pronome relativo à direita em pelo menos 11 situações ([x] *que tiene* [x]) e 3 vezes com ele ou à direita e à esquerda ou só à direita ([*que tiene que* [...]/ *tiene que* [x]). A construção *tiene que* foi empregada com *decir*, *tener* e *ver* (no infinitivo). Outra construção mais frequente foi a que traz à direita o artigo feminino *la* ([x] *tiene la* + *substantivo feminino*) (*piel*, *capacidad*, *libertad*, *percepción*, etc.)

O cálculo das colocações demonstra que é frequente a combinação *no* + *tener* + *substantivo* (*miedo*, *tiempo*, *sentido*, etc.). Uma futura investigação poderia descrever as características de tais combinações, uma vez que, em alguns dos exemplos do *corpus*, parece haver emprego de *tener* como “coringa”; mais especificamente, por ser verbo de significação mais ampla, estaria sendo usado no lugar de outros mais específicos. Em um exercício realizado como *warm up*, escrevemos na lousa a frase: *Las personas no tienen tiempo para [...]* (exemplo retirado do *corpus* II) e pedimos para que os alunos substituíssem *tiene* por outro verbo e refletissem sobre a mudança. A escolha recaiu sobre o verbo *disponer* (*de*), aumentando, dessa forma, o grau de especificidade. Uma aluna mencionou que a frase “ficou mais espanhola e objetiva”, ou seja, notou que a combinação agrega maior precisão ao período e está mais próxima de um registro formal. Tais exercícios, além de promoverem reflexão, contribuem para o enriquecimento do repertório linguístico em LE.

Na próxima seção, encontram-se observações mais detalhadas sobre o emprego das formas verbais selecionadas.

## Os usos de *hay* nas redações do *corpus* I e II

O verbo *haber*, em língua espanhola, apresenta-se com alguns usos peculiares, e seu emprego adequado pode trazer certa dificuldade para o aprendiz. Trata-se de um verbo irregular com uma forma especial na terceira pessoa do singular (*hay*); é utilizado, obrigatoriamente, como auxiliar na formação dos tempos compostos (*he sabido*, *has entendido*, etc.); em perífrases verbais (p.ex.: as de obrigação: *haber* + *de* + infinitivo); quando é importante mencionar a presença de ser(es) ou objeto(s) designados por um substantivo utilizado após *haber* (p. ex.: *Hay un cuaderno en la mesa.*), de modo a enfatizar a impessoalidade ou matizar características de um objeto qualquer.

Levando em consideração os usos mencionados, passamos a ler, por meio da ferramenta *Concord*, as 327 linhas de concordância nas quais a forma *hay* aparece. O

critério de seleção das linhas foi quantitativo (por meio do levantamento apresentado nas figuras 1 a 4 e da observação dos *clusters* mais frequentes) e qualitativo. Após a leitura, destacamos dez linhas com base em um número abrangente de casos que contemplassem nossa intuição, anotações durante as correções das redações e discussões em sala. A título de ilustração, seguem as dez primeiras linhas eleitas do *corpus* I:

**Quadro 2. Linhas de concordância geradas do *corpus* I**

Linha	Concordância
1	<i>En los dos lados de la cama hay mesillas de noches con una lámpara. Delante de la cama hay un escritorio y una silla. Sobre el escritorio hay mi ordenador, mis hojas, [...]</i>
2	<i>Hay un ropero blanco empotrado delante de la cama.</i>
3	<i>[...] cama doble con una sábana blanca con dibujos de mariposas violetas. Sobre la cama hay algunos cojines también blancos y con mariposas violetas.</i>
4	<i>Hay cuatro ventanas y una cortina blanca.</i>
5	<i>Algunas personas acreditan [sic] que ya hay vida en Marte. No sé, pero pienso que, se [sic] realmente hay, ellos están ciertos [...]</i>
6	<i>Hay un libro en la mesa de la cocina.</i>
7	<i>[...] agradables: cuidan del solo, de las plantas, viven una vida muy tranquila. No hay guerras, no hay problemas, los marcianos no erran [sic], siempre procuran reflejar</i>
8	<i>Al lado de la ventana hay un escritorio y una silla y sobre el escritorio hay un portalápices negro y también una lámpara azul.</i>
9	<i>Sin embargo, en el lado derecho de la casa está la sala de estar [sic], (hay) tiene otra sala también [...]</i>
10	<i>Rio Preto es muy grande, hay muchos coches, gentes y rascacielos.</i>

Cabe salientar que observamos as linhas e identificamos, quase na totalidade dos casos, o emprego de *hay* para mencionar objetos (ex. linhas 1 a 9) e, em outros, para aludir à existência de pessoas, atribuindo à oração um matiz de impessoalidade (ex. linha 10, *mucha gente* é complemento direto do verbo *haber*).

Na linha 1, por seu turno, verifica-se o uso da sequencia *hay mi ordenador, mis hojas [...]*, na qual o verbo *haber* foi empregado equivocadamente no lugar do verbo *estar*, pois o sintagma *mi ordenador* e *mis hojas* seriam sujeitos do verbo *estar*. O uso do pronome *mi* atribui um traço de singularidade ao objeto, impossibilitando o uso de *hay* (impessoal).

Na linha 5, há o emprego adequado do verbo *haber* e a utilização de *acreditar*, não apropriado ao contexto, pois, em espanhol, esse verbo tem o sentido de demonstrar o valor ou a autenticidade de certo objeto. Poderia ser traduzido para o português por “acreditar” (no sentido de dar crédito), “atestar” ou “assegurar”. O verbo adequado ao contexto seria *creer* (*yo creo que...*).

Na linha 7, verificamos o uso do verbo *erran*, que em língua espanhola, assim como em língua portuguesa, pode ter, pelo menos, dois sentidos: o primeiro é andar sem destino fixo ou objetivo (“vagar”) e o segundo é “desacertar”. Apesar de os sentidos terem sido registrados em dicionário, e o aluno ter utilizado uma obra de referência para argumentar que estava certo no emprego do verbo, sabíamos, por conhecimento intuitivo da língua, que o seu uso não estava adequado, ou, no mínimo, a forma não é frequente. Em buscas efetuadas no CREA, a forma conjugada *erran* registrou uma ocorrência com o sentido de

perambular sem destino ([...] *los seres de carne y hueso que, como una multitud de sombras, erran por sus novelas..* Por sua vez, a mesma forma verbal no infinitivo retornou 182 ocorrências com o sentido de “equivocar-se” (sentido constatado nas primeiras 25 primeiras linhas de concordância do CREA).

No trecho 9, há dúvida no uso de *tener* ou *haber*. Em conversas posteriores, o aluno relatou sua dificuldade em compreender e diferenciar o emprego dos dois verbos. Alguns exercícios foram oferecidos ao aluno, e, em outras redações, constatamos o emprego adequado dos dois verbos.

O quadro, a seguir, ilustra alguns usos mais frequentes do verbo *haber* pelos alunos do segundo ano de licenciatura em Letras:

**Quadro 3. Linhas de concordância geradas do corpus II**

Linha	Concordância
1	<i>Pero no es mi habitación. La habitación, aquella en que vivo, es diferente: en ella hay muchos colores. Hay el verde de mis ojos, que hoy no se puede ver, ya que [...]</i>
2	<i>Lo que devemos [sic] meter en la cabeza es que hay la parte buena y la mala en todo en la vida.</i>
3	<i>Hay seres humanos que, a través de las descubiertas científicas, hacen bombas para matar e [sic] destruir, en cuanto otros buscan la cura de enfermedades.</i>
4	<i>El paladar es muy personal, pues hay personas que no les gustan de salado [sic], hay otras que no les gustan de ácido [sic] tampoco y hay gente que les gusta de todo [sic]</i>
5	<i>Hay una alianza entre el hombre y la naturaleza, todo está interconectado</i>
6	<i>Pero hay el silencio, invisible, pero presente en la imagen.</i>
7	<i>Hay cines y centros comerciales.</i>
8	<i>Desafortunadamente, ya hay embotellamiento en mí ciudad porque esa ciudad se ha desarrollado mucho en lo últimos años.</i>
9	<i>La verdad es que no hay colores, ningún... No hay el amarillo de las falsas sonrisas, ni hay el oscuro de los ojos de los muertos. No hay, también, el rojo de las bocas de las putas. Hay sólo yo. Nadie más vive aquí.</i>
10	<i>A pesar de ser una persona extrovertida, hay veces que soy más reservada.</i>

No *corpus* II, podemos destacar alguns usos que estabelecem relações com contextos mais abstratos. Na linha 1, do Quadro 3, há o excerto de uma redação na qual o aluno descreve parte de sua casa. Verificamos o emprego do sintagma *hay muchos colores*, de forma adequada, remetendo não ao contexto físico da casa, e sim, a um estado de espírito ou o possível estado interior do aluno-autor. De modo análogo, na 6, há o uso de *hay el silencio*, sintagma utilizado em uma redação na qual o aluno descrevia uma foto com um “tom” tétrico. Na construção, sobressai o sentido abstrato e indeterminado do sintagma *el silencio*, apesar de estar sendo empregado um artigo determinado o restante da frase apresenta a contradição no plano discursivo, por meio das palavras *invisible* e *presente*. Criando, portanto, um jogo expressivo duplo, entre o conhecido e o desconhecido, entre o determinado e o indeterminado. Segundo Gili (2005, p. 159), a estrutura gramatical de uso seria: *verbo na terceira pessoa + artigo indeterminado ou um adjetivo indefinido e um substantivo*. No caso citado, há inserção de componente determinado gramatical (*hay el silencio*) e dois elementos lexicais, um que indetermina (*invisible*) e outro que acrescenta certo matiz de determinação (*presente*) ao tipo de silêncio (o uso de *pero* contribui

para o aumento do grau de indefinição). No CREA, foram encontrados dois usos para o sintagma *hay el silencio* (usos metafóricos) e mais de 150 usos para construção *hay silencio* (sem intercalação de artigo e em contextos metafóricos – *hay el silencio interno de las aves / de las leyes, de una habitación*). Cabe dizer que, de modo semelhante, o exemplo da linha 9 também explora os níveis de determinação e indeterminação (ou conhecido e não conhecido).

No exemplo da linha 2, por seu turno, o emprego do artigo determinado (... *hay la parte buena*...) não está adequado, e deveria ser substituído pelo artigo indefinido *una*, para que possa haver um nível de indeterminação e não especificação da “parte” da vida que é “boa” ou “má”. Seria difícil precisar quais seriam as “partes boas” da vida ou as “partes más”. Dessa forma, somente com o uso do indefinido a generalização dicotômica terá seu sentido expresso de forma “eficiente”. Nas linhas 3, 4, 7, 9 e 10, verifica-se o uso adequado de *hay* + substantivo plural. O verbo se mantém na terceira pessoa do singular.

### Os usos de *tiene* nas redações do *corpus* I e II

Como consta da Tabela 1, o verbo *tener*, em sua forma conjugada de terceira pessoa do singular, registrou 286 ocorrências totais (187 + 99) nos *corpora* do presente estudo. Trata-se de um verbo irregular e intransitivo com o sentido de estabelecer relação entre uma pessoa e um objeto possuído. É utilizado em inúmeros contextos em construções como: *tener a alguien de ayudante, tenerlo como ayudante, tener a alguien por idiota, tenerla tomada con alguien; tener que + infinitivo*, para demonstrar necessidade, obrigação ou determinação para fazer aquilo que está expresso por meio do infinitivo. Na construção *tener de + infinitivo*, expressa vontade, e, em *tener + participio*, ressalta ato terminativo.

Por meio da ferramenta *Concord*, foram extraídas as linhas de concordância para a forma *tiene*. O quadro 4, na continuação, apresenta algumas linhas nas quais o verbo foi utilizado no *corpus* I.

Quadro 4. Linhas de concordância geradas do *corpus* I

Linha	Concordância
1	<i>El colegio tiene y tendrá una grande [sic] parte en mi memoria gastada de mi cabeza.</i>
2	<i>Apesar [sic] de pequeña, tiene casi todo lo que los habitantes necesitan para vivir:[...]</i>
3	<i>Ultimamente [sic] tiene ganado [sic] premios por su participación activa para [sic] la preservación del medio ambiente</i>
4	<i>Mariana tiene pelo liso, corto y negro, también tiene gafas y es baja y delgada.</i>
5	<i>Nuestro amigo se llama Andrés, tiene veintiocho años, es ingeniero [sic] estudia inglés. Él vive sólo en Córdoba y no tiene ninguna mascota. Su pelo es negro</i>
6	<i>Él tiene un perro.</i>
7	<i>Claro que tiene las personas malas, pero es una ciudad pequeña, con solo 150 mil habitantes [...]</i>
8	<i>En mi habitación tiene una cama [...]</i>
9	<i>De el [sic]comedor se tiene acceso a la cocina.</i>
10	<i>Para nosotros ese lugar tiene que ser preservado de la manera que Dios lo hizo.</i>

O uso mais frequente foi no sentido de atribuir a posse de um objeto/característica/animal a um ser animado. Também registramos uso de lugares (casa e cidade) como detentores de características ou objetos (exemplos 2 e 10).

Com relação aos usos menos frequentes, podemos mencionar o primeiro exemplo, no qual o verbo está atribuindo, de forma metafórica, a posse de parte da memória ao sujeito *colegio*. Em vez de o sujeito ser representado por uma pessoa detentora de memória e, portanto, capaz de registrar os fatos independentemente, o sujeito *colegio* passa a ser dono da memória do aluno-autor.

Na linha 3, o emprego de *últimamente* pressupõe o uso da seguinte construção: *estar* + gerúndio (ou reescrevendo com *pretérito perfecto*: *En este año la ciudad ha ganado muchos premios...*). Quando a expressão é utilizada como equivalente de *haber* + participípio, pode-se constatar um caso de regionalismo galego e asturiano (cf. SECO, 1989, p. 358).

Nas linhas 7 e 8, há o uso de *tiene* no lugar de *hay* (ou *existir*). Talvez por influência da língua portuguesa, uma vez que é frequente empregarmos “ter” no lugar de “haver” em registros menos formais (exemplo: “tem um carro na garagem” / “há um carro na garagem”). E, em outros contextos em língua portuguesa, podemos permutar uma forma pela outra, em casos em que “ter” e “haver” funcionem como verbos auxiliares (eles tinham chegado / eles haviam chegado).

É importante salientar que esperávamos maior ocorrência desse fenômeno, já que, em linguagem oral, durante as atividades de conversação em sala de aula, houve emprego acentuado de *tener* no lugar de *haber*; no entanto, após a leitura das linhas de concordância, elencamos apenas 3 ocorrências nas quais essa substituição foi feita.

Outros usos observados foram os das linhas 9 (menos frequente) e 10. No primeiro caso, há o uso de “se” apassivador, e, na última, a perífrase verbal *tener que*, expressando obrigatoriedade. O último ocorreu 7 vezes no *corpus* I.

No quadro abaixo, apresentamos os exemplos de uso retirados do *corpus* II:

**Quadro 5. Linhas de concordância geradas do *corpus* II**

Linha	Concordância
1	<i>Los colores de la naturaleza tan radiantes e iluminados por los rayos del sol parece [sic] nos dar las sugerencias de que nuestra vida tiene que tener muchos colores.</i>
2	<i>Ella tiene 19 años y vive en Catanduva y todos los días viaja para Rio Preto.</i>
3	<i>Tiene su cabello negro y lleno, que siempre lleva una cola de caballo.</i>
4	<i>Cada especie tiene el suyo lugar, cada una tiene la suya responsabilidad en la vida.</i>
5	<i>Sin embargo, sabemos que todo lo que nos rodea tiene sus lados positivos y negativos.</i>
6	<i>Tiene pelo rubio, lleva gafas, es un poco baja, pero no es gorda. Es una chica muy guapa.</i>
7	<i>La vida, ese gran ordenador, que tiene mucha memoria temporaria, pero poca memoria física.</i>
8	<i>Pero una cosa más llena esas calles, algo que no puedo explicar, pues es algo que no es visible, es algo que tiene que ver con el legado que esas pequeñas y estrechas sendas cargan.</i>
9	<i>Bueno, después te cuento más cosas, es que ahora Antonio y yo vamos a conocer otro sitio de la ciudad, parece que es otra finca de su familia, donde tiene varios animalitos y todo más; y él quiere mostrarme eso.</i>
10	<i>El internet es un servicio óptimo y, con atención, solo tiene a ayudar.</i>

Tanto no *corpus* I quanto no *corpus* II os usos mais frequentes estão relacionados à atribuição de posse de um objeto ou característica a um sujeito (linhas 2,3,4,5,6 e 7). E os usos menos frequentes dizem respeito ao emprego de *tener* em perífrase verbais (linha 1).

No exemplo da linha 3, a construção com o uso do adjetivo possessivo *su* acrescenta um traço de cuidado em relação à manutenção do cabelo. Em segundo plano, teríamos a constatação das propriedades do cabelo, ou seja, o fato de seu cabelo *ser* negro (*tiene el cabello negro* expressa, em espanhol, o mesmo sentido em português de “seu cabelo é negro”). Na linha 4, há redundância de informação por meio do emprego (equivocado) da forma tônica do pronome no lugar da forma átona. Como se sabe, a forma tônica é responsável pela inserção no contexto de informações novas a respeito de determinado substantivo. O emprego correto seria *cada especie tiene su lugar*, com a forma átona, introduzindo, por conseguinte, elementos não conhecidos.

Com relação aos usos menos frequentes, podemos mencionar o exemplo da linha 8, que traz uma construção, ressaltando o estabelecimento de relações entre um objeto e outro. A expressão também é utilizada na forma negativa (*no tener algo nada que ver con otro*).

No exemplo 9, por seu turno, registramos o único caso de emprego de *tener* no lugar de *haber*.

### **Observações finais**

Ao empregarmos o arcabouço teórico-metodológico citado e compilarmos dois *corpora* de aprendizes de E/LE, pudemos acercar-nos de fenômenos colocacionais mais frequentes nos *corpora* analisados, além de iniciar a utilização dos dados levantados em nossas atividades didáticas, uma vez que alunos e docente iniciaram a observação das linhas de concordâncias selecionadas e passaram a refletir sobre o emprego dos dois verbos nas redações.

Em um dos momentos de reflexão, pudemos promover uma ampliação sobre a compreensão de aspectos relacionados ao uso do E/LE. Mais detalhadamente, a turma do primeiro ano começou a explorar as várias possibilidades de significação e as combinações de uso do verbo *haber*. Notaram que havia uma prescrição mais fechada, no material didático adotado, e um uso mais flexível e mais metafórico nos exemplos do CREA e nas próprias redações. A “descoberta”, de certo modo, surpreendeu vários alunos, pois alguns ainda acreditam (ou acreditavam) que as obras de referência trouxessem *todas* as formas de uso e normatizassem o uso do idioma estudado de maneira extensiva. Os contatos com a diversidade e a observação sistematizada do emprego inadequado do léxico passaram a despertar nos discentes uma inquietação salutar, ampliando, por conseguinte, o leque de consulta aos *corpora* eletrônicos e à *web*, além, claro, de manter as consultas aos dicionários e às gramáticas. Assim, consideramos que a primeira experiência de levar os dados gerados para serem discutidos em aula foi útil e será ampliada.

Em etapas posteriores da pesquisa – além de abordarmos a descrição de outros itens lexicais – trataremos das dificuldades no gerenciamento adequado de um *corpus* de aprendiz e de suas idiossincrasias; a importância da devolução em menor tempo das

redações aos alunos, do *feedback* consistente e do tratamento, em sala de aula, das dificuldades encontradas. Também merecerá espaço a discussão a respeito da forma de coleta, das instruções de organização da escrita para o aprendiz, entre outros aspectos.

Por fim, neste artigo, tivemos como objetivo apresentar os dados parciais de um estudo que se encontra em desenvolvimento e que tem como foco a descrição e o encaminhamento em sala de aula de questões interligadas à aprendizagem de E/LE.

Esperamos que a metodologia adotada possa ser empregada por outros pesquisadores e que os resultados possam servir de base para a reflexão sobre a aprendizagem de LE.

## REFERÊNCIAS

- BAKER, M. *Corpora in Translation Studies: an Overview and Some Suggestions for Future Research*. *Target*, Amsterdam, v. 7, n. 2, p. 223-243, 1995.
- BERBER SARDINHA, A. P. A influência do tamanho do *corpus* de referência na obtenção de palavras-chave. *DIRECT Papers*, 38; São Paulo: PUC/SP, 1999. Disponível em: <<http://sites.uol.com.br/tony4/homepage.html>>. Acesso em: 24 jun. 2005.
- \_\_\_\_\_. *Linguística de corpus*. São Paulo: Manole, 2004.
- CINTRÃO, H. Um *corpus* de tradução no par português-espanhol. *Tradução e Comunicação*, São Paulo, n. 18, p. 131-144, 2009.
- GILI, O. C. *Diccionario práctico de gramática*. Madrid: Edelsa, 2005.
- GRANGER, S. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam/Philadelphia: John Benjamins, 2002.
- JOHNS, T. From printout to handout: Grammar and vocabulary teaching in the context of Data-driven Learning. In: ODLIN, T. (Ed.). *Perspectives on Pedagogical Grammar*. New York: Cambridge University Press, 1994. p. 293-313.
- LAVI, J. *Lenguaje y nuevas tecnologías*. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI. Madrid: Cátedra, 2005.
- LEWIS, M. *The lexical approach: the state of ELT and a way forward*. Hove: Language Teaching Publications, 1993.
- PETCH-TYSON, S. Writer/reader visibility in EFL written discourse. In: GRANGER, S. (Org.). *Learner English on Computer*. Nova York: Longman, 1998. p. 107-118.
- REAL ACADEMIA ESPAÑOLA. *Corpus de Referencia del Español Actual (CREA)*. 2012. Disponível em: <<http://corpus.rae.es/creanet.html>>. Acesso em: 8 jul. 2012.
- SÁNCHEZ, A. Definición e historia de los *corpus*. In: SÁNCHEZ, A. et al. (Org.) *Cumbre: Corpus lingüístico del Español contemporáneo*. Madrid: SGEL, 1995. p. 7-24.

- SCOTT, M. *WordSmith Tools*. Version 4. Oxford: Oxford University Press, 2006.
- SECO, M. *Diccionario de dudas y dificultades de la lengua española*. Madrid: Espasa Calpe, 1989.
- SOTO BALBÁS, M. *Análise de erros, baseada na Lingüística de Corpus, da escrita de aprendizes brasileiros universitários de Espanhol como Língua Estrangeira*. 2003. 128f. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) – Pontifícia Universidade Católica de São Paulo, PUC-SP, São Paulo.
- THORNDIKE, E. L. *Teacher's wordbook*. Nova York: Columbia Teachers College, 1921.
- WEST, M. *A general service list of English words*. Londres: Longman, 1953.
- WILLIS, D. *The lexical syllabus: a new approach to language teaching*. London/Glasgow: Collins ELT, 1990.

# A tradução de obras de Patrícia Melo, um estudo à luz da Linguística de *Corpus*

(The translation of Patrícia Melo's books, a study based on Corpus Linguistics)

Elisangela Fernandes Martins<sup>1</sup>

<sup>1</sup> Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto,  
UNESP - Universidade Estadual Paulista

efmartins1d@hotmail.com

**Abstract:** We present a study about the translation of *O matador* and *Inferno* by Patrícia Melo. We reflected about how they project the violent context in the two works, and how much the translator recovered our social aspects. We believe that translators use some resources to make the target text easier to the readers. Baker (1996) suggests the investigation of translational behaviors which are characteristics of the language of translation. The purpose of this work is to identify normalization aspects presented in these translations. We looked back on Corpus Based Translation Studies (BAKER, 1993, 1996; CAMARGO, 2005, 2007), Corpus Linguistics (BERBER SARDINHA, 2004) and also on normalization studies conducted by Scott (1998). We could observe changes in language register, omissions and additions. We hope to contribute to the understanding of the tendencies presented and show how the intersection of translation studies and corpus linguistics can contribute to the development of the research in the area.

**Keywords:** Patrícia Melo; Literary Translation; Corpus Linguistics; Corpus Based Translation Studies

**Resumo:** Apresentamos um estudo da tradução de *O matador e Inferno*, de Patrícia Melo. Refletimos sobre como são projetados no exterior a realidade violenta de grandes centros urbanos brasileiros. Acreditamos que tradutores usam recursos que tornam a leitura da obra mais fácil. Baker (1996) propõe a investigação de tipos de comportamento linguístico característicos de textos traduzidos. Nosso objetivo é identificar aspectos de normalização presentes nas traduções. Recorremos aos Estudos da Tradução Baseados em *Corpus* (BAKER, 1993, 1996; CAMARGO, 2005, 2007), à Linguística de *Corpus* (BERBER SARDINHA, 2004) e aos estudos sobre normalização de Scott (1998). Pode-se observar a ocorrência de mudanças no registro de linguagem, adições e omissões. Esperamos contribuir para conscientização das tendências apresentadas e das possibilidades oferecidas pela intersecção dos Estudos da Tradução Baseados em *Corpus* e da Linguística de *Corpus*.

**Palavras-chave:** Patrícia Melo; Tradução Literária; Linguística de *Corpus*; Estudos da Tradução Baseados em *Corpus*

## Introdução

Os estudos da tradução desenvolveram-se acentuadamente nos últimos anos, valorizando o texto traduzido, antes visto como um produto secundário, passando a considerar a tradução como produtora de significados e não apenas como um transporte de significados do original para a língua alvo. Nessa perspectiva, o texto alvo sofre a intervenção do tradutor para adequar-se às necessidades do leitor de língua inglesa, e a fatores ideológicos, políticos, editoriais e mercadológicos. Os questionamentos sobre a tradução envolvem, via de regra, diferenças socioculturais que dificultam o trabalho do tradutor e requerem maior reflexão acerca dos recursos linguísticos empregados nos textos alvo.

Tais recursos deixam traços nos textos traduzidos e merecem a atenção de pesquisadores. Por considerar a natureza da linguagem da tradução diferente da linguagem do texto fonte, Baker (1996) defende a ideia de que a tradução é um evento de comunicação genuíno e propõe a investigação de características específicas que tendem a ocorrer na tradução. Entre esses traços típicos, investigamos a normalização, tendência em traduzir o texto original adequando-o aos padrões da língua e cultura alvo.

Para desenvolver esta pesquisa, selecionamos para constituir nossos *corpora* de estudo as seguintes obras da escritora Patrícia Melo: *O matador* (1995) e *Inferno* (2001), e as respectivas traduções *The Killer* (1997), *Inferno* (2001). As obras selecionadas apresentam em comum a temática da violência e criminalidade, tanto na descrição do ambiente hostil de grandes centros urbanos como São Paulo e Rio de Janeiro e na apresentação de problemas sociais brasileiros, quanto na reflexão sobre aspectos psicológicos dos personagens envolvidos com a criminalidade.

A análise das referidas traduções permite uma reflexão sobre como são projetados no exterior uma visão da realidade violenta de grandes centros urbanos brasileiros, e sobre o quanto o tradutor conseguiu resgatar da nossa sociedade no texto traduzido. Justifica-se, desse modo, a observação desses dados e uma investigação sobre tendências de normalização a serem identificadas nas duas traduções.

## Fundamentação teórica

Este estudo apoia-se no arcabouço teórico-metodológico dos Estudos da Tradução Baseados em Corpus (BAKER, 1993, 1996; CAMARGO, 2005, 2007), da Linguística de Corpus (BERBER SARDINHA, 2004); e, também, nos estudos sobre normalização de Scott (1998).

Os estudos sobre a tradução envolvem diferentes questões. Até a década de 1960, as pesquisas recaíam sobre questões da equivalência e fidelidade, e tinham caráter avaliativo ou prescritivo (BAKER, 1993). Esses estudos, no entanto, deixavam de considerar alguns tópicos relevantes que influenciam o ato tradutório, tais como: aspectos situacionais, estilo do tradutor, diferenças socioculturais, natureza da linguagem da tradução, injunções mercadológicas, entre outras.

A partir da década de 1970, uma nova corrente de pensamento passou a sugerir a necessidade de estudos descritivos que não se baseassem em características prescritivas. Essa vertente teórica dos Estudos Descritivos da Tradução serviu como base para a proposta de Baker (1993, 1995, 1996), que visa a investigar características peculiares de textos traduzidos, a partir do exame de *corpora*. A pesquisadora afirma que “os textos traduzidos registram eventos de comunicação genuínos e como tais não são nem inferiores nem superiores a outros eventos comunicativos em qualquer língua. Entretanto, eles são diferentes, e a natureza dessa diferença precisa ser explorada e registrada”<sup>1</sup> (BAKER, 1993, p. 234). Atualmente, já há uma série de pesquisas desse tipo realizadas na área dos Estudos da Tradução Baseados em *Corpus*.

---

1 No original: *Translated texts record genuine communicative events and as such are neither inferior nor superior to other communicative events in any language . They are however different, and the nature of this difference needs to be explored and recorded.* (As traduções das notas de rodapé são de nossa responsabilidade)

A proposta da teórica parte de grandes quantidades de textos compilados em *corpora*, a fim de serem analisados por meio de programas computacionais. Baker define *corpus* como: “qualquer conjunto de textos naturais (em vez de exemplos/sentenças), organizado em formato eletrônico, passível de ser analisado, preferencialmente, em forma automática ou semi-automática (em vez de manualmente)”<sup>2</sup> (BAKER, 1995, p. 226).

As pesquisas propostas por Baker (1995) envolvem três tipos principais de *corpora* em tradução, dependendo dos objetivos específicos de cada pesquisa:

- *Corpus* paralelo: composto de TOs (textos originais) em uma determinada língua (língua de origem) e suas respectivas traduções em outra língua (língua de tradução). Esse tipo de *corpus* permite pesquisar traduções consagradas de certos itens lexicais ou estruturas sintáticas, peculiares de determinado(s) tradutor(es), diferenças entre traduções de um mesmo texto, produzidas em períodos diversos, normas tradutórias, etc.
- *Corpus* multilíngue: consiste em um conjunto de dois ou mais *corpora* monolíngues, sendo cada *corpus* em uma língua diferente; e
- *Corpus* comparável: consiste em dois conjuntos de textos em uma mesma língua: um composto de TOs e outro de TTs (textos traduzidos) para a língua em questão, a partir de uma única LF (língua fonte) ou de diversas LFs. (BAKER, 1995, p. 230-234 apud CAMARGO, 2007, p. 18)

Os estudos realizados por Baker a partir de *corpora*, com o auxílio de ferramentas computacionais, permitem observar certos traços recorrentes nos textos traduzidos, a saber:

- Simplificação: tendência de simplificar a linguagem usada na tradução <sup>3</sup>(BAKER, 1996, p. 181). Pode-se observar a simplificação, também, na quebra de frases longas do texto de partida, e na mudança do uso dos sinais de pontuação para trazer maior clareza para o texto de chegada. Há, ainda, mais duas maneiras que podem evidenciar traços de simplificação: a razão forma/item (*type/token ratio*) e a densidade lexical. A primeira é a medida da variação vocabular presente num *corpus*. Uma razão forma/item mais baixa nos textos traduzidos em relação aos textos originais sugeriria maior repetição de vocábulos nas traduções. A densidade lexical, por sua vez, revela a proporção de palavras de conteúdo em oposição a palavras gramaticais de um corpus. O uso de maior quantidade de palavras gramaticais e menor de palavras lexicais indicaria uma tentativa de simplificar a mensagem para o leitor da língua de chegada.
- Explicitação: tendência de acrescentar informações ao texto traduzido, as quais estão implícitas no texto original. Observa-se um aumento do texto traduzido em relação ao texto original, por meio da inserção de palavras, locuções e, até mesmo, de frases para explicar o significado de elementos que são desconhecidos para o leitor da língua de chegada.
- Normalização: tendência de exagerar características da linguagem do texto de chegada, ou de adequá-la aos padrões da língua alvo. Pode ser observada no uso de clichês

2 No original: *Corpus mean[s] any collection of running texts (as opposed to examples/sentences), held in electronic form and analysable automatically or semi-automatically (rather than manually).*

3 No original: *The tendency to simplify the language used in translation.*

e estruturas gramaticais convencionais da língua de chegada. Baker (1995) afirma que essa tendência é influenciada pelo *status* do texto e da língua de chegada; desse modo, quanto maior for o *status* da língua ou do texto original, menor será a propensão à normalização.

- Estabilização: é “a tendência de o texto traduzido caminhar em direção ao centro de um continuum”, ou seja, afastar-se de “quaisquer dois extremos, convergindo em direção ao centro, com as noções de centro e de periferia sendo definidas a partir do próprio corpus traduzido” (BAKER, 1996, p. 184).<sup>4</sup> A estabilização não depende nem da língua de partida nem da língua de chegada. Um tradutor pode usar a língua culta no lugar de marcas dialetais, por exemplo. Das quatro características típicas da tradução, esta é a que recebeu menos atenção dos pesquisadores.

Com relação à Linguística de *Corpus*, a definição de Berber Sardinha explica que: “ela ocupa-se da coleta e da exploração de corpora, ou conjunto de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística” (BERBER SARDINHA, 2004, p. 3).

Nesta pesquisa, investigamos características de normalização, por ser importante verificar quais recursos linguístico-tradutórios são utilizados para tornar mais fácil a leitura para o público alvo. Além da proposta de Baker (1996) sobre normalização, adotaremos as características que contribuem para a normalização levantadas por Scott (1998).

Scott analisou aspectos de normalização na tradução da obra *A hora da estrela*, de Clarice Lispector, traduzida por Pontiero, identificando, na referida investigação, onze características referentes a traços de normalização:

1. Diferenças no comprimento dos textos e sentenças no texto alvo em relação ao texto fonte.
2. Diferenças no emprego da pontuação;
3. Diferenças no uso de estruturas sintaticamente complexas;
4. Diferenças de aspectos relacionados a ambiguidades;
5. Diferenças relacionadas a imprecisões de expressões;
6. Diferenças no emprego de metáforas incomuns;
7. Mudança de registro de linguagem coloquial para formal;
8. Ocorrências de omissões e/ou adições;
9. Colocações menos comuns por mais comuns;
10. Outras mudanças na tradução;
11. Padrões de repetição.

Para este artigo, procedemos um estudo utilizando os *corpora* do tipo paralelo, que permitem fazer a comparação dos textos originais alinhados às respectivas traduções,

<sup>4</sup> No original: [...] *the tendency of translated text to gravitate towards the centre of a continuum. [...] between any two extremes, converging towards the centre, with the notions of centre and periphery being defined from within the translation corpus itself.*

no intuito de observar as opções de tradução adotadas por Landers, o que possibilitou investigar as características apresentadas.

## **Procedimentos metodológicos**

Para a realização deste trabalho, foi compilado um *corpus* do tipo paralelo formado pelas obras originais *O Matador* (1995) e *Inferno* (2001) e as respectivas traduções *The Killer* (1997) e *Inferno* (2003), os quais foram alinhados para facilitar a observação dos dados.

As obras foram escaneadas por meio do OCR (*Optical Character Recognition*) e revisadas, utilizando-se o corretor ortográfico do Word. Os textos foram salvos em formato “txt” para que pudessem ser processados pelas três ferramentas do WordSmith Tools, versão 4.0, obtido pela internet, no endereço [www.lexically.net](http://www.lexically.net). O programa disponibiliza três ferramentas: WordList, KeyWords e Concord.

Com o auxílio da ferramenta WordList, foram extraídas as listas de palavras dos *corpora* de estudo, uma do texto original *O matador*, uma da tradução *The Killer*, uma do texto original *Inferno* e uma da respectiva tradução *Inferno*. Essas listas apresentam em ordem decrescente todas as palavras do *corpus* de estudo com o respectivo número de ocorrências. Dessa forma, foram obtidos os vocábulos mais frequentes nas obras originais.

O próximo passo foi a utilização da ferramenta *Concord*, para, a partir das palavras mais frequentes, verificar o sentido com que foram utilizadas na obra original. A lista de concordâncias apresenta a palavra de busca inserida em seu contexto, em todas as ocorrências no *corpus*. Por meio da observação das linhas de concordância, pudemos, então, selecionar as ocorrências dos vocábulos mais frequentes com sentido relacionado à violência, bem como excluir outros vocábulos homônimos.

Procedemos, a seguir, ao alinhamento dos textos, por meio da função Viewer & Aligner do programa, o qual exigiu alguns reajustes manuais. Depois, por meio do ícone “find” do Viewer & Aligner observamos as opções de tradução para cada vocábulo e os possíveis traços de normalização.

## **A violência na literatura**

A literatura que retrata a violência urbana tem sido objeto de muitas discussões recentemente, levando pesquisadores e estudiosos a uma reflexão sobre o caráter dessa forma contemporânea de produção literária. Há uma confluência entre o fazer literário, arte, denúncia social, consciência crítica sobre o mundo e a sociedade. Essa nova vertente requer uma reformulação do modelo de análise para que se possa compreender e explorar as diferentes formas de expressão dessas obras, geralmente marcadas pela construção de imagens, que atraem a atenção de cada vez mais leitores.

Wainberg (2005, p. 29) levanta algumas questões relacionadas ao “consumo” desse tipo de leitura: uma sobre a relação entre a ficção e a realidade, “quem imita quem, afinal?”, ou seja, tem a literatura da violência maior influência sobre a realidade, ou a realidade maior influência sobre a literatura? Há obras visivelmente influenciadas por acontecimentos como

os atentados terroristas de 11 de setembro nos Estados Unidos, bem como se pode lembrar de obras anteriores aos atentados que já abordavam situações semelhantes, como o caso de *O agente secreto*, de Joseph Conrad, publicado pela primeira vez em 1907. Outra questão levantada por Wainberg (2005, p. 30) é “por que, afinal, há um público para a violência simulada?”. O autor apresenta algumas hipóteses para responder a essa pergunta, como “busca da fantasia, o desafio e a estimulação sensorial, a pontuação proporcionada pelo fascínio do jogo e os efeitos especiais”. São, portanto, leituras motivadas pelo estímulo imagético, sensorial, que provocam medo, ansiedade, excitação, levando a um posicionamento de condenação ou aprovação.

As obras apresentam conflitos ideológicos, sociais e culturais marcados pela violência. Na literatura estrangeira, destacam-se obras sobre o terrorismo proveniente das guerras mundiais, e dos ataques terroristas motivados pelo idealismo religioso oriental. Na literatura brasileira, as obras retratam a violência criminal dos grandes centros urbanos, marcada por sequestros, roubos, assaltos, homicídios, o tráfico de drogas, o crime organizado, e a própria corrupção da polícia, o que Wainberg chama de “guerrilha não-ideológica”.

*O matador* e *Inferno* são obras que apresentam dois contextos distintos, a primeira, *O matador*, retrata a violência na cidade São Paulo; a segunda, *Inferno*, a favela e o tráfico de drogas no Rio de Janeiro.

### **Traços de normalização**

Comentaremos questões pertinentes ao comportamento tradutório de Clifford Landers que nos permitiram verificar suas tomadas de decisão diante da complexidade enfrentada ao lidar com a tradução literária. Para tanto, analisamos traços de normalização presentes nas duas obras traduzidas, tomando como base estudos de Baker (1996) e de Scott (1998) acerca das características do texto traduzido (TT).

Os traços de normalização estudados por Scott (1998) representam ações do tradutor que tornam o texto mais fluente para o leitor de chegada, podendo ser conscientes ou inconscientes, ou seja, podem ser estratégias tomadas para facilitar a leitura, ou ocorrer inconscientemente como resultado das imposições do ato tradutório em si. Por isso, é importante considerarmos aspectos pertinentes a toda e qualquer tradução, tais como a produção de textos correntes no sistema linguístico e cultural receptor, questões contextuais e socioculturais, o público-alvo e suas necessidades, entre outros fatores que influenciam de modo direto o produto final.

### **Omissões nos textos traduzidos**

O tradutor recorre a omissões nos textos traduzidos, o que se pode explicar, entre outros motivos, para evitar as repetições características da escrita de Patrícia Melo, que, muitas vezes, se tornam redundantes. Para Scott, “a omissão no texto traduzido é normalmente gerada por um traço sistêmico, no qual o tradutor não encontra correspondência para uma palavra isolada ou para uma expressão, e resolve o problema com a omissão” (1998, p. 169).<sup>5</sup>

---

<sup>5</sup> No original: *Omission in the translated text is often triggered by a systemic feature, in that the translator finds no match for a single word or an idiom and resolves the problem by omission* (SCOTT, 1998, p. 169).

Observemos os exemplos abaixo, nos quais podemos perceber a ocorrência de omissões:

- (1) A lista dos traficantes encarcerados no Presídio Padre Moraes era grande, e José Luís fez questão de conhecer o nome dos líderes antes de se encontrar com Nobre, antigo chefe do **morro** do Berimbau. (MELO, 2000, p. 234)
- (2) The list of traffickers incarcerated at Padre Moraes penitentiary was long, and José Luis made a point of learning the names of the leaders before meeting Noble, the one-time boss of Berimbau. (MELO, 2002, p. 196)

No caso acima, o tradutor omite “morro”, por estar claro tratar-se do mesmo pelo contexto. No fragmento abaixo também pode-se observar omissões,

- (3) Fez força para roubar para o outro eu, o eu-pai, mas logo descobriu que há um eu imperativo dentro dos nossos eus, um eu preocupado apenas com os próprios interesses, comodista, um eu que rouba, vence e não percebe a chegada da polícia na **favela**.

**Pa pa rá pa pa.** Quando Reizinho ouviu os tiros, já era tarde. Não adiantava mais sinalizar. Porra. Puxou a pipa, indeciso, devia voltar para casa? (MELO, 2000, p. 14)

- (4) He made an effort to cheat on behalf of the other I, the father-I, but he soon discovered that there is a primary I inside our Is, an I concerned only with its own interests, self-indulgent, an I that cheats, conquers, and doesn't notice the arrival of the police.

When Kingie heard the shots, it was already too late. It did no good to signal. Goddamn. He reeled in the kite, indecisive, should he go back home? (MELO, 2002, p. 5)

O tradutor omitiu favela, por estar claro tratar-se da mesma no contexto, e omitiu os sons dos tiros. Ao longo da obra, pode-se perceber que o tradutor prefere, muitas vezes, omitir os vocábulos que representam os sons, recurso abundante no original.

Em *O matador*, também pode-se perceber as omissões de redundâncias do original, como se pode observar no trecho abaixo,

- (5) Érica pediu para segurar, apontou o cano do **revólver** para mim. Tem bala?, ela perguntou. Negão riu. Eu ri. Ela apertou o gatilho, mirando minha cabeça. Paguei, fomos embora. (MELO, 1995, p. 46)
- (6) Érica asked to hold it, then pointed the barrel at me. Have you got a bullet? she asked. Buck laughed. I laughed. She pulled the trigger, aiming at my head. I paid, and we left. (MELO, 1998, p. 44)

No fragmento acima, o tradutor omite a palavra “revólver”, provavelmente por ser um vocábulo muito utilizado por Patrícia Melo, repetido muitas vezes.

### **Adições no texto traduzido**

Se muitas vezes o tradutor omite vocábulos para evitar as repetições e redundâncias, há ocasiões em que opta por acrescentar informações ao texto de chegada para facilitar a leitura do texto traduzido. É o que ocorre no fragmento extraído da obra *Inferno* e sua respectiva tradução, abaixo:

- (7) Vem cá babaca. Reizinho se aproximou. Miltão tirou um **revólver** da cintura, encostou o cano da arma na palma da mão do garoto e detonou. (MELO, 2000, p. 17)
- (8) C'mere, asshole. Kingie approached him. Miltão took a **38 revolver** from his belt, placed the gun barrel against the palm of the boy's hand and fired. (MELO, 2002, p. 8)

O tradutor, conhecendo o contexto do original mais detalhadamente, opta por oferecer ao leitor de chegada o detalhe sobre o modelo de arma usada para a punição do olheiro da favela quando este falha em sua função de avisar os traficantes sobre a presença da polícia. Observa-se, no mesmo fragmento, que o tradutor aproxima-se bastante do original ao utilizar a contração *c'mere* para apresentar uma linguagem bastante informal e dialetal.

Outro recurso utilizado pelo tradutor para evitar os vocábulos onomatopéicos é a adição de orações, explicando as ações que produzem os sons no original. Vejamos, abaixo, um trecho extraído de *Inferno*:

- (9) Suzana sempre dizia que sua cor preferida era amarelo, lembrou José Luís, observando os detalhes da fotografia, enquanto Marta contava como descobrira que Denilson, o novo namorado de Suzana, era um cana, e não um gerente de supermercado, ploc, como ela andara espalhando por lá, o cara era um meganha, um escroto, ploc, que desapareceu da favela assim, **pluft**. (MELO, 2000, p. 321)
- (10) Suzana always said that yellow was her favorite color, José Luís remembered, observing the details in the photograph, while Marta told how they'd discovered that Denilson, Suzana's new boyfriend, was a cop, not a supermarket manager, pop, as she had been telling everybody. The guy was an undercover cop, a piece of shit, pop, who disappeared from the favela like that, she said, **snapping her fingers**. (MELO, 2002, p. 270)

O excesso de onomatopéias pode dificultar a leitura para o leitor de chegada. No original, esses efeitos são mais familiares aos leitores da cultura de partida. Os leitores da autora têm um perfil diferenciado, provavelmente buscam uma leitura mais imagética, rápida, e com efeitos semelhantes aos da imagem televisiva.

Pode-se observar uma preferência do tradutor em explicitar os tipos de armas utilizadas na obra original, por meio do acréscimo de vocábulos específicos, como os nomes das armas. Vejamos, abaixo, no fragmento extraído de *Inferno*,

- (11) Reizinho ouvira dizer que alguns olheiros sabiam reconhecer as **armas** de combate apenas pelos disparos, AR-15, Uzi, M-16, HK-47, **armas** que chegavam a dar quinze tiros por segundo e pelas quais se pagavam até sete mil dólares, e que além de matar, estilhaçavam o inimigo. (MELO, 2000, p. 14)
- (12) Kingie had heard that some of the lookouts could recognize combat **weapons** just from their sound, **American AR-15s, South Korean Daewoos, Russian AK-47s, guns** that shot fifteen rounds a second and cost as much as seven thousand dollars and didn't just kill the enemy but ripped him to shreds. (MELO, 2002, p. 5)

Como se pode observar, Landers acrescenta até a origem das armas, sendo essas informações desconhecidas no original, percebe-se, de certa forma, a intervenção criativa do tradutor, que reforça as características de armamento ao contexto original, como em um efeito, de certa forma, até estereotipado. O tradutor pode aproximar o texto de chegada ao de guerra propriamente dita, pela valorização do arsenal utilizados pelos traficantes.

## Diferenças relacionadas a imprecisões de expressões

São observadas, nas duas traduções, tendências de facilitar a leitura de expressões imprecisas para a cultura de chegada, tais como nomes de produtos, como se pode observar no fragmento abaixo, extraído do par de obras *Inferno-Inferno*

- (13) Reizinho caminhou pelo **calçadão**, pensou em comprar um **Chica Bom**, desistiu. O tênis. Desistiu. O boné. Desistiu. Andou para lá e para cá, não comprou nada, foi até o calçadão, tanta gente na praia. Garotos jogando vôlei, um punhado de babás conversando, bebês, vento agradável. O Rio de Janeiro era uma cidade bonita de verdade. (MELO, 2000, p. 24)
- (14) Kingie moved along the **calçadão**, the walkway along the beach, thought about buying **an ice cream**, changed his mind. Sneakers. Changed his mind. The cap. Changed his mind. He wandered around, bought nothing, went back to the **walkway**, so many people at the beach. Guys playing volleyball, a group of nannies talking, babies, a pleasant breeze. Rio de Janeiro really was a pretty city. (MELO, 2002, p. 14)

No original, Patrícia usa o nome do sorvete, Chica Bom, recurso muito utilizado pela autora, que destaca marcas de produtos, evidenciando um aspecto da sociedade capitalista retratada na obra, marcada pela persuasão midiática. Na tradução, Landers suaviza essa massificação utilizando um nome genérico para o produto, *ice cream*, ao invés de adaptar para um nome de produto de impacto semelhante ao do original, ou recorrer à explicitação ou empréstimo.

Neste mesmo fragmento, pode-se observar o empréstimo para a tradução de “calçadão”, em um primeiro momento, seguido da explicitação *walkway along the beach*, e, posteriormente, o uso do correspondente próximo *walkway*, o qual fora apresentado pela explicitação. Percebe-se que o tradutor preocupa-se em facilitar a leitura ao construir o cenário urbano específico da cidade do Rio de Janeiro.

## Mudanças na estrutura

Por vezes o tradutor opta por mudar uma estrutura mais complexa por uma mais linear, simples ou convencional. No trecho abaixo, extraído do par de obras *O matador-The killer*, Landers muda a posição de um sintagma nominal dentro da frase:

- (15) Meu primo segurou minhas mãos, **na calçada**: o que há, cara? É só um carburador. É só um carro. É o Marcão, teu amigo. Ficou maluco? (MELO, 1995, p. 36)
- (16) **On the sidewalk**, my cousin gripped both my hands in his: What’s wrong with you, man? It’s just a carburetor. It’s just a car. That’s Marcão, your friend. Have you gone crazy? (MELO, 1998, p. 32)

A posição do sintagma na frase pode indicar a ordem de relevância das informações, o que vem primeiro é a informação mais importante, sendo assim, no original, a autora valoriza mais a ação do primo do narrador, enquanto o tradutor opta por organizar a cena, colocando em primeiro lugar a organização do espaço, mostrando o lugar onde os personagens se encontravam quando um segurou as mãos do outro. Esse tipo mudança evidencia uma estratégia facilitadora na tradução.

## Mudança de linguagem mais informal para mais formal

Há mudanças na tradução que tornam a linguagem informal em mais formal. Nos fragmentos abaixo extraídos do par de obras *O matador* → *The Killer*, percebe-se esse tipo de tendência facilitadora,

- (17) Disseram, os babacas, um erro jurídico é inaceitável. Erro médico, eles aceitam. E os infelizes que eles matam no hospital? Dizem que a pobreza geral do país é que gera a violência. Gera violência, gera poluição, gera doença, gera **o diabo**, mas não gera esses estupradores filhos da puta, isso não gera. Não é verdade? (MELO, 1995, p. 32)
- (18) The assholes said that a legal error is unacceptable. A medical error is fine with them. What about the poor devils they kill in hospitals? They say that it's the country's overall poverty that generates violence. Generates violence, generates pollution, generates disease, generates **whatever**, but it doesn't generate rapist sons of bitches. It doesn't generate that. Right? (MELO, 1998, p. 28)

Para a tradução da expressão “o diabo”, Landers opta pela expressão *whatever*, mais genérica do que a expressão bem coloquial utilizada no original. Também pode-se observar, ao longo da tradução, que o uso de palavrões do original é bastante atenuado pelo tradutor.

## Considerações finais

Por meio dos dados obtidos pode-se perceber que o tradutor apresenta grande autonomia em relação ao trabalho com as obras de Patrícia Melo, recorrendo a mudanças que, consciente ou inconscientemente, facilitam a leitura para os leitores de chegada.

Percebe-se, também, que o tradutor, muitas vezes, diminui marcas características do original, no caso de substituir nomes de marcas de produtos pela descrição dos mesmos, ao suavizar o uso de palavrões, ao mudar a organização dos elementos da frase, e ao omitir nomes próprios que são desconhecidos do leitor de chegada.

Com este trabalho, é possível refletir sobre a tradução de obras brasileiras que têm boa repercussão no exterior, e sobre como o tradutor lida com as peculiaridades dessa forma literária, que tem recebido maior atenção de pesquisadores recentemente.

Pretende-se demonstrar que a Linguística de *Corpus*, em intersecção com os Estudos da Tradução, pode contribuir muito para as investigações sobre a tradução literária e, espera-se que o material fornecido instigue outros pesquisadores a seguir caminho nesse domínio de pesquisa científica.

## REFERÊNCIAS DOS *CORPORA* DE ESTUDO

MELO, P. *O matador*. São Paulo: Companhia das Letras, 1995.

\_\_\_\_\_. *The Killer*. Tradução Clifford Landers. London: Bloomsbury, 1997.

\_\_\_\_\_. *Inferno*. São Paulo: Companhia das Letras, 2001.

\_\_\_\_\_. *Inferno*. Tradução Clifford Landers. London: Bloomsbury, 2003.

## REFERÊNCIAS

BAKER, M. Corpus Linguistic and translation studies: implications and applications. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (Ed.) *Text and Technology*: In honour of John Sinclair. Amsterdam: John Benjamins, 1993, p. 233-250.

\_\_\_\_\_. Corpora in translation studies: an overview and some suggestions for future research. *Target*, v. 7, n. 2, p. 223-243, 1995.

\_\_\_\_\_. Corpus-based translation studies: the challenges that lie ahead. In: SOMERS, H. *Terminology, LSP and Translation Studies in Language Engineering*: In Honour of Juan C. Sager. Amsterdam: John Benjamins, 1996, p. 175-186.

BERBER SARDINHA, T. *Linguística de Corpus*. Barueri: Manole, 2004.

CAMARGO, D. C. *Metodologia de pesquisa em tradução e linguística de corpus*. São Paulo: Cultura Acadêmica; São José do Rio Preto, SP: Laboratório Editorial do IBILCE, UNESP, 2007.

\_\_\_\_\_. *Padrões de Estilo de Tradutores*: um estudo de semelhanças e diferenças em corpora de traduções literárias, especializadas e juramentadas. 2005. 512 f. Tese (Livre-Docência em Estudos da Tradução) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto.

SCOTT, M. N. *Normalization and Reader's Expectations*: A Study of Literary Translation with Reference to Lisspector's *A Hora da Estrela*. 1998. 318 f. Tese (Doutorado em Tradução) – Universidade de Liverpool. Liverpool.

WAINBERG, Jacques A. *Mídia e terror*: comunicação e violência política. São Paulo: Paulus, 2005

# Instrumentos e atividades realizadas para a coleta de um *corpus* de aprendizes em língua inglesa para integrar o Br-ICLE (*Brazilian Portuguese Sub-corpus of ICLE*)

(Tools and activities used in data collection for a *corpus* composed of texts written by students of English language to be added to Br-ICLE (*Brazilian Portuguese Sub-Corpus of ICLE*))

Paula Tavares Pinto Paiva<sup>1</sup>

<sup>1</sup>Instituto de Biociências, Letras e Ciências Exatas –  
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)

paula@ibilce.unesp.br; paulapaivaibilce@gmail.com

**Abstract:** This paper presents and discusses initiatives taken in a public university in the state of São Paulo in order to collect and organize a corpus composed of argumentative texts to be part of Br-ICLE – a subcorpus of ICLE<sup>1</sup> – which is composed of texts produced by Brazilian students enrolled in the courses of Arts and Languages and Translation. The discussion is based on a three-year project in which we could observe underuse and overuse features in the texts produced by Brazilian undergraduate students.

**Keywords:** corpus linguistics; learner corpus; underuse and overuse features of English.

**Resumo:** O objetivo deste trabalho é apresentar e discutir as iniciativas tomadas em uma universidade pública do estado de São Paulo com o intuito de coletar e organizar um corpus de textos argumentativos que ajudará a compor o Br-ICLE – o subcorpus do ICLE, composto por textos em língua inglesa produzidos por alunos brasileiros dos cursos de Licenciatura em Letras e Bacharelado em Letras com Habilitação em Tradução. A discussão tomará como base o desenvolvimento do projeto durante três anos de ocorrência, no qual puderam-se observar características de subuso e sobreuso nos textos dos alunos brasileiros.

**Palavras-chave:** Linguística de corpus; corpora de aprendizes; Características de subuso e sobreuso da língua inglesa.

## Introdução

O objetivo deste trabalho é apresentar e discutir as iniciativas tomadas em uma universidade pública do estado de São Paulo com o intuito de coletar e organizar um *corpus* de textos argumentativos que ajudará a compor o Br-ICLE<sup>2</sup> – o *subcorpus* de textos em língua inglesa produzidos por alunos brasileiros. O projeto geral de compilação do Br-ICLE conta com a participação de equipes de pesquisa de várias universidades brasileiras, dentre elas, PUC-SP, USP, UERJ, UFMG, Unesp, Unicsul e UniCastelo.<sup>3</sup> Trata-se de um acordo feito com o Centro de Linguística de *Corpus* (doravante LC) de Inglês, da Universidade de Louvain, Bélgica, que tem coletado *corpora* de aprendizes de diversos países e visa a desenvolver estudos que descrevam a língua inglesa produzida por aprendizes de diferentes línguas maternas. No caso específico deste trabalho, objetivamos compilar um *corpus* que

1 Disponível em: < <http://www.uclouvain.be/en-cecl-icle.html>>. Acesso em: 13 fev. 2012.

2 Disponível em: < <http://www2.lael.pucsp.br/corpora/bricle/index.htm>>. Acesso em: 13 fev. 2012.

3 Para mais informações, acessar < <http://www2.lael.pucsp.br/corpora/bricle/team.htm>>. Acesso em: 13 fev. 2012.

possibilite a observação de características linguísticas da língua-alvo que são empregadas mais frequentemente (*sobreuso*) ou menos frequentemente (*subuso*) pelos aprendizes de língua inglesa dos cursos de Licenciatura em Letras e Letras com Habilitação em Tradutor, e compará-las à produção de falantes nativos de língua inglesa. A importância da realização de uma pesquisa como esta encontra-se em consonância à opinião de Biber *et al.* (1998), uma vez que os autores afirmam que um estudo sobre os desvios encontrados nas produções de aprendizes aumenta nossa compreensão sobre a aquisição de uma língua estrangeira, fornece dados para outras perspectivas relacionadas aos “erros” (ex. interlíngua e variedades não padrão da língua-alvo), e fornece evidências para decisões instrucionais. Uma das hipóteses a ser observada a fim de justificar a presença de traços de sobreuso<sup>4</sup> e/ou subuso seria a influência da língua materna dos aprendizes em suas produções escritas.

Para que o estudo fosse realizado, iniciamos, ao final do ano de 2009, a coleta dos textos produzidos pelos alunos dos cursos supracitados, que são enviados a uma das pesquisadoras responsáveis pela coleta e organização do Br-ICLE, Denise Delegá Lúcio. Importante ressaltar que os textos são enviados sem que nenhuma correção tenha sido feita, a fim de manter as características da escrita original dos alunos, entretanto, como a coleta é feita por nós no primeiro e no segundo semestre, descreveremos na seção de materiais e métodos as atividades de caráter didático-pedagógica realizadas por nós para que a escrita dos alunos seja melhorada com o passar do ano.

## Fundamentação teórica

A observação e a descrição da linguagem em uso em diferentes contextos, como no de ensino e aprendizagem de línguas, têm sido favorecidas pela LC em virtude de sua abordagem teórico-metodológica apresentar um caráter interdisciplinar. Berber Sardinha (2004) destaca as vantagens que a LC oferece tanto para o professor quanto para o aluno, por permitir o acesso à tecnologia computacional e à linguagem utilizada nos textos contidos em um determinado *corpus*.

As aplicações da LC se fazem sentir tanto na área da Lexicografia, quanto nos estudos sistemáticos do uso da língua, em trabalhos dos Estudos da Tradução, Linguística Aplicada e em Processamento de Linguagem Natural. Tal pressuposto deve-se à proposta de Sinclair (1978, apud BERBER SARDINHA, 2004) ao afirmar que a língua deveria ser estudada a partir de exemplos reais de uso e não somente a partir de textos criados com propósitos de exemplificação. A interdisciplinaridade constitutiva dessa área de estudos tem possibilitado a troca de experiências e uma real colaboração entre profissionais de diferentes áreas do conhecimento.

Berber Sardinha discute em sua obra o papel da LC uma vez que esta abordagem:

Ocupa-se da coleta e exploração de *corpora*, ou conjuntos de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador. (BERBER SARDINHA, 2004, p. 3)

4 De acordo com Delegá-Lucio (2006, p. 2), “O sobreuso de palavras diz respeito à utilização de uma mesma palavra mais vezes do que um falante nativo normalmente o faria”. O subuso seria o oposto, ou seja, um uso menos frequente de palavras que são comumente utilizadas pelos falantes nativos.

Importante destacarmos que o estudo ora proposto pretende descrever a linguagem do aprendiz de uma língua estrangeira, sendo assim, o *corpus* adequado para tal propósito é o *corpus* de aprendiz, cuja compilação e propósito é discutido por Delegá-Lucio ao afirmar que:

Um corpus de aprendiz é aquele formado por textos naturais falados ou escritos por estudantes de uma língua estrangeira, que atenda a critérios que possibilitem seu estudo e que tenha sido coletado de modo que possa ser lido por computador. Os textos de um corpus de aprendiz são considerados naturais porque são produzidos por estudantes sem controle do que escrevem, ou seja, nenhum ponto gramatical (por exemplo) específico deve ser elicitado. Há, no entanto, critérios para a coleta desse corpus de acordo com o tipo de texto que se quer representar. (DELEGÁ-LUCIO, 2006, p. 21)

A diferença entre o *corpus* de aprendiz e o *corpus* nativo se dá pelo fato de os textos serem escritos por não nativos de uma língua estrangeira. Ao se utilizar o *corpus* de aprendiz, de acordo com a autora, pode-se mostrar aos alunos, por exemplo, o uso excessivo que fazem dos mesmos adjetivos, como mostra seu próprio estudo sobre relexicalização de adjetivos (DELEGÁ-LUCIO, 2006). Dessa forma, acredita-se que os alunos serão conscientizados sobre este problema. Ao mesmo tempo, ao se usar materiais informados por um *corpus* de falantes nativos, pode-se expor aos alunos às variedades de adjetivos que se associam aos substantivos para os quais há adjetivos sobreusados, ou seja, um corpus de aprendiz e um de textos de falantes nativos podem ser usados como instrumentos de aprendizagem de língua estrangeira.

Leech (1998) também destaca algumas questões que podem ser levantadas a partir da observação de *corpora* de aprendizes, dentre elas:

- a) Quais características linguísticas da língua-alvo são empregadas com mais (sobreuso) ou menos (subuso) frequência em comparação com falantes nativos?
- b) Qual é a extensão da influência da língua nativa (transferência) na produção dos aprendizes?
- c) Em que áreas eles tendem a usar estratégias de evitação deixando de explorar a fundo o potencial da língua-alvo?
- d) Em que áreas eles tendem a demonstrar desempenho nativo ou não-nativo?
- e) Quais são as áreas nas quais os aprendizes de um dado país parecem necessitar de mais ajuda para desenvolver sua produção na língua-alvo? (LEECH, 1998, p. xiv)

*Corpora* de aprendizes são constituídos por textos de falantes não-nativos, os quais são utilizados para o estudo da linguagem que produzem. De acordo com Haan (1992), o tamanho ideal de um *corpus* de aprendizes vai depender do tipo de pesquisa a ser realizada. Ainda segundo o autor, um corpus de 20 mil palavras pode ser suficiente para se realizar um estudo. *Corpora* de aprendizes que variam de 20 a 200 mil palavras costumam ser mais específicos em relação a *corpora* maiores, no que se refere aos seus tópicos e gêneros discursivos. Para Granger (1998), 200 mil palavras foi o número ideal que se determinou para a constituição do ICLE (*International Corpus of Learner English*). De acordo com a autora, *corpora* de aprendizes maiores seriam mais destinados a pesquisadores que visam, por exemplo, a compilação de dicionários para aprendizes de uma língua estrangeira.

Tomando como base as questões destacadas acima, este estudo vem sendo guiado pelos dados fornecidos pelos *corpora* de aprendizes uma vez que procura identificar, com mais exatidão, em quais áreas os aprendizes brasileiros costumam ter mais dificuldade. Adicionalmente, citamos um estudo diacrônico baseado em *corpus*, e discutido por Biber *et al.* (1998), no qual foram evidenciados desvios de uso da língua inglesa utilizada por alunos estrangeiros em relação à concordância verbo-nominal e marcação do plural. Após quatro anos de acompanhamento desse estudo, os autores verificaram que houve uma diminuição considerável de erros relacionados à concordância de sujeito e verbo nas orações produzidas pelos aprendizes. No entanto, os erros ligados à morfologia verbo-nominal ocorriam mais frequentemente do que os erros de concordância. Mais recentemente, observamos o sobreuso do artigo definido “the” (ORENHA *et al.*, 2007), que poderia ser considerada uma das evidências observadas nas produções em língua inglesa de alunos brasileiros e que, possivelmente, é influenciada pela língua materna dos aprendizes. Por esse motivo, temos observado tais características tomando como base estudos sobre a aquisição e aprendizagem de língua estrangeira que tratam do assunto (SHUMANN, 1992; MITCHELL; MYLES, 1998).

Ainda sobre estudos a partir de *corpora* de aprendizes, baseamo-nos em diversos trabalhos realizados no exterior e no Brasil (GRANGER 1993, 1998, 2009; BERBER SARDINHA, 2004; DELEGÁ-LUCIO, 2006; MEUNIER, 2011; DUTRA; SILEIRO, 2010, 2012).

Neste ponto, cabe destacar que, em termos de nomenclatura, embora alguns pesquisadores prefiram apresentar a diferenciação entre “*corpus-based*” (*corpus* usado para provar uma teoria ou posição a priori) e “*corpus-driven*” (*corpus* utilizado para permitir contraprova a posições iniciais assumidas pelos pesquisadores em geral) Beber Sardinha (1999) enfatiza que o primeiro termo, hoje em dia, tem sido mais utilizado para se referir às pesquisas a partir de *corpora* eletrônicos.

## **Metodologia e *corpora***

Os alunos selecionados para participar da coleta do Br-ICLE são regularmente matriculados nas disciplinas de Língua Inglesa III e IV dos terceiro e quarto anos, uma vez que os textos devem ser escritos por aprendizes em nível avançado. No início de cada ano letivo, os alunos conhecem o projeto Br-ICLE por meio de uma apresentação na qual eles têm contato com a terminologia específica da Linguística de *Corpus*, ou seja, os alunos que não conhecem esta abordagem teórico-metodológica passam a conhecer a definição de *corpus* e conhecer sua bases teóricas.

Em seguida, eles preenchem o Learner Profile (anexo) que descreverá seus perfis como aprendizes de língua inglesa. Em nossa universidade, os alunos fazem as redações em casa e utilizam todas as ferramentas que quiserem como dicionários mono e bilíngues, glossários e internet. Depois de preenchidos e assinados os formulários, os alunos escolhem dois temas para serem trabalhados no primeiro semestre, e dois para o segundo. Para tanto, eles recebem a lista dos tópicos a serem utilizados como tema para suas composições, conforme sugerido no projeto de Granger (1993). Os requisitos para a escritura das composições são: a) conter no mínimo 500 palavras e no máximo 1000; b) o aluno deve ter nível avançado de inglês; c) o aluno não pode ser nativo de língua inglesa.

Os temas sugeridos para as redações do Br-ICLE são:

- (1) Crime does not pay
- (2) The prison system is outdated. No civilised society should punish its criminals: it should rehabilitate them.
- (3) All armies should consist entirely of professional soldiers: there is no value in a system of military service
- (4) Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value
- (5) A man/woman's financial reward should be commensurate with their contribution to the society they live in
- (6) In the 19th century, Victor Hugo said: "How sad it is to think that nature is calling out but humanity refuses to pay heed. "Do you think it is still true nowadays?"
- (7) Some people say that in our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion?
- (8) In the words of the old song "Money is the root of all evil"
- (9) The Gulf War has shown us that it is still a great thing to fight for one's country.
- (10) Feminists have done more harm to the cause of women than good.
- (11) In his novel *Animal Farm*, George Orwell wrote "All men are equal: but some are more equal than others" How true is this today?
- (12) The role of censorship in Western society.
- (13) Marx once said that religion was the opium of the masses. If he was alive at the end of the 20th century, he would replace religion with television.

A seguir, apresentaremos a composição do *corpus* de aprendizes de nossa universidade conforme sua coleta, iniciado em 2010. Também apresentaremos algumas amostras das palavras mais frequentes, palavras-chave<sup>5</sup> e linhas de concordância,<sup>6</sup> que são analisadas por nós e pelos alunos durante as aulas de língua inglesa com o auxílio de *corpora* computadorizados.

### **Composição do *Corpus* em 2010 e algumas análises com o AntConc<sup>7</sup>**

Abaixo mostramos a tabela 1 com os dez itens mais frequentes no primeiro sub-corpus, composto por textos de alunos do terceiro ano do curso de Licenciatura em Letras:

---

5 Palavras-chave são estatisticamente mais altas e, portanto, relevantes no corpus de estudo em comparação a um corpus de língua geral.

6 Linhas de concordância são todas as linhas de um corpus de estudo que contenham uma palavra de busca inserida em seu contexto de uso.

7 Disponível em < <http://www.antlab.sci.waseda.ac.jp/index.html>>. Acesso em: 15 fev. 2012

**Tabela 1: Número total de formas e itens do *corpus* de 2010 e os dez itens mais frequentes**

Total No. of Word Types: 4635		
Total No. of Word Tokens: 8287		
1	4398	the
2	2654	to
3	2514	of
4	2362	and
5	1810	a
6	1740	is
7	1618	in
8	1420	that
9	890	it
10	886	are

Além das palavras que são, em sua maior parte, palavras funcionais, temos o número total de formas do *corpus*, isto é, o número total de palavras não repetidas, (4.635 formas), assim como o número total de palavras, chamadas aqui de itens (8.287).

#### **Composição do *Corpus* em 2011 e algumas análises com o AntConc.**

Similarmente ao *corpus* descrito anteriormente, no ano de 2011, coletamos as redações dos alunos do terceiro ano do curso de Licenciatura em Letras, conforme demonstrado na Tabela 2:

**Tabela 2: Número total de formas e itens do *corpus* de 2011 e os dez itens mais frequentes**

Total No. of Word Types: 4639		
Total No. of Word Tokens: 8752		
1	4399	the
2	2655	to
3	2517	of
4	2363	and
5	1811	a
6	1741	is
7	1619	in
8	1421	that
9	891	it
10	887	are

Neste ano, o número de formas (4.639) e itens (8.752) foi semelhante ao ano anterior.

#### **Composição do *Corpus* em 2012 e algumas análises com o AntConc**

Diferentemente dos *corpora* anteriores, como mudamos de *campus* e de curso, desta vez, a turma que atenderia ao perfil requisitado no Br-ICLE era o quarto ano do curso de Bacharelado em Letras com Habilitação em Tradutor. Abaixo apresentamos a tabela com os dados gerados pelo *corpus* composto por sua produção:

**Tabela 3: Número total de formas e itens do *corpus* de 2012 e os dez itens mais frequentes**

Total No. of Word Types: 4639		
Total No. of Word Tokens: 9217		
1	4400	the
2	2656	to
3	2520	of
4	2364	and
5	1812	a
6	1742	is
7	1620	in
8	1422	that
9	892	it
10	888	are

Embora o número de itens tenha sido mais alto no *corpus* composto por alunos de Tradução (9.217), coincidentemente, o número de formas foi o mesmo que o empregado no ano anterior pelos alunos de Letras (4639). O próximo passo, ao analisarmos os *corpora* detalhadamente, será o de analisar se o emprego lexical também foi semelhante.

Nesta fase do estudo temos trabalhado com os *corpora* separadamente, ou seja, com as listas geradas a partir das composições coletadas nos anos de 2010, 2011 e 2012. Entretanto, ao juntarmos os três subcorpora teremos um número total de itens de 26.256 itens, que seria considerado um tamanho aceitável para a realização de uma pesquisa com *corpus* de aprendiz, de acordo com o pesquisador Haan (1992), citado anteriormente.

### **Participação de Assistentes de Língua Inglesa (ETAs) na análise manual do Br-ICLE**

No ano de 2010, recebemos em nossa universidade dois assistentes de língua inglesa (*English Teaching Assistants* – ETAs) pelo programa da Fulbright/Capes. Assim como o programa de leitores estrangeiros que auxiliam os professores de línguas estrangeiras no Departamento de Letras Modernas, os ETAs auxiliaram os professores da área de língua inglesa desempenhando várias funções, dentre as quais, destacamos a assistência na correção das composições dos alunos. Os ETAs também desempenharam um papel fundamental ao discutirem, nas aulas de língua inglesa, quais são as características mais frequentes nos textos dos alunos brasileiros de nossa universidade, que não soam naturais em língua inglesa. Nosso estudo nos mostrou, em primeiro lugar, que do total de treze temas a serem escolhidos pelos alunos brasileiros de nossa universidade, a maior parte tem optado pelos seguintes títulos:

- a) Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value;
- b) Some people say that in our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion?;
- c) In the words of the old song “Money is the root of all evil”;
- d) In his novel *Animal Farm*, George Orwell wrote “All men are equal: but some are more equal than others”. How true is this today?

Interessante notar que esses quatro temas foram repetidamente escolhidos pelos alunos no primeiro semestre dos três anos em que o projeto tem sido desenvolvido. Com o intuito de aumentar a variedade de temas utilizados e, por consequência, aumentar a diversidade lexical empregada em suas composições, pedimos aos ETAs que debatessem questões culturais da realidade estadunidense como, por exemplo, o sistema penitenciário e a pena de morte nos Estados Unidos, a seleção de jovens durante o alistamento militar, a participação dos Estados Unidos na Guerra do Golfo e outros temas mais gerais como o movimento feminista nos dias de hoje, a censura e as diferentes religiões que coexistem atualmente nos Estados Unidos.

Sabemos que a prévia discussão desses temas pode ter influenciado a escrita dos alunos em língua inglesa, assim como, tê-los incentivado a fazer uma busca mais refinada do vocabulário a ser empregado em suas redações, o que poderia de certa forma, invalidar os resultados finais do Br-ICLE. Entretanto, percebemos que as redações no início do projeto não eram bem escritas e que os temas se repetiam. A entrada dos ETAs no projeto foi justamente para que isso não acontecesse. Acreditamos que os alunos brasileiros puderam entrar em contato, de uma forma mais dinâmica e palpável, com a realidade cultural estadunidense por meio dos relatos e discussões com jovens que a vivenciaram e que tiveram seus familiares envolvidos em tais questões. Esse fato aguçou a curiosidade dos aprendizes brasileiros para a pesquisa de um vocabulário mais avançado a ser empregado em suas redações. Acreditamos que essa realidade acabou levando-os a, pelo menos, pensar sobre os diferentes títulos e, dessa forma, enriquecer seu léxico na língua estrangeira que estudavam. A nosso ver, essa “intervenção” indireta foi muito positiva para o desenvolvimento de um aprendizado mais autônomo, o que era o objetivo primordial de nossa investigação, ou seja, por meio da observação de “desvios”, características de sobreuso e subuso, pretendíamos direcionar o ensino da língua inglesa a fim de suprir lacunas do conteúdo programático das disciplinas de Língua Inglesa III e IV.

## Discussão

Até o momento, em uma breve análise das linhas de concordância, observamos o sobreuso do artigo definido “the” que, conforme destacado anteriormente, pode ter sido influenciado pela língua materna dos aprendizes, como, por exemplo, o uso do artigo com certas entidades geográficas que, na língua inglesa, não deveria ter sido usado.

Também observamos a influência da língua portuguesa nas colocações verbo-nominais, como em “have time to”, “have classes”, “have an appointment”, “have chances” e “have opportunity to” (“ter tempo para”, “ter aulas”, “ter um compromisso”, “ter chances” e “ter oportunidade para”). Este fato não é um problema, mas percebemos o subuso de outras possibilidades de combinatórias como em “have a laugh” e “have a look at”, que são altamente empregadas na língua inglesa por seus nativos.

Na análise manual, ou seja, sem o auxílio de ferramentas computacionais, observamos a presença de “dangling modifiers”, isto é, o emprego do particípio em posições erradas na oração e um desconhecimento do uso da pontuação na língua inglesa. O próximo passo será verificar esses itens no *corpus* geral com os programas de análise lexical.

Como uma possível solução para estes itens, temos apresentado aos alunos, por meio de atividades didáticas, linhas de concordância em *corpora* on-line a partir de textos

produzidos por nativos de língua inglesa que exemplificam o uso dessas estruturas. Dessa forma, os alunos entram em contato com uma diversidade maior de colocações verbo-nominais e com a estrutura da língua inglesa e se tornam cientes de outras possibilidades de combinatórias desconhecidas que são apresentadas junto ao seu contexto de uso.

## Conclusões e encaminhamentos

Uma vez que o projeto de coleta do Br-ICLE, durante três anos, já está completo em nossa universidade, enviaremos todas as composições aos coordenadores do projeto no Brasil. Em seguida, continuaremos as análises do *corpus* compilado com o auxílio de ferramentas computacionais, conforme os passos elencados no trabalho de Dutra e Sileiro (2010) utilizando o *Error Tagger 1.0*, disponível em <[www.corpuslg.org](http://www.corpuslg.org)> e a análise mais aprofundada das palavras-chave com o auxílio do programa *WordSmith Tools 4.0* (SCOTT, 2004). Acreditamos que os dados apontados até o momento e as ações tomadas no sentido de discutir os temas do Br-ICLE e propor exercícios de pesquisa com *corpora* disponibilizados gratuitamente na internet já têm rendido frutos positivos para um aprendizado mais autônomo da língua inglesa.

## REFERÊNCIAS

ANTHONY, L. *AntConc* Disponível em <<http://www.antlab.sci.waseda.ac.jp/software.html>>. Acesso em: 15 jul. 2011

BERBER SARDINHA, A. P. *Linguística de Corpus*. Barueri, SP: Manole, 2004.

BIBER, D.; CONRAD, S.; REPPERN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 1998. 300 p.

DELEGÁ-LUCIO, D. *A relexicalização de adjetivos nas redações de alunos de inglês – um estudo baseado em corpus de aprendiz*. 2006. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) – Pontifícia Universidade Católica de São Paulo. São Paulo.

DUTRA, D. P.; SILERO, R. P. Descobertas linguísticas para pesquisadores e aprendizes: a Linguística de Corpus e o ensino de gramática, 2010. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v. 10, n. 4, 2010.

\_\_\_\_\_. O uso de For: uma análise de itens linguísticos em corpus de aprendizes brasileiros. In: SHEPHERD, T. M. G.; BERBER SARDINHA, T.; PINTO, M. V. (Org.) *Caminhos da linguística de corpus*. Campinas: Mercado de Letras, 2012. p. 325-341.

GRANGER S. The International Corpus of Learner English. In: AARTS, J.; DE HAAN, P.; OOSTDIJK, N. (Org.) *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam: Rodopi, 1993. p. 57-69.

\_\_\_\_\_. *Learner English on Computer*. London and New York: Longman, 1998.

\_\_\_\_\_. The contribution of learner corpora do second language acquisition and foreign language teaching: a critical evaluation. In: ALIJMER, K. (Org.) *Corpora and language teaching*. Studies in Corpus Linguistics 33. Amsterdam: John Benjamins, 2009.

HAAN, P. The Optimum Corpus Sample Size? In: LEITNER, G. (Org.) *New Directions in English Language Corpora*. Berlin: Mouton de Gruyter, 1992. p. 3-19.

LEECH, G. Preface – Learner corpora: what they are and what can be done with them. In: GRANGER S. *Learner English on Computer*. London and New York: Longman, 1998. p. xiv – xx.

MITCHELL, R.; MYLES, F. Functional/Pragmatic Perspectives on Second Language Learning. In: *Second language learning theories*. London: Arnold, 1998.

ORENHA, A.; PAIVA, P. T. P.; CAMARGO, D. C. O uso de corpora de aprendizes no ensino de produção escrita em língua inglesa. *MOARA*, v. 26, p. 281-293, 2007.

SCOTT, Mike. *WordSmith Tools: Version 4*. Oxford: Oxford University Press, 2004.

SHUMANN, J. La adquisición de lenguas segundas: la hipótesis de la pidginización. In: LICERAS, J. M. *La adquisición de las Lenguas Extranjeras*. Madrid: Visar, 1992. p. 123-141.

# ANEXO

## LEARNER PROFILE

=====

Text code: (do not fill in)

**Essay:** .....

Title: .....

Approximate length required: +500 words

Conditions:                    timed ( )    untimed ( )

Examination:                yes ( )        no ( )

Reference tools:             yes ( )        no ( )

What reference tools?

Bilingual dictionary: .....

English monolingual dictionary: .....

Grammar: .....

Other(s): .....

=====

**Surname:** .....    **First names:** .....

**Age:** .....    Male ( )    Female ( )

**Nationality:** .....

Native language: .....

Father's mother tongue: .....

Mother's mother tongue: .....

Language(s) spoken at hom: (if more than one, please give the average % use of each)

.....

**Education:**

Primary school - medium of instruction: .....

Secondary school - medium of instruction: .....

Current studies: .....

Current year of study: .....

Institution: .....

Medium of instruction: .....

English only ( )

Other language(s) ( ) (specify) \_\_\_\_\_

Both ( )

=====

Years of English at school: .....

Years of English at university: .....

**Stay in an English-speaking country:**

Where? .....

When? .....    How long? .....

=====

**Other foreign languages in decreasing order of proficiency:**

1.....

2.....

3.....

4.....

=====

**I hereby give permission for my essay to be used for research purposes**

Date: .....

Signature: .....

# A (não) realização da vogal epentética em Português Arcaico à luz derivacional

(The non-occurrence of the epenthetic vowel in Archaic Portuguese)

Ana Carolina Freitas Gentil Almeida Cangemi<sup>1</sup>

<sup>1</sup> Faculdade de Ciências e Letras de Assis, UNESP – Universidade Estadual Paulista

carolcangemi@gmail.com

**Abstract:** This study has as main aims to discuss the motivation for the non-occurrence of the epenthetic vowel at the beginning of the word in the context of word junction in the first hundred *Cantigas de Santa Maria* (CSM – Afonso X, 1221-1284). Through this work and from a non-explored corpus (especially as for linguistic themes – SNOW, 1987, p. 478-480; MASSINI-CAGLIARI, 2005, p. 24-26), it is intended to contribute to reaching a deeper knowledge of Portuguese prosody history in the period of its recognition as an “independent” language in relation to Latin.

**Keywords:** epenthesis; sandhi processes; archaic portuguese.

**Resumo:** O objetivo desse trabalho consiste em discutir a motivação para a não-ocorrência da vogal epentética em início de palavra no contexto de junção vocabular nas cem primeiras *Cantigas de Santa Maria* (CSM – Afonso X, 1221-1284). Por meio deste trabalho e a partir de um *corpus* ainda pouco explorado (sobretudo quanto a temas linguísticos – cf. SNOW, 1987, p. 478-480; MASSINI-CAGLIARI, 2005, p. 24-26), pretende-se contribuir para o aprofundamento do conhecimento da história da prosódia do português no período de seu reconhecimento como língua “independente” do latim.

**Palavras-chave:** epêntese; sândi; português arcaico.

## Objetivo

O objetivo deste trabalho consiste em discutir a não-ocorrência da vogal epentética em início da segunda palavra no contexto de junção vocabular nas cem primeiras *Cantigas de Santa Maria* (CSM), compiladas a mando de Afonso X, o Sábio (1221-1284). À luz das teorias derivacionais – em especial a Teoria Prosódica – Nespor e Vogel (1986) e Teorias da Sílabas – Selkirk (1982) e Hogg e McCully (1987).

## Corpus: Cantigas de Santa Maria

Um problema metodológico crucial, quando o objetivo é analisar elementos prosódicos de períodos passados das línguas ou de línguas cujos falantes já desapareceram (as chamadas “línguas mortas”), consiste na escolha do tipo de texto escrito remanescente. Essa é um fator decisivo para a viabilidade da pesquisa. Com relação especificamente ao Português Arcaico (PA), sobreviveram textos escritos tanto em prosa quanto em verso. Porém, para o tipo de trabalho que aqui se propõe, a possibilidade de escolha entre material poético e não-poético não se coloca, uma vez que é extremamente difícil, senão impossível, extrair indícios seguros da realização fonética de sequências particulares, com elementos prosódicos como ritmo, entoação, tessitura, etc. Pelo fato de não ter sobrevivido registros

orais da língua da época estudada por nós (final do século XIII), optamos por focalizar nossa atenção nos textos poéticos, uma vez que a escansão das sílabas (metrificação) nos versos e a disposição nestes das sílabas acentuadas e átonas (principalmente a composição da rima poética e a distribuição dos acentos poéticos nos versos) pode indicar elementos prosódicos de forma mais confiável no que tange a sua realização fonética. Portanto, entre os tipos de textos poéticos cultivados na época, esta pesquisa optou por focalizar as CSM.

Segundo Parkinson (1998a), as CSM constituem o monumento literário de maior importância para o período medieval. São consideradas pelo autor uma obra rica, contendo texto, melodia e pintura, sendo uma das fontes mais importantes para a história da métrica e do galego-português antigo. As CSM são uma coleção de 420 cantares em louvor da Virgem Maria compiladas em galego-português por Afonso X.

The Cantigas de Santa Maria is a collection of more than four hundred poems recounting miracles worked through the intercession of the Virgin Mary or songs of praise in her honor. The text of many is illuminated in full-page miniatures. The poems were written in the language of medieval Galicia and Portugal, the medium of expression preferred by the lyric poets of that day. (O'CALLAGHAN, 1998, p. 1)

Tais cantigas chegaram até nós por meio de quatro manuscritos antigos denominados códices:

E: El Escorial, Real Monasterio de san Lorenzo, MS B.I.2 (conhecido como Escorial ou códice dos músicos) – o mais completo de todos;

T: El Escorial, Real Monasterio de san Lorenzo, MS T.I.1 (códice rico ou códice das histórias) – considerado o mais rico em conteúdo artístico (sobretudo iconográfico);

F: Firenze, Biblioteca Nazionale Centrale, Banco Rari, 20 (códice de Florença) – que forma um conjunto com o códice Escorial rico, uma vez que as cantigas que contém completam o códice T;

To: Toledo, Madrid, Biblioteca Nacional, MS 10.069 – o menor e mais antigo de todos, que contém também um índice de cem cantigas.

Segundo Ilari e Basso (2007, p. 21), a língua em 1100, no berço do Estado português, era muito parecida com o galego, por isso a denominação galego-português, ou seja, por apresentar notáveis semelhanças com a língua falada na outra margem do rio Minho. Encontramos na bibliografia especializada essa denominação aplicada à variedade de língua em que se expressou sua manifestação literária: a lírica trovadoresca. No século XIII, o galego-português foi usado como língua da poesia não só por trovadores portugueses como Dom Dinis – rei a partir de 1290 –, mas também por trovadores de outras regiões da Ibéria – por exemplo, Afonso X, o Sábio, rei de Castela.

Massini-Cagliari (2007, p.122) demonstra que o galego e o português daquela época não devem ser considerados línguas diferentes, mas sim “uma e a mesma língua”, obviamente dotada de variações internas, como qualquer outra língua natural. A autora, a partir da comparação entre as cantigas profanas (provenientes de Portugal) e as religiosas (compiladas em Toledo), ressalta que essas duas vertentes são muito próximas em relação aos elementos prosódicos e que “as distinções linguísticas [...] não são de tipologia dos fenômenos, mas de frequência. Não havendo distinções tipológicas, não há diferença de

sistema” (2007, p.122). Assim, é possível inferir que o galego-português das CSM pode ser considerado uma manifestação ancestral legítima do Português e, devido a essa constatação, é que foram escolhidas as CSM como *corpus* representante para nosso estudo. Partimos de um recorte de cem cantigas pertencentes ao códice de Toledo (To), uma vez que estas fazem parte do códice mais antigo e estão presentes em dois outros manuscritos, correspondendo à compilação original da coleção (cf. PARKINSON, 1998a). Como *corpus* de apoio, consideramos a edição crítica de Mettmann (1986-1989).

## Metodologia

Com relação ao fenômeno específico focalizado por este artigo, Henriques (2009) mostrou que até mesmo o falante nativo de línguas que permitem *onsets* complexos do tipo sibilante + oclusiva tem dúvidas na segmentação silábica das sequências /sp, st/. Em nosso trabalho, como não temos a presença física do falante, usufruímos de uma metodologia que busca abstrair dos limites das sílabas poéticas as sílabas fonéticas. No caso da observação da vogal epentética em início de palavra, partimos da escansão e da contagem das sílabas poéticas dos versos das CSM para poder elucidar dúvidas acerca da consideração de sua (não)realização. A presente metodologia busca abstrair, então, da escansão dos versos em sílabas poéticas os limites entre as sílabas fonéticas.

Essa metodologia foi inaugurada, no Brasil, por Massini-Cagliari (1995) e busca as características prosódicas de línguas mortas ou de períodos passados de línguas vivas por meio da estrutura métrico-poética da poesia.

Uma introdução à aplicação da metodologia acima descrita à análise dos dados das cantigas medievais religiosas, com vistas ao mapeamento do processo descrito, está exemplificada em (1), em que aparecem as duas primeiras estrofes da CSM 01.<sup>1</sup>

(1)	E/ non/ ar/ que/ <b>ro</b> / o/bri/dar	<b>a</b> <sup>8</sup>
	co/ <u>m</u> ' <u>an</u> /ge/os/ can/ta/da	<b>b</b> <sup>6</sup>
	lo/or/ a/ Deus/ fo/ron/ can/tar	<b>a</b> <sup>8</sup>
	e/ «paz/ en/ te/rra/ da/da»;	<b>b</b> <sup>6</sup>
	nen/ co/ <b>mo</b> / a/ con/tra/da	<b>b</b> <sup>6</sup>
	a/os/ tres/ Reis/ en/ Ul/tra/mar	<b>a</b> <sup>8</sup>
	ou/ <u>v</u> ' <u>a</u> s/tre/la/ mos/tra/da,	<b>b</b> <sup>6</sup>
	por/ que/ sen/ de/mo/ra/da	<b>b</b> <sup>6</sup>
	vẽ/e/ron/ <b>sa</b> / o/ffer/ta/ dar	<b>a</b> <sup>8</sup>
	es/tra/ <b>nna</b> / e/ pre/ça/da.	<b>b</b> <sup>6</sup>

## A não-realização da vogal epentética em Português Arcaico

Por meio do mapeamento feito do encontro de vogais em juntura vocabular em 100 CSM, encontramos em nosso *corpus* a não-realização da vogal inicial da segunda

1 As letras minúsculas a, b, c representam o padrão rimático das estrofes. Ao final do verso, as letras são seguidas de um algarismo, que representa a quantidade de sílabas poéticas do verso. As sílabas em sublinhado correspondem ao processo de elisão da última vogal da primeira palavra com a primeira vogal da segunda palavra, já as em negrito correspondem à formação do hiato da última vogal da primeira palavra com a primeira vogal da segunda palavra.

palavra, vogal epentética /e/, quando seguida por /s/ ou /n/.<sup>2</sup>

- (2) do/ de/mo/, **que s**/te/rre/ces. (CSM 20; verso 39) - 6 sílabas poéticas
- (3) nen/ d' a/gui/llo/ **a 's**/co/du/das. (CSM 31; verso 68) - 7 sílabas poéticas
- (4) e/no/ mes/ d' a/gos/to/, no/ di/**a 's**/co/llei/to (CSM 77; verso 27) - 11 sílabas poéticas
- (5) e/ di/ss': «Es/**ta 'n**/fer/me/da/de | se/me/lla/ mui/t' a/fi/ca/da. (CSM 75; verso 36) - hemistíquio de 7 sílabas poéticas
- (6) ou/tro/si/ nen/ d' al/va/yal/de,| que/ faz/ a/ ca/**ra 'n**rru/ga/da (CSM 75; verso 96) - hemistíquio de 7 sílabas poéticas
- (7) del/, a/ Re/y/na/ no/**bre s**/pi/ri/tal. (CSM 58; verso 53) - 10 sílabas poéticas
- (8) **ouv' a** strela mostrada (CSM 1; verso 38)

Nos casos (2), (3) e (4), podemos considerar, assim como o fez Massini-Cagliari (2005), baseada em Cunha (1961), que, pelo fato de os monossílabos portarem uma semiforça, fica mais restrita a possibilidade de a vogal de um monossílabo se elidir ou não com a vogal seguinte (do início da palavra seguinte), uma vez que essa está relacionada mais diretamente com o grau de tonicidade do monossílabo do que com a quantidade de sílabas das palavras envolvidas, como vemos a seguir.

**Quadro 1. Não-realização da vogal epentética**

Vogal final da primeira palavra	Vogal inicial da palavra seguinte	Quantidade	Percentual
a	e (e, ê / en)	6	50,0
e	e (e, ê / en)	6	50,0
<b>TOTAL</b>		<b>28</b>	<b>100%</b>

Massini-Cagliari (2005, p. 236) afirma que

[...] esses são processos marginais, cuja função é prioritariamente estilística, nos quais o poeta encontra apoio para obter a quantidade de sílabas poéticas desejada para o verso. Mas não se pode negar que, mesmo com função estilística, seu aparecimento está condicionado ao fato de a qualidade da primeira vogal ser /a/.

Nos casos (5) e (6), a qualidade da vogal /a/ é diferenciada das outras vogais, assim como afirmou Massini-Cagliari (2005), e elevada a um valor diferente das demais dentro do sistema, uma vez que nesses casos há sílabas átonas envolvidas e quando há palavras como uma sílaba final átona seguida de uma outra palavra iniciada por vogal acontece, *a priori*, elisão. No entanto, foi ressaltado por Massini-Cagliari (2005), e por nós em outros trabalhos (CANGEMI, 2010; 2011), que é pouco recorrente a elisão do /a/ final da primeira palavra. Então, temos que o valor de uma palavra ser terminada por /a/ é mais alto ao ser essa terminada em sílaba átona (ambas seguidas de outra vogal). Nesse sentido, quando se trata de uma vogal final /a/ seguida de outra há um rearranjo no sistema – os casos mais típicos são a resolução em hiato, os menos típicos a não-realização

<sup>2</sup> Foram mapeadas no âmbito das cem CSM todas as soluções para o encontro de vogais em junção de palavras. Os resultados encontrados foram 7296 ocorrências de encontros vocálicos intervocábulares. Desses, 3457 (47,4%) são referentes à elisão, 3262 (44,7%) referentes à hiatização de sílabas CV com vogal, 526 casos (7,2%) de hiatização de ditongos decrescentes (VV) com vogal, 11 casos (0,2%) de hiatização de ditongos crescentes (VV) com vogal, 12 (0,2%) de não-realização da vogal epentética (processo ainda não nomeados pela literatura especializada) e 28 (0,4%), processos de ditongação.

da vogal epentética. Fica evidenciado, portanto, que o fato de a vogal final da palavra ser /a/ é mais relevante que a sílaba ser átona.

No entanto, no caso (7), não podemos considerar essa explicação, pois temos um contexto diferente. Temos a sílaba final átona terminada em /e/ seguida da sibilante. Podemos considerar apenas que, igualmente aos casos anteriores (2) a (4), a sibilante se agarra na coda da sílaba anterior pertencente a primeira palavra.

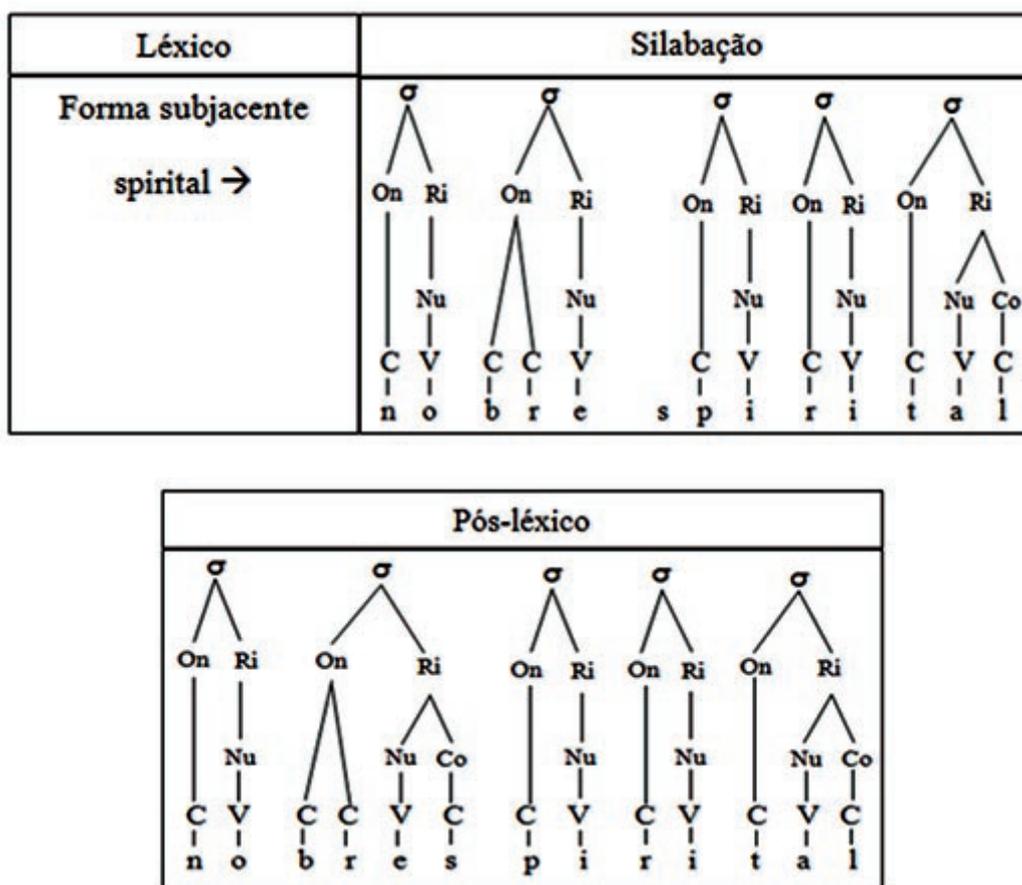


Figura 1. Silabação de *Spirital*

A respeito de (8), o primeiro caminho que tomamos ao ver essa realização foi de considerar a metrificação exposta em (9) (corroborando as afirmações que acabamos de fazer) e observar que ocorreria uma reestruturação silábica ligando a consoante inicial da segunda palavra à coda da primeira palavra.

(9) ou/v' a s/tre/la/ mos/tra/da (CSM 1; verso 38) - 6 sílabas poéticas

As consoantes perdidas (<s> ou <n>), aptas a ocupar a posição silábica de coda, se ligam diretamente ao nó silábico da sílaba final da primeira palavra, uma vez que o item lexical não conta com a vogal inicial de palavra, considerada somente *strela*. No entanto, podemos pensar em uma metrificação diferente para esse caso, que mantém as mesmas 6 sílabas poéticas; teríamos um *onset* ramificado:

(10) ou/v' a/ **stre**/la/ mos/tra/da (CSM 1; verso 38) - 6 sílabas poéticas

Para Massini-Cagliari (2005, p. 98, grifos nosso) este

[...] é um problema conhecido da fonologia, dado o fato de línguas já bastante exploradas (como o inglês e o italiano) possuírem sílabas deste tipo. No entanto, não parece ser este o caso do PA. Em todos os exemplos mapeados, as sequências de S+C(C) sempre ocorrem depois de palavra terminada em vogal. [...] **Por este motivo, é mais provável a hipótese de estar ocorrendo um processo de sândi, que apaga a vogal inicial de *estrela* (ou impede a sua inserção, no caso de modelos fonológicos que consideram essa vogal epentética), e liga o “S desgarrado” à coda da sílaba anterior.**

Embora o tipo de estrutura silábica presente na sílaba stre- (CCCV) da palavra *strela*, mostrada acima, seja menos recorrente como um tipo de sílaba fonética possível no PA, podemos acrescentá-la como uma realização possível, pois essa se mantém em outras CSM mesmo não sendo precedida de uma vogal (11) na metrificação dos versos. Nesse sentido, introdutoriamente, não haveria a possibilidade de considerar o som sibilante como pertencente à coda da sílaba anterior. Assim, teríamos uma perspectiva diferente de Massini-Cagliari (2005, p. 235), que considera nesse contexto (str-) que “a coda inicial ‘desgarrada’ /S/, então, adjunge-se ao núcleo da sílaba anterior”; para nós – nesse momento – não ocorreria uma ressilabação nesse contexto. Outro fator que vem a contribuir com essa hipótese é a realização dessa palavra em início de verso dentro nas CSM, não tendo nenhuma vogal antes de sua realização.

- (11) [[San/ta/ Ma/ri/a]I]U  
 [[**Stre**/la/ do/ di/a]I]U  
 Mos/tra/-nos/ vi/a  
 pe/ra/ Deus/ e/ nos/ gui/a  
 (CSM 100, 3 primeiros versos de 4 sílabas poéticas e o último 6 sílabas poéticas)
- (11') [Santa Maria]I  
 [Strela do dia]I

Tenderíamos a ir contra a proposta de Massini-Cagliari (2005, p. 99), que considera que “o PA não possui ataques silábicos supercomplexos”. O Princípio do Ataque Máximo para o PE, desenvolvido por Vigário e Falé (1994, p. 475), considera que “é preferível o preenchimento de Ataques ao preenchimento de Codas”. Por outro lado, a própria organização da sílaba em termos de sonoridade, mais especificamente de um crescendo de sonoridade, que coloca o problema com esse tipo de sequências (/sp/, /st/ ...), nomeadamente sequências de fricativa+oclusiva como tautossilábicas, que não são permitidas no português justamente por violarem o Princípio de Sonoridade. Blevins (1993, p. 211) observa que: “However, in English syllable initial /sp st sk/ occur, and postvocalic tautossyllabic /sp st sk/ are also found, and English is far from unique in this regard”. Selkirk (1982) defende que esses grupos consonânticos constituem uma unidade. Para Parker (2002, p. 9, grifo nosso),

A number of formal devices, some of which are painfully ad hoc, have been posited to explain away exceptional ‘sonority reversals’ like these: extrasyllabicity, syllable appendices and ‘afixes’, adjunction, non-exhaustive parsing, degenerate syllables, null or empty nuclei, language particular stipulation, complex phonetic units. [...] However, /s/ (or perhaps sibilants in general) may be a special case, due to their stridency.

Hogg e McCully (1987, p. 49) ressaltam que, por se tratar de uma unidade, não importa o valor dado ao constituinte único que esse não violará o *template* da sílaba.

Viana (1892, p. 24) observa que as sibilantes e nasais são segmentos mais complexos se comparados aos outros e afirma “há consoantes que podem constituir síllaba, funcionando como vogaes: são ellas as sibilantes se, z, as ancípites l, r, e as nasaes”. Camara Jr. (1973, p. 43) também não exclui essa possibilidade:

[...] normalmente a vogal, como o som vocal mais sonoro, de maior fôrça expiratória, de articulação mais aberta e de mais firme tensão muscular, que funciona em todas as línguas como centro da sílaba, embora algumas consoantes, particularmente as que chamamos «sonantes», não estejam necessariamente excluídas dessa posição.

Assim, sabemos que o princípio universal de escala de sonoridade inviabiliza a existência dessas sequências, de acordo com regras fonotáticas de outras línguas. A dificuldade em identificar a estrutura fonética e fonológica (existência ou não de vogal) origina problemas na segmentação silábica. Nossa proposta consiste em levantar hipóteses que são plausíveis com a metodologia empregada e com o aparecimento dessas sequências no *corpus*.

No entanto, outra análise seria se considerássemos essas ocorrências como um caso de *enjambement*, típico daqueles tempos para a poesia trovadoresca. Ocorreria uma reestruturação de U, composto por duas  $\phi$  e não somente uma. A vogal final da primeira palavra /a/ faria parte da segunda palavra – [strela] – dando como resultado a palavra fonológica [aStrela].

(12) [[San/ta/ Ma/ri/a]I, | [Stre/la/ do/ di/a]I]U,

(13) [MariaS] $\omega$  [trela] $\omega$

(13') [Mari] $\omega$  [aStrela] $\omega$

A elisão não ocorre entre Is (cf. MASSINI-CAGLIARI, 2005; CANGEMI, 2011) e essa é mais restrita em uma sílaba sem *onset*; a vogal /a/ se manteria e seria necessário outro procedimento para a resolução desse encontro, ou seja, o apagamento da vogal inicial (V<sub>2</sub>) da sequência.

(14) San/ta/ Ma/ri/a, | Stre/la/ do/ di/a,

(14') [Mari] $\omega$  [aStrela] $\omega$

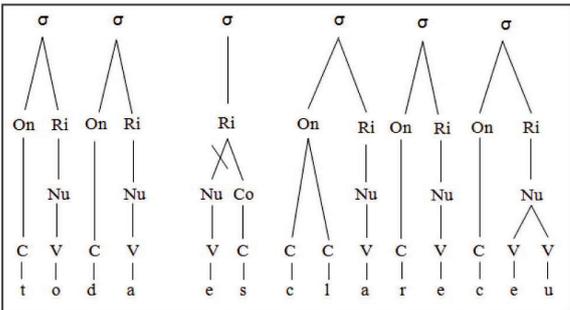
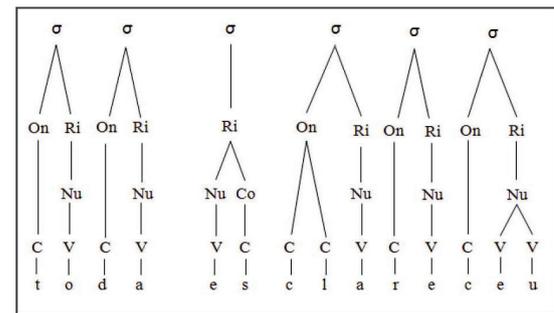
(15) que a terra toda 'sclareceu (CSM 15; verso 91)

(15') [to] $\omega$  [dasclareuceu] $\omega$

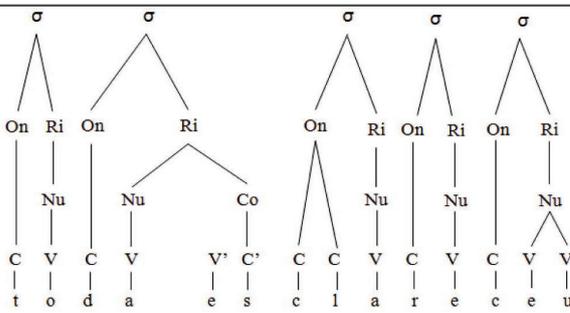
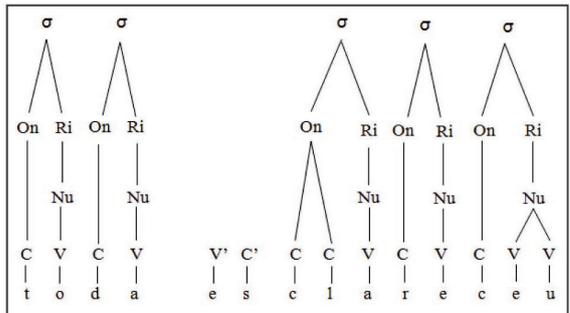
Para postular esse apagamento teríamos que considerar a seguinte  $\omega$ : [estrela] ou [esclareceu].

(16) [Maria] $\omega$  [estrela] $\omega$  → [Mari] $\omega$  [aStrela] $\omega$

(17) [toda] $\omega$  [esclareceu] $\omega$  → [to] $\omega$  [dasclareu] $\omega$



Choque  
→



Ress.  
→

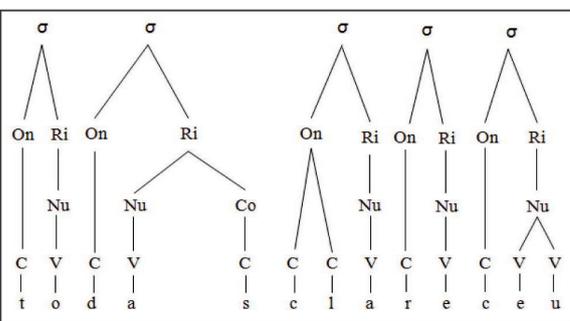


Figura 2. Silabação *toda sclareceu*

A resolução no apagamento da  $V_2$  seria possível ritmicamente também, uma vez que cada  $\omega$  portaria somente um acento:

- (18) [Mari]ω [aStrela]ω  
 (19) [to]ω [dasclareceu]ω

No entanto, considerar a elisão da vogal inicial da primeira palavra seria ir contra o princípio de economia das línguas, uma vez que nós postularíamos a inserção de uma vogal e depois o seu apagamento. Na perspectiva que adotamos, derivacional, seria postular a existência de regras coexistentes que se anulam. Seria o caso de pensarmos no tópico *Duke-of-York (DY) gambit* no trabalho de Pullum (1976): a forma da derivação seria  $A \rightarrow B \rightarrow A$ ; A como forma de base, subjacente, que passaria ao estágio B e depois retornaria a superfície como A novamente (20):

- (20) Strela  $\rightarrow$  estrela  $\rightarrow$  strela  
 A  $\rightarrow$  B  $\rightarrow$  A

O exemplo dado pelo autor é a análise do “r-dropping and r-intrusion” (apagamento de r e r intrusivo) em vários dialetos do Inglês: “final r is first deleted and then re-inserted before a vowel: Homer is  $\rightarrow$  Hom[ə] is  $\rightarrow$  Homer is (cf. Hom[ə] saw)” (MCCARTHY, 2003, p. 30). Pullum (1976) aborda esse caso e outros como ele, perguntando se derivações DY são obrigadas pelos fatos e se/como podem ser descartadas em geral. Podemos afirmar que elas são necessárias e não devem ser descartadas.

Casali (1997) diz que a elisão da primeira ( $V_1$ ) de duas vogais adjacentes é mais produtiva nas línguas, no entanto é atestada a elisão de  $V_2$  também. O autor sugere que há posições sensíveis que favorecem a permanência/queda das vogais, a saber: posição inicial da palavra, posição inicial da sílaba, conteúdo semântico e saliência fonética. O que torna o assunto difícil é o fato de que fonologicamente as duas saídas são praticamente idênticas. A silabificação resultante de ambos os processos é idêntica:

- (21) Elisão de  $V_1$  à  $CV_1 V_2CV = CV_2. CV$   
 (22) Elisão de  $V_2$  à  $CV_1 V_2CV = CV_1. CV$

Colina (1997) propõe uma análise do processo tradicionalmente chamado de epêntese e apagamento no Galego. A respeito da epêntese inicial na língua galega, a autora diz que essa é restrita à posição inicial (casos de /s/ + C) e à posição final, depois de obstruintes, e que resulta da necessidade de satisfazer a sonoridade do *onset*. Formas que não contenham a epêntese, como por exemplo [sta.bi.li.da.de], violam uma restrição altamente ranqueada, pois /st/ não é um *cluster* aceitável em início de sílaba.

Collischonn (2002) retoma que, em latim, qualquer consoante poderia aparecer em posição final de sílaba (exceto  $k^w$ ,  $g^w$ , f e h), enquanto o português arcaico (cf. SOMENZARI, 2006) e as outras línguas romance mostram uma tendência à redução ou até mesmo de eliminação de sílabas fechadas através de uma variedade de processos históricos, como a simplificação de geminadas e de grupos consonantais, a vocalização e a monotongação. O resultado disso é que as sílabas somente podiam ser fechadas por soantes ou /s/. A autora nos diz que, durante a silabação, uma consoante não-apta a ocupar uma posição silábica de ataque ou coda não seria ligada a nenhum nó silábico. A existência de uma dessas consoantes perdidas na representação fonológica desencadearia a criação de uma sílaba estrutural, desprovida ainda de núcleo vocálico, a qual permite a associação da consoante

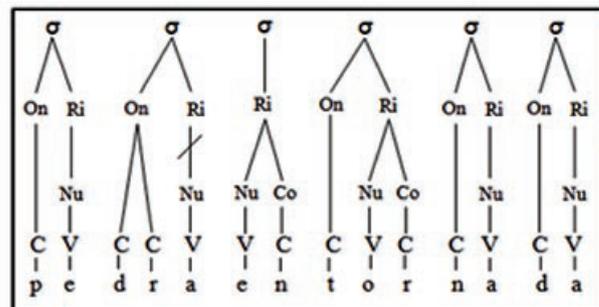
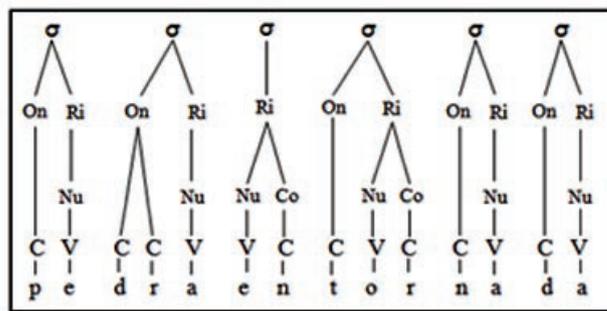
perdida em posição de ataque. Mais tarde, no pós-léxico, essa sílaba seria preenchida com uma vogal e a mora correspondente.

Se assumirmos que as línguas fazem um maior esforço para preservar material fonológico em contextos mais salientes, podemos esperar que o destaque acústico para o segmento da posição inicial da palavra lidere a preservação preferencial de elementos nesse contexto, fato que não ocorre quando  $V_2$  é elidida. Além disso, como postulou Massini-Cagliari (2005), a consoante perdida no choque das vogais pode, nesse caso, ocupar coda de  $V_1$ , se ligando ao nó silábico da primeira palavra. Nessa perspectiva, na hora de escolher qual vogal apagar, o trovador apaga  $V_2$ , vogal epentética – subespecificada – e deixa a vogal plena.

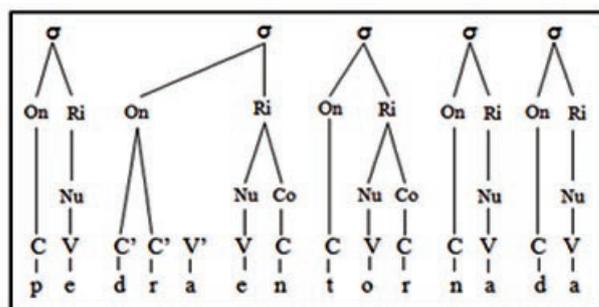
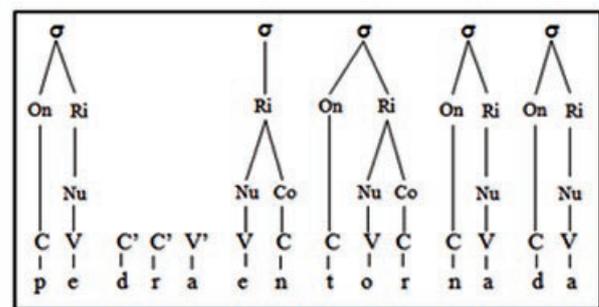
A partir da escansão dos versos em sílabas poéticas, pudemos observar os limites entre as sílabas fonéticas e constatar que nesses casos não é necessária a inserção da vogal epentética /e/ para que seja satisfeita a boa formação da sílaba e, conseqüentemente, a sonoridade do *onset*. Todavia, temos que ressaltar que há alguns casos do galego-português que o processo de resolução para essas consoantes perdidas (casos de /s/ + C ou /n/ + C) ora se realiza com a inserção de uma vogal epentética (23) e, como vimos, ora, não (exemplos de 2 a 8, acima).

(23) v̄yu/ a/ pe/dr' en/tor/na/da (CSM 1; verso 46) - 6 sílabas poéticas

Quando acontece a sua realização, temos como resultado uma elisão, isto é, a vogal final da primeira palavra é apagada. Na elisão, o choque das duas rimas dessilabifica a primeira sílaba, que deixa C' e V' flutuantes, com os seus segmentos pendentes. A ressilabificação ativada pelo Princípio de Licenciamento Prosódico faz da consoante perdida o ataque da sílaba remanescente, criando uma sílaba ótima, mas esquece a vogal, que é apagada por convenção.



Choque  
→



Ress.  
→

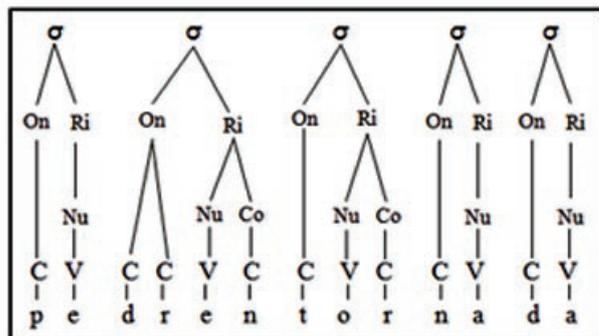


Figura 3. Silabação *pedra ntornada*

Quando isso não acontece, uma das hipóteses seria considerar essas consoantes, aptas a ocupar a posição silábica de coda, ligadas ao nó silábico da sílaba final da primeira palavra.

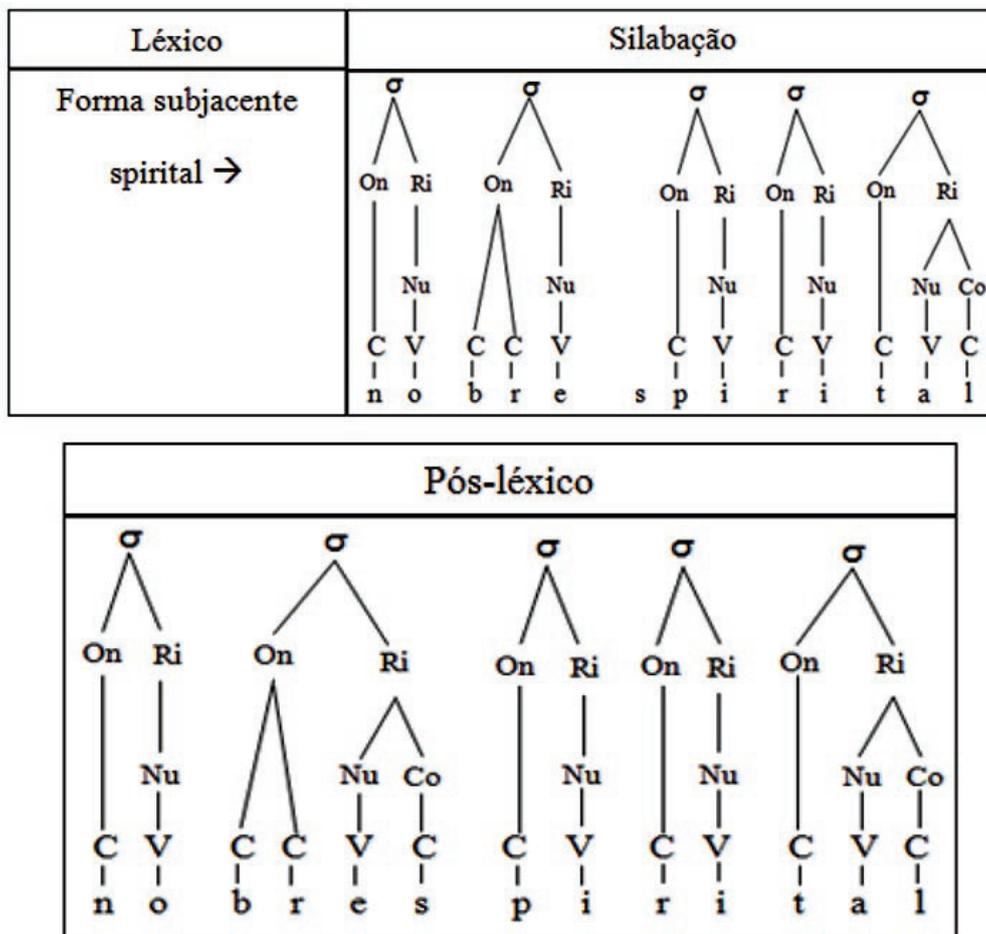


Figura 4. Silabação *nobre spirital*

Notamos também que a não-realização da vogal epentética tem um ambiente rítmico comum: a) sílaba átona seguida de uma sílaba pretônica; b) monossílabo seguido de uma sílaba pretônica.

Nessa perspectiva, podemos postular – introdutoriamente – que /s/ e /n/ podem provocar ou não, em início de palavra e sendo essa sílaba pretônica, a epêntese. Em outras palavras, a epêntese externa nem sempre ocorre para otimizar uma estrutura silábica. Parece-nos que, se as propriedades de sonoridade da sílaba estiverem sendo satisfeitas pela junção dos vocábulos, não haverá a necessidade da realização da vogal epentética.

Assim, por meio das possibilidades de escolha do trovador e identificando as resoluções dos encontros vocálicos no sistema do PA dado pelo falante podemos perceber nuances dos valores atribuídos para esse processo fonológico. É necessário ressaltar que as vogais finais envolvidas da primeira palavra são /a/ ou /e/ e a vogal inicial que não se realiza é sempre /e/.

A dificuldade em identificar a estrutura fonética e fonológica (existência ou não de vogal) origina problemas na segmentação silábica. Até mesmo o falante tem dúvidas na

segmentação silábica das sequências /sp, st/ (cf. HENRIQUES, 2009). Em nosso trabalho, como não temos a presença física do falante, usufruímos de uma metodologia que buscar abstrair dos limites das sílabas poéticas, as sílabas fonéticas.

## Conclusão

Este artigo mostrou que, em termos de sândi no PA, há a ocorrência de um processo ainda não muito explorado pela literatura especializada, que consiste na não-realização da vogal inicial da segunda palavra quando essa é precedida por /s/ ou /n/. Quando essa vogal se realiza, é grafada <e>, vogal epentética do português da época.

Aliás, a determinação do *status* fonológica da vogal “apagada” pelo processo descrito como epentética ou não (ou seja, como presente ou não na forma de base ou no *input*) é o ponto crucial do qual depende toda a análise, uma vez que a dificuldade com relação à discussão do fenômeno aqui focalizado reside principalmente em identificar a estrutura fonológica, em termos de existência ou não de vogal na sua subjacência, em termos derivacionais.

A partir dessa perspectiva, supomos que, se as propriedades de sonoridade da sílaba estiverem sendo satisfeitas pela junção dos vocábulos, não haverá a necessidade da realização da vogal epentética, sendo a questão mais crucial determinar a forma subjacente ou *input* de palavras específicas, em um contexto de análise histórica, em que os dados nem sempre são abundantes nem de fácil interpretação.

## REFERÊNCIAS

- BLEVINS, J. The syllable in phonological theory. In: GOLDSMITH, J. (Ed.) *The handbook of Phonological Theory*. Cambridge: Oxford UK Blackwell, 1993. p. 206-244.
- CAMARA JR, J. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão, 1973.
- CANGEMI, A. Algumas considerações sobre a solução dos processos de sândi vocálico externo no Português Arcaico. In: CONGRESSO INTERNACIONAL DA ASSOCIAÇÃO BRASILEIRA DE LINGUÍSTICA (ABRALIN), VII, 2011, Curitiba. *Anais...* Curitiba: Abralin, 2011. p. 159-173.
- \_\_\_\_\_. Estudos preliminares sobre os processos de sândi vocálico externo - destaque para a elisão – do Português Arcaico ao Português Brasileiro atual. In: ENCONTRO DO CÍRCULO DE ESTUDOS LINGUÍSTICOS DO SUL, IX, 2010, Palhoça: Editora Unisul. *Anais...* p. 1-14. Disponível em: <<http://www.celsul.org.br/Encontros/09/artigos/Carol%20Cangemi.pdf>> Acesso em: 10 jan. 2011.
- CASALI, R. Vowel elision in hiatus contexts: which vowel goes? *Language*, v. 73, n. 3, p. 493-533, 1997.
- COLINA, S. Epenthesis and deletion in Galician: na optimality-theoric approach. In: MORALES-FRONT, A.; MARTÍNEZ-GIL, F. (Org.) *Issues in the Phonology and Morphology of the Major Iberian Languages*. Washington: Georgetown University, 1997. p. 235-267.

- COLLISCHONN, G. A epêntese vocálica no português do Sul do Brasil. In: BISOL, L.; BRESCANCINI, C. (Org.). *Fonologia e variação do Português Brasileiro*. Porto Alegre: EDIPUCRS, 2002. p. 205-230.
- CUNHA, C. *Estudos de Poética Trovadoresca: versificação e ecdótica*. Rio de Janeiro: Instituto Nacional do Livro, 1961.
- HENRIQUES, I. A importância da sílaba: uma reflexão fonológica. *Centro de Linguística da Universidade do Porto*, Portugal, 2009. p. 37-59.
- HOGG, R.; MCCULLY, C. B. *Metrical Phonology: a coursebook*. Cambridge: Cambridge University Press, 1987.
- ILARI, R.; BASSO, R. *O português da gente*. São Paulo: Contexto, 2007.
- MASSINI-CAGLIARI, G. Legitimidade e identidade: da pertinência da consideração das Cantigas de Santa Maria de Afonso X como *corpus* da diacronia do Português. In: MURAKAWA, C.; GONÇALVES, M. (Org.) *Novas contribuições para o estudo da história e da historiografia da língua portuguesa*. São Paulo: Cultura Acadêmica, 2007. p. 101-126.
- \_\_\_\_\_. *A música da fala dos trovadores: estudos de prosódia do português arcaico, a partir das cantigas profanas e religiosas*. Tese (Livre docência em Linguística) - Faculdade de Ciências e Letras-UNESP, Araraquara, 2005.
- \_\_\_\_\_. *Cantigas de amigo: do ritmo poético ao linguístico. Um estudo do percurso histórico da acentuação em Português*. 1995. Tese (Doutorado em Linguística) - IEL/ UNICAMP, Campinas, 1995.
- MCCARTHY, J. Sympathy, Cumulativity, and the Duke-of-York Gambit. In: FÉRY, C.; VIJVER, R. (Org.). *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, 2003. p. 30-111.
- METTMANN, W. (Ed.). *Cantigas de Santa Maria (cantigas 1 a 100)*: Alfonso X, El Sabio. Madrid: Castalia, 1986.
- NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.
- O'CALLAGHAN, J. *Alfonso X and the Cantigas de Santa Maria: a poetic biography*. Boston: Brill, 1998.
- PARKER, S. *Quantifying the Sonority Hierarchy*. 2002. 372f. Tese (Doutorado em Filosofia) – Universidade de Massachussets Amherst, Massachussets.
- PARKINSON, S. As Cantigas de Santa Maria: estado das questões textuais. *Anuario de estudios literarios galegos*, Vigo, p. 179-205, 1998a.

\_\_\_\_\_. Two for the price of one; on the Castroxeriz *Cantigas de Santa Maria*. In: FLITTER, D. W.; BAUBETA, P. O. (Coord.). *Ondas do Mar de Vigo: Actas do Simposio Internacional sobre a Lírica Medieval Galego-Portuguesa*. Día das Letras Galegas. Birmingham, UK: Seminario de Estudos Galegos, Department of Hispanic Studies, University of Birmingham, 1998b. p. 72-88.

PULLUM, G. The Duke of York gambit. *Journal of Linguistics*, Grã-Bretanha, v. 12, n. 12, p. 83-102, 1976.

SELKIRK, E. O. *Phonology and Syntax*. Cambridge, Mass.: MIT Press, 1984.

SOMENZARI, T. Status fonológico dos casos de sândi de vogais idênticas. In: \_\_\_\_\_. *Estudo da Possibilidade de Geminção em Português Arcaico*. Dissertação (Mestrado em Linguística e Língua Portuguesa) - Faculdade de Ciências e Letras, Araraquara, 2006.

SNOW, J. T. Current Status of Cantigas Studies. In: KATZ, I. J.; KELLER, J. E. (Ed.). *Studies on the Cantigas de Santa Maria: Art, Music, and Poetry*. Madison: The Hispanic Seminary of Medieval Studies, Ltd., 1987. p. 475-486.

VIANA, A. *Exposição da Pronúncia Normal Portuguesa para uso de nacionais e estrangeiros*. Lisboa: Imprensa Nacional, 1892.

VIGÁRIO, M.; FALÉ, I. A sílaba do Português Fundamental: uma descrição e algumas considerações de ordem teórica. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA, IX, 1993, Coimbra. Lisboa: APL/Colibri, 1994. p. 465-478.

# Um estudo descritivo-comparativo sobre a colocação pronominal em jornais de São Paulo e de Rio Claro

(A descriptive-comparative study on the position of clitic pronouns in newspapers of São Paulo and Rio Claro)

Caroline Carnielli Biazolli<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Linguística e Língua Portuguesa – Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp)

carolbiazolli@yahoo.com.br

**Abstract:** The present study verifies the position of clitic pronouns, associated to a single verb, in clauses which are part of journalistic texts produced in São Paulo city and in Rio Claro town, located in the interior of São Paulo State. The period analyzed embraces the years 1880 to 1920. This spatial cut is due to the diverse profile of the locations, which are analyzed to verify probable similarities and particularities of the use of clitic pronouns in those areas. The study is based on the theoretical and methodological purposes from the Theory of Linguistic Variation and Change, and investigates of which independent linguistic and non-linguistic variables motivate, in a sentence, a certain position of the clitic pronouns. These positions, the most relevant ones, are indicated by the absolute results, percentage and relative weights, provided by the *GOLD-VARB X* program.

**Keywords:** clitic pronoun; paulista portuguese, paulista press; linguistic variation and change.

**Resumo:** O presente artigo averigua a posição dos clíticos pronominais, adjungidos a um único verbo, em orações presentes em textos jornalísticos produzidos na cidade de São Paulo e no município interiorano paulista de Rio Claro, no período que abrange os anos de 1880 a 1920. Opta-se pelo referido recorte espacial por se tratar de localidades com perfis, em termos, diversos, fazendo-se uma análise para verificar prováveis semelhanças e particularidades do emprego dos pronomes clíticos nessas regiões. Baseando-se nos pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística, investigam-se quais variáveis independentes linguísticas e não-linguísticas motivam, dentro da oração, determinada posição dos pronomes clíticos, apresentando-as, somente as mais relevantes, a partir dos resultados absolutos, percentuais e pesos relativos, fornecidos pelo *GOLDVARB X*.

**Palavras-chave:** pronome clítico; português paulista; imprensa paulista; variação e mudança linguísticas.

## Introdução<sup>1</sup>

Reconhece-se que um expressivo número de estudos linguísticos, distribuídos sob as mais diversas orientações teórico-metodológicas, já tiveram como objeto de investigação determinado aspecto referente aos pronomes clíticos (DUARTE, 1986; PAGOTTO, 1992; CYRINO, 1996; VIEIRA, 2002; BIAZOLLI, 2010, entre tantos outros). No entanto, quanto à posição que ocupam numa oração, no que se refere a sua abordagem nos estudos de cunho teórico-explicativo e descritivo, não se deve inferir que todas as circunstâncias, sejam elas de valor linguístico ou social, já tenham sido relatadas; pelo

<sup>1</sup> O conteúdo apresentado neste artigo compõe parte do estudo de Biazolli (2010), sob orientação da Prof<sup>ª</sup>. Dr<sup>ª</sup>. Rosane de Andrade Berlinck e financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Processo n. 08/51935-1).

contrário, acredita-se que as pesquisas que ainda pretendem discorrer sobre os clíticos pronominais, principalmente aquelas que se voltam à análise dos grupos de fatores que determinam a escolha de suas variantes – sabe-se que, adjuntos a um único verbo, os pronomes átonos podem ocupar as posições proclítica, mesoclítica ou enclítica e, adjungidos a um complexo verbal, podem se alternar nas posições pré-complexo verbal (cl V1 V2), intra-complexo verbal (V1 cl V2) ou pós-complexo verbal (V1 V2 cl) –, possam contribuir de maneira significativa ao que se conhece, até agora, sobre esses condicionamentos.

O presente artigo, sob a perspectiva da Linguística Histórica e da Sociolinguística Variacionista, debruça-se sobre o estudo da posição dos pronomes clíticos, utilizando-se, como matriz das análises, produções jornalísticas elaboradas no final do século XIX e início do século XX, particularmente entre os anos de 1880 a 1920, e oriundas das cidades de São Paulo e Rio Claro. Observa-se o comportamento desses pronomes, em contextos de um único verbo, a fim de que se permita, além de contribuir com a descrição da história do Português Brasileiro (doravante PB) e variedade paulista, averiguar as preferências de colocação, modificadas, consoante outras pesquisas apontam, num curto intervalo de tempo.

Baseado na proposta teórico-metodológica da Teoria da Variação e Mudança Linguísticas (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 1994, 2001, 2008[1972]), este estudo considera fundamental um levantamento dos possíveis elementos condicionadores, linguísticos e não-linguísticos, que estariam, desde o período observado, a controlar os usos alternados – mencionados no primeiro parágrafo desta seção – da colocação dos clíticos pronominais. Insere-se, no âmbito dos aspectos extralinguísticos elencados, considerações pertinentes acerca dos gêneros textuais, remetendo-se, assim, a pesquisa em questão, também a conceitos referentes a essa esfera de reflexão.

Opta-se pela análise de dados oriundos das cidades de São Paulo e de Rio Claro por se tratar de localidades com perfis, em termos, diversos. Ao observar, por um lado, São Paulo, com desenvolvimento mais acelerado e um fluxo mais intenso de pessoas, vindas de vários lugares, compartilhando, entre outras, as suas formas de falar, e, por outro, Rio Claro, que, embora bastante próspera no período a ser investigado, caracterizava-se por uma economia fundada na agricultura, sendo um espaço regrado e mais homogêneo, tenciona-se destacar, concomitantemente, as semelhanças entre os índices de ocorrências de cada alternativa e entre os fatores elencados como decisivos para suas aparições e as peculiaridades encontradas dentre os dados de cada região. Nessa direção, nos jornais de São Paulo, tendendo essa cidade à inovação, presume-se haver, de um modo geral – sem se atentar, nesse momento, a contextos linguísticos –, maior aceitação da próclise; enquanto na produção do interior, por manter um ideal de conservação, inclusive para a língua, imagina-se um maior uso dos pronomes clíticos na posição pós-verbal.

O material a ser utilizado como fonte das análises concentra exemplares do jornal *A Província de São Paulo* que, a partir de 1890, renomeado, passa a ser *O Estado de São Paulo*, produzidos na capital paulista, e os jornais *Correio do Oeste*, *O Tempo*, *Diário do Rio Claro*, *O Rio Claro*, *A Mocidade* e *A Semana Militar*, do município de Rio Claro. Quanto à variedade de periódicos consultados para a cidade interiorana, explica-se pela inexistência da conservação de um mesmo jornal que mantenha uma linearidade cronológica. Assim, para que os resultados averiguados nas duas localidades possam ser contrastados, são escolhidos, dentre um número restrito, os jornais de Rio Claro que mais se aproximam da época compreendida pela pesquisa e das características salientes nos jornais da cidade de São Paulo.

A seguir, após essas considerações iniciais, apresentam-se os parâmetros de análise dos dados, abarcando desde a composição do *corpus* até as variáveis dependente e independentes, linguísticas e não-linguísticas, aqui retratadas, e, também, os principais resultados, oriundos das análises referentes às variáveis independentes selecionadas pelo programa estatístico *GOLDVARB X* (SANKOFF; TAGLIAMONTE; SMITH, 2005) como as mais motivadoras para a realização de determinada posição do pronome átono. Além disso, evidenciam-se, ainda, algumas palavras finais e a lista das obras citadas.

## **Parâmetros de análise dos dados**

Nesta seção, relatam-se os procedimentos adotados para o alcance de uma análise válida e confiável que retrate o fenômeno linguístico observado em questão: a posição dos pronomes clíticos, em textos jornalísticos, no período que compreende o final do século XIX e o início do século XX. Discorre-se a respeito da constituição do *corpus* escolhido como matriz das análises, destacando-se, em seguida – por se tratar de um estudo circunscrito às premissas da Sociolinguística Variacionista –, o envelope de variação considerado, isto é, apresentam-se as variantes correspondentes à variável da posição dos clíticos pronominais – adjuntos a um único verbo –, e, também, enumeram-se os elementos condicionadores, não-linguísticos e linguísticos, vistos como possíveis responsáveis pela motivação do uso de determinada posição.

### **A composição do *corpus***

Como anteriormente mencionado, opta-se por uma amostra composta de jornais das cidades de São Paulo e de Rio Claro. Reconhece-se que as formas linguísticas em documentos históricos são muitas vezes distintas das formas vernáculas, podendo refletir usos da variedade normativa; no entanto, destaca-se a opinião defendida no presente estudo de que nos periódicos, por serem formados por vários gêneros textuais, que podem ocorrer como formas rígidas ou inovadoras ou mistas, há a convivência de formas linguísticas padrão e não-padrão. Desse modo, os textos presentes nos jornais são materiais relevantes para estudos de variação e mudança linguística.

Destaca-se, portanto, que, neste estudo, a coleta dos dados – sentenças que apresentavam o clítico pronominal – foi realizada a partir de todos os gêneros que compunham os periódicos em questão.

Os exemplares, dos dias 04/01/1880, 06/01/1885 – do jornal “A Província de São Paulo” –, 08/01/1890, 03/01/1895, 08/01/1900, 18/01/1905, 03/01/1910, 02/01/1915, 02/01/1920 – do jornal “O Estado de São Paulo” –, 02/05/1880 – do jornal “Correio do Oeste” –, 23/07/1885 – do jornal “O Tempo” –, 21/11/1894 – do jornal “Diário do Rio Claro” –, 30/11/1900, 26/12/1905, 06/01/1910, 03/01/1915 – do jornal “O Rio Claro” –, 03/04/1920 – do jornal “A Mocidade” – e 28/03/1920 – do jornal “A Semana Militar” – foram examinados minuciosamente, registrando-se, de acordo com o avanço dos anos, no caso dos periódicos da capital paulista, modificações em suas estruturas e funções comunicativas, refletindo-se, de forma justificável, também, na aparição de outros gêneros textuais, não identificados nos primeiros exemplares estudados. Quanto às produções rioclareses, devido à pluralidade de jornais analisados, notam-se, entre eles, diversos registros de gêneros, quando observados relacionados aos anos.

Na sequência, arrolam-se as variáveis, dependente e independentes, aqui consideradas.

### A variável dependente

Considera-se que a alguns vocábulos gramaticais são atribuídos os termos *átonos* ou *clíticos*, uma vez que tais vocábulos apresentam função de morfema, figurando sem acento na frase. Segundo Câmara Jr. (1985, p. 39), “isto significa que na enunciação eles são incorporados a um vocábulo contíguo, como uma ou duas sílabas a mais desse vocábulo, ficando em próclise ou em ênclise.” Desse modo, conforme Matthews (1997, p. 56), define-se *clítico* como<sup>2</sup>

Um elemento gramatical tratado como uma palavra independente na sintaxe, mas formando uma unidade fonológica com a palavra que o precede ou o segue. Por exemplo, em grego antigo *tis* é um clítico em *nêsós tis* ‘uma (certa) ilha’: ele é flexionado de forma independente (neste caso como nominativo singular), mas acentuadamente ele forma uma unidade com a palavra ‘ilha’ (basicamente *nêsós*) que o precede.<sup>3</sup>

Estabelecem-se, portanto, como variantes da variável dependente examinada neste estudo – a posição dos pronomes átonos, que são os clíticos pronominais – as posições **proclítica** (cf. exemplo 01) e **enclítica** (cf. exemplo 02), considerando-se, para as análises, os pronomes clíticos referentes à primeira, segunda e terceira pessoas do singular e do plural, não-reflexivos e reflexivos, e exercedores de diferentes funções sintáticas.

(01) *Perto da aldeia de Jesergetz, deixámos uma columna alleman aproximar-se a cerca de trezentos passos das nossas posições e em seguida a recebemos com fogo cerrado [...].* (O ESTADO DE SÃO PAULO, 1915)

(02) *Prometemos para hoje exemplos das duas enumeradas modalidades de semelhante torpeza. Desobriguemo-nos desse empenho.* (O RIO CLARO, 1900)

### As variáveis independentes

Como já sabido, as variáveis independentes (ou grupos de fatores) podem ser de natureza externa ou interna à língua. De acordo com a intensidade com que influem na realização das variantes, aumentam ou diminuem a frequência de ocorrência de determinado uso.

Observam-se, neste estudo, os seguintes grupos de fatores, não-linguísticos e linguísticos:

- Ano do jornal: 1880, 1885, 1890, 1895, 1900, 1905, 1910, 1915 e 1920;

2 Nossa tradução.

3 No original: “A grammatical element treated as an independent Word in syntax but forming a phonological unit with the Word that precedes or follows it. E.g. Ancient Greek *tis* is a clitic in *nêsós tis* ‘a (certain) island’: it is inflected independently (in this case as nominative singular) but accentually it forms a unit with the word for ‘island’ (basically *nêsós*) that precedes it” (MATTHEWS, 1997, p. 56).

- Nome do jornal: “A Província de São Paulo”, “O Estado de São Paulo”, “Correio do Oeste”, “O Tempo”, “Diário do Rio Claro”, “O Rio Claro”, “A Mocidade” e “A Semana Militar”;
- Cidade: São Paulo e Rio Claro;
- Gênero textual em que o clítico estava inserido: *edital, notícia, aviso, anúncio, classificado, editorial, artigo, resenha (ou crítica), crônica, carta do leitor, nota e comentário*;
- Tipo de clítico: *me, te, o(s)/a(s) e formas variantes (FV), lhe(s), se, nos e vos*;
- Função do clítico: *acusativo, dativo, dativo ético, dativo de posse, predicativo, apassivação, indeterminação e inerência/reflexividade*;
- Formas verbais: *presente, pretéritos e futuros do Indicativo, presente, pretérito e futuro do Subjuntivo, Imperativos e Formas Nominais*;
- Tipo de verbo, do ponto de vista lógico-semântico: *ação, processo, ação-processo e estado*;
- Presença ou ausência de elemento proclisador na oração; e,
- Verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração.

## Principais resultados

Os resultados apresentados são oriundos de observações referentes à manifestação das variáveis independentes extralinguísticas e linguísticas apontadas, pelo tratamento estatístico realizado pelo programa *GOLDVARB X*, como as que atuam de forma mais relevante sobre a posição dos pronomes clíticos na oração.

Na tabela a seguir são mostrados os números de ocorrências dos clíticos pronominais, adjungidos a lexias verbais simples, nos referidos jornais, num total de **3.331** dados.

**Tabela 1: Distribuição geral das ocorrências de clíticos pronominais nos jornais de São Paulo e de Rio Claro, de 1880 a 1920**

	Jornais/São Paulo	Jornais/Rio Claro	Total
<b>Lexias Verbais Simples</b>	2.785	546	3.331

Dentre os dados obtidos nas publicações da capital do estado, **983 (35,3%)** pronomes ocupavam a posição pré-verbal e **1.802 (64,7%)** a posição pós-verbal, enquanto nas produções rioclarenses, foram coletadas **236 (43,2%)** ocorrências de pronomes proclíticos e **310 (56,8%)** de pronomes enclíticos.

Considerando-se, de um modo geral, os percentuais dos usos dos pronomes clíticos nas posições pré-verbal e pós-verbal, nos jornais oriundos das duas localidades em

questão, sem, nesse momento, atentar-se aos contextos linguísticos em que figuravam, pode-se supor a não comprovação de uma das hipóteses desta pesquisa: por um lado, de que na cidade de São Paulo, por ser uma localidade densamente mais povoada, onde a fermentação da vida é mais intensa, seria mais expressiva a frequência do uso da próclise – apontada, naquela época, como uma tendência maior de uso do PB. Por outro lado, de que em Rio Claro, devido ao seu perfil mais conservador, o contrário seria notado: um predomínio altamente significativo do uso da ênclise.

#### Próclise x Ênclise: variáveis extralinguísticas selecionadas

Neste estudo, foram indicadas as variáveis *jornal estudado* e *gênero textual* como os grupos de fatores mais significativos quanto à motivação da colocação do pronome clítico em determinada oração. Aqui, são explicitados os resultados referentes à segunda variável mencionada, selecionada como a primeira mais relevante na análise dos dados dos jornais de São Paulo e a segunda mais significativa referente às informações dos jornais de Rio Claro, na sequência da variável *jornal estudado*.

#### Variável *gênero textual*

No que concerne à variável *gênero textual*, pode-se mencionar que, de acordo com as funções comunicativas que seus discursos desempenham e com as características textuais que apresentam, atentando-se às singularidades de cada um, os gêneros foram, primeiramente, identificados, e nomeados – ressaltando-se o caráter conflituoso dessa classificação, por se tratar de produções do final do século XIX e início do século XX –, para, então, averiguar as posições dos clíticos pronominais em cada texto selecionado.

Desse modo, a fim de que fossem encontradas respostas a uma das hipóteses estabelecidas neste artigo – a saber, a de que os textos que circulam nos jornais possam ser fonte de formas ou construções linguísticas conservadoras e inovadoras, considerando-se a presença (ou ausência) dessas unidades diretamente relacionadas com o gênero que cada um desses textos materializa –, convém reforçar a imprescindibilidade de uma caracterização pormenorizada dos gêneros textuais observados nos jornais analisados em questão.

Assim, como descrito anteriormente, foram elencados os seguintes gêneros textuais, presentes nos jornais das cidades de São Paulo e de Rio Claro: *edital*, *editorial*, *notícia*, *nota*, *comentário*, *aviso*, *artigo*, *resenha* (ou *crítica*), *crônica*, *carta do leitor*, *anúncio* e *classificado*. Segundo os traços peculiares desses gêneros textuais, esperava-se observar realidades diversas, em cada um deles, isto é, ora a predominância do pronome enclítico ora a do pronome proclítico, de acordo com o gênero textual averiguado.

Acredita-se que os gêneros *edital*, *notícia* e *aviso*, embora em escalas diferentes – já que o *edital*, por exemplo, em consequência da sua função, essencial, de documento oficial, caracteriza-se pelo seu caráter demasiadamente rígido –, asseguram, como resultado de suas estruturas organizacionais e de seus conteúdos detalhados, um uso mais representativo da forma (em geral considerada) conservadora, a ênclise.

Quanto aos gêneros *anúncio* e *classificado*, espera-se que, embora possibilitem aos indivíduos maior liberdade de criação, devido à diversidade de temas que abrangem, apresentem-se redigidos, quase, invariavelmente, sob as mesmas formas, privilegiando, também, a colocação enclítica.

Por outro lado, espera-se nos gêneros textuais *editorial*, *artigo*, *resenha* (ou *crítica*), *crônica* e *carta do leitor* a incorporação de usos linguísticos variados, destacando-se a posição pré-verbal. Tal apontamento, no caso do gênero *editorial*, deve-se por sua característica de apresentar ao leitor determinado acontecimento e persuadi-lo a adotar a sua opinião, utilizando-se de formas mais usuais, e, quanto aos demais gêneros – *artigo*, *resenha* (ou *crítica*), *crônica* e *carta do leitor* –, por poderem retratar os assuntos mais diversos e por apresentarem finalidades as mais distintas, salientando-se, muitas vezes, características associadas ao próprio escritor.

Por fim, no que concerne aos gêneros *nota* e *comentário*, se considerados o dinamismo e a brevidade de seus textos, apostando-se na transmissão clara e direta de determinada mensagem, presume-se maior aceitação da próclise; no entanto, consoante a intenção de quem os produz, pode-se verificar formas mais rebuscadas, acentuando-se, também, a ocorrência da ênclise.

Os resultados obtidos, em relação à posição dos pronomes clíticos, de acordo com os gêneros textuais, podem ser apurados nas tabelas seguintes.

**Tabela 2:<sup>4</sup> Número de ocorrências, percentuais e pesos relativos dos pronomes clíticos, de acordo com os gêneros textuais dos jornais de São Paulo, de 1880 a 1920**

	<i>Próclise</i>			<i>Ênclise</i>			<b>Total</b>	
	N	%	PR	N	%	PR	N	%
<b>Resenha</b>	29	72.5	0.859	11	27.5	0.141	40	1.5
<b>Editorial</b>	65	65	0.832	35	35	0.168	100	3.7
<b>Artigo</b>	57	60.6	0.830	37	39.4	0.170	94	3.5
<b>Carta do leitor</b>	144	57.1	0.783	108	42.9	0.217	252	9.3
<b>Comentário</b>	20	51.3	0.722	19	48.7	0.278	39	1.4
<b>Aviso</b>	104	49.1	0.702	108	50.9	0.298	212	7.9
<b>Edital</b>	78	42.9	0.639	104	57.1	0.361	182	6.7
<b>Notícia</b>	249	41	0.620	358	59	0.380	607	22.5
<b>Nota</b>	97	28.4	0.493	245	71.6	0.507	342	12.7
<b>Anúncio</b>	73	18.9	0.356	314	81.1	0.644	387	14.3
<b>Classificado</b>	12	2.7	0.062	435	97.3	0.938	447	16.5
<b>Total</b>	928	34.3		1774	65.7		2702	

Identificou-se, nos dados retirados dos jornais de São Paulo, significativa relação entre as considerações, descritas acima, e os resultados averiguados.

Quanto aos dados dos gêneros *edital*, *notícia* e *aviso* – ainda que o último apresente uma diferença bastante comedida entre as ocorrências de próclise e ênclise (cf. tabela 2) –, pode-se dizer que confirmaram a tendência de que a ênclise prevalece em textos com estrutura e papel a desempenhar mais cuidadosos, sendo produzidos, então, com maior monitoramento. No entanto, os pesos relativos apontam para a próclise.

Os resultados provenientes dos gêneros *editorial*, *artigo*, *resenha* e *carta do leitor* permitiram afirmar a predominância da próclise em textos que, muitas vezes, principalmente de acordo com os temas que retratam, procuram construir um lugar de familiaridade para a relação enunciador/enunciatório, qualificando-se como mais subjetivos.

4 O gênero *crônica* foi encontrado no periódico de 1920. Por estar presente apenas em um exemplar, foi desconsiderado da análise.

O fato dos gêneros *anúncio* e *classificado* terem apresentado predomínio relevante do uso do pronome enclítico – no caso do *classificado*, a posição pós-verbal foi praticamente categórica, apresentando tendência ao uso da próclise de apenas 0.062 – corrobora a ideia de possuírem certo grau de rigidez, através do uso de expressões cristalizadas.

O comportamento dos gêneros *nota* e *comentário* revelou orientações diversas. No gênero *nota*, as frequências indicaram expressiva diferença entre os usos dos pronomes proclíticos e enclíticos e, no gênero *comentário*, assinalaram representativo equilíbrio entre eles. Quanto aos pesos relativos, naquele gênero se notou uma tendência discreta à ênclise (0.507) e, neste, notável significância da próclise (0.722). Desses resultados, portanto, pode-se dizer que esses fatores não se mostraram relevantes, por si sós, para explicar a variação.

A Tabela 3, a seguir, refere-se aos dados provenientes dos jornais de Rio Claro.

**Tabela 3:<sup>5</sup> Número de ocorrências, percentuais e pesos relativos dos pronomes clíticos, de acordo com os gêneros textuais dos jornais de Rio Claro, de 1880 a 1920**

	<i>Próclise</i>			<i>Ênclise</i>			<b>Total</b>	
	N	%	PR	N	%	PR	N	%
<b>Edital</b>	67	62.6	0.743	40	37.4	0.257	107	19.9
<b>Artigo</b>	24	66.7	0.722	12	33.3	0.278	36	6.7
<b>Comentário</b>	14	56	0.622	11	44	0.378	25	4.7
<b>Editorial</b>	19	52.8	0.609	17	47.2	0.391	36	6.7
<b>Crônica</b>	39	45.9	0.469	46	54.1	0.531	85	15.8
<b>Notícia</b>	8	33.3	0.438	16	66.7	0.562	24	4.5
<b>Anúncio</b>	22	30.6	0.382	50	69.4	0.618	72	13.3
<b>Carta do leitor</b>	20	29	0.344	49	71	0.656	69	12.8
<b>Nota</b>	12	26.7	0.334	33	73.3	0.666	45	8.4
<b>Classificado</b>	5	18.5	0.239	22	81.5	0.761	27	5
<b>Aviso</b>	2	16.7	0.195	10	83.3	0.805	12	2.2
<b>Total</b>	232	43.1		306	56.9		538	

Os resultados, para os dados rioclarenses, apontaram, comparando-os aos resultados dos jornais paulistanos, discreta comprovação da hipótese proposta.

Os gêneros *notícia*, *aviso*, *editorial*, *artigo*, *anúncio* e *classificado* apresentaram os resultados esperados. No entanto, admira-se o fato dos gêneros *edital* e *carta do leitor* terem apresentado resultados bastante destoantes entre as ocorrências de pronomes proclíticos e enclíticos, apresentando o *edital* o maior índice à próclise (0.743) e a *carta do leitor* mostrando uma correlação maior com a ênclise (0.656). O gênero *crônica* apresentou alternância equilibrada entre os pronomes em posições pré e pró-verbais.

Quanto aos gêneros *nota* e *comentário*, observaram-se comportamentos distintos. No gênero *nota*, as frequências marcaram diferença expressiva entre os usos dos pronomes em posições pré e pós-verbal e os pesos relativos indicaram maior tendência à ênclise (0.666); acerca do gênero *comentário*, os percentuais, de forma menos destoante, assinalaram maior propensão ao uso do pronome proclítico, confirmando-se essa posição,

<sup>5</sup> Ainda foi encontrado no jornal de 1880, da cidade de Rio Claro, o gênero *resenha*. Por estar presente apenas em um exemplar, foi desconsiderado da análise.

quando atestados os pesos relativos (0.622 – próclise x 0.378 – ênclise). Em termos, a observação feita a esses dois gêneros, quando analisados nos jornais paulistanos, estendeu-se aos jornais da cidade de Rio Claro.

Assim, de acordo com as propensões aos usos da ênclise e da próclise, apresentaram-se dispostos, da forma explicitada abaixo, os gêneros textuais dos referidos jornais:

**Quadro 1: Distribuição geral dos gêneros textuais entre as predominâncias dos usos dos pronomes enclíticos e proclíticos nos jornais de São Paulo e de Rio Claro, de 1880 a 1920**

Ênclise < ----- > Próclise
<b>Jornais da cidade de São Paulo</b>
Classificado > Anúncio > Nota > Notícia > Aviso > Carta do Leitor > Artigo > Resenha ou Edital                      Comentário                      Editorial                      Crítica
<b>Jornais da cidade de Rio Claro</b>
Aviso > Classificado > Nota > Anúncio > Notícia > Editorial > Artigo Carta do leitor                      Crônica                      Comentário                      Edital

Aposta-se, em primeiro plano, para a explicação das realidades divergentes, nos gêneros textuais presentes nos jornais das cidades de São Paulo e de Rio Claro, que, embora haja certos aspectos que possibilitem determinadas caracterizações, e possíveis uniformizações, dos gêneros, há, também, de acordo com as funções e as intenções pré-estabelecidas, admissíveis particularidades entre, por exemplo, dois ou mais textos que materializam o mesmo gênero textual. Essas especificações podem ser, entre outros motivos, reflexos do exemplar em questão analisado. Assim, em segundo plano, para as diferenças no quadro acima assinaladas, tem-se que os jornais escolhidos, nesta investigação, apresentavam traços semelhantes, não idênticos. Convém, então, ressaltar a necessidade de serem consideradas as peculiaridades dos materiais utilizados e, paradoxalmente, a dificuldade encontrada, por parte dos pesquisadores, para o resgate histórico dessas informações, em estudos que lidam com o emprego de textos da mídia escrita.

### Próclise x Ênclise: variáveis linguísticas selecionadas

Quanto ao controle das variáveis independentes linguísticas, fez-se um recorte na amostra de dados analisados, compondo-se, assim, uma subamostra. Dos jornais da cidade de São Paulo, foram extraídos **495** dados. Optou-se por essa quantidade de dados, baseando-se no número de dados presentes nos jornais de Rio Claro, a fim de que o total de dados analisados de ambas as localidades fosse aproximado. Os periódicos de Rio Claro, para esse conjunto de análises, totalizaram **488** ocorrências de clíticos pronominais. Deve-se ressaltar que, para essas observações, os dados provenientes do jornal rioclareense “A Mocidade”, de 1920, foram excluídos, devido aos poucos traços, em termos funcionais, que o assemelhavam aos demais periódicos.

As primeiras rodadas, para os dados oriundos dos jornais das duas localidades, apresentaram *knockouts*, excluídos posteriormente. Assim, no total, para as análises em função das variáveis independentes linguísticas, foram computados **903** usos de pronomes clíticos, divididos entre **441**, referentes aos jornais da cidade de São Paulo, e **462**,

representantes dos jornais interioranos.

O quadro abaixo aponta quais variáveis independentes linguísticas foram selecionadas e eliminadas, das primeiras às últimas, nas análises multidimensionais, referentes aos dados dos jornais das cidades de São Paulo e de Rio Claro. Na sequência, detalham-se as variáveis *presença ou ausência de elemento proclisador na oração* e *verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração*, consideradas, depois de feitos os cruzamentos necessários entre todas as variáveis e uma análise qualitativa dos resultados, os grupos mais relevantes no que concerne ao posicionamento do pronome clítico.

**Quadro 2: Relação das variáveis independentes linguísticas selecionadas e eliminadas, ordenadas segundo os maiores valores de significância, a partir das análises dos dados presentes nos jornais de São Paulo e de Rio Claro**

<b>Variáveis independentes linguísticas selecionadas</b>	
<b>Jornais da cidade de São Paulo</b>	<b>Jornais da cidade de Rio Claro</b>
1ª. Presença ou Ausência de elemento proclisador na oração; 2ª. Formas verbais; 3ª. Verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração; 4ª. Função do clítico, e 5ª. Tipo de verbo, do ponto de vista lógico-semântico.	1ª. Presença ou Ausência de elemento proclisador na oração; 2ª. Formas verbais; 3ª. Verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração, e 4ª. Tipo de clítico.
<b>Variáveis independentes linguísticas eliminadas</b>	
<b>Jornais da cidade de São Paulo</b>	<b>Jornais da cidade de Rio Claro</b>
1ª. Tipo de clítico.	1ª. Função do clítico, e 2ª. Tipo de verbo, do ponto de vista lógico-semântico.

### **Variável *presença ou ausência de elemento proclisador na oração***

A variável *presença/ausência de atrator* se mostrou, como previsto, um relevante controlador da ordem dos clíticos pronominais, selecionada, tanto nos resultados da cidade de São Paulo quanto nos da cidade de Rio Claro, como o grupo de fatores mais significativo para a motivação da colocação pronominal.

O comportamento dos clíticos variou, consideravelmente, segundo a existência, na oração, de um elemento proclisador. Quando não presente, o uso do pronome enclítico, na maior parte das vezes, foi a opção escolhida, revelando maior tendência para a obediência da norma-padrão vigente naquela época.

As tabelas a seguir, 4 e 5, apresentam os resultados, bastante semelhantes, averiguados nos jornais paulistas analisados.

**Tabela 4. Número de ocorrências, percentuais e pesos relativos dos pronomes clíticos, de acordo com a presença, ou ausência, de elemento proclisador na oração nos jornais de São Paulo, de 1880 a 1920**

	Próclise			Ênclise			Total	
	N	%	PR	N	%	PR	N	%
<b>Presença</b>	124	67.8	0.816	59	32.2	0.184	183	41.5
<b>Ausência</b>	13	5	0.258	245	95	0.742	258	58.5
<b>Total</b>	137	31.1		304	68.9		441	

**Tabela 5. Número de ocorrências, percentuais e pesos relativos dos pronomes clíticos, de acordo com a presença, ou ausência, de elemento proclisador na oração nos jornais de Rio Claro, de 1880 a 1920**

	Próclise			Ênclise			Total	
	N	%	PR	N	%	PR	N	%
<b>Presença</b>	179	64.2	0.749	100	35.8	0.251	279	60.4
<b>Ausência</b>	12	6.6	0.159	171	93.4	0.841	183	39.6
<b>Total</b>	191	41.3		271	58.7		462	

No entanto, ressalta-se, nos dados dos jornais de ambas as localidades – São Paulo e Rio Claro – o fato de terem aparecido casos em que se optou pelo uso do pronome enclítico, mesmo com a presença do elemento proclisador – como visto nos exemplos de 03 a 06, em que o elemento atrator do pronome clítico aparece sublinhado.

- (03) [...] *conta a Província de São Paulo fazer da sua independencia o apanagio de sua força e a medida da severa moderação, sisudez, franqueza, lealdade e criterio em que fundará o salutar prestígio a que destina-se a imprensa livre e consciente.* (A PROVÍNCIA DE SÃO PAULO, 1880 – gênero Editorial)
- (04) *E' grande pechincha pois vende-se por pouco mais de um conto de réis tudo [...].* (O ESTADO DE SÃO PAULO, 1910 – gênero Classificado)
- (05) [...] *mas nesse caso faria melhor ficar calado – não sahir-nos ao encontro com rompantes de Hespanhol.* (CORREIO DO OESTE, 1880 – gênero Editorial)
- (06) *Sou filho desta terra onde sou bastante conhecido e até hoje, graças a Deus, ninguém rio-se a minha custa, por espectáculo que eu dêsse em estado interessante pelas ruas.* (O RIO CLARO, 1900 – gênero Carta do Leitor)

Deve-se lembrar que, na história do português, sempre houve a obrigatoriedade, prescrita nos compêndios gramaticais, do uso da próclise em contextos com uma série de elementos – partículas negativas, pronomes indefinidos, interrogativos e relativos, conjunções subordinativas e coordenativas, advérbios, preposições e orações optativas – nomeados **operadores de próclise**.

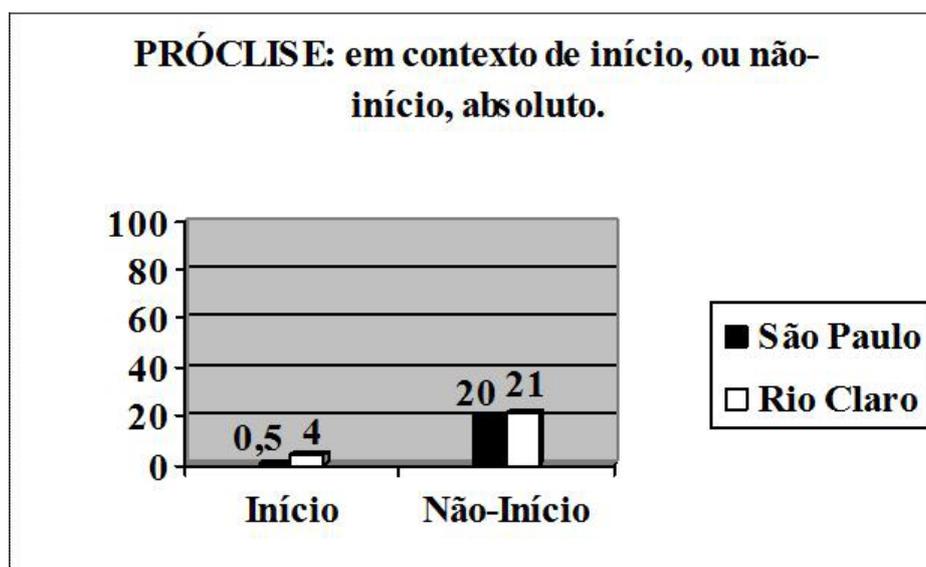
Uma vez que a ênclise era tida, e por uma maioria ainda hoje é conservada, como a posição normal dos pronomes átonos, inclusive quando o sujeito viesse imediatamente antes do verbo, em orações afirmativas, evidenciou-se, também, uso proclítico em contextos em que não havia a presença do elemento proclisador (exemplos 07 a 10, abaixo).

- (07) *A alternativa, no entanto, se dissipa logo no contexto irrefragavel do livro [...].* (O ESTADO DE SÃO PAULO, 1910 – gênero Resenha ou Crítica)

- (08) *O sabonete Rifger, conhecido ha mais de 10 annos,, se impõe como o melhor para o banho [...].* (O ESTADO DE SÃO PAULO, 1915 – gênero Anúncio)
- (09) *O honrado negociante desta praça, sr. Francisco Pereira dos Santos, nos obsequiou com duas garrafas do excellente vinho verde [...].* (O RIO CLARO, 1900 – gênero Nota)
- (10) *Como quer que seja, a recorrente se baseia em um contracto.* (O RIO CLARO, 1905 – gênero Edital)

Desse modo, puderam-se apontar inovações referentes à norma-padrão vigente naquela época e que vigora até os dias atuais.

A correlação entre esta variável – *presença/ausência de atrator* – e a variável *início (ou não-início) absoluto da oração pelo verbo hospedeiro do clítico* foi relevante. Como esperado, os dois fatores interagiram. Naturalmente, os dados que apresentaram presença de atrator se relacionaram com o não-início da oração. Nesse contexto, verificaram-se os índices de próclise em 73%, nos jornais de São Paulo, e 67%, nos jornais de Rio Claro. Nas orações em que havia a ausência do elemento proclisador, os dados puderam aparecer em início, ou não-início, absoluto, como mostrado no gráfico 1, abaixo, em relação aos usos de próclise.



**Gráfico 1.** Frequências de próclise em contextos do verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração, nos jornais de São Paulo e de Rio Claro

Observa-se que o contexto de início absoluto inibiu fortemente a posição pré-verbal, embora não a tenha impedido. Por outro lado, em contexto de não-início, mesmo sem a presença de algum elemento proclisador, mostrou-se um elevado índice do uso da próclise.

**Variável verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração**

Outra importante variável independente linguística selecionada, decisiva para o comportamento dos clíticos pronominais, foi a *posição, inicial ou não, do verbo* ao qual o pronome estava adjungido na oração.

Embora, de modo geral, os resultados tenham se apresentado bastante semelhantes entre os dados dos jornais das cidades de São Paulo e Rio Claro, observa-se, segundo os pesos relativos, que essa variável foi mais significativa nas análises dos textos paulistanos.

Contudo, tanto nos jornais de São Paulo quanto nos jornais de Rio Claro, em início absoluto, verificou-se a realização quase categórica do pronome enclítico, assim como é prescrito nos estudos normativos. Em não-início absoluto, observou-se a possível alternância entre próclise e ênclise. No entanto, embora os pesos relativos, para os resultados de ambas as localidades, tenham mostrado que a tendência em não-início absoluto é a próclise, essa foi mais marcante nos dados provenientes dos jornais de São Paulo, podendo-se reafirmar a ideia apresentada como hipótese de essa variante – a próclise – ser mais significativa nos dados paulistanos.

As tabelas 6 e 7, na sequência, ilustram esses comentários.

**Tabela 6 – Número de ocorrências, percentuais e pesos relativos dos pronomes clíticos, de acordo com o verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração nos jornais de São Paulo, de 1880 a 1920**

	Próclise			Ênclise			Total	
	N	%	PR	N	%	PR	N	%
<b>Não-início</b>	136	56.4	0.952	105	43.6	0.048	241	54.6
<b>Início</b>	1	0.5	0.027	199	99.5	0.973	200	45.4
<b>Total</b>	137	31.1		304	68.9		441	

**Tabela 7 – Número de ocorrências, percentuais e pesos relativos dos pronomes clíticos, de acordo com o verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração nos jornais de Rio Claro, de 1880 a 1920**

	Próclise			Ênclise			Total	
	N	%	PR	N	%	PR	N	%
<b>Não-início</b>	186	56.9	0.613	141	43.1	0.387	327	70.8
<b>Início</b>	5	3.7	0.246	130	96.3	0.754	135	29.2
<b>Total</b>	191	41.3		271	58.7		462	

Quanto a esta variável, ainda se deve destacar o aparecimento, embora em número pequeno, dos pronomes proclíticos em início absoluto nas orações, como exemplificados abaixo.

- (11) *Quer louvando, quer censurando, se esforçará sempre a Província de S. Paulo por ser justa: é este um dever que ella se impôs em virtude de suas condições de folha diaria [...].* (A PROVÍNCIA DE SÃO PAULO, 1880 – gênero Editorial)

- (12) *Querendo que os habitantes desta florescente cidade do Rio Claro e de sua immediações, participem dos preços baratíssimos por que vendem suas fazendas; resolverão lançar mão deste meio de publicidade para que, quando forem para Campinas e tiverem de comprar fazendas finas, ou grossas, miudezas de armario, calçado, chapéus e com especialidade roupas feitas para homem e crianças, assim como chapéus de sol etc., etc., se **dirigão** a Grande Loja do VEADO sita a Rua do Commercion 71 A para certificarem-se da veracidade do que dizem.* (CORREIO DO OESTE, 1880 – gênero Anúncio)
- (13) *Finalmente no parecer da Comissão do Senado se afirmou que a recorrente [...] ficaria im - possibilitada de amortizar o capital e juros do empréstimo que Miguel Rinaldi fez a Camara [...]. Entretanto, se **verifica** pelo doc. n. 5 que a amortização do dito empréstimo está suspensa por cinco annos.* (O RIO CLARO, 1905 – gênero Edital)
- (14) *[...] é preciso que os senhores vereadores attendam as condições preliminares afim de que uma vez approved o contracto, se **comprehenda** a aprovação d'essas mesmas condições.* (O RIO CLARO, 1910 – gênero Edital)
- (15) *Si a estrada de ferro diminue a distancia e o telegrapho quasi a destróe, a caridade a desco - nhece, a **anniquila**, faz com que não exista de forma alguma.* (A SEMANA MILITAR, 1920 – gênero Artigo)
- (16) *Dia a dia, de toda a parte, **nos chegam** novas pessoas que vêm á procura do restabelecimento de sua saude abalada por qualquer moléstia que os atormenta. [...]* (A SEMANA MILITAR, 1920 – gênero Anúncio)

Tais comportamentos, transgredindo os preceitos impostos pelas instituições normativas, apontaram, como já sabido, um caráter relevante, e inovador, da realidade linguística do português brasileiro: o uso da próclise em primeira posição na oração, produto de uma norma linguística brasileira.

## Conclusões

Constatou-se, referente aos condicionamentos extralinguísticos, maior significância da interferência da variável *gênero textual* na colocação pronominal. Deve-se destacar, novamente, no que se refere às diversidades observadas entre os resultados provenientes dos jornais das cidades de São Paulo e de Rio Claro, o fato de haver a possibilidade de um mesmo gênero textual ser representado por textos com aspectos comunicativos diferentes, propiciando, assim, estruturas organizacionais e usos linguísticos não idênticos. Ressalta-se, ainda, que essas diferenças também podem se dever a aspectos de natureza linguística diferenciados, consoante as características particulares de cada contexto no qual o clítico pronominal estava inserido.

Quanto aos gêneros textuais, percebe-se a relevância de um maior aprofundamento em busca de traços que melhor definam as suas naturezas, uma vez que ainda são poucas as discussões, no meio acadêmico, que retratam essa questão, a fim de que sejam somadas, cada vez mais, informações que possam auxiliar na interpretação, mais fidedigna, dos resultados. Contudo, acredita-se, a partir dos resultados apresentados neste artigo, ter mostrado que as relações entre variação, mudança linguística e gêneros textuais devem ser consideradas na investigação dos processos ocorridos na história de uma língua.

Considerando-se a subamostra composta para a análise das variáveis independentes linguísticas, de um modo geral, averiguou-se a predominância da ênclise, com

percentuais de 68.9%, nos jornais de São Paulo, e 58.7%, nos jornais de Rio Claro, não se comprovando a hipótese de que nos dados extraídos dos jornais paulistanos prevaleceria, através do uso mais acentuado da próclise, um caráter mais inovador; ao mesmo tempo em que, nos resultados de Rio Claro, diante de um perfil mais conservador da localidade, predominaria a colocação do pronome enclítico. No entanto, quando observados os pesos relativos referentes à variável *verbo hospedeiro do pronome clítico em início, ou não-início, absoluto na oração*, mostrou-se mais acentuadamente a tendência para a próclise (0.952), em não-início absoluto, nos dados provenientes da cidade de São Paulo, podendo-se restabelecer, novamente, a ideia apresentada como hipótese de essa variante ser mais significativa nos dados paulistanos.

Ainda que o decorrer do século XIX, principalmente as suas últimas décadas, e o início do século XX tenham sido marcados pela busca por uma identificação com o modelo europeu, adotando-se também a norma-padrão do português de além-mar para os mais diversos usos linguísticos, revelou-se através dos dados uma das características de qualquer língua, a de que é infinitamente variada, não existindo uma norma única, mas sim uma pluralidade de normas. Através dos dados apresentados – por exemplo, ainda que eu número bastante restrito, aqueles com o pronome átono em início absoluto na oração –, atestaram-se características próprias do PB, traços inovadores.

A partir do levantamento de todas as variáveis – extralinguísticas e linguísticas – discutidas no correr deste estudo, conclui-se, portanto, que a presente investigação pôde contribuir com a descrição da variedade do Português Paulista, assinalando, concomitantemente, um avanço em relação aos estudos já existentes sobre esta mesma temática – a posição dos clíticos pronominais –, principalmente no que concerne à análise da variável independente extralinguística *gênero textual*, que, de acordo com os propósitos esclarecidos neste estudo, ainda é pouco examinada, e a relevância desses trabalhos e de tantos outros que, possivelmente, virão a se realizar, acerca desse rendoso assunto.

## REFERÊNCIAS

BIAZOLLI, C. C. *Clíticos pronominais no português de São Paulo: 1880 a 1920 – uma análise sócio-histórico-linguística*. 2010. 230 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Unesp, Araraquara.

CÂMARA Jr., J. M. *Dicionário de Linguística e Gramática*. 12. ed. Petrópolis: Vozes, 1985.

CYRINO, S. M. L. Observações sobre a Mudança Diacrônica no Português do Brasil: objeto nulo e clíticos. In: ROBERTS, I; KATO, M. (Org.). *Português Brasileiro: uma viagem diacrônica*. 2 ed. Campinas: Editora da Unicamp, 1996. p. 163-175.

DUARTE, M. E. L. *Variação e sintaxe: clítico acusativo, pronome lexical e categoria vazia no Português do Brasil*. 1986. 73 f. Dissertação (Mestrado em Ciências – Linguística aplicada ao ensino de línguas) – Pontifícia Universidade Católica de São Paulo, São Paulo.

LABOV, W. *Padrões sociolingüísticos*. São Paulo: Parábola, 2008 [1972].

\_\_\_\_\_. *Principles of Linguistic Change*. v. 2: Social factors. Cambridge: Blackwell, 2001.

\_\_\_\_\_. *Principles of Linguistic Change*. v. 1: Internal factors. Cambridge: Blackwell, 1994.

MATTHEWS, P. H. *Oxford Concise Dictionary of Linguistics*. Oxford: Oxford University Press, 1997.

PAGOTTO, E. G. *A posição dos clíticos em português: um estudo diacrônico*. 1992. 168f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Unicamp, Campinas.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

VIEIRA, S. R. *Colocação pronominal nas variedades européia, brasileira e moçambicana: para a definição da natureza do clítico em português*. 2002. 441 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

WEINREICH, V.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança lingüística*. São Paulo: Parábola, 2006 [1968].

# O comportamento prosódico dos monossílabos no Português arcaico por meio das *Cantigas de Santa Maria*

(The prosodic behavior of monosyllables in Medieval Portuguese through the *Cantigas de Santa Maria*)

Daniel Soares da Costa<sup>1</sup>

<sup>1</sup>Faculdade de Ciências e Letras de Araraquara (FCL-UNESP)

dan.fono@gmail.com

**Abstract:** This paper aims to conduct a study on the prosodic behavior of monosyllables in Medieval Portuguese. This research is based on an innovative methodology that observes the relation between musical and linguistic prominences and its contributions for data collection regarding the prosodic structure of the language in study. There are in Medieval Portuguese heavy (tonic) monosyllables and light (atonic) monosyllables with diverse syllabic structure. The onset is irrelevant for the syllabic weight. Monosyllables with a consonant in the coda are heavy unless this consonant represents an inflexion element. Monosyllables with no consonant in the coda are normally light but there are also heavy monosyllables with no consonant in the coda.

**Keywords:** monosyllables; prosody; Medieval Portuguese; *Cantigas de Santa Maria*.

**Resumo:** O objetivo deste trabalho é fazer um estudo sobre o comportamento prosódico dos monossílabos no Português Arcaico. Esta pesquisa tem como base uma metodologia que observa a relação entre proeminências musicais e linguísticas e suas contribuições para a coleta de dados sobre a estrutura prosódica da língua em estudo. Há, no Português Arcaico, monossílabos pesados (tônicos) e monossílabos leves (átomos) com estruturas silábicas diversas. O *onset* é irrelevante para o peso silábico. Os monossílabos com uma consoante na coda são pesados, a menos que essa consoante represente um elemento de flexão. Os monossílabos sem nenhuma consoante na coda são normalmente leves, porém, também há nessa língua, monossílabos pesados sem travamento silábico.

**Palavras-chave:** monossílabos; prosódia; português arcaico; *Cantigas de Santa Maria*.

## Introdução

O objetivo deste trabalho é apresentar uma análise sobre o comportamento prosódico dos monossílabos do português arcaico (de agora em diante, PA). A coleta de dados foi feita por meio de uma metodologia inovadora baseada na observação das proeminências musicais da melodia de textos poéticos musicados, na observação das proeminências linguísticas do texto desses poemas, junto com a observação da sua estrutura métrico-poética.

A proposta centra-se no mapeamento das coincidências entre o tempo mais forte do compasso musical (o primeiro tempo) e a sílaba do texto do poema a ele anexada. Dessa forma, é possível verificar em que proporção a proeminência musical marca a acentuação de uma sílaba tônica no texto ou exerce alguma influência sobre ela, fazendo com que o acento mude de posição.

O *corpus* se constitui de um recorte das cem primeiras *Cantigas de Santa Maria* (de agora em diante, CSM), tomadas a partir da edição feita por Anglés (1943), a qual traz

as pautas musicais transcritas para a notação musical atual. As *Cantigas de Santa Maria* são um conjunto de 420 composições musicais com textos poéticos metrificados, escritos em galego-português (fins do século XII até meados do século XIV).

Como resultado, obtivemos um percentual de 63,32% de coincidências entre as proeminências musicais e linguísticas, ou seja, a proeminência musical marca preferencialmente uma sílaba tônica no nível textual, servindo, assim, como uma ferramenta auxiliar na busca de pistas sobre a prosódia de línguas que não possuem mais falantes.

Em relação aos monossílabos especificamente, temos um percentual de 31,3% de coincidências entre proeminências musicais e sílabas desse tipo, sendo 23,56% de coincidências entre a proeminência musical e monossílabos tônicos e 7,74% de coincidências entre a proeminência musical e monossílabos átonos. Em termos de quantidade de palavras, foi levantado um total de 38.018 palavras, das quais 18.633 (49%) são monossílabos, sendo 12.846 (33,78%) monossílabos tônicos e 5.787 (15,22%) monossílabos átonos.

A metodologia utilizada neste trabalho, bem como os dados aqui apresentados, são frutos da tese de doutorado de Costa (2010a), intitulada *A interface música e linguística como instrumental metodológico para o estudo da prosódia do português arcaico*. Trata-se de um extenso trabalho de pesquisa, desenvolvido ao longo de quatro anos, que rendeu (e ainda tem rendido) frutos desde então, fazendo-se necessária, aqui, a referência aos trabalhos publicados ao longo do desenvolvimento dessa tese, tanto pelo autor quanto pela sua orientadora. Sobre esse tema, temos os trabalhos de Massini-Cagliari (2008a; 2008b; 2008c; 2008d; 2009; 2010) e Costa (2007<sup>1</sup>; 2008; 2009; 2010b, 2011).

## Síntese da metodologia

A metodologia utilizada para a coleta de dados baseia-se na observação de coincidências entre proeminências musicais e linguísticas em textos poéticos musicados, além da observação da sua estrutura métrica.

Como dissemos anteriormente, partimos das transcrições das pautas musicais para a notação musical atual feitas por Anglés (1943), já que a música medieval não apresenta, de maneira visível, a divisão dos compassos musicais (o que não quer dizer que essa divisão não exista). Essa edição de Anglés foi feita por meio de um trabalho de musicologia bastante sério e confiável, sendo a edição mais conhecida e utilizada até o momento.

A partitura musical, nas CSM, vem anexada apenas ao refrão e à primeira estrofe de cada cantiga, o que sugere que essa melodia serve para o canto de todas as demais estrofes da mesma cantiga, uma vez que a estrutura métrica estipulada na primeira estrofe é rigorosamente seguida em todas as demais estrofes.

A organização do *corpus* se dá por meio da elaboração de fichas de análise para cada uma das cantigas selecionadas. As fichas são elaboradas da seguinte maneira: depois de digitalizada a versão da cantiga a ser analisada, recorta-se cada uma das linhas da sua partitura; em seguida, anexa-se o texto de todas as estrofes à sua respectiva linha melódica. Dessa forma, podemos verificar as sílabas e as palavras em que a proeminência musical recai.

---

<sup>1</sup> Trabalho apresentado em forma de comunicação no 55º *Seminário do GEL*, com o título “Da notação musical às proeminências da fala: uma proposta metodológica para o estudo do ritmo linguístico das Cantigas de Santa Maria de Afonso X”, no ano de 2007.

Recortadas todas as linhas das pautas musicais e anexadas todas as partes do texto da cantiga à sua respectiva melodia, atribuímos, para cada tipo de coincidência entre a proeminência musical e as sílabas, um tipo de marcação, que poderá ser feita por meio da atribuição de cores diferentes, por exemplo. Dessa forma, se a proeminência musical marcar uma sílaba tônica de uma palavra com mais de uma sílaba no nível textual, essa sílaba receberá a cor vermelha; se marcar um monossílabo tônico, receberá a cor azul; se for um monossílabo átono, receberá a cor verde; se for pretônica, laranja; se for postônica final, receberá a cor rosa; e, por fim, se for postônica não-final, receberá a cor roxa.

A Figura 1, abaixo, traz um exemplo de como fica o texto das estrofes depois que as sílabas que coincidem com proeminências musicais são marcadas, seguindo o esquema de cores estipulado.<sup>2</sup>

Ca e-la et sseu Fi - llo son jun - ta - dos d'a-mo[r],<sup>1)</sup> que par - ti - dos per  
 Ca e-la et sseu Fi- llo son jun- ta- dos d'a- mor, que par- ti- dos per  
 Da-quest' a- vê- o, tem- pos sson pa- ssa- dos gran-des, que o Con- de de  
 Poi- los mon- ges fo- ron en- de ti- ra- dos, mui ma- as con- pa- nnas se  
 Mas hũ- a mo- ller, que por seus pe- ca- dos en- tra- ra na ci- gre- ja,  
 O ta- fur, quan- d' es- to vya, con y- ra- dos e- llos a ca- tou, e co-  
 E deu no Fi- llo, que am- bos al- ça- dos tí- i- a seus bra- ços en  
 May- o- res mi- ra- gres ou- v' y mos- tra- dos Deus, que san- gui cra- ro fez  
 E de- mais ou- ve os o- llos tor- na- dos tan bra- vos, que quan- tos a  
 Ou- tros dous ta- fu- res de- mo- ni- a- dos ouv' y, por- que fô- ran a-  
 O Con- de, quan- do' est' o- yu, con ar- ma- dos ca- va- lei- ros vê- o e  
 Pois es- to di- sse, per- nas e cos- ta- dos e a ca- be-

Figura 1 – Quarta linha da transcrição da CSM 38 (ANGLÉS, 1943, p. 46, “parte musical”) com as demais estrofes anexadas e as sílabas que aparecem em posição de proeminência musical coloridas

Terminada a elaboração da ficha de análise, parte-se para a contagem (que pode ser feita muito facilmente por meio dos recursos do próprio programa Microsoft Office Word) dos tipos de coincidências entre as proeminências nos níveis musical e linguístico, e a criação de um quadro quantitativo para a cantiga analisada.

## Dados quantitativos e análise

A Tabela 1, abaixo, mostra a quantificação geral das coincidências entre proeminências musicais e linguísticas na análise das cem primeiras CSM através da metodologia aqui adotada. Pode-se observar que se trata de um levantamento bastante grande, ultrapassando trinta e duas mil coincidências entre proeminências musicais e sílabas textuais, distribuídas de acordo com a pauta acentual das sílabas.

Somando o percentual de coincidências entre a proeminência musical e sílabas tônicas de palavras polissilábicas com o percentual de coincidências entre proeminências

<sup>2</sup> É importante dizer que as sílabas que aparecem no refrão são contadas apenas uma vez, já que ele se repete com o mesmo texto e a mesma música após cada estrofe.

musicais e monossílabos tônicos, temos um total de 63,32%, confirmando a nossa expectativa inicial de que a proeminência musical tende a marcar sílabas tônicas nas palavras dos textos.

**Tabela 1. Quantificação geral de coincidências entre proeminências musicais e linguísticas**

Coincidências com tônica	12997	39,76%
Coincidências com monossílaboônico	7703	23,56%
Coincidências com monossílabo átono	2531	7,74%
Coincidências com pretônica	4221	12,91%
Coincidências com postônica final	5189	15,87%
Coincidências com postônica não-final	44	0,13%
Total de proeminências	32685	100%

A Tabela 2, a seguir, mostra a quantificação dos dados relativos às pautas acentuais das palavras encontradas no *corpus*. Chama a atenção o número total de monossílabos (18633) que corresponde a praticamente metade das palavras encontradas. Isso demonstra a alta frequência do emprego dessas palavras nessa língua.

**Tabela 2. Palavras de acordo com a pauta acentual linguística**

Oxítonas	4733	12,44%
Paroxítonas	14539	38,24%
Proparoxítonas	113	0,29%
Monossílabos tônicos	12846	33,78%
Monossílabos átonos	5787	15,22%
Total	38018	100%

Vimos que as proeminências no nível musical (notas presentes no tempo mais forte do compasso musical, o primeiro tempo) marcam preferencialmente proeminências no nível linguístico (sílabas tônicas de palavras com mais de uma sílaba ou monossílaboônico).

Partindo dessa constatação e focalizando a questão dos monossílabos que apareceram na análise do PA referente às cem primeiras CSM, poderíamos pensar que monossílabos considerados tônicos (por outros estudos, como em Cunha (1961), por exemplo; ou mesmo monossílabos que já apareceram em posição de final de verso, estabelecendo rima com palavras oxítonas, o que indicaria que são tônicos) têm preferência em aparecerem em posição de proeminência musical, sendo mais recorrentes nessa posição do que fora dela. Já os monossílabos considerados átonos seriam mais recorrentes fora da posição de proeminência musical do que nela.

A fim de constatar se esse pensamento realmente procede dentro do *corpus* analisado, além da quantificação dos monossílabos que apareceram em posição de proeminência musical, fizemos também a quantificação das vezes em que esses mesmos monossílabos apareceram fora da posição de proeminência musical de modo a comparar os percentuais. Por exemplo, o monossílabo “ca”, que é consideradoônico por Cunha (1961), apareceu no *corpus* num total de 198 vezes, aparecendo 134 vezes em posição de proeminência musical (aproximadamente 67%) e 64 vezes fora da posição de proeminência musical (aproximadamente 33%). Esse tipo de comparação entre os percentuais corrobora a afirmação

de Cunha (1961) de que esse monossílabo é tônico, uma vez que segue a preferência de sílabas tônicas aparecerem marcadas com proeminência no nível musical.

Vários outros monossílabos considerados tônicos apareceram em posição de proeminência musical com um percentual de mais ou menos 60%, que é próximo ao percentual geral de coincidências entre as proeminências musicais e sílabas tônicas do texto das cantigas.

Na Tabela 3, a seguir, são apontados alguns desses casos.

**Tabela 3. Relação de alguns monossílabos tônicos em posição de proeminência musical e fora dela**

<b>Monossílabo</b>	<b>nº total de vezes que apareceu</b>	<b>nº de vezes que apareceu em posição de proeminência musical (percentual)</b>	<b>nº de vezes que apareceu fora da posição de proeminência musical (percentual)</b>
ai	18	11(61,11%)	7 (38,88%)
al	38	29 (76,31%)	9 (23,68%)
ás	14	9 (64,28%)	5 (35,71%)
az	7	6 (85,71%)	1 (14,28%)
bel	4	4 (100%)	0 (0%)
ca	198	134 (67,67%)	64 (32,32%)
cas	5	3 (60%)	2 (40%)
cras	5	3 (60%)	2 (40%)
cruz	10	10 (100%)	0 (0%)
dá	13	10 (76,92%)	3 (23,07%)
dar	48	34 (70,83%)	14 (29,16%)
Deus	220	130 (59,09%)	90 (40,9%)
don	14	12 (85,71%)	2 (14,28%)
dous	14	9 (64,29%)	5 (35,71%)
e	2034	1411(69,37%)	623 (30,63%)
ei	23	16 (69,57%)	7 (30,43%)
et	77	55 (71,43%)	22 (28,57%)
eu	125	70 (56%)	55 (44%)
ey	9	7 (77,78%)	2 (22,22%)
fal	4	4 (100%)	0 (0%)
fas	4	4 (100%)	0 (0%)
faz	36	25 (69,44%)	11 (30,56%)
fe	4	4 (100%)	0 (0%)
fez	191	118 (61,78%)	73 (38,22%)
fis	4	4 (100%)	0 (0%)
fiz	6	5 (83,33%)	1 (16,67%)
for	14	11 (78,57%)	3 (21,43%)
foy	33	20 (60,61%)	13 (39,39%)
ir	37	24 (64,86%)	13 (35,14%)
jaz	20	18 (90%)	2 (10%)
luz	5	5 (100%)	0 (0%)
mais	217	144 (66,36%)	73 (33,64%)

mal	114	76 (66,67%)	38 (33,33%)
mas	170	141 (82,94%)	29 (17,06%)
mays	4	3 (75%)	1 (25%)
mes	8	6 (75%)	2 (25%)
meus	20	17 (85%)	3 (15%)
mil	10	7 (70%)	3 (30%)
min	5	4 (80%)	1 (20%)
nen	123	81 (65,85%)	42 (34,15%)
ou	24	15 (62,5%)	9 (37,5%)
pan	15	9 (60%)	6 (40%)
par	32	26 (81,25%)	6 (18,75%)
paz	11	11 (100%)	0 (0%)
pez	6	6 (100%)	0 (0%)
poi	37	24 (64,86%)	13 (35,14%)
poy	3	2 (66,67%)	1 (33,33%)
pran	6	6 (100%)	0 (0%)
praz	19	17 (89,47%)	2 (10,53%)
pres	9	6 (66,67%)	3 (33,33%)
prez	18	18 (100%)	0 (0%)
prol	4	3 (75%)	1 (25%)
qual	12	8 (66,67%)	4 (33,33%)
que	1836	1078 (58,71%)	758 (41,29%)
quen	65	38 (58,46%)	27 (41,54%)
quer	34	22 (64,71%)	12 (35,29%)
quis	87	52 (59,77%)	35 (40,23%)
rei	25	15 (60%)	10 (40%)
ren	73	51 (69,86%)	22 (30,14%)
rey	16	10 (62,5%)	6 (37,5%)
sal	7	4 (57,14%)	3 (42,86%)
san	35	20 (57,14%)	15 (42,86%)
sen	176	102 (57,95%)	74 (42,05%)
seus	105	65 (61,9%)	40 (38,1%)
si	14	8 (57,14%)	6 (42,86%)
so	6	5 (83,33%)	1 (16,67%)
son	29	22 (75,86%)	7 (24,14%)
ssi	14	9 (64,29%)	5 (35,71%)
sson	5	4 (80%)	1 (20%)
sur	4	4 (100%)	0 (0%)
ten	13	11 (84,62%)	2 (15,38%)
teus	12	12 (100%)	0 (0%)
tol	3	2 (66,67%)	1 (33,33%)
tres	13	9 (69,23%)	4 (30,77%)
u	159	106 (66,67%)	53 (33,33%)
val	14	12 (85,71%)	2 (14,29%)

vay	9	7 (77,78%)	2 (22,22%)
ven	12	10 (83,33)	2 (16,67%)
vez	30	29 (96,67%)	1 (3,33%)
vi	10	8 (80%)	2 (20%)
vil	4	4 (100%)	0 (0%)
viu	79	46 (58,23%)	33 (41,77%)
vos	187	107 (57,22%)	80 (42,78%)
voz	14	8 (57,14%)	6 (42,86%)
vyu	21	18 (85,71%)	3 (14,29%)
y	126	75 (59,52%)	51 (40,48%)
yr	5	3 (60%)	2 (40%)

Em alguns casos, os dados se mostraram insuficientes para sustentar uma afirmação a respeito da tonicidade dos monossílabos. São casos em que o monossílabo apareceu uma única vez no *corpus* todo ou apareceu muito poucas vezes, não sendo possível se estabelecer a comparação dos percentuais.

No entanto, de uma maneira geral, pudemos perceber que, em relação aos monossílabos considerados tônicos, existe uma preferência para que ocorram mais em posição de proeminência musical do que fora dessa posição, conforme o que podemos observar na tabela 4, abaixo, em que temos a quantificação geral desses monossílabos nessas duas posições, apresentando um percentual de 59,56% para ocorrência desses monossílabos em posição de proeminência musical contra 40,04% para a ocorrência desses monossílabos fora dessa posição.

**Tabela 4. Quantificação geral dos monossílabos tônicos em relação à sua ocorrência em posição de proeminência musical ou fora dela**

<b>Monossílabos tônicos em posição de proeminência musical</b>	7703	59,96%
<b>Monossílabos tônicos fora da posição de proeminência musical</b>	5143	40,04%
<b>Total de monossílabos tônicos</b>	12846	100%

Já em relação aos monossílabos considerados átonos, esperava-se um comportamento contrário ao comportamento dos tônicos em relação às proeminências musicais, isto é, a expectativa era a de os monossílabos considerados átonos aparecessem mais fora da posição de proeminência musical do que nela.

Em alguns casos pudemos verificar essa possibilidade, conforme o que apresentamos na Tabela 5, a seguir.

**Tabela 5. Relação de alguns monossílabos átonos em posição de proeminência musical e fora dela**

<b>Monossílabo</b>	<b>nº total de vezes que apareceu</b>	<b>nº de vezes que apareceu em posição de proeminência musical (percentual)</b>	<b>nº de vezes que apareceu fora da posição de proeminência musical (percentual)</b>
che	17	2 (11,76%)	15 (88,23%)
cho	4	0 (0%)	4 (100%)
de	1000	405 (40,5%)	595 (59,5%)

dos	83	36 (43,37%)	47 (56,62%)
la	40	17 (42,5%)	23 (57,5%)
las	11	4 (36,36%)	7 (63,64%)
le	9	0 (0%)	9 (100%)
lle	454	133 (29,30%)	321 (70,7%)
los	22	8 (36,36%)	14 (63,64%)
me	134	39 (29,1%)	95 (70,9%)
no	81	31 (38,27%)	50 (61,73%)
nos	145	64 (44,14%)	81 (55,86%)
sa	115	50 (43,48%)	65 (56,52%)
sse	158	29 (18,35%)	129 (81,65%)
te	72	29 (40,28%)	43 (59,72%)
xe	20	6 (30%)	14 (70%)
xo	3	1 (33,33%)	2 (66,67%)

Novamente constatamos casos de monossílabos que apareceram muito poucas vezes no *corpus* todo, dificultando, assim, uma sustentação mais forte a respeito de sua tonicidade ou atonicidade.

Podemos ressaltar que a respeito dos monossílabos considerados átonos apenas oito deles apareceram em posição de proeminência musical, com um percentual próximo aos sessenta por cento, que é considerado aqui um número base para a análise, uma vez que representa o número de coincidências entre proeminências musicais e linguísticas no *corpus* como um todo, representando, também, as chances de a sílaba ser considerada tônica, ou servindo como um reforço para a ratificação da sua tonicidade.

Também vale a pena observar que apenas um dos monossílabos considerados átonos só apareceu em posição de proeminência musical, não aparecendo nenhuma vez fora dessa posição.

Em síntese, podemos afirmar, a respeito dos monossílabos considerados átonos, que a maioria deles apareceu, na análise das cem primeiras CSM, com maior frequência fora da posição de proeminência musical, ou manteve-se um equilíbrio entre os percentuais das duas posições (em posição de proeminência musical ou fora dela); além disso, uma minoria apareceu na posição de proeminência musical com um percentual maior ou próximo a sessenta por cento e apenas um desses monossílabos apareceu somente em posição de proeminência musical.

A Tabela 6, a seguir, nos mostra que, em relação aos monossílabos átonos, na sua totalidade dentro do *corpus*, temos uma preferência para que apareçam fora da posição de proeminência musical, num percentual de 56,26%.

**Tabela 6. Quantificação geral dos monossílabos átonos em relação à sua ocorrência em posição de proeminência musical ou fora dela**

<b>Monossílabos átonos em posição de proeminência musical</b>	2531	43,73%
<b>Monossílabos átonos fora da posição de proeminência musical</b>	3256	56,26%
<b>Total de monossílabos átonos</b>	5787	100%

Partindo desse tipo de análise, podemos, também, fazer algumas considerações sobre o comportamento dos clíticos no PA.

Bisol (1996, p. 251) aponta dois tipos de clíticos: aqueles que se comportam como uma só unidade fonológica junto da palavra adjacente; e aqueles que possuem certa independência fonológica e que se submetem às mesmas regras da palavra fonológica. Como base para essa afirmação, a autora aponta o fato de os clíticos no português brasileiro se submeterem à regra de neutralização da vogal átona final independentemente de serem proclíticos ou enclíticos, como em “me leve”, que se realiza [mi leve] e “o leque”, que se realiza [u leki].

Em relação aos clíticos do português arcaico que apareceram no *corpus*, pudemos verificar certo equilíbrio nos percentuais de ocorrências desses monossílabos em posição de proeminência musical e fora dela em alguns casos; ou um percentual mais alto de ocorrências desses monossílabos fora da posição de proeminência musical do que nela. Vejamos alguns exemplos da ocorrência desses clíticos na Tabela 7, a seguir.

**Tabela 7. Ocorrência de clíticos no *corpus***

<b>Monossílabo</b>	<b>nº total de vezes que apareceu</b>	<b>nº de vezes que apareceu em posição de proeminência musical (percentual)</b>	<b>nº de vezes que apareceu fora da posição de proeminência musical (percentual)</b>
a	1244	632 (50,80%)	612 (49,19%)
as	86	43 (50%)	43 (50%)
che	17	2 (11,76%)	15 (88,23%)
la	40	17 (42,5%)	23 (57,5%)
las	11	4 (36,36%)	7 (63,64%)
lla	21	12 (57,14%)	9 (42,86%)
llas	5	3 (60%)	2 (40%)
lle	454	133 (29,30%)	321 (70,7%)
lles	67	34 (50,75%)	33 (49,25%)
llo	22	11 (50%)	11 (50%)
llos	2	1 (50%)	1 (50%)
lo	72	34 (47,22%)	38 (52,78%)
los	22	8 (36,36%)	14 (63,64%)
me	134	39 (29,1%)	95 (70,9%)
na	111	59 (53,15%)	52 (46,85%)
nas	10	5 (50%)	5 (50%)
no	81	31 (38,27%)	50 (61,73%)
nos	145	64 (44,14%)	81 (55,86%)
o	820	403 (49,15%)	417 (50,85%)
se	182	74 (40,66%)	108 (59,34%)
sse	158	29 (18,35%)	129 (81,65%)
te	72	29 (40,28%)	43 (59,72%)
xe	20	6 (30%)	14 (70%)
xo	3	1 (33,33%)	2 (66,67%)

Observando o comportamento desses clíticos, podemos afirmar que há uma tendência maior para que essas palavras sejam consideradas átonas, uma vez que, na maior parte dos casos, tais palavras apareceram mais fora da posição de proeminência musical do que nela. No entanto, devido ao equilíbrio apresentado nos percentuais de ocorrência de uma parte considerável desses monossílabos, tanto em posição de proeminência musical quanto fora dela, podemos concluir que, apesar de não termos a certeza de que os clíticos do PA tinham caráter tônico, é possível afirmar que há indícios de sua relativa independência prosódica, uma vez que, em certos contextos, eles podem assumir proeminência, o que se aproxima do apontamento feito por Bisol (1996, p. 251), em relação ao português brasileiro.

Para finalizar a análise do comportamento prosódico dos monossílabos no PA, podemos dizer que há, nessa língua, monossílabos pesados, constituídos de sílaba travada, tais como *bel*, *cruz*, *Deus*, *mal* etc.; monossílabos que não são constituídos de sílaba travada, mas que são comprovadamente tônicos (de acordo com a afirmação de estudiosos medievalistas e a análise do percentual de incidência de proeminências musicais, sobre tais formas, que confirma a tonicidade das mesmas - cf. Tabela 3), tais como *ca*, *que*, *si*, *u*, entre outros; e, por fim, monossílabos leves, sem travamento silábico ou travados por elemento representativo de flexão, o que não os torna pesados, tais como *de*, *dos*, *la*, *las*, *te*, *nos* etc.

## Conclusão

Por meio do que foi apresentado neste trabalho, constatamos que a metodologia desenvolvida por Costa (2010a) mostra-se bastante eficaz no trabalho de coleta de dados relativos à prosódia do PA, já que permite a localização de acentos de palavras em versos de poemas metrificados e musicados, por meio da análise das proeminências musicais das melodias que acompanham esses poemas.

Em relação à análise dos monossílabos, foco deste trabalho, pudemos constatar que essas palavras constituem praticamente a metade das palavras analisadas. Também vimos que os monossílabos considerados tônicos têm preferência por aparecerem mais em posição de proeminência musical do que fora dela, ao passo que, com os monossílabos átonos, essa preferência se inverte.

Também foi possível fazer alguns apontamentos relativos ao comportamento de clíticos no PA. Esse tipo de monossílabo apresentou um equilíbrio considerável relativo ao seu aparecimento em posição de proeminência musical e fora dela, o que nos permite concluir que, apesar de não termos a certeza de que os clíticos do PA tinham caráter tônico, é possível afirmar que há indícios de sua relativa independência prosódica, uma vez que, em certos contextos, eles podem assumir proeminência.

O PA possui monossílabos pesados (tônicos), constituídos de sílaba travada ou não, e monossílabos leves (átonos), sem travamento silábico ou travados por elemento representativo de flexão, que não os torna pesados.

## REFERÊNCIAS

- ANGLÉS, H. *La música de las Cantigas de Santa María del Rey Alfonso el sabio*: facsímil, transcripción y estudio crítico por Higinio Anglés. Barcelona: Diputación Provincial de Barcelona; Biblioteca Central; Publicaciones de la Sección de Música, 1943-1964.
- BISOL, L. Constituintes prosódicos. In: BISOL, L. (Org.). *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1996. p. 247-261.
- COSTA, D. S. Proparoxítonas no português arcaico: uma análise por meio da música e do texto das Cantigas de Santa Maria. *Estudos Linguísticos*, São Paulo, n. 40, v. 1, p. 773-785, 2011.
- \_\_\_\_\_. *A interface Música e linguística como instrumental metodológico para o estudo da prosódia do português arcaico*. 2010a. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Unesp, Araraquara.
- \_\_\_\_\_. Estudo sobre o acento secundário e a tonicidade dos monossílabos em português arcaico por meio da música e da métrica das *Cantigas de Santa Maria*. *Estudos Linguísticos*, São Paulo, n. 39, v. 1, p. 21-34, 2010b.
- \_\_\_\_\_. Música e texto: uma metodologia para o estudo da prosódia de línguas mortas. *Estudos Linguísticos*, São Paulo, n. 38, v. 2, p. 211-221, 2009.
- \_\_\_\_\_. Música e linguística: uma metodologia para estudos da prosódia do português arcaico. In: SIMCAM4. 2008. São Paulo. *Anais eletrônicos...* Disponível em: <[http://www.fflch.usp.br/dl/simcam4/downloads\\_anais/SIMCAM4\\_Daniel\\_Costa.pdf](http://www.fflch.usp.br/dl/simcam4/downloads_anais/SIMCAM4_Daniel_Costa.pdf)>. Acesso em: 28 maio 2009.
- \_\_\_\_\_. Da notação musical às proeminências da fala: uma proposta metodológica para o estudo do ritmo linguístico das *Cantigas de Santa Maria* de Afonso X. Comunicação apresentada no 55º Seminário do Gel, Franca, Unifran, 2007.
- CUNHA, C. F. *Estudos de poética trovadoresca - versificação e ecdótica*. Rio de Janeiro: MEC/Instituto Nacional do Livro, 1961.
- MASSINI-CAGLIARI, G. From Musical Cadences to Linguistic Prosody: How to Abstract Speech Rhythm of the Past. In: PARTRIDGE, John (Org.). *Interfaces in language*. 1. ed. Newcastle upon Tyne: Cambridge Scholars, 2010. v. único, p. 113-134.
- \_\_\_\_\_. A notação musical como fonte para o estudo do ritmo linguístico no período trovadoresco do português: as cantigas de amor de D. Dinis. In: MASSINI-CAGLIARI, G.; MUNIZ, M. R. C.; SODRÉ, P. R. (Org.). *Série Estudos Medievais 2 Fontes*. 1. ed. Araraquara: GT de Estudos Medievais da Anpoll, 2009. v. 1, p. 63-79.
- \_\_\_\_\_. Do ritmo musical para o ritmo linguístico, a partir da análise de uma *Cantiga de Santa Maria* de Afonso X. In: SIMCAM4. 2008a. São Paulo. *Anais eletrônicos...* Disponível em: <[http://www.fflch.usp.br/dl/simcam4/downloads\\_anais/SIMCAM4\\_Gladis\\_Cagliari.pdf](http://www.fflch.usp.br/dl/simcam4/downloads_anais/SIMCAM4_Gladis_Cagliari.pdf)>. Acesso em: 28 maio 2009.

\_\_\_\_\_. Interface Fonologia-Poesia-Música: Uma análise do ritmo lingüístico do Português Arcaico, a partir da notação musical das *Cantigas de Santa Maria*. *Estudos Lingüísticos*, São José do Rio Preto, n. 37, v. 1, p. 9-20, 2008b. Disponível em: <[http://www.gel.org.br/estudoslinguisticos/volumes/37/EL\\_V37N1\\_01.pdf](http://www.gel.org.br/estudoslinguisticos/volumes/37/EL_V37N1_01.pdf)>. Acesso em: 28 maio 2009.

\_\_\_\_\_. Contribuição para a análise do ritmo lingüístico das cantigas profanas e religiosas a partir de uma interface Música-Lingüística. Comunicação apresentada no IX Congresso da Associação Internacional de Lusitanistas. Funchal, Madeira, Portugal, Universidade da Madeira, de 4 a 9 de agosto de 2008c.

\_\_\_\_\_. Das cadências musicais para o ritmo lingüístico: Uma análise do ritmo do Português Arcaico, a partir da notação musical das Cantigas de Santa Maria. *Revista da ABRALIN*, Natal, v. 7, n. 1, p. 9-26, jan./jun. 2008d. (ISSN 1678-1805)

# Percurso diacrônico de completivas introduzidas por *como* no português

(Diachronic path of complement clauses introduced by *como* in Portuguese)

Gisele Cássia de Sousa<sup>1</sup>

<sup>1</sup>Instituto de Biociências, Letras e Ciências Exatas, UNESP – Universidade Estadual Paulista  
gcsousa@ibilce.unesp.br

**Abstract:** As a more extensive issue, this paper deals with the polysemy attested in all Romance languages between interrogative pronouns and subordination markers in complex sentences. It investigates, specifically, the behavior of objective complement clauses introduced by “*como*” (“*how*”) in constructions in which they occur as equivalent to the meaning of prototypical declarative complement clause introduced by the conjunction “*que*” (“*that*”). It analyzes complement clauses introduced by “*como*” and “*que*” occurring in representative Portuguese texts from the thirteenth to the twentieth centuries. The results of the comparative analysis between the two forms of clausal complement show that in archaic Portuguese “*como*” was used primarily to introduce complement clauses of factive verbs. When the complement clause is introduced by “*como*”, the content that is presupposed as true due to the meaning of matrix verb has reinforced this factuality. In the passage from the archaic to the modern period, there was a drastic decrease in the frequency of occurrences of complement clauses with “*como*”, which are replaced by the form of complementive introduced by the conjunction “*que*” in the context of factual constructions. In contemporary Portuguese, this replacement is fully established. Complement clauses with “*como*” remain rare and limited to specific contexts in which “*como*” maintains its reinforcement function of factual meaning.

**Keywords:** complement clauses; *wh*-complementizer; factuality.

**Resumo:** Este trabalho aborda, como temática mais ampla, a polissemia atestada em todas as línguas românicas entre pronomes interrogativos e marcadores de subordinação em sentenças complexas. Investiga-se aqui, especificamente, o comportamento de orações completivas objetivas introduzidas por “*como*” que, nas construções em que aparecem, exibem significado equivalente ao de uma completiva declarativa prototípica, iniciada pela conjunção “*que*”. Analisam-se dados de completivas, tanto com “*como*” quanto com “*que*”, ocorrentes em textos representativos do português dos séculos XIII a XX. Os resultados da análise comparativa entre as duas formas de complemento oracional revelam que, no português arcaico, “*como*” era empregado preferencialmente para introduzir completivas declarativas de verbos factivos. Ao ser introduzido por “*como*”, alternativamente à conjunção “*que*”, o complemento oracional, cujo conteúdo é pressuposto verdadeiro por força do significado do verbo matriz, tem essa sua factualidade reforçada. Na passagem do período arcaico ao período moderno, observa-se queda drástica na frequência de ocorrências de oração completiva com “*como*” que, no contexto de construções factuais, é substituída pela forma de completiva introduzida pela conjunção “*que*”. No português contemporâneo, essa substituição mostra-se completamente implantada, permanecendo rara a forma de completiva com “*como*”, limitada a contextos específicos nos quais esse elemento introdutor mantém sua função de reforço de significado factual.

**Palavras-chave:** orações completivas; complementizador interrogativo; factualidade.

## Considerações iniciais

O objetivo deste trabalho é descrever o percurso diacrônico de orações completivas declarativas introduzidas por “*como*” do português, com enfoque para os aspectos

mais relevantes associados à mudança por que passam essas orações do português arcaico (séculos XIII a XV) até o português contemporâneo (séculos XVIII a XX).

Especificamente, constituem objeto de análise deste trabalho orações iniciadas por “como” que, nas diferentes fases do português, ocorrem em posição de complemento verbal, com função equivalente ao de uma completiva introduzida pela conjunção “que”, conforme exemplifica a construção em (01), uma ocorrência do português arcaico, período em que essa forma de complemento oracional se mostra mais abundante:

- (01) Foi a deteença em esto tam grande que sse fazia ja muito tarde; e veendo *como* o achar nom *podiam*, levarom a besta e foram-sse ao logar, nom sabendo que cuidassem em tal feito. (século XV)
- (01') E vendo *que não o podiam achar...*

Conforme se buscará demonstrar, em construções como a de (01), o elemento “como” compete com “que” na introdução de completivas objetivas ao longo do período arcaico do português, até ser por ela praticamente suplantada no período moderno. Conjuntamente à descrição desse percurso histórico, focaliza-se, neste trabalho, a função cumprida pelo elemento “como” enquanto introdutor de orações completivas declarativas, alternativamente à conjunção “que”.

Os dados que serviram às análises que aqui se apresentam foram extraídos de três *corpora* de língua escrita distintos. Para a coleta de dados representativos do português arcaico (séculos XIII a XV) e moderno (séculos XVI e XVII), foram utilizados textos que integram o *Córpus Diacrônico do Português*, organizado por Tarallo (1991), e o “Banco Informatizado de Textos”, composto no âmbito do “Projeto para a História do Português”, por pesquisadores da Universidade Federal da Bahia (BIT-PROHPOR/UFBa). Para os dados do português contemporâneo (séculos XVIII a XX), foram investigados textos integrantes do “Córpus de Língua Escrita do Brasil”, pertencente ao “Centro de Estudos Lexicográficos”, da Universidade Estadual Paulista, câmpus de Araraquara.

Na seção seguinte deste texto, é focalizada a mudança, atestada em diferentes línguas românicas, de marcadores de interrogação, a exemplo do advérbio *como* do português, a introdutores de oração completiva. Posteriormente, apresenta-se o comportamento de completivas com *como* ao longo das diferentes fases do português investigadas e, na última seção do texto, apresentam-se as conclusões do trabalho.

## O desenvolvimento de interrogativos a introdutores de oração completiva

Heine e Kuteva (2006, p. 209) apontam que a polissemia entre formas marcadoras de interrogativas parciais (ou “interrogativas de conteúdo”) e formas introdutoras de orações subordinadas em sentenças complexas é fenômeno atestado em um conjunto significativo de línguas. Segundo os autores, em todas as línguas românicas e eslavas, em algumas línguas germânicas, no grego moderno, no húngaro e no georgiano, encontra-se o emprego de uma mesma forma para introduzir **interrogativas parciais simples** (02a), **orações completivas indefinidas** (02b), **completivas definidas** (02c) e **orações relativas com antecedente expresso** (02d), conforme ilustram as construções abaixo, exemplos do inglês dados pelos autores.

- (02) a. **Who** came?  
Quem veio?

- b. I don't know **who** came.  
Não sei quem veio.
- c. Y ou also know **who** came.  
Você também sabe quem veio.
- d. Do you know the woman **who** came?  
Você conhece a mulher que veio?

Conforme também demonstram Heine e Kuteva (2006), nas línguas em que é atestada, essa polissemia é resultado de desenvolvimento histórico de marcadores de interrogação que passam a atuar como subordinadores em sentenças complexas. O processo, segundo os autores, é unidirecional nessas línguas e se caracteriza por gramaticalização do item interrogativo envolvendo quatro estágios, conforme esquematizados no Quadro 1, a seguir.<sup>1</sup>

**Quadro 1: Evolução de interrogativos a subordinadores (HEINE; KUTEVA, 2006, p. 210)**

Estágio	Função do marcador de interrogação
I	Marcar interrogativas
II	Introduzir orações completivas indefinidas
III	Introduzir orações completivas definidas
IV	Introduzir orações relativas com antecedente expresso

Heine e Kuteva (2006, p. 210) caracterizam esses estágios de evolução do seguinte modo:

- (i) **Estágio I:** o marcador de interrogação é usado para iniciar interrogativas de conteúdo, como em (02a). Esse estágio distingue-se dos demais por envolver sentenças simples (mono-oracionais), ao passo que todos os outros envolvem combinação de orações.
- (ii) **Estágio II:** o uso do marcador de interrogação é estendido para a introdução de orações subordinadas indefinidas, de que (02b) é um exemplo.
- (iii) **Estágio III:** além de orações indefinidas, o marcador de interrogação passa a introduzir orações completivas definidas, como em (02c), que podem também ser interpretadas como relativas sem antecedente; mas o elemento não é usado para introduzir orações relativas com antecedente expresso.
- (iv) **Estágio IV:** o marcador de interrogação, nesse último estágio, introduz, além de orações relativas sem antecedente, relativas com antecedente expresso, tal como a de (02d).

Conforme também atestam Heine e Kuteva (2006), já no latim os marcadores de interrogação se encontravam completamente gramaticalizados. Além de introduzirem interrogativas parciais, marcavam orações completivas, tanto definidas quanto indefinidas, e orações relativas do latim, como ilustram as ocorrências em (03), presentes nos escritos de Plautus, dadas em Heine e Kuteva (2006, p. 218).<sup>2</sup>

1 Heine e Kuteva (2006) incluem na descrição desses estágios de gramaticalização o desenvolvimento dos marcadores de interrogação também a introdutores de oração adverbial, não mencionado aqui por razões de enfoque.

2 Esses diferentes empregos de palavras interrogativas no latim são atestados também por Ernout e Thomas (1972).

- (03) a. Miser est **qui** amat. **Estágio II**  
*Someone who is in love is to be deplored.*  
 Pobre é quem (de quem) ama.
- b. **Quod** credidisti reddo. **Estágio III**  
*What you have lent (me) I return (it to you)*  
 Eu devolvo o que você me emprestou.
- c. Eae litterae **quas** mihi scripsisti mihi jucundissimae fuerunt. **Estágio IV**  
*The letters that you wrote me were very pleasant to me.*  
 Estas cartas que você me escreveu me foram muito agradáveis.

Nas línguas românicas, a polissemia entre marcadores de interrogação e introdutores de orações subordinadas completivas (e também de relativas) é, portanto, conforme concluem Heine e Kuteva, reflexo direto da instauração do processo de gramaticalização de palavras interrogativas a subordinadores no latim (2006). Esse processo também explica, assim, a atuação dos elementos interrogativos em construções subordinadas do português, tanto completivas quanto relativas, conforme ilustram as traduções dos exemplos em (02), com os pronomes *quem* e *que*.<sup>3</sup>

No tocante especificamente à atuação de *como* enquanto introdutor de oração completiva, foco deste trabalho, no entanto, esse processo tem seus reflexos mais profícuos apenas no período arcaico do português, conforme também aponta, aliás, Dias (1933), ao registrar que, no português arcaico: “As orações substantivas são [...] orações introduzidas pela conjunção *que* (e às vezes também por *como*)” (DIAS, 1933, p. 257; grifo acrescentado).

A partir do período moderno, os contextos dessa atuação de *como* restringem-se drasticamente, especialmente no tocante a orações completivas definidas, tais como a de (02c), conforme se passa a demonstrar na seção a seguir.

## O percurso diacrônico de completivas com *como* em português

Nos textos representativos dos três períodos do português investigados, o elemento “como” aparece introduzindo orações completivas tanto **definidas** quanto **indefinidas**, conforme exemplificam, respectivamente, as ocorrências em (04a) e (04b), relativas ao português moderno.

- (04) a. E sei bem que se o rogares de bom coração, que haverás alegria e honra na batalha. E **eu te direi como este sinal [da cruz] te guarecerá (=livrará) de morte e te dará honra e alegria.** (séc. XVI)
- b. mas esta vossa falta me causa estar eu ainda em parte que vos possa escrever, ainda que **eu não sey como isto digna**, porque dou a entender ho que em my nam há como devia. (séc. XVI)

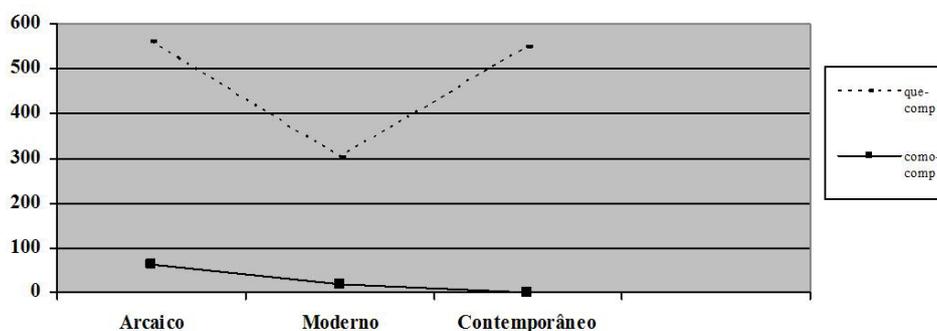
Observe-se que, nesses casos, “como” equivale a um pronome indicativo de modo, e a oração que ele introduz também pode ser interpretada como uma relativa sem antecedente. Em construções nas quais “como” equivale à conjunção “que” (as quais constituem foco principal das análises neste trabalho), entretanto, a oração que “como” introduz é predominantemente do tipo **definida**, e o significado de modo, diferentemente do que ocorre em (04a-b), não se revela prevacente para o elemento “como”, conforme evidenciam as ocorrências e suas paráfrases em (05a) e (05b):

<sup>3</sup> Uma proposta de distinção entre relativas sem antecedente e completivas introduzidas por palavras –*QU* do português, de um ponto de vista funcional, encontra-se em Sousa (2011).

- (05) a. Quando Nosso Senhor disse isto, tomou Josefes pela mão direita e tirou-o apos si, assi que todos os cristãos o vírão tirar e vírão a ele e aos anjos, que estavam derredor, e **vírão todos como Josefes estava ante ele e como ele fazia sobre ele o sinal da cruz.** (séc. XVI)
- a'. E viram todos **que** Josefes estava ante ele e **que** fazia sobre ele o sinal da cruz.
- b. E **vos bē sabedes commo el rei Bucar vē sobre nós e traz cōsigo XXXVI reys.** E, pois que elle traz tã grãde poder e eu tã çedo hey de morrer, vós outros catade se poderedes defēder Valêça. (séc. XIV)
- b'. E vos bem sabedes **que** o rei Bucar vem sobre nós e traz consigo 36 reis.

A maior parte de completivas declarativas com “como” ocorrentes nos dados analisados é de construções **factuais**, isto é, do tipo em que o conteúdo da oração completiva é pressuposto como verdadeiro (KIPARSKY; KIPARSKY, 1970). As ocorrências em (05) também assim se comportam quanto à factualidade, isto é, em (05a) é interpretado como verdadeiro o fato de que “Josefes estava diante de Cristo e que fazia o sinal da cruz”; e, em (05b), é pressuposto como verdadeiro o fato de que “o rei Bucar vem e traz consigo 36 reis”. Essa propriedade que têm as orações introduzidas por “como” de se comportarem predominantemente como factuais, quando introduzindo orações declarativas equivalentes a completivas com “que”, cumpre importante papel na mudança por que passam essas orações do português arcaico ao português contemporâneo, como se demonstrará adiante.

Conforme revela a distribuição das ocorrências no gráfico abaixo, do período arcaico ao contemporâneo, observa-se queda drástica no número de ocorrências de completivas introduzidas por “como” equivalentes a “que-complemento”. Completivas iniciadas pela conjunção “que”, por outro lado, mantêm seu número de ocorrências sempre superior ao de completivas com “como” ao longo dos três períodos investigados.



**Figura 1 - Frequência de ocorrência de completivas com “como” e com “que” nos três períodos investigados**

Nos textos do período arcaico, encontram-se 62 dados de completivas com “como” (que perfazem 75% do total de ocorrências desse tipo de oração), frente a 582 ocorrências de completivas com “que”. No período moderno, somente 20 ocorrências de completivas com “como” são encontradas (que correspondem a 25% do total de dados), ao passo que completivas com “que”, em textos desse período, totalizam 326 ocorrências. No período contemporâneo, por fim, encontra-se uma única ocorrência de oração completiva introduzida por “como”, frente a 602 ocorrências de completivas iniciadas pela conjunção “que”.

Uma vez que os dados de oração com “como” são de ocorrências dessa oração em construções nas quais ela equivale a uma completiva com “que”, a frequência inalterada dessa última forma de oração, ao longo dos períodos investigados, frente à queda drástica na frequência de orações com “como”, revela que a conjunção “que” suplanta diacronicamente a forma “como”, substituindo-a praticamente por completo, na passagem do português arcaico ao moderno, na função de introduzir completivas declarativas.

Análise mais detalhada dos ambientes sintáticos em que aparecem, nos dados analisados, essas duas formas de oração completiva oferece evidência adicional dessa substituição de “como” pela conjunção “que”. Duas propriedades das construções mostram-se particularmente relevantes nesse sentido: o **tipo semântico** e o **valor pressuposicional** do predicado matriz, isto é, se o predicado determina (factivo), ou não (não-factivo), como factival (pressuposto como verdadeiro) o conteúdo na oração completiva (KIPARSKY; KIPARSKY, 1970).

Predicados que aparecem nos dados como matriz de completivas tanto introduzidas por “como” quanto iniciadas por “que” são de quatro tipos: predicados de cognição (como *saber*), predicados de atitude (mental, como *lembrar*, ou emocional, como *lastimar*, *admirar-se*), de manipulação (como *mandar*) e de elocução (como *dizer*).<sup>4</sup> Conforme demonstram os resultados nas tabelas 1, 2 e 3, a distribuição dos dados ao longo dos períodos revela que gradativamente predicados que apareciam completados por oração tanto iniciada por “como” quanto introduzida por “que” passam a ocorrer exclusivamente com complementos na forma de oração iniciada pela conjunção “que”.

**Tabela 1: Forma da oração completiva vs. tipo de predicado matriz no português arcaico**

Forma da oração Tipo de predicado matriz	<i>Que-comp</i>	<i>Como-comp</i>	Total
Cognição	83/75%	28/25%	111
Atitude	123/86%	20/14%	143
Manipulação	206/100%	----	206
Elocução	148/91%	14/9%	162
Total	560	62	622

**Tabela 2: Forma da oração completiva vs. tipo de predicado matriz no período moderno**

Forma da oração Tipo de predicado matriz	<i>Que-comp</i>	<i>Como-comp</i>	Total
Cognição	38/95%	2/5%	40
Atitude	42/76%	13/24%	55
Manipulação	89/98%	2/2%	91
Elocução	134/98%	3/2%	137
Total	303	20	323

<sup>4</sup> Essa classificação dos predicados matrizes de oração completiva é baseada em proposta de Gonçalves, Sousa e Casseb-Galvão (2008).

**Tabela 3: Forma de completiva vs. tipo de predicado matriz no período contemporâneo**

Forma da oração	<i>Que-comp</i>	<i>Como-comp</i>	Total
Tipo de predicado matriz			
Cognição	93/100%	---	93
Atitude	182/99%	1/1%	183
Manipulação	122/100%	---	122
Elocução	150/100%	---	150
Total	547	01	548

Os únicos tipos de construção com completiva introduzida por “como” que, do período arcaico para o período moderno, exibem um comportamento de elevação, em vez de queda, na frequência de ocorrência são construções em que esse tipo de completiva se encaixa a predicados de **atitude**, que passam de 14% no período arcaico a 24% do total de dados no período moderno, e aquelas em que essa oração complementa predicados de **manipulação**, que não ocorrem nos textos do período arcaico e aparecem representando 2% do total de ocorrências no período moderno. O número de ocorrência de “como” com predicados de manipulação é, porém, muito baixo (apenas duas ocorrências, que seguem transcritas em (06a-b)), e sua ausência nos dados do período arcaico pode dever-se à natureza dos textos investigados para esse período.

- (06) a. Nesta viagem aconteceo que, chegando o Padre com seus companheiros a hum lugar de muyta gente, **quis dar ordem como se fizesse ali outra casa**, por a terra ser pera yso aparelhada. (séc. XVI)
- b. E, quando comigo quiserdes falar, abrireis a arca, onde quer que tu fores, e tu soo verás a escudela, ca eu nom quero que a outrem toque, senam tu e teu filho Josefes. Ora te vai e **ordena a teu povo como pouse** e faz assi como eu te mandei! (CBP/16/JA-20)

A elevação na frequência de ocorrências de orações com “como” que funcionam como completivas de predicados de atitude, por outro lado, revela-se interessante na medida em que é com predicado matriz desse tipo a única ocorrência de completiva introduzida por “como” encontrada nos textos do período contemporâneo. A ocorrência em questão encontra-se transcrita em (07), abaixo.

- (07) Meu caro Barão, a sua carta deu-me grande prazer como a primeira esperança de vê-lo ainda em Londres este mês. Estou preparando-me para ir recebê-lo a Charing Cross com o seu estado maior de legação. **Não sei realmente como se pode assim desertar de Londres durante o único tempo em que há merecimento em ficar.** (séc. XIX)
- Não consigo aceitar que se possa assim desertar de Londres durante o único tempo em que há merecimento em ficar.*
- Não acredito que se possa assim desertar de Londres durante o único tempo em que há merecimento em ficar.*

Note-se que, apesar de *saber* ser tipicamente um verbo de cognição, na ocorrência em (07), em que ele aparece negado, com flexão de 1ª pessoa do singular, e complementado por oração introduzida por “como”, seu significado equivale a “não acredito que...”, “não consigo aceitar que...”, ou “não me conformo que...”, conforme demonstram as paráfrases da construção em (07), constituindo-se, portanto, um predicado de atitude, indicativo de uma indignação do falante.

Observe-se ainda, a partir da ocorrência em (07), que, do mesmo modo que na maior parte das ocorrências de completivas com “como” em todos os períodos investigados, o conteúdo do complemento em (07) é do tipo **factual**, e essa factualidade é inclusive reforçada pela presença de “como”, em vez de “que”, enquanto introdutor da oração completiva, conforme também evidenciam as paráfrases da construção em (07) com oração completiva iniciada por “que”.

A factualidade da construção corresponde, de fato, até o período moderno, propriedade relevante para a ocorrência de oração completiva introduzida por “como” alternativamente a uma completiva iniciada por “que”, conforme demonstram os resultados nas tabelas 04, 05 e 06 abaixo.

**Tabela 4: Forma da completiva vs. factualidade do predicado matriz no período arcaico**

Forma da oração			
Valor pressuposicional do verbo matriz	<i>Que-comp</i>	<i>Como-comp</i>	Total
Factivo	73/59%	51/41%	124
Não-factivo	509/98%	11/2%	520
Total	582	62	644

**Tabela 5: Forma da completiva vs. factualidade do predicado matriz no período moderno**

Forma da oração			Total
Valor pressuposicional do verbo matriz	<i>Que-comp</i>	<i>Como-comp</i>	Total
Factivo	22/55%	18/45%	40
Não-factivo	304/99%	02/1%	306
Total	326	20	346

**Tabela 6: Forma da completiva vs. factualidade do predicado matriz no período contemporâneo**

Forma da oração			Total
Valor pressuposicional do verbo matriz	<i>Que-comp</i>	<i>Como-comp</i>	Total
Factivo	38/98%	01/2%	39
Não-factivo	564/100%	---	564
Total	602/99%	01	603

Conforme revelam esses resultados, interpretados em direção horizontal da tabela, do português arcaico ao contemporâneo, a função de complementar predicados factivos gradativamente vai sendo assumida exclusivamente pela oração introduzida pela conjunção *que*. Nos períodos arcaico e moderno, cerca de 45% de ocorrências com predicados factivos é complementada por oração introduzida por “como”, ao passo que, no português contemporâneo, esse percentual é de apenas 2%, representativo de apenas uma ocorrência.

## Considerações finais

O principal propósito deste trabalho foi analisar, a partir de dados representativos de diferentes períodos do português, o comportamento de completivas declarativas iniciadas por “como”, com função claramente equivalente ao de uma completiva prototípica, introduzida pela conjunção “que”. Apresentam-se, neste texto, os resultados que se mostraram mais significativos na análise do comportamento dessa oração, comparativamente ao de orações iniciada por “que”, nos textos investigados.

Em termos gerais, esses resultados revelam que a ocorrência de “como” enquanto introdutor de orações completivas, resultante de gramaticalização desse elemento no latim, é motivada essencialmente pela factualidade da construção com complemento oracional. Até o período moderno, conforme indicam esses resultados, “como” exercia, paralelamente à conjunção “que”, a função de atribuir factualidade ao conteúdo de completivas de predicados tanto factivos (com os quais essa factualidade era, portanto, reforçada) quanto não-factivos. Do período moderno ao período contemporâneo, entretanto, essa função passa a ser cumprida quase exclusivamente pela conjunção “que”, permanecendo a forma “como” enquanto introdutora de completivas apenas em contextos altamente restritos, exercendo prioritariamente, afinal, a função que caracteriza sua atuação nos dois períodos anteriores do português, ou seja, a de reforçar a factualidade do conteúdo na oração que introduz.

Um ponto ainda a ser investigado, a partir desses resultados, diz respeito à identificação de propriedades ligadas a “como” interrogativo, tanto gramaticais, quanto semânticas e possivelmente semântico-cognitivas, que possivelmente explicariam o recrutamento dessa forma, e não de outras formas de interrogativos (como “por que”, “quando”, “onde”, etc.), para a atuação em completivas factuais, alternativamente à conjunção “que”. Em outras palavras, caberia investigar, a partir dos resultados que aqui se apresentam, quais propriedades associadas a “como” teriam motivado a extensão, verificada especialmente no português arcaico, de um significado como o de modo (significado básico da proforma “como”) para uma função como a de indicar factualidade de uma proposição, tendo sido aparentemente essa extensão barrada a significados como os de tempo (indicado por “quando”), lugar (representado por “onde”) e razão (indicado pela forma “por que”).

## REFERÊNCIAS

DIAS, A. E. S. *Syntaxe historica portuguesa*. 3. ed. Lisboa: Livraria Clássica Editora, 1933.

ERNOUT, A. S.; THOMAS, F. *Syntaxe latine*. Paris: Klincksieck, 1972.

GONÇALVES, S. C.; SOUSA, G. C.; CASSEB-GALVÃO, V. C. As construções subordinadas substantivas. In: ILARI, R.; NEVES, M. H. M. (Org.). *Gramática do português falado culto no Brasil: classe de palavras e processos de construção*. Campinas: Editora da Unicamp, 2008. p. 1021-1084.

HEINE, B.; KUTEVA, T. *The changing languages of Europe*. New York: Oxford University Press, 2006.

KIPARSKY, P.; KIPARSKY, C. Fact. In: STEINBERG, D. D.; JAKOBOWITS, L. A. *An Interdisciplinary Reader in Philosophy*. Cambridge: Cambridge University Press, 1970. p. 345-369.

SOUSA, G. C. Interrogativas encaixadas vs. relativas sem antecedente: contribuições da Gramática Discursivo-Funcional. *Estudos Linguísticos*, São Paulo, v. 40, n. 1, p. 123-132, 2011.

TARALLO, F. *Cópus diacrônico do português*. Campinas, 1991.

# A realização do sujeito pela elite paulistana do início do séc. XX

(The realization of subject by the São Paulo elite  
in the beginning of the twentieth century)

Hélcio Batista Pereira<sup>1</sup>

<sup>1</sup> FFLCH/USP; Universidade Paulista – Unip

helcius@usp.br

**Abstract:** This paper analyses the trajectory of empty subjects and overt subjects in the standard language of the São Paulo elite from the late twentieth century to 1930s, when that city received a large influx of immigrants, who were speakers of languages characterized by the use of null subject. The analysis of the linguistic corpora led us to conclude that the elite, despite making more use of empty subject, was already boosting the use of overt subjects, a phenomenon well advanced in the caipira dialect. This behavior is explained by structural factors internal to language, and also expressed, from the social-historical standpoint, a clear choice of a linguistic form, which was opposed to the standard language of immigrants, as established by the *habitus* of the elite.

**Keywords:** null subject parameter; social history of language; linguistic standard language of the São Paulo elite; linguistic *habitus*; Pierre Bourdieu.

**Resumo:** O presente trabalho analisa a trajetória dos sujeitos vazios e expressos na norma linguística efetivamente utilizada pela elite paulistana, em fins do século XIX até a década de 1930, quando São Paulo recebeu grande afluxo de imigrantes usuários de “línguas de sujeito nulo”. A análise de *corpora* linguísticos nos levou a concluir que a elite, apesar de fazer um maior uso do sujeito vazio, já incrementava o uso dos expressos, fenômeno já bem avançado no dialeto caipira. Tal comportamento é explicado por fatores estruturais internos à língua, e também expressa, do ponto de vista histórico-social, uma clara opção por uma forma linguística que se opunha à norma dos imigrantes, tal qual previsto no *habitus* da elite.

**Palavras-Chaves:** parâmetro do sujeito nulo, História Social da linguagem; dialeto da elite paulistana, *habitus* linguístico; Pierre Bourdieu.

## Introdução

Muitos estudos já foram realizados sobre a realização do sujeito nulo no Português Brasileiro (doravante PB). Este trabalho dialoga com tais trabalhos, mas busca um tratamento que vai além da análise linguística em perspectiva histórica. Procuraremos avaliar o tema em termos de História Social da Língua. Nesse modelo, é necessário trazer para o interior do modelo de análise os aspectos sócio-históricos, utilizado para interpretar os resultados linguísticos com os quais dialoga.

O desafio aqui é contribuir para que o objeto seja lido em uma perspectiva que não se limite aos campos tradicionalmente identificados como da Linguística e da Sociolinguística, mas seja também iluminado pelos instrumentos da História e da Sociologia.

Como veremos adiante, teremos que buscar um arcabouço teórico apropriado a tal procedimento. Partindo da tradição que herdamos – a chamada sociolinguística paramétrica, cuja formulação se deu a partir de Tarallo e Kato (2007), nos apoiaremos na teoria sociológica de Pierre Bourdieu, obtendo um apoio teórico adequado para nosso programa de

pesquisa. Não se trata aqui de negar ou refutar a tradição que herdamos, mas de adequá-la para o objeto de pesquisa que pretendemos analisar.

### **A realização do sujeito nulo como objeto de estudo da História Social do Português de São Paulo**

Vários estudos sobre a realização do sujeito pronominal no PB evidenciaram um processo de mudança na distribuição de nulos e plenos, o qual foi marcado pela ampliação do uso da forma expressa sobre a oculta. A leitura de dois dos mais importantes estudos sobre o tema em perspectiva diacrônica localiza o século XX como o período em que o fenômeno teria se desenvolvido. Assim, analisando *corpora* de cartas, diário e peças, Tarallo (1985) constatou que a produtividade dos plenos teria subido de 32%, em 1880, para 79%, na década de 1980. Já Duarte (1996),<sup>1</sup> que trabalhou com peças teatrais brasileiras, constatou que a forma nula respondia em 1918 por algo em torno de 70% das ocorrências; atingiu 50% em 1937; chegou a 30% nos materiais de 1975 e, finalmente, reduziu-se a apenas 25%, nos textos analisados da década de 1990.

Se considerarmos os dados encontrados principalmente por este último estudo, chegaremos à conclusão de que foi mais exatamente nas primeiras décadas do século passado que a realização do sujeito com pronomes expressos ganha relevância, quebrando a hegemonia da forma oculta, que marca as línguas chamadas de línguas de sujeito nulo. Partindo dessa constatação procuramos investigar a questão sob o ponto de vista do dialeto das elites de origem nacional que se transferiram para São Paulo a partir da década de 1880 até a década de 1930, advindos principalmente do interior do Estado.

Muito nos interessou saber como o fenômeno se desenvolveu para essa camada da população que fixou residência em uma cidade que recebeu massas populacionais de imigrantes, constituídas principalmente por falantes de línguas maternas que faziam amplo uso do sujeito nulo. Teriam sido 1,2 milhões de imigrantes que teriam desembarcado no porto de Santos de 1908 a 1936 – especialmente italianos, portugueses e espanhóis, sendo que boa parte desse contingente teria se fixado na capital paulista imediatamente ou após experiência no interior, trabalhando principalmente nas fazendas de café.

Cercados por tal volume de falantes que preferiam construir sentenças omitindo os sujeitos pronominais, como teriam se comportado os indivíduos das classes que detinham o poder econômico e social em São Paulo? Teriam seguido o movimento mais geral do PB de intensificação do uso do sujeito pleno ou teriam sido influenciados a manter a forma nula com elevados índices de produtividade?

Antes de mostrarmos os resultados que encontramos, explicitaremos aspectos de nossa metodologia e do arcabouço teórico que nortearam nosso trabalho.

---

<sup>1</sup> Estudo publicado pela primeira vez em 1993.

## Aspectos teóricos e metodológicos

Partimos do arcabouço teórico já consolidado nos quadros da pesquisa linguística brasileira, proposto por Tarallo e Kato (2007), que realizou o “casamento” do modelo laboviano com o gerativismo de Chomsky. Como se sabe, a questão do chamado “parâmetro do sujeito nulo” é ponto importante da pesquisa gerativista de uma maneira geral, mas foi a partir do uso da metodologia da sociolinguística laboviana que o fenômeno da intensificação do uso de plenos pode ser melhor descrito no âmbito do Português Brasileiro.

Buscando um arcabouço teórico que incluísse a História Social no modelo explicativo, agregamos a este escopo a teoria sociológica de Pierre Bourdieu. Muito crítico ao “estruturalismo” de Saussure, às propostas do modelo chomskyano e a toda perspectiva teórica com enfoque exclusivo nas “estruturas objetivas”, Bourdieu cunhou sua teoria também avesso às correntes de cunho “subjetivista”. Para o sociólogo, as ações dos indivíduos – incluindo aí as ações linguísticas – não são fruto exclusivo nem de estruturas autônomas (sociais, culturais, etc.), nem é o mero resultado da criação e produção de um indivíduo que é livre para agir e realizar suas escolhas de forma independente de qualquer determinismo social. Aposta na relação dialética entre o nível individual e o social, entre a interioridade e a exterioridade (BOURDIEU, 1994, p. 60). Assim, os indivíduos fazem suas escolhas sociais (e linguísticas) a partir do *habitus*, um dispositivo gerador de ações, construído historicamente, interiorizado já nos primeiros contatos sociais. Trata-se de

[...] um sistema de disposições duráveis e transponíveis que, integrando todas as experiências passadas, funciona a cada momento como matriz de percepções, de apreciações e de ações – e torna possível a realização de tarefas infinitamente diferenciadas, graças às transferências analógicas de esquemas, que permitem resolver os problemas da mesma forma, e às correções incessantes dos resultados obtidos, dialeticamente produzidos por esses resultados. (BOURDIEU, 1994, p. 65)

O *habitus* é o mecanismo pelo qual todos os atos (inclusive os linguísticos) passarão a ser julgados. Do ponto de vista do uso da língua, é o dispositivo que atuará para o julgamento de uma forma ou discurso linguístico como “correta” ou “legítima”. O processo de legitimação de determinados usos da língua está sujeito ao jogo concorrencial que se dá em um “mercado” ou “campo” linguísticos. Cada grupo social, a partir do *habitus* próprio a este associado, lança no “mercado linguístico” seus produtos (ou seja, seus discursos e formas linguísticas). O grupo social que possui mais “capital linguístico” consegue impor os seus produtos e dele auferir um maior lucro nas relações de troca, impondo o seu “poder simbólico”. É a partir disso, que os todos os grupos sociais passaram a tomar os produtos linguísticos deste grupo como “legítimo”, ou seja, como o parâmetro a partir do qual todas as outras formas linguísticas serão julgadas. Somente os produtos linguísticos tomados por “legítimos” terão *status* de formas prestigiadas, restando às que mais lhe opõem a situação de estigmatizadas.

Partindo desse arcabouço teórico, propusemos a realizar primeiro uma avaliação mais propriamente linguística a partir de *corpus* que pudessem nos mostrar a realização do sujeito pronominal por membros da elite paulistana no início do século XX. Escolhemos a família Mesquita, por constituírem família importante não somente no campo econômico – proprietários que se tornaram de *O Estado de São Paulo*, mas também pela importância que tinham do ponto de vista cultural – exatamente por dirigirem um dos

mais importantes veículos de imprensa do período. Dividimos este *corpora* em dois, mapeando duas gerações: a 1ª geração dos Mesquita, com 9 cartas produzidas por Júlio Mesquita ao amigo e político Washington Luís, e a 2ª geração da família, com 23 missivas trocadas entre Júlio de Mesquita Filho e sua esposa Marina Mesquita no final da décadas de 1920 e início 1930, em função em primeiro lugar da separação imposta pela ida do jornalista para o front da guerra da chamada “Revolução Constitucionalista” e de seu consequente exílio do país, após sua captura.

Esse uso foi comparado aos resultados que obtivemos do “dialeto caipira”, uma vez que, como a biografia dos Mesquita nos contou, estes, assim como quase toda a elite nacional paulistana do período, tinha origens no interior do Estado. Para tanto, constituímos *corpus* formado com textos de caráter ficcional produzidos por Valdomiro Silveiro, dada a ausência de melhor material para comparar caracterizar este dialeto – a camada caipira não era letrada, nem há registros orais de seus usos linguísticos para o período considerado nesta pesquisa – e à ambição “realista” deste autor de constituir uma prosa que imitasse a fala do interior de maneira fidedigna, evitando tratá-la de forma estilizada ou pela via da paródia.

Para melhor caracterizar o *habitus* da elite utilizaremos, além de pesquisas realizadas no campo da História Social de São Paulo, das análises dois artigos de Monteiro Lobato “Urupês” e “Velha Praga”, da crônica “Xenofobia”, de Jorge Americano, e o conto “Sociedade”, de Alcântara Machado. Todos esses autores são membros da elite e com sua análise buscamos as visões da elite sobre os imigrantes e os caipiras.

## O que os dados linguísticos nos contam

Os resultados que encontramos para a realização do sujeito pronominal pela elite paulistana – representada aqui pelos usos feitos pela família Mesquita – mostram que, nas primeiras décadas do século XX, esse grupo social ainda não fazia amplo uso do sujeito pleno. Como podemos ver na Tabela 1 abaixo, as duas gerações da família proprietária de *O Estado de S. Paulo*, privilegiava construções com a forma vazia. Entretanto, chamamos atenção que a geração mais nova – a de Júlio Mesquita Filho e sua esposa – realiza um aumento da produtividade dos pronomes expressos, fazendo este uso em 30% das ocorrências contra 20% da 1ª geração.

**Tabela 1: Corpora Elite Paulistana – Distribuição de sujeitos vazios e expressos por geração**

Forma Pronominal	1ª Geração	2ª Geração
Vazio	61 (80%)	387 (70%)
Expresso	15 (20%)	168 (30%)

Apenas para exemplificar, a sentença (1) abaixo mostra uso de Júlio Mesquita (o pai), sempre com sujeito vazio; (2) e (3) são casos da geração mais novas.

- (1) Pessoalmente, cumpre-me declarar mais uma vez, depois de agradecer ao Sr. a delicadesa do convite, que me cv dirigiu, e a honra da missão, de que me cv encarregou - que por motivos por mim ao Sr. já expostos, não me é possível voltar á cadeira do Senado, que cv perdi, nada tendo perdido com isso nem República, nem o Estado. (WL 19801010)

- (2) O Ruy recebeu cartões do Caco, Lia e da Lucila, ficou radioso. Ele pede a você para trazer-lhe dos livros de Gedon. (MESQUITA FILHO, 2006, p. 59)
- (3) Mequinha seguiu dia 21. cv Foi tristíssima, por ter deixado as filhas. (MESQUITA FILHO, 2006, p. 59)

Estes dados podem ser contrastados ao que encontramos no *corpus* que representa o uso do dialeto caipira, consolidados na Tabela 2. Assim, os caipiras de Valdomiro Silveira já optam pela forma expressa na posição de sujeito em 61% das sentenças que analisamos, uma das quais apresentamos em (4).

**Tabela 2: *Corpus* Dialeto Caipira: Distribuição de Vazios e Expressos**

Forma Pronominal	Qtd.
Vazio	105 (39%)
Expresso	165 (61%)

- (4) Quando eu trouxe a Tuca p'ra morar comigo (você na certeza ainda se alembra), a mó que eu vinha trazendo uma riqueza: eu tava alegre por demais (SILVEIRA, 17/01/1906).

Desses resultados chegamos a duas conclusões importantes que contam sobre o *habitus* linguístico da elite. Em primeiro lugar, a despeito do enorme contato que possa ter tido com populações que faziam amplo uso do sujeito nulo, a elite paulistana do período leva a cabo os primeiros movimentos do PB, de uma maneira mais geral, para intensificar o uso da forma expressa. Ou seja, não se deixou influenciar pelo contato linguístico com imigrantes italianos, portugueses e espanhóis e, embora ainda preferissem a forma vazia, já dão mostras de que este quadro era passível de alteração.

A segunda constatação que chegamos é que o processo de intensificação da forma plena estava em estágio avançado nos usos das camadas populacionais ditas “caipiras”. Em São Paulo, pode-se dizer que o fenômeno atingiu primeiro esse grupo social para depois, ao longo do século XX, se expandir para outros estratos sociais.

Mas se essas duas constatações são pertinentes, chegaríamos então à conclusão de que a elite paulistana do início do século XX optou por um uso próprio do dialeto caipira e refutou o uso que provavelmente fazia a camada de imigrantes. Para esclarecer melhor este ponto, devemos passar a avaliar os dispositivos que constituíam o *habitus* da elite paulistana, de uma forma mais geral, o que certamente definia também o seu *habitus* linguístico.

### **Em busca do *habitus* e do *habitus* linguístico da elite**

A análise dos textos de Alcântara Machado e Jorge Americano nos permitiu constatar que o *habitus* da elite continha dispositivos que avaliavam negativamente os imigrantes, fossem eles pobres ou ricos. Em “Xenofobia”, encontramos na boca de um personagem a aversão ao imigrante estrangeiro, fato que deveria ser comum entre os membros da elite paulistana “nacional”, como notamos a seguir:

- Isto aqui está ficando um acampamento de aventureiros. Não há mais lugar para os paulistas dos velhos troncos. Você sai a rua e só ouve falar estrangeiro. (AMERICANO, 1962, p. 279)

Esse mesmo preconceito aparece em “Sociedade”. No conto, a esposa do Conselheiro Jose Bonifácio de Matos e Arruda, uma família da elite dos “velhos troncos paulistas”, anuncia, logo no início da narrativa, que sua filha não se casaria com filho de “carcamano” (ALCÂNTARA MACHADO, 1971, p. 25). Ela resiste em aceitar o matrimônio de sua filha com o filho da “nova elite” “estrangeira”, que enriquecida procurava reconhecimento social, e exige a aprovação e o apoio do marido. Mas a família acaba por aceitar em função de necessidades financeiras, depois de o conselheiro ter se tornado sócio do imigrante. O “carcamano” se converte em “capital”, conforme mostra ironicamente o narrador, o que podemos ver no trecho abaixo:

O capital acendeu o charuto. O conselheiro coçou os joelhos disfarçando a emoção. [...] O capital levantou-se. Deu dois passos. Parou. (ALCÂNTARA MACHADO, 1971, p. 27)

Este aspecto do *habitus* da elite, ou seja, a sua recusa de conceder ao imigrante um espaço social privilegiado é atestado também pela historiografia. Basta lembrar que

[...] até 1918, por exemplo, não era permitido o ingresso de certos imigrantes nos clubes sociais de maior prestígio, o Jockey, o Automóvel e o São Paulo, e ainda hoje [década de 1970] não se encontram muito amiúde nomes de imigrantes em suas diretorias. (DEAN, 1971, p. 85)

Por possuir um maior “capital social”, a elite manteve um uso que divergia do feito pelo imigrante, não se deixando influenciar pelo contato com esse grupo linguístico. Esse “capital social”, do ponto de vista linguístico, implicava um “capital linguístico” suficiente para impor os seus usos como “legítimos” – ainda que em determinado momento o seu “capital financeiro” houvesse diminuído ou reduzido.

Mas se o *habitus* da elite tem elementos para afastar – até quando pode – o imigrante, e isso explicaria por que o dialeto da elite se comportou de modo a não legitimar a forma de realização do sujeito própria dos idiomas de origem de tais estrangeiros, o que dizer sobre o seu comportamento em relação ao caipira? A avaliação de “Velha Praga” e “Urupês”, de Monteiro Lobato, nos dá mais evidência disso. Monteiro, intelectual e antigo proprietário de terra do Vale do Paraíba, mostra um profundo preconceito e recusa social em relação às camadas mais pobres do interior do Estado. Em “Velha Praga”, a destruição da Serra da Mantiqueira é colocada na conta dos caipiras, em uma visão elitizada dos fatos que nem sequer menciona a destruição das matas promovidas pelo avanço do café nas terras paulistas. Em “Urupês”, Lobato atribui ao caipira a característica de estar sempre alheio aos fatos históricos, diante dos quais, estaria “sempre de cócoras”, e que somente ajeitado sobre os seus calcanhares é que conseguiria “destravar” sua língua e sua inteligência. O caipira seria preguiçoso, o que explicaria, por exemplo, a rudeza da habitação em que mora, o que faria “[...] sorrir aos bichos que moram em tocas e gargalhar ao João-de-barro” (LOBATO, 1976, p. 148).

Mas se a depreciação dos produtos culturais e linguísticos advindos da camada caipira estava também previsto no *habitus* da elite, por que optou em aceitar fenômeno – a intensificação do sujeito pleno – tão amplamente desenvolvido no dialeto caipira? A resposta, nós encontramos na História. A cidade de São Paulo e sua elite, a despeito das transformações modernizadoras, não substituíram a velha São Paulo por uma nova, livre da camada caipira de uma hora para outra:

São carroceiros (essenciais para o transporte de algumas cargas na cidade), os carregadores, os ambulantes, os coletores de lixo, as lavadeiras, os trabalhos domésticos. Ou seja, são trabalhadores manuais sem qualquer qualificação e que sobrevivem nas franjas da economia urbano-industrial em expansão. (SAES, 2004, p. 234)

Além disso, é preciso lembrar que esse processo de modernização da economia, em grande medida, foi tributário de um negócio que não deixava que a cidade rompesse com o seu passado rural, embora também a sintonizasse com os mercados mundiais: a “economia cafeeira”. O café forneceu muito do capital e dos empreendedores imprescindíveis para o desenvolvimento das áreas industrial, comercial e bancária da cidade. Muitos membros da elite passavam parte do ano na capital, parte no campo para acompanhar as questões de produção agrícola e, finalmente, “mudavam-se” para a Europa. A cidade e sua elite se mantiveram sintonizada a um só tempo com a produção no campo e com as atividades econômicas tipicamente urbanas, relacionadas ou não à exportação de café. Apenas posteriormente, e muito lentamente, é que o café cederia seu lugar à indústria, com a intensificação do processo de industrialização, “libertando” a cidade de suas faces “rurais”.

Nesse contexto é preciso, então, perceber a origem “rural” da elite e os vínculos que mantinha com o campo. Esse aspecto explica por que pôde partilhar com a camada “caipira” de um de seus usos linguísticos. Não partilha de todos os usos do dialeto caipira, já que como vimos, prevê em seu *habitus* elementos para refutar os “produtos linguísticos” dessa variedade. Continuará, assim, estigmatizando, por exemplo, a concordância nominal e verbal encontrada na fala caipira. Para esses usos não aceitos, seu *habitus* linguístico continuará reservando o estigma e o preconceito. Mas para o sujeito pleno reservará espaço para sua legitimação.

## Conclusões

Este trabalho procurou analisar a realização do sujeito pela elite paulistana do início do século XX. Nossos resultados mostraram que o sujeito pleno ainda que não houvesse se transformado na opção preferida para a constituição de sentenças por esse grupo social, já mostrava sinais de intensificação. Esse comportamento mostra pontos de contato com o “dialeto caipira” – que pelos nossos dados já havia feito do sujeito expreso a forma mais comum – e diverge da gramática da maior parte dos imigrantes que se transferiram para São Paulo no início do século.

Buscamos explicação para o fato no *habitus* que refutava os imigrantes e que, apesar de também depreciar os grupos caipiras, compartilhava de aspectos do universo rural, por estarem ainda ligados ao campo em função da economia cafeeira. O seu *habitus* linguístico valida, então, o sujeito pleno, que pode iniciar o seu processo de intensificação, substituindo a forma vazia em grande parte dos contextos sentenciais.

Nossa pesquisa, por fim, procurou colaborar para a definição de um arcabouço teórico-metodológico adequado ao tratamento de temas na área de História Social da Língua. A solução que oferecemos optou pela conciliação da teoria sociológica de Pierre Bourdieu ao escopo teórico que herdamos, ou seja, a sociolinguística paramétrica. Essa inclusão se deu para que pudéssemos considerar na totalidade as questões sócio-históricas em nosso modelo interpretativo. Não quisemos com isso desconsiderar ou desqualificar os pilares teóricos que tantos frutos trouxeram para a pesquisa linguística brasileira, filiados

à proposta de Tarallo e Kato (2007), mas adequarmo-nos aos nossos objetivos de realizar pesquisa na área de História Social da Língua.

## REFERÊNCIAS

### Fontes Metodológicas

Correspondência entre Júlio de Mesquita e Washington Luís:

Instituição: Arquivo do Estado de S. Paulo

Coleção: Arquivo Privado Washington Luís

Referência do documento: 19801005 S. Paulo, 28-01-1907

Referência do documento: 19801012 S. Paulo 28-11-1907

Referência do documento: 19801013 S. Paulo, 12-12-1907

Referência do documento: 19801002 S. Paulo, 25-3- 1911

Referência do documento: 19801010 S. Paulo, 30 -03-1922

Referência do documento: 19801008 S. Paulo, 07-10- 1922

Referência do documento: 19801007 S. Paulo - 26-12-1922

Referência do documento: 19801006 Rio de Janeiro, 07-10-1923

Referência do documento: 19801004 Rio de Janeiro, 15-11-1923

MESQUITA FILHO, Júlio; MESQUITA, Marina. *Cartas do exílio*. São Paulo: Terceiro Nome, 2006.

SILVEIRA, Valdomiro. As frutas. *O Estado de São Paulo*, São Paulo, p. 1, 08/01/1906.

\_\_\_\_\_. Mau Costume. *O Estado de São Paulo*. São Paulo, p. 1, 17/01/1906.

\_\_\_\_\_. Tal e qual. *O Estado de São Paulo*. São Paulo, p. 1, 25/01/1906.

\_\_\_\_\_. Eu, no Sertão. *O Estado de São Paulo*. São Paulo, p. 1, 28/06/1906.

\_\_\_\_\_. Trama. *O Estado de São Paulo*. São Paulo, p. 1, 31/07/1906.

\_\_\_\_\_. A avinha má. *O Estado de São Paulo*. São Paulo, p. 1, 10/08/1906.

\_\_\_\_\_. A Pantomima. *O Estado de São Paulo*. São Paulo, p. 1, 25/08/1906.

\_\_\_\_\_. No sertão. *O Estado de São Paulo*. São Paulo, p. 1, 12/09/1906.

\_\_\_\_\_. Enredos. *O Estado de São Paulo*. São Paulo, p. 1, 26/10/1906.

### Fontes bibliográficas

ALCÂNTARA MACHADO, Antonio de. Sociedade. *Novelas paulistanas*. Rio de Janeiro: José Olympio, 1971.

- AMERICANO, Jorge. *São Paulo nesse tempo (1915-1935)* São Paulo: Melhoramentos, 1962.
- BOURDIEU, Pierre. Economia das Trocas Lingüísticas. In: ORTIZ, Renato (Org.) *Pierre Bourdieu: Sociologia*. São Paulo: Ática. Coleção Grandes Cientistas Sociais, 1994. v. 39.
- CHOMSKY, Noam. *Novos horizontes no estudo da linguagem e da mente*. São Paulo: Editora Unesp, 2005.
- DEAN, Warren. *A industrialização de São Paulo*. 2. ed. São Paulo: Difel Editorial, 1971.
- DUARTE, Maria Eugênia Lamoglia. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: KATO, Mary Aizawa; ROBERTS, Ian (Org.). *Português Brasileiro: uma viagem diacrônica*. 2. ed. Campinas: Editora da Unicamp, 1996.
- LOBATO, Monteiro. *Urupês*. 20. ed. São Paulo: Brasiliense, 1976.
- SAES, Flávio. São Paulo republicana: vida econômica. In: PORTA, Paula (Org.). *História da cidade de São Paulo: a cidade na primeira metade do século XX*. São Paulo: Paz e Terra, 2004. v. 3. p. 215 - 245.
- TARALLO, Fernando. The filling of the gap: pro-drop rules in brazilian portuguese. In: KING, Larry; MALEY, Catherine. A. (Ed.) *Selected papers from the XIIIth linguistic symposium on romance languages*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 1985. p. 355-375.
- TARALLO, Fernando; KATO, Mary Aizawa. Harmonia trans-sistêmica: variação intra- e inter-lingüística. *Diadorim: Revista de Estudos Linguísticos e Literários*, Rio de Janeiro, UFRJ, n, 2. p. 13-42, 2007.

# Um enfoque no domínio da junção: a gramaticalização de *mesmo assim*

(An approach for the junction domain: the grammaticalization of *mesmo assim*)

Lúcia Regiane Lopes-Damasio<sup>1</sup>

<sup>1</sup>Instituto de Linguagens – Universidade Federal de Mato Grosso (UFMT)

luciaregiane@bol.com.br

**Abstract:** This paper presents a qualitative and quantitative analysis and is focused on the process of grammaticalization of the junctive phrase *mesmo/ainda assim* of contrastive value in correlation to its function in Discursive Traditions, investigated from the 18<sup>th</sup> to 21<sup>th</sup> centuries. The results point out to a close relation between the traditions of investigated texts and a path of change that has been revealed to be still on going.

**Keywords:** grammaticalization; junction; contrast; discursive tradition.

**Resumo:** Este artigo focaliza, numa análise de natureza qualitativa e quantitativa, o processo de gramaticalização da locução conjuntiva de valor contrastivo *mesmo/ainda assim* em correlação ao seu funcionamento em Tradições Discursivas nos séculos XVIII a XXI. Os resultados apontam para a uma estreita relação entre as tradições investigadas e um caminho de mudança que se revela ainda em curso.

**Palavras-chave:** gramaticalização; junção; contraste; tradição discursiva.

## Introdução

Neste artigo,<sup>1</sup> resultado de trabalho desenvolvido no âmbito do Projeto para a História do Português Paulista (Projeto Caipira), analiso o funcionamento da locução conjuntiva *mesmo/ainda assim*, priorizando a abordagem de seu desenvolvimento, via processo de gramaticalização (GR), em diferentes Tradições Discursivas (TDs), no recorte temporal compreendido entre os séculos XVIII e XXI.

O conceito de TD é aqui definido a partir de um laço que se estabelece entre *atualização* e *tradição via repetição* de um texto, ou de uma forma textual, que caracterize uma maneira particular de escrever ou falar, significável e, portanto, com valor de signo propriamente dito. Uma TD pode se formar em relação a qualquer finalidade de expressão ou com qualquer elemento de conteúdo (KABATEK, 2005, p. 159), a partir da reduplicação do nível histórico de Coseriu (1979).

A GR, por sua vez, é entendida com base na concepção de Traugott (1982, 2003), Traugott e König (1991), entre outros, priorizando a análise dos aspectos semânticos e pragmáticos envolvidos no processo. Considero que duas principais abordagens fundamentam a apreensão de trajetórias de GR, focalizando a mudança categorial, a saber: (i) uma que considera o percurso *léxico* > *gramática*, muito postulado desde Meillet (1965 [1912]), no qual um item lexical torna-se gramatical ou um item gramatical torna-se ainda mais gramatical, e (ii) outra que considera o percurso *discurso* > *morfossintaxe*, postulado por Givón (1979). Este trabalho, cf. Traugott (1995), fundamenta-se na combinação dessas

<sup>1</sup> Parte de tese de doutorado defendida em 05/2011, na UNESP-IBILCE (Fapesp/Proc. n. 07/07955-5).

duas abordagens, segundo a qual (iii) *um item lexical* > *discurso*. Nessa direção, a GR é definida como a pragmatização gradual do significado, que envolve estratégias de caráter inferencial e metafórico.

O enfoque na mudança semântica, segundo a trajetória *concreto* > *abstrato*, leva a motivações de ordem cognitiva e comunicativa que se envolvem mutuamente no processo. As motivações cognitivas estão relacionadas à busca por resoluções de problemas de representação, no âmbito léxico, enquanto as comunicativas, à resolução de problemas associados ao intuito do falante de ser informativo e relevante, no âmbito, portanto, do discurso/texto. Para a resolução dos dois tipos de problemas atuam as mudanças viabilizadas metafórica e metonimicamente (TRAUGOTT; KÖNIG, 1991; HEINE et al., 1991; HOPPER; TRAUGOTT, 1993).

Dessa forma, de acordo com Heine et al. (1991), a GR pressupõe a existência de forças distintas, sendo uma motivada pragmaticamente, de natureza metonímica, associada à emergência de funções e significados que se sobrepõem numa estrutura em cadeia; e outra motivada cognitivamente, de natureza metafórica, relacionada a domínios cognitivos distintos. Assim, cada enfoque atua cooperativamente de modo a sinalizar aspectos distintos do processo de mudança focalizado.

A hipótese que dirige este trabalho parte, pois, de uma vinculação teórica entre GR e TD e baseia-se na aceitação de que a adoção de novas TDs tem servido, na história da língua, como motor para inovações e mudanças, via GR, o que leva ao pressuposto de que, a cada nova tradição, se dá uma busca por meios linguísticos apropriados, podendo levar tanto à conservação do que já existe no sistema, como à criação de algo novo. Dessa forma, é hipotetizada uma relação entre TD e o caminho de mudança da locução *mesmo/ainda assim*, capaz de apontar para especificidades desse processo de mudança no que tange à implementação de seus mecanismos fundamentais, a partir do pressuposto de que as características internas e externas de cada TD são determinantes para o processo de inovação linguística.

Nessa direção, com o intuito de evitar uma homogeneização do material de análise e, conseqüentemente, resultados inconsistentes, o estudo proposto deixa de ser apenas de um século a outro e passa a comparar resultados analíticos de textos diferentes, em “tempos” diferentes, segundo uma concepção de diacronia *não ideal* que acarreta um novo ângulo de focalização do processo de GR. Essa postura torna necessária uma análise em correlação a cada TD a fim de determinar o funcionamento do item, de modo a refletir suas transformações nas TDs focalizadas.

### **Material de análise: um recorte teórico-metodológico**

O *corpus* organiza-se segundo: (i) as TDs investigadas; (ii) a delimitação geográfica (Estado de São Paulo); e, (iii) a delimitação temporal (séculos XVIII a XXI). Foram analisadas amostras compostas por 274 textos referentes a cada TD dos *corpora*:

- (A) diacrônicos, representativos das TDs *carta* e *editorial*. O *corpus* da TD *carta* divide-se em: (i) *Administração Privada* (séc. XVIII e XIX); (ii) *Documentos Pessoais* (séc. XIX ao XX); e (iii) *Cartas de leitores e redatores de jornais* (séc. XIX e

XX). O *corpus* da TD *editorial* compõe-se de textos do jornal *O Estado de S. Paulo* desde a sua fundação até 1964; e

- (B) sincrônicos, compostos por: (i) TDs que constituem o Banco de Dados Iboruna (amostras de fala do Noroeste Paulista), a saber: TDs narrativa de experiência, narrativa recontada, opinativa, injuntiva e descritiva;<sup>2</sup> e (ii) TD *e-mail*.

Nas seções seguintes, apresento, numa perspectiva teórico-analítica, o domínio, caracterizado por flutuações semântico-cognitivas e categoriais, em que a locução *mesmo/ainda assim* será analisada, qualitativa e quantitativamente, mediante exposição de seus usos prototípicos e não-prototípicos, com o intuito de flagrar indícios de seu processo de GR, para, nas Considerações finais, traçar um quadro geral da GR dessa locução em correlação às TDs focalizadas.

### **As relações contrastivas: no domínio da flutuação**

Há uma grande flutuação existente entre as construções *concessivas* e *adversativas*. Segundo Koch (2001), do ponto de vista semântico, os jutores adversativos e os concessivos têm funcionamentos semelhantes: opõem enunciados de perspectivas diferentes, que orientam para conclusões contrárias, instaurando uma relação de *contraste*. Também para Halliday e Hasan (1976), tanto as construções adversativas quanto as concessivas são mediadas por uma relação contrastiva, cujo significado básico é a contrariedade à expectativa. Nessa linha, Chen (2000) afirma que semanticamente a noção de concessão é muito similar à de contraste e inerentemente relacionada à de adversidade. Como Quirk *et al.* (1985) pontuam, existe frequentemente uma mistura de contraste e concessão. Assim, é possível argumentar que uma relação concessiva implique uma relação adversativa. Diante disso, pressupõe-se que alguns mecanismos e/ou contextos relacionados à emergência de jutores concessivos podem ser fundamentais para a emergência de adversativos, em nível sintático, e, ao mesmo tempo, casos de grande flutuação em nível semântico-pragmático.

Segundo König (1985, p. 5), as adversativas, em nível da coordenação/parataxe, são definidas como relações entre proposições que suportam conclusões contraditórias, sendo que o ponto principal do falante se expressa pela segunda proposição (Q). As concessivas, ao contrário, são baseadas em relações condicionais/causais entre eventos em nível da subordinação/hipotaxe. Apesar dessas diferenças, o autor destaca que não deve ser obscurecido um ponto essencial de similaridade: em ambos os casos, *algo* é suspenso – na relação adversativa, o significado da primeira proposição (P) como argumento em favor de uma conclusão R, à qual o falante/escritor quer que o ouvinte/leitor chegue, e, na concessiva, a aplicabilidade de certas normas concernentes às relações causais ou condicionais para uma situação corrente.

Além disso, segundo König (1985, p. 7), alguns jutores (*apesar de, ainda que, por exemplo*) somente expressam relações concessivas, enquanto outros (*embora, por exemplo*) podem ser usados para os dois tipos de relações. Tomando esse fato juntamente

---

2 Considero a formação de uma TD a partir da combinação de funções comunicativas, cognitivas e instruccionais, bem como suas peculiaridades linguístico-estruturais. Assim, a TD diferencia-se do gênero textual, mas pode incluí-lo, bem como os tipos textuais. Portanto, da mesma forma que *cartas* e *editoriais* constituem TDs, podendo englobar outras TDs, a partir da composicionalidade da tradição que representam, os diferentes tipos textuais também constituem TDs.

com o de que nem todas as línguas apresentam juntores concessivos enquanto todas parecem ter seus adversativos – como *mas* –, sugere-se que a relação de adversidade seja mais geral e básica e que a relação de concessão seja uma variedade específica da primeira. Diante disso, pressuponho que, a partir de uma relação basicamente adversativa, pode-se inferir, por conta da própria fluidez semântico-pragmática, uma relação concessiva, mais abstrata, a depender do contexto.

A fluidez da relação entre adversidade e concessão fundamenta-se ainda a partir do exame cuidadoso das fontes de GR de juntores concessivos, que, segundo Chen (2000, p. 91), apresentam-se sob uma grande variedade, relacionada, de uma forma ou de outra, às propriedades sintático-semântico-pragmáticas das concessivas, em expressões que: (i) concedem a existência de uma situação adversa, como os SVs *admitir/conceder*; (ii) enfatizam a adversidade de uma situação concedida, como *mesmo/ainda assim*; (iii) afirmam a ineficácia de uma situação, como *apesar de/todavia*; (iv) declaram a concomitância de duas situações expressas pela antecedente (P) e pela conseqüente (Q), como *embora*; (v) ajudam a afirmar a factualidade da situação expressa pela conseqüente (Q), como *apesar disso*.<sup>3</sup>

Neves (1999) afirma que a relação que se estabelece entre essas construções não é simples, da mesma forma como também não o é a implicação que pode ser apontada entre as relações causais, condicionais e concessivas, “todas elas expressivas de uma conexão ‘causal’ *lato sensu* entendida, e todas elas explicáveis em dependência da (não) satisfação de necessidade/de suficiência de determinadas condições” (p. 545).<sup>4</sup> Torna-se evidente a ligação entre concessão, não satisfação de condições e frustração de uma causalidade possível. Aqui, admitindo a fluidez da relação entre *adversativas* e *concessivas*, tomarei como pressuposto o fato de que possíveis implicações para o desenvolvimento de uma podem ser relevantes para o desenvolvimento de outra.<sup>5</sup>

Em relação à manobra argumentativa, nas *adversativas*, emprega-se, segundo Guimarães (1987), a *estratégia do suspense*, que faz vir à mente do interlocutor/leitor uma conclusão R, para depois introduzir um (conjunto de) argumento(s) que levará à conclusão ~R. Na *concessiva*, o locutor utiliza a *estratégia da antecipação*, anunciando de antemão que o argumento introduzido por esse jutor será anulado. É importante lembrar, mais uma vez, que a possibilidade de alguns juntores prototipicamente concessivos, como *embora*, surgirem encabeçando Q torna, novamente, a relação entre adversativas e concessivas marcada pela flutuação.

Esse raciocínio sinaliza a base argumentativa da construção concessiva que representa, segundo Neves (1999, p. 562-563), o ponto exato em que podem ser examinadas as similaridades e diferenças entre concessivas e adversativas. Para ilustrar essa colocação, parto de uma adaptação dos exemplos expostos pela autora a fim de correlacionar ocorrências concessivas, em que se *refuta* uma objeção, com possíveis formulações adversativas, em que se *admite* uma proposição:

3 Vale lembrar que *apesar disso* permite paráfrase por *mesmo/ainda assim*.

4 Para Hermodsson (1994, p. 93 apud NEVES, 1999, p. 546), a concessiva qualifica-se como negação, de nível sintagmático, da relação normal suposta entre as proposições citadas na premissa maior e na menor.

5 Harris (1988) afirma que existe uma relação semântica que se estende no seguinte *continuum*: orações causais > condicionais > condicionais-concessivas > concessivas. Enquanto no extremo à esquerda a relação de causa é afirmada, no extremo à direita, esse vínculo causal é negado. Assim, uma causa ou condição é expressa na oração concessiva, mas o que se espera a partir dela é negado na oração nuclear.

- (a) *embora* muito forte, eles fazem um molho muito gostoso.
- (b) (o molho) é muito forte, *mas (mesmo/ainda assim)* é muito gostoso.

Neves (1999, p. 563-564) ilustra essa operação argumentativa com os esquemas:

*Concessivo:*

- (a') alguém *pode me objetar* que o molho que eles fazem é muito forte, e eu *não desconheço* isso;
- (a'') (*mesmo/ainda assim*) o molho é muito gostoso.

*Adversativo:*

- (b') eu *admito* que o molho que eles fazem é muito forte;
- (b'') (*mesmo/ainda assim*) o molho é muito gostoso.<sup>6</sup>

Chamo a atenção, também, para aspectos de natureza sintática, relacionados à integração oracional, que caracterizam a diferença tradicionalmente observada entre essas orações, já que as *concessivas* equivalem, prototipicamente, às subordinadas adverbiais, no âmbito da *hipotaxe*, enquanto as *adversativas*, às coordenadas, no da *parataxe*. Por hipótese, as orações paratáticas são gradualmente menos vinculadas sintaticamente do que as hipotáticas e desse fato decorrem algumas distinções em cada um dos tipos de relações interproposicionais. Para ilustrar essas considerações, apresento as orações abaixo, adaptadas de Martelotta (1998, p. 42):

- (c) Ele correu muito, *mesmo/ainda assim* não se cansou.
- (d) *Mesmo/Ainda que tenha corrido muito* não se cansou.

Em (c), ocorre a parataxe com o conectivo expresso em Q, sendo que a vinculação entre P e Q limita-se à referenciação anafórica realizada por *assim*, na locução em Q, responsável por *retomar* P como um todo. Já em (d), P assume o lugar do elemento anafórico, penetrando em Q e indicando um grau maior de integração. Se (c) exemplifica *adversidade* e (d), *concessividade*, tais colocações ilustram diferenças entre essas duas acepções no que diz respeito ao grau de integração oracional.

Segundo Martelotta (1998, p. 43), um reflexo dessa distinção relaciona-se às incidências adverbiais, isto é, nas adversativas, um advérbio (como *com certeza*) incide apenas sobre P, constituindo Q um argumento referente à P, mas distinto dele e fora do escopo adverbial, o que não ocorre nas concessivas, em que esse escopo incide sobre toda a construção, referindo-se à relação semântica global estabelecida entre as orações:

- (e) Com certeza, *a vida é boa*, mesmo/ainda assim devemos levá-la a sério.
- (f) Com certeza, *embora a vida seja boa* devemos levá-la a sério.

Também a presença de formas reduzidas indica um grau maior de vinculação oracional, por isso tende a não ocorrer nas paratáticas. Além disso, é possível a inversão da ordem

6 O misto concessivo-adversativo torna-se mais evidente quando ambos os juntores vêm expressos na mesma oração. Nesses casos, a concessiva, ao mesmo tempo, expressa refutação a uma possível objeção e assentimento referente a alguma validade dessa objeção. Nesse tipo de construção, determina-se uma ordem fixa, na qual a refutação precede a admissão, diferentemente das construções tradicionalmente concessivas, em que a ordem é livre.

sintática quando se trata de uma concessiva (característica das hipotáticas),<sup>7</sup> mas não de uma adversativa, cf. as paratáticas, cuja ordenação obedece a uma relação icônica entre a ordem dos fatos e dos argumentos lógicos.<sup>8</sup>

### Aspectos da gramaticalização de *mesmo/ainda assim*

Considerando as distinções entre a natureza *concessiva* e *adversativa* das construções e também as flutuações verificadas nesse domínio, incluindo aí, expressivamente, também as relações *causais* e *condicionais* (cf. CHEN, 2000; GUIMARÃES, 1987; HALLIDAY; HASAN, 1976; KOCH, 2001; KÖNIG, 1985; MARTELOTTA, 1998; NEVES, 1999; HARRIS, 1988), denominarei a relação *P, mesmo/ainda assim Q*, em seus usos mais gramaticalizados e prototípicos, como *contrastiva* a partir da constatação do traço de contra-expectativa.<sup>9</sup>

A locução juntiva *mesmo/ainda assim* apresenta, em seus usos mais gramaticalizados, aqui também denominados de mais prototípicos, as seguintes características semântico-formais:

- (a) tipo de unidade articulada: oração;
- (b) sem ocorrência com outra conjunção;
- (c) posição inicial de Q;
- (d) sem incidência adverbial sobre toda a onstrução;
- (e) sem ocorrência de formas verbais reduzidas;
- (f) impossibilidade de inversão da ordem;
- (g) relação de independência entre P e Q;
- (h) não compartilhamento de estruturas diversas;
- (i) relação de sentido: valor contrastivo.

(LOPES-DAMASIO, 2011, p. 252)

A ocorrência (1), extraída de um editorial, ilustra um uso prototípico da locução:

- (1) Quer nos parecer que hoje há vergonha| em confessar-se a verdade inteira. || **Ainda assim**, estas informações não [levam] | o sr. conselheiro director das terras | e colonização a formular um desmentido á | imprensa paulista, e particularmente a nós. [APSPXIX-1876/026]

Em (2),<sup>10</sup> *ainda assim* articula as orações P e Q, em posição inicial de Q, sem coocorrência de outra conjunção ou incidência adverbial, sendo P uma reduzida:

7 A inversão da ordem sintática de P e Q não é gratuita, já que, conseqüentemente, passa-se do domínio do conteúdo para o domínio epistêmico (SWEETSER, 1990).

8 Para Sweetser (1990), nossa expressão linguística do mundo interno é modelada a partir de nossa expressão do mundo real. Assim, a anterioridade epistêmica, visualizada, por exemplo, na premissa antes da conclusão, reflete, iconicamente, a anterioridade temporal desses eventos no mundo. Portanto, espera-se que, na relação contrastiva, a quebra de expectativas (Q) se dê em relação àquilo que vem antes (P).

9 Segundo Heine et al. (1991, p. 192), a noção de expectativa relaciona-se ao fato de que as línguas apresentam formas de expressão para codificar *situações que correspondem às normas compartilhadas e situações que se desviam dessas normas*, sendo apenas estas codificadas gramaticalmente. Os elementos marcadores de contra-expectativa (*contrastivos*) codificam esse segundo tipo de situação. Seu uso implica uma comparação entre o que é dito e o que se espera/pressupõe em relação à norma.

10 As ocorrências de (2) a (6) foram extraídas de cartas.

- (2) É neste tempo que eu chego defora em 6 do|corrente e sendo já nomeado por esta Villa a dita Junta, vime obrigado|a tomar parte nos negocios, temendo, que o excesso de enthusiasmo não|degenerasse em males internos, e externos. Estando Membros d'|algũa Villas, *ainda assim* tenho feito, que se não tenha installado|a Junta, a espera de Sorocaba, que ainda não o quiz mandar (apezar |de que está tambem com bons sentimentos) [...]. [BNXIX-11/03]

Não existe, nesse contexto, uma relação *contrastiva* veiculada pela locução, isto é, P descreve uma situação factual/verdadeira em relação à qual Q também descreve outra situação factual/verdadeira que não representa uma situação inesperada/incompatível com a informação em P, mas uma consequência do que foi aí enunciado. Assim, há uma relação de causa-consequência, que permite a paráfrase de *ainda assim* por *por isso*, mas não por *apesar disso*, a partir da acepção de *assim* não reanalisada com *ainda*, conforme seguinte reescrita do segmento, que apresenta uma única modificação na localização do sinal de pontuação (vírgula) que, normalmente, acompanha essa acepção funcional de *assim* (hoje!):

- (2') Estando membros de algumas vilas ainda, *assim*, tenho feito[...]

A ocorrência ilustra as relações entre causa e concessão (CHEN, 2000; NEVES, 1999; KÖNIG, 1985), refletidas, na diacronia, em contextos que, embora sintagmaticamente favoreçam a leitura contrastiva, cognitivamente revelam a relação causal, significando que um contexto marcado pela noção de causa pode favorecer o desenvolvimento de traços contrastivos e, dessa forma, a mudança linguística.

A presença da reduzida e da relação semântico-conclusiva impossibilita a inversão da ordem das orações e, por conta disso, estabelece-se certo grau de dependência entre P e Q, no sentido da completude semântico-formal.

Em (3), abaixo, novamente não se verifica o *contraste* marcado por *assim mesmo*, mas uma relação, não totalmente gramaticalizada, de coordenação conclusiva veiculada por *e assim*, seguida por uma relação condicional veiculada por *mesmo se*. Não se trata, portanto, de uma leitura reanalisada da locução contrastiva *assim mesmo*.

- (3) [...] faltas de equi= | dade em manobras mais sublimes, vejo eu que me cauzaõ pena, efa= | zem viver por ca com desgosto, eamuito tempo estaria em Lisboa, naõ obitive sa= | ber muito bem os termos, emque está de caristia de viveres, se me naõ os pesados | grilhoens damesma familia, *easim mesmo* se mefoce pocivel, cobrar dehum | golpe oque me devem, hiria acabar os meos dias em Lisboa perto do | meo Principe, epodendo euaqui dizer a V. S<sup>a</sup>. muitas couzas, mere= meto ao silencio. [BNXIX-19/11]

Dessa forma, *mesmo* integra com *se* uma locução condicional, que estabelece a relação *mesmo se P, Q*, em que P apoia-se em uma hipótese, nessa ocorrência, observada em *me fosse possível*, representativa de uma *condição de realização* de Q, entendida como o resultado da condição enunciada em P (*iria acabar os meus dias em Lisboa*). Ambas as relações, a de conclusão (não-prototípica), expressa por *e assim*, e a de condição, expressa por *mesmo se*, compartilham a mesma oração Q.

Esse contexto é relevante para o disparo da emergência da reanálise estrutural de [assim] [mesmo] a [assim mesmo], bem como para o da emergência do significado *contrastivo* da locução, favorecendo a reinterpretação dos limites semântico-formais induzida pelas

contingências contextuais, isto é, por processo metonímico (pela contiguidade formal, em relação à reanálise estrutural, e pela contiguidade dos significados – contrastividade e condicionalidade – em relação à reanálise semântica).<sup>11</sup>

A ocorrência (4) demonstra contexto de ambiguidade semelhante a (1):

- (4) O dito quintal seacha imaberto, por que se||hefazem algum concerto, como ja vi praticado peloz Indios, para vedar as criaSsoens dediferentes vezinhos, que|aly entraó apastar todo anno, vem o fogo Com so|me o mato, ou feital, que antes estava cobrindo as|ruinas, ou fraqueza dos valos, reconheassem as cri|aSsoens as partes por onde podem fazer assua en|=|trada, e por estes passáo quando emcontraó segu|ranSsa nas partes retificadas, oque naó aconte|se hoje ou há annos, por que todo seacha fran|co e se conserva, como campo comum; esse|alguns Indios tem alguns retalhinhos do mesmo|quintal atacados com cercas saó taó lemitados|osterrenos, que nomeo conseito so daó para pouco|mais que huma orta de repolhos, ficando *assim|mesmo* parte da Aldeã, eaIreja [sic] sem defeza|das criaSsoens, [AIXVIII-24/62]

Em (4), *assim mesmo* ocorre após o verbo da reduzida que constitui Q, em posição medial, sem a coocorrência de outra conjunção ou incidência adverbial. A ambiguidade configura-se pela possibilidade de duas interpretações distintas para o mesmo contexto. Na primeira, cf. (1), a relação *contrastiva* não é observada, já que P e Q descrevem situações factuais/verdadeiras e que não representam uma situação inesperada/incompatível. Segundo essa interpretação, o fato de os índios terem *retalhinhos* de quintal, atacados com cercas, é insuficiente para modificar o fato de as suas criações continuarem sem defesas. Dessa forma, a locução *assim mesmo* pode ser parafraseada por *do mesmo modo* e permite uma relação em que todo o ST, em P, em especial *porque tudo se acha franco e se conserva como campo comum*, funciona numa relação de causa-consequência com Q: *ficando da mesma forma parte da aldeia e a igreja sem defesa das criações*. Na segunda interpretação, a relação *contrastiva* é agora veiculada pela locução se for considerado um contexto mais específico do tópico, marcado pela oração condicional *se alguns índios tem alguns retalhinhos do mesmo quintal, atacados com cerca*, em P. Essa informação funciona como base para a inferência de que há *defesa das criações, já que os índios têm seus quintais atacados com cercas*, contrastada pela afirmação contrária em Q. Dessa forma, a locução *assim mesmo* resgataria, em Q, exatamente o conteúdo da condicional, podendo ser parafraseada por *apesar disso*.

Toda a complexidade semântico-formal, principalmente de P, nesse caso, gera a ambiguidade da acepção e constitui um ambiente propício para a mudança, uma vez que mesmo que se queira transmitir A, não se pode evitar, a partir de um tópico como esse, que se interprete B. Soma-se a isso a correlação entre causa, condição e concessão, novamente relacionada ao desenvolvimento da acepção contrastiva.

Em (5), *assim* e *mesmo*, no início de Q, sem coocorrência de outra conjunção ou incidência adverbial, articulam P e Q, já reanalisados na locução com valor *contrastivo*.

<sup>11</sup> Harris (1988, p. 75) apresenta seis padrões que auxiliam na compreensão da natureza dos conectivos e, sobretudo, dos contextos que refletem a relação de contrastividade. São eles: (i) O padrão indefinido, portanto total (*Whatever you may do, nevertheless...*); (ii) O padrão total explícito (*All that you may do, nevertheless...*); (iii) O padrão volitivo (*Let it be as you wish/ let it be so, nevertheless...*); (iv) A marca explícita de posição escalar extrema (*Most X that it may be, nevertheless...*); (v) Padrão temporal, frequentemente reforçado (*While/when...nevertheless...*) e; (vi) O padrão condicional, frequentemente reforçado (*Even if/ though He did it, nevertheless...*). Aqui, interessa especificamente o padrão (vi).

- (5) Eoutro Ssim por que hay huãs profecias | de Como vossaexelencia não hade hir a Cujaba SenoSso | Senhor ouvir osRogos dospeccadores aSsim mesmo | hadeSer por que me parece não ha' quem | Sedescuide Com esta deligençia inda que | Seria Com muito dispendio de vossaexelencia Seadita pro | feçia sahir Certo[...] [AIXVIII-11/48]

Sintaticamente, essa construção pode resultar da elisão da forma verbal *sendo* (*mesmo [sendo] assim*, cf. LOPES-DAMASIO, 2011; 2012). Nesse contexto, *mesmo* funciona como operador de subordinação com interpretação concessiva/contrastiva, semelhante àquela que se obtém com *embora*, enquanto *assim* retoma anaforicamente a proposição precedente, P, em que há uma asseveração com a admissão de V. Exa. *não ir a Cuiabá*. Em Q, por sua vez, está a não aceitação da inferência realizada a partir de P, gerando o contraste. Novamente, observa-se, em P, a condicional *se nosso Senhor ouvir os Rogos dos pecadores*, antes da inserção de Q. A relação de quebra de expectativas que gera o contraste, via negação de inferência existente entre P e Q, é perpassada pela condição, revelando, mais uma vez, a importância do contexto – condicional/contrastivo – para o desenvolvimento dessa última acepção, ainda que em condições distintas daquela analisada em (3), uma vez que aqui a locução já está reanalisada.

O segmento *há de ser* de Q ativa o compartilhamento da informação *ir a Cuiabá*, de P, desfazendo a negativa presente nesse segmento e estabelecendo o contraste, marcado pela locução.

Frente a essas constatações, a relação de ordenação entre P e Q é marcada, tanto em nível formal, pela foricidade estabelecida por Q em relação a P, quanto em nível semântico, pela colocação em P que funciona como base para a inferência que contrasta com a colocação em Q, sendo esse *contraste* marcado pela locução *assim mesmo*. Torna-se, portanto, agramatical uma inversão da ordem de P e Q.

Em (6), *assim mesmo* coocorre, no final de Q, com *mas*, responsável pela articulação de P e Q, em contexto de inserção parentética em Q:

- (6) Além disso desde março que|estou lecionando Inglês e Português; o que|mais custa é ser à noite, mas como|não tinha nada, aceitei *assim mes-|mo*. [FFXX-54f/136]

Como já destaquei, em P, há uma asseveração, com a admissão de um fato e, em Q, a não aceitação da inferência a partir do que foi veiculado em P. Diferentemente de (5), a insuficiência da asseveração, em P (cuja forma não é a de uma asseveração propriamente dita), para permitir a inferência contrastiva, é indicada por *assim mesmo*, justificando a ocorrência de *mas*, que marca o contraste e, no final de Q, a de *assim mesmo*, que reforça esse contraste. No segundo membro coordenado, o parêntese insere um argumento não considerado *como não tinha nada*, avaliado como mais relevante do que o anterior, no primeiro membro, embora esse argumento, em P, seja admitido, justificando a relação contrastiva e, assim, o uso de *mas* e de *assim mesmo*.<sup>12</sup>

A inversão da ordem de P e Q torna a construção agramatical graças: (i) à articulação realizada por *mas*, indicando o princípio icônico de ordenação desses enunciados coordenados, e (ii) à relação fórica estabelecida por *assim mesmo*, ainda que possa haver mobilidade da locução em Q após *mas* (pré ou pós verbal).

12 Além da relação contrastiva, mais abstrata, a locução *assim mesmo* permite paráfrase por *do mesmo modo*, revelando a ambiguidade semântica (MODO > CONTRASTE) que perpassa seus usos não-prototípicos, com acepção mais concreta.

Nas distintas TDs presentes no IBORUNA, ocorrem casos cf. (6), em que, tanto em posição final de Q (7), como após *mas*, em posição pré verbal (8), a locução indica a não aceitação da inferência realizada a partir da asseveração, em P, e marcada inicialmente por *mas*, enfatizando a acepção contrastiva do complexo:

- (7) um po(u)quinho só... e:: cada duas horas eu tenho que comê(r) um po(u)quinho de coisa porque meu estômago tá pequenininho né?... mas eu me conformo *MESmo assim* porque eu saí dos cento e trinta quilos... [AC-152/NE1222]
- (8) chegamo(s) em Noronha pegamo(s) um aviãozinho pequeno -- óh eu dentro dos aviões de novo -... só que aquele eu senti medo... tomei Dramim tudo pra num vomitá(r) mas *mesmo assim* passei muito mal... [AC-051/DE095]

Esse tipo de ocorrência representa um estágio em que a locução, atreladamente ao seu elo anafórico – que, como germe de sua função juntiva, reforça a relação entre Q e P, já marcada por *mas* –, desempenha a função de reforçar o contraste, veiculado por este item, prototipicamente responsável pela junção com acepção contrastiva.<sup>13</sup> Ressalto, novamente, a ambiguidade, nesse uso, perpassado pela acepção modal.

O *corpus* da TD *e-mail* apresenta uma única ocorrência, exposta abaixo, relacionada a esse padrão e à situação de ambiguidade semelhante àquela analisada cf. (6), (7) e (8). No dado de *e-mail*, entretanto, *mesmo assim* encontra-se no final da coordenada Q, sem a presença de outro elemento para o estabelecimento do contraste a partir da não aceitação da inferência baseada na asseveração em P. Apesar disso, também são possíveis ambas as leituras, contrastiva e modal, a partir de paráfrases por *apesar disso* e *do mesmo modo*:

- (9) Oi, T., olha eu de novo te enchendo!!!! || Em algumas cartas q são e-mail, não aparece o local, || apenas o e-mail da pessoa. Devo utilizá-las *mesmo assim*? [...] [EMAILXXI-13/21D]

A TD editorial também oferece contextos importantes para a descrição do desenvolvimento do funcionamento dessa locução juntiva, no que diz respeito à contiguidade sintagmática e às estratégias semântico-cognitivas realizadas por ela, mas, em todas as ocorrências, já apresenta a locução reanalisada. Nessa TD, foram constatadas, por exemplo, ocorrências relacionadas a certa fluidez argumentativa, como pode ilustrar (10):

- (10) *Dado mesmo que* haja reflectido e mudado| de opinião [ininteligível] os srs. ministros da fazenda| e presidente do conselho, *ainda assim* ha| ahi uma questão melíndrosa e gravíssima a| decidir: é licito á corôa *demittir* um ministro| de estado estando o parlamento a funcionar| e sem que este se tenha manifestado? [APSPXIX-1879/058]

Em (10), *ainda assim* encontra-se em contexto que articula Q a P, sendo P iniciada por construção parafraseável por *embora*. A concessiva, iniciada por *dado mesmo que*, ao mesmo tempo, expressa refutação a uma possível objeção e assentimento referente a alguma validade dessa objeção, na oração iniciada por *ainda assim*. Dessa forma, o contexto mostra uma forte mescla entre o funcionamento de *embora*, aí representado por outra construção, e *ainda assim*, uma vez que P inicia-se apontando para a antecipação da negação da inferência que será realizada *a posteriori* em Q, sendo que, em Q, gramaticaliza-se novamente a negação dessa inferência por *ainda assim*. Nessa ocorrência, a

<sup>13</sup> Portanto, considero usos como (7), em que *mesmo assim* localiza-se no final de Q, como representativos de contextos importantes para o desenvolvimento da função juntiva da locução, desempenhada prototipicamente aí pelo juntor *mas*.

paráfrase da locução por *mas* não é aconselhável, o que reforça a importância do contexto para a emergência das características mais concessivas no funcionamento de *ainda assim*.

Outros contextos dessa locução, na TD editorial, revelam o mesmo tipo de fluidez, por sua vez, com o item *mas*, cf. (11):

- (11) Tudo isso explica-se; muita coisa desculpa-se e tolera-se, até certo ponto: *mas ainda/assim* é preciso que a população estranha álfesta tenha certeza de que anda resguardada|pela vigilancia policial, [...] [APSPXIX-1876/017]

Nesse contexto, já observado em ocorrências extraídas das TDs do IBORUNA, ocorre a estratégia argumentativa de suspense, como nos usos adversativos, em que P encaminha para uma inferência conclusiva R e Q leva à sua negação, acarretando não-R. *Ainda assim*, como reforço argumentativo, mostra que, apesar de se admitir a proposição em P, é em Q que há a asseveração do fato mais relevante. Assim, as relações contextuais dessa locução mostram, mais uma vez, o caráter fluido no que tange às características concessivas e adversativas.

Na ocorrência (12), também extraída de um editorial, além de o contexto apresentar a coocorrência de *ainda assim* e *mas*, a condicionalidade está também presente. A primeira oração apresenta um argumento *maior seria também a sua riqueza e prosperidade* sob uma condição, indicada por *Si*. No período iniciado por *mas*, contraria-se a inferência realizada a partir do argumento anterior e, por fim, o último período, articulado por *ainda assim*, enuncia um argumento não considerado que, novamente, admitindo o anterior (indicado por *mas*), apresenta o que é mais relevante. Aqui, cf. constatado em ocorrências da TD carta, a relação de quebra de expectativas que gera o contraste, via negação de inferência, relaciona-se com a aceção de condição, próxima contextualmente.

- (12) *Si*, pois, a provincia fosse possivel ap-| plicar maior parte de sua renda a esses| dous poderosos factores do seu progresso, | maior seria também a sua riqueza e pros-| peridade. || *Mas* a centralisação pêa-lhe os movi-| mentos, abafa as aspirações e impõe uma | distribuição de rendas em grande parte| com destino improdutivo, mesmo com re-| lação ao todo. || *Ainda assim* devem os paulistas em| cada anno fazer a conta do Deve e Haver| com o imperio para ter certeza do quan-| to vale na união, como esta estabelecida| no regimen actual, que parece não com-| prendel-a sem as fortes cadêas da cen-| tralisação. [APSPXIX-1881/072]

Além dos contextos de ambiguidade já expostos, nas ocorrências das TDs do IBORUNA, observam-se outros, igualmente importantes para o processo de desenvolvimento, refletido sincronicamente, em que a interpretação dos itens como componentes de uma locução com aceção contrastiva depende, imprescindivelmente, de sua realização prosódica:

- (13) foram seis meses... saindo::... aí voltei com e::le... meio na escond/ escondida assim... minha avó num sabe *ainda:: assim* alguns parentes sabem... meu pai e minha mãe sa::bem tudo [AC-052/NE163]
- (14) mas é uma organização mui::to bacana... as ruas bem largas... pra gente fazê(r) caminhada e dá pra passá(r) os carros *assim mesmo*... porque lá a gente num caminha pelas calçadas... [AC-152/DE1239]

Em (13), a sintaxe, a semântica e até a pragmática permitem a interpretação de P como “minha avó não sabe” e Q, introduzida pela locução contrastiva, “*ainda assim* alguns parentes sabem...”, baseando a leitura na quebra de expectativas que representa

outros parentes saberem do namoro se a avó não sabe. Ter-se-ia um uso prototípico da locução. Entretanto, apenas a audição do inquérito e uma análise prosódica revelam um uso de *ainda* e *assim* não reanalisado na locução, de forma que *ainda* mantém sua acepção temporal, associando-se ao contorno entoacional descendente de “minha avó não sabe *ainda*”, com pausa separando esse enunciado prosódico do seguinte, iniciado por *assim*, realizado com contorno entoacional ascendente, sinalizando cataforicamente o enunciado “alguns parentes sabem... meu pai e minha mãe sa::bem tudo”. Apesar disso, esse contexto evidencia um ambiente propício para a reanálise formal, na constituição da locução, dado que, semanticamente e pragmaticamente, a acepção contrastiva pode ser inferida a partir dele, cf. sugere leitura do tipo: *minha avó num sabe (ainda)... ainda assim alguns parentes sabem...*

De forma semelhante, em (14), apenas a análise prosódica garante, agora, a interpretação contrastiva da locução no complexo, com P, “as ruas bem largas... pra gente fazê(r) caminhada”, e Q “e dá pra passá(r) os carros *assim mesmo*”, ao invés da análise de *assim* como Sinalizador de construção de Quadro Mental (cf. LOPES-DAMASIO, 2011), associado à descrição “e dá pra passá(r) os carros *assim*”, e *mesmo* ligado a *porque* em “*mesmo* porque lá a gente num caminha pelas calçadas”. Ainda a leitura não atestada pela análise prosódica,<sup>14</sup> não deixa de evidenciar o significado contrastivo do complexo, representando, novamente, importante contexto para a constituição e solidificação da locução contrastiva *assim mesmo*.

Portanto, nos usos mais gramaticalizados, que estou chamando aqui de usos prototípicos, a locução é empregada em orações consideradas livres e funcionalmente independentes. Para uma melhor ilustração desse funcionamento, recupero os pressupostos de Halliday (1985), conforme quadro abaixo, em que à notação 1 (inicial) e 2 (continuação), na parataxe, e  $\alpha$  (dominante) e  $\beta$  (dependente), na hipotaxe, acrescento P e Q:

**Quadro 1: Parataxe e Hipotaxe**

	PRIMÁRIA	SECUNDÁRIA
PARATAXE	1P (inicial)	2Q (continuação)
HIPOTAXE	$\alpha$ P (dominante)	$\beta$ Q (dependente)

De acordo com a análise realizada neste artigo, em seus usos mais prototípicos, em 2Q, *mesmo/ainda assim* representa uma combinação de *realce* com a *parataxe* (1P x 2Q), produzindo o que é também um tipo de coordenação. A oração introduzida por *mesmo/ainda assim* integra a categoria *causal-condicional* apresentada por Halliday (1985, p. 213), marcando um significado de concessão-consequência, que, aqui, denominei de relação *contrastiva*, ao invés de concessiva, a fim de marcar a flutuação existente entre esses usos e os concessivos sem deixar, no entanto, de sinalizar a diferença entre eles: se a oração em realce preceder, tornando-se temática, no complexo oracional, a relação é hipotática, e, portanto, concessiva, se a oração em realce proceder, como é o caso de 2Q introduzida por *mesmo/ainda assim*, a relação é paratática. Dessa forma, estou considerando toda a fluidez semântico-cognitiva que caracteriza esses usos, conforme exposição realizada nesta e na seção anterior, como indicativa do processo de mudança por que passa a locução.

<sup>14</sup> Prosodicamente, constata-se uma frase entoacional (I) em Q, marcada por contorno final descendente e por pausa que a separa da próxima I *porque lá a gente num caminha pelas calçadas*. Portanto, *assim mesmo* realiza-se no interior da mesma I, sem pausa separando os itens que constituem a locução.

Note que aqui, como preconiza Halliday, embora haja a independência característica das orações paratáticas, a sequência/ordenação não pode ser alterada.

Por fim, a Tabela 1 traz a frequência de usos prototípicos (P) e não-prototípicos (N-P) da locução analisada de acordo com cada TD:

**Tabela 1: Frequência de usos prototípicos e não-prototípicos de P *mesmo/ainda assim* Q**

	P	%	N-P	%
TD carta	01	(20%)	04	(80%)
TD editorial	06	(60%)	04	(40%)
TD <i>e-mail</i>	0	(0%)	01	(100%)
TDs do Iboruna	TD narrativa de experiência	02	05	(71,43%)
	TD narrativa recontada	0	0	(0%)
	TD opinativa	02	03	(60%)
	TD injuntiva	01	01	(50%)
	TD descritiva	01	01	(50%)

Como mostra a Tabela 1, P *mesmo/ainda assim* Q é mais recorrente, diacronicamente, na TD editorial (10 ~ 5,34% frente a 05 ~ 3,93% na TD carta). Na perspectiva sincrônica, o maior número de casos é constatado na TD narrativa de experiência (7 ~ 2,38%). Vale destacar que, na diacronia, a TD carta apresenta 80% das ocorrências em contextos não-prototípicos e favorecedores da mudança em comparação com 40% de suas ocorrências não-prototípicas nos dados da TD editorial, nessa mesma perspectiva. Sincronicamente é também elevada a frequência de seus usos não-prototípicos.

### Considerações finais

A análise dos dados aqui realizada mostra que o processo de mudança, refletido neles, sofre influência dos contextos linguístico-pragmáticos que o subjazem. Estes, por sua vez, estão associados diretamente às características de TDs específicas.

De modo geral, na TD carta, as análises mostraram vários exemplos em que, embora em contextos relevantes para a emergência da aceção e do funcionamento do junctor contrastivo, foram constatadas situações de não reanálise ou de ambiguidade em relação às características dessa construção. Por sua vez, nos editoriais, embora a reanálise de *ainda/mesmo* e *assim* estivesse concluída (sendo a locução usada, mais frequentemente, de forma próxima ao seu funcionamento prototípico), os contextos de sua utilização continuaram permeados por aspectos concessivos, adversativos e/ou condicionais, por vezes até de forma contígua, o que, além de indicar o próprio desenvolvimento da locução, justifica sua acentuada fluidez entre concessão e adversidade. Por sua vez, os dados de *e-mail* e das TDs do Iboruna mostraram contextos relevantes para a identificação de ambiguidades, principalmente, nestas últimas, em que se constataram relações de contiguidade sintagmática na possível implementação da mudança refletida sincronicamente, associada, muitas vezes, a fatores que extrapolam a análise sintática, semântica e pragmática, exigindo uma análise de cunho prosódico.

A partir desses resultados, é possível afirmar, no que tange ao processo de GR da locução *mesmo/ainda assim* em correlação às TDs focalizadas, que:

- (i) com base na análise dos usos não-prototípicos da locução, há relações entre as TDs carta, *e-mail* e TDs do IBORUNA, apesar das distintas bases semióticas que as subjazem, indiciando, portanto, aspectos sintomáticos da heterogeneidade constitutiva da escrita, nos moldes de Corrêa (1997),<sup>15</sup>
- (ii) existe um processo de mudança que ainda se revela em trânsito, mesmo na TD editorial, que favorece os usos prototípicos da locução, já que 40% de suas ocorrências equivalem a usos não-prototípicos, sendo que mesmo os prototípicos se encontram em ambientes ainda fortemente marcados por traços semântico-cognitivos e formais característicos de contextos que favorecem a emergência da locução.

Ou seja, os usos da locução aqui investigada, desde o século XVIII até a sincronia atual, comprovam que, sincronicamente, sentidos adjacentes são também diacronicamente adjacentes (TRAUGOTT, 1982). Mostram, assim, que os sentidos que caracterizam a fonte de derivação, na diacronia, refletem-se como fonte de polissemia e ambiguidade, na perspectiva sincrônica. Entretanto, nessa relação entre as perspectivas, no que tange ao processo de GR, ganha destaque aqui a implementação da mudança em cada uma das TDs focalizadas. Embora revelem usos da locução em estágios distintos do processo (TDs carta, *e-mail* e TDs do IBORUNA, com usos menos prototípicos e TD editorial, com os mais prototípicos), apresentam, da mesma forma, fortes indícios de um processo de mudança que ainda se revela em curso.

## REFERÊNCIAS

- CHEN, G. The grammaticalization of concessive markers in Early Modern English. In: FISCHER, O. (Org.) *Pathways of change, grammaticalization in English*. Philadelphia, PA, USA: John Benjamins Publishing Company, 2000. p. 85-108.
- CORRÊA, M. L. G. *O modo heterogêneo de constituição da escrita*. Campinas. SP. 435f. 1997. Tese de doutorado – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas.
- COSERIU, E. *Sincronia, diacronia e história: o problema da mudança linguística*. Rio de Janeiro: Presença, 1979.
- GIVÓN, T. *On understanding grammar*. London: Academic Press, 1979.
- GUIMARÃES, E. R. J. *Texto e argumentação*. Campinas: Pontes, 1987.
- HALLIDAY, M. A. K. Above the clause: the clause complex. In: \_\_\_\_\_. *An introduction to Functional Grammar*. New York: Arnold, 1985.
- HALLIDAY, M. A. K.; HASAN, R. *Cohesion in English*. London, Longman, 1976.

---

<sup>15</sup> Cf. a relação entre fala/oralidade e escrita/literalidade assumida aqui (CORRÊA, 1997), adoto não a compartimentação de gêneros em um contínuo, mas um *modo heterogêneo de constituição da escrita*, fundamentado na apreensão de TDs escritas transpassadas por traços de oralidade e vice-versa.

HARRIS, M. Concessive clauses in English and Romance. In: HAIMAN, J.; THOMPSON, S. *Clause combining in grammar and discourse*. Amsterdam: John Benjamins, 1988. p. 71-99.

HEINE, B.; CLAUDI, U.; HÜNNEMEYER, B. *Grammaticalization: a conceptual framework*. Chicago: The University of Chicago Press, 1991.

HOPPER, P.; TRAUGOTT, E. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.

KABATEK, J. Las tradiciones discursivas del español medieval: historia de textos e historia de la lengua: *Iberoamericana*, n. 62, p. 28-43, 2005.

KOCH, I. G. V. *A inter-ação pela linguagem*. São Paulo: Contexto, 2001.

KÖNIG, E. On the history of concessive connectives in English. Diachronic and synchronic evidence. *Lingua*, v. 66, p. 1-19, 1985.

LOPES-DAMASIO, L. *Diacronia dos processos constitutivos do texto relativos a assim: um novo enfoque da gramaticalização*. São Paulo: Cultura Acadêmica, 2011.

\_\_\_\_\_. Tradição discursiva e mudança linguística: uma abordagem da gramaticalização. *Estudos linguísticos*, São Paulo, v. 41, n. 3, 2012.

MARTELOTTA, M. Gramaticalização e graus de vinculação sintática em cláusulas concessivas e adversativas. *Veredas: revista de Estudos Linguísticos*, Juiz de Fora, v. 2, n. 3, p. 37-56, 1998.

MEILLET, A. L' évolution des formes grammaticales. *Scientia* 12, v. 26, n. 6. Reimpresso em A. MEILLET *Linguistique historique et linguistique générale*. Paris: Librairie Honoré Champion, 1965 [1912].

NEVES, M. H. M. As construções concessivas. In: \_\_\_\_\_. (Org.) *Gramática do português falado*. São Paulo: Humanitas FFLCH/USP, 1999. Vol. II: Novos estudos, p. 545-591.

QUIRK, R. et al. *A comprehensive Grammar of the English Language*. London-New York: Longman, 1985.

SWEETSER, E. *From etymology to pragmatics*. Cambridge: Cambridge University Press, 1990.

TRAUGOTT, E. Constructions in Grammaticalization. In: JOSEPH, B.; JANDA, R. D. (Org.). *The handbook of historical linguistics*. Oxford: Blackwells, 2003.

\_\_\_\_\_. *The role of the development of discourse markers in a theory of grammaticalization*. Paper presented at ICHL XII, Manchester, 1995. Disponível em: <<http://www.stanford.edu/~traugott/traugott.html>>. Acesso em: 20 fev. 2004.

\_\_\_\_\_. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In: LEHMMAN, W.; MALKIEL, Y. (Ed.) *Perspectives on historical linguistics*. Amsterdam: John Benjamins, 1982. p. 245-271.

TRAUGOTT, E.; KÖNIG, E. The semantic-pragmatics of grammaticalization revisited. In: TRAUGOTT, E.; HEINE, B. (Ed.) *Approaches to grammaticalization*. Vol. 1. Amsterdam: John Benjamins Publishing Company, 1991.

# As fricativas sibilantes nas *Cantigas de Santa Maria*

(The sibilant fricatives in the *Cantigas de Santa Maria*)

Mariana Moretto Gementi<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista “Júlio de Mesquita” (Unesp)

mariana\_moretto@hotmail.com

**Abstract:** The aim of this study is mainly to map and analyze the sibilant fricatives in coda position of the *Cantigas de Santa Maria*, Afonso X, the Wise King from the thirteenth- XIII century. Words containing consonants spelled with <s>, <z> and <x> in syllable-final position will be considered as object of study. From a still unexplored corpus, we intend to look into some aspects of the evolution of Portuguese language that refer to the sibilants. Thus, the importance of the description of the relationship between letters and sounds will be shown regarding the possible medieval lyric spellings and regarding the position of the segment in the syllable.

**Keywords:** sibilant fricatives; coda; archaic Portuguese; *Cantigas de Santa Maria*.

**Resumo:** O presente estudo tem como objetivos principais o mapeamento e a análise das fricativas sibilantes na posição de coda nas *Cantigas de Santa Maria*, compiladas por Afonso X, o rei Sábio – século XIII. Serão consideradas como objeto de estudo as palavras contendo consoantes grafadas com <s>, <z> e <x>, em posição de travamento silábico. Por meio deste trabalho e a partir de um *corpus* ainda pouco explorado, pretendemos observar alguns aspectos de evolução da língua portuguesa ao que se refere às sibilantes. Dessa forma, mostraremos a relevância da descrição da relação entre letras e sons com relação às grafias possíveis da lírica medieval, no que diz respeito à posição silábica.

**Palavras-chave:** fricativas sibilantes; coda; português arcaico; *Cantigas de Santa Maria*.

## Introdução

O objetivo principal deste estudo é analisar as fricativas sibilantes nas *Cantigas de Santa Maria* (CSM), que são cantigas religiosas medievais galego-portuguesas, mandadas compilar pelo Rei Afonso X de Castela, no último quartel do século XIII. Além disso, pretende-se observar seu comportamento em relação à posição que ocupam na sílaba, observando se existem as oposições apontadas pelos estudiosos entre fricativas surdas e sonoras e entre dentais e alveolares, nas posições mediais e finais da sílaba, ou se essas oposições são condicionadas pela posição da consoante na sílaba. O estudo investiga se, naquele momento, os processos de neutralização das fricativas existiam ou não no português, para, então, estabelecer se esses grafemas representavam sons de caráter distintivo ou não no contexto de início e de travamento silábico.

O fato de escolher as fricativas sibilantes dá-se, primeiramente, pela grande produtividade no *corpus* das *Cantigas de Santa Maria* e, além disso, pelo fato de haver controvérsias quanto à consideração da oposição entre fricativas e entre os autores que vêm estudando o assunto.

A metodologia utilizada neste trabalho baseia-se na possibilidade (ou não) de variação gráfica na representação das consoantes e na possibilidade (ou não) de rima entre essas palavras específicas para determinar sua possível realização fonética naquela época.

Para estabelecer se há ou não oposição entre os sons representados pelos grafemas focalizados, os dados serão analisados a partir do modelo estruturalista de Cagliari (2002).

### ***Corpus: Cantigas de Santa Maria***

O *corpus* escolhido para a realização desta pesquisa foram as *Cantigas de Santa Maria*, por ser a maior coletânea medieval de poesias em louvor à Virgem Maria. O Rei de Castela, Afonso X, as compôs para cantar os feitos da Virgem, e ele próprio as mandou compilar no princípio dos anos sessenta do século XIII. A obra mariana é um monumento literário, musical e artístico de mais elevada importância (cf. PARKINSON, 1998, p. 179), além de ser o mais rico cancionário em louvor à Virgem Maria da Idade Média (cf. METTMANN, 1986a, p. 7; BERTOLUCCI PIZZORUSSO, 1993, p. 142).

O conjunto das cantigas afonsinas é composto por 420 poemas marianos musicados, divididos entre cantigas narrativas ou de milagres, as quais contam os feitos milagrosos da Virgem Santa Maria, a respeito de ajuda com enfermidades, socorro e perigos, ou também na ajuda às decisões do Rei D. Afonso X, e cantigas líricas, de louvor à Virgem Maria como auxiliadora, mediadora e interventora (BERTOLUCCI PIZZORUSSO, 1993, p. 143).

Mettmann (1986a, p.7) afirma que, das 420 *Cantigas de Santa Maria* (descontadas sete repetições), 356 são narrativas e relatam os milagres da Virgem, e as demais, excetuando-se uma introdução e dois prólogos, são cantigas de *loor* (louvor) ou referem-se a festividades do calendário cristão, relativas a episódios da vida de Santa Maria ou de seu Filho Jesus Cristo.

No que diz respeito à proporção das cantigas de milagre e das cantigas de louvor na coletânea, Leão (2007, p. 24) afirma que há predominância das primeiras sobre as segundas, em uma relação de nove para um. Assim, a cada grupo de nove cantigas de milagre segue uma cantiga de louvor, numerada com uma dezena inteira.

A estruturação das cantigas obedece, pois, a um ritmo regular, em que as cantigas de louvor ocupam sempre as dezenas, enquanto as de milagre têm números terminados pelas unidades de um a nove, comparando-se esse sistema, aproximadamente, ao de um rosário. (LEÃO, 2007, p. 24)

Nas cantigas de milagres, D. Afonso agradece às curas de suas enfermidades, e também pode ser uma testemunha do fato miraculoso, ou ainda, um conhecedor que dele teve notícia por leitura ou por ouvir dizer, como podemos observar abaixo, no resumo da cantiga 7, feito por Ângela Vaz Leão (2007, p. 26):

Uma abadessa, que não era benquista pelas monjas suas subordinadas, engravida-se de um homem de Bolonha. As monjas, alegres por isso, vão acusá-la ao bispo de Colônia (diocese a que pertencia o convento), que vem, em comitiva, apurar o fato *in loco*. Reunida a comunidade diante o bispo, que pede explicações à monja, esta permanece silenciosa no interrogatório. Depois, retira-se e vai orar à Virgem diante do altar, chorando. Aí começa o milagre: ela se deita e adormece profundamente. Durante o sono, é salva pela Virgem Maria, num parto miraculoso, operado por dois anjos do séquito mariano, que, como se vê pelas iluminuras (mas não pelo texto), depois levam o recém-nascido para ser criado fora dali, por um ermitão, certamente devoto da Virgem. A abadessa acorda e, imune assim às acusações das monjas, vai se apresentar ao bispo. Este, após determinar que ela

se dispa, examina-lhe os seios desnudos e a isenta de culpa, dizendo que nela nada pode encontrar para acusá-la.

Além de ser um exemplo de cantiga de milagre, Ângela Vaz Leão explica que a partir deste milagre se pode avaliar a importância das *Cantigas de Santa Maria* como testemunho social da época, pois explica a vida num convento e, a partir disso, podemos pensar que as cantigas de milagre são uma valiosa fonte histórica para o conhecimento do viver e do morrer, das doenças e das calamidades, do jogo e da prostituição, dos ofícios e dos lazeres, das crenças e das religiões, da vida quotidiana e do imaginário popular, enfim de toda a cultura ibérica, na Idade Média (LEÃO, 2007, p. 27).

As cantigas de louvor constituem a parte essencialmente lírica da coletânea, mostram sempre o Rei-trovador diante da Virgem Maria, exaltando-lhe as qualidades, louvando sua beleza ou oferecendo-lhe a sua devoção. Ângela Vaz Leão (2007, p. 29) afirma que Dom Afonso X foi um apaixonado “trovador da Virgem”, pois, nas cantigas de louvor, o trovador da Virgem Maria tem um comportamento masculino semelhante ao das *cantigas d’amor*, ao mesclar os ideais do amor cortês com os do Cristianismo, e a mulher amada se sublima em Santa Maria.

As *Cantigas de Santa Maria* foram escritas em galego-português e contêm notação musical, acompanhadas por pautas musicais que eram cantadas. A autoria é atribuída a Dom Afonso X, o Rei Sábio, porém há autores que contestam que o rei tenha a autoria total da obra. Além da notação musical, as cantigas contêm, também, *iluminuras* - desenhos miniaturizados que representam o conteúdo que está sendo narrado na respectiva cantiga.

Parkinson (1998, p. 179) afirma que as CSM constituem um monumento literário, musical e artístico da mais elevada importância. Para Bertolucci Pizzorusso (1993, p. 144), a coletânea das CSM é uma obra para ser vista e ouvida, na qual “uma milagristica por imagens junta-se à milagristica em versos”. Há um perfeito equilíbrio entre texto, melodias e pintura e, assim, Mettmann (1986b, p. 8) afirma que as CSM ocupam um lugar privilegiado na literatura medieval e revelam que, para seu principal idealizador, o Rei Afonso X, a música e a pintura não eram menos importantes do que o “contar”, o “trovar” e o “rimar”. Sobre esse assunto, Leão (2007, p. 30-31) declara: “Conforme se reconhece hoje, os textos, as iluminuras e as notações musicais, em conjunto, fazem das *Cantigas de Santa Maria* uma das mais ricas de toda a Idade Média - o que justifica que tenha sido caracterizada por Menéndez y Pelayo como ‘a Bíblia estética do século XIII’”.

A escolha de textos poéticos como *corpus* da pesquisa deve-se ao fato de que, por meio da análise das rimas, é possível obter pistas satisfatórias sobre a realização fônica de vogais e consoantes em momentos passados da língua, dos quais não se têm registros orais.

Nesta pesquisa, optou-se por trabalhar com as cantigas religiosas (e não com as profanas), porque estudos revelam que as CSM, em termos de léxico e de rima, são mais ricas do que as cantigas profanas (cf. LEÃO, 2007, p. 152-153).

## Forma de análise dos resultados

Como forma de ilustração, apresenta-se abaixo um exemplo dos procedimentos de mapeamento dos dados utilizados nesta pesquisa. O exemplo citado (1) é um fragmento

da cantiga de número onze (CSM11). Apresentamos os versos de 3 a 17, de um total de 98 versos da cantiga. As palavras que possuem fricativas sibilantes serão grafadas em negrito:

- (1) ESTA É DE COMO SANTA MARIA TOLLEU A ALAMA DO MONGE  
QUE SS'AFFOGARA NO RIO AO DEMO, E FEZE-O RESSOCITAR.

Macar ome per folia	3
aginna caer	4
pod'en pecado,	5
do bem de <b>Santa</b> Maria	6
non dev'a <b>seer</b>	7
<b>desasperado.</b>	8
Poren direi todavia	9
com'en hũa abadia	10
un <b>tesoureiro</b> avia,	11
monge que trager	12
con mal recado	13
a <b>ssa fazenda sabia,</b>	14
por a <b>Deus</b> perder,	15
o malfadado.	16
Macarome per folia...	17
[...]	

Depois de mapeadas as ocorrências das fricativas sibilantes, são montados quadros, nos quais os grafemas <s>, <z>, <x>, <c>, <ç>, <sc>, <ss> são divididos entre *onset* (início e meio de palavra) e *coda* (meio e fim de palavra), assim, podemos observar e quantificar todas as ocorrências. Os resultados das ocorrências são analisados tanto quantitativa, quanto qualitativamente.

Em seguida, dentro desta perspectiva, a análise dos resultados é feita no interior do quadro teórico inaugurado pelas teorias fonológicas não lineares, em especial os modelos de Geometria de Traços (CLEMENTS; HUME, 1995; para o Português Brasileiro, CAGLIARI, 1998).

## Resultados

A coleta de dados no *corpus* deste estudo possibilitou o mapeamento das fricativas sibilantes (grafadas por <s>, <z> e <x>, <c>, <ç>, <sc>, <ss>) nas 50 primeiras CSM. Apresentaremos a seguir os resultados obtidos.

Conforme podemos observar na Tabela 1, foi coletado um total de 7.151 de ocorrências de consoantes fricativas:

**Tabela 1: Quantificação das ocorrências nas consoantes fricativas mapeadas no *corpus***

Consoantes sibilantes	Quantidades (percentual)
<i>Onset</i>	3.524 (49,3%)
Coda	3.627 (50,7%)
<b>Total</b>	7.151 (100%)

Desse total, dividimos os grafemas e quantificamos na posição de *onset* e coda, como podemos observar nas tabelas 2 e 3, abaixo:

**Tabela 2: Consoantes sibilantes em posição de *onset***

Grafema/ posição do som na sílaba	Posição Inicial de palavra	Posição Medial de palavra	Subtotal
<s>	1637 (79,3%)	427 (20,7%)	2.064 (100%)
<ss>	182 (26,4%)	508 (73,6%)	690 (100%)
<sc>	0	0	0
<ç>	1 (0,40%)	240 (99,6%)	241 (100%)
<sç>	0	0	0
<z>	0	325 (100%)	325 (100%)
<x>	14 (20,6%)	54 (79,4%)	68 (100%)
<c>	83 (33,4%)	165 (66,6%)	248 (100%)
<b>Subtotal</b>	1.917	1.719	<b>3.636</b>

**Tabela 3: Consoantes sibilantes em posição de coda**

Grafema/ posição do som na sílaba	Posição Medial de palavra	Posição Final de palavra	Subtotal
<s>	766 (23,8%)	2.455 (76,2%)	3.221 (100%)
<ss>	0	0	0
<sc>	0	0	0
<ç>	0	0	0
<sç>	0	0	0
<z>	7 (2,4%)	290 (97,6%)	297 (100%)
<x>	0	2 (100%)	2 (100%)
<c>	0	0	0
<b>Subtotal</b>	773	2.747	<b>3.520</b>

### Fricativas sibilantes nas *Cantigas de Santa Maria*

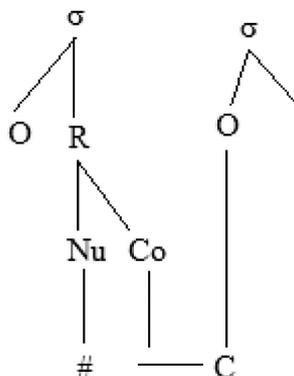
Abaixo iremos discutir os dados obtidos na coleta feita na posição de coda nas cinquenta primeiras CSM.

Em primeiro lugar, foi feito um mapeamento das ocorrências das consoantes fricativas sibilantes do *corpus*, levando em consideração sua posição na sílaba (na coda); em seguida, tendo como objetivo apresentar o sistema das consoantes fricativas empregado pelos trovadores que compuseram as cantigas religiosas em galego-português, foram analisadas todas as ocorrências de fricativas sibilantes na posição de coda encontradas no *corpus*.

### Sibilantes em coda em posição medial de palavra

Abaixo podemos observar as ocorrências de sibilantes em coda em posição medial de palavras mapeadas (2) no *corpus*, que correspondem ao contexto representado:<sup>1</sup>

(2)



<s> - Contexto: entre vogal e consoante: Conpostela [CSM A, v.3]; batismo [CSM 4, v.100]; Cristo [CSM 13, v.3].

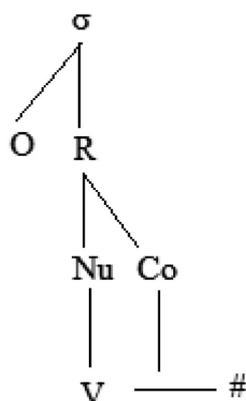
- Contexto: entre vogal e consoante, em contexto de elisão: [...] que sterreces [CSM 20, v.39]; [...] a 'scodudas [CSM 31, v.68]; [...] que speciais [CSM 34, v.27]; [...] toda 'sclareceu [CSM 15, v.91].

<z>-Contexto: entre vogal e consoante: fazfeiro [CSM 45, v.26]; dezpraz [CSM 47, v.15]; azcũa [CSM 22, v.25]; vezỹos [CSM 45, v.12].

### Sibilantes em coda em posição final de palavra

Abaixo podemos observar as ocorrências de sibilantes (3) em coda em posição final de palavras mapeadas no *corpus*, que correspondem ao contexto representado:

(3)



<sup>1</sup> A sílaba é definida como uma estrutura hierarquicamente organizada em constituintes silábicas, que se ramifica em *onset* (O) e *rima* (R). A *rima* se ramifica em *núcleo* (Nu) e *coda* (Co), sendo que o *núcleo*, no português, aloja as vogais e constitui o pico silábico e a *coda*, as consoantes da *rima*. O *onset* é o elemento que precede o núcleo de uma sílaba e é geralmente formado por uma ou mais consoantes, e a *coda*, é a consoante ou as consoantes em posição pós-nuclear dentro de uma sílaba, ou seja, após a vogal nuclear. A *coda*, juntamente com o *núcleo*, se denomina *rima*, que não é totalmente necessária em uma sílaba.

<s> - Contexto: depois de vogal (são maioria das palavras), como cantares [CSM B, v.41]; miragres [CSM A, v.24]; coitas [CSM 5, v.25].

<z> - Contexto: depois de vogal no final de palavra como em: faz [CSM 3, v.3]; cruz [CSM 5, v.174]; fiz [CSM 28, v.122].

<x> - Contexto: depois de vogal no final de palavra. Foram encontradas apenas duas ocorrências: aprix [CSM 25, v.16]; fix [CSM 47, v.13].

### **Análise dos dados na posição de coda no meio e no final da palavra**

Na posição de coda no meio da palavra, foram encontrados apenas os grafemas <s> e <z>, os quais estão no contexto entre vogal e consoante, como nas palavras: *Conpostela* [CSM A, v.3]; *batismo* [CSM 4, v.100]; *Cristo* [CSM 13, v.3].

M. Said Ali (1905, p.8) explica que o *s* antes de outra consoante terá valor fonético ora surdo ora sonoro, conforme for surda ou sonora a consoante imediata como nas palavras encontradas no *corpus*: *batismo* (CSM 4, v.100), *Bispo* (CSM 7, v.35), *Cristo* (CSM 13, v.3), *Conpostela* (CSM 26, v.64). No caso de sufixo de plural e de desinência pessoal de 2º pessoa escreve-se *s*, e não *z*, como nas palavras encontradas nas *CSM cantares* (CSM 8, v.28), *saberdes* (CSM 5, v.174), *comerás* (CSM 15, v.54). O fato de essa escrita vir com <s> e não com <z> mostra que a pronúncia [s] era típica nesses contextos.

Outra questão interessante encontrada na análise do *corpus* diz respeito ao grafema <s>, do tipo <s + oclusiva> no início de palavra, como nos exemplos *sterreces* [CSM20, v.39], *'scodudas* [CSM 31, v.68], *speciais* [CSM 34, v.27], *'sclareceu* [CSM 15, v.91]. Nesses três exemplos, assim como afirma Massini-Cagliari (2005, p. 98), a forma métrica do verso está correta, não há a necessidade da inserção da vogal inicial para acertar a contagem da sílaba poética. No entanto, na palavra *[e]splandores* (CSM 384, v.18), o editor (METTMANN, 1989, p. 282) achou melhor acrescentar a vogal “e” para que o verso tivesse o número de sílaba exigido pela métrica do poema. Portanto, assim como Massini-Cagliari (2005, p. 98) concluiu, a vogal não era pronunciada nas palavras transcritas.

Uma outra questão levantada pela autora e observada em nossas análises foi qual seria a estrutura inicial das palavras *sterreces*, *'scodudas*, *speciais*, *'sclareceu*. Estaria na posição de *onset* complexo ou coda? Sobre essa questão Massini-Cagliari (2005) explica:

Em todos os exemplos mapeados, as seqüências de S+C(C) sempre ocorrem depois de palavra terminada em vogal, a não ser no caso do exemplo *[e]splandores*, em que o editor postula a necessidade da vogal inicial. Por este motivo, é mais provável a hipótese de estar ocorrendo um processo de sândi, que apaga a vogal inicial de *estrela* (ou impede a sua inserção, no caso de modelos fonológicos que consideram essa vogal epentética), e liga o “S desgarrado” à coda da sílaba anterior. (p. 98)

Os exemplos *'scodudas* e *'sclareceu* mostram casos em que processos de sândi (crase, no primeiro caso, e elisão, no segundo) ocasionam o “desgarramento” do S da sílaba a que originariamente pertenceria, dado o apagamento do núcleo, e a sua adjunção ao núcleo da sílaba imediatamente anterior.

Portanto, analisando esses casos, podemos concluir que o PA não possui ataques silábicos supercomplexos, uma vez que em nenhum dos casos pode-se afirmar que o “S desgarrado” se realiza no *onset* da sílaba (MASSINI-CAGLIARI, 2005, p. 98).

Os grafemas <s>, <x> e <z> aparecem em posição de coda simples e todos com a representação sonora de [s], como em: *miragres, fix, fez, fiz*. Isso ocorre por haver neutralização entre os fonemas /s/ e /z/ na posição de coda. Por isso, pode-se representá-los através do arquifonema /S/, assim como fez Câmara Jr. (1995[1970]) para o PB (PINHEIRO, 2004, p. 70).

Em Massini-Cagliari (2005), a autora mostrou que em relação ao preenchimento silábico, dada a maior presença de palavras estrangeiras associadas às histórias milagrosas, há maior ocorrência de codas irregulares constituídas de oclusivas nas cantigas religiosas do que nas profanas.

Apesar de no PA predominarem as sílabas abertas, como afirmou Biagioni (2002, p. 87-88), o travamento silábico é permitido. No entanto, as possibilidades são restritas quando se referem às consoantes que podem ocupar a posição de coda no PA.

Consoante na posição de coda simples:

**Quadro 1. Consoantes na posição de coda simples (BIAGIONI, 2002, p. 83)**

Consoante	Grafemas correspondentes	Exemplos
/S/	<s>, <x>, <z>	Varões, diz, quis, Deus, jaz, fez, paz, luz, emperadriz, aprix, fix
/r/	<r>	Lazerar
/l/	<l>	Tal
/N/	<m> e <n>	bem, razon

Biagioni (2002, p. 83) afirma que /r/, /l/, /S/ e /N/ são consideradas consoantes, formando codas simples, isso quer dizer que o PA proíbe codas complexas. Além disso, a autora diz que os segmentos /S/ e /N/ são considerados arquifonemas por duas razões: o primeiro porque possui mais de uma realização fonética e o segundo por haver neutralização da oposição, em contexto de travamento silábico, oposição que havia no contexto de início de sílabas entre os sons /m/ e /n/. Em exemplos como *bem* e *razon*, admite-se que os grafemas <m> e <n> não eram pronunciados como consoantes [m] e [n], mas indicavam apenas a presença de uma vogal nasalizada.

Na posição final da palavra, podemos identificar, ao analisar nossa coleta de dados do *corpus*, buscando todas as palavras que possuem consoantes fricativas na posição de coda silábica nas CSM, rimas silábicas grafadas com <s>, rimas grafadas com <x> e rimas grafadas com <z>. Pode-se perceber que não há oposição entre os sons representados por esses grafemas, pois não há pares mínimos que indiquem oposição. Portanto, na posição de coda, essa oposição desaparece; por exemplo, na vigésima terceira estrofe da CSM 57, aparecem os seguintes versos: “E pois que os ladrões / ant’o altar trouxeron, / por eles orações / e pregairas fezeron”. Nesse exemplo, encontram-se na posição focalizada, finalizadas por grafemas representativos de consoantes fricativas, as palavras *ladrões* e *orações*.

No livro *Análise Fonológica* (CAGLIARI, 2002, p. 46-48), o autor diz que, em relação à distribuição dos fonemas, podemos notar que dois sons foneticamente semelhantes ocorrem

em oposição fonológica em certos contextos, mas não estão em oposição fonológica em outros contextos. Ou seja, a oposição fonológica que ocorre num contexto se neutraliza em outro contexto.

Algumas razões para a neutralização de uma oposição fonológica são: a não ocorrência de um dos membros do par de fonemas; a ocorrência complementar deles, caso em que um fonema ocorre num contexto e o outro em outro tipo de contexto; ocorrência de variação livre, envolvendo os sons em questão. (CAGLIARI, 2002, p. 46)

Segundo M. Said Ali (1905, p. 8), a ortografia da sibilante final do português atual é determinada pela pronúncia da vogal que a precede, como no caso de a vogal ser precedente da tônica, escreve-se com *z*: *paz, nariz, paiz, timidez, Pariz, cruz*. Nessa regra, há as exceções dos pronomes (*nós, vós*) e os vocábulos *tres, Moysés, aliás, bis, cris, cós, Jesus, jus, pus (substantivo), mas, cis*. Outra hipótese é que se a vogal precedente for átona, escrevemos com *s*: *Marques, Venus, onus, Tunis, lapis, iris, Deus, mais, dois, pois* (SAID ALI, 1905, p. 8).

Segundo Pinheiro (2004, p.70), os grafemas <s>, <z> e <x> aparecem em posição de coda simples, todos com a representação sonora de [s] (valendo /S/), por exemplo: *ondas, fez e fax*. Justamente por haver neutralização entre os fonemas /s/ e /z/ na posição de coda, não há certeza quanto à atualização fonética dos grafemas <z>, <s> e <x> nessa posição. Pode-se, por esse motivo, representá-los através do arquifonema /S/.

A partir da coleta de dados do *corpus* das CSM presenciamos a mesma palavra, com duas ou mais grafias diferentes, *fis* [CSM 5, v.181], *fiz* [CSM 28, v.122], *fix* [CSM 47, v.13], entretanto, quanto à pronúncia das mesmas, sabemos que são variantes [s, z, ʃ], mas é uma hipótese sem provas. O mais provável era uma neutralização fonológica nessa posição, refletida na escrita, de um modo geral. Nos dados, não encontramos nenhuma palavra que tenha <x> em posição de coda no meio da palavra, como “Lixboa”. Por isso, é muito provável que naquela época <x> tinha o som de [s]. Em Edwin Williams (1973, p.102) o autor diz: “O [k] de x final se tornou iode e o s permaneceu: *sex = seis*”. Ele reconhece que em final, o *x* do latim podia ser pronunciado [ks].

Na posição final de palavra, encontramos casos de palavras como *mays* [CSM 5, v.185]; *poys* [CSM 2, v.60]; *depoys* [CSM 3, v.7]; *Seixons* [CSM 41, v.6], <s> em posição de coda em final de sílaba, depois de consoante (semivogal <y> ou nasal <n>). Nesses casos, ocorre a ramificação da coda, ou seja, o primeiro elemento será ou uma nasal ou um glide e o segundo elemento sempre será uma fricativa (*de.poys, sei.xons*). É interessante notar que os elementos /N,L,R,S/ representam grandes classes de modo de articulação de uma consoante: nasal, lateral, vibrante, fricativa. Há líquidas: /L, R/; sonorantes (vozeamento intrínseco): /N, L, R/; e um segmento tipicamente surdo /S/ (CAGLIARI, 1997, p. 35).

## Conclusão

Os grafemas <s>, <x> e <z> em posição de travamento silábico correspondem provavelmente a um arquifonema fricativo, ou seja, a um som especificado apenas com o traço fricativo, sem especificação necessária quanto ao traço de vozeamento, porque, apesar de haver a certeza de que se trata de um segmento fricativo, não é possível saber com certeza qual a sua realização fonética exata, nesse contexto, dada a ausência de

variação dos grafemas focalizados nos mesmos dados ou de evidências de oposição. O arquifonema representa a neutralização de uma oposição fonológica estabelecida em outros contextos, ou seja, a oposição fonológica que ocorre num contexto se neutraliza em outro contexto.

Enfim,

[...] cada contexto tem sua estrutura e o que acontece num caso não precisa acontecer do mesmo modo em outros. Por exemplo, há oposição fonológica no Português entre [s] e [z], quando ocorrem entre vogais, como em *caça* e *casa*. Porém, em final de palavras, diante de pausa, só ocorre o [s] e nunca o [z], como se pode observar em palavras como *fiz*, *paz*, *vês*, *avôs*, *todos*, *eles*, etc. (CAGLIARI, 2002, p. 25)

## REFERÊNCIAS

BERTOLUCCI PIZZORUSSO, V. Cantigas de Santa Maria. In: LANCIANI, G.; TAVANI, G. (Org.). *Dicionário da literatura medieval galega e portuguesa*. Lisboa: Caminho, 1993. p. 36-41.

BIAGIONI, A. B. *A sílaba em português arcaico*. Dissertação (Mestrado em Linguística) – Faculdade de Ciências e Letras, Unesp, Araraquara, 2002.

CAGLIARI, L. C. *Análise fonológica: Introdução à teoria e à prática com especial destaque para o modelo fonêmico*. Campinas: Mercado de Letras, 2002.

\_\_\_\_\_. A escrita do português arcaico e a falsa noção de ortografia fonética. In: EARLE, T. F. (Org.). *Actas do V Congresso da Associação Internacional de Lusitanistas*. Coimbra: AIL, 1998. p. 57-69.

\_\_\_\_\_. *Fonologia do Português: análise pela Geometria de Traços*. Campinas: Edição do Autor, 1997.

CÂMARA JR, J. M. *Estrutura da língua portuguesa*. 21. ed. Petrópolis: Vozes, 1995. [1. ed., 1970].

CLEMENTS, G. N.; HUME, E. V. The internal organization of speech sounds. In: GOLDSMITH, J. A. *The handbook of Phonological Theory*. Cambridge MA, Oxford UK: Blackwell, 1995. p. 245-306.

LEÃO, Â. *Cantigas de Santa Maria, de Afonso X, o Sábio: aspectos culturais e literários*. Belo Horizonte: Veredas & Cenários, 2007.

MASSINI-CAGLIARI, G. *A música da fala dos trovadores: estudos de prosódia do português arcaico, a partir das cantigas profanas e religiosas*. 2005. Tese (Livre docência em Linguística) – Faculdade de Ciências e Letras, Unesp, Araraquara.

METTMANN, W. (Ed.). *Cantigas de Santa Maria (Cantigas 261 a 427): Alfonso X, el Sabio*. Madrid: Castalia, 1989.

\_\_\_\_\_. Affonso X e elSabio. *Cantigas de Santa Maria* (Cantiga 1 a 100). Madrid: Castalia, 1986a.

\_\_\_\_\_. Introducción. In: AFFONSO X, EL SABIO. *Cantigas de Santa Maria* (Cantiga 1 a 100). Madrid: Castalia, 1986b. p. 7-42.

PARKINSON, S. Layout and Structure of the Toledo Manuscript of the *Cantigas de Santa Maria*. In: PARKINSON, S. (Ed.). *Cobras e Son: Papers on the Text Music and Manuscripts of the "Cantigas de Santa Maria"*. Oxford: Legenda, University of Oxford, 2000. p. 133-153.

\_\_\_\_\_. As Cantigas de Santa Maria: estado das cuestións textuais. *Anuário de Estudos Literarios Galegos*, Vigo, 1998.

PINHEIRO, M. H. D. O sistema consonantal do português arcaico visto através das cantigas profanas. 2004. Dissertação (Mestrado em Linguística) – FCL, Unesp, Araraquara.

SAID ALI, M. *Gramática Histórica da Língua Portuguesa*. Casa filial em São Paulo: USP, 1905.

WILLIAMS, E. B. *Do latim ao português: fonologia, morfologia históricas da língua portuguesa*. Tradução de Antônio Houaiss. 2. ed. Rio de Janeiro: Tempo Brasileiro, 1973.

# Análise morfológica de elementos lexicais em textos jornalísticos do português brasileiro

(Morphological analysis of lexical elements in journalistic texts in Brazilian Portuguese)

**Teresa Leal Gonçalves Pereira**

Instituto de Letras – Universidade Federal da Bahia

tlgpereira@gmail.com

**Abstract:** This is a study of lexical-semantic values, which is part of a broader research project currently in progress, about variation and change in journalistic texts in Brazilian Portuguese. The purpose of this text lies mainly in the analysis of lexical expressions marked in static and dynamic relations. The research is focused on neology and processes of word formation in a historical perspective of morphology.

**Keywords:** Neologisms; derivatives; compounds; morphology; Brazilian Portuguese.

**Resumo:** Este é um estudo dos valores léxico-semânticos, como parte de um projeto de pesquisa mais amplo que se encontra em desenvolvimento, acerca da variação e mudança em textos jornalísticos do português brasileiro. O objetivo desta comunicação reside, preponderantemente, na análise das expressões lexicais marcadas quer nas relações estáticas, quer nas relações dinâmicas. A investigação é voltada para a neologia e para os processos de formação de palavras, em uma perspectiva da morfologia histórica.

**Palavras-chave:** Neologismos; derivados; compostos; morfologia; português brasileiro.

## Introdução

Trata-se da amostragem de um levantamento de ocorrências de unidades lexicais e da sua mudança semântica, em textos jornalísticos atuais da cidade de Salvador, no Estado da Bahia, a partir de um *corpus* constituído de exemplos encontrados no período de 2011 a 2012 em alguns números da Revista *Muito*, encarte editado aos domingos, do jornal *A Tarde*.

A fim de privilegiar e demonstrar a vitalidade linguística, procede-se à verificação de expressividade e originalidade, na comunicação. São observados os processos de prefixação e derivação, especialmente os que decorrem de novos valores. Além do mais, são levados, também, em consideração alguns exemplos de processo de composição que refletem, na maioria das vezes, hábitos da fala característica da comunidade.

Observa-se a tendência de importação de palavras estrangeiras e a influência de termos provindos dos textos digitais e dos vários tipos da mídia eletrônica e televisiva. Essa importação confere a renovação do léxico da estrutura linguística, em geral, tão importante para a efetiva capacidade de comunicação das novas gerações de falantes, ao contrário do que preconizam os puristas e os gramáticos tradicionais.

Como ponto de partida, verifica-se o sentido de base a fim de se estabelecer a diferença entre as formas polissêmicas e as metáforas, levando-se em conta os princípios que regem

a codificação linguística e as assimetrias que, porventura, podem ser detectadas. Atenção especial é conferida aos esquemas imagéticos e demais motivações que poderão ser diagnosticados para as denominações registradas, com o intuito de identificar o seu papel na estrutura conceptual. Destarte, admite-se por princípio que as denominações linguísticas refletem o modo como se percebe e se experimenta a realidade do mundo externo e que a estrutura dos elementos polissêmicos não é arbitrária.

## A derivação lexical

Os neologismos morfológicos formados através da derivação lexical pela adição de sufixos, a partir de bases nominais e verbais, são bastante frequentes em textos midiáticos, uma vez que a expressividade é a marca essencial dos gêneros textuais que caracterizam a comunicação jornalística, radialista ou televisiva. A esse propósito, torna-se pertinente a observação seguinte acerca da relevância semântica inerente a tais formas:

Além da expressividade que se obtém pelo rompimento do bloqueio lexical, as formações derivadas podem ser expressivas quando a união entre afixo e base, mesmo sendo previsível pelo sistema, resulta em uma formação inesperada (*azulidade, despensamento*). (CARDOSO, 2010, p. 235)

É de evidência incontestável que os sufixos produzem alterações formais e semânticas quando agregados, tanto no que concerne ao sentido da base, quanto ao seu próprio significado.

Alguns autores, como Coutinho (1971), Travaglia (1998) e Rio-Torto (2008), chamam a atenção para os problemas que podem surgir na derivação sufixal e a conseqüente existência de Regras de Formação de Palavras (RFPs) e das respectivas restrições que delimitariam a criação e emprego desses neologismos.

As RFPs foram estabelecidas para possibilitar a constituição de uma lista de constituintes, tais como raízes, prefixos e sufixos, porém, como nem sempre essas regras podiam ser aplicadas, além da existência de sentidos e relações divergentes, houve a necessidade de serem criadas as chamadas “restrições” ou “bloqueios”, como ressalta Basílio (1987).

Travaglia (1998, p. 62), citando Aronoff, considera como bloqueio a não aceitação de adjetivos como *\*convencioso* e *\*religional*, em virtude da existência de *convencional* e *religioso* e ressalta que as restrições não linguísticas seriam aquelas decorrentes de: a) utilização apenas coloquial da base; b) medo do erro e do ridículo; c) possibilidade e/ou necessidade de analisar a condição intrínseca (a qualidade x) de algo ou alguém.

Seja qual for a restrição, sempre há casos em que ela é inoperante. Além disso, restrições como as não lingüísticas [...] sugerem que a questão da formação de palavras obedece a algo mais do que simplesmente fatores estritamente lingüísticos. O próprio bloqueio apenas lingüístico traz em si traços de uso social da língua. De qualquer modo todos os mecanismos propostos estão longe de resolver totalmente a questão da produtividade lexical [...]. (TRAVAGLIA, 1998, p. 63)

Para citar apenas um exemplo de derivação morfológica incompatível na língua portuguesa em Portugal, como *sair* / *\*saideiras*, observa-se que a sua ocorrência é verificada no Brasil. Até mesmo se encontra registrada na Gramática de Cunha e Lindley Cintra

(1984, p. 247-248) a restrição de grau de intensidade para uma forma como *paternal* / *\*muito paternal*, que, também, é facilmente encontrada em uso no português do Brasil.

Como se sabe, a variação e mudança semânticas dependem muito das relações sociais, históricas, culturais e da visão de mundo das comunidades linguísticas.

Este é claramente um domínio em que a realidade sociológica se reflecte na reordenação e na mudança sufixal, motivando-a e determinando-a. E é também um sub-sector da formação de palavras em que, fruto da coexistência de diversos afixos, com funcionalidades próximas e que se intersectam em alguns casos, não é linear simplificar as coisas e dizer que *-dora* representa o feminino gramatical de *-dor* e *-deira* o feminino gramatical de *-deiro*. A descrição da estrutura lexical dos derivados portadores destes sufixos acusa uma notória influência dos aspectos e dos papéis sociais e funcionais atribuídos aos indivíduos ou secundarizados ou sobrevalorizada, tem aqui lugar de eleição, que não pode ser ignorada ou cotornada. (RIO-TORTO; RENCA, 2007, p. 114)

Do levantamento realizado, na Revista *Muito* (M), destacam-se alguns exemplos de neologismos sufixais, tais como:

*-Ança*, sufixo aumentativo, do lat. *-antia*, que se encontra documentado em romance, entre os autores literários, semi-literários e, também, parcialmente, no uso popular, com reflexos românicos: fr. *-ance*, esp. ant. *-ança*, esp. mod. *-ancia/-encia*, port. *-ança/-ença/-ancia/-encia*.

Na edição de Piel (1942) do *Leal Conselheiro* (LC) de Dom Duarte, encontram-se documentadas, ao lado dos exemplos clássicos e que sobrevivem até os dias atuais, algumas formas arcaicas, concorrendo com outros derivados sufixais e que caíram em desuso, como *desvairança*, *igualdança*, *husança*, *trigança*:

- (i) *E assy em nossas obras e em nossa vyda nom devemos de husar em desvairanças*. LC, 248.9;
- (ii) *E de todallas cousas que som fremosas nom ha hi outra que o mais seja que hũa igualdança de toda a vyda*. LC, 248.2;
- (iii) *A husança das pirollas comũus pera esto achey muyto proveitosa, [...] LC, 77.5;*
- (iv) *Nom afroxando per fraqueza de vontade, nem nos torvando por trigança [...] LC, 105.15.*

Com a significação contextual de ‘grande quantidade’, empregada no sentido de ‘exagero no costume de se embegar’, sv. *Cachaçada* encontram-se registradas as seguintes acepções: ‘muita cachaça’, ‘bebedeira’, ‘borracheia’, no dicionário de Aulete, e também não há registro para o derivado nominal *borbulhança*, tal como se encontra documentado no exemplo (2).

- (1) [...] coma um tacacá na Gisela ou ao Parque 10, na *catchança* do Dadá, um dos melhores bares da cidade. (M, 18.03.2012).
- (2) A tarde vai se espreguiçando e a vontade que dá é de ficar ali, rindo com esse poeta das “*borbulhanças*”. (M, 25.03.2012).

*-Ista*, sufixo, cujo sentido é o do ‘que é partidário’:

Nos exemplos (3) e (4), *Butoísta* e *Fashionista* são produtos de formação híbrida que combinam uma base de origem estrangeira com um sufixo da língua portuguesa. *Butoísta* é derivado de “Butô ou Butoh, estilo de dança criado por Tahsumi Hijitaka, que carrega forte influência da dança expressionista alemã dos 1930 e da filosofia zen-budista”.

(3) Foi muito bom para mim ter decidido. Porque isso é o que me faz um artista, um dançarino, um *butoísta*. Só a chance de poder fazer isso já vale a pena”. (M, 18.03.2012) [ENDO, Tadashi, bailarino].

(4) De noivas a *fashionistas*. (M, 22 04.2012) [título de uma seção].

O mesmo processo de formação híbrida revela-se no exemplo da forma verbal *googuei*, com referência ao *site* do *Google*. Observa-se, nesse caso, por meio da projeção de esquemas imagéticos, a elaboração cognitiva entre a corporalidade neural e o artefato do computador.

(5) Inexistirá programa de humor melhor do que o PR, o PTB e congêneres na oposição? Qual será o discurso deles? “*Googuei*” Partido da República e corrupção juntos e aspeados, [...] (M, 25.03.2012).

*-Izar*, sufixo com valor aspectual causativo, factitivo ou resultativo, empregado como expressão de um processo, em “(6) *maximizar*, do inglês *to maximize* e em (7) *fidelizar*, do latim *fidelis*, ilustra formações de anglicismo e latinismo, respectivamente.

(6) A Pós FTC é a qualificação que você precisa para *maximizar* o seu sucesso profissional (M, 25.03.2012) [anúncio da FTC, p. 17].

(7) *EXOPLASTIA* CAPILAR: A SAÍDA INTELIGENTE PARA FUGIR DOS RISCOS DO FORMOL E *FIDELIZAR* SUA CLIENTE (M, 25.03.2012) [Anúncio de EXO original Hair, p. 14].

*-Iscar* é um sufixo portador de função produtiva de valor aspectual, imprimindo ao verbo o aspecto iterativo, que caracteriza situações durativas, como em (9).

(8) Aberta há quatro meses, a Barbearia e Choperia Cabelo & Barba oferece tudo o que a ala masculina mais gosta. [...] Para *mordiscar*, salgadinhos e tábua de frios. (M, 22.04.2012).

*-Ismo*, o significado desse sufixo, inicialmente, era próprio do nome de um partido dos carlistas, do que ou de quem defendia o direito alegado por Carlos de Bourbon, irmão de Fernando VII, de seus descendentes para ocuparem o trono espanhol, ou a Carlos X da França. No *corpus* em apreço, indica os afiliados a Antonio Carlos Magalhães, político de grande influência e poder, durante anos, no Estado da Bahia:

(9) Uma partida jogada por 40 anos entre ACM e seus adversários, o xeque-mate de Jaques Wagner, em 2006, foi uma jogada de Lula, que se tornou mais forte que ACM, para 3.242.336 eleitores (votos válidos) e abateu o *carlismo* com ACM vivo (M, 18.03.2012).

(10) Mas ACM não vai ressuscitar para jogar xadrez, e o *Carlismo* é um sistema político do século XX que ainda resiste em alguns estados, por pouco tempo, esperamos (M, 18.03.2012).

*-Eiro*, sufixo que exprime a ideia de sentido avaliativo intensivo, relativo aos adeptos da dança de estilo *reggae* ou da balada.

(11) [...] um dia precisei correr, com Márcia Ganem e Gil Maciel, porque quando um *regueiro* cantou que Selassie era um rei que gostava de reggae, eu gritei que além de reggae, ele gostava de matar e torturar. (M, 22.04.2012).

- (12) Aberto no fim do verão deste ano, o espaço é bem simplesinho e, em muito, destoa do perfil *baladeiro* do Rio Vermelho – [...] (M, 03.06.2012).

Os sufixos *-il*, *-ária* e *-ódromo* estão registrados no *corpus* em análise:

- (13) [...] Cristiane Rocha capricha nas perguntas para o casal interessado em adotar um dos filhotes disponíveis no *gatil*: [...] (M, 22.04.2012).
- (14) Despentei na Avenida Dendezeiros, uma das mais bonitas de Salvador, e vi, em meio às árvores *decenárias*, a igreja imponente. (M, 25.03.2012).
- (15) Em entrevista do Caderno 2+ desta terça-feira, o cantor baiano Lucas Santtana disse ser favorável à criação de um “*axeódromo*” (uma espécie de circuito privado) no Carnaval baiano, segundo a proposta idealizada por Durval Lelys. (M, 25.03.2012).
- (16) O que vocês acham do Durval, Ivete, Daniela e C&A montarem um “*Axelódromo*” em São Paulo e outro no Rio de Janeiro [...] (M, 25.03.2012).

Os sufixos *-ão* e *-zão*, com valor aumentativo, são usados para a formação de nomes como *folgão* e *axeção*.

- (17) Faz tempo que a bancária Rita Azevedo, 47, professora independente, não tem um *folgão*. (M, 22.04.2012).
- (18) Airthon (fev.18th 2010 at 3.00 [...] Proibiram-se os trios e as cordas, o *axeção* foi no Bonfim Light, [...] (M, 17.02. 2010). Disponível em: <<http://revistamuito.atarde.uol.com.br/?=4320>>. Acesso em 11 abr. 2012.

Na formação de verbos e de nomes, documentam-se os sufixos *-izar* e *-ização*.

- (19) Existem dois caminhos que vêm sendo trilhados por outras orquestras no País. Uma é *publicizar*, outra é concurso. Não estou certo da *publicização*, mas é um caminho. (M, 17.06. 2012).

*-Mento* – Sufixo que documenta, no *corpus*, uma formação antiga, correspondendo ao sinônimo *pertença* (AULETE, [2008]):

- (20) Hoje olho com certa melancolia para a decadência da península itapagipana, seu casario aristocrático profanado por lajes mal ajambradas, e me apavoro com a sensação de *pertencimento*. (M, 25.03.2012).

Os sufixos *-inho/ -inha*, formadores de diminutivo, são registrados com valor depreciativo em *vira-latinha* e *mundinho*. Há, também, a ocorrência de outro sufixo de caráter avaliativo: *-udo*, que se refere à representação de uma característica exagerada:

- (21) A sorte virou para a *vira-latinha* em dezembro de 2010 (M, 22.04.2012).
- (22) Os jardins verticais, também chamados quadros vivos, foram criados como solução para esta que é uma característica comum no nosso *mundinho* urbano. (M, 6.05.2012).
- (23) Não por acaso, foi a primeira MPB *classuda* pescada pela pirataria (M, 15.04.2012).

É digno de registro que há mudança de classe gramatical, a partir da derivação sufixal, como se pode formalizar no seguinte esquema:

N – V	Google –	<i>googuei</i>
	Máximo –	<i>maximizar</i>
	Público –	<i>publicizar</i>
A – V	Fiel (pelo lat. <i>fidel-</i> )	<i>fidelizar</i>
N – A	Balada –	<i>baladeiro</i>

O uso de prefixos está ilustrado em menor proporção relativamente à formação sufixal. Documenta-se o prefixo de intensidade, como se verifica em (23), que registra uma formação híbrida no uso de *super-*, acoplado ao ingl. *fly*, cujo significado implica movimento por meio de asas dos pássaros e aviões (GURALNIK, 1974):

- (24) Mais tarde, formou uma banda, a *Superfly*, que fez sucesso em Salvador nos anos 1889 e chegou a viajar para shows no Sudeste. (M, 17.06.2012).

Quanto ao prefixo *re-*, que indica repetição ou reiteração, encontra-se registrado nos seguintes exemplos:

- (25) [...] Wilson, que é professor da Escola de Nutrição da Ufba, reforça com os alunos a idéia da *refeitura* de comidas perdidas, pois elas são parte da gente. (M, 15.04.2012).
- (26) “Ao voltar aos palcos, milagrosamente, ficou boa. Como o grupo *refundou-se* na faculdade de Filosofia, lá foi ela. (M, 17.06.2012).

## Composição lexical

O processo de composição não se constitui como mais produtivo no *corpus* em apreço. Verificou-se a possibilidade de ser formado por duas palavras autônomas NN *samba-reggae*, *videoartista*, *carro-chefe*; NA *político-ideológica*, AN *afrobaile*, VN *vira-latinha*, além da redução, como na expressão de *sialata* ‘cruzamento de gato siamês com vira-lata’ ilustrada em (30). Importa assinalar que há o registro de *afro-cepticismo*, entre alguns casos de composição culta (ANTUNES; CORREIA, 2010, p. 157) considerados como “unidades não-autônomas” em um artigo relativo a neologismos de procedência da imprensa do português europeu, porém, no português brasileiro, *afro* é um adjetivo, muito comum para identificar a *música afro*, os *blocos afros* etc. Encontra-se documentado por Aulete (2008), como referente ao substantivo *afro* “o mesmo que africano” e como adjetivo de dois gêneros “Que é típico da África negra (*músicas afro*, *penteados afro*).”

- (27) O *samba-reggae* é uma música *político-ideológica* da comunidade negra, pulsante e dinâmica. (M, 13.11.2011).
- (28) [...] A dupla começa hoje (15) como o *Afrobaile*, que tem no repertório canções de Jorge Bem Jor [sic], Tim Maia, Rita Lee, Beбето, Tinhoãs, entre outros, [...] (M, 15.04.2012).
- (29) *Voltar-se para dentro*, como quem busca o conflito compartilhado, o *videoartista* Danilo Barata, 35, se propôs na instalação *O corpo como inscrição de acontecimento*. (M, 06.11.2011).
- (30) A sorte virou para a *vira-latinha* em dezembro de 2010 (M, 22.04.2012).
- (31) É um ‘*sialata*’, brinca, referindo-se à mistura das raças siamesa e vira-lata que enche o bichano de charme (M, 22.04.2012, *ibid.*).

- (32) Tendo Valentim como *carro-chefe*, as três ficarão em cartaz de 12 de dezembro a 16 de fevereiro (M, 13.11.2011).

É oportuno salientar que, no processo de composição, quando há dois radicais, pelo menos um não dispõe de autonomia, e não há marca de concordância s no segundo elemento, exceto quando se trata de formações de caráter coordenativo, com dois nomes (NN), em que o último assume a função determinante, como no exemplo (31). A marca do plural recairá obrigatoriamente, apenas, no primeiro nome *carros-chefe*.

### Aspectos semânticos

Na teoria da Semântica cognitiva, responsável pela mudança de paradigmas dos estudos linguísticos, consideram-se os Modelos Cognitivos Idealizados (MCIs), que se caracterizam pela projeção de categorias de nível básico para as formações mais abstratas.

Os MCIs estruturam espaços mentais que, na teoria do modelo cognitivo, substituem “as situações e os mundos possíveis”. Tal projeção tem a sua origem nos chamados “esquemas imagéticos”, compreendidos como meios elementares pelos quais se organizam as experiências, em grande parte, pela consciência humana do próprio corpo e da sua expansão no espaço e no tempo.

Torna-se perceptível o estabelecimento de uma estrutura holística, denominada “enquadramento” (ingl. *frame*), na forma verbal “*Googuel*” do exemplo (5), correlacionando a entidade apreendida do esquema imagético da enciclopédica computadorizada, *Google*, através do mapeamento metafórico da capacidade de apreensão do significado compreendido pelo contexto situacional e linguístico. Dessa forma, é que se pode estabelecer a metáfora conceptual PENSAMENTO É MÁQUINA.

No exemplo (13), encontra-se o esquema imagético do *container*, expresso na forma *gatil* ‘espaço onde se acomodam os gatos’, resultante da projeção da entidade percebida em um espaço convencional para os cães (*canil*), em relação a outras entidades compreendidas pela expansão de MCIs que se encontram no campo abrangente de animais de estimação.

Já nos exemplos (15) “*axeódromo*” e no (16) “*Axelódromo*”, o que se observa é o modelo idealizado denominado de “*cenário*” (ingl. *script*), em que há um evento por outro evento, através do mapeamento metonímico. A relação da estrutura holística, em que se efetua essa transposição cognitiva de enquadramento, provém da percepção do espetáculo promovido pela *performance* das Escolas de Samba do Rio de Janeiro – no *sambódromo* – e da realidade da música baiana, o Axé.

Não se pode negar a importância dos aspectos sociais, históricos e culturais, além do contexto discursivo, para a produção de sentido dos termos linguísticos, no entanto, é importante considerar, como assinala Lakoff (1987), que a língua é um domínio da cognição humana e, para melhor compreender como os seres humanos produzem os vários e numerosos tipos de categorização, é necessário admitir que essa é a forma inerente às línguas naturais.

Como processos morfológicos, a derivação e a composição desempenham um relevante papel na renovação linguística, especialmente no tocante à produção de novas formas lexicais e mudanças semânticas, resultante da criatividade e da imaginação humanas. Nesse particular, há que se considerar que novas categorizações surgem, a qualquer momento,

motivadas pela necessidade de se estar conectado com as criações de natureza cultural, artística e científica, reveladoras da visão de mundo das mais diversas comunidades.

Sabe-se que, além da capacidade da percepção objetiva das coisas e dos seres existentes na natureza, assim como das criações humanas, existe sempre a possibilidade de se produzirem extensões da realidade por intermédio da subjetividade, imaginação, fantasia, crenças e símbolos, que são responsáveis pela organização do conhecimento e formação das mais diversas concepções.

Ao se analisarem aspectos concernentes à neologia, detectam-se muitas das estratégias que presidem a variedade estrutural das línguas e, nos dias atuais, é importante levar em consideração a natureza cognitiva dos seres humanos. Segundo Lakoff (1987), as categorias podem ser consideradas em termos de um domínio para outro domínio a partir do esquema CONTAINER, em que se verifica um espaço bem delimitado e que pode ser transposto, através de mapeamento metafórico, em outro espaço de natureza semelhante ou, até mesmo, um espaço de configuração abstrata, fisicamente sem essas características, como os limites que são atribuídos a entidades como florestas e clubes, dentre outros tipos, com relação aos quais pode-se atribuir uma “entrada” ou um “saída”. Na categorização do esquema PERCURSO, o mapeamento metafórico admite a atribuição de fonte – desenvolvimento – destino.

A mente e corpo se encontram intimamente relacionados e processos motores interferem na imagem mental que surge no cérebro. Desde cedo, as imagens do espaço físico são apreendidas do próprio corpo humano. Como se encontra na obra fundamental de Lakoff e Johnson (2002), os conceitos espaciais de nível básico: – dentro-fora, em cima-embaixo, direita-esquerda – são assimilados, por extensão, dos limites e da estrutura do corpo humano e transpostos, das mais variadas formas metafóricas, a concepções mais abstratas.

Os esquemas imagéticos podem ser conceptualizados a partir de suas próprias estruturas, a fim de se relacionarem a estruturas mais complexas. Para o exemplo encontrado no *corpus*, em que o verbo *googuei* é metaforicamente transposto de uma categoria utilizada em um material computadorizado para uma entidade corporificada, especificamente, a capacidade cerebral de raciocinar, pode-se remontar ao que Gibbs (2006) se refere como “propriedades básicas dos espaços mentais”, tais como o poder de conter entidades mentais, ser estruturadas por modelos cognitivos relacionados a outros espaços e a capacidade de expandir-se por intermédio de entidades adicionais ou por modelos cognitivos idealizados.

Nos exemplos (15) e (16), verifica-se o mapeamento metonímico, quando foi identificado o esquema imagético CENÁRIO, caracterizado por um estado inicial, sequência de eventos e estado final, em que se percebe uma ocorrência incluída em um domínio conceptual único, estruturado por um Modelo Cognitivo Idealizado (MCI), no qual um elemento A é identificado como um membro ou subcategoria de um elemento B: do conceito de *sambódromo*, extraem-se *axeódromo* e *axelódromo*.

Dessa forma, procura-se demonstrar a mudança lexical provocada pela derivação e composição, assim como a natureza das transferências metafóricas e metonímicas encontradas no *corpus* constituído de exemplos da Revista *Muito*.

## REFERÊNCIAS

- ANTUNES, Mafalda; CORREIA, Margarita, Novos formantes da língua portuguesa: análise dos fractoconstituintes presentes no ONP. In: ALVES, Ieda Maria. (Org.) *Neologia e neologismos em diferentes perspectivas*. São Paulo: Paulistana, 2010. p. 147-172.
- AULETE, *Dicionário da língua portuguesa*. [2008] Disponível em: <[http://aulete.uol.com.br/site.php?mdl=aulete\\_digital](http://aulete.uol.com.br/site.php?mdl=aulete_digital)>. Acesso em: 8 set. 2012.
- BASÍLIO, Margarida. *Teoria lexical*. São Paulo: Ática, 1987.
- CARDOSO, Elis de Almeida. A criação neológica estilística. In: ALVES, Ieda Maria. (Org.) *Neologia e neologismos em diferentes perspectivas*. São Paulo: Paulistana, 2010. p. 229-249.
- COUTINHO, Ismael de Lima. *Gramática histórica*. 6. ed. Rio de Janeiro: Acadêmica, 1971.
- CUNHA, Celso; LINDLEY CINTRA, Luís Felipe. *Nova gramática do português contemporâneo*. Lisboa: João Sá da Costa, 1984.
- GIBBS Jr., Raymond W. *Embodiment and cognitive science*. Cambridge: Cambridge University Press, 2006.
- GURALNIK, David B. *Webster's new word dictionary of the american language*. New York: William Collis, 1974.
- LAKOFF, George. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago/London: The University of Chicago Press, 1987.
- LAKOFF, George; JOHNSON, Mark. *Metáforas da vida cotidiana*. Tradução de Grupo de estudos da indeterminação e da metáfora (GEIM), sob a coordenação de Mara Sophia Zanotto. Campinas: Mercado das Letras, 2002.
- RIO-TORTO, Graça. Para uma gramática do adjetivo. *Alfa*, São Paulo, v. 50, n. 2, p. 103-139, 2006. 2008. Disponível em: <<http://seer.fclar.unesp.br/alfa/article/view/1414>>. Acesso em: 12 abr. 2012.
- RIO-TORTO, Graça; RENCA, Nuno. Nomes em *-deira* em português, *Filologia e Linguística*, São Paulo, v. 9, p. 103-116, 2007. Disponível em: <[http://www.fflch.usp.br/dlcv/port/flp/images/arquivo/FLP9/Riotorto\\_Renca.pdf](http://www.fflch.usp.br/dlcv/port/flp/images/arquivo/FLP9/Riotorto_Renca.pdf)>. Acesso em: 12 abr.2012.
- TRAVAGLIA, Luiz Carlos. Hipótese de explicação discursiva para a mudança de significado e a formação de palavras. In: IX CONGRESSO INTERNACIONAL DA ASSOCIAÇÃO DE LINGÜÍSTICA E FILOLOGIA DA AMÉRICA LATINA (ALFAL), Campinas, 1990. *Atas...* Campinas: Instituto de Estudos da Linguagem da Unicamp, 1998. p. 58-77.

# A constituição e desconstituição dos termos *linguagem*, *língua* e *fala* de Saussure

(The constitution and deconstitution of the terms *langage*, *langue* and *parole* of Saussure)

Thayanne Raísa Silva e Lima<sup>1</sup>

<sup>1</sup>Instituto de Letras e Linguística – Universidade Federal de Uberlândia (UFU)

thayannerslima@hotmail.com

**Abstract:** This article presents the constitution and deconstitution of the terms *langage*, *langue* and *parole* of Ferdinand de Saussure. We will analyze the construction of this tripartition visiting firstly one of his manuscripts called *First Conference at the University of Geneva* (1891), and secondly, the manuscript entitled *Notes for course II* (1909), making notes of the delimitation and definition process of these three terms. However, we have noticed that the work of Saussure may have been deconstituted from the translation of the *Course* to English as there are different translations for these three terms. Therefore, we divided the work into four different moments: first: the presentation of the oldest manuscripts; second: manuscripts closer to the dates of the courses; third: the consolidation of the terms in the *CLG (Curso de Linguística Geral)*, and fourth: the deconstituted translations.

**Keywords:** Saussure, manuscripts, tripartition, translation.

**Resumo:** Neste artigo serão apresentadas a constituição e a desconstituição dos termos *linguagem*, *língua* e *fala* de Ferdinand de Saussure. Para analisarmos a constituição dessa tripartição, passamos primeiramente por um de seus manuscritos, chamado *Primeira Conferência na Universidade de Genebra*, de 1891, e, em segundo, pelo manuscrito intitulado *Notas para o Curso II*, de 1909, fazendo nota do processo em que surgiram as delimitações dessa tripartição. Contudo, notamos que o trabalho de Saussure pode ter sido desconstituído a partir das traduções do Curso para a língua inglesa, uma vez que elas apresentam distintas traduções para esses três termos. Portanto, dividimos o trabalho em quatro momentos diferentes. Primeiro: apresentação dos manuscritos mais antigos; segundo: manuscritos mais próximos às datas dos cursos; terceiro: a consolidação dos termos no *CLG (Curso de Linguística Geral)*; e quarto: a desconstituição nas traduções.

**Palavras-chave:** Saussure, manuscritos, tripartição, tradução.

## Introdução

Ferdinand de Saussure (1857-1913), pai da Linguística Moderna, trabalhou para criar um objeto para a Linguística. E assim trabalhou bastante tempo ministrando aulas, escrevendo sobre suas ideias, ou projetos de ideias, que circundavam os estudos linguísticos, como, por exemplo, a crítica ao que estava sendo estudado como ciência linguística até então. Em 1907, o genebrino repassou ideias e novos conceitos aos estudantes do curso que passara a ministrar na Universidade de Genebra. Posteriormente à sua morte, dois de seus alunos, Charles Bally e Albert Sechehaye, resolveram organizar e editar as notas que os discentes fizeram durante os três cursos, surgindo, portanto, o livro conhecido como *Curso de Linguística Geral*.<sup>1</sup>

---

<sup>1</sup> Doravante CLG.

O CLG repercutiu em todo o mundo, tendo hoje tradução para mais de vinte línguas. Além de ser uma das principais referências aos estudos linguísticos, ele se tornou o fundador da linguística moderna e do Estruturalismo. Porém, dentre essas diversas traduções existem algumas com bastantes peculiaridades, em que o processo de tradução teve seus embates e críticas, como as traduções para o russo e o chinês.<sup>2</sup> Contudo, neste momento, focaremos somente na tradução feita para a língua inglesa.

A tradução do *Curso* para a língua inglesa foi feita primeiramente em 1959 por Wade Baskin e, posteriormente, em 1983, por Roy Harris. A primeira foi bastante criticada, como o próprio Harris afirma: “Saussure has on the whole been poorly served by his English translator and commentators”<sup>3</sup> (1983, p. xiii, Prefácio), apontando, portanto, como essa tradução de Baskin foi vista por aqueles que estudavam Saussure na época. Em seguida, Harris fez uma nova tradução do CLG, mas que também foi alvo das críticas, como justifica Evans (1991, p. 191), ao optar fazer citações da tradução de Baskin em seu livro: “Escolhi usar a tradução do *Cours* de Wade Baskin, em vez da recente tradução de Roy Harris. Enquanto a tradução de Harris é, em muitos aspectos, uma melhoria, em pontos cruciais ela implica uma interpretação em desacordo com o texto original” (tradução nossa).

Faremos, portanto, uma análise do trabalho de anos que Saussure levou para constituir limites entre os conceitos de *linguagem*, *língua* e *fala*. Para tanto, utilizaremos os manuscritos *Primeira Conferência na Universidade de Genebra*, datado de 1891, e outro escrito um pouco mais tarde, de 1909, intitulado *Notas para o Curso II*. Em ambos, apontaremos as diferenças e semelhanças apresentadas nessa tripartição de Saussure até a constituição final dos termos no *Curso de Linguística Geral*.

Posteriormente a essa análise, faremos nota do processo de desconstituição desses conceitos presentes nas controversas traduções do *Curso* para a língua inglesa. Isso se deve ao fato de Baskin ter traduzido *língua*, *linguagem* e *fala* de uma forma e Harris ter escolhido traduzi-los de uma maneira completamente distinta.

De acordo com tais diferenças, partimos do pressuposto de que todo aquele processo de constituição dos limites entre os conceitos – notado nos manuscritos de Saussure – foi desconstituído por essas diferentes traduções. Analisaremos tais questões a partir dos primeiros esboços desses termos, até que virassem conceitos próprios da linguística estruturalista.

### **A constituição dos conceitos *linguagem*, *língua* e *fala***

*Linguagem*, *língua* e *fala* são conceitos que há muito circundavam os estudos da linguística. Desde a gramática comparada já se falava neles, mas não havia uma distinção entre os mesmos. Foi somente com Saussure, no início do século XX, que os mesmos foram definidos e constituídos de forma que conseguíssemos notar limites que os separavam, e os conceituavam. Assim, Normand (2009) afirma que “estes [conceitos] são elaborados com o cuidado de operar as demarcações fundadoras, elas vão definir o ponto de vista que o linguista deve adotar e que, simulando aquele do locutor, distancia-se do que Saussure chama de “ciências conexas” (história, sociologia...)” (p. 49). Logo, a

2 Ver Lima (2011).

3 “Saussure tem sido, sobretudo, mal servido por seu tradutor inglês e comentaristas” (tradução nossa).

partir dessa delimitação, a tripartição saussuriana tomou grande importância nos estudos linguísticos como afirma Silveira (2007, p. 20): “as questões que o livro coloca sobre a língua, a fala e a linguagem marcam a linguística que, a partir daí, não está mais diante do mesmo objeto”.

Dessa forma, abordaremos os manuscritos para nos situar em relação a esse processo de constituição da teoria saussuriana, detendo-nos principalmente na elaboração dos termos *linguagem*, *língua* e *fala* que são essenciais na compreensão de toda a teoria supracitada.

### **O manuscrito *Primeira Conferência na Universidade de Genebra***

O primeiro manuscrito que trabalharemos, intitulado por Godel<sup>4</sup> como *Primeira conferência na Universidade de Genebra*, está datado, também por ele, como de novembro de 1891, exatos 16 anos antes do primeiro *Curso de Linguística Geral*, e de acordo com Silveira (2011), é bastante provável que esse manuscrito, seguido de outros dois que são nomeados como segunda e terceira conferência, “[...] trata-se das três primeiras aulas de Saussure em Genebra, de um curso que se chamava ‘Fonética do Grego e do Latim’ (SILVEIRA, 2011, p. 4).

Apesar de o nome dos manuscritos e do curso não fazerem sequer referência aos termos que nos propomos a investigar, no decorrer da leitura de tais obras conseguimos identificar aspectos que se referiam a eles, ou até mesmo rasuras feitas por Saussure, em que pudéssemos ler, por debaixo delas, os termos que procuramos.

Já na primeira página desse manuscrito nos deparamos com a palavra *langage* suprimida pelas rasuras de Saussure: “[...] cette j’aurais à définir la place qui revient dans le cercle des connaissances humaines à la science ~~du langage~~” (aula inaugural, f. 1; linhas 12-13 apud 3951-1/PS/BGE)<sup>5,6</sup>

Notamos, portanto, já uma primeira hesitação do autor em colocar a palavra *linguagem*, em que nos parece que ele tivera a intenção de apagar essa questão de ser “[...] uma ciência da linguagem”, deixando somente o conceito “[...] dos conhecimentos humanos a essa ciência”. Dessa forma, podemos inferir desse fragmento uma primeira incerteza de Saussure quanto ao conceito de *linguagem*.

Na página 8 do manuscrito, vemos o ponto que confirma explicitamente a não distinção dos termos. Saussure aponta: “Langue et langage ne sont qu’une même chose; l’un est la généralisation de l’autre” (SAUSSURE, f. 8; linhas 23-25).<sup>7</sup> Vemos que esse conceito não condiz com os ensinamentos saussurianos abordados nos cursos do início do século XX, em que o autor faz uma distinção clara entre tais termos, como veremos a seguir. As palavras “[...] uma é a generalização da outra” caracterizam *língua* e *linguagem* de uma forma ainda bem distante da que conhecemos no CLG.

4 Apesar de utilizarmos neste trabalho o livro *Escritos de Linguística Geral*, com a transcrição dos manuscritos, quem originalmente nomeou tal conjunto foi Godel, que também os catalogou.

5 “[...] de estabelecer o lugar que ela ocupa no círculo dos conhecimentos humanos [...]” (SAUSSURE, 2002, p. 126).

6 Os manuscritos aqui utilizados nos foram cordialmente cedidos pela Prof. Dra. Eliane Mara Silveira.

7 “Língua e linguagem são apenas uma mesma coisa: uma é a generalização da outra” (SAUSSURE, 2002, p. 128).

Essa afirmação é, na verdade, totalmente contrária ao conceito final apresentado por Saussure.

Na página 5, notamos novamente que o genebrino ainda não tinha bem definida a distinção entre *língua* e *linguagem*, e ainda acrescenta o terceiro termo *tripartição*. Apesar de o fragmento estar rasurado por ele, notamos: “~~Le langage? Mais la parole? C’est une chose que nous oublions généralement parce que [...]~~” (SAUSSURE, f. 5; linhas 21-23).<sup>8</sup>

Vemos aqui que Saussure refere-se ao termo *parole* que ainda não aparecera no manuscrito, porém no próximo parágrafo, já na página 6, temos a formulação de uma frase que suprime o termo *fala*: “Le langage ou la langue peut-il passer pour un objet qui appelle par lui-même, l’étude?” (SAUSSURE, f. 6; linhas 1-3).<sup>9</sup>

Lembramos o que Silveira (2007) ressalta acerca das rasuras desse mesmo manuscrito: “[...] o que se rompe retorna como repetições ou mesmo integrado no texto, o que já aponta para um deslocamento na elaboração de Saussure” (SILVEIRA, 2007, p. 125). Entretanto, no caso dessa última rasura, ele optou por não retornar, como vimos no segmento acima. Seguindo nesta página, o autor aponta: “[...] ~~parole~~ [...] ou come on dit le ~~langage~~ articulée” (SAUSSURE, f. 6; linhas 11-12).<sup>10</sup>

Notamos que o autor rasura a palavra *langage*, depois acrescenta a palavra *parole* acima dela, e ainda assim a risca, demonstrando-nos exatamente essa distinção que ainda não existia para o próprio Saussure. Apesar de os editores do *Escritos de Linguística Geral* optarem pela palavra *fala* na transcrição dessa passagem, Saussure não opta por nenhuma das duas na elaboração do seu manuscrito. Portanto, concordamos com Vinhais (2011, p. 4), que afirma que “[...] os três termos, fala, língua e linguagem ainda estavam num emaranhado, isto é, apenas num *a posteriori* Saussure delimitaria seus conceitos, ou pelo menos, suas diferenças”.

Notamos, no manuscrito acima citado, um movimento de constituição dos termos. Saussure demonstra-se incerto sobre a expressão que seria mais adequada para a construção de suas frases; ele deixa muitas delas incompletas e até retira algumas do seu corpo de texto. Essa primeira conferência nos mostra um momento claro de elaboração dessa tripartição.

## O manuscrito *Notas para o Curso II*

Ao partir para a análise do outro manuscrito, intitulado *Notas para o Curso II (1908-1909): Dualidades*,<sup>11</sup> conseguimos notar uma grande aproximação do que encontramos no CLG. Vemos, por exemplo, quando ele discorre sobre a *linguagem*: “Le langage ne promettrais que l’idée d’une multiplicité, elle-même composée de faits hétérogènes [...]” (SAUSSURE, 3951-22, f. 1).<sup>12</sup>

8 A linguagem, mas a fala? Isso é algo que costumamos esquecer, porque [...]” (tradução nossa).

9 “Pode-se considerar a linguagem ou a língua como um objeto que pede por si mesmo, esse estudo?” (SAUSSURE, 2002, p. 128).

10 “Ou como se diz, a fala articulada” (SAUSSURE, 2002, p. 128).

11 Notamos que esses foram escritos anos depois do primeiro, e que as datas coincidem com o momento em que Saussure ministrava aulas sobre Linguística em Genebra.

12 A linguagem não promete uma ideia de multiplicidade; ela própria é composta de fatos heterogêneos (tradução nossa).

Percebemos aqui a *linguagem* já vista como heterogênea. Depois temos uma abordagem do conceito de *língua* seguindo-se como:

La langue et consacrée socialement et ne depend pas de l'individu. Est de l'individu, ou de la Parole:

- a) Tout ce qui est Phonation
- b) Tout ce qui est combination
- c) Tout ce qui est Volonté

Dualité Parole/Langue Volonté individual/Passivité sociale.<sup>13</sup>  
(SAUSSURE, 3951-22)

Nesse fragmento é apresentado o discurso abordado no *Curso de Linguística Geral*, como afirmações de que a *língua* é social e de que a *fala* é individual, por exemplo, conceitos essenciais para a construção da teoria saussuriana. As conhecidas dicotomias já se apresentam nesse momento como dualidades, reafirmando essa aproximação da elaboração final do *Curso de Linguística Geral*.

Notamos, portanto, que no primeiro manuscrito que trabalhamos há uma grande distância acerca do que conhecemos da tripartição saussuriana, já nesse segundo manuscrito há uma aproximação do que nos é apresentado posteriormente no *Curso de Linguística Geral*. Contudo, eles ainda não são exatamente como no curso.

### **A tripartição no *Curso de Linguística Geral***

Depois de fazermos nota de todo esse processo em que nos fica claro que Saussure trabalhou durante muitos anos para chegar aos conceitos finais de *linguagem*, *língua* e *fala*, apontaremos como essa tripartição se apresenta no CLG com conceitos bem distintos entre si, apesar de estarem correlacionados, e, se valerem uns dos outros para serem explicados, conseguimos diferenciá-los. É nesse momento que Saussure marca os estudos da linguística, mostrando um estudo diferente dos que estavam sendo feitos na época.

Primeiramente veremos como o conceito de *linguagem* está no CLG (SAUSSURE, 2006):

A linguagem é multiforme e heteróclita; o cavaleiro de diferentes domínios, ao mesmo tempo física, fisiológica e psíquica, ela pertence além disso ao domínio individual e ao domínio social [...] (p. 17). Enquanto a linguagem é heterogênea, a língua assim delimitada é de natureza homogênea [...] (p. 23). O estudo da linguagem comporta, portanto, duas partes: uma, essencial, tem por objeto a língua [...] outra, secundária, tem por objeto a parte individual da linguagem, vale dizer, a fala. (p. 27)

Ao compararmos as formulações de Saussure nessas notas com o que temos no CLG, podemos notar o quanto os termos e conceitos se aproximam na explicação – Saussure inclusive usa a mesma palavra (“heterogênea”) para caracterizar a *linguagem*.

<sup>13</sup> “A língua é consagrada socialmente e não depende do indivíduo. É do indivíduo, ou da fala: a) Tudo o que é Fonação; b) tudo o que é combinação – tudo o que é Vontade. Dualidade: Fala/Língua e Vontade individual/passividade social” (SAUSSURE, 2002, p. 258).

Já para o conceito de *língua* presente no CLG (SAUSSURE, 2006), temos o seguinte:

Ela é um objeto bem definido no conjunto heteróclito dos fatos da linguagem. (p. 22) [...]. A língua, distinta da fala, é um objeto que se pode estudar separadamente (p. 22) [...]. É, ao mesmo tempo, um produto social da faculdade da linguagem e um conjunto de convenções necessárias, adotadas pelo corpo social para permitir o exercício dessa faculdade dos indivíduos. (p. 17)

Com tais conceitos percebemos novamente uma congruência com o manuscrito *Notas para o Curso II*, que utiliza novamente a palavra “social” para definir a *língua*, mas ao mesmo tempo, completamos a ideia daquele contraponto citado anteriormente, em que Saussure coloca no manuscrito da primeira conferência que “[...] língua e linguagem são apenas uma mesma coisa: uma é a generalização da outra”, o que fica completamente diferente do que aparece no CLG. Isso nos mostra exatamente esse movimento de constituição do conceito que podemos perceber durante todos esses anos, pois de 1891 a 1913, os conceitos saussurianos foram amadurecendo e tomando outros rumos.

Para concluir, trazemos a definição de *fala* presente no CLG (2006, p. 22):

Com o separar a língua da fala, separa-se ao mesmo tempo: 1º, o que é social do que é individual; 2º, o que é essencial do que é acessório e mais ou menos acidental [...]. A fala, é ao contrário [da língua], um ato individual de vontade e inteligência, no qual convém distinguir: 1º, as combinações pelas quais o falante realiza o código da língua no propósito de exprimir seu pensamento pessoal; 2º, o mecanismo psicofísico que lhe permite exteriorizar essas combinações.

Dessa forma, fechamos com a certificação de que novamente os conceitos de *fala* não se diferem bastante das notas feitas pelo genebrino. Contudo, a palavra *parole* aparecia timidamente entre as rasuras de Saussure, e no *Curso de Linguística Geral* ela também não foi ponto de grande referência, mas esse já se torna outro tópico de discussão que não nos diz respeito neste momento.

Até então notamos dois processos distintos: num primeiro momento, vimos como Saussure elaborava os termos que analisamos em seus manuscritos, criando conceitos e frases que foram rasuradas e desenvolvidas com novos termos; num segundo momento, notamos no CLG que esses conceitos já estavam estabelecidos e explicados ao longo do livro, com muitos exemplos e esclarecimentos. A seguir, veremos como essa tripartição que, como notamos, levou anos para ser constituída, foi somente uma tradução desconstituída que se perde em conceitos não muito bem elaborados.

## **A desconstituição nas traduções para o inglês**

Nas traduções para a língua inglesa, deparamo-nos com uma complicação nesses termos que se apresentaram como essenciais na compreensão das ideias saussurianas. Como dito antes, temos duas versões de tradução do CLG na língua inglesa, e veremos aqui como os termos *linguagem*, *língua* e *fala* podem ter sido desconstituídos.

A primeira tradução, feita em 1959 por Wade Baskin, suscitou várias críticas dos estudiosos de Saussure, sendo que muitos linguistas americanos, ao citar o mestre genebrino, preferiam citá-lo na sua língua original – vemos isso nas obras de Chomsky (1964)

e Labov (1972), por exemplo. Além disso, no prefácio da segunda edição, Harris (1983) aponta vários problemas na tradução de Baskin, como:

Suffice it to say that the varied catalogue of mistranslations available for public inspection runs the whole gamut from the trivial to the grossly misleading (*langage* rendered as “speech”). On crossing the Channel Saussure has been made to utter such blatantly un-Saussurean pronouncements as “language is a form, not a substance”.<sup>14</sup> (HARRIS, 1983, p. xiii)

Harris (1983) propôs uma nova tradução, em que prometia ser “Primeiramente ao leitor que não é um especialista em linguística” (p. xiv), o que presumimos ser uma leitura tranquila para aqueles que se interessam por Saussure; contudo, a nova tradução também foi alvo de críticas, uma vez que Harris foi acusado de ser interpretativo, como vimos anteriormente. É interessante notar que, em seu próprio livro, publicado após a tradução do CLG, ele mesmo aponta esses problemas, e decide utilizar os termos em francês, afirmando, portanto, que:

Anyone who chooses to write about Saussure in English immediately takes up various self-inflicted burdens [...]. Be that as it may, there soon comes a point at which any commentator must have recourse to citing the French texts available. I have done this wherever it seemed to me that the original wording was important or indispensable for the clarification of a particular issue (which is most of the time).<sup>15</sup> (HARRIS, 2001, Prefácio)

Para explicitar como essa desconstituição acontece nas duas traduções, elaboramos uma tabela com os termos em português e em ambas as traduções.

**Tabela 1 – Traduções de *Linguagem, língua e fala* de acordo com Saussure (1959, 1983)**

	<b>Baskin</b>	<b>Harris</b>
Linguagem	Speech	Language
Língua	Language	1 – linguistic structure
		2 – language system
		3 – a language/the language
Fala	Speaking	Speech

Esse quadro nos indica que os termos indicados à esquerda dele se tornaram diferentes tanto em uma como em outra tradução – nos parece que a proposta de distinção

14 “Basta dizer que o catálogo variado de erros de tradução, disponível para consulta pública, corre toda a gama do trivial à mais grosseiramente enganosa (*langage* traduzida como ‘speech’). Ao cruzar o Canal Saussure, têm sido feitas declarações descaradamente antissaussurianas como ‘linguagem é uma forma, não uma substância’” (tradução nossa).

15 “Qualquer pessoa que escolhe escrever sobre Saussure em inglês ocupa, imediatamente, vários encargos autoinfligidos [...]. Seja como for, lá em breve chega a um ponto em que qualquer comentador deve recorrer a citar os textos disponíveis em francês. Eu tenho feito isso sempre que me pareceu que a redação original era importante ou imprescindível para o esclarecimento de uma questão específica (que é na maioria das vezes)” (tradução nossa).

que Saussure elabora no CLG fica perdida. Notamos que as traduções parecem confundir os termos ou nos levam a um processo de leitura que requer muito cuidado e atenção. Aquela tripartição proposta por três termos diferentes em francês (*langage, langue e parole*) apresenta-se, pois, com dificuldade na língua inglesa, uma vez que não temos os três termos no idioma em questão.

## Conclusão

Diante dos aspectos supramencionados, podemos concluir que o processo de constituição dos termos *linguagem, língua e fala* levou anos para ser consolidado. Ferdinand de Saussure demonstra titubear em relação a tais conceitos em seus manuscritos de 1891, mas já nas notas para seu curso vemos as delimitações bastante claras entre os termos, assim como observamos no CLG essas demarcações em acordo com as notas. Assim criou-se uma teoria que se tornou conhecida mundialmente, sendo traduzida para mais de vinte línguas.

Entretanto, os tradutores do CLG em inglês encontraram grandes problemas durante a tradução, desconstituindo os conceitos de Saussure e fazendo com que os leitores pudessem interpretar de maneira equivocada a tripartição saussuriana. Isso pode desencadear uma má compreensão de elementos posteriores na leitura do *Curso de Linguística Geral*.

Acreditamos que tanto Baskin quanto Harris se depararam com questões complicadas no momento de tradução de tais conceitos. Assim, defendemos que, diante de todas essas complicações, os conceitos de Saussure podem não ter ficado bem claros para os leitores do Curso em língua inglesa, e isso causou uma desconstituição da tripartição de Saussure; uma vez que isso acontece com termos tão essenciais da obra dele, acreditamos também que a leitura do CLG pode ficar comprometida e assim criar leitores, e até linguistas, que não compreendem uma obra tão importante para os estudos da linguagem, como aponta Sanders (2000, p. 355): “[...] the CGL had become a canonical text which was known but still misunderstand and misquoted”.<sup>16</sup>

Desse modo, o processo de constituição do genebrino parece ser desconstituído na primeira tradução e, apesar das críticas feitas a ela pelo próprio Harris, ainda permanecemos com uma versão do CLG em inglês com termos que desconsideram todo o processo que Saussure levou para estabelecê-los, os quais são fundamentais para a Linguística.

## REFERÊNCIAS

CHOMSKY, Noam. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press, 1964. p. 23

EVANS, Joseph Claude. *Strategies of deconstruction*. Minneapolis: University of Minnesota Press, 1991, 208 p.

---

<sup>16</sup> O CLG tornou-se um texto canônico, bastante conhecido, mas que ainda não era bem compreendido ou erroneamente citado (tradução nossa).

HARRIS, R. *Saussure and his interpreters*. Edinburgh: Edinburgh University Press, 2001. 224 p.

\_\_\_\_\_. Translator's Introduction. In: SAUSSURE, Ferdinand de. *Course in General Linguistics*. London: Open Court Classics, 1983. p. ix – xvi.

LABOV, W. *Sociolinguistic Patterns*. Pennsylvania: University of Pennsylvania Press, 1972. 347 p.

LIMA, T. R. S. Os problemas de tradução no Curso de Linguística Geral. In: SIMPÓSIO INTERNACIONAL DE LETRAS E LINGUÍSTICA, 2., n. 2, 2011, Uberlândia. *Anais...* do SILEL. Uberlândia: EDUFU, 2011.

\_\_\_\_\_. *Saussure*. Tradução de Ana de Alencar e Marcelo Diniz. São Paulo: Estação Liberdade, 2009. 183 p.

NORMAND, C. *Convite à Linguística*. Organizado por Valdir Flores e Leci Borges Barbisan. Tradução de Cristina de Campos Velho Birck. São Paulo : Contexto, 2009

SANDERS, Carol. Saussure Translated. *Historiographia Linguistica*, XXVII 2/3, p. 345-358. 2000.

SAUSSURE, F. *Course in General Linguistics*. Tradução de Wade Baskin. Nova Iorque: Philosophical Library, 1959. 260 p.

\_\_\_\_\_. *Course in General Linguistics*. Tradução de Roy Harris. EUA: Open Court Classics, 1983. 236 p.

\_\_\_\_\_. *Escritos de Linguística Geral*. Organizados e editados por Simon Bouquet e Rudolf Engler. Tradução de Carlos Augusto Leuba Salum e Ana Lucia Franco. São Paulo: Cultrix, 2002. 299 p.

\_\_\_\_\_. *Curso de Linguística Geral*. Tradução de Antônio Chelini, José Paulo Paes e Izidoro Blikstein. São Paulo: Cultrix, 2006.

SILVEIRA, E. M. *As marcas do movimento de Saussure na fundação da Linguística*. Campinas: Mercado de Letras, 2007. 168 p.

\_\_\_\_\_. Uma leitura preliminar de dos manuscritos de Ferdinand de Saussure: ‘*Conférences à l’Université*’ e ‘*L’essence Double du langage*’. In: SIMPÓSIO INTERNACIONAL DE LETRAS E LINGUÍSTICA, 2., 2011, Uberlândia. *Anais...* Uberlândia: EDUFU, 2011.

VINHAIS, E. A. A noção de fala em Ferdinand de Saussure: dos manuscritos ao Curso de Linguística Geral. In: SIMPÓSIO INTERNACIONAL DE LETRAS E LINGUÍSTICA, 2., 2011, Uberlândia. *Anais...* Uberlândia: EDUFU, 2011.

# O processo de aquisição de verbos irregulares no português brasileiro

(The acquisition process of irregular verbs in Brazilian Portuguese)

Aline Garcia Rodero Takahira<sup>1</sup>

<sup>1</sup> Departamento de Linguística – Universidade de São Paulo (USP)

alinegr@usp.br

**Abstract:** In this paper, we discuss the irregular verb formation and acquisition process in Brazilian Portuguese, considering cases of regularization as *fazi/trazi* for *fazer/trazer*. This acquisition process is frequently explained by the binary mechanism model of inflection that involves two distinct representational systems (CLASHEN et al., 2002). Our main goal is to point to a treatment of this phenomenon in a theory that avoids the presupposition that the same verb needs two processes to be formed. We consider Siddiqi (2009) that defends that the root, the head *v* and the past feature merge forming a complex functional head, and the vocabulary item (VI) enters in only one place. We explain the regularization process as a case in which the child does not form this complex functional head, then the regularized form: the root of *fazer/trazer*, the head *v* and the VI that occurs in past form *-i* or *-u* are separately inserted.

**Keywords:** acquisition; verbal inflection; Distributed Morphology; irregular verbs.

**Resumo:** Neste trabalho, discutimos o processo de aquisição e formação de verbos irregulares no português brasileiro, abordando casos de regularização como *fazi/trazi* para *fazer/trazer*. Esse processo de aquisição muitas vezes é explicado pelo modelo do mecanismo binário de flexão que envolve dois sistemas representacionais distintos (CLASHEN et al., 2002). Nosso objetivo principal é apontar para um tratamento desse fenômeno dentro de uma teoria que evita a pressuposição de que um mesmo verbo necessite de dois processos para ser formado. Para tanto, nos valem do trabalho de Siddiqi (2009), que defende que a raiz, o núcleo *v* e o traço de passado se fundem formando um nó funcional complexo, e o item de vocabulário (IV) entra em um lugar só. Explicamos o processo de regularização como o caso em que a criança ainda não forma esse nó funcional complexo, então a forma regularizada: a raiz de *fazer/trazer*, o núcleo *v* e o IV que se realiza no passado *-i* ou *-u* são inseridos separadamente.

**Palavras-chave:** aquisição; flexão verbal; Morfologia Distribuída; verbos irregulares.

## Introdução<sup>1</sup>

Verbos irregulares passam por um processo de regularização durante o processo de aquisição. Kirn (1986) afirma que, uma vez que a criança desenvolve uma regra, ela a generaliza aos aspectos irregulares da língua, como a regra do uso de *-ed* no tempo passado generalizada em “I goed” em vez de “I went”. Esse processo de aquisição muitas vezes é explicado pelo modelo do mecanismo binário de flexão que envolve dois sistemas representacionais distintos. Clashen *et al.* (2002) usam dados do espanhol e explicam esse processo de aquisição como: i) um grupo de entradas lexicais listadas associativamente na memória (formas irregulares); e, ii) regras para formar expressões linguísticas maiores (formas regulares). Os autores fazem uma separação entre a formação do radical e o domínio da flexão, assim, no caso dos verbos irregulares, quando a criança superaplica a regra, ela

---

<sup>1</sup> Agradeço à Capes pelo financiamento da pesquisa e à Professora Doutora Ana Paula Scher pela orientação.

aciona uma forma irregular memorizada e depois aplica uma regra de formação de verbos regulares, usando dois processos diferentes para formar um único verbo irregular.

Neste trabalho, discutimos o processo de aquisição e formação de verbos irregulares no português brasileiro (PB), abordando casos de regularização dos verbos irregulares no pretérito perfeito do indicativo. Usamos dados apontados em Maldonade (2003), Lorandi (2006) e Figueira (2010). Tratamos de dados como *fazer* e *trazer*, que, durante o processo de aquisição, se realizam como *fazi* e *trazi*, na primeira pessoa do singular do pretérito perfeito do indicativo, em vez de *fiz* e *trouxe*; e *fazeu* e *trazeu*, na terceira pessoa do singular do mesmo tempo e modo, em vez de *fez* e *trouxe*.

Os objetivos deste trabalho são:

i) trazer uma discussão sobre tratamentos da regularização de verbos irregulares no pretérito perfeito do indicativo no PB; e,

ii) apontar para um tratamento desse fenômeno dentro de uma teoria não-lexicalista, que utiliza apenas um componente gerativo da gramática.

O segundo objetivo nos leva a uma análise que evita a pressuposição de que um mesmo verbo, o irregular regularizado pela criança, necessita de dois processos para serem formados.

Para tal análise, nos valem do modelo da Morfologia Distribuída (MD) (HALLE; MARANTZ, 1993), uma teoria não-lexicalista, na medida em que considera que o que em modelos lexicalistas era resolvido no léxico está, de fato, distribuído pela derivação. Esse modelo trabalha com a noção de que nos nós terminais há feixes de traços que são especificações dos traços que o item de vocabulário (IV) requerido deve conter. Há uma competição entre IVs e o mais especificado, que não apresenta traços divergentes dos requeridos no nó terminal, ganha a competição e será inserido. Dentro da MD, seguimos o trabalho de Siddiqi (2009), que defende que a raiz, o núcleo *v* e o traço de passado se fundem formando um nó funcional complexo, e o IV entra em um lugar só.

Chamamos atenção para algumas questões já colocadas na literatura:

1- Por que as crianças cometem erros de regularização?

2- O que faz esse tipo de erro acabar?

Mais uma questão se coloca:

3- Dentro de um modelo não-lexicalista (como a MD), como explicamos o fato de a criança passar por um período no qual ela produz *fiz* e depois passar a usar *fazi* e voltar a usar *fiz*? No primeiro momento, ela ainda não adquiriu, de fato, a forma irregular?

Buscamos mostrar como a MD possibilita explicar o processo de aquisição de verbos irregulares. Nesse sentido, vamos investigar a hipótese de que o processo de regularização pelo qual a criança passa é o caso em que a criança ainda não forma um nó funcional complexo, então, a forma regularizada: a raiz de *fazer/trazer*, o núcleo *v* e o IV que se realiza no passado *-i* ou *-u* são inseridos separadamente.

## A regularização dos verbos irregulares e a aquisição de morfologia

Muitos dos verbos no português brasileiro são regulares, ou seja, são formados seguindo os padrões que podemos observar em (1):

(1) T (R + VT) + SF (SMT + SNP)

Em (1), temos a constituição morfológica do verbo português definida por Câmara Jr. (1991 [1970]; 2004 [1972], p. 132-133). Os sufixos modo-temporal e número-pessoal “se aglutinam intimamente num global sufixo flexional (SF), que se adjunge ao tema do verbo (T), constituído pelo radical (R) seguido da vogal temática (VT) da conjugação correspondente” (CÂMARA JR., 1991 [1970]).

Esses verbos regulares se contrapõem aos verbos irregulares. Os casos estudados neste trabalho são aqueles nos quais o Pretérito Perfeito do Indicativo é formado de uma outra forma com alguma modificação na raiz.

Os adultos conhecem as formas do Pretérito Perfeito do Indicativo dos verbos irregulares, como *fazer – fiz*, e também sabem que os padrões regulares não podem ser aplicados nesses casos, ou seja, *fazi* não é uma forma aceitável para o Pretérito Perfeito de *fazer*. Para chegar nesse ponto, na gramática adulta, a criança passa por um processo de aquisição chamado curva em formato de U (ERVIN, 1964, apud PETAKOVIC, [s.d.]): uma época na qual ela usa as formas corretas, *fazer – fiz*, uma época na qual ela regulariza, *fazer – fazi*, e uma época na qual ela para de regularizar e volta a usar a forma correta.

Aronoff *et al.* (2006) discutem como as crianças depreendem recursos morfológicos do seu *input* linguístico. Eles assumem que um dos primeiros passos para se adquirir o sistema morfológico é descobrir que sequências fonéticas correspondem aos morfemas. Essas sequências fonéticas podem, então, ser mais analisadas para se determinar seus privilégios gramaticais e contribuição para o significado e, aí, chegar ao domínio do sistema morfológico.

Petakovic (s.d.) afirma que, no estágio inicial do processo de aquisição, as crianças adquirem todas as formas de passado por memorização, tanto verbos regulares, *walk – walked*, como irregulares, *break – broke*. Nesse estágio elas não fazem diferença psicológica entre os dois tipos de verbos. Considerando a discussão de Aronoff *et al.* (2006), esse deve ser um período no qual a criança ainda não descobriu quais sequências fonéticas correspondem aos morfemas.

Petakovic (s.d.) explica que o início do segundo estágio, rumando para a regularização, é quando as crianças determinam que há uma regra produtiva em ação, no caso, a regra de adicionar o sufixo *-ed* para formar o passado no inglês. As crianças passam, então, a superaplicar as regras, produzindo formas como *breaked* e *goed*, por exemplo. O terceiro estágio é quando a criança toma consciência da irregularidade de alguns verbos e para de regularizar.

O estudo da aquisição de morfologia envolve o entendimento de habilidades da criança para segmentar e reconhecer morfemas. Após o segundo ano de vida essa segmentação fica mais apurada o que permite às crianças reconhecerem os morfemas da língua (TITONE, 1983; FIGUEIRA, 1995; SANTOS, 2001).

A segmentação, esse processo de aquisição de morfologia, fica evidente observando-se os erros morfológicos, como a regularização, *trazi*, ou troca de sufixos flexionais, *mexei* e *suji* (LORANDI, 2010). Lorandi (2006, 2010) não chama de erros essas formas de

regularização. Uma vez que elas co-ocorrem com a gramática adulta, ela as classifica como formas morfológicas variantes (FMVs).

### Regularização: dados para análise

Consideramos alguns erros de regularização ou generalização que encontramos nos dados anexos em Maldonade (2003), Lorandi (2006) e Figueira (2010).

#### Maldonade (2003) – dados

- (2) 2;04.26 (D) FAZER  
I: Que foi isso na sua perna?  
M: A Ana Cláudia **fazeu** dodói ni mim. (MALDONADE, 2003, p. 152)
- (3) 2;05.02 (D) FAZER  
M: Mãe, ôi o que eu **fazi** aqui.  
(M havia derrubado seu copo de leite todo sobre a mesa)  
S: Ai, ai, ai, Fica aí. Não pisa no chão. (MALDONADE, 2003, p. 152)
- (4) 2;11.15 CABER  
I: Que sapatilha que sua mãe mandô mostrá pra mim nova?  
I: Deixo vê, minha baixinha.  
M: Cabe em mim, ó.  
I: Coube?  
M: **Caibe**. (MALDONADE, 2003, p. 157)
- (5) 2;11.29 CABER  
I: Ih, Marcela! Na caixa, acho que não cabe não.  
M: Coloca.  
M: **Cabeu** viu! ... (MALDONADE, 2003, p. 158)
- (6) 2;08.14 (D) SABER  
I: Será que eu sei, Marcela, fazê uma casa de massinha? Vamo vê!  
M: Eu **sabo**.  
I: É, você é danada! Que cor você qué?  
M: Azul. (MALDONADE, 2003, p. 146)

#### Lorandi (2006) – dados<sup>2</sup>

- (7) *fazeu* (G. 2:7) (para *fez*) (LORANDI, 2006, p. 104)
- (8) *fazi* (FRA. 2:6, 2:9, 2:11) (para *fiz*) (LORANDI, 2006, p. 105)
- (9) *trazeu* (R. 3:11) (H. 2:3) (para *trouxe*) (LORANDI, 2006, p. 102 e 109)
- (10) *trazi* (B. 3:1) (para *trouxe*) (LORANDI, 2006, p. 109)

---

2 O contexto dos dados de (7) a (10) não aparecem nos trabalhos.

### **Figueira (2010) – dados**

- (11) (J coloca o fio da vitrola na tomada; ouve-se um barulhão)  
J. *Fazeu* barulho, mãe! (D – 2;7.1) (FIGUEIRA, 2010, p. 142)
- (12) (J coloca o apontador dentro da pasta)  
J. *Cabeu*, mãe.  
(a mãe duvidava que coubesse)  
M (corrigindo). Nãoé “cabeu”, é “coube”.  
J. *Coube*. (D – 3) (FIGUEIRA, 2010, p. 142)

### **Estudos de aquisição no espanhol: Clashen *et al.* (2002)**

Clashen *et al.* (2002) identificaram três tipos de erros da flexão verbal produzidos por uma criança entre 2;2 (dois anos e dois meses) e 2;8 (dois anos e oito meses) no espanhol:

i) erros morfológicos – a criança produziu um afixo ou radical regular para um verbo que requer uma forma irregular (*\*pusí* em vez de *puse – colocar*);

ii) erros de classe de conjugação – a criança flexionou incorretamente um verbo de 2ª ou 3ª conjugação de acordo com a 1ª (*\*abré* em vez de *abri – abrir*); e,

iii) superaplicação de formas de 3ª pessoa do singular nos contextos que requerem formas de 1ª ou 2ª pessoa do singular ou plural (*\*Tu presta tus monedas?*).

Eles apontam uma dissociação entre processos regulares e irregulares da flexão verbal nas produções de crianças falantes de espanhol, o que fica evidente tanto pela distribuição dos erros quanto pelo desenvolvimento de regularizações. Assumem que essas afirmações são compatíveis com a diferença apontada pelo modelo do mecanismo binário entre representações baseadas em regras (formas regulares) e baseadas na memória (formas irregulares) para palavras complexas morfológicamente.

Eles mostram que sufixos regulares e temas não marcados, não alternantes, são estendidos para verbos irregulares nos erros de flexão das crianças, mas não vice-versa. Os erros de generalização não são tão frequentes. Eles acontecem apenas quando a criança falha em usar a forma irregular. O período de regularização é precedido por um estágio sem erros e o grupo de regularizações está conectado à emergência de marcas finitas obrigatórias. Eles explicam esse processo pelo modelo do mecanismo binário de flexão que envolve dois sistemas representacionais distintos:

i) um grupo de entradas lexicais que são listadas associativamente na memória; e,

ii) um grupo de operações simbólicas ou regras para formar expressões linguísticas maiores. Ou seja, esse mecanismo no qual as flexões regular e irregular são dissociadas na gramática das crianças funciona basicamente da mesma forma que acontece na gramática adulta.

## A Morfologia Distribuída (MD)

### O modelo teórico

A MD é um dos desenvolvimentos mais recentes da Teoria Gerativa que assume que a formação de palavras e sentenças se dá no único componente gerativo da arquitetura da gramática, o componente sintático. Nesse modelo não há um componente lexical com caráter gerativo. O que, em modelos anteriores, era resolvido no léxico, na MD, é distribuído pelos outros componentes da gramática. A gramática é organizada como mostra a figura abaixo. A Estrutura Morfológica (MS, do inglês *Morphological Structure*) faz interface com a sintaxe e com a fonologia. A fonologia é vista como o componente interpretativo que realiza representações sintáticas fonologicamente.

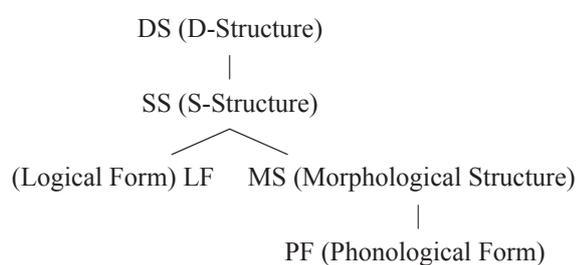


Figura 1: Arquitetura da gramática (HALLE; MARANTZ, 1993, p. 114)

Segundo Harley e Noyer (1999), há três propriedades centrais que definem essa teoria:

i) Inserção Tardia – categorias sintáticas são puramente abstratas, não têm conteúdo fonológico. A expressão fonológica dos nós terminais sintáticos é inserida apenas no mapeamento para a Forma Fonológica (PF, do inglês *Phonological Form*);

ii) Subespecificação de Itens de Vocabulário – as expressões fonológicas não precisam ser completamente especificadas para as posições sintáticas onde serão inseridas. Os Itens de Vocabulário (IVs) podem ser formas *default* inseridas quando um item mais especificado não estiver disponível; e,

iii) Estrutura Hierárquica Sintática *All The Way Down* – elementos dentro da sintaxe e da morfologia entram nos mesmos tipos de estrutura constituintes.

A sintaxe opera (*move e merge*) sobre os traços morfossintáticos selecionados do inventário disponível. Os IVs competem para inserção nos nós terminais e o item mais especificado ganha essa competição.

Halle e Marantz (1993) mostram que dois tipos de competição podem ocorrer na inserção de vocabulário: inserção livre de contexto e dependente de contexto, ou alomorfa condicionada. Os autores apresentam os IVs de passado em (13):

- (13) [+past] ↔ -Ø / [+forte] \_\_\_\_  
      [+past] ↔ /-t/ / [-forte] \_\_\_\_  
      [+past] ↔ /-d/ (HALLE; MARANTZ, 1993, p. 123)

Nessa análise, os alomorfes de passado *-t* e *Ø* têm precedência sobre *-d*, pois, *-t* e *Ø* impõem condições sobre a raiz verbal, enquanto *-d* é a forma inserida nos demais casos.

## Siddiqi (2009)

Siddiqi (2009) adota o aparato teórico da DM, mas toma uma perspectiva um pouco diferente em relação às operações em MS e *spell-out*. Ele adiciona uma condição de economia na gramática: Exponência Minimizada – a derivação mais econômica será aquela que realiza maximamente todos os traços formais da derivação com o menor número de morfemas.

(14) a. Entrada de vocabulário para *speak*<sup>3</sup>

√SPEAK → speak  
[v] /spik/

b. Entrada de vocabulário para *spoke*

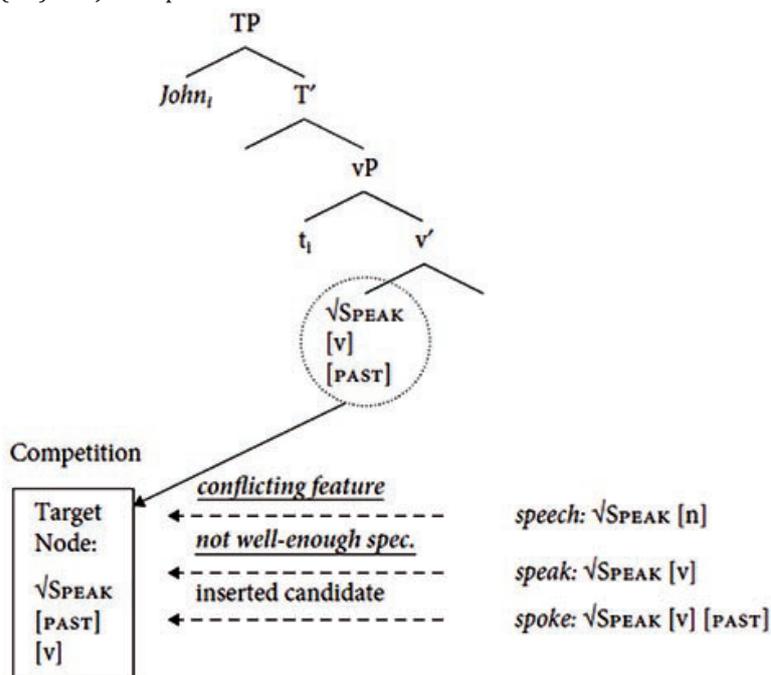
√SPEAK → spoke  
[v] /spowk/  
[past]

c. Entrada de vocabulário para *speech*

√SPEAK → speech  
[n] /spit/

(SIDDIQI, 2009, p. 39)

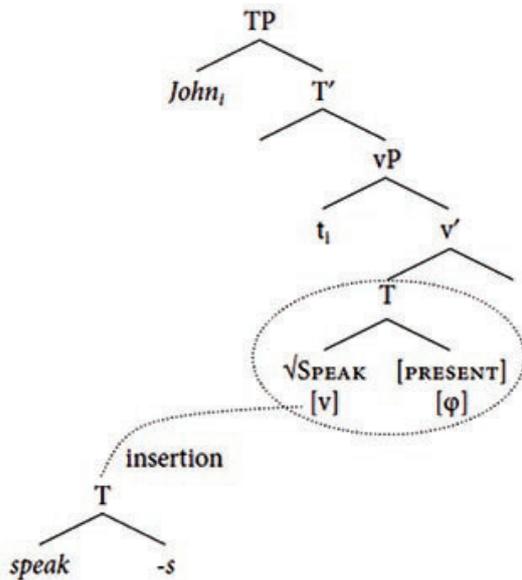
(15) John spoke.



(SIDDIQI, 2009, p. 39)

<sup>3</sup> Tradução nossa.

(16) John speaks.



(SIDDIQI, 2009, p. 48)

Se a raiz  $\sqrt{\text{SPEAK}}$  pode deixar de fundir com [PRESENT] para resultar em *speaks*, prevemos que ela também pode falhar e não fundir com [PAST] e resultar em *\*speaked*. Assim, fusão deve ser bloqueada para disparar morfologia regular, mas deve ser permitida em casos de supleção.

Siddiqi (2009) propõe uma especificação  $\neg$  para indicar que um dado IV não pode aparecer na presença de certos traços.

(17) Entrada de vocabulário para *speak*<sup>4</sup>

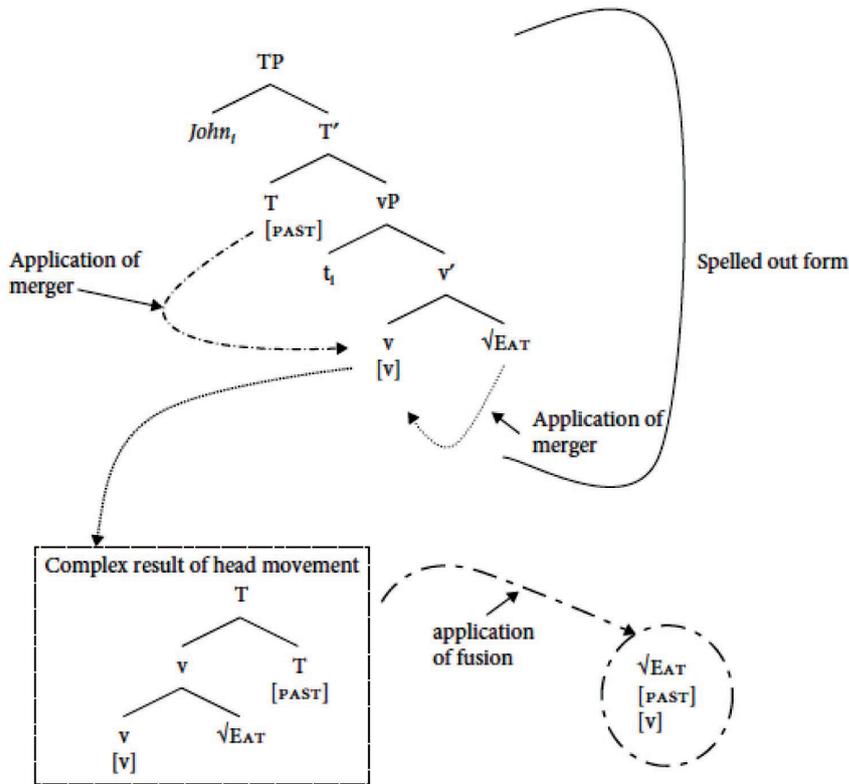
$\sqrt{\text{SPEAK}} \rightarrow \text{speak}$   
 [v] /spik/  
 $\neg$  [3sg]

(SIDDIQI, 2009, p. 51)

Essa análise leva a crer que há duas derivações diferentes competindo: uma na qual o traço [3sg] funde, e uma onde ele não funde. Deve haver essas duas derivações, uma que funde, outra que não funde, para todas as sentenças.

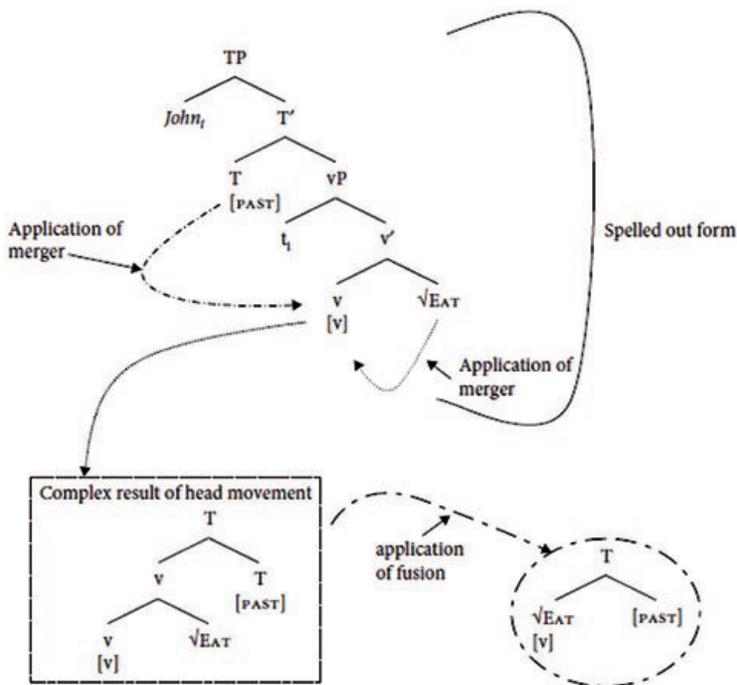
4 Tradução nossa.

(18) John ate.



(SIDDIQI, 2009, p. 52)

(19) \*John eated.



(SIDDIQI, 2009, p. 53)

Exponência Minimizada é uma restrição usada para a escolha da derivação mais econômica, que vai convergir entre as duas derivações dos mesmos traços formais.

## Análise do processo de aquisição de verbos irregulares

Fazemos algumas considerações a seguir que levarão às respostas para as questões propostas no início deste trabalho. Em (20), mostramos a definição dos IVs que realizam os traços morfossintáticos do paradigma flexional do PB, conforme Bassani e Lunguinho (2011) para o Indicativo (Presente, Pretérito Perfeito e Imperfeito):

(20) O paradigma flexional do verbo português: Itens de Vocabulário

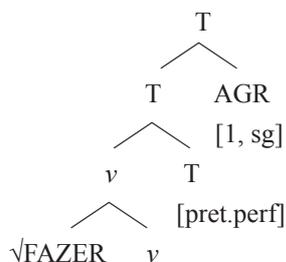
- /a/ ↔ [c1]
- /e/ ↔ [c2]
- /i/ ↔ [c3]
- /va/ ↔ [pret.imp] / [c1]
- /ia/ ↔ [pret.imp] / [c2] ou [c3]
- /u/ ↔ [pret.perf, sg]
- /ran/ ↔ [pret.perf, pl]
- /o/ ↔ [pres, 1, sg]
- /i/ ↔ [1, sg]
- /mos/ ↔ [1, pl]
- /N/ ↔ [pl]

(BASSANI; LUNGUINHO, 2011, p. 11)

Considerando a proposta de Siddiqi (2009) e Bassani e Lunguinho (2011), em um nó terminal com o feixe de traços [1, sg], o item de vocabulário mais especificado, sem nenhum traço divergente daqueles requeridos no nó terminal, será inserido, como em (21):

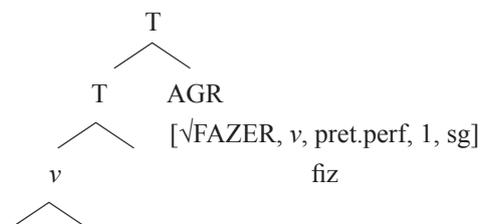
(21) Verbo antes da fusão

a) fazer



Verbo após a fusão

b) fiz



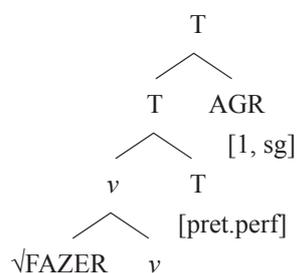
Esse processo respeita Exponência Minimizada, a derivação mais econômica é aquela que é realizada com menos IVs (22b).

## Conclusões

Considerando Siddiqi (2009), que defende que a raiz, o núcleo *v* e o traço de passado se fundem formando um nó funcional complexo, e o IV entra em um lugar só, podemos explicar o processo de regularização pelo qual a criança passa como o caso em que a criança ainda não forma esse nó funcional complexo, então, a forma regularizada: a raiz de *fazer/trazer*, o núcleo *v* e o IV que se realiza no passado *-i* ou *-u*, são inseridos separadamente.

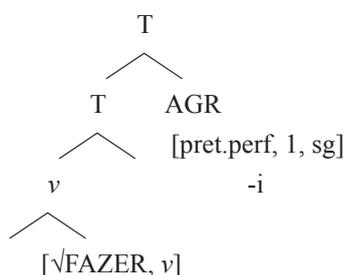
(22) Verbo antes da fusão

a) fazer



Verbo após a fusão

b) \*fazi



Como vemos em (22b), o que acontece nos casos de generalização é que nesse momento a criança ainda não domina totalmente a restrição de Exponência Minimizada.

Assim, podemos dizer que, enquanto teorias como a proposta por Clashen *et al.* (2002) precisam lançar mão de um mecanismo binário para dar conta dessas formações, na MD conseguimos explicá-las utilizando apenas um componente gerativo da gramática, a sintaxe.

## REFERÊNCIAS

ARONOFF, Mark; GIRALT, Nuria; MINTZ, Toben H. Stochastic Approaches to Morphology Acquisition. In: KLEE, Carol A.; FACE, Timothy L. (Ed.). *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Somerville, MA: Cascadilla Proceedings Project, 2006. p. 110-121.

BASSANI, Indaiá Santana; LUNGUINHO, Marcus Vinicius. Revisitando a flexão verbal do português à luz da Morfologia Distribuída: um estudo do presente, pretérito imperfeito e pretérito perfeito do indicativo. *ReVEL*, edição especial, n. 5, 2011. Disponível em: <<http://www.revel.inf.br>>. Acesso em: 04 jan. 2012.

CÂMARA JR., Joaquim Mattoso. Para o estudo descritivo dos verbos irregulares. In: *Dispersos de J. Mattoso Camara Jr.* 3. ed. Rio de Janeiro: Lucerna, 2004. [1972]. p. 131-146.

\_\_\_\_\_. A flexão verbal portuguesa: o padrão geral. In: \_\_\_\_\_. *Estrutura da língua portuguesa*. 20. ed. Petrópolis: Vozes, 1991. [1970]. p. 104-110.

CLASHEN, Harald; AVELEDO, Fraibet; ROCA, Iggy. The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language*, v. 29, issue 03, Cambridge, Cambridge University Press, p. 591-622, 2002.

ERVIN, Susan. Imitation and structural change in children's language. In: LENNEBERG, Eric H. (Ed.). *New directions in the study of language*. Cambridge, Mass.: MIT Press, 1964. p. 163-180.

FIGUEIRA, Rosa Attié. O que a investigação sobre o erro na fala da criança deve a Saussure. *Caderno de Estudos Linguísticos*, Campinas, v. 52, n. 1, p. 115-143, 2010.

\_\_\_\_\_. A palavra divergente. Previsibilidade e imprevisibilidade nas inovações lexicais da fala de duas crianças. *Trabalhos em Lingüística Aplicada*, Campinas, n. 26, p. 49-80, jul./dez. 1995.

HALLE, Morris; MARANTZ, Alec. Distributed Morphology and the Pieces of Inflection. In: HALE, Kenneth Locke; KEYSER, Samuel Jay (Ed.). *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press, 1993. p. 111-176.

HARLEY, Heidi; NOYER, Rolf. State-of-the-Article: Distributed Morphology. *GLOT*, v. 4.4, p. 3-9, April 1999.

KIRN, John. Acquisition/Development of Morphology: Specific Aspects of Noun- and Verb-Phrases. *Developmental Psychology*. 1986. Disponível em: <<http://www.hoboes.com/FireBlade/Politics/Texas/Morphological%20Development/>>. Acesso em: 04 jan. 2012.

LORANDI, Aline. Formas morfológicas variantes na aquisição da morfologia: evidências da sensibilidade da criança à gramática da língua. *Letrônica*, Porto Alegre, v. 3, n. 1, p. 81, 2010.

\_\_\_\_\_. *Formas morfológicas variantes na gramática infantil: um estudo à luz da Teoria da Otimidade*. 185 p. 2006. Dissertação (Mestrado em Linguística Aplicada) – PUC-RS, Porto Alegre.

MALDONADE, I. R. *Erros na aquisição da flexão verbal: uma análise interacionista*. 167 p. 2003. Tese (Doutorado em Linguística) – Unicamp, Campinas.

PETAKOVIC, Elizabeta. The Logical Problem of Recovery from Overregularization. [s.d.]. Disponível em: <[http://www.elizabeta.net/NOTES/notes\\_language%20acquisition.html](http://www.elizabeta.net/NOTES/notes_language%20acquisition.html)>. Acesso em: 04 jan. 2012.

SANTOS, Raquel Santana. *A aquisição do acento primário no português brasileiro*. 316 p. 2001. Tese (Doutorado em Linguística) – Unicamp, Campinas.

SIDDIQI, Daniel. *Syntax within the word: economy, allomorphy, and argument selection in Distributed Morphology*. Amsterdam: John Benjamins, 2009.

TITONE, R. *Psicolinguística aplicada: introdução psicológica à didática das línguas*. São Paulo: Summus Editorial, 1983. 219 p.

# Trajecória da constelação sufixal: *-ismo, -ista, -ístico(a)* e *-ística* – desde a origem grega às línguas modernas

(Sufixal constellation trajectory: *-ismo, -ista, -ístico(a)* e *-ística* – from its Greek origin until Modern Languages )

Nilsa Areán-García<sup>1</sup>

<sup>1</sup>Grupo de Morfologia Histórica do Português (GMHP) – Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH) – Universidade de São Paulo (USP)

nilsa.arean@gmail.com, nilsa.garcia@usp.br

**Abstract:** The present paper, a result of researches carried out by the GHMP (GMHP in Portuguese) – Group of Historical Morphology of the Portuguese Language – aims at studying both the synchronic and diachronic aspects of the constellation formed with the suffixes *-ismo, -ista, -ístico(a)* and *-ística*. For such, initially we work with its Greco-Latin genesis and, after that, we investigate its evolution into modern languages. With that, we evaluate the importance of languages of culture and the importance of processes of translations in the international dissemination of the suffixes in question.

**Keywords:** Historical morphology; suffixation; suffixal constellation.

**Resumo:** O presente artigo, resultado das pesquisas do GMHP – Grupo de Morfologia Histórica do Português –, visa a estudar os aspectos sincrônicos e diacrônicos da constelação formada com os sufixos *-ismo, -ista, -ístico(a)* e *-ística*. Para tanto, inicialmente averigua-se a sua gênese greco-latina e, posteriormente, investiga-se sua evolução nas línguas modernas. Desse modo, pode-se avaliar a importância das línguas de cultura e dos procedimentos de tradução na disseminação internacional dos sufixos em questão.

**Palavras-chave:** Morfologia histórica; sufixação; constelação sufixal.

## Introdução

De acordo com a metodologia desenvolvida pelo Grupo de Morfologia Histórica do Português – GMHP – e descrita em Viaro (2011), inicialmente foi feito um estudo das origens da constelação sufixal no grego e no latim. Para tanto, faz-se necessária uma pesquisa em obras históricas e etimológicas sobre os elementos estudados nestas duas línguas clássicas, com o qual se observa que *-ιστικός, -ιστής* e *-ισμός* não são sufixos no grego, mas terminações concatenadas, respectivamente aos sufixos *-ικός, -τής* e *-μός* quando da derivação de palavras provenientes de formas verbais geralmente terminadas em *-ίζω*, representado morfologicamente pela partícula *-ισ-* que compõe as terminações. Embora não encontremos menções a *-ística* ou a *-ιστική* nas obras teóricas consultadas, pôde-se verificar a sua efetiva ocorrência no grego na formação de palavras substantivas, bem como a sua ocorrência na constelação de terminações: *-ίζω, -ισμός, -ιστής, -ιστικός, -ιστική*. Não obstante as terminações se mostrem produtivas no grego abrangendo vários campos semânticos, ao serem transpostas para o latim, perdem sua grande produtividade e se especializam no campo semântico filosófico e religioso, como também, em muitos casos, perdem a sua associação verbal, tornando-se denominais nessa transposição. Em decorrência da perda da associação verbal na transposição à língua latina, a partícula *-ισ-*

cristaliza-se junto aos sufixos gregos, originando, assim, os sufixos da constelação ora estudada.

Sabendo-se que não são exclusivos das línguas portuguesa, pois têm como origem as terminações gregas *-ισμός*, *-ιστής*, *-ιστικός* e *-ιστική* que foram incorporadas à língua latina sob a forma *-ismus*, *-istēs*, *-isticus* e *-istica*; procurou-se também explorar a internacionalidade dos sufixos *-ismo*, *-ista*, *-ístico(a)* e *-ística*, nas suas várias formas cognatas e, por meio de um estudo de suas formações, verificar o comportamento e a expansão internacional dos sufixos e das constelações com eles formadas. Constatou-se que a grande maioria das associações ocorrem entre *-ismo* e *-ista*, por isso é a constelação mais referida, mas se constata também a ocorrência da constelação ternária composta pelos sufixos: *-ismo*, *-ista* e *-ístico(a)* com menor frequência. Além disso, ocorre a constelação quaternária, não mencionada pelas obras teóricas de apoio consultadas, formada por: *-ismo*, *-ista*, *-ístico(a)* e *-ística*. Entretanto, praticamente não se verifica a associação verbal, tal qual no grego. Nota-se também que a retomada da produtividade dos sufixos não se dá pela sua transposição ao latim clássico, mas por uma retomada do grego nas formações de novos vocábulos no âmbito formal e acadêmico e pela sua veiculação por meio de línguas de influência cultural.

Com esses dados, então, consegue-se distinguir, além do processo genealógico desde o grego às línguas modernas, também algumas características que definem as constelações sufixais *-ismo*, *-ista*, *-ístico(a)* e *-ística*, bem como a importância das traduções na sua incorporação lexical, da produtividade e de sua disseminação internacional.

### Constelação sufixal

Convém lembrar que o termo *constelação*, ora usado, significa, na sua associação por extensão de sentido, segundo Houaiss (2001), entre outras acepções, “conjunto de elementos que formam um todo coerente, ligados por algo em comum”. Tal acepção fornece subsídio ao uso do vocábulo nas mais diversas áreas, não mais apenas restrito à astronomia e à astrologia, tampouco restrito apenas ao significado de “um grupo de estrelas próximas ou afins”. Nesse sentido, também a área da linguística se apoderou do termo, conforme Houaiss (2001), com a designação de “grupo de palavras ligadas por associação semântica”. Embora não esteja definido no dicionário consultado, consideramos que o termo *constelação sufixal*, também usado no seu verbete do sufixo *-ismo*, conforme as designações anteriores, seja a aplicação da definição linguística de *constelação* ao vocábulo *sufixo*, que neste estudo é tido como um morfema provido, além de função gramatical, também de significado; de maneira que o termo designa, por especialização do sentido linguístico, *um grupo de sufixos ligados pela associação semântica de suas funções*.

Assim, uma constelação sufixal aplicada a uma dada base formará uma família de palavras, derivadas com os sufixos pertencentes ao grupo, e que constituirá uma instância da constelação sufixal. Por exemplo: *urbanismo*, *urbanista*, *urbanístico(a)* e *urbanística* é uma família de palavras que constitui uma instância da constelação sufixal: *-ismo*, *-ista*, *-ístico(a)* e *-ística*; tomando como base a palavra *urbano*.

## A origem da constelação no grego clássico

Para se conhecer a origem da constelação no grego clássico, foram localizadas e estudadas todas as palavras terminadas em *-ισμός*, *-ιστής*, *-ιστικός* e *-ιστική*, encontradas no dicionário etimológico Chantraine (1968), bem como os verbos terminados em *-ίζω* e suas famílias de palavras criadas por derivação sufixal, além de consultas a obras gramaticais e etimológicas sobre o assunto.

Segundo Casevitz (1985, p. 21-30), no grego arcaico, considerado pelo autor a época de Homero e de Hesíodo, já podem ser encontrados verbos em *-ίζω*, muitos dos quais tomam como suas raízes palavras concretas pertencentes ao período micênico e ao domínio agrícola, evoluindo, no período arcaico, em direção ao domínio da construção e ao urbano, bem como ao militar e da guerra. Convém lembrar ainda que o sufixo *-τής*, na formação de *nomina agentis*, é comum na escrita de Homero (período arcaico da língua grega), suplantando a forma jônico-ática mais antiga e concorrente, *-τηρ*, e contribuindo para as associações entre *-ίζω* e *-ιστής*. De acordo com Juret (1937, p. 148), os verbos gregos formados a partir de *-ίζω* podem denotar ruídos, em geral, onomatopéicos ou sons musicais, muitas vezes provenientes do instrumento tomado como base; em outros casos são verbos que podem indicar ‘imitadores de’, em geral, tomando um gentílico como base, e ainda, verbos que estão associados a ‘ser seguidores de’ ou ‘ser partidários de’. O sufixo *-ίζω* é amplamente conhecido por formar derivados verbais que denotam o senso factivo e/ou iterativo, por exemplo, *ἀκοντίζω* (‘atirar flechas’) e *ὄνειδίζω* (‘ultrajar, afrontar, insultar’), segundo Fleury (1947, p. 175). Conforme Pharies (2002, p. 373), *-ίζω* é extremamente produtivo na língua grega, tendo formado centenas de verbos.

Neste sentido, conforme Juret (1937, p. 83), *-ισμς* e *-ισμα* são terminações designadoras de formações *nomina actionis* a partir de verbos terminados em *-ίζω*, considerando que os derivados formados com a primeira indicam uma ação em curso, ao passo que os formados com a segunda indicam uma ação já terminada, ou o resultado desta. Porém, segundo Fleury (1947, p. 32), o sufixo *-μα*, ao contrário de *-μς*, não é produtivo no grego, aparecendo associado a um pequeno número de verbos e em um número isolado de palavras.

Constatou-se, ainda, em Chantraine (1968), que normalmente a ocorrência do verbo é mais antiga que as ocorrências dos nomes de ação, agentes, adjetivos e aos substantivos terminados em *-ιστική*; sugerindo que estes sejam derivações posteriores. Outrossim, nota-se que em geral a abonação dos nomes de ação é mais antiga sugerindo que, normalmente, surjam com anterioridade aos nomes de agentes, que, por sua vez, geralmente, são abonados com anterioridade aos nomes adjetivais e estes mais antigos que os substantivos femininos terminados em *-ιστική*.

Etimologicamente pode-se considerar, então, ao observar os dados obtidos em Chantraine (1968) com suas respectivas datações, que as constelações formadas com as terminações estudadas começam, no grego, em geral, por uma formação verbal terminada em *-ίζω*, a partir da qual se formam palavras funcionalmente *nomina actionis* terminadas em *-ισμα*, que indicam o resultado da ação, e terminadas em *-ισμός*, que indicam o processo em curso; além das formações *nomina agentis* terminadas em *-ιστηρ*, cuja concorrência com *-ιστής* faz a terminação se especializar na formação de *nomina instrumenti*. Nessa linha, partindo da formação *nomina agentis* em *-ιστής* acrescida do sufixo *-ικός*, formam-se as palavras adjetivas terminadas em *-ιστικός*, que, desse modo, se mostra como a

concatenação de *-ιστής* com o sufixo *-ικός*. Somente, então, pela conversão funcional de palavras adjetivas formadas com *-ιστικός* no gênero feminino em substantivas, formam as terminadas em *-ιστική*, que, como terminação, denota a conversão substantiva e feminina de *-ιστικός*. O processo etimológico de formação das constelações na língua grega está ilustrado na Figura 1.

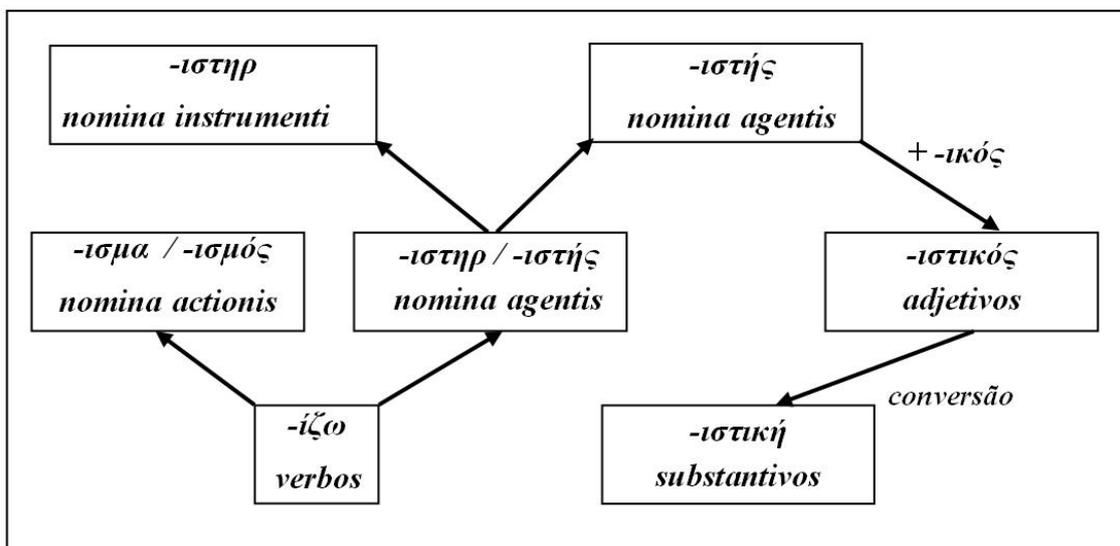


Figura 1. Processo etimológico de formação das constelações no grego

*Grosso modo*, pode-se admitir que o sufixo *-ίζω* indica uma ação factiva ou iterativa, a terminação *-ισμα* indica o produto da ação, *-ισμός* indica o processo da ação em curso, *-ιστηρ* indica o instrumento requerido pela ação ou processo, *-ιστής* indica o agente da ação ou processo, a terminação *-ιστικός* designa o adjetivo relacionado à ação ou ao processo e a terminação *-ιστική* indica a ciência ou arte que envolve a ação ou processo. Convém destacar que esta acepção semântica adquirida pela terminação deve-se ao frequente uso de suas formações adjetivas ao lado das palavras femininas *τέχνη* (‘método, técnica, arte, prática, competência’) e *έπιστήμη* (‘ciência, habilidade, conhecimento, saber, estudo’) com a posterior elisão da forma substantiva, promovendo a substantivação do adjetivo feminino. Por exemplo, *άγωνιστική τέχνη* passou a ser escrita já no próprio grego como *άγωνιστική* nos contextos em que o substantivo *τέχνη* está claramente entendido; analogamente *λογιστική έπιστήμη* passou a ser referida como *λογιστική*.

A título ilustrativo, segue-se uma instância da constelação formada com as terminações *-ίζω*, *-ισμα*, *-ισμός*, *-ιστής*, *-ιστικός*, *-ιστική*: *άκοντιζω* – ‘lançar o dardo’; *άκοντισμα* – ‘alcance do dardo lançado’; *άκοντισμός* – ‘lançamento do dardo’; *άκοντιστής* – ‘lançador de dardos’; *άκοντιστικός* – ‘relativo ao lançamento do dardo’; *άκοντιστική* – ‘a arte de lançar o dardo’.

Pode-se observar uma instância de constelação formada com *-ίζω*, *-ισμα*, *-ισμός*, *-ιστηρ*, *-ιστής*, *-ιστικός*, *-ιστική*: *όρίζω* ‘delimitar, demarcar fronteiras, limites’; *όρισμα* ‘fronteira, limite’; *όρισμός* ‘delimitação, definição’; *όριστηρ* ‘marco, pedra que se usa para demarcar limites territoriais’; *όριστής* ‘demarcador de fronteiras; aquele que determina, estabelece, juiz’; *όριστικός* ‘que/quem serve para delimitar, para definir’; *όριστική* ‘a arte de delimitar, na gramática é o modo indicativo’.

Com os exemplos anteriormente elucidados, nota-se que a derivação em uma constelação nem sempre é previsível, concomitantemente e nem sempre segue apenas as regras do processo derivativo, pois pode haver as acomodações fonéticas, as necessidades dos falantes em prescindir ou não de determinados termos e suas funções, há também as formas de terminações concorrentes que se podem cristalizar em determinado dialeto em detrimento de outras. Ademais, acredita-se que, no caso da língua grega, muitos textos perdidos teriam de maneira concreta, possivelmente, como confirmar a existência de palavras consideradas hipotéticas, formadas com as terminações aqui analisadas.

Embora a Figura 1 esquematize o processo etimológico de formação das constelações com as terminações estudadas, a partir do momento em que cada terminação se torna produtiva passa a haver uma associação entre elas sem que as formações sigam exatamente os mesmos passos do processo. Assim, nota-se que pode não haver a existência concomitante de todas as formas previstas, mas apenas de algumas delas devido às associações estabelecidas entre as terminações e/ou a produtividade própria de cada terminação. A título de justificativa, as terminações *-ισμα* e *-ιστηρ* mostraram-se com baixíssima produtividade, menos de 1%, nos dados coletados de Chantraine (1968), por isso poucas são as instâncias de constelações que as contemplam. De qualquer modo, no grego, a formação das constelações está geralmente centrada nas formas verbais, como sua origem. A título de ilustração, a figura 2 esquematiza, no grego, as relações entre as terminações estudadas nas constelações, mantendo o verbo como o centro destas nas relações.

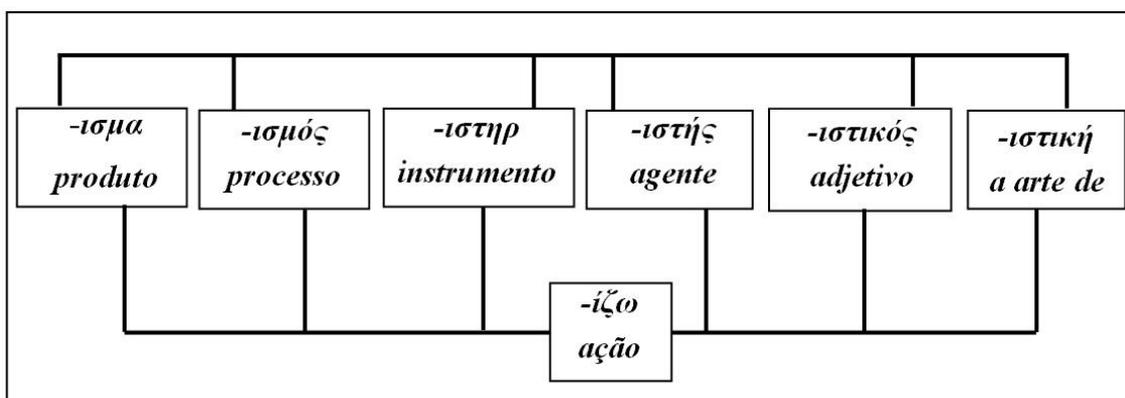


Figura 2. Esquema das relações entre as terminações estudadas nas constelações

Com o total de 247 palavras obtidas em Chantraine (1968) pode-se notar que, quanto à forma, no grego, a constelação mais ocorrente é a instância composta pelo verbo (*-ίζω*), nome de ação verbal (*-ισμός*) e nome de agente (*-ιστής*), apresentando 35% do total de ocorrências. A segunda constelação que mais se evidencia é a quaternária, apresentando 16% do total, cujas instâncias são formadas pelo verbo (*-ίζω*), nome de ação verbal (*-ισμός*), nome de agente (*-ιστής*) e nome adjetival (*-ιστικός*). Notou-se ainda que as instâncias de constelações formadas com nomes substantivos terminados em *-ιστική* são pouco representativas e ao todo somam 6% do total. Por outro lado, as com a forma adjetival em *-ιστικός* mostram-se bastante representativas e apresentam um total de 35% das ocorrências.

## Ocorrências das constelações no latim

De modo análogo ao que foi feito no grego, no latim foram levantadas todas as palavras terminadas em *-ismus*, *-istēs*, *-istīcus* e *-istīca*, encontradas nos dicionários Gaffiot (1934) e Munguía (2010), bem como os verbos terminados em *-izō* e suas famílias de palavras criadas por derivação sufixal. Ao consultar a obras gramaticais latinas, nada se pode encontrar com relação aos sufixos estudados. Acredita-se que tal fato se deva ao pouco uso feito dos afixos estudados pelos autores clássicos, cujos textos têm servido como base às gramáticas latinas.

Assim, com os dados coletados, obteve-se um total de 79 palavras, 25 delas totalmente isoladas, ou seja, sem pertencer a alguma constelação. Assim, nem todos os tipos de combinações com os sufixos estudados ocorrem nas constelações do latim, havendo uma grande preponderância, 50%, nas ocorrências isoladas das formações com os afixos, e 40% das ocorrências apresentando associação com verbo. Convém notar que não foram encontradas palavras com o sufixo *-isma* ou *-istēr* e foi encontrada apenas uma palavra isolada com o sufixo *-istīca*. Os dados referentes às constelações estão dispostos na Tabela 1.

**Tabela 1: Tipos de constelações em Gaffiot (1934) e Munguía (2012)**

Tipo de constelação	Ocorrências	Porcentagem
<i>-ismus + -izō</i>	9	18%
<i>-istēs + -istīcus</i>	3	6%
<i>-istēs + -izō</i>	7	14%
<i>-istēs + -ismus</i>	2	4%
<i>-ismus + -izō + -istēs</i>	3	6%
<i>-ismus + -izō + -istīcus</i>	1	2%
Palavras isoladas	25	50%

Ao observar os dados pôde-se constatar que a maioria das ocorrências no grego se dá sem vinculação com o verbo. Assim, na transposição das terminações gregas para o latim, a partícula *-is-*, no grego *-ισ-*, deixa de significar a deverbalidade na derivação e passa a ser incorporada aos sufixos: *-ismus*, *-istēs* e *-istīcus*. Ainda assim, pode-se notar que *-ismus* indica um processo em curso e/ou um ideal filosófico-religioso, *-istēs* indica um agente ou seguidor de ideal, *-istīcus* indica um adjetivo relativo ao agente ou ao ideal e *-izō* indica uma ação. Verificou-se também que a maioria das palavras encontradas no latim são transliterações gregas do campo religioso e filosófico, usadas como termos e cuja terminação enfatiza sua proveniência.

## Comparação das ocorrências entre o grego e o latim

Pode-se notar que há, em números absolutos, muito mais ocorrências de *-ισμός/-ismus*, *-ιστής/-istēs* e *-ιστικός/-istīcus* no grego que no latim. Além disso, a maioria das palavras encontradas no latim, segundo Munguía (2010) e Gaffiot (1934), é importada diretamente da língua grega. Desse modo, *-ισμός/-ismus*, *-ιστής/-istēs* e *-ιστικός/-istīcus* apresentam uma boa produtividade na língua grega, porém não na latina, deixando fortemente marcado o traço de sua origem helênica nas ocorrências desta. Ademais, nota-se que, embora timidamente, no grego aparece a produtividade de *-ιστική/-istīca* na formação de nomes substantivos em algumas ocorrências, ao contrário do latim, no qual apenas

se encontra uma única palavra com tal terminação em Gaffiot (1934, p. 919), associada pelo autor à terminação *-ιστικός*, notadamente importada da língua grega, e portanto, não seja quiçá reconhecida como sufixo.

Quanto ao estudo das constelações, notou-se que no grego a maioria de suas ocorrências está vinculada a um verbo, ao passo que a preponderância latina é para as ocorrências isoladas. Acredita-se, então, que inicialmente a importação do grego ao latim se deu palavra a palavra de acordo com a necessidade de uso, de modo que nem sempre foi necessária a importação de todas as formações de uma instância de constelação. Somado a isto, pode-se conjecturar que o fato de a língua latina denotar o verbo e seus objetos em palavras separadas, tenha propiciado que a partir do latim os sufixos estudados comecem a ser denominais e não mais deverbais, amalgamando, desse modo, a partícula *-is-* aos sufixos. Assim, generalizando o parecer de Pharies (2002, p. 356), pode-se afirmar que o latim os adota como sufixos, porém deixa de seguir a regra gramatical grega nas suas poucas, porém, novas formações.

### A internacionalização das constelações estudadas

Sabe-se que os sufixos estudados são internacionais e formam palavras de ampla circulação e que são importadas até mesmo por línguas nas quais não há a produtividade destes, por exemplo, no japonês. Para justificar a asserção da internacionalização de ditos sufixos, foi feito um levantamento da existência de formas cognatas em diversas línguas, não apenas as românicas, após pesquisa e consultas a obras linguísticas, gramáticas, dicionários e a especialistas em cada uma das línguas consideradas. Com os dados obtidos elaboraram-se as tabelas exemplificativas a seguir com as formas cognatas dos sufixos *-ismo*, *-ista*, *-ístico(a)* e *-ística* encontradas, a título ilustrativo, nas línguas românicas: romeno, italiano, francês, catalão, valenciano, castelhano, galego e português; nas línguas germânicas: inglês, alemão, sueco, holandês e norueguês; nas eslavas: tcheco, polonês, eslovaco, búlgaro e russo; nas bálticas: letão e lituano; em outras línguas europeias: *euskera*, húngaro, albanês e turco; nas línguas semitas: hebraico e árabe, bem como nas línguas orientais: japonês e chinês, conforme está exposto nas tabelas 2 a 8, a seguir.

**Tabela 2: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, nas línguas românicas**

romeno	italiano	francês	catalão, valenciano	castelhano galego, português
<i>-ism</i>	<i>-ismo</i>	<i>-isme</i>	<i>-isme</i>	<i>-ismo</i>
<i>-ist(ă)</i>	<i>-ista</i>	<i>-iste</i>	<i>-ista/-iste</i>	<i>-ista</i>
<i>-istic(ă)</i>	<i>-ístico(a)</i>	<i>-istique</i>	<i>-ístic</i>	<i>-ístico(a)</i>
<i>-istică</i>	<i>-ística</i>	<i>-istique</i>	<i>-ística</i>	<i>-ística</i>

**Tabela 3: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, nas línguas germânicas**

inglês	alemão	sueco	holandês	norueguês
<i>-ism</i>	<i>-ismus</i>	<i>-ism</i>	<i>-isme</i>	<i>-isme</i>
<i>-ist</i>	<i>-ist(in)</i>	<i>-ist</i>	<i>-ist/-iste</i>	<i>-ist</i>
<i>-istic</i>	<i>-istisch</i>	<i>-istika</i>	<i>-istische</i>	<i>-istik(e)</i>
<i>-istics</i>	<i>-istik</i>	<i>-istika</i>	<i>-istiek</i>	<i>-istik</i>

**Tabela 4: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, nas línguas eslavas**

tcheco	polonês	eslovaco	búlgaro	russo
<i>-ismus</i>	<i>-izm</i>	<i>-izmus</i>	<i>-изъм</i> [izəm]	<i>-изм</i> [izm]
<i>-ista(ka)</i>	<i>-ista(ka)</i>	<i>-ista(ka)</i>	<i>-ист(ка)</i> [ist]	<i>-ист(ка)</i> [ist]
<i>-isticý(é)</i>	<i>-istyczne(ych)</i>	<i>-isticý(é)</i>	<i>-истически</i> [isti]eki]	<i>-истический</i> [isti]ekii]
<i>-istika</i>	<i>-istyka</i>	<i>-istika</i>	<i>-истика</i> [istika]	<i>-истика</i> [istika]

**Tabela 5: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, nas línguas bálticas**

letão	lituano
<i>-isms</i>	<i>-izmas</i>
<i>-ists</i>	<i>-istas</i>
<i>-istikas</i>	<i>-istinis</i>
<i>-istika</i>	<i>-istika</i>

**Tabela 6: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, em outras línguas europeias**

euskera	húngaro	albanês	turco
<i>-ismo</i>	<i>-izmus</i>	<i>-izëm</i>	(não há)
<i>-ista</i>	<i>-ista/-ist</i>	<i>-ist</i>	
<i>-ístico(a)</i>	<i>-isztikus</i>	<i>-istik(e)</i>	
<i>-ística</i>	<i>-isztika</i>	<i>-istikë</i>	

**Tabela 7: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, em línguas semitas**

hebraico	árabe
<i>/ism/</i>	(não há)
<i>/izm/</i>	
<i>/ist/</i>	
<i>/isti/</i>	
<i>/istikah</i>	

**Tabela 8: Cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística*, em línguas orientais**

japonês	chinês
<i>/isumo/</i>	(não há)
<i>/isuto/</i>	

Não obstante sejam internacionais, nota-se que em algumas línguas os sufixos não apresentam formas cognatas. É o caso do turco, do árabe e do chinês que, em geral, apresentam poucos ou nenhum sufixo derivacional. Nas línguas em que aparecem, porém, nem sempre estes apresentam produtividade própria, muitas vezes, surgem apenas na formação de palavras adquiridas como empréstimos adaptados foneticamente. A título de ilustração, no japonês as palavras */pianisuto/*, */nudisuto/* e */paulisuto/* são empréstimos das palavras ocidentais *pianista*, *nudista* e *paulista*, adaptados foneticamente à língua japonesa, mas sem que */isuto/* seja produtivo, pois a língua não faz uso de sufixos nominais, e, portanto, não é reconhecido pelos falantes como tal. Nota-se também que, no japonês, não foram encontradas, nas consultas a dicionários, palavras formadas com cognatos de *-ístico(a)* e *-ística*. Além do japonês, no hebraico, em euskera, húngaro, albanês, letão e lituano, apesar de apresentarem palavras formadas com os cognatos dos quatro sufixos estudados, não se observa produtividade própria, apenas se verificam empréstimos.

Notam-se, ademais, que as línguas das tabelas de 2 a 6, excetuando-se o turco, apresentam constelações com as formas cognatas dos sufixos, em sua grande maioria denominais. Utilizando-se como *corpus* dicionários das línguas estudadas, pôde-se verificar que a distribuição das constelações com os quatro sufixos é bem similar nestas, de forma que a constelação formada com os cognatos de *-ismo* e *-ista* é a mais representativa com uma variação média de 60% a 75% das ocorrências dependendo da língua; a constelação ternária formada com os cognatos de *-ismo*, *-ista* e *-ístico(a)*, apresenta uma variação média de 3% a 10% das ocorrências, destacando-se a língua alemã com 20%; a constelação quaternária formada com os cognatos de *-ismo*, *-ista*, *-ístico(a)* e *-ística* apresenta uma variação média de 1% a 4% das ocorrências. Considerando-se também a forma verbal com cognatos de *-izar*, pouquíssimas constelações foram encontradas apresentando uma variação média menor que 1% nas línguas consultadas.

Justifica-se, então, por meio da análise de porcentagens, que a maioria das associações encontradas em obras teóricas da área seja entre *-ismo* e *-ista*, seguida das associações entre *-ismo*, *-ista* e *-ístico(a)* e, ainda que não seja mencionada pelas obras de apoio teórico consultadas, a associação quaternária entre *-ismo*, *-ista*, *-ístico(a)* e *-ística*, dada a pequena porcentagem de palavras formadas com *-ística*, pode explicar a ausência de menções a esse sufixo em muitos estudos consultados.

*Grosso modo*, observando-se semanticamente os dados, pode-se inferir que, para uma determinada base, a derivação com os cognatos de *-ismo* indica o sistema e/ou processo feito com ou a partir dela; a derivação com os cognatos de *-ista* indica o agente de tal sistema/processo, a derivação com *-ístico(a)* indica o adjetivo de pertinência ao sistema/processo, e a derivação com *-ística* indica a disciplina e/ou ciência que estuda o sistema/processo. Nota-se também que, ao contrário do grego, no qual as constelações partem de um verbo, a partir do latim, na maioria dos casos as constelações passam a ser denominais, e é o que geralmente ocorre nas constelações internacionais observadas, pela quase ausência da forma verbal nelas. Convém observar que, em geral, as formas verbais ocorrem internacionalmente em constelações do âmbito religioso e são empréstimos gregos, por exemplo, no português: *batizar*, *catequizar*, *evangelizar*.

## Considerações finais

Sabe-se que, embora *-ιστής* e *-ισμός* tenham sido duas terminações produtivas no grego, apresentam pouca produtividade ao serem transpostas para o latim sob a forma *-istēs* e *-ismus*. Já a terminação *-ιστικός*, que se mostra presente em vários vocábulos adjetivos do grego, importada sob a forma latina *-isticus*, apresenta muito raras ocorrências em palavras latinas. Outrossim, ainda que não se tenham encontrado menções teóricas, verificou-se no grego a existência produtiva da terminação *-ιστική* na formação de palavras substantivas que designam arte ou ciência, embora, sob a forma latina *-ística*, se tenha encontrado apenas uma ocorrência em palavra importada diretamente do grego. Pelo observado em Munguía (2010), a maioria das palavras formadas com as terminações estudadas não é usada pelos autores do latim clássico, o que justificaria, em parte, sua pequena produtividade latina na escrita.

É notório que no grego as constelações têm como ponto de partida uma designação verbal, à qual se associam. Assim, as terminações gregas *-ισμός*, *-ιστής*, *-ιστικός*

e *-ιστική* estão associadas a verbos nominativos e/ou factivos terminados em *-ίζω*, de tal forma que podem denotar ruídos, em geral, onomatopeicos ou sons musicais, muitas vezes provenientes do instrumento tomado como base, neste caso já se notando a origem da predileção dos sufixos estudados no âmbito artístico, principalmente da música. Exemplos: *κιθαρίζω* ('tocar a cítara') > *κιθαριστής* ('quem toca a cítara'), *κιθαριστικός* ('adjetivo relacionado a tocar a cítara'), *κιθαριστική* ('a arte ou técnica de tocar a cítara'). Em outros casos, são verbos que podem indicar 'imitadores de', em geral, tomando um gentílico como base, e, por generalização, acabaram por designar 'ser seguidores de' ou 'ser partidários de', indicando a gênese dos sufixos estudados na designação genérica de pertinência a grupos (dos mais variados movimentos: filosóficos, científicos, artísticos, políticos, religiosos, etc.) e seus partidários. Exemplos: *ἀτικίζω* ('imitar os atenienses', 'falar como os atenienses', 'ser do partido ateniense') > *ἀτικισμός* ('adesão ao partido ateniense'), *ἀτικιστής* ('imitador dos atenienses', 'quem fala como os atenienses', 'partidário dos atenienses'). Os verbos com esta terminação grega podem designar ainda o senso factivo e/ou iterativo, indicando ações repetitivas, por exemplo, atuando já desde o grego de maneira evidenciada, no âmbito esportivo: *ἀκοντίζω* ('lançar o dardo') > *ἀκοντισμός* ('lançamento do dardo'), *ἀκοντιστής* ('lançador de dardo'), *ἀκοντιστικός* ('relativo a lançar o dardo'), *ἀκοντιστική* ('a arte de lançar o dardo').

Observa-se, ainda, que *-ισμός*, *-ιστής*, *-ιστικός* e *-ιστική* são terminações produtivas concatenadas aos sufixos gregos *-μός*, *-τής*, *-ικός* e *-ική*, quando associadas a verbos terminados em *-ίζω*, nas respectivas funções de *nomina actionis*, *nomina agentis*, designação de adjetivos relacionais e designação de substantivos femininos que nomeiam ciência, arte e/ou técnica. Entretanto, ao serem incorporadas ao latim por importação de suas formações, percebe-se que, muitas vezes, não houve a importação da instância inteira da constelação, apenas das palavras necessárias. Nessas importações, notoriamente é evidenciada a falta do verbo, promovendo, por um lado, o início da produção denominal dos sufixos, repassada às línguas românicas, por outro, uma maior não concomitância na possível rede de associações sufixais.

Pode-se concordar também que, de fato, etimologicamente, *-ístico(a)*, o sufixo formador de adjetivos, é um cultismo proveniente da concatenação da terminação grega *nomina agentis -ιστής* e do sufixo grego *-ικός*, importado pela língua latina sob a forma pouco produtiva *-isticus*. De modo similar, *-ística*, sufixo formador de substantivos, é proveniente da conversão substantiva de adjetivos femininos formados com *-ístico(a)*, e, portanto, indiretamente também é fruto etimológico da concatenação de *-ιστής* e *-ικός*.

Observou-se que, apesar de o latim apresentar ocorrências escassas destes sufixos, nas línguas modernas observadas estes se evidenciam em número relevante. Assim, infere-se que a maioria observada dessas ocorrências internacionais não é de origem latina clássica, mas houve uma posterior retomada, por uma ou mais línguas europeias que veicularam as palavras grego-latinas e seus modelos. Nesse sentido, convém lembrar que, para Ferreiro (2001, p. 156), os sufixos *-ista* e *-ismo* tiveram uma grande produtividade nas línguas francas europeias, principalmente nos séculos XIX e XX, por meio da expansão dos movimentos culturais, uma das quais o galego, entre outras línguas, que tomam como empréstimo um número muito grande de palavras com eles formadas. Said Ali (1930, p. 19) explica que as palavras mais antigas formadas com *-ismo*, sufixo associado a *-ista*, na língua portuguesa, devem-se ao latim da Idade Média, algumas formadas de

empréstimos latinos (em geral, provenientes do grego) e outras já formadas analogicamente. De acordo com o autor, o fato inicial de popularização dessas palavras deve-se à ação da Igreja Católica, e ressalta que, ainda que tenham sido difundidas no vulgo, não apresentaram produtividade na formação de novos derivados. O autor ressalta ainda que esta disposição só veio a mudar nos séculos XVIII e XIX, sob influência dos movimentos intelectuais franceses, e então muitas palavras não só foram formadas e importadas do estrangeiro, mas também se iniciou a sua formação a partir de bases próprias do português. Como os sufixos *-ismo*, *-ista*, *-ístico(a)* e *-ística* estão associados, infere-se que, não apenas *-ismo*, mas também os demais foram veiculados, de modo análogo, por línguas francas europeias.

Conforme visto, os sufixos estudados tiveram a sua origem na língua grega, adentraram o latim e, por meio dessa língua, as línguas românicas. Entretanto, somente a partir do século XVIII começaram a ter produtividade própria nas românicas, devido à sua propagação pelo latim científico, no âmbito técnico e acadêmico, mas também em outros âmbitos por influência de línguas de cultura, tais como o francês, italiano e alemão, dentre outras. Nota-se também que há línguas veiculares específicas para determinadas áreas, por exemplo, entre outras, a influência do italiano na música e pintura, o alemão na filosofia, o francês na moda da alta costura, e da língua inglesa, a partir do século XX, nos mais variados âmbitos. Convém lembrar que o papel das traduções foi fundamental na disseminação das formas cognatas dos sufixos nas mais variadas línguas, inicialmente com palavras cultas nos âmbitos acadêmicos e culturais.

Podemos, então, concluir que na língua, à semelhança das palavras, seus constituintes também se transformam formal e semanticamente, se associam entre si, são traduzidos e inseridos por meio de línguas veiculares a línguas naturais, circulam em âmbitos específicos, são decalcados, emprestados e adaptados foneticamente a línguas diferentes da de sua origem e nas quais podem vir a ser produtivos abrangendo novos campos semânticos, formando e ampliando constelações próprias.

## REFERÊNCIAS

CASEVITZ, M. *Le vocabulaire de la colonisation en grec ancien*. Paris: Klincksieck, 1985.

CHANTRAINE, P. *Dictionnaire étymologique de la langue grecque. Histoire des mots*. Paris: Klincksieck, 1968.

*Dictionary Reference*. s.d. Disponível em: <<http://dictionary.reference.com/>>. Acesso em: jul. 2012.

FERREIRO, M. *Gramática histórica galega*. Noia: Laiovento, 2001.

FLEURY, E. *Morphologie historique de la langue grecque*. Paris: J. de Gigord, 1947.

GAFFIOT, F. *Dictionnaire Latin Français*. Paris: Hachette, 1934.

HOUAISS, A.; VILLAR, M. (Org.). *Dicionário eletrônico Houaiss da língua portuguesa*.

sa. Versão 1. Rio de Janeiro: Objetiva, 2001. CD.ROM, v. 1.0.

IECAT. *Diccionari de la llengua catalana* (DIEC). L'Institut d'Estudis Catalans. Disponível em: <<http://www.dlc.iec.cat/>>. Acesso em: jul. 2012.

JURET, A. –C. *Formation de noms et des verbes en latin et en grec*. Paris: Les Belles Lettres, 1937.

MAURO, T. de; MANCINI, M. *I Dizionario Medi Garzanti. Parole Straniere nella lingua italiana*. Milano: Garzanti Linguistica, 2003.

MUNGUÍA, S. S. *Nuevo diccionario etimológico Latín-Español y de voces derivadas*. 4. ed. Bilbao: Deusto, 2010.

PHARIES, D. *Diccionario Etimológico de los Sufijos Españoles*. Madri: Gredos, 2002.

RAE – REAL ACADEMIA ESPAÑOLA. *Diccionario de la lengua española*. 22. ed. Madri: Real Academia Española, 2001.

RIO-TORTO, G. M.; VIEIRA, A. Nomes em *-ismo* no português dos séculos XVIII a XX. *Cuadernos del Instituto de Historia de la Lengua*, n. 2, p. 115-137, 2009.

SAID ALI, M. *Gramática histórica da língua portuguesa*. Rio de Janeiro: Melhoramentos, 1930.

\_\_\_\_\_. *Dicionário Prático Russo-Português*. Moscou: Russki Jazyk, 1986.

VIARO, M.E. *Etimologia*. São Paulo: Contexto, 2011.

## **BIBLIOGRAFIA CONSULTADA**

AREÁN-GARCÍA, N. *Estudo comparativo de aspectos semânticos do sufixo -ista no português e no galego*. 2007. 463 f. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

\_\_\_\_\_. *Aspectos sincrônicos e diacrônicos do sufixo -ístico(a) no português e no galego*. 2012. 520 f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

CARBALLEIRA ANLLO, X. M. *Gran Dicionario Xerais da Lingua*. Santiago de Compostela: Xerais, 2009.

COROMINAS, J. *Breve Dicionario etimológico de la lengua castellana*. 3. ed. Madri: Gredos, 1961.

CORTELAZZO, M.; ZOLLI, P. *Il nuovo etimológico. DELI – Dizionario Etimológico della Lingua Italiana*. Segunda Edição. Bologna: Zanichelli, 2009.

CUNHA, A. G. *Dicionário Etimológico Nova Fronteira da Língua Portuguesa*. 2. ed. Rio de Janeiro: Nova Fronteira, 1997.

DUDEN. *Duden in zwölf Bänden. Das Standardwerk zur deutschen Sprachen. Das Herkunftswörterbuch. Etymologie der deutschen Sprache*. Band 7. Mannheim: Dudenverlag, 2007.

FRINTA, A. *Bulharko-Český Slovník*. Praga: Nakladatelství Československé Akademie Věd, 1959.

GIANASTACIO, V. *A presença do sufixo -ismo nas gramáticas da língua portuguesa e sua abrangência dos valores semânticos, a partir do Dicionário de Língua Portuguesa Antônio Houaiss*. 2009. 176f. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

LAPESA, R. *Historia de la lengua española*. Real Academia Española. Colección Nebrija y Bello. Madrid: Espasa Calpe, 6ª Ed., 2003.

LÁSZLO, D. *Česko-Maďarský Slovník / Cseh-Magyar Szótár*. Praga: Akadémiai Kiadó, 1960.

LE ROBERT. *Dictionnaire de la langue française*. Paris: Le Robert, 1985.

MOCANU, P. *Dictionar Român-Portuguez*. București: Științifică și enciclopedică, 1981.

MUÑOZ ARMIJO, L. *La historia de los derivados en -ismo e -ista en el español moderno*. 2010. 708 f. Tese (Doutorado em Filologia Espanhola) – Facultad de Filosofía y Letras, Universidad Autónoma de Barcelona, Bellaterra.

ONIONS, C. T.; FRIEDRICHSEN, G. W. S.; BUCHFIELD, R. W. *The Oxford Dictionary of English Etymology*. London: Oxford, 1966/1994.

RIO-TORTO, G. M.; VIEIRA, A. Nomes em -ismo no português dos séculos XVIII a XX. *Cuadernos del Instituto de Historia de la Lengua*, n. 2, p. 115-137, 2009.

STANISŁAWSKI, J. *The great Polish-English Dictionary*. Varsóvia: Wiedra Powsechna, 1969.

STARETS, S.; VOINOVA, N. *Dicionário Prático Português-Russo*. Moscou: Russki Jazyk, 1986.

WAHRIG, G. *Wörterbuch der deutschen Sprache*. München: Deutscher Taschenbuch, 1999.

# Anáforas associativas baseadas em relações eventivas

(Associative anaphoras based on eventive relations)

Alex de Britto Rodrigues<sup>1</sup>

<sup>1</sup>Mestre e Doutorando em Estudos Linguísticos pelo Departamento de Pós-Graduação em Letras da Universidade Federal do Paraná (UFPR)

britto4r@gmail.com

**Abstract:** This paper investigates the semantic proprieties presented in associative anaphora related to eventualities. Therefore, the theoretic scope consists of some elements which clarify this relation, such as: aspectual classes, verbal decomposition, participants of eventualities and qualia structure.

**Keywords:** associative anaphora; lexicon; eventualities.

**Resumo:** Este artigo visa a investigar as propriedades semânticas presentes nas anáforas associativas relacionadas a eventualidades. O quadro teórico seguido leva em consideração, portanto, alguns elementos que elucidam essa relação, tais como: classes aspectuais, decomposição verbal, participantes das eventualidades e estrutura qualia.

**Palavras-chave:** anáfora associativa; léxico; eventualidades.

## Introdução

Objetivamos nesta pesquisa analisar um subtipo das anáforas associativas (AAs), as relacionadas a propriedades de eventualidades, de modo a identificarmos os elementos que garantem as associações lexicais promovidas por esse fenômeno. Em outras palavras, buscamos elucidar o que garante, do ponto de vista semântico-lexical, a configuração dessas AAs. Antes de adentrarmos no percurso deste trabalho e esclarecermos como pretendemos chegar a tal objetivo, convém expor como entendemos o fenômeno a ser analisado.

As AAs são consideradas neste trabalho um subtipo das anáforas indiretas. Estas, sendo um fenômeno mais amplo, apresentam duas características: 1) seu referente é identificado por meio de informações presentes no contexto anterior, o que lhes garante a denominação de “anáforas”; 2) não possuem o mesmo referente de seu antecedente. Vejamos um exemplo de anáfora indireta:

(1) O casamento foi lindo. **Ela** estava radiante.

Nesse exemplo, como em uma anáfora direta, o referente da expressão anafórica “ela” é identificado em razão da instauração de um antecedente (“o casamento”). Porém, diferentemente de casos envolvendo anáforas diretas, as duas expressões não possuem a mesma referência, ou seja, não denotam a mesma coisa.

As AAs, sendo um fenômeno mais específico, são aquelas expressões que, por meio de *uma associação lexical* com um antecedente, garantem sua definitude, como no exemplo seguinte (traduzido de KLEIBER, 2001, p. 316):

(2) Uma velha senhora foi assassinada. **O homicida** não foi encontrado.<sup>1</sup>

Nesse exemplo, a expressão “o homicida” é definida em razão de sua relação com “assassinada”, ou seja, uma relação entre itens lexicais.

Ao afirmarmos que a expressão anafórica associativa é definida e estabelecida por meio de uma associação lexical, ancoramos nosso trabalho em uma perspectiva que se atém a questões semânticas (KLEIBER, 2001; KORZEN; BUCH-KROMANN, 2011) (há outras que tratam desse fenômeno, como a perspectiva mais ampla da linguística textual, mas que não abordaremos aqui). É essa perspectiva que guiará o desenvolvimento de nosso quadro teórico e a nossa análise.

## Quadro teórico

Com o objetivo de elucidar propriedades das AAs relacionadas a eventos (ou eventualidades), discutimos, primeiramente, o conceito “eventualidade” (derivado do conceito de “evento”), desenvolvido na filosofia da linguagem e trabalhado na semântica. Na sequência, levamos em conta discussões a respeito da relação entre o léxico e as eventualidades, o que envolve as propriedades dos predicados verbais, a participação dos argumentos dos predicados verbais nas eventualidades denotadas pelos verbos e o acionamento da estrutura *qualia* de alguns nomes quando uma eventualidade é implicada. À medida que esse quadro for exposto e explicado, algumas considerações a respeito das AAs são feitas para que o recorte teórico seja justificado e a análise de nosso *corpus* realizada na sequência seja delineada.

## Eventualidades e seus participantes

Uma das perspectivas que conduziram o desenvolvimento de uma semântica de eventos pós-Davidson<sup>2</sup> é a de Parsons (1990). Entre os componentes teóricos que o autor sistematiza, estão: a identificação de uma variável de evento para estados, colocando-os, com os eventos e as atividades, em uma classe maior, a das eventualidades; a identificação de propriedades das eventualidades, que as dividem em partes; e a consideração de participantes,<sup>3</sup> que recebem papéis temáticos de acordo com a função que desempenham nas eventualidades.

As ideias de Parsons a respeito dos dois primeiros componentes são desenvolvimentos das propostas de Vendler (2005[1967]) e Dowty (1979). Às classes vendlerianas (estados, processos, eventos *accomplishments* e eventos *achievements*), que são tipos de eventualidades denotadas pelos verbos, são adicionadas as ideias de culminância (Cul) e

1 No original : “Une vieille dame a été assassinée. **Le meurtier** n’a pas été retrouvé”.

2 O trabalho de Davidson (2001[1967]) é considerado a principal fonte do surgimento de uma semântica de eventos. A noção de “evento” então conhecida era a de um argumento verbal denotado por alguns tipos de verbos. Em perspectivas pós-Davidson, todos os verbos passaram a ser considerados, sendo não mais o conceito de “evento” mas, sim, o de “eventualidade” (mais amplo), estabelecido como um argumento denotado. De um modo simplificado, uma eventualidade é a denotação mais básica do verbo, é a “situação” implicada, o que pode envolver diferentes participantes.

3 É como iremos nos referir aos argumentos de um predicado verbal, pois eles são “participantes” de um evento.

*holding* (Hold). De modo resumido, podemos dizer que estados e processos são caracterizados por não terem culminância, apenas *holding*, ao contrário de eventos, que possuem culminância. Para nossos propósitos, é interessante considerarmos todas as quatro classes vendlerianas, mas as distinções entre elas só serão relevantes para nós na medida em que mostrarem a que a configuração das AAs é sensível. Vejamos um exemplo de anáfora associativa envolvendo cada classe para ilustrarmos esse ponto:

- (3) A cidade ficou deserta. **Os habitantes** viajaram no feriado. (habitar = estado).
- (4) Havia uma piscina enorme no ginásio. **Os nadadores** adoraram. (nadar = processo).
- (5) Esse livro é muito bom. **O escritor/o autor** merece os parabéns. (escrever = *accomplishment*).
- (6) A vitória foi merecida. **Os vencedores** ainda comemoram. (vencer = *achievement*).

Como podemos notar, as AAs podem ser baseadas em relações promovidas por qualquer uma das classes. Resta-nos investigar de que modo as propriedades e os elementos de tais classes contribuem para a configuração do fenômeno analisado, o que será realizado nas seções seguintes. Antes disso, estabeleceremos o limite entre as considerações de Parsons e Vendler, dentro do âmbito filosófico, e nossa abordagem, desenvolvida no âmbito de uma investigação lexical.

Ao tratar dos participantes das eventualidades, Parsons (1990, p. 72) se baseia na noção de papéis temáticos. Partindo de considerações de linguistas como Fillmore (2003[1968]), o autor entende um papel temático como o significado atribuído a um sintagma nominal em sua relação com um predicado verbal. Os papéis temáticos serão mais bem discutidos posteriormente. No momento, apenas constatamos que eles já foram, dentro do âmbito da filosofia (por PARSONS, 1990), relacionados a eventualidades, relação esta que manteremos.

Depois de esboçar considerações linguísticas sobre os papéis temáticos, Parsons passa a se ater ao seu objetivo central, estabelecer relações em uma ontologia, diferente do nosso. Depois de descrever alguns papéis, o autor comenta que “O que está faltando é como tudo isso é correlacionado com o domínio semântico – como os papéis se relacionam com o mundo” (PARSONS, 1990, p. 72, tradução nossa). Não nos interessa, portanto, o desenvolvimento da ideia do autor a partir disso.

Voltando às classes vendlerianas, também podemos estabelecer um limite entre uma perspectiva filosófica e a nossa. Parsons (1990, p. 34, tradução nossa) comenta que, ao classificarmos as eventualidades, “não é claro se a classificação é de coisas no mundo ou de partes da linguagem. Quando partes da linguagem estão sendo classificadas, às vezes verbos, às vezes VPs, e às vezes sentenças inteiras são então classificadas”.<sup>4</sup> Depois disso, por sua perspectiva ser filosófica, afirma que está preocupado em classificar, principalmente, coisas não linguísticas – eventualidades.

Sobre a que parte da linguagem as eventualidades correspondem, Vendler (2005, p. 22) diz, primeiramente, que as classes de eventualidades se relacionam a classes de verbos. Porém, o autor reconhece que uma eventualidade depende não apenas do verbo,

---

4 No original: “It is sometimes unclear whether the classification is meant to classify things in the world or pieces of language. When pieces of language are being classified, sometimes verbs, sometimes VPs, and sometimes whole sentences are so classified”.

mas também do contexto sintático. Entre os exemplos que o autor cita para ilustrar essa ressalva está “fumar”, uma atividade, e “pintar”, um *accomplishment*. Ambos os verbos podem ter uma leitura derivada de estado (VENDLER, 2005, p. 27), como em “ele fuma” e “ele pinta”, o que pressupõe uma categoria “original” desses verbos (atividade e *accomplishment*, respectivamente). Em suma, para Vendler, cada verbo corresponde a uma classe de eventualidade, mas a leitura dela pode ser alterada a depender do contexto sintático desse verbo.

É importante notar que, embora o autor não mencione o termo “léxico”, a noção de que o verbo relaciona-se, originalmente, a uma classe de eventualidade corresponde à noção de “aspecto lexical”, desenvolvida dentro da semântica lexical (cf. ROTHSTEIN, 2004). Por aspecto lexical entendemos o modo como as propriedades estruturais de uma eventualidade são lexicalizadas. Dentro dessa perspectiva, já fora do âmbito da filosofia e dentro do quadro teórico seguido diretamente por nós, o significado verbal pode passar por *shiftings*, operações em que o significado “original” do verbo incorpora uma nova estrutura, o que faz a classe verbal mudar (ROTHSTEIN, 2004, p. 34). Desse modo, traçamos a correspondência entre verbos e eventualidades.

### Decomposição verbal em primitivos

A caracterização das classes aspectuais foi sugerida por Dowty (1979) por meio dos operadores DO, CAUSE e BECOME, a serem explicados a seguir. Para o autor, a classe dos *accomplishments* se caracteriza por apresentar duas partes, uma em que há o operador CAUSE, e outra em que há o operador BECOME. Vejamos o exemplo em (7) e sua representação em (8):

(7) Maria quebrou o vaso.

(8) [[Maria fez algo] CAUSE [BECOME [o vaso quebrado]]].

O operador BECOME corresponde à parte do evento que contém uma mudança de estado, que, no caso de 7, é a mudança de não quebrado para quebrado. Já o operador CAUSE corresponde à parte do evento que contém um agente causador da mudança de estado em que BECOME opera.

Em *achievements*, não há um agente causador; portanto, na representação proposta por Dowty, não há o operador CAUSE. Há apenas o operador BECOME, como em (9), representado em (10):

(9) O vaso quebrou.

(10) BECOME [o vaso quebrado].

O operador BECOME captura a semelhança entre (7) e (10), correspondente à mudança de estado de um participante no evento.

O operador DO é o que distingue atividades de estados. Ambos os tipos de eventualidades não apresentam mudança de estado, ou seja, não têm em suas representações o operador BECOME. Além disso, Dowty não comenta sobre a possibilidade de haver CAUSE, pois esse operador só aparece exemplificado em construções *causativas* em que

há uma mudança de estado (BECOME) *causada*. O fator semântico com que DO contribui é a presença de controle pelo agente, controle este normalmente chamado, de modo não preciso, de “volição”/“intenção”<sup>5</sup> (cf. DOWTY, 1979, p. 114, 118). A classe aspectual representada apenas com o operador DO, a das atividades, possui um agente controlador, ao passo que a classe dos estados, sendo a mais primitiva, não possui nenhum desses operadores em sua representação. Vejamos alguns exemplos de atividades e estados citados por Dowty:

- (11) John beijou Maria.<sup>6</sup> (DOWTY, 1979, p. 115) = atividade.
- (12) John está sendo educado.<sup>7</sup> (DOWTY, 1979, p. 114) = atividade (com a leitura intencional, sem a leitura de voz passiva possibilitada em português).
- (13) John é educado.<sup>8</sup> (DOWTY, 1979, p. 115) = estado.

O evento descrito em (11) não apresenta problemas em ser descrito como uma atividade, conforme o autor. No entanto, (12) e (13) recebem classificações diferentes, mesmo contendo o mesmo predicado “ser educado”. O que torna (12) uma atividade é o controle do agente, ou seja, é possível estar propositadamente em um estado, sendo DO lexicalizado como “ser” (cf. DOWTY, 1979, p. 115). Segundo o autor, é o gerúndio que permite essa lexicalização. No entanto, consideramos que (12) se trata de um caso de *shifting*, sendo o verbo “ser” classificado, originalmente, como estado, isto é, sem DO. Esse verbo, então, em razão do processo de *shifting*, passa a incorporar esse operador.

A proposta de Dowty é mais adequada ao nosso objetivo do que outras que, em vez de propor uma decomposição em primitivos, estabelecem uma estrutura eventiva, como a sugerida por Pustejovsky (1996). Esse tipo de estrutura permite que visualizemos a relevância das partes e de algumas propriedades dos eventos não apenas como primitivos, mas como elementos estruturados. Porém, não tira o mérito de decomposições como a sugerida por Dowty, pois essa decomposição, ao estabelecer primitivos, revela características não apreensíveis diretamente em uma estrutura como a de Pustejovsky (1996). Por exemplo, agentividade é uma propriedade que Pustejovsky (1996) não considera relevante para determinar tais estruturas, mas ela está presente na proposta de Dowty, e, como será visto, essa propriedade pode ser justificada por estar relacionada diretamente na configuração de algumas AAs.

## Estrutura qualia

A estrutura *qualia*, conforme proposta de Pustejovsky (1996), é o que dá força relacional aos itens lexicais. Corresponde a informações que promovem restrições de uso e que são selecionadas conforme o contexto sintático ou semântico. Essa estrutura é composta por quatro aspectos (ou papéis) do significado:

---

5 Lembremos que a noção de “volição” como um componente de alguns tipos de eventos já havia sido apontada por Davidson, conforme vimos anteriormente.

6 No original: “John kissed Mary”.

7 No original: “John is being polite”.

8 No original: “John is polite”.

- Constitutivo: papel que relaciona o objeto e as suas partes. Corresponde a elementos como o material, o peso e aos elementos definidos como partes.
- Formal: distingue o objeto dentro de um domínio maior. Corresponde a características como orientação, magnitude, forma, dimensão, cor e posição.
- Télico: compreende a proposta ou a função do objeto. Refere-se à proposta de uma ação realizada por um agente ou à função que determina algumas atividades.
- Agentivo: trata-se de fatores envolvidos na origem ou na criação de um objeto. Compreende elementos como criador e artefato.

Consideraremos, neste trabalho, de modo mais detido, apenas os papéis télico e agentivo por permitirem explicar certos casos de AAs baseadas em eventualidades. Os papéis constitutivo e formal, porém, devem ser considerados na análise de outros casos de AAs, como os que envolvem relações nominais (como a relação parte-todo, nos termos de Kleiber, 2001), o que não é nosso objetivo.

Os papéis *qualia*, acionados em contextos sintático-semântico específicos, sugerem interpretações de acordo com esses contextos (PUSTEJOVSKY, 1996, p. 87). Por exemplo, o item “livro” possui como papel télico o predicado verbal “ler”, acionado no contexto de (14).

(14) Bill gostou do último livro do Steven King.<sup>9</sup> (PUSTEJOVSKY, 1996, p. 88) (ler)

Nesse exemplo, há uma elipse do predicado verbal “ler”, recuperado na interpretação da sentença graças à presença do item “livro”, ou seja, o que Bill gostou foi de “ler” o livro. Observamos que o verbo “gostar” seleciona apenas eventualidades como complemento, ou seja, se um argumento com núcleo nominal está na posição de complemento sintático de gostar, uma eventualidade precisa ser inferida.

Em (15), temos um exemplo de AA baseado na relação entre “livro” e seu papel agentivo, que é “escrever”.

(15) Este livro está em inglês, mas o autor é russo.<sup>10</sup> (KLEIBER, 2001, p. 344)

Nesse exemplo, “livro” é um participante do evento “escrever”, correspondente a seu papel agentivo. Só é possível inferir “autor”, outro participante desse evento, se assumirmos que há um evento estabelecido. A relação entre um evento e seus participantes será tratada na próxima seção.

A AA em (15) é classificada por Kleiber como “funcional”, pois o referente da expressão anafórica desempenha uma função em relação ao referente do antecedente. Essa explicação do autor, porém, não é capaz de demonstrar como essa função desempenhada é inferida, nem qual é o tipo de relação entre anafórico e antecedente. Esse problema é contornado quando consideramos a estrutura *qualia*, pois é por meio do conteúdo correspondente a ela que podemos inferir um evento e, posteriormente, os participantes desse evento.

<sup>9</sup> No original: “Bill enjoyed Steven King’s last book”. (reading)

<sup>10</sup> No original: “Ce livre est en anglais, mais l’auteur est russe”.

## Participantes das eventualidades

A relação entre uma eventualidade e seus participantes já foi considerada para os casos de AAs funcionais. As AAs actanciais, na classificação de Kleiber (2001), compreendem essa relação de modo mais direto, pois correspondem a uma expressão anafórica que recupera um actante de uma eventualidade presente no antecedente. Nessas AAs, esses actantes, definidos na grade argumental dos verbos, são identificados por suas propriedades projetadas pelos verbos, e não por suas posições sintáticas. Portanto, esses tipos de AAs justificam a consideração das propriedades dos actantes relacionadas às eventualidades em vez da posição na sentença determinada por projeções argumentais. Em outros termos, essas AAs requerem que consideremos propriedades dos argumentos que garantem sua relação com predicados verbais fora de uma projeção sintática. Para isso, vemos um argumento como um participante de uma eventualidade, e a identificação do papel que ele desempenha nessa eventualidade é o que garante a relação promovida pelas AAs actanciais. Esse papel corresponde ao papel temático de um argumento.

A noção de papéis temáticos, de acordo com Cançado (2008), foi cunhada por Gruber (1965), Fillmore (2003[1968]) e Jackendoff (1972). A proposta desses autores é explicar relações estabelecidas entre argumentos e seus predicados verbais não capturadas por noções estritamente sintáticas. Conceitos como sujeito e objeto, ou como argumento interno, argumento externo e adjunto, não traduzem a função semântica desses elementos. Por exemplo, a denominação de “sujeito” para “John” e “a chave” nos exemplos (16) e (17) não explicita que esses argumentos exercem funções semânticas diferentes, correspondentes, respectivamente, a “agente” e “instrumento”.

(16) John abriu a porta.<sup>11</sup> (FILLMORE, 2003, p. 47)

(17) A chave abriu a porta.<sup>12</sup> (FILLMORE, 2003, p. 47)

Contudo, algumas configurações sintáticas podem ser associadas aos papéis temáticos. Por exemplo, em uma sentença em que o sujeito é um instrumento, não é possível um sintagma preposicional adjunto formado pela preposição “com”, ao passo que em uma sentença cujo sujeito é agente, esse tipo de adjunto é possível:

(18) \*A chave abriu a porta com o machado.

(19) John abriu a porta com a chave.

A explicação para a agramaticalidade de (18) é que a função de instrumento, ocupada pelo argumento sujeito, impede outro argumento (“o machado”) de ocupar a mesma função. Como esse argumento, por se constituir em um sintagma preposicional nucleado por “com”, desempenha prototipicamente a função de instrumento, mas não pode ocupar tal função, ficando sem papel temático, então a sentença passa a ser agramatical, já que todo argumento DP precisa receber um papel temático (cf. CHOMSKY, 1981).

Ao contrário de considerações estritamente sintáticas, para as quais não é necessário determinar os tipos de papéis temáticos, identificar esses papéis torna-se relevante quando o objetivo é analisar relações semânticas estabelecidas entre predicados verbais

11 No original: “John opened the door”.

12 No original: “The key opened the door”.

e seus argumentos. A nossa motivação direta para identificarmos os papéis temáticos em nossa análise das AAs baseadas em relações eventivas é que a mera identificação do tipo de argumento não é suficiente para a explicitação de relações estabelecidas fora da sentença. Um exemplo de AA actancial de Kleiber (2001) demonstra isso:

(20) Uma velha senhora foi assassinada. **O homicida** não foi encontrado.<sup>13</sup> (KLEIBER, 2001, p. 317)

Em 20, “o homicida” é um participante do evento denotado pelo predicado verbal “matar”. Sua participação nesse evento não é capturada pela identificação de sua posição argumental, mas sim pelo seu papel temático. Exemplos como (20), baseados na relação entre um evento e um participante, podem ser vistos como uma justificativa empírica para a introdução de papéis temáticos, feita por Parsons (1990), em uma semântica de eventos. Em outros termos, conforme Rothstein (2004, p. 4), papéis temáticos denotam funções dos participantes estabelecidas pelos eventos, e são essas funções que garantem associações como a encontrada no exemplo anterior.

Entretanto, identificar os papéis temáticos, tarefa importante para revelar as associações em exemplos como o anterior, é algo problemático. Como diz Dowty (1991, p. 547), “talvez não haja conceito em teoria sintática e semântica moderna que esteja tão frequentemente envolvido em tantos contextos, mas a respeito do qual haja tão pouco consenso sobre sua natureza e sua definição, como o conceito de papéis temáticos”.<sup>14</sup> Essa falta de consenso é atestada pelas diferentes listas de papéis temáticos sugeridas.

No momento, para nossas considerações a respeito dos papéis temáticos, é suficiente a lista geral sugerida por Cançado (2008, p. 111-112), elaborada com base na literatura sobre o assunto: agente (o desencadeador de alguma ação, capaz de agir com controle), causa (o desencadeador de alguma ação, sem controle), instrumento (o meio pelo qual a ação é desencadeada), paciente (a entidade que sofre o efeito de alguma ação, havendo mudança de estado), tema (a entidade deslocada por uma ação), experienciador (ser animado que mudou ou está em determinado estado mental, perceptual ou psicológico), beneficiário (a entidade que é beneficiada pela ação descrita), objetivo – ou objeto estativo – (a entidade à qual se faz referência, sem que esta desencadeie algo ou seja afetada por algo), locativo (o lugar em que algo está situado ou acontece), alvo (a entidade para onde algo se move) e fonte (a entidade de onde algo se move)

Entendemos que, do ponto de vista teórico, não há como definir uma lista precisa de papéis temáticos, mas do ponto de vista metodológico, é possível estabelecer uma relação de papéis temáticos relevantes para uma análise específica. Não é relevante estabelecermos uma lista fixa de papéis temáticos a serem atribuídos aos participantes de uma eventualidade, mas, sim, identificarmos a relação entre um participante e uma eventualidade.

13 No original: “Une vieille dame a été assassinée. Le meurtrier n’a pas été retrouvé”.

14 No original: “There is perhaps no concept in modern syntactic and semantic theory which is so often involved in so wide a range of contexts, but on which there is so little agreement as to its nature and definition, as THEMATIC ROLE”.

## Análise de dados

Com base na discussão a respeito de relações estabelecidas entre eventualidades e suas propriedades e/ou participantes, podemos estabelecer quatro tipos de AAs organizados da seguinte maneira: 1) os que recuperam um primitivo (DO, CAUSE ou BECOME), 2) os que recuperam uma propriedade comum a todas as eventualidades, 3) os que recuperam um participante<sup>15</sup> e 4) os que recuperam um primitivo ou um participante baseando-se na estrutura *qualia*. Os quatro exemplos a seguir, que servem como ponto de partida para a discussão proposta neste trabalho, correspondem, respectivamente, a esses quatro tipos:

- (21) Durante oito meses, eles elaboraram os textos, revisaram o material coletivamente e fizeram novas versões. **A intenção** era produzir uma antologia de relatos e socializá-las com a comunidade escolar. (FERNANDES, Elisângela. Propostas ruins geram desinteresses pela escrita. **Nova Escola**, Fala Mestre, São Paulo, mar. 2011, n. 240, p. 28-29)
- (22) Há um tempo, o Centro Acadêmico resolveu fazer uma Assembleia. **O motivo** era uma votação num departamento sobre “criar um curso novo” ou “manter um curso único, devidamente reformado”. (ORNELAS, Claudio Luciano. Bullying no ensino superior – capítulos de uma novela mexicana. *Boca do Inferno*, Curitiba, ed. 25, p. 12-13, abr./jun. 2011)
- (23) Depois de revolucionar processos industriais, como no setor automotivo, os robôs chegaram às salas de cirurgia. Em um cenário que até há pouco seria definido como futurista, **o cirurgião** instala-se em frente de um console e opera manejando os braços do robô. (Cirurgia robótica vale mesmo a pena? Página Einstein. Disponível em: <<http://www.einstein.br/pagina-einstein/Paginas/cirurgia-robotica-vale-mesmo-a-pena.aspx>>. Acesso em: 25 ago. 2010)
- (24) Eu sonhei que estava numa **pizzaria** pedindo uma pizza meio calabresa, meio extravagância. E **o garçom** ainda me perguntava se a metade extravagância era com cebola ou não. (PRATA, Mario. Extravagância. In: **Cem melhores crônicas**. São Paulo: Planeta do Brasil, 2007, p. 21)

Na sequência, por meio da análise de dados, buscamos identificar algumas limitações da proposta desses tipos de AAs.

### AAs que recuperam primitivos

Em (21), “a intenção” recupera uma propriedade presente no antecedente, configurado por três eventualidades: “elaboram os textos”, “revisaram o material” e “fizeram novas versões”. A hipótese é de que o termo “intenção” apareça apenas quando houver a propriedade de “volição” nas eventualidades, o que implica a presença do primitivo DO. Isso prevê que, ao mudar as eventualidades no antecedente para alguma que não tenha esse primitivo, a sequência se torne inaceitável:

- (25) ?Durante oito meses, eles apanharam, sofreram e sentiram muita dor. **A intenção** era aprender com os erros.<sup>16</sup>

---

15 Chamado de *actante* por Kleiber (2001), que faz referência ao fato de ser um argumento verbal, ou de *papel temático*, por Korzen e Buch-Kromann (2011), que fazem referência ao fato de a expressão anafórica ser identificada pela propriedade semântica que a relaciona a um predicado verbal. Escolhemos a denominação *participante*, como Parsons (1990), para fazermos referência à participação do elemento denotado pela expressão anafórica em uma eventualidade e para promovermos uma generalização que consiste em todas as AAs tratadas aqui serem relacionadas a eventualidades.

16 Essa sequência seria aceitável se “apanharam”, “sofreram” e “sentiram” fossem interpretados como “propositais”, deixando, nos termos de Dowty (1979), de ser estados e passando a ser atividades, implicando o primitivo DO. Não nos parece ser a interpretação mais comum.

A única leitura possível de (25) é a que infira algum evento que permita a presença de volição. Porém, caso haja essa leitura, não se trataria mais de uma relação entre itens lexicais nem de um caso de anáfora, pois esse evento seria inferido do contexto. Assim, o primeiro tipo de AAs apresentado no começo do capítulo é justificado. Porém, ainda falta investigarmos a possibilidade de AAs que recuperem os primitivos CAUSE e BECOME.

Alguns dados parecem ser candidatos a recuperarem o primitivo CAUSE. No exemplo a seguir, “a razão”, assim como “o motivo” em (22), aparentemente, recupera essa propriedade:

- (26) Segundo o estudo, a Educação é a que mais eleva o Produto Interno Bruto (PIB) e, junto com a saúde e o Bolsa Família, lidera a diminuição da desigualdade na distribuição de renda [...]. **A razão** é simples: 60% do orçamento vai para o pagamento de professores e funcionários. (SALLA, Fernanda; LIMA, Laize. Gasto em Educação eleva PIB. *Nova Escola*, Em dia, São Paulo, n. 240, p. 32-33, mar. 2011)

Em nenhum dos dois exemplos, porém, a expressão anafórica recupera uma propriedade presente no antecedente. O que essas expressões sugerem é que os eventos presentes no antecedente participam de cadeias causais em que a causa do evento (não o primitivo CAUSE), presente no antecedente, seja expressa na segunda sentença. Se essa “causa”, expressa por termos como “razão” e “motivo”, não é a propriedade associada a um protoagente que causa algo, isto é, não é uma propriedade interna a uma eventualidade, podemos estipular que qualquer eventualidade tenha a propriedade de fazer parte desse tipo de cadeia causal. Vejamos eventualidades diferentes para confirmarmos essa hipótese:

- (27) João corre todos os dias. **O motivo/a razão** é a implicância de sua esposa.  
(28) João ama Maria. **O motivo** é o caráter da moça.  
(29) O vento destruiu a cidade. **O motivo/a razão** é a pressão atmosférica.  
(30) O morro desmoronou. **O motivo/a razão/a causa** foi a forte chuva dos últimos dias.

Essas expressões apenas indicam a vinculação a uma cadeia causal. Não encontramos em nossos dados uma expressão que recuperasse o primitivo CAUSE do antecedente, o que deixa a pergunta em aberto: por que um primitivo como DO pode ser recuperado em uma AA e um primitivo como CAUSE (a princípio) não pode? Poderia ser dito que DO é um primitivo mais saliente que CAUSE, mas isso ainda não explicaria por que um primitivo é mais saliente que o outro. Provavelmente, isso se deve ao fato de que CAUSE nunca aparece sem BECOME, ou seja, talvez BECOME nunca permita CAUSE ser saliente o suficiente para ser recuperado.

A respeito da possibilidade de recuperação do primitivo BECOME em uma AA, encontramos o exemplo a seguir:

- (31) Desde que ela entrou aqui, senti que ganhou confiança para caminhar e até correr. **A mudança** foi visível. (GALVEZ, Camila. Mães pedem vaga em escola especial. *Diário da Grande ABC*. Disponível em: <<http://www.dgabc.com.br/News/5959811/maes-pedem-vaga-em-escola-especial.aspx>>. Acesso em: 06 jun. 2012.)

Em (31), supomos que a expressão “a mudança” recupera BECOME do *achievement* “ganhou confiança”. Para verificarmos se essa expressão realmente recupera esse primitivo, consideremos um caso em que não há esse primitivo:

(32) ?João amou Maria. **A mudança** foi visível.

Como pode ser notado, a presença da mesma expressão torna a sequência em (32) inaceitável (em uma leitura, obviamente, em que tal expressão é anafórica associativa, em que a dependência interpretativa não se dê pragmaticamente).

Assim como DO, portanto, BECOME também é saliente o suficiente para ser recuperado em uma AA. A não recuperabilidade de CAUSE<sup>17</sup> ainda precisa ser analisada, o que requer uma investigação que vai além do propósito deste trabalho.

### **AAs que recuperam propriedades comum a todas as eventualidades**

A respeito do segundo tipo de AA considerado por nós, o que recupera uma propriedade comum a todas as eventualidades, já verificamos o caso em que a AA relaciona-se à possibilidade de participar de cadeias causais (exemplos 22 e 26). Do mesmo modo, estar vinculada a tempo e a espaço é uma propriedade de todos os tipos de eventualidades, o que sugere que expressões de tempo e de lugar sejam recuperadas de qualquer eventualidade. Vejamos os seguintes exemplos:

(33) Abriu os olhos. **O lugar** estava escuro. (VIANCO, Andre. *Bento: Herói humano, inimigo vampiro*. São Paulo: Novo Século, 2003)

(34) Com esse espírito, Arouca foi nomeado coordenador-geral da 12ª Conferência Nacional de Saúde. **O momento** era de grande esperança. (FIOCRUZ. Secretaria de Gestão Participativa. Disponível em: <[http://bvsarouca.icict.fiocruz.br/sanitarista03\\_03.html](http://bvsarouca.icict.fiocruz.br/sanitarista03_03.html)>. Acesso em: 01 jan. 2012)

A hipótese de essas expressões destacadas poderem estar presentes em AAs relacionadas a qualquer eventualidade, porém, não se confirma. A expressão anafórica presente em (33) é um exemplo:

(35) João morreu. **O lugar** ficou marcado.

(36) João corre todos os dias. **O lugar** nunca se repete.

(37) ?João ama Maria. **O lugar** sempre fica contagiado.

A não aceitabilidade de (37) se deve ao fato de estados, com exceção daqueles em que um locativo não seja um adjunto, não estarem vinculados a um lugar específico. Se João ama Maria, ele a ama independentemente do lugar. Porém, o exemplo a seguir, em que não há um estado, mas uma atividade, também não é aceitável:

(38) ?João corre muito rápido. **O lugar** nunca se repete.

A hipótese que explica a não aceitabilidade de (38) é que esse exemplo não se trata, na realidade, de uma atividade. “Correr muito rápido” é interpretado como uma propriedade de “João”, ou seja, a eventualidade é interpretada como um estado (algo como “João é muito rápido”). Mesmo que no léxico de “correr” consideremos que atividade esteja marcada, lembremos que a interpretação de uma classe de eventualidade não depende só do verbo, podendo ser resultante também da presença de outros elementos associados a ele, como é o caso do adjunto “muito rápido” em 38. Portanto, a respeito do

17 Caso algum dado venha a demonstrar a possibilidade da recuperação de CAUSE em AAs, ainda seria notável a baixa ocorrência dessa recuperação em relação aos outros dois primitivos.

locativo “o lugar” presente em AAs, concluímos que a única restrição de sua ocorrência é em estados.<sup>18</sup>

No que se refere às expressões temporais, a expressão “o momento” parece sofrer o mesmo tipo de restrição de “o lugar”:

- (39) João morreu. **O momento** foi oportuno.  
(40) João correu muito. **O momento** foi glorioso.  
(41) ?João amou Maria. **O momento** foi adequado.

Assim como a expressão “o lugar”, “o momento” não ocorre como AA vinculada a um estado. Isso demonstra a impossibilidade de verbos de estado implicarem uma unidade de tempo de modo a torná-la saliente, o que pode funcionar como um teste para identificarmos essa classe verbal.

Como vimos, a recuperação de unidades de espaço e de tempo é restringida pelas propriedades das classes de eventualidades. Nesse sentido, mesmo que ocorrer em um espaço e em um tempo seja uma característica comum a todas as eventualidades, a recuperação dessa característica está sujeita ao sentido da expressão anafórica. Embora a vinculação a unidades de tempo e de espaço seja garantida em um nível de representação ontológica (o que não é considerado na análise do fenômeno em questão), a vinculação a tais unidades não é garantida em um nível lexical. Em outros termos, a possibilidade de uma AA relacionada a tais unidades depende da informação lexical que o antecedente permite ser acionada.

### **AAs que recuperam participantes**

Sobre as AAs que recuperam participantes, lembremos que exemplos como 3 já foram tratados por Kleiber (2001) e Korzen e Buch-Kromann (2011). Porém, esses autores não fazem distinção entre os participantes no que se refere à recuperabilidade por AAs. Acabam por considerar um participante com papel temático de instrumento da mesma maneira como consideram um com papel de agente ou tema. No entanto, um participante correspondente a um adjunto de um predicado verbal, posição canônica de instrumento, parece ser menos recuperável do que um agente, por exemplo. Vejamos o seguinte exemplo:

- (42) Suspeito de matar procuradora em MG é encontrado morto em motel

A faca estava perto do corpo pouco tempo depois de a mulher, a procuradora Ana Alice Moreira de Melo, ter sido encontrada morta em casa. (BOM dia Brasil. Suspeito de matar procuradora em MG é encontrado morto em motel. *Bom dia Brasil*, 03 fev. 2012. Disponível em: <<http://g1.globo.com/bom-dia-brasil/noticia/2012/02/suspeito-de-matar-procuradora-em-mg-e-encontrado-morto-em-motel.html>>. Acesso em: 14 fev. 2012)

Um evento como o denotado pelo verbo “matar”, que apresenta o primitivo CAUSE, licencia um instrumento como participante. Isso colabora para recuperarmos a noção de instrumento em uma AA, mas não a noção específica do tipo de instrumento. Diferentemente

---

<sup>18</sup> Lembremos que o antecedente de uma AA relacionado a uma eventualidade pode ser o interpretado a partir de uma combinação estrutural (em que o léxico do verbo é relevante, mas não suficiente).

de (23), em que “cirurgião” é um item cujo conteúdo semântico o identifica como agente de “operar” (ou “fazer cirurgia”), dados como (42) apresentam o item “faca”, que não possui a informação de “ser um instrumento de matar”. Explicamos casos como o (42) do seguinte modo: o antecedente licencia um instrumento como participante; logo, uma expressão genérica como “o instrumento” poderia ser recuperada; porém, uma expressão específica é utilizada, e a informação específica, não implicada pelo antecedente, é introduzida por meio de processos discursivos não considerados por nós. Sendo assim, a classificação de casos desse tipo como AAs deve ser posta sob suspeita. Por outro lado, (43) demonstra-se um caso legítimo de AA (ou seja, é garantido por relações semânticas):

- (43) Homem esfaqueia mulher no centro de Rio Claro  
[...] Guardas municipais foram chamados e encontraram a vítima, uma mulher de 48 anos, com ferimentos na face e no abdômen – **a faca** estava jogada a seu lado. (JORNAL da Cidade. Homem esfaqueia mulher no centro de Rio Claro. *Jornal da Cidade*, 01 nov. 2010. Disponível em: <<http://jornalcidade.uol.com.br/rioclaro/seguranca/assassinatos/67920-Homem-esfaqueia-mulher-no-centro-de-Rio-Claro>>. Acesso em: 14 fev. 2012)

Esse exemplo apresenta-se como uma AA em razão de o item “faca” poder ser recuperado de “esfaqueia”. O participante “a faca”, portanto, corresponde a um argumento *default* do predicado verbal presente no antecedente.

#### **AAs que recuperam um primitivo ou um participante baseando-se na estrutura *qualia***

Korzen e Buch-Kromann (2011) apontam que exemplos como 24 acionam a estrutura *qualia* do antecedente. De “pizzaria”, no antecedente, é acionado o papel télico “servir”, do qual “o garçom” é um participante. Consideremos 44, que também aciona o papel télico:

- (44) Na manhã do dia 5, um avião fretado paquistanês caiu pouco depois de decolar do aeroporto de Karachi, no Paquistão [...]. **Os 21 tripulantes**, incluindo os dois pilotos, morreram. (MACHADO, Juliano. Uma semana de desastres. *Época*, O Filtro, São Paulo, 08 nov. 2010, n. 651, p. 92)

Nesse exemplo, novamente o papel télico do antecedente (“um avião”) é acionado. Esse papel corresponde ao predicado verbal “tripular”, isto é, a proposta de “um avião” é “tripular”. Esse predicado verbal denota uma eventualidade de que “os 21 tripulantes” é um participante. O tipo de AA em (44) é o mesmo que Kleiber (2001) considera como funcional.

Se de um antecedente é possível recuperar, por meio da estrutura *qualia*, uma eventualidade, e desta, um participante, então é previsível que o mesmo processo permita recuperar, em vez de um participante, um primitivo. Esse é o caso de (45):

- (45) Aí podem me dizer: mas o livro é sobre o Louis. Sim, é para ser sobre ele. Mas já que **a intenção** era meramente contar a história do personagem, bastava escrever um livro, e poderia até ser narrado em primeira pessoa, sem a pretensão de criar uma entrevista entre dois personagens (que inexistente). (ADORNORULES. *Entrevista Com o Vampiro, Anne Rice*. 1 out. 2011. Disponível em: <<http://adornorules.wordpress.com/2011/10/01/entrevista-com-o-vampiro-anne-ric/>>. Acesso em: 5 fev. 2012)

Nesse exemplo, de “livro” o papel agentivo “escrever” é acionado, do qual é recuperado o primitivo DO, correspondente à expressão anafórica “a intenção”. Assim como nos casos em que é possível recuperar, de uma eventualidade no antecedente, participantes e

primitivos, nos casos em que é acionada da estrutura *qualia* no antecedente uma eventualidade, é possível recuperar, também, participantes e primitivos. Desse modo, traçamos um paralelo entre os dois tipos gerais de AAs que recuperam primitivos e participantes (os que acionam e os que não acionam estrutura *qualia*): o processamento de ambos é, parcialmente, o mesmo, consistindo na recuperação de um elemento a partir de uma eventualidade; o que diferencia os dois tipos é a complexidade adicionada com o acionamento da estrutura *qualia*.

## Conclusão

Os tipos de AAs discutidas neste trabalho podem ser aproximadas em um mesmo processo: uma eventualidade no antecedente (expressa ou acionada por meio da estrutura *qualia*) permite que um elemento (um participante ou um primitivo) implicado por ela seja retomado na expressão anafórica. Essa aproximação promove uma generalização capaz de demonstrar um processo produtivo na configuração do fenômeno analisado, que é a orientação de uma eventualidade para um elemento implicado por ela.

Porém, a discussão realizada, não pretendendo ser definitiva, não responde por que o primitivo CAUSE não pode ser recuperado em uma AA enquanto os outros podem. Além dessa questão, outras podem ser derivadas do que foi exposto: Quais outras propriedades específicas de cada eventualidade restringe a realização de AAs? Quais outros tipos de expressões recuperam propriedades comuns a todas as eventualidade e quais outros tipos recuperam propriedades específicas?

Mesmo que a análise realizada aqui possa deixar diversas questões sem resposta, ela permite uma organização dos dados levando em conta critérios relevantes nem sempre considerados pelos autores citados. Por exemplo, se as AAs são definidas por relações lexicais entre antecedente e anafórico, as considerações a respeito da recuperação de participantes e de primitivos devem levar em conta as informações estruturadas no antecedente, tais como a estrutura argumental e a estrutura *qualia*. Além disso, devem considerar apenas participantes ou primitivos de fato implicados.

## REFERÊNCIAS

CANÇADO, M. *Manual de semântica*. 2. ed. Belo Horizonte: Editora UFMG, 2008.

CHOMSKY, N. *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris, 1981.

DAVIDSON, D. *Essays on Actions and Events*. Oxford: OUP, 2001.[1967]

DOWTY, D. R. Thematic proto-roles and argument selection. *Language*, v. 67, n. 3, p. 547-519, 1991.

\_\_\_\_\_. *Word Meaning and Montague Grammar*. Boston: Reidel, 1979.

FILLMORE, C. Case for case In: \_\_\_\_\_. *Form and meaning in language*. Stanford: CSLI, 2003. [1968]. v. 1 (Papers on semantic roles, 1).

GRUBER, J. S. *Studies in Lexical Relations*. 1962. 310 f. Tese (Doutorado em Linguística) – Massachusetts Institute of Technology, Massachusetts, 1965.

JACKENDOFF, R. *Semantic interpretation in generative grammar*. Cambridge: MIT Press, 1972.

KLEIBER, G. *L'anaphore associative*. Paris: Presses Universitaires de France, 2001.

KORZEN, I.; BUCH-KROMANN, M.. Anaphoric Relations in the Copenhagen Dependency Treebanks. In: KORZEN, I.; BUCH-KROMANN, M. (Org.) *Proceedings of the Workshop Beyond Semantics: Corpus-based investigations of Pragmatic and Discourse Phenomena*. Göttingen, Alemanha: Bochumer Linguistische Arbeitsberichte, fev. 2011. p. 83-98.

PARSONS, T. *Events in the semantics of English: a study in subatomic semantics*. Cambridge: MIT press, 1990.

PUSTEJOVSKY, J. *The generative lexicon*. Cambridge: The MIT Press, 1996.

ROTHSTEIN, S. *Structuring events: a study in the semantics of lexical aspect*. Oxford: Blackwell, 2004.

VENDLER, Z. Verbs and times In: MANI, I.; PUSTEJOVSKY, J.; GAIZAUSKAS, R. (Ed.). *The language of time*. Oxford: OUP, 2005. [1967]. p. 21-32.

# A integração entre *affordances* e restrições gramaticais no processo de compreensão de sentenças

(The integration between *affordances* and grammatical constraints in the process of sentence comprehension)

Paulo Henrique Duque<sup>1</sup>

<sup>1</sup>Programa de Pós-graduação em Estudos da Linguagem –  
Universidade Federal do Rio Grande do Norte (UFRN)

ph.duque@uol.com.br

**Abstract:** This paper presents and discusses the assumptions of the relationship between grammar and meaning construction, adopting Constructions Grammar and Mental Simulation as the theoretical premises. The challenge of a perspective of usage-based grammar is to eliminate the gap between the grammar abstraction, embodied experiences and social interaction. In other words, the perspective proposed here can help us to reveal how the conceptual structure is embodied. In this paper we present, discuss and exemplify the relationship between language, cognition and embodiment in order to answer that challenge.

**Keywords:** cognitive linguistics; constructions grammar; affordances.

**Resumo:** Este trabalho apresenta e discute os pressupostos da relação entre gramática e construção de sentido, adotando como pressupostos teóricos a Gramática de Construções e a Simulação Mental. O desafio de uma perspectiva de gramática baseada no uso da linguagem é o de eliminar o fosso que existe entre a abstração da gramática e as experiências corporificadas e de interação social. Em outras palavras, a perspectiva aqui proposta pode nos ajudar a revelar como a estrutura conceitual é corporificada. Neste artigo, apresentamos, discutimos e exemplificamos as relações entre linguagem, cognição e corporalidade a fim de respondermos esse desafio.

**Palavras-chave:** linguística cognitiva; gramática de construções; affordances.

## Introdução

Todos os dias, ouvimos e lemos dezenas de milhares de palavras. Essas palavras nos chegam com rapidez e são estruturadas em padrões específicos. Por mais complexo que esse mecanismo possa parecer, na maior parte das vezes somos capazes de compreender o que nos é falado ou escrito de forma eficiente. Identificamos com clareza a quem ou a quem o falante está se referindo. Mais do que isso, inferimos e pressupomos situações e coisas que não tenham sido mencionadas. Quanto a esse aspecto em particular, consideremos, por exemplo, um trecho extraído do conto “Amor”, de Clarice Lispector:

(01) Enquanto não chegou à porta do edifício, parecia à beira de um desastre. (CLARICE LISPECTOR, 1998 [1974], p. 17)

Em (01), o local de onde a personagem veio não está explícito, mas sabemos que ela se deslocou de um ponto de origem até a porta do edifício. Até mesmo o trecho figurativo *à beira de um desastre* parece não prejudicar o processo de compreensão do trecho. É notório, pois, que mesmo não reconhecendo grande parte dos processos cognitivos subjacentes à produção e à compreensão da linguagem, somos capazes de produzi-la e compreendê-la

sem muito esforço, ou seja, fazemos isso automaticamente. Mesmo com todo o avanço da ciência da computação e da inteligência artificial, ainda não foi realizada uma implementação de programas computacionais capazes de processar pelo menos uma fração do que os seres humanos são capazes de produzir e compreender. O problema não está no poder de processamento dessas máquinas, mas na maneira como o tal processamento ocorre.

Ao considerarmos as relações entre linguagem, cognição e corporalidade situada, somos levados a admitir que, em vez de um mero conjunto de comandos simbólicos, a linguagem está diretamente vinculada aos sistemas sensorial e motor, responsáveis respectivamente pela percepção (real ou simulada) e pelos movimentos do corpo (reais ou simulados). A seguir, apresentaremos algumas reflexões sobre a integração entre linguagem, percepção e movimento.

### **O papel da linguagem na construção da percepção**

Ao investigar os efeitos cognitivos da direção automobilística em situações de risco, Jin (2008) conduziu um experimento em que dezesseis participantes dirigiram um carro virtual num simulador de direção em condições de risco, usando (ou não) um telefone celular. Os resultados revelaram que o uso do telefone celular reduz a percepção, a compreensão e a habilidade de projeção dos motoristas e, conseqüentemente, seu desempenho operacional, tático e estratégico. De acordo com os resultados da pesquisa, “parece que o uso do telefone celular distraiu a atenção e a percepção do motorista em termos do reconhecimento das informações essenciais da estrada” (2008, p. 64). O autor verificou que os motoristas enfrentaram distrações ao processar informações envolvendo o uso da memória de trabalho e da capacidade de processamento visual, o que afetou a compreensão. Isso parece evidenciar que dependemos da percepção, da atenção e da memória de trabalho para alcançarmos a compreensão de uma dada situação.

Da mesma forma que a conversação (por exemplo, em um telefone celular) prejudica a direção de automóveis, Becic et al. (2010) demonstraram que a produção e a compreensão da linguagem são afetadas durante a direção. Para demonstrar isso, os pesquisadores solicitaram que quarenta e oito participantes executassem tarefas de ouvir, contar e recontar histórias num simulador de direção. Os resultados evidenciaram uma considerável imprecisão das histórias contadas (e recontadas) pelos motoristas e grande prejuízo na produção linguística em situações de direção arriscada.

Em linhas gerais, esses experimentos demonstraram que, durante a produção e a compreensão de enunciados, as pessoas ativam regiões do cérebro relacionadas à percepção do mundo e ao movimento do corpo. Experimentos que, como os apresentados acima, denunciam a interferência entre linguagem e demais capacidades cognitivas têm contribuído para o fortalecimento de abordagens linguísticas perspectivadas na não autonomia da linguagem.

Outro aspecto que essas pesquisas salientam é a relevância de experiências sensório-motoras para processos de produção e compreensão da linguagem. A ativação seletiva dos sistemas perceptuais e motores durante a produção e compreensão de enunciados descrevendo cenas ou ações reflete a simulação mental do conteúdo das sentenças. Isso explica porque, de acordo com Yaxley e Zwann (2006, p. 2), “a compreensão de eventos descritos é frequentemente acompanhada de uma rica e aparentemente palpável sensação de se ‘estar lá’”.

O exemplo (02) pode revelar como o processamento de sentenças leva à realização de simulações mentais.

(02) Quem foi, quem foi/ Que falou no boi voador/ Manda prender esse boi/ Seja esse boi o que for.

Em (02), ao nos depararmos com a expressão “boi voador”, construímos a imagem mental de um animal que se parece, em vários aspectos, com um boi, ou seja, consideramos a anatomia, o modo de se locomover, a aparência e o que esse animal come, por exemplo. Precisamente, imaginamos um boi com asas (não mais do que isso), misturando determinadas características presentes na categorização de bois e de pássaros. O mais interessante é que, ao misturarmos características dos dois animais, nós nos restringimos, quase sempre, a incorporar um par de asas ao animal. Por que não incorporamos outras características salientes nos pássaros, como bicos e penas? Como duas simples palavras (boi e voador) possibilitam a construção de imagens visuais tão vívidas e repletas de detalhes na nossa mente? Exemplos como (02) nos permitem afirmar que a compreensão é um processo ativo e produtivo. Além disso, parte dele parece envolver a (re) criação e a combinação de experiências de percepção e de ação no mundo.

A partir dessa rápida reflexão acerca do exemplo (02), podemos estabelecer uma associação entre o boi voador, de Chico Buarque e Ruy Guerra, e as pesquisas sobre o uso do celular à direção. De acordo com Bergen e Heeyeon (2010), se para compreendermos palavras somos induzidos a ativar as imagens das coisas descritas, para interagirmos linguisticamente não nos valemos apenas das nossas capacidades linguísticas. Como vimos, para podermos comunicar coisas não presentes como bois voadores, por exemplo, ativamos áreas do cérebro destinadas à percepção e ao movimento. Quanto à interferência entre linguagem e percepção encontrada nas pesquisas sobre o uso de aparelhos celulares durante a direção, é possível que os informantes tenham encontrado dificuldades para usar sua visão (não virtual) na percepção de sinais e objetos que estivessem à sua frente, devido ao esforço de visualização virtual das cenas e ações sendo relatadas enquanto dirigiam.

Ainda com relação ao “boi voador”, podemos visualizar o animal utilizando uma capa de super-herói. Ele estaria numa roupa colante com um símbolo no peito. Em se tratando de um herói, a forma de voar, a sua postura e o seu movimento seriam diferentes da postura e da forma de deslocamento do boi com asas imaginado anteriormente. Somos capazes até de imaginar as patas dianteiras erguidas por sobre a cabeça e as patas traseiras esticadas nos moldes dos braços e pernas do super-homem em voo. Esse jogo de imaginação sobre bois voadores e bois super-heróis, por mais estranho que possa parecer, leva-nos a algumas reflexões:

1<sup>a</sup> – Por meio da linguagem, acionamos imagens mentais que variam substancialmente. Como vimos, é possível que, ao lerem as palavras *boi* e *voador*, os leitores tenham imaginado um boi de asas ou um boi com capa. Essa variação do processo de ativação de significados pela linguagem parece ter relação com o fato de sermos constituídos de diferentes experiências, expectativas e interesses;

2<sup>a</sup> – As imagens mentais construídas durante o processo de compreensão da linguagem não são arbitrárias. Note que, apesar de o *boi voador* possuir asas de pássaro, ele tem chifres, nariz, olhos, dentes, orelhas, pés e rabo de boi. Por que realizamos essa combinação tão específica? Tudo indica que isso se deva a peculiaridades da visão, como a exigência de

um foco de atenção: por focalizarmos o voo, não precisamos transformar a boca em bico, nem o pelo em penas, por exemplo. Parece que as imagens mentais e a geração de inferências acionadas pela linguagem acessam (e dependam de) o conhecimento sobre como o mundo funciona, como, por exemplo, a forma de voar dos pássaros e dos super-heróis;

3º – Comprometemos o nosso sistema visual na compreensão da linguagem de forma criativa. Podemos acionar percepções experienciadas previamente (tais como a aparência dos bois) e ações (tais como voar) e formar novas combinações. Nessa perspectiva, a construção do significado de *boi voador* depende do *blending*<sup>1</sup> (FAUCONNIER; TURNER, 2002) de representações conceituais independentes. Apesar de não termos visto o *boi voador*, o significado atribuído à expressão é resultante de um processo de integração conceptual criativo. Segundo Kosslyn *et al.* (2001, p. 635),

Imagens mentais não precisam resultar exclusivamente da recordação de objetos ou eventos anteriormente percebidos. Elas também podem ser criadas através da combinação e modificação de informações armazenadas em novas formas de percepção.

Assim como Feldman (2008), acreditamos que esse processo mental altamente criativo deva também ser considerado na investigação de casos mais simples envolvendo a construção de sentido de coisas que existem no mundo real, como *cavalo branco*, *sala comprida* ou *nota manchada*;

4º – O fato de compreendermos o que é um *boi voador*, apesar de essa criatura não existir é bastante esclarecedor. Ora, muitos teóricos defendem que a construção do significado dependa de associações entre as palavras e as coisas do mundo. Nessa concepção representacional de significado, a palavra *boi* estaria se referindo a um conjunto de coisas definidas objetivamente no mundo e; *voador* estaria se referindo a uma das propriedades dessas coisas. Nessa linha de raciocínio, a combinação de *boi* e *voador* deveria receber significado recorrendo-se a uma correspondência no mundo.

Em parte construímos o significado de *boi voador* recorrendo a entidades e propriedades que já tenhamos experienciado no mundo, mas é a nossa capacidade de criar imagens a partir dessas experiências que nos permite compreender criações fantásticas. Em outras palavras, o mecanismo cognitivo que produz essas imagens está fundamentado na corporalidade, ou como sugere Johnson (1987, p. 29), nos “nossos movimentos corporais através do espaço, nossa manipulação de objetos e nossas interações perceptuais”. Nesse sentido, as experiências corporificadas recorrentes são filtradas como padrões, que Johnson (1987) e Lakoff (1987) denominam *esquemas*. Na seção seguinte, trataremos da relevância da percepção e do controle motor na estruturação linguística.

## **O papel da percepção e do controle motor na construção da linguagem**

A noção de simulação mental (BARSALOU, 1999), apresentada de forma geral na seção anterior, tem sido fundamental para entendermos como o significado é construído gradativamente a partir das nossas experiências perceptuais e motoras. A simulação mental envolve a construção ativa de novas experiências motoras, perceptuais e afetivas, com base em percepções, ações e sentimentos anteriores e, o quanto é limitada e alimentada por tais experiências. Apesar de só percebemos a capacidade de simular mentalmente

<sup>1</sup> Mesclagem conceptual.

em situações em que somos convidados a criar imagens visuais, como a do *boi voador*, nós simulamos o tempo todo, mesmo em situações corriqueiras, como na visualização de rostos, na reconstrução mental de um lance de futebol ou ao contarmos histórias, fictícias ou não. Em especial, interessa-nos saber como a simulação mental participa do processo de produção e compreensão da linguagem, uma vez que a construção do sentido parece estar inextricavelmente relacionada ao acionamento de esquemas de percepção e de ação.

### **Simulações perceptuais e motoras**

Durante o exercício de imaginação do que seria um *boi voador*, fizemos mais do que simplesmente ativar imagens visuais previamente estocadas na memória de longo termo. Como vimos, procuramos simular detalhes da composição daquele ser fantástico. Da mesma forma, quando simulamos eventos simples, como girar uma maçaneta, estamos simulando a direção e a força a ser aplicada durante a execução da ação. Nesse sentido, somos capazes de reproduzir experiências de ver, fazer alguma coisa ou ouvir algum som (ou ruído) sem que necessariamente o estímulo esteja sendo captado pelo nosso campo visual e auditivo ou estejamos, de fato, executando uma determinação ação ou movimento.

Como a simulação mental é construída no cérebro por meio dos ecos de experiências prévias, isto é, de ressonâncias atenuadas dos padrões mentais ativados durante as experiências perceptuais e motoras reais, não é difícil deduzir que fazemos uso de simulações mentais numa grande quantidade de tarefas cognitivas diferentes. Algumas pesquisas (DRISKELL et al., 1994; WEINBERG, 2008) demonstraram, inclusive, que à medida que desportistas veem ações como chutes a gols ou arremessos de bola na cesta bem sucedidos, a simulação mental se torna mais precisa, contribuindo para o aperfeiçoamento dessas habilidades.

### **O papel da linguagem no refinamento das simulações mentais**

De acordo com Bergen e Chang (2005), compreendemos enunciados por meio da realização de simulações mentais, ou seja, acionamos nossos sistemas perceptuais e motores para criarmos experiências internas semelhantes àquelas vivenciadas caso estivéssemos experienciando as cenas e situações descritas de fato. Nesse sentido, acreditamos que, uma vez incorporada às perspectivas teóricas cognitivistas, a noção de simulação mental possa trazer importantes contribuições para os estudos das construções gramaticais, tendo em vista que a dimensão gramatical da linguagem parece orientar a maneira como as situações e acontecimentos devam ser simulados mentalmente.

Com essa mesma perspectiva de investigação, Bergen e Chang (2005) afirmam que, durante a leitura ou a audição de histórias, vamos simulando mentalmente as cenas e as ações que estejam sendo descritas. Graças aos recursos linguísticos, a simulação mental pode ser mais ou menos detalhada, como demonstraremos através do exemplo (03):

(03) Antes de enfiar um parafuso na parede, mergulhe-o em água fervente... (retirado de <<http://dicas-deconsertosemcasa.blogspot.com>>).

Alguns experimentos (AZIZ-ZADEH et al., 2006; BERGEN; WHEELER, 2010) evidenciaram que, para compreendermos enunciados, simulamos ações considerando informações específicas como orientação espacial, cores e formatos dos objetos envolvidos. Em (03), o parafuso, peça de cor metálica sulcada em espiral com um formato côncavo que

possui uma base superior adaptada para diversos tipos de chaves de fenda, encontra-se numa posição perpendicular à parede.

Para Bergen *et al.* (2007), o fato de a linguagem nos levar a realizar simulações mentais tão detalhadas tem relação com as próprias palavras. Segundo os autores, formas substantivas (como parafuso e parede) acionam simulações perceptuais de objetos e cenas, e formas verbais (como enfiar) acionam simulações mentais de ações motoras. No entanto, alertam-nos que isso é apenas parte da explicação, tendo em vista que para haver a simulação mental do conteúdo integral de uma sentença, como (03), os elementos perceptuais e motores, acionados pelas palavras, devem ser combinados de maneira apropriada (o parafuso foi enfiado na parede, não a parede no parafuso!). Tudo indica que, à medida que as palavras de uma sentença sinalizam os tipos de experiência que devem ser simuladas mentalmente (parafuso, parede etc.), a estrutura gramatical orienta o processo de compreensão quanto ao modo de organizar as experiências dentro da simulação mental, posicionando, por exemplo, as palavras em determinadas ordens. É por isso que simulamos cenas distintas quando nos deparamos com os exemplos (04) e (05).

(04) O montador enfiou o parafuso na parede.

(05) O parafuso enfiou o montador na parede.

As simulações mentais em (04) e (05) são diferentes, não por causa das palavras, que são as mesmas em ambas as sentenças (*montador*, *enfiou*, *parafuso* e *parede*), mas devido à ordem dos termos. Com isso, constatamos que a gramática desempenha uma função relevante para a compreensão da linguagem, a de arquitetar a simulação mental da sentença como um todo. No entanto, restringir o papel da gramática à mera ordenação de palavras numa sentença não constitui uma novidade nos estudos linguísticos. A tradição gramatical sempre atribuiu esse papel à gramática, definindo-a inclusive como o conjunto de regras e princípios que determinam a ordenação de palavras numa sentença. Na perspectiva da simulação mental, a gramática parece fazer bem mais do que simplesmente combinar palavras individuais em uma dada sequência ordenada. De acordo com Bergen (2007), a gramática:

1º – contribui fornecendo conteúdos para a simulação mental.

(06) Conca lançou a bola para Washington.

(07) O advogado passou o processo para o cliente.

As sentenças (06) e (07), por exemplo, compartilham a mesma organização gramatical, ou seja, em cada uma delas há um verbo (*lançou* e *passou*) seguido de um objeto direto (*bola* e *processo*) e um objeto indireto (*Washington* e *cliente*). Durante o processo de simulação mental das cenas descritas por essas sentenças, identificamos alguém transferindo um objeto para um recipiente: em (06), Conca agiu de modo a fazer com que Washington recebesse a bola e, em (07), o advogado agiu de modo a fazer com que o cliente recebesse o processo. Nos exemplos (06) e (07), é fácil perceber que a noção de *transferência de posse* não é acionada por palavras isoladas:

a) não é acionada pelo verbo, uma vez que há muitas formas de *lançar* e de *passar* sem que necessariamente haja algum tipo de transferência de coisas (*lançar um produto*,

*lançar um alerta, lançar a nota no diário de classe; passar roupas, passar de ano, passar por um lugar etc.);*

b) não é acionado pelos substantivos *Conca* e *advogado*, uma vez que estes podem executar outras tarefas que não tenham relação com a transferência de coisas;

Se a noção de transferência de posse não é acionada pelo verbo nem pelos substantivos, tudo indica que tal significado seja ativado pela estrutura gramatical da sentença como um todo. Nesse sentido, o esquema genérico *x faz com y receba z* é o verdadeiro responsável pela construção do sentido de transferência de posse.

2º – contribui indicando como a simulação mental deve ocorrer. Por exemplo, ao lermos ou ouvirmos uma sentença, somos induzidos a adotar uma perspectiva particular.

(08) Você está lendo um livro.

(09) O livro está sendo entregue a você.

(10) O livro está em cima da mesa.

Em (08) assumimos a perspectiva da pessoa que está realizando uma ação, mas, em (09), assumimos a perspectiva de uma pessoa para quem uma ação é realizada ou, em (10), a perspectiva de um observador externo. As palavras que integram cada uma dessas sentenças não são suficientes para revelar a perspectiva adotada. Parece que a construção gramatical é a verdadeira responsável por atribuir instruções sobre o modo de se realizar a simulação mental. Em (08) e (09), orientando-nos a adotar uma perspectiva de 1ª pessoa e, em (10), orientando-nos a adotar uma perspectiva de 3ª pessoa.

Vimos até aqui que a estrutura gramatical desempenha o papel de refinar o processo de simulação mental durante a produção e compreensão da linguagem. Tal especificação ocorre de três maneiras: por meio da ordenação dos elementos na sentença e por meio do fornecimento de conteúdo e de perspectivas para a simulação mental. Em seguida, demonstraremos como a gramática modela a simulação mental durante a produção e compreensão da linguagem.

## **O papel de uma gramática cognitiva baseada no uso**

Na seção anterior, demonstramos que a noção de simulação mental integrada a de gramática de construções fornece um aparato teórico que nos permite investigar restrições formais e semânticas em um único modelo de análise. A construção *x faz com y receba z*, subjacente aos enunciados apresentadas em (08) e (09) são especificadas com a restrição formal de que o *slot x* seja preenchido por uma expressão de referência e com a restrição semântica de que tal construção deva codificar (literalmente ou metaforicamente) a noção de transferência de posse. De forma mais detalhada, *x* e *z* acionam o esquema imagético CONTÊINER de *y*, seu conteúdo (antes e depois da transferência de posse, respectivamente). É possível identificarmos, ainda, um deslocamento de *y* (de *x* para *z*), manifesto pelo acionamento do esquema ORIGEM-CAMINHO-META, sendo *x*, a origem, e *z*, a meta.

De acordo com Bergen (2007), por abranger restrições de caráter formal e semântico, esse tipo de abordagem apresenta algumas vantagens para o pesquisador:

1º – acrescenta uma dimensão semântica à descrição gramatical e representa conhecimentos linguísticos (geral e específico) como construções, abrangendo ampla variedade de fenômenos linguísticos (normalmente ignorados por outras teorias de gramática) como: expressões idiomáticas, padrões morfológicos, construções de estrutura argumental e palavras simples;

2º – integra esquemas de forma e de sentido numa única arquitetura de gramática, potencializando o poder explicativo da teoria, uma vez que não se limita a uma explanação exclusivamente formal. Esse modelo toma por base o sentido e a experiência corporificada, ativados e acionados pelos elementos formais da gramática.

Ainda no que diz respeito à construção do sentido, a abordagem construcional apresentada até aqui exige uma noção coerente com a de simulação mental, isto é, a de que os mesmos mecanismos cognitivos que as pessoas acionam durante o movimento e a percepção são também utilizados para simular ações e eventos respectivamente. A simulação mental não é, portanto, um conjunto ilimitado de possibilidades. Segundo Bergen e Chang (2005), os dados linguísticos fornecem possibilidades limitadas de ativação e acionamento de esquemas à simulação mental.

Nesse sentido, para Bergen et al. (2003), Bergen e Chang (2004), Bergen *et al.* (2004) e Feldman e Narayanan (2004), compreender uma pequena porção da linguagem é o suficiente para desencadear a realização de simulações mentais de percepção e de movimento. Isso quer dizer que os significados das palavras e de suas configurações gramaticais são precisamente as contribuições que esses elementos linguísticos fornecem para a constituição de simulações mentais. O estudo sobre como os diferentes aspectos da linguagem podem contribuir para a construção de imagens e a teoria do significado linguístico enquanto um conjunto de especificações linguísticas sobre o que e como simular constituem a semântica da simulação.

Durante a compreensão, os esquemas servem como meio de as palavras ativarem os componentes perceptuais e motores da simulação. Os esquemas mais simples projetam *gestalts* baseadas na experiência, no domínio da forma ou do sentido, como atesta o exemplo (11).

(11) Maria acabou de sair.

Em (11), a construção de movimento intransitivo *x mover-se para y* aciona os esquemas imagéticos ORIGEM-CAMINHO-META (Maria sai de algum lugar com destino a outro) e CONTÊINER (o lugar de onde Maria saiu e o lugar para onde Maria vai), embora não haja pistas explícitas linguisticamente acerca do ponto de origem e de chegada do trajetor (Maria).

### **O pareamento de esquemas de forma e de sentido: a construção gramatical**

O domínio canônico da forma linguística é o som. Os padrões de sons mais comuns podem ser descritos em termos de sequências de fonemas, tons e acentos, ou mais concretamente como os padrões acústicos e os movimentos articulatórios que os produzem. A noção de forma, num sentido mais amplo, no entanto, pode se estender a qualquer tipo de significante, incluindo a escrita, os gestos e as expressões faciais, os ícones pictóricos e os padrões táteis.

Como vimos anteriormente, as relações entre as formas podem variar em relação à ordem dos segmentos.

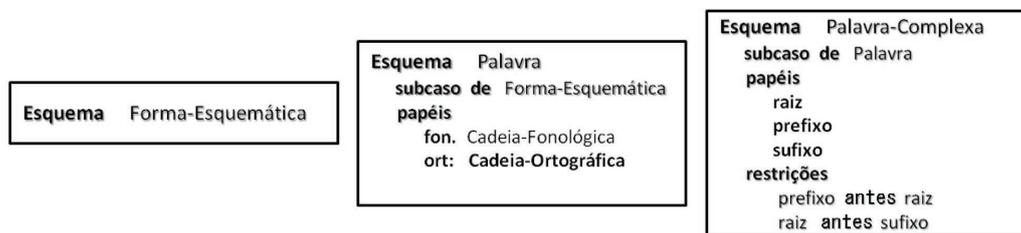


Figura 1: Esquemas formais

Os quadros da figura 1 representam graficamente três esquemas, do mais genérico (mais abstrato) ao menos genérico. O esquema *Palavra* é um subcaso (simplificado) do esquema *Forma-Esquemática* e apresenta papéis indicativos de suas possibilidades de realização (por meio de cadeias fonológicas ou gráficas). Já o esquema *Palavra-Complexa* é herdado (“subcaso de”) do esquema *Palavra* e apresenta, além de papéis relacionados à sua estruturação, restrições formais relacionadas à ordem dos segmentos que o constituem.

O esquema do significado, por sua vez, engloba todas as noções conceptualizáveis, desde entidades concretas, categorias de relação e ação a categorias mais abstratas, usualmente associadas à gramática. Abaixo, apresentamos esquemas de categorias concretas, distribuídos em (a), indivíduos com nomes específicos, em (b), esquemas de ação, em (c) e alguns esquemas imagéticos como TRAJETOR-MARCO, CONTÊINER e OCM (ORIGEM-CAMINHO-META), em (d).

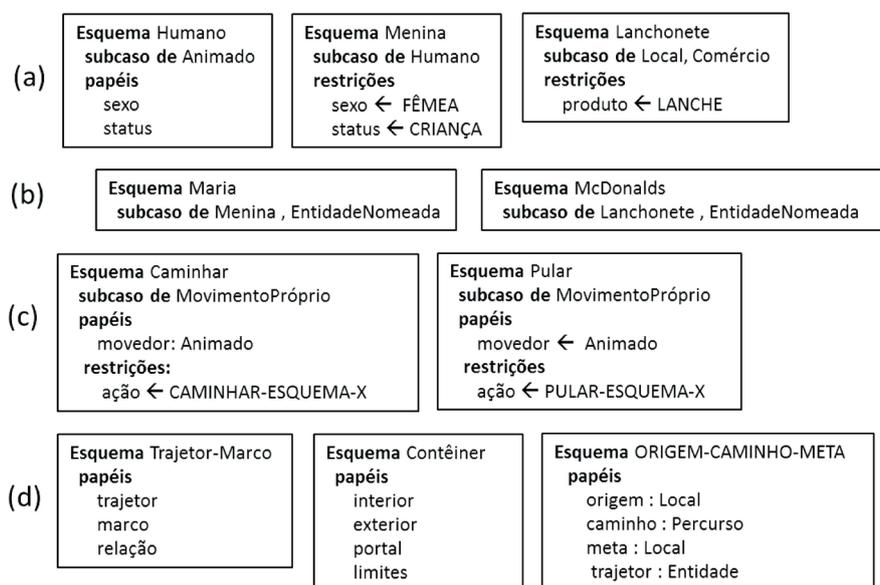
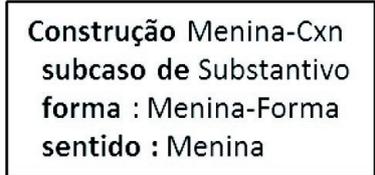


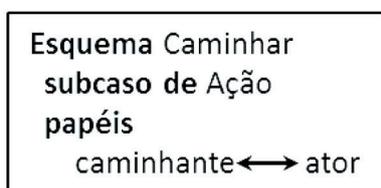
Figura 2: Esquemas de significado: (a) entidades; (b) entidades nomeadas; (c) ações; (d) esquemas de imagens

Os esquemas de forma e de sentido são vinculados por *construções* (Cxn). Nesse sentido, a representação gráfica das construções deve explicitar os esquemas da forma e do sentido, como ilustra a figura 3.



**Figura 3: Construção “Menina”**

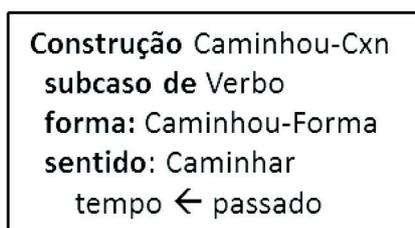
Há esquemas e construções mais complexos que os apresentados nas figuras 1, 2 e 3. Preposições, verbos e adjetivos, cujos significados são computados junto com os de outras palavras envolvem estruturas múltiplas, cujos componentes internos, muitas vezes, são fornecidos pelo contexto linguístico, discursivo ou situacional durante o uso da linguagem. A palavra “caminhou”, por exemplo, ativa e aciona entidades animadas que executam a ação de caminhar, como ilustra a representação gráfica abaixo (figura 4).



**Figura 4: Esquema “Caminhar”**

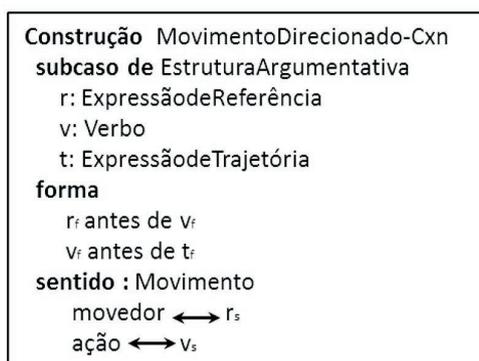
Na representação gráfica, verificamos que o esquema *caminhar* é um subcaso (herança) do esquema *ação*, categoria conceptual de eventos cujos agentes animados executam ações intencionalmente. O esquema é definido como apresentando um participante especificado como o *caminhante* (na área reservada aos papéis). O *caminhante* é identificado como o “ator” (herdado do esquema de ação), o que é indicado pela flecha de duas pontas.

Na representação gráfica, a construção “caminhou” é especificada como o pareamento de uma forma específica com um esquema de sentido. Tal construção é um subcaso (herança) da construção verbo (figura 5).



**Figura 5: Construção “caminhou”**

A construção *caminhou* herda o papel tempo da construção *verbo*. Tal restrição é representada pela seta de uma ponta. Exemplos desse tipo apenas fornecem uma visão geral de como construções e esquemas se inter-relacionam com outras construções e esquemas por meio de papéis e subcasos (relações de herança). O que distingue construções gramaticais complexas de palavras simples, no entanto, é o fato de aquelas exibirem uma estrutura interna constituída por constituintes múltiplos, cada qual instanciando uma projeção própria de forma e sentido. Do ponto de vista da simulação, isso significa que as construções gramaticais contribuem com o acionamento de esquemas conceptuais variados para a modelagem das simulações mentais.



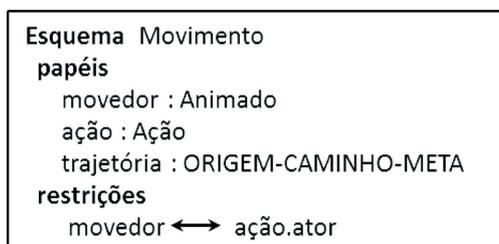
**Figura 6: Construção de Movimento Direcionado**

A construção representada graficamente na figura 6 é uma construção de estrutura argumental que licencia sentenças como (12) e (13).

(12) O gato correu para debaixo da mesa.

(13) O menino pulou para baixo.

Em (12) e (13), um *movedor* (designado pelas Expressões de Referência – *gato* e *menino*) se desloca ao longo de uma trajetória (designada pela Expressão de Trajetória – *para debaixo da mesa* e *para baixo*) por meio de alguma ação (designada por um verbo – *correu* e *pulou*). O polo da forma dessa *Construção de Movimento Direcionado* (CxMD) estabelece que a *Expressão de Referência* deve preceder o *Verbo* que, por sua vez, deve preceder a *Expressão de Trajetória*. No polo do significado, a CxMD denota um movimento especificado pelo esquema *Movimento* (representado graficamente na figura 7).



**Figura 7: Esquema de movimento**

Na CxMD, os papéis do Esquema de Movimento são explicitamente vinculados à área correspondente aos polos de sentido dos constituintes da Expressão de Referência, do Verbo e da Expressão de Trajetória. Em (12), por exemplo, o sentido da Expressão de Referência *gato* é restrito ao papel de *movedor* do Esquema de Movimento.

Uma construção gramatical pode alinhar diferentes aspectos dos sentidos de seus constituintes, reunindo contribuições diversas para a simulação mental. Por exemplo, uma construção de estrutura argumental pode designar um desses constituintes (por exemplo, a Expressão de Referência) como o *movedor* envolvido no movimento denotado pelo verbo ao longo de uma trajetória (denotada pela Expressão de Trajetória). Isso possibilita a identificação dos movimentos dentro da construção que podem afetar as simulações mentais.

Uma construção gramatical fornece conteúdo à simulação mental, por acionar categorias de experiências que ajudam na representação dos eventos descritos. Por exemplo,

uma CxMD pode induzir leitores ou ouvintes a simularem eventos como se esses envolvessem um *movedor* se deslocando ao longo de uma *trajetória*, mesmo nos casos em que o verbo não especifica explicitamente deslocamento, como em (14).

(14) O gato miou estrada abaixo.

Além disso, as construções gramaticais podem apenas modular as propriedades da simulação, fornecendo uma perspectiva ou um ponto a ser focalizado durante a compreensão. Por exemplo, as sentenças ativas devem levar o ouvinte ou leitor a simular um evento da perspectiva perceptual ou motora do agente, enquanto uma sentença passiva deve levar à simulação da perspectiva do paciente.

De acordo com a perspectiva apresentada neste artigo, a gramática contribui para a simulação mental por criar vínculos entre os esquemas conceptuais e esquemas formas e por impor restrições de segunda ordem sobre como a simulação deve ser realizada. Dessa forma, acreditamos que uma teoria cognitivista que se proponha a investigar como as pessoas usam a gramática exige uma análise bastante cuidadosa de como as especificidades da construção podem afetar o processo de simulação mental.

### ***Affordances* e gramática**

De acordo com Kaschak e Glenberg (2000), em sua *Hipótese Indexical*, a compreensão de enunciados envolve processos de indexação, derivação de *affordances* e combinação de restrições fornecidas pelos *affordances* e pela gramática. Acrescentamos a noção de *frame* a esses processos:

- a) A indexação de palavras e expressões linguísticas a representações mentais dos referentes descritos;
- b) A assimilação dos *affordances*<sup>2</sup> desses referentes junto com o acionamento de *frames*<sup>3</sup> contextuais;
- c) A combinação (“meshing”, p. 508) de restrições físicas e biológicas fornecidas pelos *affordances*, de aspectos focalizados do *frame* e de restrições fornecidas pela gramática.

Dos três processos usados na compreensão da linguagem, o primeiro é a indexação de palavras e sintagmas: aos referentes no ambiente e às representações mentais análogas. Para os autores, a indexação estabelece o conteúdo da linguagem: sobre quem ou o que se está falando. Como evidência para esse processo, Glenberg e Robertson (1999) demonstraram que a oportunidade de indexar instruções às ações apropriadas é um fator crítico na determinação de quão bem os participantes adquirem os conhecimentos necessários para realizar uma tarefa (p.ex. aprender a usar uma bússola, aprender a andar de bicicleta etc.).

---

2De acordo com Gibson (1979), um *affordance* é uma possibilidade de ação disponível no ambiente, independente da capacidade do indivíduo de perceber essa possibilidade.

3Segundo Minsky (1974), um *frame* é uma estrutura de dados para representar algum tipo de ambiente ou uma situação estereotipada, tal como estar em uma cozinha, ou ir para um restaurante. Junto com cada *frame* estão vários tipos de informação: algumas são sobre como usar o *frame*, algumas são sobre o que se espera que aconteça em decorrência do uso do *frame*, e outras são sobre o que fazer se alguma dessas expectativas não for confirmada.

O segundo processo é a derivação de *affordances* a partir dos referentes já indexados. Segundo os autores, o termo *affordance* se refere aos meios como indivíduos podem interagir com coisas em seus ambientes. Eles ilustram o conceito de *affordance* com as possíveis interações com o objeto *muleta* (*crutch*). Segundo os autores, uma pessoa pode interagir com uma muleta de modos particulares: o objeto pode ser usado por alguém ferido, para golpear alguma coisa ou empurrar um objeto através de uma fenda estreita aberta na parede. Todas as possibilidades de interação extraídas de objetos particulares refletem as maneiras como podemos interagir e os objetivos que adotamos em relação a esses objetos, num momento particular. Se a pessoa tem uma perna ferida, por exemplo, o *affordance* “auxiliar na caminhada” vai ser extraído mais facilmente, ao passo que, se a pessoa quer empurrar alguma coisa através de uma fenda estreita, o *affordance* “empurrar um objeto” vai ser extraído mais facilmente. De acordo com Gibson (1979), o *affordance* tem uma relação estreita com as capacidades de ação de atores particulares. Nesse sentido, uma escada pode apresentar um *affordance* de “escalar” (e, subseqüentemente, chegar ao andar superior, por exemplo), mas pode simplesmente apresentar um *affordance* de “sentar”, para uma pessoa cansada. No entanto, para o autor, *affordances* não devem ser vistos como formas de interação alicerçadas na cultura, em conhecimento prévio ou nas expectativas do indivíduo. Nesse sentido, ao desconsiderar o conhecimento prévio e as expectativas do ator, Gibson contempla apenas as capacidades motoras do ator e não as capacidades perceptivas e mentais. Vale enfatizar que, ao contrário disso, acreditamos que as informações que especificam o *affordance* dependem das experiências do indivíduo e da cultura em que está inserido.

O terceiro e último processo diz respeito à combinação de *affordances* a partir das restrições impostas pelos próprios objetos e as restrições impostas pela construção gramatical. Esse processo é responsável pela combinação dos *affordances* originando padrões de ação que visam atingir objetivos determinados. A estrutura do objeto muleta permite tocar ou empurrar objetos, por exemplo. Esses movimentos (tocar e empurrar) podem ser combinados com outras ações, tais como direcionar uma maçã para a entrada de uma fenda e empurrá-la através da passagem. Nesse sentido, ao combinarmos os *affordances* da *muleta*, da *maçã* e da *fenda*, somos capazes simular mentalmente com precisão o conteúdo de uma sentença como (15).

(15) Lisa empurrou a maçã através de uma fenda usando uma muleta.

Os três processos apresentados acima ocorrem de forma dinâmica e ao mesmo tempo, não de forma algorítmica.<sup>4</sup> Ao lermos o exemplo (15), os referentes para *Lisa*, *muleta*, *maçã* etc, são indexados e usados para configurar um modelo mental. Quando os *affordances* são extraídos de *Lisa* e *maçã*, o processo de combinação tem início. À medida que a sentença é lida/ouvida, os *affordances* vão sendo combinados em uma simulação mental (BARSALOU, 1999) do evento retratado na sentença. A construção gramatical vai atuar em todos os estágios tanto para ajudar no processo de indexação (p.ex. formando sintagmas nominais) quanto para fornecer limitações sobre como a simulação mental deve operar. Alterações em qualquer um desses passos, como uma mudança repentina na estrutura sintática da sentença, por exemplo, resultaria numa alteração da simulação a fim de acomodar essas mudanças. Cumpre ressaltar que, no discurso, as representações para *Lisa*, *muleta*, etc., podem ser indexadas em sentenças anteriores. Isso não ocorre, porém,

4 Um algoritmo é uma sequência finita de instruções que devem ser cumpridas uma após a outra.

se os *affordances* não puderem ser combinados da forma orientada pela sintaxe (p.ex. empurrar uma maçã com um fio).

## Conclusão

A proposta apresentada neste artigo, de integrar o conceito de construções gramaticais com o de simulação mental e *affordances*, parte do pressuposto de que o processo cognitivo de produção e compreensão da linguagem utiliza os sistemas perceptuais e motores do cérebro. Essa pressuposição repercute diretamente na maneira de abordarmos a noção de sentido e a maneira como ele é construído. Ele passa a ser estruturado no cérebro concretamente por meio de representações específicas de percepção e ação (em vez de meros símbolos abstratos). De acordo com esse enfoque, ao invés de a linguagem ser concebida como um módulo separado da mente, ela é analisada em conjunto com os sistemas cognitivos dedicados à percepção e ação.

Do ponto de vista prático, as pesquisas sobre os efeitos da gramática na simulação mental podem apresentar novas perspectivas às abordagens de aquisição e de uso da linguagem natural. Quanto aos modelos computacionais de processamento da linguagem, sabemos que o seu sucesso depende de entendermos de que forma ocorre a produção e a compreensão da linguagem por humanos. Além disso, a compreensão de como a gramática afeta a simulação mental pode ser a chave para as abordagens que enfoquem a relação entre linguagem e cognição. Defendemos que o sistema sensorio-motor desempenha um papel crucial na organização de funções cognitivas superiores, mostrando que a gramática, capacidade cognitiva altamente abstrata presente exclusivamente no homem, conecta-se com sistemas neurais preexistentes dedicados à percepção e à ação.

## REFERÊNCIAS

AZIZ-ZADEH, L.; WILSON, S.; RIZZOLATTI, G.; IACOBONI, M. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, Cell Press, v. 16, p. 1818-1823, 2006.

BARSALOU, L. Perceptual symbol systems. *Behavioral and Brain Sciences*, v. 22, p. 577-609, 1999.

BECIC, E.; DELL, G.; BOCK, K.; GARNSEY, S.; KUBOSE, T. Driving impairs talking. *Psychonomic Bulletin & Review*, University of Illinois at Urbana-Champaign, USA, v. 17, n. 1, p. 15-21, 2010.

BERGEN, B. Experimental methods for simulation semantics. In: GONZALEZ-MARQUEZ, M.; MITTELBERG, I.; COULSON, S.; SPIVEY, M. J. (Ed.). *Methods in cognitive linguistics*. Amsterdam/Philadelphia: John Benjamins, 2007. p. 277-301.

BERGEN, B.; CHANG, N. *Embodied construction grammar in simulation-based language understanding*. In: ÖSTMAN, J.-O.; FRIED, M. (Ed.). *Construction grammar(s): Cognitive grounding and theoretical extensions*. Amsterdam/Philadelphia: John Benjamins, 2005. p. 147-190.

BERGEN, B.; CHANG, N.; NARAYAN, S. Simulated action in an embodied construction grammar. In: ANNUAL CONFERENCE OF THE COGNITIVE SCIENCE SOCIETY MAHWA, 26. *Proceedings...* NJ: Erlbaum, 2004. p. 108-113.

BERGEN, B.; HEEYEO, D. Language-driven motor simulation is sensitive to social context. In: ANNUAL MEETING OF THE COGNITIVE SCIENCE SOCIETY, 32, Mahwa. *Proceedings...* NJ: Erlbaum, 2010.

BERGEN, B.; LINDSAY, S., MATLOCK, T., NARAYANAN, S. Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, n. 31, v. 5, p. 733-764, 2007.

BERGEN, B.; NARAYAN, S.; FELDMAN, J. Embodied verbal semantics: evidence from an image-verbmatching task. In: *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum, 2003, p. 139-144.

BERGEN, B.; WHEELER, K. Grammatical aspect and mental simulation. *Brain & Language*, v. 112, p. 150-158, 2010.

BONVILLIAN, J.; RAEBURN, V.; HORAN, E. Talking to children: the effects of rate, intonation, and length on children's sentence imitation. *Journal of Child Language*, v. 6, p. 459-467, 1979.

DRISKELL, J.; COPPER, C.; MORAN, A. Does mental practice enhance performance? *Journal of Applied Psychology*, v. 79, p. 481-492, 1994.

FAUCONNIER, G.; TURNER, M. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books, 2002.

FELDMAN, J. *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, Mass: MIT Press, 2008.

FELDMAN, J.; NARAYANAN, S. Embodied Meaning in a Neural Theory of Language. *Brain and Language*, Elsevier Press, v. 89, p. 385-392, 2004.

GIBSON, J. J. *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

GLENBERG, A. M.; ROBERTSON, D. A. Indexical understanding of instructions. *Discourse Processes*, v. 28, p. 1-6, 1999.

JIN, S. *The Effect of Driver Cognitive Abilities and Distraction on Situation Awareness and Performance under Hazard Conditions*. (Dissertation under the direction of Dr. David B. Kaber), North Carolina State University, 2008.

JOHNSON, M. *The body in the mind: the bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press, 1987.

KASCHAK, M. P.; GLENBERG, A. M. Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, v. 43, p. 508-529, 2000.

KOSSLYN, S.; GANIS, G.; THOMPSON, W. Neural foundations of imagery. *Nature Reviews, Neuroscience*, v.2, p. 635-642, December, 2001.

LAKOFF, G. *Women, fire, and dangerous things*. Chicago: University of Chicago Press, 1987.

MINSKY, M. L. A framework for representing knowledge. *MIT AI Memo 306*, 1974.

WEINBERG, R. Does imagery work? Effects on performance and mental skills. *Journal of Imagery Research in Sport and Physical Activity*, v. 3, n. 1, p. 1-21; 2008.

YAXLEY, R.; ZWAAN, R. Simulating visibility during language comprehension, *Cognition*, v. 105, n. 1, p. 229-236, out. 2006.

# Condições de opcionalidade do objeto direto no português do Brasil<sup>1</sup>

(Optionality conditions of the direct object in Brazilian Portuguese)

Maria Madalena Loredo Neta<sup>1</sup>

<sup>1</sup>Programa de Pós-graduação em Estudos Linguísticos da Faculdade de Letras – Universidade Federal de Minas Gerais (UFMG)

madalena@loredo.com.br

**Abstract:** Adopting a descriptive approach, this study focuses on what traditionally is called the direct object of Brazilian Portuguese, especially the conditions of its omission. The possibility of occurrence of a NP as an object divides the verbs between those that admit and those that do not admit direct object, hence the importance of these conditions for the formulation of verb valencies. In this paper, the analysis is limited to simple sentences; anaphoric situations are not considered as the omission of the direct object in anaphoric situation applies to any verb annulling subclassification.

**Keywords:** direct object; direct object omission; verbal valences; Brazilian Portuguese.

**Resumo:** Dentro de uma abordagem descritiva, este estudo se concentra no que se chama tradicionalmente objeto direto no português do Brasil, especialmente nas condições de sua omissão. A ocorrência de SNs na função de objeto divide os verbos entre os que admitem e os que não admitem objeto direto, daí a relevância dessa função sintática na formulação das valências verbais. Esta análise irá se limitar a orações simples, e situações anafóricas não são consideradas, pois a omissão do objeto direto em situação anafórica vale para qualquer verbo, de modo que a subclassificação fica anulada.

**Palavras-chave:** objeto direto; omissão do objeto direto; valências verbais; português do Brasil.

## Introdução

Dentro de uma abordagem descritiva, este estudo se concentra no termo que se chama tradicionalmente objeto direto, mais especificamente nas condições de sua omissão. A definição tradicional é adotada porque ainda não dispomos de uma conceituação realmente satisfatória do que vem a ser objeto direto.

O tema “objeto” tem ocupado espaço expressivo nos estudos linguísticos, em geral sugerindo a conclusão de que há uma dificuldade (ou impossibilidade) de generalizações universais a respeito do conjunto de propriedades comumente atribuídas ao SN que ocupa essa função. Mencionamos aqui a antologia *Objects*, editada por Plank (1984), composta por diversos artigos, nenhum chegando a uma conceituação satisfatória de objeto. Muitos teóricos concluem que definir objetividade (*objecthood*) não é tarefa simples até mesmo com relação a uma língua particular. Beth Levin (1999) afirma que a noção de objeto continua a representar um desafio para a teoria linguística.

A partir do modelo de construções – o pareamento forma-significado – já formulado por Goldberg (1995) e ampliado e adaptado ao português brasileiro por Perini (2008a),

<sup>1</sup> Este trabalho é parte de minha pesquisa de Doutorado, desenvolvida sob orientação do Prof. Dr. Mário Alberto Perini.

propõe-se um inventário das diáteses verbais. Cada verbo é associado a um conjunto de construções nas quais ele pode ocorrer; e cada construção na qual um verbo pode ocorrer e outros verbos não podem é uma diátese. A diátese é, assim, uma construção que subcategoriza os verbos, pois há construções nas quais certos verbos não podem ocorrer. A valência de um verbo é o conjunto de suas diáteses. O objetivo da formulação de diáteses e da associação de cada uma delas a um conjunto de verbos é caracterizar as suas propriedades gramaticais – o que cada verbo contribui para a sintaxe e a semântica das orações de que participa – e assim tornar possível a elaboração de um dicionário de valências dos verbos do Português.

A relevância da função sintática “objeto direto” na formulação das valências verbais está em que a ocorrência de SN nessa função divide os verbos entre os que admitem e os que não admitem objeto direto: aparentemente, há verbos que recusam esse complemento, há os que permitem a sua omissão, enquanto há verbos que exigem a sua presença. Vejamos alguns exemplos. Tomando-se o verbo *comer*, observamos que ele ocorre nas seguintes construções:

(01) José comeu o bombom.  
 SN V SN  
 Agente Paciente

(02) José comeu.  
 SN V Ø  
 Agente Paciente

(03) O bombom, José comeu.  
 SN SN V  
 Paciente Agente

Essas são construções em que o verbo *comer* pode ocorrer. Porém (03) não é uma diátese desse verbo, pois qualquer verbo que admita objeto também admite o objeto topicalizado. Assim, as diáteses representadas em (01) e (02) pertencem à valência do verbo *comer*. Por outro lado, a construção ergativa não se inclui na valência do verbo *comer*, uma vez que não se constrói *O garoto comeu* com o SN *O garoto* no papel de Paciente; porém com *engordar* a construção ocorre e faz parte da valência deste verbo:

(04) O garoto engordou.  
 SN V  
 Paciente

Como se vê, as diáteses são representadas por uma análise sintática sumária, mais os papéis temáticos dos termos relevantes. Essa notação é a proposta por Perini (2008a), e será utilizada aqui.

A oração simples será tomada como base de nossa análise: frases com sujeito mais um SN (como em *José comeu o bombom*); com sujeito mais dois SNs (*Eu acho Maria uma ótima colega*); ou, ainda, sem sujeito, com verbo seguido de um ou dois SNs, (*Comi o bombom*; *Acho Maria uma boa pessoa*). A seguir vamos discutir como definir o objeto direto e analisar as condições de sua omissão.

## Definir sujeito para definir objeto direto

Procuraremos mostrar que o processo de identificação do objeto se dá em conexão com o de identificação do sujeito, o que nos leva a caracterizar o objeto praticamente por sua posição em relação ao verbo, mais o fato de que não se trata do sujeito. Em Português, o sujeito vem em geral imediatamente antes do verbo, na ordem SVO, como em frases *Maria assou os bolos*; e, às vezes, após o verbo em orações como *Chegaram as encomendas*. O sujeito participa da concordância verbal, enquanto o objeto não participa. A concordância se dá numa correlação formal entre pessoa e número do SN a que chamamos sujeito e o sufixo de pessoa-número do verbo. Por exemplo, tomando-se o verbo *assar*, teremos a seguinte compatibilidade:<sup>2</sup>

- *Eu*: *asso, assei, assarei...* : Eu *assei* o bolo.
- *Nós, eu* mais *SN*: *assamos, assávamos, assaremos...* : Maria e eu *assamos* o bolo.
- *Qualquer outro SN*: *assa, assou...* : Maria *assou* o bolo./ Ela *assou* o bolo.

A valência do verbo *assar* inclui um Agente e um Paciente. Pelo sufixo verbal *-ou*, o falante identifica um referente de 3ª pessoa do singular – *ela/Maria* – que indica aquele que receberá o papel temático (nesse caso, Agente); nessa frase, o SN *os bolos* é incompatível com o sufixo verbal *-ou* e não pode ser o sujeito. Em *Os bolos assaram*, o SN *Os bolos*, compatível com o sufixo verbal *-am*, poderá ser o sujeito, de uma construção ergativa,<sup>3</sup> na qual esse SN – na função de sujeito – recebe o papel temático de Paciente.

Para que o receptor possa relacionar corretamente o SN ao sufixo de pessoa-número do verbo, a ordem dos sintagmas é fundamental: quando há mais de um SN na frase, o sintagma que se relaciona com o sufixo pessoa-número do verbo é o que vem imediatamente antes do verbo. A esse sintagma chamamos sujeito. Assim, para a identificação do sujeito, o usuário o faz com base na sequência formal que ouve ou lê e nos seus conhecimentos da ordem dos sintagmas na oração.

O usuário do Português sabe que a valência do verbo *matar* compreende um Agente (o matador) e um Paciente (o que morre) e sabe ainda que o Agente é representado pelo SN que vem antes do verbo. Isso é fundamental para distinguir as frases *A onça matou o caçador* de *O caçador matou a onça*, em que apenas o sufixo pessoa-número do verbo não é suficiente para a identificação do sujeito, e a ordem dos elementos é decisiva. Aqui é tomada a regra de identificação do sujeito elaborada por Perini:

Condição prévia: o sujeito é um SN cuja pessoa e número sejam compatíveis com a pessoa e número indicados pelo sufixo de pessoa-número do verbo.

- (i) Se na oração houver um SN nessas condições, esse SN é o sujeito.
- (ii) Se houver mais de um SN, então o sujeito é o SN que precede imediatamente o verbo.
- (iii) Mas se o SN em questão for um clítico (*me, te, nos, se*), ele não conta e o sujeito é o SN precedente. (2008a, p. 108)

Submetamos alguns exemplos à regra de identificação:

<sup>2</sup> As formas *tu* e *vós* não foram consideradas por não serem usadas regularmente.

<sup>3</sup> Essa construção é chamada incoativa ou inacusativa por alguns autores.

(05) Costurei a bolsa.

Na frase (05) não há um SN pré-verbal, e é preciso procurar a indicação do referente que pode receber o papel temático indicado no sufixo pessoa-número do verbo *-ei*. O sufixo indica uma compatibilidade com a 1ª pessoa do singular, a forma *eu*, que não está na frase, não há a presença concreta de um SN antes do verbo. Não foi atendida, assim, a condição prévia de identificação do sujeito, portanto o SN presente na oração – *a bolsa* – não é o sujeito; essa é uma oração sem sujeito; porém com um Agente, sinalizado pelo sufixo verbal *-ei*: eu.

Já em:

(06) Elias costurou a bolsa.,

o SN *Elias* elabora, redundantemente, o referente que recebe o papel temático compatível com aquele da pessoa e número do sufixo verbal, atendendo à condição prévia da regra de identificação do sujeito; a alínea (ii) dita que, se há dois SNs nessa condição, o sujeito é o SN que precede imediatamente o verbo. Portanto, *Elias* é o sujeito da oração. Essas mesmas condições é que nos autorizam a analisar igualmente *Elias* como sujeito na frase seguinte:

(07) A bolsa, Elias costurou.

Quando houver apenas um SN na oração, como em:

(08) Elias costura.

(09) Chegaram as encomendas.

Em (08) e (09) a condição prévia e a alínea (i) são atendidas, isto é, apenas um SN presente na frase, de pessoa e número compatíveis com a pessoa e número do sufixo verbal; portanto os SNs *Elias* e *as encomendas* são os sujeitos dessas orações. Essas condições contemplam orações com sujeito posposto como (09) e outras, por exemplo: *Sumiu um documento; Aconteceram coisas horríveis; Ocorreu um engavetamento há pouco na BR-381*.

Perini, na obra *Describing verb valency: practical and theoretical issues*, considera que a distinção de um SN pós-verbal em orações como *Maria é minha professora*, analisado na gramática tradicional como predicativo, é irrelevante sintaticamente, e o SN *minha professora* seria também um objeto direto, isto é, um SN pós-verbal não sujeito.

[...] o complemento tradicionalmente chamado ‘predicativo’ não é uma função sintática autônoma, e pode ser analisado como um simples SN não sujeito [...]. Cada característica tradicionalmente apresentada para distingui-lo do objeto direto, ou a partir do adjunto adverbial ou complemento, é mais adequadamente descrita como o efeito de traços semânticos das construções ou itens lexicais envolvidos. (PERINI, em fase de elaboração. Tradução e grifo nossos)<sup>4</sup>

4 No original: “[...] the complement traditionally called “predicative” is not an autonomous syntactic function, and may be analyzed simply as a nonsubject NP [...]. Every characteristic traditionally presented that would distinguish it from the direct object, or from the adverbial adjunct or complement, is more adequately described as the effect of semantic features of the constructions or lexical items involved.

A regra de identificação do sujeito ainda precisa ser refinada. Há situações em que a identificação do sujeito é problemática, como nos casos que envolvem a polissemia da 3ª pessoa do plural, que tem um comportamento próprio e pode denotar um sujeito esquemático (indeterminado). Frases como *Encheram os tanques* é ambígua, pode significar que alguém encheu os tanques, o SN *os tanques* na função de objeto direto e com papel temático de Paciente; ou, ainda, que os tanques encheram, construção ergativa, o SN *os tanques* no papel temático de Paciente, na função sintática de sujeito. Segundo a regra – condição prévia e (i) – “*Se na oração houver um SN nessas condições [compatibilidade entre o SN e o sufixo verbal], esse SN é o sujeito.*”; assim, em frases como *Comeram os docinhos*, o SN *os docinhos* teria que ser o sujeito da mesma forma que o SN *as encomendas* é analisado como sujeito em *Chegaram as encomendas*; e não é o que ocorre, pois na primeira frase o SN *os docinhos* é analisado como objeto direto. Esse ponto ainda requer maior investigação empírica ao longo da pesquisa e revisão da regra de identificação. Parece que a semântica do verbo – de mudança de estado externamente causada, por exemplo, – estaria determinando essa questão, o que tentaremos comprovar com nossos levantamentos ao longo do estudo.

### **Definindo objeto direto**

Para alcançar o objetivo deste estudo – determinar as condições de omissão do objeto direto – procuramos chegar a uma definição de “objeto (direto)” no Português do Brasil, respeitando estritamente a separação forma-significado. Isto é, estamos considerando o objeto uma função sintática, a ser definida em termos exclusivamente formais.

A partir dessa definição sintática, torna-se possível fazer afirmações importantes de ordem simbólica. Por exemplo, em Português um objeto pode ser Paciente, mas nunca Agente, ao contrário do sujeito, que pode receber ambos os papéis temáticos.

O que chamamos tradicionalmente objeto direto é sempre representado por um SN e em geral ocorre na mesma posição na frase; além do fato de não ser o sujeito. O objeto pode às vezes vir antes do verbo, quando for um clítico, *Maria me viu na praça*; ou quando topicalizado, em frases como *Os bolos, Maria assou*. Chegamos a essa resposta pela regra de identificação do sujeito, vista anteriormente.

### **A omissão do objeto direto**

Para este estudo da omissão do objeto direto, situações anafóricas não são consideradas, pois nessas situações as lacunas deixadas pelo complemento omitido podem ser preenchidas para qualquer verbo; e uma construção não é diátese quando qualquer verbo pode ocorrer nela. Nos exemplos a seguir, o objeto omitido – representado por  $\emptyset$  – é recuperável anafóricamente; e inúmeros verbos poderiam formar uma frase colocando-se no lugar de *ler* outros verbos como *levar, emprestar, pedir, devolver, escrever, rasgar, comprar* etc.:

(10) João não leu o livro hoje. E vai ler [ $\emptyset$ ] depois.

$\emptyset$  representa o SN *o livro*, numa recuperação anafórica.

Assim, será considerada omissão o que o falante não produziu formalmente e que não pode ser deduzido por retomada do seu texto, sendo, entretanto, acessível à intros-

pecção. Isto é, este estudo não inclui os objetos suprimidos que já foram introduzidos no discurso prévio. Tratamos, portanto, da omissão de complementos não colocados em discurso antecedente; que não podem ser recuperados anaforicamente.

As orações em que ocorre o objeto direto ou nas quais o objeto foi omitido serão as focalizadas em nosso estudo, como nos exemplos a seguir. O objeto omitido será representado pelo símbolo Ø.

(11) Elias costura o boné.  
SN      V      SN  
Agente              Paciente

(12) Elias costura.  
SN      V      Ø  
Agente              Paciente

O objetivo deste estudo é determinar que fatores controlam a omissão do objeto (não anafórico) e quais as consequências disso para a determinação das diáteses. A possibilidade de omissão do objeto é governada pelo verbo (enquanto item lexical) ou por fatores independentes do verbo (fatores semânticos ou pragmáticos mais gerais)? Quando ouve a frase *Eu já comi*, o ouvinte consegue preencher o objeto suprimido por um Paciente esquemático (algo que se pode comer); já em *\*Eu já tirei da estante*, a omissão do SN objeto direto gera uma frase agramatical; não é possível, nesse caso, omitir o objeto, a não ser em contexto anafórico que, como foi dito, não será considerado neste estudo. O que determina essa diferença de comportamento dos verbos chamados transitivos, quanto à presença/ausência do complemento chamado objeto direto?

A ocorrência de SN na função de objeto divide os verbos entre os que admitem e os que não admitem objeto direto, por isso esse fato é relevante na formulação das valências verbais. O objeto direto é representado por um SN cuja ocorrência em uma oração não é livre, uma vez que há verbos que recusam esse complemento, como nos exemplos seguintes: *\*Estela ia uma blusa*; *\*Túlio gosta seus pais*; *As chaves apareceram*; *Surgiu um probleminha na sua inscrição*. Com esses verbos, não pode ocorrer um SN além do sujeito. A ocorrência de objeto não é livre também porque, ao que se nota, há verbos que exigem a sua presença, como *pôr*, *colocar*, *tirar* e alguns outros: *\*Ele tirou da gaveta*.; *\*Joana empurrou para um canto*. Com a supressão do objeto, essas frases são agramaticais.

### A omissão do objeto direto e os verbos de movimento

Uma hipótese que levantamos é que verbos em orações cujo objeto direto seja Tema, isto é, entidade que se desloca em um evento de movimento ou entidade localizada em um evento de localização, não aceitam a omissão do objeto direto. Nos exemplos seguintes o SN objeto direto encontra-se grafado em itálico:

(13) Meu pai levou *o celular*.

(14) Os professores distribuem *muitos livros*.

(15) O governo enviará *uma resposta*.

- (16) Meu avô tem *uma fazenda*.  
(17) Maria tirou *o bolo* da forma.  
(18) O vendedor trouxe *o aparelho*.

Nos exemplos dados, a omissão do objeto direto gera frases agramaticais, como se segue:

- (13a) \* Meu pai levou.  
(14a) \* Os professores distribuem.  
(15a) \* O governo enviará.  
(16a) \* Meu avô tem.  
(17a) \* Maria tirou da forma.  
(18a) \* O vendedor trouxe.

Parece-nos, ainda, que com verbos de movimento ou de causação de movimento não é possível a omissão do objeto direto quando este desempenha os papéis de Fonte e Meta. Vejamos alguns exemplos em que o objeto direto é Fonte. São poucos verbos que cabem nessa construção, uma vez que a maioria dos complementos no papel temático de Fonte estrutura-se com Sprep, em frases como *Ele saiu de casa*.

- (19) Marido entediado abandona *o lar*.  
(20) A torcida deixou *o campo* antes do final da partida.  
(19a) \* Marido entediado abandona.  
(20a) \* A torcida deixou antes do final da partida.

Alguns exemplos com o objeto direto no papel temático de Meta:

- (21) A pipa alcançou *o céu*.  
(22) Os caminhantes atingiram *a metade do trajeto*.  
(21a) \* A pipa alcançou.  
(22a) \* Os caminhantes atingiram.

As observações acima são importantes porque apontam para a presença de fatores semânticos na omissibilidade do objeto: nesses casos ela não é uma idiossincrasia do verbo (a ser consignada na diátese), mas decorre de um princípio geral que proíbe a omissão de objetos com certos papéis temáticos. Esse princípio não dá conta de todos os casos de omissão do objeto, mas precisa ser levado em conta como fator importante na descrição do fenômeno.

## A omissão do objeto direto e os efeitos de sentido

Trataremos a seguir de questões relacionadas à omissão do objeto, procurando questionar que efeitos semânticos ela acarreta. Dada a frase *Minha mãe costurava*:

(23)	Minha mãe	costurava	
	SN	V	Ø
	Agente		Paciente

Ao ouvir essa frase, apesar da ausência do objeto direto, o receptor elabora em sua mente, de forma esquemática, um complemento para o verbo *costurar*: vestidos, bolsas, ternos, bonés etc. Já em frases como:

(24) Ele cuspiu na rua.

(25) Ele mijou na rua.

elabora-se, com base pragmático-contextual, um complemento para preencher o objeto direto omitido: *cuspe, saliva* em (24); *urina* em (25).

Talvez isso seja em razão da complexidade semântica de cada verbo (dizemos talvez porque ainda não aprofundamos o estudo sob esse aspecto). No dicionário Michaelis, *Costurar/coser* significa “ligar, unir com pontos de agulha”; *cuspir* é “expelir cuspo, lançar da boca cuspo ou outra coisa” (MICHAELIS, 1998). Note-se que para *costurar* o dicionário não menciona a “coisa costurada”, mas para *cuspir* menciona a “coisa cuspada”, presumivelmente porque aqui temos um objeto prototípico. Isso pode ser expresso, dentro da teoria dos esquemas (RUMELHART; ORTONY, 1976), dizendo que *costurar* tem as variáveis rotuladas Agente e Paciente, porém esta última em aberto; e *cuspir* tem as mesmas variáveis, mas com o Paciente já “pré-preenchido”, prototipicamente com “cuspo”, cuja expressão é aparentemente dispensada; mas há também a possibilidade de se cuspir outra coisa, e nesses casos o objeto precisa ser expresso:

(26) O artista cuspia *fogo*.

(27) O doente cuspiu *sangue*.

(28) O marido cuspiu *ofensas* durante a discussão.

(29) O bebê cuspia *o leite dormido*.

O mesmo se dá com o verbo *urinar*, que é expelir urina; quando o objeto é outra coisa diferente, é preciso explicitar: urinar sangue, por exemplo. Aqui temos, portanto, um novo fator na omissibilidade do objeto: o caráter prototípico do objeto denotado em relação ao evento expresso pelo verbo.

Goldberg (2001) observa que a omissão de complementos requer que a ação denotada pelo verbo seja entendida como aplicável reiteradamente: iterativamente ou genericamente – *Eu compro, compro e não pago; Eu como à noite antes de dormir*. Porém, o que levaria o ouvinte/leitor a preencher o complemento da frase a seguir:

(30) Meu marido bebe.,

com o Paciente *bebida alcoólica*, um Paciente específico? Aqui já não ajuda lançar mão do esquema, porque na definição de *beber* não se inclui a natureza da coisa bebida, nem sequer se pode dizer que bebida alcoólica seja a coisa que mais frequentemente se bebe. No caso da frase acima, o Paciente do verbo *beber* foi preenchido no esquema Agente/Paciente por bebida alcoólica, mas poderia ter sido por outro líquido qualquer, isto é, por um Paciente esquemático; o Paciente privilegiado bebida alcoólica só aparece em certos contextos. O que leva o ouvinte, nesse caso, a preencher esse Paciente com bebida alcoólica, um Paciente privilegiado, mas não em: *É aqui que os animais bebem?* Conhecimento de mundo, que animais não bebem bebida alcoólica? Mas desde já podemos apontar um terceiro fator que condiciona a omissão do objeto: com certos verbos, parece haver um “Paciente privilegiado”, que aparece em certas circunstâncias ainda a explicitar. Que circunstâncias são essas é uma das perguntas a que tentaremos dar resposta no decorrer da pesquisa.

### Considerações finais

Vimos acima que um dos fatores que governam a possibilidade de omissão do objeto é o papel temático que ele veicula. Essa é certamente uma solução parcial, mesmo porque há casos de objetos de papéis temáticos idênticos e que se comportam diferentemente quanto à omissão, como, por exemplo, *beber*, cujo objeto, quando omitido, pode ser entendido de maneira elaborada como “bebida alcoólica”, em contraste com *comer*, que, sem objeto, tem Paciente esquemático. Assim, vamos precisar de outras hipóteses, que não podem ser examinadas aqui, mas que constituem as perguntas básicas na continuação desta pesquisa. Damos a seguir algumas possibilidades:

I – A omissão é livre em princípio, e o referente do objeto direto omitido é esquemático:

(31) Eu nunca como à noite.

II – A omissão parece que não é livre com verbos como *fazer*, *colocar*, *tirar*, que aparentemente exigem objeto explícito, mesmo esse objeto sendo esquemático, como nos exemplos seguintes. Isto é, *fazer*, *tirar*, *colocar*, *pôr* é sempre fazer, tirar, colocar, pôr *qualquer coisa*; porém a omissão parece não ser possível, gerando agramaticalidade:

(32) \*Ele tirou da geladeira (o leite, o doce, o peixe etc.).

(33) \*Mamãe fazia (careta, docinhos, o exame etc.).

(34) \*Ele colocou no carro (o livro, o alarme, os pés, as malas ...).

III – A omissão é imprevisível em princípio, e cada verbo precisa ser marcado. Disso decorre que cada verbo que admite a omissão do objeto precisa ser marcado, isto é, precisa ter duas diáteses, uma com o termo e a outra sem o termo em questão.

IV – A omissão é livre, mas sujeita a condições pragmáticas ou de maior ou menor elaboração do verbo, como *comer* e *devorar*, quando o primeiro permite e o segundo não permite a omissão do objeto.

(35) A onça já comeu.

(36) \*A onça já devorou.

V – Com verbos leves – alguns verbos possuem uma acepção “leve”, basicamente quando participam de construções nas quais o elemento nominal (e não o verbo) parece ser o verdadeiro responsável pela denotação do evento da sentença (SHERR, 2003, p. 206). Quando alguns desses verbos estão em sua forma não leve, a omissão do objeto direto ocasiona uma alteração de sentido (38), ou gera uma frase agramatical (40):

(37) Minha avó sofreu uma queda quando era criança.

(38) Minha avó sofreu quando era criança.

(39) Eles realizaram o assalto em plena luz do dia.

(40) \*Eles realizaram em plena luz do dia.

Como vimos, em alguns casos a omissão do objeto acarreta mudança de sentido expresso pelo verbo. Dizemos em alguns casos, pois ocorre com *sofrer* (leve/não leve), mas não com *comer*, por exemplo. Esses casos podem ser delimitados semanticamente, ou vão ter de ser marcados um a um nas diáteses?

O que se pretende é a ampliação deste estudo para pesquisar as condições em que se verifica a opcionalidade do objeto direto no Português do Brasil, se e quando resulta em alteração de sentido; se a ocorrência depende de traços semânticos do verbo, do seu grau de elaboração; ou do aspecto; se é idiossincrática; como se comportam os verbos semanticamente semelhantes, ou outros fatores.

## REFERÊNCIAS

GOLDBERG, A. E. Patient arguments of causative verbs can be omitted: the role of information structure in argument distribution. *Language Sciences*, Cambridge, v. 23, p. 503-524, 2001.

\_\_\_\_\_. *Constructions: a construction grammar approach to argument structure*. Chicago: The University of Chicago Press, 1995.

LEVIN, B. Objecthood: an event structure perspective. *CLS*, Chicago, n. 35, v. 1, p223-247, The Main Session, 1999.

MICHAELIS: moderno dicionário da língua portuguesa. São Paulo: Melhoramentos, 1998. (Dicionários Michaelis).

PERINI, M. A. *Estudos de gramática descritiva: as valências verbais*. São Paulo: Parábola, 2008a.

\_\_\_\_\_. *Describing verb valency: practical and theoretical issues*. Belo Horizonte: Em fase de elaboração.

PLANK, F. (Org). *Objects: towards a theory of grammatical relations*. London: Academic Press, 1984.

RUMELHART, D. E.; ORTONY, A. *The representation of knowledge in memory*. La Jolla, California: Center for Human Information Processing, University of California, San Diego, 1976.

## **BIBLIOGRAFIA NÃO CITADA**

CULICOVER, P. W.; JACKENDOFF, R. *Simpler syntax*. Oxford: Oxford University Press, 2005.

FILLMORE, Charles J. *Form and meaning in language: papers on semantic roles*. v. I. Stanford: CSLI Publications, 2003.

GOLDBERG, A. E. *Constructions at work*. Oxford: Oxford University Press, 2006.

HOPPER, P. J.; THOMPSON, S. A. Transitivity in grammar and discourse. *Language*, Stanford, v. 56, n. 2, p. 251-299, jun. 1980.

LEVIN, B. *More on semantic determinants of argument realization: evidence from transitivity*. August 2006. Disponível em: <<http://www.stanford.edu/~bclevin/dgfs06.html>>. Acesso em: 02 abr. 2012.

\_\_\_\_\_. *Is aspect a semantic determinant of argument realization?* 2005. Disponível em: <<http://www.stanford.edu/~bclevin/lisa05.html>>. Acesso em: 02 abr. 2012.

McKOOK, G.; MACFARLAND, T. Externally and internally caused change of state verbs. *Language*, Stanford, v.76, n. 4, p. 833-858, 2000.

PERINI, M. A Por uma descrição gramatical mais concreta: as funções sintáticas. In: CONGRESSO INTERNACIONAL DA ASSOCIAÇÃO BRASILEIRA DE LINGÜÍSTICA, V., 2008, Belo Horizonte. *Conferências do V...*, Belo Horizonte: UFMG. 2008b, p. 29-44.

\_\_\_\_\_. *Gramática descritiva do português*. São Paulo: Ática, 1995.

PERINI, M. A.; OTHERO, G. de A. *Corpus, introspecção e o objeto da descrição gramatical*. 2010. Disponível em: <<http://online.unisc.br/seer/index.php/signo/article/viewFile/1348/1255>>. Acesso em: 04 jul. 2011.

SAUSSURE, F. de. *Curso de linguística geral*. 20. ed. Tradução de Antônio Chelini, José Paulo Paes e Izidoro Blikstein. São Paulo: Cultrix, 1995. [1916].

SCHER, A. P. Quais são as propriedades lexicais de uma construção com verbo leve? In: MÜLLER, A. L.; NEGRÃO, E. V.; FOLTRAN, M. J. (Org.). *Semântica formal*. São Paulo: Contexto, 2003. p. 205-219.

WACHOWICZ, T. C.; FOLTRAN, M. J. Sobre a noção de aspecto. *Caderno de estudos linguísticos*, Campinas, v. 48, n. 2, p. 211-232, 2006.

# Construções de *small clauses* complexas em português brasileiro

(Complex small clauses constructions in Brazilian Portuguese)

Patrícia Rodrigues,<sup>1</sup> Maria José Foltran<sup>2</sup>

<sup>1,2</sup>Universidade Federal do Paraná (UFPR)

rodriguespatriciaa@gmail.com, mariajose.foltran@gmail.com

**Abstract:** This work is concerned with copular clauses in Brazilian Portuguese as (1a) *A Maria bêbada é chata* e (1b) *A Maria bêbada é chato*. In the sentence (1a), in the relevant reading, the adjective *bêbada* is a predicative adjective, and the sentence can be paraphrased as *A Maria quando está bêbada é chata*. The sentence in (1b) has as its subject a constituent that refers to a situation, and can be paraphrased as *A Maria estar bêbada é chato*. The current standard analysis for copular sentences claims that the copula selects a small clause (SC) complement. The agreement marks in the sentence (1a) suggest that the subject of the selected SC is a DP. Under this view, we will argue that the adjective *bêbada* in (1a) must be analyzed as a secondary predicate. For the sentence in (1b), we argue that a SC subject analysis – *Ser [<sub>SC</sub> [<sub>SC</sub> a Maria bêbada] chato]]* – can explain the agreement facts observed.

**Keywords:** small clause; secondary predication; copula.

**Resumo:** Este trabalho examina as seguintes sentenças do português brasileiro: (1a) *A Maria bêbada é chata* e (1b) *A Maria bêbada é chato*. Na sentença (1a), na leitura relevante, o adjetivo *bêbada* é um adjetivo predicativo, e a sentença pode ser parafraseada como *A Maria quando está bêbada é chata*. A sentença (1b) tem como sujeito um constituinte que denota uma situação e pode ser parafraseada como *A Maria estar bêbada é chato*. A análise corrente para construções copulares propõe que a cópula seleciona uma *small clause* (SC) complemento. As marcas de concordância em (1a) sugerem que o sujeito da SC selecionada é um DP. Assim, propomos neste trabalho que o adjetivo *bêbada* em (1a) deve ser analisado como um predicado secundário. Para a sentença (1b), argumentamos que uma análise da SC como sujeito – *Ser [<sub>SC</sub> [<sub>SC</sub> a Maria bêbada] chato]]* – pode explicar os fatos de concordância observados.

**Palavras-chave:** small clause; predicação secundária; cópula.

## Introdução

Este trabalho examina construções copulares do português brasileiro como as ilustradas em (1).

- (1) a A Maria bêbada é chata.  
b A Maria bêbada é chato.

Uma primeira diferença a ser observada entre essas construções é a forma do adjetivo “chato”: em (1a), ele é marcado para o gênero feminino, enquanto em (1b), apresenta forma neutra (masculino singular). Além disso, as interpretações de (1a) e de (1b) são diferentes. A construção em (1a) é ambígua entre uma leitura, parafraseada em (2a), em que o adjetivo *bêbada* é modificador do nome *Maria* (pressupõe-se que há uma outra Maria que não está bêbada), e uma leitura, parafraseada em (2b), em que o adjetivo *bêbada* é predicativo. Nessa interpretação, que é a relevante para este trabalho, o adjetivo

*chata* atribui uma propriedade ao DP *A Maria*. Já na sentença em (1b), o adjetivo *chato* atribui uma propriedade a uma situação; essa leitura está parafraseada em (3).

- (2) a A Maria que está bêbada é chata  
b A Maria é chata quando está bêbada
- (3) A Maria estar bêbada é chato

Ainda, no caso de (1a), o adjetivo *bêbada* pode ocupar diferentes posições na sentença sem que haja mudança no significado, como ilustrado em (4a-b). Isso não ocorre no caso da sentença (1b), como pode ser visto em (5a-b)..

- (4) a A Maria é chata bêbada.  
b Bêbada, a Maria é chata.
- (5) a \*A Maria é chato bêbada.  
b \*Bêbada, a Maria é chato.

Compararemos as construções em (1) com a sentença em (6), que apresenta o mesmo tipo de concordância neutra observado em (1b). O que as diferencia é o formato do constituinte sujeito: em (6) não temos SC. Uma abordagem mais aprofundada sobre (6) está fora do escopo deste trabalho.<sup>1</sup>

- (6) Alunas que bebem é chato.

O objetivo deste trabalho é examinar as propriedades das sentenças em (1) e discutir em que medida essas propriedades estão ligadas à estrutura sintática dessas sentenças. Para apoiar nossa descrição, partimos de uma análise das construções copulares já consagrada na literatura que postula que a cópula é um verbo inacusativo que seleciona uma *small clause* (SC). É importante salientar que este estudo possui um objetivo essencialmente descritivo. Procuramos fornecer uma boa caracterização das construções em tela, para que estudos mais aprofundados possam ser posteriormente desenvolvidos.

Na primeira seção do artigo, discutiremos as propriedades da construção em (1a), que apresenta uma concordância marcada. As marcas de concordância e os fatos relativos à ordem ilustrados em (5a) sugerem que o sujeito da SC selecionada é o DP *a Maria* e que o adjetivo *bêbada* seria um predicado secundário. Veremos que existe na literatura uma distinção que separa os predicados secundários depictivos típicos dos circunstanciais. Mostraremos que, para a leitura apontada em (2b), o adjetivo *bêbada* em (1a) pode ser classificado como um predicado secundário do tipo circunstancial, e discutiremos as implicações dessa classificação para a determinação da estrutura da sentença. Em seguida, na segunda seção, discutiremos as propriedades das sentenças em (1b). No que diz respeito à sentença em (1b), propomos que o sujeito da SC selecionada pela cópula é também uma SC, como representado em (7). Realizaremos, em seguida, uma breve discussão sobre a sentença em (6). Por fim, apresentaremos nossas considerações finais.

- (7) Ser [<sub>SC</sub> [<sub>SC</sub> a Maria bêbada] chato].

<sup>1</sup> A esse respeito, ver Foltran e Rodrigues (2012).

## A concordância marcada

Como discutido na introdução, o adjetivo *chata* na sentença (1a) predica sobre o DP *a Maria*. Considerando a análise em que a cópula seleciona uma SC, propomos que a ordem de base da sentença (1a) seria a da sentença (5a). Assim, a cópula *SER* selecionaria a SC *A Maria chata* e o adjetivo *bêbada* funcionaria como um predicado secundário orientado para o sujeito: a propriedade denotada pelo predicado secundário (*bêbada*) se sobrepõe temporalmente à propriedade expressa pelo predicado matriz (*chata*). A variação apresentada em (5a-a'), no entanto, não é uma característica dos predicados secundários como um todo, como é possível observar em (8) e (9), que trazem exemplos de sentenças com predicados secundários também orientados para o sujeito.

- (8) a A Maria chegou em casa cansada.  
b ??Cansada, a Maria chegou em casa.  
c ??A Maria cansada chegou em casa.
- (9) a A Maria comeu a sobremesa bêbada.  
b ??Bêbada, A Maria comeu a sobremesa.  
c ??A Maria bêbada comeu a sobremesa.

De modo geral, na literatura, os predicados secundários são divididos em depictivos e resultativos. Himmelmann e Schultze-Berndt (2005), contudo, com base em Nichols (1981) e Simpson (2005), reconhecem ainda uma subdivisão no domínio dos depictivos entre depictivos em sentido estrito (*depictives proper*) (10) e circunstanciais (11).

- (10) a They left angry.  
b She died young.
- (11) a This food is not supposed to be nice cold.  
b I can't work hungry.

Os autores afirmam que esses dois tipos de predicado expressam um estado que se sobrepõe temporalmente ao estado expresso pelo predicado principal, mas que os circunstanciais fornecem informação adicional com relação à circunstância do evento principal (tempo, condição, concessão). Os autores afirmam igualmente que, em algumas línguas, inclusive em inglês, esses dois tipos de predicado se distinguem no que diz respeito à interação com o foco da sentença. Os predicados secundários típicos (*depictives proper*) fariam parte do domínio focal da oração em que aparecem, ao passo que os predicados secundários circunstanciais contribuiriam com a informação pressuposta. Por exemplo, nos casos ilustrados em (10), o foco cai sobre os adjetivos *angry* e *young*. Esse não é o caso dos adjetivos *cold* e *hungry*,<sup>2</sup> em (11).

Isso explicaria, de acordo com os autores, por que, diferentemente dos circunstan-

<sup>2</sup> De acordo com Himmelmann e Schultze-Berndt (2005), a sentença em (i) só seria aceitável em um contexto contrastivo, como em (ii).

??This tea isn't good COLD.  
This tea isn't good COLD, it's good HOT.

ciais, os depictivos em inglês não são normalmente aceitos no início da sentença, a não ser em contextos contrastivos, e por que eles entram no escopo da negação sentencial, como pode ser observado em (12). Não é possível interpretar a negação em (12) como uma negação apenas do predicado principal. A negação tem escopo ou sobre os dois predicados (12a) ou somente sobre o predicado secundário (12b-c). Já os predicados secundários circunstanciais aparecem fora do escopo da negação, como ilustrado em (11). Ou seja, é possível negar o predicado principal sem negar o predicado secundário.

- (12) a No! They didn't leave outraged, they didn't leave at all.  
 b They didn't leave outraged (they left happy).  
 c She didn't die young, she died in her seventies.

Essa distinção entre dois tipos de predicados secundários depictivos também é encontrada em português. Foltran (1999) observa que as sentenças em (13a-b), ambas com um adjetivo no papel de predicado secundário orientado para o sujeito, têm comportamentos distintos com relação às paráfrases possíveis para cada uma (13a'-b') e também com relação ao alcance da negação (13a''-b''). Outro exemplo pode ser visto em (14), que é ambígua entre duas leituras, conforme o adjetivo *bêbada* esteja (15) ou não (16) no domínio focal da sentença. Essas leituras estão respectivamente parafraseadas em (17).

- (13) a João chegou cansado.  
 a' João estava cansado quando chegou.  
 a'' João chegou cansado? #Não, ele não chegou.  
 b João é insuportável bêbado.  
 b' João é insuportável quando está bêbado.  
 b'' Ele é insuportável bêbado? Não, ele não é.

(14) A Maria dança valsa bêbada.

(15) (— *Como que a Maria dança valsa?*)  
 — A Maria dança valsa bêbada.

(16) (— *O que a Maria faz bêbada?*)  
 — A Maria dança valsa bêbada.

- (17) a Quando a Maria dança valsa, ela está bêbada.  
 b Quando a Maria está bêbada, ela dança valsa.

Considerando a classificação de Himmelmann e Schultze-Berndt (2005), o predicado secundário em (14) pode ser descrito seja como um depictivo em sentido estrito (15), seja como um circunstancial (16). É possível observar, nesses casos, as propriedades relativas a cada tipo de predicado secundário, como discutidas pelos autores.

Em primeiro lugar, há uma diferença relacionada com as posições que o adjetivo *bêbada* pode ocupar na sentença. Quando o adjetivo é deslocado, como em (18), a única leitura possível é a parafraseada em (17b).

- (18) a A Maria bêbada dança valsa.  
b Bêbada, a Maria dança valsa.

Ainda, as duas leituras da sequência em (14) estão relacionadas a uma diferença de escopo da negação sentencial. Para a sentença em (19a), em que o escopo da negação inclui apenas o adjetivo *bêbada*, a única leitura possível é a leitura em (17a). Já para a sentença em (19b), em que o escopo da negação inclui apenas o predicado principal, a única leitura possível é a de (17b).

- (19) a A Maria não dança valsa bêbada, ela dança valsa sóbria.  
b A Maria não dança valsa bêbada, ela toca guitarra bêbada.

Geuder (2000), em discussão sobre depictivos em inglês, examina casos como o da sentença em (20a), que estamos chamando de depictivos típicos, como sugerido por Himmelmann e Schultze-Berndt (2005). Para o autor, os depictivos ocorrem sempre em posição final e são ancorados ao evento principal. Ele afirma que, quando construções com depictivos são parafraseadas por *when-clauses*, o adjetivo regularmente aparece como predicado principal, como em (20b). Essa correspondência resultaria do fato de que os depictivos, em posição final, atraem o acento focal; assim, como a *when-clause* normalmente pressupõe que o evento descrito é em certo sentido “conhecido”, é a oração principal que contém informação nova.

- (20) a John left the party sad  
b John was sad when he left the party

Para Geuder (2000), os depictivos em inglês, tanto os orientados para o sujeito, como os orientados para o objeto, são adjungidos à direita no nível do VP. Ele justifica a adjunção nesse nível mostrando que a negação sentencial tem, invariavelmente, escopo sobre o depictivo. A sentença (21), por exemplo, não significa que Bill, bravo com John, não foi embora, mas que Bill não estava bravo com John quando foi embora. Geuder conclui que, se a adjunção ocorresse no nível de IP, esse fato seria inesperado. Com relação a esse aspecto, ele distingue os depictivos de outros elementos adjungidos à direita, que podem ser interpretados dentro ou fora do escopo da negação, como nos exemplos em (22). A desambiguação é feita pela entoação: se o adjunto está fora do escopo da negação, o VP *kiss his wife* recebe o acento sentencial não marcado e o adjunto recebe um acento próprio. Quando o acento é distribuído sobre os elementos pós-verbais, isso é normalmente considerado como indicação de que esses elementos estão contidos no VP.

- (21) Bill didn't leave angry at John.  
(22) a John didn't kiss his wife, [because he loves her].  
b John didn't kiss his wife (,) deliberately.

Geuder estabelece um paralelo ainda entre depictivos e adverbiais temporais adjungidos ao VP. Muitos dos adverbiais seriam, para o autor, ambíguos com relação a uma anco-

ragem no tempo de referência e uma ancoragem no tempo do evento.<sup>3</sup> Para os adverbiais em início de sentença, somente a ancoragem no tempo de referência seria possível (23a), mas, para os adverbiais em posição final, ambas as ancoragens seriam possíveis, como exemplificado em (23b-c) (a distinção entre (23b) e (23c) é realizada pela entoação). As sentenças em (23a) e (23b) são sinônimas – às dez horas é o tempo depois da partida. Já no caso de (23c), a única leitura disponível é aquela em que a hora da partida é identificada como às dez horas – nesse caso, o adverbial seria adjungido ao VP, como os depictivos.

- (23) a At ten, John had léft.  
 b John had léft at ten.  
 c John had left at tén.

Ainda que Geuder não discuta casos como os mencionados por Himmelmann e Schultze-Berndt (2005), listados em (11), ou casos como os do português, exemplificado em (14), é possível estender o paralelo que faz o autor e afirmar que os depictivos circunstanciais, diferentemente dos depictivos típicos, não estão adjungidos no nível de VP.

Retomando a discussão sobre a sentença (1a), objeto deste trabalho, cabe observar que ela possui uma leitura preferencial, cuja paráfrase, vista em (2), está repetida em (24a). Nessa leitura, o adjetivo *bêbada* não está no domínio focal da sentença. A leitura em que o foco cai sobre esse adjetivo é mais difícil de ser obtida (24b). Considerando-se ainda que, na leitura preferencial, esse adjetivo pode ocupar posições variadas na construção (25) e que pode ficar fora do escopo da negação sentencial (26a), ele pode ser descrito como um predicado secundário circunstancial. Assim, com base no que foi discutido acima, propomos que, para a leitura preferencial de (1a), o adjetivo *bêbada* não está adjungido no nível do VP, mas do IP. Se, no entanto, considerarmos a leitura não preferencial (24b), o adjetivo *bêbada* deve ser analisado como um depictivo típico, estando, portanto, adjungido a VP. Nesse caso, (25b-c) seriam inaceitáveis, e o foco da negação recairia sobre o adjetivo *bêbada* (26b).

- (24) a A Maria é chata quando está bêbada. (cf. A Maria é CHATA bêbada)  
 b #A Maria quando fica chata, está bêbada. (cf. A Maria é chata BÊBADA)
- (25) a A Maria é chata bêbada.  
 b A Maria bêbada é chata.  
 c Bêbada, a Maria é chata.
- (26) a A Maria não é chata bêbada, ela é divertida.  
 b #A Maria não é chata bêbada, ela é chata sóbria.

Como observado na introdução do trabalho, nosso objetivo não é aprofundar a discussão sobre as construções examinadas.<sup>4</sup> Assim, não discutiremos em detalhes a estrutura dessas construções. Acreditamos, no entanto, que as propriedades apontadas são importantes e devem ser consideradas em estudos futuros.

<sup>3</sup> Para Geuder (2000), *tense is taken as a relation between utterance time and reference time (i.e., the time a claim is made about)*.

<sup>4</sup> Para maiores detalhes, ver Foltran et al (2012).

## A concordância neutra

Nesta seção, discutiremos as sentenças em (1b) e em (6), que apresentam concordância neutra. No caso da sentença (1b), propomos que a concordância neutra se deve ao fato de o sujeito da construção ser uma SC; ou seja, haveria uma concordância *default* típica de casos com sujeitos oracionais.

Não analisamos o sujeito em (1b) como um DP, porque, se fosse o caso, isso significaria que o adjetivo *bêbada* estaria incluído nele. No entanto, como vimos na introdução deste trabalho, numa leitura em que o adjetivo *bêbada* é modificador do nome *Maria* (em que se pressupõe que há uma outra *Maria* que não é/está bêbada), o adjetivo *chato*, predicado da construção copular, deve concordar em gênero e número com o DP *a Maria*. Com a concordância neutra, como em (1b), essa leitura não é possível. Vimos também que, na leitura em que o adjetivo *bêbada* está fora do DP sujeito e funciona como um predicado secundário, o adjetivo *chato* também deve concordar com o DP *a Maria* (leitura parafraseada em (2) e repetida em (27)); além disso, nessa leitura, vimos que o adjetivo *bêbada* pode ocupar outras posições na sentença. Na sentença em (1b), a única interpretação possível é aquela em que o adjetivo *chato* atribui uma propriedade ao que podemos chamar de situação (leitura parafraseada em (3) e repetida em (28)), e é o constituinte *a Maria bêbada* (parafraseado como *A Maria estar bêbada*) que denota essa situação. Assim, a configuração que defendemos para a sequência *a Maria bêbada* em (1b) é a de uma SC sujeito da SC selecionada pela cópula, como representado em (7), e repetido em (29).

(27) A *Maria* é chata quando está bêbada. (leitura preferencial de (1a))

(28) A *Maria* estar bêbada é chato.

(29) Ser [<sub>SC</sub> [<sub>SC</sub> *a Maria bêbada*] chato]].

Outros autores reconhecem a possibilidade de uma SC poder ocupar a posição de sujeito, como ilustrado em (30).<sup>5</sup>

(30) a Will President Botha set him free to allow him to take part in that process? Or does he fear that Mandela free would pose a greater threat than Mandela behind bars. (AARTS, 1992, p. 45)

b Mandela behind bars would undoubtedly suit the government. (HAEGEMAN; GUÉRON, 1999, p. 120)

c Mary sad would be hard to deal with. (CONTRERAS, 1987, p. 233)

Além disso, essa sequência pode aparecer como complementos de predicados que selecionam SCs, como o verbo *querer* ou *preferir* (31a) (CARREIRA, 2008). Nesses casos, os argumentos levantados para se defender que a sequência *a Maria bêbada* é uma SC dizem respeito à falta de acarretamento necessário observado na sentença em (31b) e à possibilidade de movimentação da sequência como um todo, como em (31c).

(31) a O João quer / prefere a *Maria bêbada*.

b #O João quer / prefere a *Maria*.

c A *Maria bêbada*, o João quer / prefere.

<sup>5</sup> Grifo nosso.

É possível argumentar igualmente que essa mesma sequência pode aparecer como sujeito de verbos como *surpreender*, que selecionam um argumento externo que pode se realizar sintaticamente como uma oração (32a). O exemplo em (32c) mostra que a sequência *A Maria bêbada* em (32b) pode ser analisada como um constituinte oracional, que não desencadeia concordância verbal. Ou seja, em (32c), se o sujeito fosse um DP, o verbo apresentaria marca de plural; na nossa opinião, a ausência dessa marca implica, nesse caso, que o sujeito deve ser analisado como uma SC. É possível, assim, estender essa análise para a sentença em (32b), postulando que a concordância observada no verbo é neutra.

- (32) a Que a Maria tenha bebido surpreendeu todo mundo.  
b A Maria bêbada surpreendeu todo mundo.  
c As alunas bêbadas surpreendeu todo mundo.

Dessa forma, propomos que a sequência *A Maria bêbada* em (1b) seja analisada como uma SC.<sup>6</sup> Isso explicaria a concordância neutra que aparece na cópula e no adjetivo. Contudo, o caso da sentença em (6), repetida em (33), não pode ser explicado da mesma maneira. Nessa sentença, o sujeito *alunas que bebem* – um nominal modificado por uma relativa – é claramente um DP.

- (33) Alunas que bebem é chato.

Essa possibilidade foi originalmente apontada por Baltin (1998), em argumentação contra a análise da construção do inglês dada em (34), proposta por Safir (1983). Safir afirma que a concordância da cópula no singular exclui a possibilidade de que a sequência *workers angry about the pay* seja analisada como um NP seguido de um adjetivo atributivo (i.e. *the workers who were angry about the pay*). Para Safir, essa sequência deve ser interpretada como uma situação e analisada como um constituinte oracional (*a clausally interpreted construction*). Safir propõe que esse tipo de construção pode servir como uma evidência para a existência de SCs (1983).

- (34) Workers angry about the pay is just the sort of situation that the ad campaign was designed to avoid. (SAFIR, 1983, p. 732)

Baltin (1996) argumenta contra a análise de Safir com base nos dados em (35).<sup>7</sup> De acordo com Baltin, o sujeito em (35a) só pode ser analisado como um nominal com um adjetivo pré-nominal, e o sujeito em (35b) só pode ser analisado com um nominal modificado por uma relativa. Assim, para o autor, esses constituintes obviamente não são SCs. Ele conclui então que a sequência *workers angry about the pay* em (34) constitui um DP e que a concordância no singular não serve como diagnóstico para SCs. Para ele, esses fatos mostram que a concordância deve ser, ao menos parcialmente, um fenômeno semântico, e não puramente sintático.

- (35) a Several angry workers is just the sort of situation that the ad campaign was designed to avoid.  
b Workers who are angry about the pay is just the sort of situation that the ad campaign was designed to avoid.

<sup>6</sup> Maiores detalhes sobre a constituição dessa SC estão em Foltran et al (2012).

<sup>7</sup> Exemplos (5a-b), p. 514.

Ainda que estejamos de acordo com Baltin quando ele afirma que a concordância no singular observada por Safir não constitui um argumento para se postular uma SC no caso de (34), haja vista os casos em (35), discordamos de sua conclusão de que *workers angry about the pay* é necessariamente um DP. O fato de que a concordância neutra não seja evidência para analisar essa sequência como uma SC não significa que o sujeito em (34) não possa ser uma SC. De acordo com o que foi discutido acima sobre a sentença (1b), acreditamos que é plausível postular que uma construção copular possa apresentar como sujeito uma SC. No entanto, a questão levantada por Baltin (1996) é pertinente, e os casos como o da sentença em (33) devem ser explicados.

Retomando o que foi visto até agora, teríamos, em princípio, dois casos de construções copulares em que aparece a concordância neutra: a sentença em (1b), repetida em (36a), e a sentença em (33), repetida em (36b). No caso de (36a), a análise do constituinte sujeito como uma SC explicaria os fatos de concordância. Contudo, essa explicação não estaria disponível para (36b), já que o constituinte sujeito forma claramente um DP. O que deve ser explicado, então, é como constituintes que formam claramente DPs plurais podem não desencadear concordância na cópula e no adjetivo. Além disso, também pede uma explicação o fato de que, em (36b), há desencadeamento de concordância dentro do DP (entre *alunas* e *bebem*). É necessário observar que a concordância neutra não é obrigatória nesses casos. No entanto, vale lembrar que, quando a concordância é marcada, a interpretação que se obtém é diferente (37). Nas sentenças em (37), não se fala mais de situações, mas de indivíduos.

- (36) a A Maria bêbada é chato.  
b Alunas que bebem é chato.

- (37) a A Maria bêbada é chata.  
b Alunas que bebem são chatas.

Assim, de um ponto de vista descritivo, podemos afirmar que tanto SCs quanto DPs podem denotar uma situação, e que os casos em que a concordância é neutra parecem coincidir, em princípio, com os casos em que o constituinte sujeito denota uma situação. A questão que se coloca é a de saber se SCs e DPs têm propriedades em comum que, por um lado, permitem que esses constituintes denotem situações, e que, por outro, fazem com que a concordância não seja disparada. Com o objetivo de entender melhor o fenômeno da concordância neutra, discutiremos a seguir algumas propriedades dessas construções. Em primeiro lugar, examinaremos os predicados passíveis de aparecerem como predicado principal nessas sentenças copulares. Em seguida, examinaremos os seus constituintes sujeitos, discutindo, com relação à SC sujeito, os tipos de sujeito e de predicado que a constituem, e, com relação ao DP, quais características que o capacitariam a denotar uma situação. Voltamos a salientar que nosso objetivo é tão somente fornecer uma descrição dessas construções, sem nos aprofundar na determinação da estrutura de cada um desses constituintes nem na determinação do mecanismo que explicaria os fatos de concordância.

A fim de embasar minimamente nossa descrição dos fatos observados, recorremos ao estudo de Hegarty (2003) sobre entidades abstratas. O autor, com base em Asher (1993), situa proposições, fatos, situações, razões, eventos e estados em um espectro de

imanência ao mundo, “according to the degree to which they are entities in the world, interacting with other entities in space and time” (HEGARTY, 2003, p. 893). Para o autor, fatos, situações e razões ocupariam uma posição intermediária entre eventos e proposições. Desse ponto de vista, “situations have locations in time and space, rather like events and states, but situations include ramifications to an extent that events and states do not, making them less definitively bounded in space and time” (p. 892). Considerando as propriedades de uma situação, Hegarty (2003) propõe que predicados como *(in)tolerable*, *(ir)reversible* or *can(’t) be remedied* tomariam situações como argumentos, pois uma situação pode incluir efeitos persistentes, que são passíveis, por exemplo, de serem tolerados (ou não). Com base nessa definição, podemos acrescentar outros predicados à lista, como *temporário*, *passageiro*, *inconcebível*, *incontornável*.

Dessa forma, é plausível afirmar que as sequências *a Maria bêbada*, na sentença em (36a), e *alunas que bebem*, em (36b), podem referir-se a situações, como pode ser inferido dos exemplos em (38). As sentenças em (39) nos mostram, ao contrário, que essa sequência dificilmente pode fazer referência a um evento (39a) ou a uma proposição (39b).

- (38) a A Maria bêbada é intolerável / inconcebível / passageiro / incontornável / irremediável.  
 b Alunas que bebem é intolerável / inconcebível / passageiro / incontornável / irremediável.
- (39) a \*A Maria bêbada / alunas que bebem aconteceu ontem.  
 b \*A Maria bêbada / alunas que bebem é verdade.

Em (40), pode-se confirmar que esses constituintes sujeitos não podem se combinar com predicados que atribuem propriedades exclusivamente a indivíduos, como *orgulhoso* ou *preguiçoso*. Nesses casos, a concordância é obrigatória (41).

- (40) a \*A Maria bêbada é orgulhoso / preguiçoso.  
 b \* Alunas que bebem é orgulhoso / preguiçoso.
- (41) a A Maria bêbada é orgulhosa / preguiçosa.  
 b Alunas que bebem são orgulhosas / preguiçosas.

A partir dessa breve discussão, pode-se estabelecer uma relação, nos casos examinados, entre o fenômeno da concordância neutra e predicados que, segundo Hegarty (2003), selecionam uma situação como argumento. Descreveremos a seguir os tipos de sujeitos passíveis de aparecerem como sujeito dessas construções.

### A SC como sujeito da cópula

Com relação à SC sujeito da cópula, é possível observar uma interação entre o tipo do predicado da SC e o tipo do predicado principal. Por exemplo, com o predicado principal *intolerável*, a SC não pode ser construída com um predicado *individual-level*, a não ser que se atribua a esse predicado uma leitura de propriedade transitória, como pode-se ver em (42). A sentença em (42a) só pode ser interpretada como (42c), e não como (42b).

- (42) a ??A Maria inteligente é intolerável / passageiro.

- b A Maria ser inteligente é intolerável.
- c ?A Maria estar inteligente é intolerável.

Já no caso de predicados como *inconcebível*, existe a possibilidade de os predicados da SC serem do tipo *individual-level* (43a). Nesse caso, (43a) pode ser interpretada como (42b) ou como (43c).

- (43) a A Maria inteligente é inconcebível.
- b A Maria ser inteligente é inconcebível.
- c ?A Maria estar inteligente é inconcebível.

Uma discussão detalhada sobre a ontologia das entidades abstratas está fora do escopo deste trabalho. Ainda assim, gostaríamos de observar que, nesses casos, é possível falar de dois tipos de situação. De um lado, tem-se uma situação que pode ser descrita como *realis*, no sentido de que a situação a que se faz referência já ocorreu ao menos uma vez; essa entidade apresentaria uma compatibilidade maior com predicados como *intolerável*. De outro, tem-se uma situação que pode ser descrita como *irrealis*, no sentido de que a situação a que se faz referência existe como possibilidade; essa entidade apresentaria uma compatibilidade maior com predicados como *inconcebível*.

Com relação ao sujeito da SC, não parece haver restrições: eles podem ser genéricos ou existenciais, definidos ou não (44).

- (44) (As/minhas) aluna(s) bêbada(s) é inconcebível / intolerável.

## O DP como sujeito da cópula

Nos casos em que o sujeito da construção copular é um DP,<sup>8</sup> é possível notar que há uma restrição contra DPs definidos específicos (45). Contudo, ainda que o DP seja composto por nominais genéricos, esses nominais, quando são concretos, devem apresentar algum tipo de modificação (46).

- (45) \*A(s) aluna(s) que bebe(m) é intolerável / inconcebível.
- (46) a \*Aluna / alunas é intolerável / inconcebível
- b Aluna(s) que bebe(m) é intolerável / inconcebível.
- c Uma aluna que bebe é inconcebível.
- d Tristeza é passageiro.

Além da modificação pela relativa, outros casos de modificação são igualmente possíveis, como em (47). Considerando a possibilidade de que o sujeito da cópula possa ser também uma SC, como discutido acima, levantamos a hipótese de que a sentença em (47) é ambígua, com seu sujeito podendo ser analisados como uma SC ou como um DP. Ou seja, o adjetivo *bêbada* poderia ser analisado como predicado de uma SC ou como modificador de *alunas*. Explicando melhor, no caso de uma SC, como vimos acima, o seu sujeito pode ser específico ou genérico (44). Já nos casos em que o sujeito é um DP, ele

<sup>8</sup> Há propostas de se analisar esse DP como uma infinitiva reduzida. Ver em Foltran e Rodrigues (2012) evidências contra essa análise.

deve ser genérico (46). Assim, o nominal da sentença em (47), que é genérico, poderia ser tanto sujeito de uma SC quanto um DP modificado. Os exemplos em (48) mostram mais claramente a diferença entre os dois tipos de sujeito. Em (48a), *chatas* é um modificador de *aulas*, o sujeito da sentença é um DP definido específico, a sentença é inaceitável; já em (48b), *chatas* funciona como predicado em uma SC que tem como sujeito *as aulas do Pedro*. Essa SC é sujeito da sentença, nesse caso, aceitável.

(47) a Alunas bêbadas é inconcebível. (alunas estarem bêbadas / alunas que bebem)

(48) a \*As aulas chatas do Pedro é inconcebível.

b As aulas do Pedro chatas é inconcebível.

Vale notar que os adjetivos presentes nos DPs sujeitos dessas construções devem ser do tipo predicativo. DPs construídos com os chamados adjetivos argumentais ou não predicativos (NEGRÃO et al., 2008; RIO-TORTO, 2006) não são passíveis de denotar uma situação, como ilustrado em (49)-(51).

(49) a \*Decisões presidenciais é intolerável / inconcebível.

b Decisões precipitadas é intolerável / inconcebível.

(50) a \*Flores campestres é intolerável / inconcebível.

b Flores perfumadas é intolerável / inconcebível.

(51) a \*Sistemas solares é intolerável / inconcebível.

b Sistemas vulneráveis é intolerável / inconcebível.

Devemos observar ainda que os dados discutidos acima se distinguem dos dados em (52a), (53a) e (54a), em que se tem igualmente o fenômeno da concordância neutra. Nesses casos, as expressões que ocupam a posição de sujeito não podem ser analisadas como uma situação, como pode ser visto em (52b), (53b) e (54b), ainda que os predicados presentes nessas sentenças sejam compatíveis com SCs e com os tipos de DP apresentados acima, como ilustrado em (55). Além disso, o DP não precisa ser modificado, como no caso de denotação de uma situação, mas deve ser genérico (ver exemplos em (c)). No caso das sentenças em (52)-(54), os predicados se aplicam a uma atividade relacionada ao sujeito.

(52) a Água é bom para a saúde.

b \*Água é intolerável / passageiro

c \*A água da mina do meu sítio é bom para a saúde.

(53) a Coca-cola é perigoso.

b \*Coca-cola é intolerável / irreversível

c \*A coca-cola que eu comprei ontem é perigoso.

(54) a Crianças é divertido / cansativo.

b \*Crianças é intolerável / inconcebível.

c \*As crianças do meu vizinho é cansativo.

- (55) a Alunas que bebem é cansativo.  
b A Maria bêbada é divertido.

Vimos nesta seção, primeiramente, que o constituinte sujeito da sentença em (1b) pode ser analisado como uma SC, explicando dessa forma a concordância neutra observada. Em seguida, notamos que, além de SCs, DPs também podem aparecer como sujeito de construções copulares que apresentam concordância neutra. Não discutiremos neste artigo a estrutura desses DPs, nem o que permitiria essa concordância neutra. As propriedades descritas aqui, no entanto, apontam questões que devem ser consideradas em estudos futuros. Vimos que SCs e DPs sujeitos dessas construções podem denotar uma situação. Vimos também que DPs podem denotar uma atividade/eventualidade. Assim, será necessário explicar o que esses constituintes têm em comum que permite a produção dos efeitos semânticos observados; isto é, que permite que eles sejam compatíveis com predicados que teriam como argumento uma situação ou uma atividade. Além disso, se a concordância deve ser tratada em termos de valoração de traços-phi presentes no DP, no adjetivo e na cópula (CHOMSKY, 2000, 2001), deve-se explicar por que não há concordância entre o DP, de um lado, e a cópula e o adjetivo, de outro, mas há concordância dentro do DP, entre determinante, nome e adjetivos.

### Considerações finais

O objetivo desse trabalho foi discutir os dois tipos de sentenças com cópula apresentados em (1). Vimos que as duas sentenças se estruturam sintática e semanticamente de maneiras diferentes. Em (1a), o adjetivo *bêbada* é um predicado secundário e o predicado *chata* se aplica ao indivíduo Maria. Observamos que *bêbada* tem como leitura preferencial a de um predicado depictivo circunstancial, com desdobramentos em escopo de foco e negação. Assumimos que, nessa interpretação, o predicado secundário está adjungido a IP. Na interpretação de depictivo típico, menos saliente, o predicado secundário está adjungido a VP.

Por sua vez, a sentença em (1b) não expressa uma predicação sobre um indivíduo, mas trata-se de uma predicação sobre uma situação. SC são constituintes que podem denotar situações e é essa a estrutura que assumimos para o sujeito de (1b). Considerando-se a análise clássica de que a cópula seleciona uma SC, temos aí um exemplo de uma SC dentro de outra SC, ou seja, a SC selecionada pela cópula *ser* tem como sujeito uma SC. Essa análise, no entanto, não explica o fato da concordância neutra ou *default*. O exemplo em (6) mostra que essa concordância especial se dá também com constituintes diferentes de SC. Mostramos que a sentença em (6) é diferente de (1b) e deixamos em aberto a explicação técnica para o padrão especial de concordância.

### REFERÊNCIAS

- AARTS, B. *Small Clauses in English: the nonverbal types*. Berlim: Mouton de Gruyter, 1992.
- ASHER, N. *Reference to abstract objects in discourse*. Dordrecht: Kluwer Academic Publishers, 1993.

- BALTIN, M. R. A nonargument for Small Clauses as constituents. *Linguistic Inquiry*, Cambridge, v. 29, n. 3, p. 513-515, 1998
- CARREIRA, M. *Diagnósticos de constituição para construções predicativas adjetivais*. 2008. 116 f. Dissertação (Mestrado em Linguística) – Universidade Federal do Paraná, Curitiba.
- CHOMSKY, N. Derivation by phase. In: KENSTOWICZ, M. (Org.). *Ken Hale: a Life in Language*. Cambridge, Mass.: MIT Press, 2001. p. 1-52.
- \_\_\_\_\_. Minimalist inquiries: the framework. In: MARTIN, R.; MICHAELS, D.; URIAGEREKA, J. (Org.). *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press, 2000. p. 89-155.
- CONTRERAS, H. Small Clauses in Spanish and English. *Natural Language and Linguistic Theory*, New York, v. 5, n. 2, p. 225-243, 1987.
- FOLTRAN, M. J. *As construções de predicação secundária no português do Brasil: aspectos sintáticos e semânticos*. 1999. 206 f. Dissertação (Doutorado em Linguística) – Universidade Federal do Paraná, Curitiba.
- FOLTRAN, M. J.; RODRIGUES, P. *On denoting abstract entities*. (In)Definites and Weak Referentiality. Florianópolis, 2012. (Comunicação oral).
- FOLTRAN, M. J.; RODRIGUES, P.; CARREIRA, M.; MOREIRA, T. D. *Complex Small Clauses constructions in Brazilian Portuguese*. IX Workshop on Formal Semantics. Rio de Janeiro, 2012. (Comunicação Oral).
- GEUDER, W. *Oriented adverbs: issues in the lexical semantics of event adverbs*. 2002. 220 f. Tese (Doutorado em Linguística) – Tübingen: Universidade de Tübingen.
- HAEGEMAN, L.; GUÉRON, J. *English grammar: a generative perspective*. Oxford: Blackwell, 1999.
- HEGARTY, M. Semantic types of abstract entities. *Lingua*, Cambridge, v. 13, p. 891-927, 2003.
- HIMMELMANN, N. P.; SCHULTZE-BERNDT, E. Issues in the syntax and semantics of participants-oriented adjuncts: an introduction. In: HIMMELMANN, N. P.; SCHULTZE-BERNDT, E. (Org.). *Secondary predication and adverbial modification: the typology of depictives*. Oxford, New York: Oxford University Press, 2005. p. 1-68.
- NEGRÃO, E.V.; MÜLLER, A. L.; NUNES-PEMBERTON, G.; FOLTRAN, M. J. O adjetivo. In: ILARI, R.; NEVES, M. H. M. (Org.). *Gramática do português culto falado no Brasil*. Campinas: Editora Unicamp, 2008. p. 371-396.
- NICHOLS, J. *Predicate Nominals: a partial surface syntax of Russian*. Berkeley: University of California Press, 1981.

RIO-TORTO, G. Para uma gramática do adjetivo. *Alfa*, São Paulo, v. 50, n. 2, p. 103-129, 2006.

SAFIR, K. On Small Clauses as constituents. *Linguistic Inquiry*, Cambridge, v. 14, n. 4, p. 730-735, 1983.

SIMPSON, J. Depictives in English and Warlpiri. In: HIMMELMANN, N. P.; SCHULTZE-BERNDT, E. (Org.). *Secondary predication and adverbial modification: the typology of depictives*. Oxford, New York: Oxford University Press, 2005. p. 69-106.

# A subversão da norma e a produção de sentidos nos deslocamentos enunciativos

(The subversion of normativity and the production of meaning in enunciative displacements)

Stéfano Grizzo Onofre<sup>1</sup>

<sup>1</sup> Programa de Pós-Graduação em Linguística – Universidade Federal de São Carlos (UFSCar)

tefo42@hotmail.com

**Abstract:** A grammatical revision treaty about Machado de Assis's literary production conceives the asymmetry between the category of person and subject only as an infraction of syntactic rules. That posture disregards the meanings implicated in the use of those pronouns. By problematizing a normative view, we intend to discuss the viability of creating labels for every shift of personal pronouns as a mistake. The grammatical revision treaty of Machado de Assis's work, developed by the grammarian Lima (1973), disregards the possibility of non-existent meaning problems in occurrences that possess syntactic asymmetry. In general, the grammatical treaties disregard the uttering activity. Such posture illustrates how normativity, in the context of learning and teaching, can devalue student's linguistic constructions paying attention only to the form and the syntactic rules. The work of Lima represents the grammatical purism and, therefore, disregards that an utterance can be acceptable even if it is not in a syntactic pattern. With such a patterning difficulty in the innumerable grammars, we pose the following question: how can one demand from the students the appropriation of their texts if one deprives them of the right to operate with language. We defend, therefore, that there is a difficulty of defining a distinction, before the materialization and analysis of an utterance, if a linguistic occurrence is a subversion that derives from the ignorance of the linguistic code or if it is a creative subversion of the same linguistic code.

**Keywords:** language activity; uttering; theory and linguistic analysis; learning and teaching.

**Resumo:** Um tratado de revisão gramatical sobre a obra de Machado de Assis concebe a não concordância da categoria de pessoa deflagrada pelo verbo em relação a seu respectivo sujeito apenas como uma infração das regras sintáticas. Dessa forma, desconsidera os efeitos de sentidos que tal mudança acarreta. Por meio da problematização dessa visão normativa pretendemos discutir se é produtivo etiquetar como erro toda mudança de pronomes pessoais que infrinja uma norma. O trabalho de revisão da obra de Machado de Assis, desenvolvido pelo gramático Lima (1973), desconsidera a possibilidade de que a assimetria sintática no emprego de pronomes pode, em certas ocasiões, não comprometer o efeito de sentido do texto. Dessa forma o trabalho da gramática desconsidera a enunciação. Tal postura ilustra como a normatividade, no contexto de ensino e aprendizagem, pode desvalorizar a construção linguística dos alunos apenas considerando a forma e as regras sintáticas. O trabalho de Lima representa o purismo gramatical e, portanto, desconsidera que um enunciado pode ser aceitável mesmo que ele não esteja sintaticamente padronizado. Tendo em vista essa dificuldade de padronização, questionamos como é possível exigir dos alunos a apropriação de seus textos se, muitas vezes, é-lhe privado o direito de operar com as marcas linguísticas? Defendemos, portanto, pautados na Teoria das Operações Predicativas e Enunciativas (TOPE), a dificuldade em se demarcar com precisão, antes da manifestação de um enunciado, se uma ocorrência procede de um desconhecimento do código linguístico ou de uma subversão criativa do próprio código.

**Palavras-chave:** linguagem; enunciação; teoria e análise linguística; ensino e aprendizagem.

## Introdução

No presente artigo objetivamos discutir como a diferença entre enunciado e frase, proposta por Culioli (1999a), contribui para reorientar o valor que se pode atribuir a ocorrências de deslocamentos enunciativos da categoria de pessoa. Essa reorientação leva-nos a considerar que um conjunto de coordenadas enunciativas pode tornar aceitáveis ocorrências linguísticas que não estão sintaticamente padronizadas. O embate entre frase e enunciado pressupõe diferentes concepções sobre a noção do fazer criativo na linguagem. Pensamos ser importante propor uma articulação de frase e enunciado porque ela abre caminhos para se pensar em ocorrências de deslocamentos enunciativos em produções textuais de vestibular não apenas como um “defeito”. Para situar a discussão, partimos de um exemplo em que uma abordagem de orientação tradicional condena uma passagem do texto de Machado de Assis (apud LIMA, 1973). Pensamos que esse diálogo é importante porque o modo de considerar uma ocorrência de deslocamento enunciativo presente no texto de Machado de Assis mobiliza uma noção de criatividade que desconsidera a possibilidade de haver valor em uma representação linguística que infringe parâmetros sintáticos.

## A criatividade na linguagem do ponto de vista da Teoria das Operações Predicativas e Enunciativas (TOPE)

O termo criatividade abre um leque de concepções teóricas que podem ultrapassar o domínio da linguística. A noção de criatividade não se restringe apenas ao domínio da língua. Inúmeras manifestações humanas podem ter alguma relação com a noção de criatividade. Restringimos nossa discussão apenas ao domínio linguístico. No entanto, pretendemos abordar noções da ordem da linguagem articuladas às noções de ordem da língua. A criatividade é considerada, portanto, resultado de um modo de se conceber a atividade de produção e reconhecimento de textos. Na tradição de estudos culioliana, toda forma linguística remete a um conjunto de operações que podem ser recuperadas a partir de um conjunto de coordenadas enunciativas que envolvem uma relação entre uma léxis<sup>1</sup> e um Sit (traço de orientação situacional). A recuperação dessas operações demonstra o infinito potencial de organização dos textos orais e escritos. Conceber os textos orais e escritos como traços dessas operações é equivalente a conceber a linguagem como um trabalho de produção de significado. Nesse contexto, a criatividade é constitutiva da atividade de linguagem.

Para matizar como a teoria culioliana concebe a criatividade apoiamos-nos em Rezende (2010), para quem a criatividade não é considerada apenas o ápice de um processo. Na reflexão de Rezende (2010), a criatividade é também a capacidade de se orientar em relação às representações linguísticas construídas seja compreendendo, seja reformulando os materiais disponíveis à representação. Muitas vezes exclui-se da noção de criatividade o olhar para mecanismos que são responsáveis por gerar<sup>2</sup> uma dada representação linguística. A concepção de criatividade presente em Rezende aproxima-se também das considerações de Franchi (2006), para quem a linguagem é uma atividade estruturante. O olhar para os

1 Para Culioli (1999a), a léxis é similar a um conteúdo proposicional e garante a união sintática e semântica do predicado presente em um dado fragmento de texto. A léxis garante a plasticidade do enunciado e, ao mesmo tempo, sua identidade.

2 Não no sentido gerativista, mas no sentido das operações de representação, referenciação e regulação envolvidas na teoria de Culioli (1999a).

processos constitutivos da linguagem coloca em cheque a desarticulação entre sintaxe e semântica e, conseqüentemente, entre erro e criatividade.

A partir da Teoria das Operações Predicativas e Enunciativas almejamos olhar para as ocorrências linguísticas buscando revelar como elas se estruturam. Desse modo, pretendemos mostrar que existe uma simultaneidade de cenários enunciativos deflagrados pelos agenciamentos de marcas linguísticas. A simultaneidade de cenários enunciativos pode tanto comprometer o texto quanto valorizá-lo.

### **A diferença entre enunciado e frase**

O seguinte predicado em francês ilustra uma insuficiência da gramática segundo os parâmetros da Teoria das Operações Predicativas e enunciativas: (1) *Un chien aboie*.<sup>3</sup> Essa espécie de predicado é o ponto de partida da reflexão gramatical. Ela sustenta como centro de sua concepção teórica frases genéricas assim como (2) *O homem é mortal* ou (1) *Un chien aboie*, e, dessa maneira, prioriza a modalidade assertiva. Em (1) e (2), é possível, por meio de uma comparação grosseira entre predicados genéricos, o ensino de algumas propriedades gramaticais clássicas; como organização dos constituintes da frase em sujeito e predicado, transitividade etc. Em línguas ocidentais essa concepção gramatical pode induzir ao equívoco, pois parece apagar distâncias significativas entre as propriedades psico-físico-culturais e entre as representações de mundo sustentadas culturalmente.

Segundo Milner (1992), para Culioli (1999b) um enunciado como (1) *Un chien aboie* é mal formado. Isso significa que é muito maior a probabilidade de aparecer situado enunciativamente na fala do dia a dia como (3) *il y a du chien qui aboie*,<sup>4</sup> ou (4) *C'est un chien qui aboie*.<sup>5</sup> A marca *il y a* é o traço de orientação situacional (Sit) que aponta a existência de um enunciador sustentando sua enunciação. Esses enunciados são quase paráfrases e possuem diferenças significativas, pois é bem diferente dizer (3) *Tem um cachorro que late* e (4) *Esse é um cachorro que late muito* ou (5) *Olha, lá está um cachorro que late*.

Ao se levar em consideração que um enunciado é significativo somente quando aponta para um cenário enunciativo, defendemos que o ponto de partida para se conceber o ensino e aprendizagem de produção/interpretação de textos é demonstrar como o léxico e a gramática do português só têm propósito quando estão submetidos a marcas espaciais, temporais e pessoais. Nesse tipo de reflexão a gramática não está restrita aos parâmetros do analista, mas sim aos parâmetros de boa formação de enunciados orientados pelo sujeito enunciador. Esses parâmetros são construídos no próprio texto em uma relação dialógica entre analista e as operações enunciativas deflagradas pelas marcas linguísticas.

### **Breve descrição da revisão de Machado de Assis**

Antes de analisar o exemplo do estudo gramatical de Lima (1973), pensamos ser pertinente contextualizar a obra com uma breve resenha das intenções do autor em relação à obra de Machado de Assis.

3 No original: Um cachorro late.

4 No original: Tem um cachorro que late.

5 No original: Esse é um cachorro que late.

A revisão de Machado Assis (apud LIMA, 1973) agrupa inúmeras passagens da obra do autor e as submete a críticas. São expostos problemas sintáticos, problemas de regência verbal, problemas de assimetria no uso dos pronomes pessoais, etc. O gramático Lima (1973), após apontar a suposta infração, cita inúmeros exemplos de autores consagrados da literatura portuguesa e brasileira para demonstrar como o texto de Machado de Assis subverte o uso “cultivado” da língua. O início do livro deixa claro o objetivo da revisão de Machado de Assis: desconstruir a imagem de que o romancista, conforme Rui Barbosa (apud LIMA, 1973), foi modelo de pureza e correção no que diz respeito ao uso da língua. Lima (1973) cita também o crítico literário Veríssimo, que diz:

Outra distinção, que por assim dizer salta aos olhos, da sua obra poética, desde seus primeiros versos, é o da sua língua e metrificacão. Exceptuando sempre Gonçalves Dias, a sua língua é incomparavelmente mais pura, mais rica, mais copiosa, e a sua versificação mais correcta, mais difficil, mais elegante de que a qualquer daquelles poetas. (VÉRISIMO, 1904, p. 91 apud LIMA, 1973, p. 24)

Após a exposição de outros escritores, alguns criticando a escrita de Machado de Assis, outros elogiando, Lima (1973) pretende demonstrar que, embora Machado de Assis tenha tido seus méritos literários, ele não foi “[...] letrado de sintaxe indiscutível [...]” (p. 30).

O trabalho de Lima (1973) representa o polo extremo da concepção tradicional da gramática. O trabalho veicula noções de linguagem já rechaçadas pela linguística contemporânea, como, por exemplo, a noção de que a infringência do padrão de uso das formas linguísticas aproxima o literato, que em tese deveria ser um homem cultivado, do homem tosco, ignorante, etc.

O prólogo do texto de Lima (1973) tenta dar um caráter objetivo às correções do texto de Machado de Assis:

Serei justo, fazendo exame imparcial de construções menos puras, nas quais o autor de Dom Casmurro não soube ou não quis polir as frases. Manifesto é que não irei analisar oração por oração, para mostrar erros que lhe maculem a glória. (LIMA, 1973, p. 19)

A passagem possui uma contradição: como é possível existir imparcialidade se o autor qualifica de antemão as construções como “menos puras”? E, ainda, há imparcialidade quando Lima (1973) conjectura a possibilidade de que o autor desconheça o polimento das frases? A imparcialidade de Lima (1973) perde força na passagem em que afirma que não irá analisar oração por oração para não macular a glória do autor. De certa forma, é possível recuperar a ideia de que a glória de Machado de Assis em algum momento dependeu da força da análise do gramático.

O prólogo da revisão Lima (1973) expõe também a preocupação de que suas observações sobre os problemas do texto de Machado de Assis poderiam ser entendidas como vaidade ou pedanteria. O texto de Lima (1973) deixa traços dessa preocupação no momento em que o gramático afirma não ter animosidade contra Machado de Assis:

Animosidade contra Machado de Assis nunca a tive. Também não sou mordido de inveja, picado de vaidade, ou mal ferido da ambição que põe no escândalo a base da popularidade. Não sou literato, não sou acadêmico, não frequento grêmio literário. Sou apenas estudioso anônimo, simples leitor que busca conhecer os mestres da língua pátria. (LIMA, 1973 p. 19)

Uma análise atenta do texto de Lima (1973) revela a imagem de um sujeito enunciad<sup>6</sup> (que é diferente do locutor) dialogando com um possível leitor do tratado da revisão. Isso implica dizer que a argumentação do prefácio de Lima (1973) ao construir a negação de algumas características sobre o próprio sujeito enunciad<sup>6</sup> constrói a afirmação delas também. Por exemplo, quando (LIMA, 1973) afirma “não sou mordido de inveja, picado de vaidade”, é possível recuperar o enunciado complementar “sou mordido de inveja, picado de vaidade”. Quando se nega algo, o acesso à afirmação não está bloqueado, ele se mantém como uma possibilidade. E, embora o enunciad<sup>6</sup> se esforce para se afastar da penderia e da vaidade, ela é realçada na conclusão de seu prólogo:

Este livro encerra assunto erudito. Eruditos devem ser também os termos, se não todos, pelo menos alguns.

Entre dois modos de compor a frase, preferi o mais culto, dizendo frutos de vez, no lugar de frutos sazonados.

Ao escolher o vocábulo menos vulgar, e a redação próxima da clássica não o fiz por penderia, senão para me ombrear com a cultura do leitor. Não é pedante o que usa as galas da boa linguagem, mas o que põe na escrita atavios de ouropel. (LIMA, 1973, p. 21)

Para Lima (1973), ser pedante é colocar enfeites fajutos na “boa linguagem”. Fica claro que Lima, de modo sutil, qualifica o autor Machado de Assis como pedante porque finge saber escrever quando na realidade não o sabe. Dessa forma, ao julgar Machado de Assis, a modéstia de Lima fica comprometida.

Não pretendemos negar ou afirmar os méritos de Machado de Assis como escritor porque essa discussão diz respeito à crítica literária. Nossa intenção é trabalhar com o material linguístico e mostrar que o linguista não deve apenas qualificar o texto como criativo ou não criativo, mas sim forjar explicações racionais para os fenômenos em observação. O intuito, portanto, é, por meio do conceito de paráfrase de Culioli (1999a), mostrar como as correções que Lima sugere para o texto de Machado de Assis desconsideram que a mudança de um constituinte linguístico em favor da adequação sintática transforma o sentido presente na ocorrência do texto do escritor literário.

Antes de transcrever a análise de Lima, colocamos uma discussão do modo como a significação é vista na Teoria das Operações Predicativas e enunciativas.

## Noção

O modelo enunciativo de Culioli (1999b) está ancorado em uma teorização da atividade de produção de significação ofertada pela linguagem. Para compreender qual a proposta do autor é necessário recuperar a ideia de que a linguística tem por objetivo apreender a linguagem por meio das línguas naturais. Para dar conta desse objetivo há que se reconhecer a complexidade presente no diálogo. A teoria leva em consideração que a

---

6 Referimo-nos a Lima (1973) como sujeito enunciad<sup>6</sup> no intuito de demarcar a diferença entre o sujeito Lima de carne e osso e o sujeito enunciad<sup>6</sup> que recuperamos em seu texto. A análise das marcas linguísticas do prólogo constrói um sujeito enunciad<sup>6</sup> que, ao desqualificar o enunciatário, acaba qualificando a própria imagem do enunciad<sup>6</sup>. O texto, ao tentar construir uma imagem de modéstia, acaba paradoxalmente negando a modéstia. Isso significa que, apenas pelo texto, não se pode julgar que a pessoa de carne e osso Lima não seja modesta. Além disso, deve-se respeitar o contexto em que essa análise foi produzida, pois ela ilustra um modo muito corrente de se conceber a linguagem escrita no ensino.

relação entre dois sujeitos está sempre em construção. Na verdade, existe uma distância permanente entre os sujeitos. Nesse contexto, a língua encontra-se em uma relação complexa de interioridade e exterioridade. A linguística orientada a partir da tradição saussuriana, em vez de abarcar o jogo complexo instaurado pela linguagem, calibrou-o em benefício da construção de um objeto estático. Podemos encontrar essa calibração no clássico esquema de comunicação saussuriano em que o diálogo é concebido como a troca mecânica de informações que vai linearmente do locutor ao interlocutor.

Para abarcar a complexidade do diálogo, Culioli (1999b) expõe que a produção de um enunciado gera retroação e, quando o som atinge o outro locutor, não há recepção e sim uma nova construção. Ignorar essa condição complexa do diálogo significa dar atenção ao código linguístico neutro e independente da atividade de linguagem dos locutores. O código linguístico, embora possuindo uma relativa estabilidade, está organizado de modo diferente para cada locutor. Essa condição é bem singularizada por Culioli (1999b) quando ele a resume por meio do seguinte aforismo: a compreensão é um caso particular de mal-entendido. Essa ideia revela que a compreensão é uma conquista e, ao mesmo tempo, a incompreensão também é sempre uma possibilidade.

No âmago dessa relação de troca entre sujeitos, Culioli (1999b) identifica a atividade epilinguística. A atividade epilinguística é uma atividade metalinguística pré-consciente, presente em todo sujeito. Essa atividade é responsável pela capacidade do sujeito dialogar consigo mesmo e com o outro. A atividade epilinguística é regida pelas operações de representação, referenciação e regulação. O sujeito tem acesso a noções psico-físico-culturais, que são apreendidas pelo material linguístico e em seguida reguladas pelo diálogo do sujeito consigo mesmo e pelo diálogo do sujeito com outro sujeito.

Uma noção não está disposta de modo especular em relação ao léxico. A noção coloca virtualmente um conjunto de traços abertos que são acessados pela sua presença em um enunciado. Para Culioli (1999b) uma noção apenas pode ser apreendida em qualidade sendo sua quantificação responsável por sua materialização em um enunciado. A hipótese central da teoria reside na aposta de que o enunciado é resultado de um conjunto de operações associadas à atividade metalinguística pré-consciente dos sujeitos enunciadoreis.

A língua comporta, então, a determinação e a indeterminação. Na Teoria das Operações Predicativas e Enunciativas a determinação advém da manifestação de noções<sup>7</sup> em uma materialidade linguística e a indeterminação, da necessidade de fragmentação dessa noção em forma de predicados. Essa operação é a quantificação de uma qualidade, ou seja, os textos orais e escritos são traços de operações de linguagem corporificadas em um sistema de referência linguístico. Assim, na palavra vaca são arrolados, virtualmente, todos os tipos de ocorrências possíveis da noção ser vaca ( $A r ( )$ ). Apenas a fragmentação (quantificação) dessa noção pelo empírico possibilita a existência de uma noção qualitativa do que é vaca. Por exemplo, as ocorrências *A vaca pasta no campo*, *A vaca dá leite*, *A vaca é um animal sem penas*, *A vaca é um mamífero* demonstram que a noção vaca inclui tudo o que pode ser predicado como vaca. Ao derivar os possíveis predicados de vaca encontramos

---

7 Uma massa de fenômenos que Culioli (1999b) diz ser um conjunto de associações semânticas estabilizados pela experiência. Essas cadeias são elaboradas de diversas maneiras, por exemplo, pela atividade onírica ou emblemática, por imagens, etc. A noção propõe a articulação entre o linguístico e o não linguístico. A noção é uma representação sem materialidade, ou uma representação em que a materialidade não é acessível ao linguista.

apenas algumas ocorrências típicas da noção vaca. Se o trabalho com a linguagem tomar como válido apenas as ocorrências típicas de uma noção, perdemos a alteridade, o impossível, as ocorrências diferenciais que reorganizam uma representação, que, na verdade, por uma relação de alteridade, garantem a própria identidade de uma noção. Ocorrências que soam estranhas como, por exemplo, *A vaca é uma mulher*, são imprescindíveis para a reorganização de uma noção. Elas trazem a ideia de que é apenas o que não é vaca que pode garantir tudo o que está na noção ser vaca. Além disso, essas ocorrências podem ser validadas em contextos precisos de uso, como, por exemplo, *Aquela mulher é uma vaca*. Esse predicado ainda estaria relacionado a ocorrências típicas da propriedade típica do animal vaca? Saberíamos ainda identificar a origem histórica do uso da palavra vaca em relação à mulher? Talvez até possamos reencontrar essa origem histórica, mas o interessante da linguagem é que, mesmo desconhecendo a origem, o termo pode comportar os valores particulares do emprego dessa noção. Isso atesta para a inexistência de uma relação simétrica entre propriedades psico-físico-culturais e uma noção. Portanto, colocar como determinada uma relação entre tais propriedades e uma marca linguística é apenas um ponto provisório, um ponto de chegada, o qual, de modo extremamente fugaz, advém de todo um trabalho linguageiro praticamente excluído das reflexões linguísticas tradicionais.

### **Descrição de operações fundamentais para sustentar a análise**

Para construir nosso referencial de análise necessitamos de expor as operações fundamentais que o quadro da TOPE oferece. Essas operações são generalizações criadas por Culioli (1999a, 1999b), elas se situam no quadro metalinguístico da teoria. Além disso, as relações primitiva, predicativa e enunciativa são o resultado formal<sup>8</sup> de um olhar para as formas linguísticas. O nível metalinguístico oferecido por Culioli (1999a, 1999b) não está em uma relação termo a termo com os constituintes linguísticos. Todas as formalizações do estudioso oferecem uma síntese de problemas sintáticos e semânticos, sem a preocupação de estabelecer cortes entre as duas dimensões. Ao proceder dessa maneira apreendemos as ocorrências por meio da simulação parafrástica dos enunciados. Mas essas simulações situam-se em um espaço híbrido entre nossa atividade epilinguística pré-consciente e uma atividade epilinguística orientada pela TOPE e, portanto, metalinguística.

Essas operações de linguagem fazem-se visíveis com o auxílio de um sistema de representação metalinguístico que dá suporte para o analista manipular o texto. Munido da sua subjetividade, o analista faz aflorar as operações postas em jogo pela enunciação. Os meta-operadores, forjados pelo sistema de representação metalinguístico culioliano, servem de suporte para que o analista realize um cálculo dos elementos postos em causa pelos marcadores linguísticos. Dessa forma, o conceito de léxis é um representante direto de relações primitivas derivadas diretamente da noção. A noção coloca em jogo um pacote de relações psico-físico-culturais que estão disponíveis à experiência do enunciador. O trabalho do analista resume-se, portanto, na explicitação da atividade linguagística e na busca de invariantes subjacentes às variáveis superficiais dos marcadores linguísticos. A léxis é similar a um conteúdo proposicional e garante a união sintática e semântica do predicado presente em um dado fragmento de texto. A léxis garante a plasticidade do enunciado e, ao mesmo tempo, sua identidade. O analista deve manipular um enunciado

---

<sup>8</sup> No sentido de uma representação metalinguística construída conscientemente pelo linguista

para extrair as operações postas em jogo. A fórmula de Culioli (1999a, 1999b)  $\lambda \in \langle \text{Sit}_2 (S_2, T_2) \in \text{Sit}_1 (S_1, T_1) \in \text{Sit}_0 (S_0, T_0) \rangle$  ilustra bem a relação entre um Sit e uma léxis. A léxis se desdobra em dois argumentos, por exemplo X e Y, unidos por um relator  $\pi^9$ ; por exemplo, o enunciado *Plácido escreveu a carta* pode ser representado por três termos que a teoria nomeia de léxis: a notação  $\lambda$  equivale a  $\langle \text{Plácido}, \text{escrever}, \text{carta} \rangle$ . Uma léxis pode sofrer alterações mediante sua localização em relação traço de orientação situacional, ou seja, um Sit. Nada impede que o Sit seja outra léxis. A léxis é um termo que permite visualizar o enunciado como sendo parte de famílias parafrásticas. Um exemplo seria:

Plácido escreveu uma carta  
Plácido está escrevendo a carta  
Plácido gostaria de escrever uma carta  
Plácido! Escreveu uma carta? Jamais!  
Foi Plácido que escreveu a carta  
Talvez tenha sido Plácido que escreveu esta carta.  
A carta de Plácido foi escrita em duas semanas.

A léxis garante a unidade de todas as proliferações acima.

O Sit sempre dá suporte para qualquer léxis, no entanto, ele pode ser mais ou menos explicitado pelos marcadores linguísticos que veiculam a noção de pessoa. As marcas morfológicas presentes no verbo *escrever* orientam o enunciado em relação ao tempo-espaco e ao enunciador. Todo enunciado está orientado (operação de orientação é notada por  $\in$ ) em relação a um parâmetro abstrato notado de Sit<sub>0</sub> (S<sub>0</sub>, T<sub>0</sub>). Esse parâmetro indica a pessoa e o tempo-espaco origem da enunciação.

As dissertações argumentativas são prenes de estratégias discursivas que visam a dar suporte para o Sit<sub>0</sub> ao mesmo tempo em que escondem a categoria de pessoa. Esse fenômeno é conhecido como objetividade.

Uma relação primitiva está no mesmo nível de uma léxis, porém é possível perceber que entre  $\langle \text{Plácido}, \text{escrever}, \text{carta} \rangle$  existe potencial de organização. O potencial de organização é uma relação primitiva, pois permite a visualização de noções como agentividade, quantificação e qualificação. Note-se o estranhamento causado por uma inversão na lexis  $\langle \text{carta}, \text{escrever}, \text{Plácido} \rangle$ . O sentido em *A carta escreveu Plácido* tem menos orientação do que a ordem contrária. Para validar a léxis  $\langle \text{carta}, \text{escrever}, \text{Plácido} \rangle$  há que considerar cenários enunciativos em que carta possa ter agentividade. Por exemplo, *Leio a carta de Plácido e vejo ele escrito em suas páginas, quase como se a carta escrevesse Plácido*. Note-se como o *escrevesse* e o *quase* remetem ao Sit<sub>0</sub>. Essa remissão passa a imagem de um enunciador imaginando um cenário em que carta troca de papel com o escritor.

A relação primitiva permite a visualização de propriedades fundamentais que atuam em um enunciado. Para Culioli (1999b), uma relação primitiva configura-se como uma hipótese do funcionamento complexo de uma relação abstrata e geral entre um termo e outro termo. O olhar para um constituinte linguístico apresenta um conjunto de combinações possíveis que se situam entre propriedades extralinguísticas e propriedades culturais. Para um dado enunciado, Culioli postula que há uma relação entre os termos e que os mesmos

9 O relator não precisa ser um verbo.

possuem uma orientação relativamente estável. A estabilidade, no entanto, não pode ser considerada algo absoluto, pois ela varia de cultura para cultura. Em uma relação primitiva encontramos um esquema de organização de uma léxis. A relação primitiva deflagra uma dada orientação dos termos, pois podemos identificar um termo de partida, a origem e o objetivo da relação.

A predicação é o arranjo de uma léxis e, conseqüentemente, de uma relação primitiva sob a forma de uma relação predicativa. Nessa operação verifica-se que dois termos serão postos em relação. Um termo será eleito como ponto de partida em relação ao outro. Por exemplo, Pedro Álvares Cabral descobriu o Brasil, O Brasil foi descoberto por Pedro Álvares Cabral, A descoberta do Brasil por Pedro Álvares Cabral. Nos três exemplos, verifica-se um termo de origem sendo localizado em relação ao resto do enunciado construído. Logo

Toda localização se resume em “identificar” e, então, extrair um objeto ou uma situação entre outras, e desse modo, construir a referência a um certo tipo em domínio determinado. Trata-se de estabelecer na relação de diferenciação baseada na alteridade: aquilo que é, será preciso sempre considerá-lo em relação ao que é outro. (REZENDE, 2000, p. 102)

A enunciação estabelece-se sobre a predicação. Esse conceito depende de um conjunto de parâmetros que definem a relação do sujeito do enunciado (S) e do sujeito enunciador (E) e, por outro lado, do momento ao qual se refere o enunciado (T) e do momento da enunciação (E). Em um dado enunciado essas referências serão validadas conforme o ponto de vista daquele que enuncia em relação ao que o mesmo supõe ser a posição de seu interlocutor. Nessa operação, ganham grande importância as modalidades, pois estas asseguram os jogos de deslocamentos enunciativos.

Há quatro tipos de modalidade na acepção de Culioli (1999a) cuja descrição será exposta conforme o texto de Rezende (2000). As chamadas *modalidades 1* são de asserção, de interrogação e de ênfase. Rezende (2000) postula que as *modalidades 1* são essenciais para o a tematização (ou localização) que se apresenta no nível predicativo. Nas *modalidades 2* verificam-se os efeitos do necessário, do possível, do eventual e as da certeza. A união desses dois níveis de modalidades resulta em enunciações que finalizarão “[...] sobre julgamentos universais (é necessário que...) ou sobre julgamentos estritamente localizados (é possível ou é provável que em certas circunstâncias...)” (REZENDE, 2000, p. 105). As *modalidades 3* são responsáveis pelos efeitos de apreciação e de afetividade que se encontram centradas no sujeito enunciador. Os componentes dessa modalidade são responsáveis por construir os efeitos de afastamento e aproximação nos jogos enunciativos.

O aspecto é o conceito responsável por revelar, em um dado enunciado, as operações de tempo e espaço. Tais operações são responsáveis por situar o enunciador em relação ao seu próprio discurso, em relação ao discurso do outro e em relação à percepção do tempo que se ancora no espaço. A união dessas características funda os jogos temporais e espaciais operados pelos sujeitos na atividade de linguagem.

## A análise de Lima e a infração sintática do texto de Machado de Assis

Lima coloca em análise a seguinte passagem:

Plácido foi á secretária, escreveu o verso, e deu-lhe o papel, mas já então Santos advertira que mostral-o á mulher era confessar a consulta spirita, e naturalmente o perjúrio. Referiu ao amigo os escrúpulos de Natividade e pediu que **calassem** tudo. (MACHADO DE ASSIS, 1962, p. 112 apud LIMA, 1973, p. 61)

Em seguida identifica a insuficiência gramatical do período:

O engano está na parte final do período: Santos “REFERIU O AMIGO (Plácido) OS ESCRUPULOS DE NATIVIDADE E (Santos) PEDIU QUE CALASSEM TUDO.” Santos pediu a quem? A Plácido, evidentemente. O sujeito está no singular, e o verbo (calassem) no plural! Discrepância flagrante, irrefutável, estranha até em colegial. (LIMA, 1973, p. 61-62)

Pela ótica tradicional nada mais correto do que colocar em concordância o sujeito e o predicado. A regularidade das regras gramaticais e a rigidez que ela possui torna esse tipo de olhar para os constituintes linguísticos incontestáveis. Não se questiona a lógica. Cada coisa em seu lugar, a língua deve ser o espelho da realidade. E assim, com essa concepção, tudo o que não presta encontra-se fora da discussão. Apenas em raras situações a não padronização gramatical é considerada, e, quando é, ela figura como uma permissão. Mas permissão de quem? Do gramático? Do professor? Nesse universo a língua figura também como uma entidade controlável e neutra. As contradições que a linguagem possui são postas em suspenso. O texto é observado a partir de um núcleo neutro em que a sintaxe invalida qualquer construção que infrinja seus parâmetros rígidos. Porém como explicar que a ocorrência em questão é compreensível? Quando produzimos e reconhecemos textos há uma atividade sustentando essa produção e ela não é simétrica, porém é recuperável pela atividade epilinguística.

O olhar da Teoria das Operações Predicativas e Enunciativas permite-nos simular o processo de compreensão do texto e de postular quais operações enunciativas sustentam a ocorrência em questão. Dessa forma, não consideramos apenas o nível sintático como parâmetro de boa formação do enunciado. Na análise de Lima (1973), os pressupostos clássicos da gramática aparecem em conjunto com a própria atividade de compreensão do período. Lima, ao identificar o suposto problema no período de Machado de Assis, passa, sem se dar conta, por uma reformulação das marcas linguísticas. Essa reformulação é explicitada em parte porque o gramático não opera com ela nem busca gerar coordenadas das possibilidades semânticas e gramaticais que o texto deflagra. Dessa forma, a análise de Lima desconsidera que a mudança do relator **calassem** para o **calasse** implica tanto a mudança da marcação da categoria de pessoa quanto a reorientação do cenário enunciativo em questão.

Usando referencial teórico criado por Culioli tentaremos mostrar que existe uma diferença significativa no sentido do enunciado do texto de Machado de Assis, que embora seja pequena, acreditamos ser extremamente importante, pois revela o âmago do trabalho literário, qual seja, o de trabalhar com singularidade das palavras para atingir a expressão estética. Essa propriedade central da linguagem revela-se também promissora para o contexto de ensino e aprendizagem de língua materna, pois incita um olhar para o texto como um espaço de apropriação e construção de significados.

A adaptação do texto conforme as orientações de Lima gera uma mudança no cenário enunciativo forjado por Machado de Assis:

**(Santos)** Referiu ao amigo os escrúpulos de Natividade e pediu **(ao amigo, Plácido)** que **calasse** tudo. (MACHADO DE ASSIS, 1962, p. 112 apud LIMA, 1973, p. 61)

Assim, podemos manipular o enunciado com algumas paráfrases. Destacamos o seguinte enunciado:

(5) Santos pediu a Plácido que **calasse** tudo.

No enunciado (5) podemos ressaltar que o relator **calar** localiza Plácido como o responsável pelo calar. Assim Santos encontra-se excluído dessa relação, pois o calar depende apenas da ação de Plácido e não de Santos.

Colocando o enunciado em discurso direto podemos, ainda, recuperar o apelo de Santos a Plácido:

(6) Santos disse: Plácido, **(eu e você) cale** tudo.

O interessante da abordagem enunciativa é que a marcação do subjuntivo deflagra dois cenários enunciativos possíveis e não apenas uma orientação de tempo verbal:

(7) Sit0 (eu-aqui- agora digo que) Santos pediu ao amigo que> (8) **ele e o amigo calassem tudo**>(9) **pois o amigo não estava calado, naquele momento**> (10) **porque havia a possibilidade do amigo não calar tudo, Santos pediu ao amigo que ambos calassem tudo** (11) **o que foi dito naquele momento.**

No caso a desinência do tempo verbal no subjuntivo incide sobre a reorientação da situação enunciativa e da relação predicativa (Plácido, calar, tudo). Assim os enunciados (5) e (6) (7), (8), (9) (10) e (11) são meta-enunciados localizáveis a partir da manipulação parafrástica do enunciado de Machado de Assis. O texto em negrito em (8) e (9), (10) e (11) são meta-enunciados que estão envolvidos no domínio nocional da noção calar. Assim o subjuntivo constitui-se por uma relação complementar p/p'. A notação p/p' é o mesmo que criar um espaço validável da noção calar/não calar. O imperfeito do subjuntivo deflagra uma noção modal e aspectual; por meio delas recuperamos que o enunciador constrói uma projeção da atitude de Plácido no tempo-espaço do enunciado em relação ao tempo-espaço de Santos. A projeção deixa marcas de um diálogo e de um apelo ao co-enunciador, pois o não calar era uma possibilidade para Santos, por isso ele pediu a Plácido para que o não calar fosse invalidado tanto no horizonte do enunciador quanto no horizonte do co-enunciador. Da aparente simplicidade do enunciado encontramos uma cascata de enunciadores (narrador e personagens) que atestam para a complexidade da situação enunciativa do texto de Machado de Assis. Assim, as elipses deixadas por Machado de Assis podem ser preenchidas pela atividade de compreensão dos leitores. O mais interessante, ainda, é que Lima utiliza esse recurso de paráfrase para poder identificar a suposta construção correta do texto de Machado de Assis. Lima, ao mudar a marcação do *calassem* para o *calasse*, transforma cenário enunciativo do texto de Machado de Assis e descaracteriza a proposta do texto em prol de uma norma sintática.

Tendo em vista os questionamentos levantados para com a postura tradicional, inserimos uma redação de vestibular da Vunesp de nosso *corpus* de mestrado para demonstrar que essas ocorrências estão presentes nos textos dos alunos também. Porque essas ocorrências aparecem em redações acreditamos ser pertinente uma nova contextualização das noções de erro e criatividade quando consideramos ocorrências de deslocamentos enunciativos.

**O homem** é animal inteligente que já passou por outras dificuldades, portanto, ele vai passar por essa crise momentânea e não vai destruir em que **vivemos**. (redação da Vunesp)

Na passagem acima encontramos o mesmo problema sintático identificado por Lima (1975) no texto de Machado de Assis. A comparação de ambos os textos poderia induzir-nos a sustentar que o texto do escritor literário e o texto do aluno se diferenciam pela presença de um projeto artístico no primeiro e sua ausência no segundo. Embora reconheçamos a plausibilidade da distinção entre o artístico e o não artístico, essa maneira de ver as ocorrências pode instaurar outro tipo de categorização que, em vez de se pautar em uma análise linguística ancorada em pressupostos teóricos, pode conduzir a discussão para um terreno não linguístico. Terreno em que surge uma disparidade, pois a figura histórica de Machado tem um peso maior que a do aluno que produziu a redação de vestibular. Nesse universo, as etiquetas estão colocadas de antemão. Isso pode desfigurar o foco do ensino de produção e interpretação de textos conforme as orientações de uma linguística preocupada com os processos de construção de significação. Na orientação enunciativa de Culioli (1999a), que leva em conta o processo, há um interesse flagrante não pelos resultados, mas sim pelos caminhos que as marcas léxico-gramaticais oferecem para o trabalho linguístico-cognitivo dos alunos.

O foco em operações de linguagem incita-nos a olhar para a ocorrência em questão não apenas como uma imprecisão, mas como uma ocorrência que oferece diferentes potenciais de organização.

Se tomarmos enunciado readaptado do texto do aluno:

(12) **O homem** não vai destruir o planeta em que **vivemos**.

Podemos identificar que existem caminhos que são abertos pela concorrência de planos enunciativos deflagrados pela não concordância do sujeito com o predicado.

(12a) **O homem** não vai destruir o planeta em que **vive**.

A ocorrência parafraseada em (12a) não coloca em jogo o mesmo cenário enunciativo que em (12). Em (12), mesmo não havendo um padrão sintático, percebe-se que o sujeito enunciativo não quer incluir o **nós (eu+eles)** nas ações do homem:

(13) O homem vive no planeta> nos vivemos no planeta> mas existem alguns homens> diferentes de nós> que querem destruir planeta em que vivemos> dessa forma o planeta em que vivemos não é o mesmo planeta que alguns homens vivem> Aqueles homens> que vivem em um planeta diferente> não vão destruir o planeta em que nós vivemos.

O exemplo (13) força a compreensão do enunciado (12), porém a concorrência de planos enunciativos nos leva a postular que (12a) e (13) são dois caminhos possíveis que podem auxiliar no refinamento da produção de sentidos no contexto de ensino e aprendizagem de

língua materna. Se se seguisse a elaboração do texto (13) poder-se-ia mostrar que o valor de planeta para o sujeito enunciativo não é similar ao valor de planeta para o homem, porém o homem deve ser quantificado e qualificado<sup>10</sup> de modo restrito como, por exemplo:

- (14) **há alguns homens** ou **Aqueles homens**> que vivem em **um mundo** diferente do nosso> pois só querem destruir o planeta em que nós vivemos.

Quando mantemos a marca **o homem** em (12), a visibilidade da manipulação parafrástica do texto (13) torna-se menos visível, mas a concorrência de planos enunciativos e a especificação de **homens versus nós** traz o embrião de um enunciado bem formado enunciativamente.

## Conclusão

Partimos de orientações teóricas que prezam pela articulação entre enunciado e frase e, conseqüentemente, entre língua e linguagem. Inserimos também um olhar para a noção de criatividade a partir do referencial da Teoria das Operações Predicativas e Enunciativas. A partir dessa reflexão mostramos que o olhar para duas ocorrências de deslocamento enunciativo da categoria de pessoa possuem elementos em comum que não devem ser considerados apenas um “defeito” ou um “equivoco”, pois, na verdade, essas ocorrências precisam ser explicitadas para a mobilização de um olhar complexo para os constituintes linguísticos. É tão importante saber quais caminhos tomamos para construir representações linguísticas boas como é importante saber quais caminhos evitar. A exclusão da dimensão do erro invalida o trabalho com a linguagem e, ainda, pode cercear a capacidade dos seres humanos de construir novas representações linguísticas. Compreendemos que Lima (1973) sustenta seus argumentos a partir de uma base histórica, assim o valor sintático orienta-se a partir de uma tradição. Mas seria realmente válido julgar os méritos linguísticos de Machado de Assis, majoritariamente, a partir uma tradição? Pensamos que Lima (1973), ao polarizar o problema, desmerece a capacidade linguística do autor literário. Seguindo esse olhar radical que se mostra tão incisivo para com escritores consagrados, imaginamos o quão nocivo pode ser esse tipo de concepção quando caso seja transposta para o contexto de ensino e aprendizagem.

## REFERÊNCIAS

CULIOLI, A. *Pour une linguistique de l'énonciation: formalisation et opérations de repérage*. Paris: Ophrys, 1999a. v. 2.

\_\_\_\_\_. *Pour une linguistique de l'énonciation: domaine notionnel*. Paris: Ophrys, 1999b. v. 3.

FRANCHI, C. Criatividade e gramática. In: FRANCHI, C.; NEGRÃO, E.; MÜLLER, A. L. *Mas o que é mesmo gramática?* Organização de Sírio Possenti. São Paulo: Parábola Editorial, 2006. p. 11-33.

<sup>10</sup> Reenviamos o leitor para os conceitos de quantifiabilização presentes nos mecanismos de discreto, denso e compacto da Teoria das Operações Predicativas e Enunciativas.

LIMA, C. *Revisão de Machado de Assis: exames de erros e ardis literários*. Rio de Janeiro: Americana, 1973.

MILNER, J-C. De quelques aspects de la théorie d'Antoine Culioli projetés dans un espace non-énonciatif. In: *La théorie d'Antoine Culioli: ouvertures et incidences*. Paris: Editions OPHRYS, 1992.

REZENDE, L. M. Contribuições da teoria das operações predicativas e enunciativas para o ensino de línguas. In: JORENTE, J.; REZENDE, L. M.; ONOFRE, M. B. (Org.). *Versão Beta: Jornada de estudos enunciativos*. São Carlos: UFSCar, 2010.

\_\_\_\_\_. *Léxico e gramática: aproximação de problemas lingüísticos com educacionais*. v.1. 2000. Tese (Livre Docência ) – Faculdade de Ciências e Letras, Unesp, Araraquara.

# **Crenças e atitudes linguísticas no Sudoeste do Paraná: tendências de reação frente às diferentes línguas e etnias**

(Beliefs and attitudes in the Southwest of Paraná: tendencies of reaction towards different languages and ethnicities)

**Clarice Cristina Corbari<sup>1</sup>, Aparecida Feola Sella<sup>2</sup>**

<sup>1,2</sup> Universidade Estadual do Oeste do Paraná (UNIOESTE)

ccorbari@yahoo.com.br, afsella1@yahoo.com.br

**Abstract:** This article presents partial results of a research on linguistic beliefs and attitudes of speakers from three localities in the Southwest of Paraná, on the border of Argentina: Santo Antônio do Sudoeste, Pranchita and Capanema. This region experienced legal, political and social disputes engendered by the process of land occupation, whose implications may have been reflected on the relations established among the various ethnic groups that contacted each other in that territory, fostering both positive and negative manifestations towards the local speeches. The study focuses on the tendencies of reaction towards neighborly, affective and professional relationships with members of different ethnic groups that belong to the population of the localities. The discussions are guided by theoretical and methodological principles of Sociolinguistics, Sociology of Language and Social Psychology related to the topic.

**Keywords:** linguistic beliefs and attitudes; languages in contact; border area.

**Resumo:** Este artigo apresenta resultados de pesquisa sobre crenças e atitudes linguísticas de falantes de três localidades do Sudoeste do Paraná, na fronteira com a Argentina: Santo Antônio do Sudoeste, Pranchita e Capanema. Essa região foi palco de conflitos jurídicos, políticos e sociais engendrados pelo processo de ocupação das terras, cujas implicações podem ter se refletido nas relações estabelecidas entre os diversos grupos étnicos que entraram em contato nesse território, propiciando manifestações tanto positivas quanto negativas frente aos falares locais. O estudo focaliza as tendências de reação frente às relações de vizinhança, afetivas e profissionais com membros das diversas etnias que compõem a população das localidades. Norteiam as discussões os princípios teórico-metodológicos da Sociolinguística, da Sociologia da Linguagem e da Psicologia Social referentes ao tema.

**Palavras-chave:** crenças e atitudes linguísticas; línguas em contato; contexto de fronteira.

## **Introdução**

Os contatos estabelecidos em contextos de fronteira fazem desse espaço um complexo cenário sociolinguístico e cultural, propiciando sobremaneira o estudo das línguas em contato e das crenças e atitudes relacionadas a essas línguas e seus usuários. De acordo com Sturza,

As fronteiras geográficas são preenchidas de conteúdo social. Se as fronteiras são sociais, se nelas vivem diferentes etnias [...] o contato linguístico é uma consequência inevitável, e a situação das práticas linguísticas nessas regiões, de um modo geral, um campo pouco explorado pela linguística brasileira. (2005, p. 47)

Existe uma estreita relação entre língua e identidade, de tal forma que, conforme Moreno Fernández (1998), uma variedade linguística pode ser interpretada como um tra-

ção definidor da identidade. Assim, segundo o autor, as atitudes em relação a grupos com certa identidade são, em parte, atitudes em relação às variedades linguísticas usadas por esses grupos e em relação aos próprios usuários. Entende-se, portanto, que o estudo das crenças e atitudes linguísticas presentes em localidades multilíngues seja pertinente pela possibilidade de fornecer indícios para a análise do comportamento linguístico dos falantes em relação à variação, revelando os elementos que atuam nas relações sociais entre os diferentes grupos.

No caso do Sudoeste do Paraná, ainda são poucos e recentes os estudos sobre as crenças e atitudes linguísticas (SILVA-PORELI, 2010; PASTORELLI, 2011), fazendo valer para esse contexto específico a reflexão de Sturza sobre as fronteiras do Brasil:

Quase dois séculos depois de conflitos, solucionados pelas armas ou pela diplomacia, ainda desconhecemos muito da situação de contato das línguas portuguesa e espanhola nas zonas fronteiriças do Brasil com os demais países hispano-americanos. [...] mesmo onde os agrupamentos são menores e menos populosos, a fronteira efetivamente é complexa pela natureza de sua formação e pelo modo como se estabelecem ali as relações sociais das diferentes etnias que nela habitam. (STURZA, 2005, p. 47)

Primeiramente habitada por caboclos, que ocuparam a terra na condição de posseiros, a região Sudoeste do Paraná recebeu posteriormente argentinos e paraguaios, que chegaram a compor 25% da população da fronteira no auge da exploração da erva-mate. Com a diminuição dessa atividade extrativa na região, os argentinos e paraguaios começaram a se evadir, chegando a menos de 1% da população regional nos anos 1940. Nessa década e na seguinte, levas de descendentes de imigrantes, predominantemente de origens alemã e italiana, instalaram-se na região, vindos de colônias previamente formadas no Rio Grande do Sul e em Santa Catarina, atraídos pela possibilidade de compra de terras a preços irrisórios (WACHOWICZ, 1985).

Por ser uma região fértil e rica, o Sudoeste do Paraná foi disputado tanto pela Argentina e pelo Brasil, quanto pelos Estados do Paraná e Santa Catarina. Segundo Lazier (2003, p. 146), “[...] essa desavença pela posse das terras envolveu também a Cia. de Estradas de Ferro São Paulo-Rio Grande, a CITLA, o Governo Federal, o Governo do Paraná e, principalmente, posseiros”.

O conflito entre os dois Estados pela posse dessa região se originou ainda no Brasil colônia e continuou com a criação da Província do Paraná, em 1853. Após a Guerra do Contestado, os dois Estados assinaram o acordo de fronteira, ficando para Santa Catarina a maior parte das terras em litígio. Foi só a partir de 20 de outubro de 1916 que a região passou a pertencer ao Estado do Paraná (LAZIER, 2003).

Na questão da disputa entre os dois países, Lazier (2003) informa que, ao ser definida a linha divisória entre eles, a Argentina reivindicou a região onde hoje é o Sudoeste do Paraná; o Brasil, por sua vez, defendia a conformação que a fronteira tem atualmente. A disputa só foi decidida em 1889, quando Cleveland, o então presidente dos Estados Unidos, escolhido como juiz da questão, deu ganho de causa ao Brasil. Porém, a divisa só veio a ser definitivamente estabelecida em 1898, por meio de um tratado. Tendo em vista essa situação, é possível que o conflito entre os dois países tenha gerado certa animosidade entre brasileiros e argentinos.

Olhar para a sócio-história pode fornecer pistas para a compreensão da natureza das relações linguístico-culturais estabelecidas em determinado espaço geossocial. No Sudoeste do Paraná, o modo de ocupação das terras, ao engendrar diversos conflitos jurídicos, políticos e sociais, pode ter trazido implicações que se refletiram nas relações instauradas entre os diversos grupos étnicos que entraram em contato nesse território. Este artigo constitui uma tentativa de verificação do contexto cultural e linguístico de uma região que, além de histórico recente, constituiu-se na soma de culturas. Por isso, apresenta alguns resultados de pesquisa descritiva sobre crenças e atitudes linguísticas de falantes de três localidades situadas nessa região, na fronteira com a Argentina: Santo Antônio do Sudoeste, Pranchita e Capanema. O estudo se concentra nas tendências de reação dos informantes frente às relações de vizinhança, afetivas e profissionais com membros das diversas etnias que compõem a população das localidades.

No contexto linguístico e cultural atual dessas localidades, além da diversidade étnica resultante da colonização e da mistura entre hispânicos e brasileiros (incluindo aqui os caboclos), há ainda o contato entre brasileiros e argentinos que atravessam continuamente a fronteira entre Brasil e Argentina, especialmente em Santo Antônio do Sudoeste, gerando uma permanente interação linguístico-cultural. Todas essas características tornam as localidades um espaço cultural e linguisticamente complexo, favorecendo o estudo das culturas e línguas em contato.

A fronteira constitui algo mais que o mero fato geográfico: os sujeitos envolvidos nesse espaço fazem dela também um fato social. Por essa razão, esse cenário de línguas e culturas em contato se torna, inevitavelmente, campo propício para a manifestação de crenças e atitudes, sejam elas positivas ou negativas, em relação à língua e à cultura do “outro”, que, em última instância, são crenças e atitudes em relação aos próprios sujeitos que se identificam com tal língua e cultura.

### **Crenças e atitudes linguísticas: pressupostos teóricos**

Diversas áreas do conhecimento, em especial a Psicologia Social, a Sociolinguística e a Sociologia da Linguagem, contribuem para o estudo das crenças e atitudes linguísticas. A tarefa da Psicologia Social é fornecer subsídios para o estudo dos papéis que os motivos, as crenças e a identidade exercem no comportamento linguístico do indivíduo (LAMBERT; LAMBERT, 1966). A Sociolinguística, por sua vez, colabora com a tarefa de pesquisar a diferença entre a maneira como as pessoas fazem uso da(s) língua(s), bem como suas crenças a respeito de seu próprio comportamento linguístico e o dos demais falantes (MORENO FERNÁNDEZ, 1998; GÓMEZ MOLINA, 1996; BLANCO CANALES, 2004). Já a Sociologia da Linguagem focaliza toda a gama de tópicos relacionados à organização social do comportamento linguístico, incluindo não apenas o uso da língua em si, mas também as atitudes explícitas em relação à língua e aos seus usuários (FISHMAN, 1972).

As observações e interpretações do comportamento social, cultural ou linguístico das pessoas com quem entramos em contato diariamente resultam em atribuições que formam a base de nossas atitudes, as quais, por sua vez, influenciam nosso comportamento em relação aos membros de nosso ou de outro grupo social. (PADILLA, 1999). No que se refere ao comportamento linguístico:

[...] existe na sociedade o que poderíamos chamar de olhares sobre a língua, de imagens da língua, em uma palavra, *normas* que podem ser partilhadas por todos ou diferenciadas segundo certas variáveis sociais [...] e que geram sentimentos, atitudes, comportamentos diferenciados. (CALVET, 2002, p. 72, grifo do autor).

O conceito de atitude, primeiramente elaborado pelos psicólogos sociais, também pode ser aplicado às atitudes linguísticas. Lambert e Lambert (1966, p. 77) definem a atitude como “uma maneira organizada e coerente de pensar, sentir e reagir em relação a pessoas, grupos, questões sociais ou, mais genericamente, a qualquer acontecimento ocorrido em nosso meio circundante”. Para esses psicólogos, a atitude é formada por três componentes: pensamentos e crenças, sentimentos ou emoções, e tendências de reação. Porém, a definição da estrutura componencial da atitude está longe de consenso. Bem (1973), por exemplo, acrescenta o componente social: para o autor, as crenças e atitudes humanas se fundamentam em quatro atividades do homem – pensar, sentir, comportar-se e interagir com os outros –, que correspondem aos quatro fundamentos psicológicos das crenças e atitudes – cognitivos, emocionais, comportamentais e sociais. Já López Morales (1993) identifica na atitude apenas o componente conativo, separando os conceitos de crença e atitude e situando-os em níveis diferentes.

As abordagens das atitudes refletem o conceito que se tem de atitude. Duas perspectivas sobressaem: (a) a mentalista, que concebe a atitude como uma entidade complexa, compreendendo os elementos cognitivo ou cognoscitivo, afetivo e conativo – embora, como lembra Moreno Fernández (1998), existam discrepâncias no interior dessa abordagem para determinar como se relacionam entre si esses elementos –, e (b) a behaviorista ou comportamentalista, em que a atitude é composta de um elemento único, geralmente afetivo ou de valoração. Neste estudo, utiliza-se a abordagem mentalista, com foco no componente conativo.

Conforme descrevem alguns autores (GÓMEZ MOLINA, 1996; MORENO FERNÁNDEZ, 1998; BLANCO CANALES, 2004), para a perspectiva mentalista, de natureza psicológica, a atitude é uma disposição mental em relação a condições ou fatos sociolinguísticos concretos, razão pela qual não é possível medi-la ou observá-la diretamente, mas apenas deduzi-la a partir de certa informação psicossociológica, sendo necessário recorrer a técnicas indiretas. Já a concepção comportamentalista ou behaviorista interpreta a atitude como uma conduta, uma reação ou resposta a um estímulo – uma variedade linguística, por exemplo –, de modo que pode ser observada diretamente a partir do comportamento do indivíduo dentro de certas situações sociais.

No âmbito das atitudes de forma geral, as atitudes linguísticas constituem uma categoria particular, uma vez que o objeto da atitude não são as línguas, mas os grupos que as falam. As atitudes linguísticas representam, assim, um componente fundamental da identidade linguística do falante e possibilitam a leitura e compreensão do próprio comportamento linguístico. Nessa perspectiva, o estudo das crenças e atitudes linguísticas precisa estar fundamentado na relação entre língua e identidade étnica, pois, segundo Liebkind (1999), usar a língua influencia a formação da identidade de grupo, que, por sua vez, influencia os padrões de atitude e uso linguísticos.

Para Moreno Fernandez (1998), as atitudes linguísticas têm a ver com as línguas mesmas e com a identidade dos grupos que as manejam. Existe, portanto, uma estreita relação entre língua e identidade, que, não raro, manifesta-se nas atitudes dos indivíduos

em relação a essas línguas e a seus usuários. Uma variedade linguística pode ser interpretada como um traço definidor da identidade, daí que as atitudes em relação aos grupos com certa identidade sejam em parte atitudes em relação às variedades linguísticas usadas nesses grupos e em relação aos usuários de tais variedades (MORENO FERNÁNDEZ, 1998; GROSJEAN, 1982). Nesse sentido, as “atitudes linguísticas são as armas usadas pelos residentes para demarcar seu espaço, sua identidade cultural, seu perfil de comunidade, de grupo social separado” (TARALLO, 1985, p. 14). Aguilera (2008) alinha-se a esse pensamento ao afirmar que a língua não está desvinculada de seu contexto social, principalmente na sua condição de aspecto constituidor da identidade de um determinado grupo étnico. Decorre daí que, “na maioria das vezes, ao caracterizar um grupo ao qual não pertence, a tendência é o usuário fazê-lo de forma subjetiva, procurando preservar o sentimento de comunidade partilhado e classificando o outro como diferente” (AGUILERA, 2008, p. 106). É a língua que simboliza os limites que separam o “nós” e os “outros”, uma vez que a língua que falamos identifica a nossa origem, nossa história, nossa cultura, o grupo a que pertencemos.

## Metodologia da pesquisa

Este estudo parte de dados coletados pelo Projeto *Crenças e atitudes linguísticas: um estudo da relação do português com línguas de contato* (AGUILERA, 2009) em regiões de fronteira e imigração, no Paraná. O projeto envolveu oito municípios, sendo seis fronteiriços ao Paraguai e à Argentina, e dois situados na região central do Estado. Dentre esses municípios, esta pesquisa focaliza três, conforme já anunciado: Santo Antônio do Sudoeste, Pranchita e Capanema.

O aporte metodológico das pesquisas sobre crenças e atitudes linguísticas advém principalmente da Psicologia Social. Para o projeto em questão, adotou-se uma metodologia baseada na abordagem mentalista, na perspectiva de que, conforme Blanco Canales (2004), apesar das evidentes desvantagens dessa proposta, que demanda um mecanismo que permita inferir e medir as atitudes, é a mais bem aceita devido à sua capacidade de prever o comportamento verbal e, portanto, converter-se em modelos sistemáticos.

O instrumento de coleta de dados seguiu a orientação de Lambert e Lambert (1966), que propõem medir as atitudes por meio de um questionário contendo itens que representem os três componentes da atitude: o cognitivo, o afetivo e o conativo. Desse modo, elaborou-se um questionário para as entrevistas com base em tais critérios, adaptados à realidade sociolinguística e cultural das comunidades de fala investigadas, com 48 perguntas específicas para avaliar crenças e atitudes linguísticas em relação às línguas em contato e ao português de cada localidade.

A seleção dos informantes foi planejada com base em três dimensões: (a) a diageracional, contemplando três faixas etárias: 18 a 30 anos, 31 a 50 anos, e 51 a 70 anos; (b) a diastrática, optando-se pela escolaridade como parâmetro definidor de classe social, resultando na definição de três níveis de escolaridade: fundamental, médio e superior; e (c) a diassexual, contemplando sujeitos dos gêneros/sexos feminino e masculino. Da combinação das variáveis resultou a seleção de dezoito informantes para cada localidade pesquisada.

Para o desenvolvimento da análise aqui empreendida, foram selecionadas as perguntas do questionário cujas respostas pudessem revelar, principalmente, tendências de reação dos informantes diante da possibilidade de contato (profissional, afetivo e de vizinhança) com membros das diversas etnias que convivem na localidade. Trata-se dos seguintes grupos de questões: (a) “Se você fosse comprar uma casa num bairro onde só houvesse argentinos / paraguaios / alemães / italianos, você compraria?”, (b) “Você namoraria ou se casaria com um(a) argentino(a) / paraguaio(a) / alemão(ã) / italiano (a)? Por quê?” e (c) “Se precisasse de um médico ou dentista, procuraria um argentino / paraguaio / alemão / italiano? Por quê?”.<sup>1</sup>

O uso das designações ‘alemão’ e ‘italiano’ para a denominação dos falantes não se refere propriamente à nacionalidade desses falantes, mas à sua origem étnica, ou seja, trata-se de filhos ou netos de alemães e italianos. A opção por essas designações se justifica pelo fato de os próprios descendentes assim se autodenominarem como forma de se distinguirem dos “brasileiros”, isto é, daqueles nascidos no Brasil e sem descendência europeia.

Os dados obtidos nas respostas foram, primeiramente, contabilizados e representados em termos percentuais.<sup>2</sup> Os dados numéricos estão dispostos em tabelas, para melhor visualização, seguindo três categorias: respostas positivas, respostas negativas e outras respostas – nessa última categoria, foram consideradas em bloco as não respostas (oriundas de perguntas não formuladas) e os raros casos de dúvida ou desvio de foco por parte do informante. No entanto, para a discussão dos resultados, são consideradas apenas as respostas positivas e negativas, indicativas dos índices de aceitação e rejeição, mesmo que, em alguns casos, as perguntas não formuladas representem uma parcela significativa do *corpus*, especialmente nas perguntas sobre os paraguaios. Vale informar que, na análise, a expressão “índice de aceitação” sempre se referirá às tendências de reação positiva (representadas pelas respostas positivas), e a expressão “índice de rejeição”, às tendências de reação negativa (representadas pelas respostas negativas).

Por fim, é importante esclarecer que as tendências de reação frente a situações hipotéticas não garantem a efetivação das atitudes reativas correspondentes; apenas representa um indicativo de como o indivíduo poderia se comportar frente às situações, caso se efetivassem.

## **Análise dos dados**

Com relação ao primeiro grupo de questões analisado – “Se você fosse comprar uma casa num bairro onde só houvesse argentinos / paraguaios / alemães / italianos, você compraria?” –, os resultados expressos na tabela mostram que, em todas as situações, as respostas positivas superaram as negativas<sup>3</sup>.

1 Em Pranchita, as questões incluíam os poloneses, mas, para esta análise, foram selecionadas apenas as perguntas referentes aos grupos étnicos comuns aos questionários de todas as localidades.

2 Os percentuais foram arredondados para compor as tabelas, o que pode resultar em uma diferença de 1%, em algumas células, para porcentagens correspondentes ao mesmo número de informantes, de forma a garantir o total de 100%. Por exemplo, o percentual de 5,55% (correspondente a um informante) é normalmente representado como 6% na tabela; eventualmente, porém, esse número pode figurar como 5% para que a soma dos percentuais para determinada questão em dada localidade não ultrapasse os 100%.

3 As porcentagens informadas estão sempre em relação com o universo dos informantes (em que 100% corresponde a dezoito informantes), e não em relação com o total de perguntas formuladas e/ou respostas obtidas.

**Tabela 1: Reação dos informantes frente à possibilidade de comprar uma casa num bairro onde só houvesse membros de determinada etnia**

<i>Grupo étnico / Respostas</i>	<i>Localidade</i>	<b>Sto. Antônio do Sudoeste</b>	<b>Pranchita</b>	<b>Capanema</b>
	Comprariam	100%	78%	83%
<b>Argentinos</b>	Não comprariam	-	11%	6%
	Outras respostas / NF*	-	11%	11%
	Comprariam	50%	39%	44%
<b>Paraguaios</b>	Não comprariam	11%	33%	17%
	Outras respostas / NF	39%	28%	39%
	Comprariam	67%	61%	83%
<b>Alemães</b>	Não comprariam	28%	17%	6%
	Outras respostas / NF	5%	22%	11%
	Comprariam	100%	72%	89%
<b>Italianos</b>	Não comprariam	-	22%	-
	Outras respostas / NF	-	6%	11%

\* NF: Questões não formuladas.

No que diz respeito à possibilidade de comprar uma casa em um bairro onde morassem apenas argentinos e descendentes, observa-se alto índice de respostas positivas nas três localidades, especialmente em Santo Antônio do Sudoeste, onde todos os informantes se mostraram dispostos a morar num bairro onde houvesse apenas membros dessa etnia. Uma possível explicação para essa unanimidade é o fato de a travessia de um país para o outro ocorrer nessa localidade, por meio da ponte que a liga à cidade vizinha de San Antonio, e, portanto, concentrar-se aí a maior ocorrência de interações entre brasileiros e argentinos, o que pode colaborar para uma atitude de maior aceitação em relação aos argentinos. As justificativas para as respostas positivas, nas três localidades, giraram em torno da facilidade de trato social com os argentinos: “[...] eles são muito educados” (Inf. 17 – Pranchita); “[...] são gente, tudo gente boa” (Inf. 5 – Capanema).

As rejeições foram consideradas baixíssimas: apenas dois informantes de Pranchita e um de Capanema não se mostraram dispostos a morar num bairro onde houvesse apenas argentinos. As justificativas apresentadas foram as seguintes:

- (01) Que nem, **a gente é costumado aqui**, com as pessoas daqui, com os mesmo, né? **A mesma idioma, conversa as mesmas coisas**, então eu acho que... (Inf. 4 – Pranchita)
- (02) Por causo que **não é o nosso país**, sabe. [...] mas me parece assim que eles têm... sempre têm a suspeita, porque é brasileiro tem... pode ser ladrão, pode ser bandido, pode ser assaltante, pode ser isso. É bem por isso então, o espanhol, ele tem, assim no meu modo de pensar, esses argentino, **eles são muito desconfiado**. (Inf. 10 – Pranchita)
- (03) Eu me sentiria... **sem saber nem falar com eles, me sentiria estranho**, né, sozinho lá no meio do argentino. (Inf. 1 – Capanema)

Com relação aos paraguaios, é preciso informar que as perguntas constavam originalmente dos questionários de todas as localidades, mas, a partir das informações de muitos informantes sobre a raridade das interações com membros dessa etnia, os inquiridores, muitas vezes, deixavam de formular as perguntas relacionadas aos paraguaios.

Mesmo assim, as respostas obtidas podem dar indícios importantes sobre as crenças e atitudes com relação aos paraguaios.

Embora as respostas positivas sempre superassem as negativas, verifica-se que a diferença entre elas é menor que a verificada para os argentinos. Os informantes que justificaram a opção por não morar em bairro de paraguaios apresentaram razões ligadas a diferenças de língua e cultura, que podem, na visão dos informantes, dificultar o convívio.

- (04) Isso é **muito difícil**, né, **porque o povo fala muito guarani**, né, e a coisa fica muito complicado, acho que daí eu não... (Inf. 11 – Santo Antônio do Sudoeste)
- (05) Paraguaio não. Acho **um povo muito... diferente do nosso**. (Inf. 12 – Santo Antônio do Sudoeste)
- (06) Eles são uma descendência de gente que **a gente não pode se enfiar neles**... Paraguai... o guarani, ele... não pode se enfiar neles, eu não (inint.) Eu nem no Paraguai não vou passear. Conheço, tive lá umas duas, três vezes, mas não vou porque **eu sei o que eles fizeram, o que eles são**. (Inf. 5 – Pranchita)
- (07) Tipo, agora, **o sistema deles, assim, parece ser um dos mais ruinzinhos assim de compreender** e de você... parece que irrita aquela... não sei que língua é mesmo que eles falam, é **o guarani**, não sei se é o modo de cada pessoa falar, mas o jeito que muitos, principalmente os que eu vi, **é muito enjoado**, assim, parece que não é uma coisa que... eu acho que desde que você não entende, eles já param e ficam falando só coisas, né? (Inf. 7 – Pranchita)
- (08) Eu acho que eu teria um pouco mais de **preconceito**. (Inf. 14 – Pranchita)
- (09) Tudo bem, ó, eu não conheço os paraguaios, eu fui duas vezes lá no Paraguai, em Foz do Iguaçu, eu estive ali, **não conheço pra dentro, quem sabe seja muito bom, mas ali na Foz do Iguaçu, meu amigo, aí é complicado**, não sei se vocês conhecem... [...] De repente, eu tô entrando numa pegadinha, mas acho que **o Paraguai precisa aprender muita coisa** [...]. Ah, mas **tem esgoto correndo lá, cano que sai assim dos prédios e sai assim na rua**... [...] Ah, eu achei **horrível**... (Inf. 17 – Pranchita)
- (10) Eles já **são diferente**, né, **uma cultura diferente**, né. Pelo que a gente conviveu, né, que a gente aprendeu com eles, assim, conheceu deles, é um povo, **uma cultura bem diferente**. (Inf. 3 – Capanema)
- (11) Eu não sei, é... eu acho que já existe uma... uma ri... não é uma **rivalidade**, isso já vem, é... por causa dessa... desse confronto **que existia antigamente entre Brasil e Paraguai**, eu acredito que eles já... como até nós tava comentando ontem, por causa dessa guerra, eu acredito que eles devam ensinar nas escolas que o país que mais prejudicou o Paraguai deve ter sido o Brasil, né. (Inf. 9 – Capanema)
- (12) **Eu ficava um pouco desconfiado do paraguaio**. Ficava um pouco desconfiado devido à nossa fronteira [...], **talvez se fosse mais no centro, capital, mais cultura**, né, mas aqui na fronteira, eu não compraria. (Inf. 11 – Capanema)

Índices mais baixos de reação positiva (em relação aos dados sobre os argentinos) também foram verificados com relação aos alemães. Além da barreira da língua, os motivos apresentados têm também relação com uma suposta atitude discriminatória atribuída aos alemães, como mostram as respostas a seguir:

- (13) Mais difícil, porque **nós somos discriminados**, sabe? **Isso vem do povo**, sei lá, como **discriminação**, sabe? Tem de cor, raça, sabe... aqui mesmo, a gente tem pessoas que são alemãs, que discriminam certas pessoas, tá. (Inf. 10 – Santo Antônio do Sudoeste)
- (14) **Eles são racistas**. (Inf. 12 – Santo Antônio do Sudoeste)
- (15) **Por causa da descendência deles**. E cê tem outra... outra... vamo supor, outra nação de gente, **eles não se acertam, tem que ser o alemão com alemão**. Então não... (Inf. 5 – Pranchita)

Já com relação aos italianos, os resultados foram similares aos verificados com relação aos argentinos, com alto índice de aceitação. Dentre os que rejeitaram a possibilidade de morar em bairro onde vivessem apenas italianos, está o informante abaixo, a quem desagrada o uso do turpilóquio,<sup>4</sup> que já é considerado expressão étnica e elemento cultural ítalo-brasileiro (cf. FROSI; FAGGION; DAL CORNO, 2008):

- (16) Por causa da idioma deles, que **tem muitas palavras que não me agrada**, quem nem esse... Então, antes de dizer ‘*porco Dio*’, então diga ‘*porco*’, eles dizem ‘*porco Dio*’, que é Deus... chama Deus de porco, e eu já não... então **eu não me adapto**, eu não adapto muito com eles por causa disso. (Inf. 5 – Pranchita)

Os resultados para o grupo de questões “Você namoraria ou se casaria com um(a) argentino(a)/paraguaio(a)/alemão(ã)/italiano (a)? Por quê?” estão apresentados na tabela a seguir.

**Tabela 2: Reação dos informantes frente à possibilidade de relacionamento afetivo (namoro ou casamento) com membros de determinada etnia**

<i>Grupo étnico / Respostas</i>		<i>Localidade</i>	<b>Sto. Antônio do Sudoeste</b>	<b>Pranchita</b>	<b>Capanema</b>
<b>Argentinos</b>	Namorariam / Casariam		94%	44%	72%
	Não namorariam ou casariam		6%	39%	11%
	Outras respostas / NF		–	17%	17%
<b>Paraguaios</b>	Namorariam / Casariam		33%	33%	56%
	Não namorariam ou casariam		17%	33%	28%
	Outras respostas / NF		50%	34%	16%
<b>Alemães</b>	Namorariam / Casariam		61%	39%	83%
	Não namorariam ou casariam		11%	39%	–
	Outras respostas / NF		28%	22%	17%
<b>Italianos</b>	Namorariam / Casariam		89%	61%	94%
	Não namorariam ou casariam		–	11%	–
	Outras respostas / NF		11%	22%	6%

Com relação à possibilidade de relacionamento afetivo com argentinos, observa-se em Santo Antônio do Sudoeste um alto índice de aceitação: quase a totalidade dos informantes se mostrou disposta a namorar ou se casar com alguém dessa nacionalidade

4 O turpilóquio diz respeito ao emprego de termos torpes (blasfêmias, imprecções, etc.), especialmente fazendo referência a nomes sagrados, que caracteriza a comunidade italiana e seus descendentes.

(alguns até o fizeram de forma efusiva). Já em Pranchita, quase houve empate (a diferença é de um informante a mais para a resposta positiva). Em Capanema, o resultado para respostas positivas foi bastante superior ao de respostas negativas.

As justificativas para as respostas positivas geralmente apontavam para a beleza do(a) argentino(a) e, principalmente, para a simpatia em relação aos argentinos, como mostram estes exemplos:

- (17) [...] **acho elas bonita, eu acho elas bonita**. O modo delas falar, **acho bonito o modo delas falar**. (Inf. 5 – Santo Antônio do Sudoeste)
- (18) Porque eu sinto aquela **simpatia** com a Argentina. (Inf. 17 – Santo Antônio do Sudoeste)

Já as justificativas para as respostas negativas se relacionam principalmente à questão das diferenças de língua e cultura:

- (19) É, **por causa do idioma** também, tudo... e **a fama deles namorador** (risos). (Inf. 4 – Pranchita)
- (20) Porque... ah, sei lá, **o tipo deles de viver é muito diferente do nosso**. (Inf. 8 – Pranchita)
- (21) **Eles são diferentes de nós**. (Inf. 4 – Capanema)

Quanto aos paraguaios, a pergunta não foi formulada à metade dos informantes de Santo Antônio do Sudoeste; entretanto, mesmo considerando a baixa porcentagem de respostas positivas em relação ao universo dos informantes, elas representam o dobro em relação às negativas. Em Pranchita, verificou-se um empate. Em Capanema, apesar de, numericamente, o resultado parecer ser mais representativo (56% e 28%) em comparação ao resultado de Santo Antônio do Sudoeste, a relação entre respostas positivas e negativas se mostrou idêntica à verificada nessa localidade, ou seja, o dobro de respostas positivas em relação às negativas.

Algumas respostas dos capanemenses mostram que o principal motivo para não namorar ou se casar com paraguaios está ligado às diferenças culturais:

- (22) Sei lá, um pouco mais diferente, **uma cultura diferente, né, da gente**, criado... **os costumes são diferentes** também, né. (Inf. 3 – Capanema)
- (23) Não, não. Eu queria até porque... a gente é, **pela origem deles mesmo, né, eles não tem muita, muito... esse asseio**, coisa assim, né, então... (risos). Não querendo ser assim, discriminar, né. Mas não. (Inf. 9 – Capanema)
- (24) **Acho que por costume, né? [...] É diferente**. (Inf. 12 – Capanema)

Quanto aos alemães, as justificativas para as respostas negativas, quando apresentadas, relacionavam-se à questão da língua.

Com relação aos italianos, também aqui se observa alto índice de aceitação. Em Pranchita, houve grande índice de perguntas não formuladas, e foi nessa localidade que se registraram as duas respostas negativas (sem razões apresentadas).

A tabela a seguir apresenta os resultados para o último grupo de questões – “Se precisasse de um médico ou dentista, procuraria um argentino / paraguaio / alemão / italiano? Por quê?” –, mostrando que as respostas positivas superaram as negativas em

relação a todas as etnias, com exceção dos resultados de Pranchita para os italianos, em que se verificou empate.

**Tabela 3: Reação dos informantes frente à possibilidade de recorrer a profissional de saúde de determinada etnia**

<i>Grupo étnico / Respostas</i>		<i>Localidade</i>	<b>Sto. Antônio do Sudoeste</b>	<b>Pranchita</b>	<b>Capanema</b>
<b>Argentinos</b>	Consultariam		78%	61%	61%
	Não consultariam		11%	28%	28%
	Outras respostas / NF		11%	11%	11%
<b>Paraguaios</b>	Consultariam		33%	50%	56%
	Não consultariam		17%	17%	33%
	Outras respostas / NF		50%	33%	11%
<b>Alemães</b>	Consultariam		78%	67%	83%
	Não consultariam		11%	22%	6%
	Outras respostas / NF		11%	11%	11%
<b>Italianos</b>	Consultariam		89%	33%	78%
	Não consultariam		11%	33%	11%
	Outras respostas / NF		–	34%	11%

Os informantes que se mostraram dispostos a procurarem um médico ou dentista argentino, de modo geral, alegaram que o fariam porque o que importa é a competência, o profissionalismo, e não o pertencimento a dada etnia. As justificativas para respostas negativas estão geralmente relacionadas à falta de preparo dos profissionais argentinos em relação aos brasileiros, devido principalmente à suposta situação de “atraso” em que se encontrariam as cidades argentinas próximas à fronteira.

- (25) [...] o problema da Argentina é que **eles tão bem aquém da nossa realidade**, eles tão... vamos dizer assim, dez a vinte anos, aqui na nossa região, aquém da nossa realidade. (Inf. 15 – Santo Antônio do Sudoeste)
- (26) **Não teria confiança.** [...] É a maneira deles, são **meio largadão**. Eu acho que não teria aquela confiança que tenho num médico brasileiro. (Inf. 5 – Santo Antônio do Sudoeste)
- (27) [...] eu acho que os curso deles não são que nem o nosso aqui no Brasil, eu acho que **eles não são muito bem preparados**. Eu acho que não, acho que **eu não confiaria**. (Inf. 4 – Pranchita)
- (28) **Não confiaria.** (Inf. 4 – Capanema)
- (29) Por causa que eu vejo falar tanto da medicina deles, que **não é de boa qualidade**, e pode ser que eu esteja errada, né, mas... (Inf. 12 – Capanema)
- (30) Agora entramos em um assunto que é porco. A maioria do pessoal da Argentina vem, até pra dentista, tudo aqui, loja, atendimento médico, posto, dentista é o que mais tem. [...] Não, não pela sensação, **não pela origem, mas pela situação que tá a cidade mais próxima aqui**. (Inf. 14 – Capanema)

Quanto aos paraguaios, desconsiderando-se o alto índice de perguntas não formuladas, principalmente em Santo Antônio do Sudoeste, pode-se considerar que, de forma geral, o índice de aceitação em relação aos profissionais dessa etnia é significativo. Po-

rém, as justificativas para respostas negativas dão indícios de uma visão de atraso político, educacional e cultural do Paraguai (da mesma forma que ocorre em relação aos municípios argentinos fronteiriços ao Brasil), o que se refletiria no preparo insuficiente do profissional paraguaio e no deficiente investimento em tecnologias na área da saúde.

- (31) Menos ainda [que o argentino]. [...] **A gente tem uma má impressão**, sei lá por quê... (Inf. 12 – Santo Antônio do Sudoeste)
- (32) De um paraguaio, daí complica (risos). Acho que mais brasileiro, acho que **a gente confia sempre, né, nos da gente**, mais assim, né? (Inf. 14 – Santo Antônio do Sudoeste)
- (33) Também por **falta de conhecimento, talvez falta de confiabilidade**, né, e de **local** que a gente conhece que eles trabalham lá, e **equipamentos** também, né. (Inf. 3 – Capanema)
- (34) Eu acho um pouco por causa da política, né [...] que nem eu falei, **a estrutura, eu acho que Paraguai falta muito crescer**, né, é o ensino, o... parte funcional deles, né. (Inf. 9 – Capanema)
- (35) [...] pode ser que eu esteja enganada, mas é... menos, **tem menos desenvolvimento lá, menos escolaridade do que aqui**, né... então eu teria **falta de confiança**. (Inf. 12 – Capanema)

Em relação aos alemães, o índice de aceitação foi bastante significativo. Ao contrário dos paraguaios e argentinos, os alemães são vistos, de modo geral, como mais competentes e dedicados, como exemplificam as respostas a seguir:

- (36) Com certeza, **teria informação bem mais...** até pela... **pela cultura** e pela... **pela descendência**, acredito que **seria bem melhor do que um argentino**. (Inf. 15 – Santo Antônio do Sudoeste)
- (37) Eu sim. E por cima, **eles são muito inteligente**. (Inf. 10 – Pranchita)
- (38) Ah, porque também **são bem confiável** nessa parte, né. (Inf. 3 – Capanema)

Por fim, os índices de aceitação em relação à possibilidade de consultar profissionais de origem italiana foram bastante significativos nas localidades de Santo Antônio do Sudoeste e Capanema. Em Pranchita, houve empate entre respostas afirmativas e negativas (um terço cada), e alto índice de perguntas não formuladas (também um terço (33%).

## Considerações finais

Os resultados mostram, de modo geral, tendências de reação positiva por parte dos informantes com relação a todas as etnias, embora com diferentes índices de aceitação e rejeição entre uma etnia e outra. As justificativas apresentadas para as respostas positivas geralmente dizem respeito à desvinculação da etnia às atitudes dos informantes (morar em bairro x, namorar pessoa y, consultar médico z), ou seja, o pertencimento étnico do(s) indivíduo(s) não seria em nada decisivo para as escolhas do informante. As respostas negativas, por sua vez, dizem respeito muito mais à possibilidade de não compreender o idioma do outro do que propriamente uma rejeição linguística/cultural com relação a indivíduos e grupos de outras etnias.

No entanto, algumas respostas – embora pouco representativas – dão margem a interpretações no sentido de que determinados grupos são mais rejeitados que outros, com recurso, inclusive, a estereótipos relacionados aos membros das diferentes etnias. São as

respostas desse tipo que se procurou destacar neste estudo, pois representam uma espécie de delimitação entre o “nós” e “os outros”. Verifica-se, portanto, que as relações afetivas, profissionais etc. são mediadas por questões identitárias. Entende-se que o conceito de identidade, definido tanto de forma objetiva quanto subjetiva, envolve a questão da língua, porque uma comunidade também se caracteriza pela(s) variedade(s) linguística(s) usada(s) em seu seio, e também porque a percepção do comunitário e do diferencial se faz especialmente evidente por meio dos usos linguísticos (MORENO FERNÁNDEZ, 1998).

Conforme Aguilera (2008, p. 106), “[...] a atitude linguística de um indivíduo é o resultado da soma de suas crenças, conhecimentos, afetos e tendências a comportar-se de uma forma determinada diante de uma língua ou de uma situação sociolinguística”. Nesse sentido, se as questões de caráter conativo atuam coerentemente com as de caráter cognitivo e afetivo, como propõe a abordagem mentalista, os resultados deste estudo permitem prever que também as crenças e a valoração atribuídas às variedades faladas nas localidades do Sudoeste do Paraná e aos seus falantes serão majoritariamente de prestígio, especialmente no que diz respeito aos argentinos e italianos e suas respectivas variedades.

## REFERÊNCIAS

AGUILERA, V. A. *Crenças e atitudes linguísticas: um estudo da relação do português com línguas de contato*. 2009. [Projeto desenvolvido pela autora. Digitado].

\_\_\_\_\_. *Crenças e atitudes linguísticas: o que dizem os falantes das capitais brasileiras. Estudos Linguísticos*, São Paulo, v. 2, n. 37, p. 105-112, maio/ago. 2008.

BEM, D. J. *Convicções, atitudes e assuntos humanos*. Tradução de Carolina Martuscelli Bori. São Paulo: EPU, 1973. (Coleção Ciências do Comportamento).

BLANCO CANALES, A. *Estudio sociolingüístico de Alcalá de Henares*. Alcalá de Henares, Madrid: Servicio de Publicaciones de la Universidad de Alcalá, 2004.

CALVET, L. J. *Sociolingüística: uma introdução crítica*. Tradução de Marcos Marcioni-lo. São Paulo: Parábola, 2002.

FISHMAN, J. A. *The sociology of language: an interdisciplinary social science approach to language in society*. Rowley, Massachusetts: Newbury, 1972.

FROSI, V. M.; FAGGION, C. M.; DAL CORNO, G. O. M. Turpilóquio: o léxico do falar na linguagem oral da região de colonização italiana do nordeste do Rio Grande do Sul. In: ENCONTRO DO CELSUL, 8., 2008, Porto Alegre. *Anais...* Pelotas: Educat, 2008. v. 8, p. 255-256.

GÓMEZ MOLINA, J. R. Actitudes lingüísticas en Valencia y su área metropolitana: evaluación de cuatro variedades dialectales. In: CONGRESO INTERNACIONAL DE LA ASOCIACIÓN DE LINGÜÍSTICA Y FILOLOGÍA DE LA AMÉRICA LATINA – ALFAL, 11, 1996, Las Palmas de Gran Canaria. *Actas...* Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria, 1996. v. 2, p. 1027-1042.

GROSJEAN, F. *Life with two languages: an introduction to bilingualism*. Harvard: Harvard University Press, 1982.

LAMBERT, W. W.; LAMBERT, W. E. *Psicologia social*. Tradução de Álvaro Cabral. Rio de Janeiro: Zahar, 1966.

LAZIER, H. *Paraná: terra de todas as gentes e de muita história*. 3. ed. Francisco Beltrão: Grafit, 2003.

LIEBKIND, K. Social psychology. In: FISHMAN, J. A. (Ed.). *Handbook of language and ethnic identity*. New York: Oxford University Press, 1999. p. 140-151.

LÓPEZ MORALES, H. *Sociolingüística*. 2. ed. Madrid: Gredos, 1993.

MORENO FERNÁNDEZ, F. *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel, 1998.

PADILLA, A. M. Psychology. In: FISHMAN, J. A. (Ed.). *Handbook of language and ethnic identity*. New York: Oxford University Press, 1999. p. 109-121.

PASTORELLI, D. S. *Crenças e atitudes linguísticas na cidade de Capanema: um estudo da relação do português com línguas em contato*. 2011. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2011.

SILVA-PORELI, G. A. *Crenças e atitudes linguísticas na cidade de Pranchita – PR: um estudo das relações do português com línguas em contato*. 2010. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2010.

STURZA, E. R. *Línguas de fronteira: o desconhecido território das práticas linguísticas nas fronteiras brasileiras*. *Ciência e Cultura*, São Paulo, v. 57, n. 2, p. 47-50, abr./jun. 2005.

TARALLO, F. *A pesquisa sociolingüística*. São Paulo: Ática, 1985.

WACHOWICZ, R. C. *Paraná, sudoeste: ocupação e colonização*. Curitiba: Lítero-Técnica, 1985.

# ***V-ra* no português: uma análise diacrônica**

(*V-ra* in Portuguese: a diachronic analysis)

**Kellen Cozine Martins,<sup>1</sup> Maria da Conceição de Paiva<sup>2</sup>**

<sup>1,2</sup>Programa de Pós-graduação em Linguística – Universidade Federal do Rio de Janeiro (UFRJ/CNPq)

kellencozine@yahoo.com.br, paiva@club-internet.fr

**Abstract:** The morpheme *-ra* to indicate pluperfect (*mandara*) practically disappeared from Brazilian Spoken Portuguese, being replaced by other variants, as the periphrase *ter/haver* + past participle (*tinha mandado*) and the simple past (COAN, 1997; MARTINS, 2010). Traces of this verb form remain in formulaic constructions, such as *quem me dera, tomara, quisera, pudera eu*, derived from the possibility of the uses of the morpheme with modal sense. The aim of this research is to discuss the contexts of modal readings of the morpheme *-ra*, during the period between sixteenth and twentieth centuries, and its idiomatization in contemporary Portuguese.

**Keywords:** morpheme *-ra*; diachronic analysis; schematization.

**Resumo:** O morfema *-ra* para a indicação de pretérito mais-que-perfeito (*mandara*) praticamente desapareceu do português brasileiro falado, cedendo espaço a outras formas concorrentes, como a perífrase *ter/haver* + particípio passado (*tinha mandado*) e o perfeito simples (*mandou*) (COAN, 1997; MARTINS, 2010). Resquícios dessa forma verbal resistem em construções optativas cristalizadas, como *quem me dera, tomara, quisera, pudera eu*, originadas na possibilidade de usos do morfema *-ra* com valor modal [-*realis*]. O objetivo deste artigo é discutir os contextos de leitura modal do morfema *-ra*, no período compreendido entre os séculos XVI a XX e sua idiomatização no português contemporâneo.

**Palavras-chave:** morfema *-ra*; análise diacrônica; esquematização.

## **Introdução**

Uma característica conservadora do sistema modo-temporal do português em relação a outras línguas românicas é a preservação da forma simples (*amara*), herdeira do mais-que-perfeito latino (*amaveram*), para a expressão de anterioridade a um ponto de referência passado. Estudos diacrônicos mostram, no entanto, que, pelo menos a partir do século XV, essa forma já encontrava concorrência na perífrase *ter/haver* + particípio passado, e mesmo na forma de pretérito perfeito<sup>1</sup> (cf. p. ex. MATTOS E SILVA, 2001; BROCARDO, 2010). Produtivo em períodos anteriores do português, como testemunham numerosos exemplos do português medieval, o morfema *-ra* de mais-que-perfeito se torna marginal no português brasileiro falado, preservando-se, ao que tudo indica, apenas em registros literários ou em estilos escritos mais formais.

Além do seu valor temporal, o morfema *-ra* admitia usos modais [-*realis*], igualmente herdados do Latim, equivalentes ao do imperfeito do subjuntivo e ao do futuro do pretérito, já destacados por diferentes autores (SAID ALI, 1964; BECKER, 2008; BROCARDO, 2010). Também esses usos, comuns entre autores quinhentistas e seiscentistas, decrescem

---

<sup>1</sup> No decorrer do presente artigo, são usadas as seguintes abreviaturas: PMQPS para se referir a pretérito mais-que-perfeito simples; PMQPC, para pretérito mais-que-perfeito composto; PPS, para pretérito perfeito simples; V para verbo; EsC para estado de coisas.

ao longo dos séculos, desaparecendo quase inteiramente no português contemporâneo. No entanto, resquícios desses valores modais podem ser encontrados em expressões cristalizadas ou formulaicas com valor optativo, tais como *quem me dera, tomara, quisera, pudera eu*, como já atestado por Coan (1997).

O objetivo deste artigo é discutir algumas evidências diacrônicas de que essas construções optativas cristalizadas com *-ra* resultam dos empregos do morfema *-ra* com valor [-*realis*]. Através de uma análise de documentos escritos do século XVI ao século XX<sup>2</sup> e de dados de fala do português brasileiro contemporâneo,<sup>3</sup> defendemos que a cristalização/idiomatização dessas construções optativas se insere no quadro mais amplo da perda de produtividade da desinência *-ra* ao longo dos séculos. Adotando pressupostos teóricos propostos por Bybee (2010), interpretamos essa perda de produtividade em termos do conceito de esquematização. Acreditamos que a construção optativa, em estágios anteriores, admitia um leque mais amplo de formas verbais e foi gradualmente se convencionalizando pela ação do processo cognitivo denominado *chunking*. De acordo com Bybee (2010), quando duas ou mais palavras são frequentemente usadas juntas, elas podem desenvolver uma relação sequencial (*chunking*) e a força de entrincheiramento (conexão) resultante da repetição faz com que essa sequência de unidades seja estocada na memória como uma única unidade de processamento linguístico e produzidas de forma automática pelos falantes. Acrescenta a autora que “todos os tipos de expressões convencionalizadas, de pré-fabricadas às idiomáticas e às construções, podem ser consideradas *chunks* para fins de processamento e análise” (BYBEE, 2010, p. 35, tradução nossa).

Este artigo compreende três seções: na primeira, tecemos considerações a respeito dos usos da desinência *-ra* em português; na segunda, focalizamos os diferentes contextos

2 Para o levantamento diacrônico, foram utilizados os seguintes textos: SÉCULO XVI: Textos extraídos do Corpus Internacional de Língua Portuguesa, são eles: Cartas de D. João III, Carta Jesuítas, Carta de Caminha, JBarros Dil Vic Vergonha; SÉCULO XVII: Textos extraídos do Corpus Histórico do Português Tycho Brahe, são eles: A arte de furtar (1645), Cartas familiares (1646), Tácito (1646), Côrte na Aldeia e Noites de Inverno (1607), Gazeta (1642), História do futuro (1667), Monarchia Lusitana (1619), Poesia e Pintura (1635), Sermões (1667); SÉCULO XVIII: Textos extraídos do Corpus Internacional de Língua Portuguesa, são eles: Carta Aldeamento de índios, Correspondência da administração pública do estado da Bahia, Cartas Comerciais do estado do Maranhão, do Rio de Janeiro, de Minas Gerais, Textos da administração pública (CILP1BMPBC), Representação oficial (CILP1BMRERJ), Carta-relatório (CILP1PMCORJ). Textos extraídos do Projeto para a História do Português Brasileiro (PHPB): Cartas Oficiais, Cartas Pessoais, Cartas Comerciais, Documentos Oficiais e Documentos Particulares. Textos extraídos do Corpus Histórico do Português Tycho Brahe: Cartas Brasileiras; SÉCULO XIX: Textos extraídos do Corpus Internacional de Língua Portuguesa, são eles: Cartas Comerciais do estado do Paraná, Cartas Aldeamento de índios do estado de São Paulo, Cartas Pessoais dos estados de Minas Gerais e Rio de Janeiro. UNESP (CGM, DMA, IRP, RNG e RV) e Cartas de leitores dos estados (Paraná, Bahia, Minas Gerais, Pernambuco, Rio de Janeiro e São Paulo).

3 Os dados de fala do português contemporâneo reúnem textos extraídos do Projeto Censo/PEUL e do Projeto NURC-RJ (Projeto da Norma Urbana Oral Culta do Rio de Janeiro). O Projeto Censo é composto pelas amostras Censo 80, representativa da variedade carioca falada no início da década de 80, e Censo 00, gravada no início dos anos 2000. A Amostra 80 reúne 64 entrevistas sociolinguísticas e foi estratificada consoante as variáveis independentes: a) idade (7 a 14 anos, 15 a 25 anos, 26 a 49 anos e acima de 50 anos); b) escolaridade (1º e 2º ciclo do ensino fundamental e médio); c) gênero (feminino e masculino). A Amostra 00 é constituída de 32 entrevistas organizadas de acordo com os mesmos parâmetros de estratificação da Amostra 80, que lhe serviu de base. Visando a compatibilizar as duas amostras, para este estudo, foram utilizadas apenas 32 entrevistas da Amostra Censo 1980. O acervo do Projeto NURC-RJ reúne entrevistas, gravadas durante as décadas de 1970 e 1990, com informantes de nível superior completo e nascidos no estado do Rio de Janeiro.

modais da construção *V-ra* ao longo de cinco séculos; na terceira, interpretamos as evidências depreendidas à luz dos princípios cognitivos envolvidos na idiomatização dessas construções optativas. Seguem-se as considerações finais.

## Usos de *V-ra*

A construção focalizada, neste estudo, é constituída por um verbo seguido do morfema *-ra* (*cantara*), que, por definição, indica a anterioridade a um estado de coisas passado, como no exemplo (01), do PB escrito contemporâneo.

- (01) “Ela me contou que, como não tinha lesões aparentes, os policiais a orientaram a representar no 7o DP (*onde **registrara** outro B.O. 15 dias antes*), porque na ocasião tinha feito exame de corpo de delito. No dia seguinte, ela foi ao 7o DP e representou”, diz Ana Maria. (*Época*, 10/04/2009 – Edição n. 569)<sup>4</sup>

Segundo diversos autores (MAURER JR., 1959; THIBAUT, 1993; BECKER, 2008; BROCARDO 2010), o português preservou o sentido latino do morfema *-ra*, distinguindo-se, assim, de outras línguas românicas em que essa desinência desapareceu, seguindo, de acordo com hipótese de Maurer Jr. (1959), uma tendência que remonta ao Latim Vulgar. Ainda de acordo com esse autor, a concorrência das formas analíticas (*laudatum habebam*) teria contribuído de forma decisiva para a decadência da desinência *-ra*.

Todavia, mesmo no português, o baixo rendimento da forma simples de pretérito mais-que-perfeito se faz notar não só na modalidade falada, mas até mesmo em registros escritos mais formais, como já observado por Mattoso Camara Jr. (1984). Segundo o autor, em lugar do PMQPS, “ou se emprega o pretérito perfeito, que não está formalmente marcado [...] ou se substitui por uma locução de participio com o verbo auxiliar *ter* no pretérito imperfeito (tinha cantado em vez de – cantara)” (MATTOSO CAMARA, 1984, p. 10).

Pesquisas, principalmente de orientação variacionista, permitem confirmar essa perda da produtividade de *V-ra* para indicação de passado do passado (cf. p. ex. COAN, 1997, 2003; MARTINS, 2010, 2011). Coan (1997) constata que, no português falado, a variação se estabelece, basicamente, entre as variantes pretérito perfeito simples e pretérito mais-que-perfeito composto (com o auxiliar *ter*), não sendo registradas, na amostra controlada pela autora, ocorrências de mais-que-perfeito simples. Martins (2010, 2011) atesta apenas 12% de *V-ra* em textos escritos do domínio jornalístico, mostrando que essa forma é largamente suplantada pelo emprego de PPS (55%) ou pela perífrase com *ter/haver* (31%). A autora pode constatar, ainda, maior limitação nos contextos linguísticos de ocorrência da forma simples de pretérito mais-que-perfeito, tal como a existência de um outro marcador de referência temporal (advérbios ou sintagmas preposicionais de referência temporal, orações adverbiais temporais) que reforça a anterioridade de um estado de coisas em relação a outro.

Como constata Martins (2010, 2011), o caráter mais formal da forma simples de pretérito mais-que-perfeito é atestado, na medida que a variante apresenta uma distribuição

4 No exemplo (01), trata-se de um contexto de relações temporais complexas, caracterizado por apresentar diferentes EsC passados que estão relacionados a um ponto de referência particular, a forma verbal *contou*. Como explica Martins (2010, p. 26): o verbo *contou* expressa anterioridade à origem, tornando-se ponto de referência para a situação descrita pelo verbo *orientaram*, que, por sua vez, é convertido em referência sobre o qual se ancoram as situações expressas pelos verbos *registrara* e *tinha feito*.

diferenciada não só em relação aos diferentes jornais e revistas analisados, como também aos diferentes gêneros textuais. A maior recorrência da forma de PMQPS é observada em jornais e revistas voltados para um público-alvo menos popular, como o jornal *O Globo* e a revista *Época*, que conservam uma escrita mais formal. Relativamente ao gênero discursivo, há maior produtividade de PMQPS nas crônicas, provavelmente, em virtude da natureza literária desse gênero, e em notícias, o que pode ser explicado pela própria estrutura essencialmente narrativa desse gênero, e, conseqüentemente, o predomínio de outros elementos demarcadores de tempo e espaço.

Como já destacado por outros autores (cf. p. ex. SAIDALI, 1964; BECKER 2008), a desinência *-ra* admitia também, já desde o período medieval, usos modais, equivalentes ao de futuro do pretérito ou de imperfeito do subjuntivo, ilustrados, respectivamente, por (02) e (03):

- (02) A decima batalha foy con o alcaide Genyo Filnio que era adeantado dos Roma~a~os. E foy ve~çudo Genyo e *perdeo* hy toda sua co~panha e elle **ouvera** hy a seer morto e escapou per ventuyra (Crónica Geral de Espanha / Séc. XIV – Cap. 54 - 294).
- (03) E fogyo Varro, consul, a Veneza co~ quinhe~tos cavalleyros. E desta batalha e das outras que avya vençudas dos Roma~a~os *e~vyou* a Afryca tres moyos de sortelhas d'ouro que forom tomadas aos senadores e aos outros home~e~s honrrados. E tanto foron coytados e desmanhados os Roma~a~os desta batalha que estiverom a ponto de leixar Roma e Ytalia e desemparalas se no~ **fora** Cepion, o Ma~cebo (Crónica Geral de Espanha / Séc. XIV – Cap. 54 - Texto 294).

Essa possibilidade de emprego do PMQPS com valor [-*realis*],<sup>5</sup> segundo Ravizza (1958, p. 265), já era atestada em latim, principalmente com os verbos modais *poder* e *dever*, expressões modalizadoras como “ser necessário”, “ser desejável”, com os verbos *credo*, *puto*, *arbitror*, *opinor* e análogos, especialmente se precedidos de negação (Ex: **non crediderat** = não teria acreditado, **melius fuerat** = teria sido melhor), e na apódose de períodos condicionais (BROCARDO, 2010).

Esse uso [-*realis*] da desinência *-ra* sobreviveu em outras línguas românicas, como destaca Maurer Jr. (1959):

Com vestígios do sentido latino, mas geralmente com valor de condicional, este tempo se encontra no espanhol, no catalão e no provençal antigos, e na Itália meridional (na idade média, também em Roma e na Úmbria); no francês, nos primeiros textos, com valor de simples pretérito: e.g., treze exemplos no Saint Léger: *laissera* (v. 126), *estera* = *esteve* (v.230), *fisdra* = *fêz* (v.121), *vindre* < *venerat* (v.202), etc. Sobrevive finalmente no sardo antigo, e.g., *levarat*, *fekerat*, *poserat*. Um vestígio desse tempo no Oriente subsiste certamente na 3ª pessoa do plural do perfeito rumeno, *cântarǎ*, *tăcurǎ* de *cantarant* e *tacúerant* respectivamente. (p. 126)

No entanto, ao que tudo indica, no português, esses empregos foram diminuindo progressivamente ao longo do tempo (BECKER, 2008). Comparando a trajetória divergente do morfema *-ra* em espanhol e em português, Becker (2008) afirma que, em espanhol, o emprego de *-ra* modal foi forte o suficiente para invalidar a interpretação tradicional temporal do morfema, ilustrando um tipo (ou *cline*) de mudança gramatical em que uma

5 Um EsC [+*realis*] é um que realmente aconteceu ou realmente está acontecendo, enquanto um EsC [-*realis*] refere-se a situações hipotéticas no passado ou no futuro, incluindo também desejos, metas e previsões sobre o futuro (v. LYONS, 1977; COMRIE, 2004 [1985]).

categoria de subjuntivo é derivada de uma de indicativo. No português, também são detectados usos de *-ra* em contextos tipicamente condicionais e modais, restritos ao subjuntivo *-sse* e ao futuro do pretérito *-ria*. Contudo, enquanto, no espanhol, a mudança avança, no português, há a perda de vitalidade de usos modais de *-ra*, que “retrocedeu durante o século 19 a contextos condicionais, transformando o uso de *-ra* em contextos *irrealis* uma estratégia estilística altamente marcada no reino da retórica pomposa” (BECKER, 2008, p. 166, tradução nossa). Ainda nos termos do autor:

em português, os desenvolvimentos tomaram a direção oposta: a contra-evidência disponível, que poderia ter negado o estatuto mais-que-perfeito de *-ra* foi simplesmente insuficiente. Os usos modais, que não eram predominantes, puderam ser reduzidos como extensões de função temporal com o efeito de que nenhuma mudança a uma reinterpretação [*-realis*] tenha ocorrido. Como vimos, algumas intrusões em outros contextos [*-realis*] estavam no caminho, mas esses movimentos permaneceram como tentativas e não foram repetidos pelo resto da comunidade de fala. (BECKER, 2008, p. 173-174, tradução nossa)<sup>6</sup>

Embora também o uso modal da desinência *-ra* se torne mais e mais rarefeito ao longo dos séculos,<sup>7</sup> resquícios da forma de PMQPS podem ser atestados nas construções optativas, idiomatizadas, que expressam uma expectativa futura, como desejo ou esperança. Como nota Othon M. Garcia (2010 [1967]):

O pretérito mais-que-perfeito (simples), além do seu sentido fundamental (de fato passado anterior a outro), pode, com certos verbos, conotar desejo ou esperança (linguagem optativa): “Ah! Quem me dera recuperar o tempo perdido! Prouvera a Deus que tal coisa fosse possível! Quisera ter hoje a idade de meus filhos!” Às vezes, tem sentido difícil de bem caracterizar, valendo como uma estrutura: “Você foi reprovado. Pudera! Não estuda!” Subtendendo-se o “pudera” como: nem poderia ser de outra forma, era coisa de prever. (p. 93)

Na seção seguinte, focalizamos, de forma mais detalhada, os usos modais do morfema *-ra* no período compreendido entre os séculos XVI e XX com o objetivo de identificar seus contextos e tecer algumas considerações a respeito da perda de esquematização de *V-ra*. Evidentemente, dado o recorte temporal deste estudo, não nos atrevemos a traçar uma trajetória dessas construções.

## Usos modais da construção *V-ra*

Em textos datados do século XVI, o morfema *-ra*, além do seu emprego canônico para marcar um distanciamento temporal, fazendo alusão a um momento passado anterior a outro, admitia valor equivalente ao do futuro do pretérito ou do imperfeito do subjuntivo, em contextos bastante próximos daqueles atestados já no latim, quais sejam, em orações complemento de verbos epistêmicos, como *crer* e *parecer*, e nas orações constituintes de períodos condicionais, como nos exemplos (04) e (05):

6 Face a isso, o autor indaga se devemos assumir um “limiar crítico” de dados responsáveis para mudança ou se temos de aceitar o simples fato de que resultados opostos podem ocorrer, mesmo em línguas muito relacionadas como o espanhol e o português.

7 A última ocorrência atestada por Said Ali (1964) do uso das terminações *-ara*, *-era* e *-ira* substituindo o imperfeito do subjuntivo e o condicional data do século XVII.

- (04) me sprevaes se vos parece bem o modo em que lhe respondy ao parecer e conselho que me deu no negocio de Franca, ou se vos *parece* que **fora** melhor d'outra maneira e quejando vos *parece* que **devera** ser, e a estas tres cartas deradeiras o que vos parece que lhe devo responder. (Cartas D. João III, s. XVI)
- (05) e por que nõ leva outro mandado meu senam este pera navegar esa fazemda, que se nom atreveria a tomar seobre sy nenhũa obrigaçam que podese trazer allguũ perjuizo ao carreguo que leva; pelo qual lhe pede que ho ajam asy por bem, avemdo por çerto que, *se* d'outra maneira **fora**, **follgara** muyto de lhe daar cõserva e se **ofercera** pera iso. (Cartas D. João III, s. XVI)<sup>8</sup>

As formas destacadas nos exemplos acima (*fora*, *devera*, *follgara* e *ofercera*) situam o estado de coisas como uma possibilidade, afastando-se, assim, do emprego [+ *realis*] do PMQPS. Nesses contextos, passam a indicar um distanciamento ou não comprometimento do falante quanto à veracidade ou validade do EsC descrito.<sup>9</sup> Adquirem, portanto valor [+ subjetivo], de acordo com a trajetória proposta por Traugott (2010, 2011) e, Traugott e Dasher (2002).

Uma interpretação de *V-ra* como futuro do pretérito é possível, ainda, em orações relativas em que se faz presente um verbo modal precedido de negação (SN [que não Vmodal-*ra*]):

- (06) E quamto á fazenda da urca, pello meu feitor de Framdes se pode fazer imteyrame)te a prova da que era minha; e a fara com a brevidade que lhe spreverdes que o faça, porque se caregou por elle, e estaa ase)tado em livro a que se daa fee em outras mayores cõtias; e a demasya que he das partes, pois elle o spreve em numero çerto, deve saber pellas adições de cujo he; que d'outra maneira o *nam* **podera** sprever em numero çerto. (CILP1BMCDJ3 - Cartas D. João III, s. XVI - p. 38)

A combinação de verbo modal e de um operador de negação (*não* ou *nunca*) constitui uma estratégia de avaliação ou de justificação para um dado EsC que, de outro modo, não poderia ser realizado, pois seria contrário às evidências ou, no mínimo, às preferências do falante.

O valor de imperfeito do subjuntivo pode ser atestado, no mesmo período, numa estrutura variante do esquema condicional que é a oração comparativa-condicional ([como se (não) + V-*ra*]), como no exemplo (07):

- (07) acabado ocomer metemo nos todos no batel e eles cõ nosco ·/ deu hu)u) gromete ahu)u) deles hu)u)a armadura grande de porco montes bem Reuolta e tamto que atomou meteo a logo no beicho e por que se lho no) queria te)e)r · derã lhe hu)u)a pequena decera vermelha e ele corejeo lhe detras seu aderemço pera se te)e)r e meteo a no beicho asy Reuolta pera cjma e vijnha tam comtente com ela *como se* **teuera** hu)u)a grande joya (CILP1PACDCM - Carta de Caminha XVI - p. 21).

8 No exemplo (05) a forma *ofercera* apresenta uma certa ambiguidade, admitindo também uma interpretação temporal.

9 Segundo Lyons (1977, p. 797, tradução nossa), “qualquer enunciado em que o falante explicitamente qualifica seu comprometimento com a verdade da proposição expressa pela sentença que ele declara, seja essa qualificação feita explicitamente no componente verbal (como *may* ou *it's possible*) ou no componente prosódico ou pragmático, é um enunciado epistemicamente modal ou modalizado”.

Destaca-se, ainda, no século XVI, a alta frequência (22 ocorrências) de *V-ra* [-*realis*] em contexto contrastivo ([V (modal)<sup>10</sup> -*ra* → (mas X)], como nos exemplos (08) e (09):

- (08) Espero em Noso Senhor que vos dee imteyra saude. E mãey loguo despachar este coreo pera vos levar esta minha carta; o quall nam vay a outra cousa. E **prazera** a Noso Senhor que me trara tam boas novas de vos, e de vosa saude, como eu agora e sempre as querya saber (CILP1BMCDJ3 - Cartas D. João III, s. XVI - p. 43).
- (09) Espero em Noso Senhor que tudo aproveyte de maneira que ho xarife se allevante cõ muyto dano seu. Eu **folgara** muyto de vos *mãdar* lloguo vyr; porem, pollas cartas que me agora trouxe Vicente Reynell, que partio de Çafim nos Xi d’este mes, e pollo que elle diz do xarife estar de vagar, cõ sua opiniã me parece neçesaria e muyto meu serviço esperardes llaa, ate vyr o primeiro Recado de dõ Garçia, e isso que estaa prestes nesa cidade estar como estaa (CILP1BMCDJ3 - Cartas D. João III, s. XVI - p.128).

No exemplo (08), o uso do desiderativo *prazer* [do latim *placere*] equivale a “Deus queira, queira Deus” ou na base de um esquema condicional “Se a Deus prouver”. No exemplo (09), a construção *V-ra* apresenta um desejo não-realizado do falante e o contraste é expresso explicitamente, pela presença de uma oração adversativa que apresenta a(s) condição(ões) que inviabilizou(-ram) a realização do desejo. O enunciado como um todo funciona como uma condição de polidez, um pedido de desculpa por algo que o falante se julgava na obrigação (ou condição) de realizar.

Necessário destacar que, no esquema contrastivo, predomina a forma verbal *prazer* (19 ocorrências), embora outros verbos, também de natureza modal/desiderativa possam ocorrer tais como *querer*, *folgar*.

Em documentos do século XVII, embora se mantenham os esquemas já presentes no século anterior, algumas observações podem ser feitas. Parece haver uma expansão do esquema condicional, de modo que a interpretação [-*realis*] pode ser atestada não apenas em orações condicionais e comparativo-condicionais, mas também em orações causais, concessivas, consecutivas, conformativas, como no exemplo (10):

- (10) Se as gotas de sangue que derramei por ti, tiveram conta, nem de uma só me puderas dar boa conta, *ainda que padecera* por mim mil mortes; mas os milhares e os milhões foram das vezes que pizaste o mesmo sangue, sacrificando o infinito valor e merecimento d’elle, aos idolos do teu appetite. (Sermões – p. 55)

Neves (1999) destaca a proximidade entre relações condicionais, causais e concessivas, na medida que são “todas elas expressivas de uma ‘conexão causal’ lato *sensu* entendida, e todas elas explicáveis em dependência de satisfação/não-satisfação de necessidade/ de suficiência de determinadas condições” (p.545). No exemplo acima, a realização da

---

10 Os modais incluem tanto os deônticos como os epistêmicos. A origem da modalidade deôntica é frequentemente buscada na função desiderativa (isto é, expressão ou indicação de um desejo) e instrumental (conseguir que algo seja feito pela imposição do desejo de alguém a outros agentes) da linguagem. Lyons (1977) aponta três diferenças principais entre enunciados epistêmicos e deônticos: a) o primeiro tem a ver com a verdade de proposições; já o segundo com a necessidade ou possibilidade de atos realizados por agentes moralmente responsáveis; b) há uma conexão intrínseca entre modalidade deôntica e futuridade, de maneira que a verdade de uma proposição modalizada deonticamente é determinada relativa a algum estado de mundo ( $w_t$ ) posterior ao estado de mundo ( $w_s$ ) em que a obrigação se sustenta; e este não pode preceder, embora possa ser simultâneo, ao estado de mundo ( $w_o$ ) em que a obrigação se impõe; c) necessidade deôntica geralmente procede, ou deriva, de alguma fonte ou causa.

proposição descrita na oração concessiva (padecer por mim mil mortes) não constitui condição suficiente para satisfazer a proposição da oração principal (poder dar boa conta das gotas de sangue que derramei por ti). Em outras palavras, o resultado da principal independe do que é descrito na concessiva.

Neste século, podemos verificar, ainda, a forte redução do verbo *prazer*, no esquema contrastivo [V (modal) *-ra* → (mas X)], com apenas 1 ocorrência. Mantêm-se produtivos os modais *querer* (3 ocorrências), *desejar* (1 ocorrência) e *poder* (3 ocorrências). Destaque é dado ao emprego do verbo *tomar* (4 ocorrências), indicando vontade de tomar posse ou de gozar de algo. Nesse caso, o traço desiderativo, que não é uma característica do verbo, emerge do contexto.

- (11) O primeiro instrumento da prática é a voz; e, para essa ser engraçada no falar, há-de ter estas propriedades: ser clara, branda, cheia e compassada; porque a voz escura confunde as palavras; a áspera e seca tira-lhe a suavidade; a muito delgada e feminina faz imprópria a acção do que fala; a muito apressada empeça e revolve as razões que per si podem ser muito boas; não trato das que a natureza inabilitou para esta perfeição, como é a voz do gago, do cicioso e do rústico grosseiro; mas na do cortesão **tomara** eu êstes atributos, porque há alguns que falam com a voz tão metida por dentro que deixam as palavras para si e os ouvintes às escuras, que lhes é necessário estar espreitando o que lhes querem dizer (Côrte na Aldeia e Noites de Inverno - p. 55).

Uma outra alteração verificada no século XVII diz respeito à expansão do esquema relativo para orações de polaridade afirmativa. O elemento de negação torna-se facultativo no esquema (SN [que (não) *V-ra*]), como mostra o exemplo (12):

- (12) Raynuncio tambem oppositor já era bisneto na linha do Infante Dom Duarte; mas não se fez caso da sua opposição, por ser defunta sua mãe, que a **devera** fazer, e por não constituir linha diferente da em que se achava a Senhora Dona Catharina, em melhor gráo que elle (A arte de furto - p. 58).

O exemplo (12) ilustra um caso em que o conteúdo da relativa não pode se sustentar no mundo real seja porque o referente não existe seja porque é incapaz de realizar dada ação devido a determinadas circunstâncias.

Para além dos contextos encontrados no século anterior, há, ainda, o registro de outros dois esquemas: o esquema comparativo composto por um sintagma de superioridade + *V-ra* [Melhor *V-ra* do que X] e o de pergunta retórica introduzida pelo interrogativo *quem* [Quem *V-ra* X?], exemplificados, respectivamente, em (13) e (14):

- (13) Poderia ser assim: Hum Ministro, que tinha por officio pagar quartéis de juros, e tenças a todo o mundo, foy sonegando muito a titulo de não haver dinheiro; e em poucos annos com esta, e outras industrias taõ maliciosas, como esta, ajuntou mais de cem mil cruzados, de que deu oitenta mil a ElRey nosso Senhor, gabando-se que os poupava aos poucos, e que eraõ frutos (*melhor dissera* furtos) da pontualidade, e primor, que guardava em seu Real serviço (A arte de furto - p.105).
- (14) Se soubesses que vos não restava de vida mais que um mez, haviéis de chorar: e rides, e andaes alegre e contente, podendo ser que vos não reste um dia inteiro. **Quem dissera** a el-rei Balthasar, quando com tanta festa e alegria estava brindando aos seus idolos nos proprios vasos sagrados de oiro e prata que Nabucodonosor seu pae tinha roubado ao templo de Jérusalem: **quem** lhe **dissera** que a mesma noite d'aquella cêa fatal, era a ultima da sua vida, e dasua corôa? (Sermões - p.19)

Nos séculos XVIII e XIX, quase todos os esquemas de *V-ra* [*-realis*] atestados nos séculos anteriores se mantiveram, com exceção do esquema [V epistêmico + *V-ra*]. Mantém-se, também, promissor o emprego dos modais *poder*, *dever*, *querer* e *desejar*.

Cabe ressaltar, no entanto, que, diferentemente do que foi verificado no século anterior, há a fixação do verbo *dar* no contexto de pergunta retórica [Quem *V-ra* X?], como ilustra o exemplo (15):

- (15) Aquelle que tem mulher e filhos outro remedio não tem senão procurar-lhe o alimento trabalhando para não ser pesado aos outros; é o que me acontece como pae de familia. Indo eu procurar em algumas casas de negocio licenças para tirar, respondeo-me um dono da casa: - meu amigo, as minhas licenças eu as dou a um guarda da camara para as tirar, porque só assim fico descansado de multas e outros incommodos, embora pague por cada uma 3 ou 4 \$ réis, porque elles lá se entendem uns com os outros; e dizendo a esse bom homem que eu as tirava a 1 \$ réis, elle callou-se como preferindo pagar quadruplicadamente, e eu disse com os meus botões: ah, **quem me dera ser guarda!** (Cartas de leitores dos estados – CapítuloIV/Pernambuco - p. 21, século XIX)

Destaque também é dado ao emprego camaleônico do verbo *poder* no esquema [V (modal) *-ra* → (mas X)], que, com base no contexto, ora pode ser interpretado com um valor deôntico ora com um valor epistêmico-avaliativo. No primeiro caso, *poder* envolve valores como desejo, capacidade (cf. (16)). No segundo, *poder* implica a avaliação de uma dada situação, na base de alguma crença ou evidência de conhecimento do falante (cf. (17)):

- (16) Amaveis leitores. (Estou falando de pince nez). Já cá estou. De partida para Bello Horizonte, deixei em meu logar o Valerio interino; volto hoje a occupar o logargulho que o amavel Redactor me concedeu pondo o meu substituto no olho da rua. Daqui estou piscando os olhos e atirando abraços. (Em Bello Horizonte aprendi a dar beijos e abraços pelo moderno systema. Um beijo para as senhoritas (**pudera!**); mil abraços para a rapaziada (Cartas de leitores dos estados – Capítulo II/ Minas Gerais - p. 40, séc XIX).
- (17) Não começo por perguntar-lhe noticia de sua saude, porque pela ultima que me escreveo fiquei sabendo que está rijo como um cerne, fresco como uma alface, e alegre como um medico em tempo de epidemia. Tambem **pudera** não ser assim. O compadre passa um vidão, mora no meio da abundancia, sente o aroma das flores, e das arvores, bebe boa e cristalina agua (...) Ora realmente felicissimo compadre, uma vidinha destas é para chegar com certeza á idade do defunto Mathuzalen, que nem eu, nem o compadre conhecemos (Cartas de leitores dos estados – CapítuloVI/São Paulo - p. 36, séc XIX).

No século XX, da variedade de contextos em que era possível emergir uma leitura [-*realis*] para *V-ra*, apenas aquelas com valor semântico desiderativo parecem ter sobrevivido. Dois esquemas, particularmente, se fixaram e se expandiram por toda a comunidade de fala do PB contemporâneo.<sup>11</sup> São eles: [V (modal) *-ra* → (mas X)] e [Quem *V-ra* X?], restritos aos verbos *querer*, *tomar*, *poder* e *dar*.<sup>12</sup> Veja os exemplos:

- (18) F: Hum. E aí você... Sim! mas aí você... deve ter planos assim...  
F: Ah!  
E: ...profissionais...  
F: Sim eu...  
E: Você acha...  
F: Eu às vezes fico até achando: “**Tomara** que eu... seja contratada da Globo!...” Aí que eu... de repente eu seja uma correspondente no exterior!... Ou então até por aqui mesmo, porque... eu gosto. Mas... **Tomare** que eu ganhe bem... Essa é minha meta: ganhar bem, muito bem. Ter dinheiro pra fazer tudo o que eu quero, né? (Censo-RJ / Amostra 2000)

11 Na análise das amostras Censo-RJ e Nurc-RJ, além dos empregos [-*realis*] de *-ra*, foram encontrados três casos do morfema com valor temporal de passado anterior a outro.

12 A não ocorrência do verbo *poder* em enunciados exclamativos com um valor epistêmico-avaliativo, por exemplo, pode ser decorrência da limitação das amostras, visto que seu emprego parece não causar estranhamento.

(19) E: O que aconteceria... se você conseguisse realizar todos os seus sonhos?

F: O que aconteceria?...

E: Ahn.

F: Ah, eu acho... se eu conseguisse... que eu acho que eu... faria me sentiria realizada, né? Ainda mais porque... Eu acho que eu... não sei se eu vou conseguir... realizar meus sonhos, todos eles! porque se for ver a cada ano que passa você... tem um sonho, né? Aí quando cê ati... assim, você... atinge a meta de um, quando cê vai ver: “Aí!... Poxa!... **quem dera** eu viajar... pra não sei aonde.” Aí quando você viajar, volta: “Poxa, mas... também **quem dera** eu ter uma casa não sei aonde!” Aí quando você tem, você tá sempre cheia de sonhos, né? Mas eu acho que eu... me sentiria muito realizada. Uma pessoa realizada!... Estaria... muito contente (Censo-RJ /Amostra 2000).

Longe de serem uma exclusividade da fala, essas construções desiderativas/optativas podem ser encontradas também na escrita formal, para expressar um desejo, uma vontade, embora seu emprego seja pouco frequente nesta modalidade:

(20) E não foi detectada morte em suínos, uma característica da Gripe Espanhola. Além de humanos, por onde passou ela atingiu porcos, gado, cães e gatos. “Não acho que esta gripe será grave como foi a Espanhola”, diz Furtado. **Tomara** que ele esteja certo (Época – 02.05.2009 – Edição nº 572 - A ameaça da gripe suína).

Os exemplos considerados até este ponto permitem discutir dois aspectos de natureza estrutural. O primeiro está relacionado, especificamente, ao esquema contrastivo [V (modal) *-ra* → (mas X)]. Em estágios anteriores, os traços [+desiderativo] e [+contraste] eram expressos explicitamente na forma de verbo volitivo e de oração adversativa. A recorrência de *tomar*, verbo tipicamente [-desiderativo], em contexto em que emergia o valor volitivo levou, ao que tudo indica, à semantização desse traço. Ressalta-se, também, que, no PB atual, a relação contrastiva pode não ser explicitada. Visto que desejo é algo que só pode ser realizado no futuro, o único contraste que pode ser estabelecido é em relação ao estado de mundo em que o desejo se impõe. Ex: “Tomara que eu... seja contratada da Globo! *Mas no atual momento eu não sou*” (grifos nossos).

O segundo aspecto está relacionado às propriedades modo-temporais do verbo da oração complemento que contribui para o emprego dessas expressões em contextos pragmáticos particulares. Assim, o predicador da oração complemento das expressões *Quisera eu* e *Quem me dera* tende a ocorrer no imperfeito do subjuntivo, como ilustra o exemplo abaixo:

(21) F: A Barra antigamente era um lugá bom de morá, tendeu? A gente andava qualqué hora na rua, tava sempre andano. Hoje em dia, por exemplo, como eu vejo até assim nos repórte assim, jornal e coisa, como já fui lá pra fazê algum serviço, aí eu vejo que não é a merma coisa, não é um lugá que a gente é reconhecido, [que]... que a gente daquela época, né? que a gente andava a cavalo, andava de tudo, agora... hoje em dia, a gente não tem nada pra fazê isso. É mais um... tá um lugá mais violento, a gente quando vê lá é... muita coisa que a gente vê lá que escuta que... não é a merma coisa, mas não é mesmo! **Quem dera** que a Barra pudesse voltá aquilo que era, eu voltá aquele poço do (inint) pra tomá aquela água gelada (Censo-RJ / Amostra 2000).

*Quisera eu* e *Quem dera* remetem a um desejo do falante, cuja possibilidade de realização fica restrita a um mundo imaginário em que as coisas se desenvolveriam de acordo com suas preferências, como um “sonho” (cf. exemplo (19)). Enunciados introduzidos por essas expressões apresentam uma interpretação similar à do mais-que-perfeito do subjuntivo em Latim, tempo que, no português “passou a ser usado como imperfeito do

mesmo modo” (COUTINHO, 1967, p. 275). Empregava-se esse tempo, em enunciados optativos, especificamente, para indicar desejo, coisa ou ação que não se espera que aconteça no presente ou se sabe não ter acontecido no passado, como ilustram os exemplos (22) e (23) (cf. RAVIZZA, 1958):

(22) **Utinam te servavissent** = prouvera a Deus que eu te tivesse salvado (mas sei que não te salvei).

(23) **Vellem redisses** = quisera que tu tivesses voltado (mas sei que tu não voltaste).

Nos enunciados introduzidos por *Tomara*, há uma forte tendência de que o predicador da oração complemento esteja no presente do subjuntivo (v. exemplo (18)). Dado que, no Latim, não se empregava o mais-que-perfeito do subjuntivo para indicar coisa ou ação que se julga possível de realizar atualmente ou no passado, a expressão *Tomara* particula- riza-se por remeter a um desejo iminente, com possibilidade de realização no mundo real, como uma meta a ser alcançada.

## Conclusão

Esta pesquisa focou, especificamente, os usos modais do morfema *-ra*. A análise desenvolvida permitiu-nos depreender os contextos de leitura [*-realis*] da construção num determinado período da história do português e sua idiomatização no estágio atual. O que nossa análise permite dizer é que a construção *V-ra* de natureza temporal (isto é, categoria cognitiva mais básica, mais objetiva), em sua trajetória no português, gradativamente, estendeu seus contextos de uso a categorias e domínios mais abstratos e subjetivos. Podemos presumir que a repetição de *V-ra* em domínios condicionais e com verbos modais afetou a representação cognitiva da construção, levando ao surgimento de expressões com implicações pragmáticas específicas. A recorrência de *V-ra* no domínio condicional, especificamente entre os séculos XVI e XVII mostra que o emprego modal de *-ra*, fadado ao desaparecimento, resiste principalmente em seu contexto mais prototípico herdado do latim.

Os dados apresentados e discutidos, neste artigo, sugerem que a construção *V-ra*, em estágios anteriores, era mais esquemática e mais produtiva, uma vez que a posição V podia ser preenchida por uma variedade maior de verbos (como *Folgara eu! Desejara eu! Quem crêra! Quem dissera! Quem pudera imaginar!* etc). Ao longo dos cinco últimos séculos, essa produtividade vai se reduzindo e determinados itens lexicais (mais especificamente, verbos de alta frequência) se conservam nessa posição, resultando em expressões formulaicas que são estocadas como unidade indissolúvel. A perda de esquematicidade levou à perda de produtividade. À medida que os falantes deixaram de reconhecer o morfema *-ra* como uma parte de um esquema mais geral [*V + ra*], concebem as sequências de *chunks* “Quem dera”, “Tomara que” e “Quisera eu” como um todo autônomo e entrincheirado. No entanto, conclusões mais definitivas acerca da trajetória do morfema *-ra* e da sua perda de produtividade exige uma análise que considere os diferentes usos dessa desinência em estágios anteriores, principalmente, no português medieval.

## REFERÊNCIAS

BECKER, Martin. From temporal to modal: divergent fates of the Latin synthetic pluperfect in Spanish and Portuguese. In: DETGES, Ulrich; WALTEREIT, Richard (Ed). *The Paradox of Grammatical Change: perspectives from romance*. Amsterdam: John Benjamins, 2008. p. 147-180.

BROCARD, Maria Teresa. Portuguese pluperfect: elements for a diachronic approach. *Estudos Linguísticos / Linguistic Studies*, Lisboa, v. 5, p. 117-130, 2010.

BYBEE, Joan. *Language, usage and cognition*. Cambridge, UK: CUP, 2010. 252 p.

COAN, Márluce. *As categorias tempo, aspecto, modalidade e referência na significação dos pretéritos mais-que-perfeito e perfeito: correlações entre função (ões)-forma(s) em tempo real e aparente*. 2003. 231 f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis.

\_\_\_\_\_. *Anterioridade a um ponto de referência passado: pretérito (mais-que) perfeito*. 1997. 177 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis.

COMRIE, Bernard. *Tense*. 8. ed. London: Cambridge University Press, 2004. [1985]. 139 p.

COUTINHO, Ismael de Lima. *Pontos de gramática histórica*. 6. ed. Rio de Janeiro: Livraria Acadêmica, 1967. 357 p.

GARCIA, Othon Moacyr. *Comunicação em prosa moderna: aprenda a escrever, aprendendo a pensar*. 27. ed. Rio de Janeiro: Editora FGV, 2010 [1967]. 548 p.

LYONS, John. *Semantics*. v. 2. London: Cambridge University Press, 1977. 897 p.

MARTINS, Kellen Cozine. A variação entre o pretérito mais-que-perfeito simples e composto em textos jornalísticos. *Diadorim*, Rio de Janeiro, v. 8, p. 15-30, 2011.

\_\_\_\_\_. *A expressão variável de anterioridade a um ponto de referência passado na escrita midiática*. 2010. 131 f. Dissertação (Mestrado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro.

MATTOS E SILVA, Rosa Virgínia. *O português arcaico: morfologia e sintaxe*. 2. ed. São Paulo: Contexto, 2001.

MATTOSO CAMARA JR., Joaquim. *Estrutura da língua portuguesa*. 14. ed. Petrópolis: Vozes, 1984. 125 p.

MAURER JR., Theodoro Henrique. *Gramática do latim vulgar*. Rio de Janeiro: Livraria Acadêmica, 1959. 298 p.

NEVES, Maria Helena de Moura. As construções concessivas. In: NEVES, Maria Helena de Moura (Org.). *Gramática do português falado*. 2. ed. São Paulo: Humanitas/FFLCH/USP; Campinas: Editora da Unicamp, v. 7, 1999. p. 545-591.

RAVIZZA, P. João. *Gramática latina*. 14. ed. Niterói: Escola Industrial Dom Bosco, 1958. 560 p.

SAID ALI, Manuel. *Gramática histórica da língua portuguesa*. 3. ed. Rio de Janeiro: Livraria Acadêmica, 1964. 375 p.

THIBAUT, André. *Formes synthétiques et analytiques de praeteritum das la Romania*. In: CONGRÈS INTERNATIONAL DE LINGUISTIQUE ET PHILOLOGIE ROMANES, XX, 1992, Zurique. *Actes...* tome 3. Tübingen: A. Francke Verlag 1993. p. 397-410.

TRAUGOTT, Elizabeth C.; DASHER, Richard B. *Regularity in semantic change*. Cambridge: Cambridge University Press, 2002. 341 p.

TRAUGOTT, Elizabeth C. Grammaticalization and mechanisms of change. In: HEIKO, Narrog; BERND, Heine (Ed.). *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press, 2011. p.19-30.

\_\_\_\_\_. (Inter) subjectivity and (inter) subjectication: a reassessment. In: DAVIDSE, Kristin et alli (Org.). *Subjetification, intersubjetification and grammaticalization*. Berlin/ NewYork: Mouton de Gruyter, 2010. p. 29-74.

# Linguística 0800: uma ação de divulgação científica

(Linguistics toll-free telephone call: an action of science popularization)

Ricardo Joseh Lima<sup>1</sup>

<sup>1</sup> Departamento de Estudos da Linguagem – Universidade do Estado do Rio de Janeiro (UERJ)

rjlimauerj@gmail.com

**Abstract:** This paper aims to present a proposal of construction of actions of science popularization in the realm of Sociolinguistics. The proposal is based on the principles postulated by Labov (1982 apud CATELLAN, 1996) and Wolfram (1993) concerning the role of the linguist in communicating research findings to society. Criticisms raised by Rajagopalan (2003) to such a task are approached and some solutions are presented in order to deal with the problems that the author points out. To this end, we argue that the merging of three ideas may serve as the foundations for solid actions of science popularization: the concept that Teaching, Research and Extension should be viewed as inseparable; the usage of medias as internet; and the theoretical bases of action-research.

**Keywords:** sociolinguistics; science popularization; action research.

**Resumo:** Este artigo busca apresentar uma proposta de construção de ações de divulgação científica no âmbito da Sociolinguística. A proposta está baseada nos princípios de Labov (1982 apud CATELLAN, 1996) e Wolfram (1993), que versam sobre o papel do linguista em relação à comunicação com a sociedade. Problemas levantados por Rajagopalan (2003) a essa tarefa são abordados e algumas soluções são apresentadas na tentativa de dar conta desses problemas. Ao fim, argumentamos que a união entre o conceito de indissociabilidade entre Ensino, Pesquisa e Extensão, a utilização de novos meios de interação com a sociedade, e as bases da pesquisa-ação podem servir de fundamentos para ações de divulgação bem-sucedidas.

**Palavras-chave:** sociolinguística; divulgação científica; pesquisa-ação.

## Introdução

A preocupação de linguistas em criar meios de comunicação com a sociedade a respeito de suas pesquisas e descobertas não é recente. Desde Saussure, pelo menos, podemos traçar um histórico dessa preocupação. O surgimento da Sociolinguística, na década de sessenta do século passado, trouxe novos ingredientes para a questão da comunicação entre linguistas e a sociedade. Afinal, pela primeira vez, havia evidências empíricas robustas para derrubar mitos como o de que alguns dialetos são inferiores a outros.

Alguns esforços foram realizados na tentativa de estabelecer canais de comunicação com a sociedade e é possível ver em Labov (1982<sup>1</sup>) uma proposta de sistematização dessa tarefa do linguista. Wolfram (1993) retoma a proposta de Labov (1982), acrescentando um princípio de ação que possui como características diferenciadoras seu aspecto positivo e proativo. Denominado “Gratuidade Linguística”, esse princípio serviria para

<sup>1</sup> apud CATELLAN, J. O gênero discursivo religioso católico: uma materialização discursiva previsível. 1996. 230 f. (Mestrado em Linguística) – Universidade Federal do Paraná, Curitiba. Disponível em: <http://dspace.c3sl.ufpr.br/dspace/bitstream/handle/1884/24416/D%20-D20CATELLAN,%20JOAO%20CARLOS.pdf?sequence=1..> Acesso em: 28 ago.2012.

orientar trabalhos de linguistas que ofereciam retorno de suas pesquisas às comunidades de onde eles obtiveram seus dados.

Rajagopalan (2003) apresenta algumas críticas ao projeto de sociolinguistas de estabelecerem um contato com a sociedade para disseminar suas descobertas e, por conseguinte, contribuir para auxiliar a comunidade em situações de seu interesse. Essas críticas podem ser resumidas no tipo de pesquisa que é feito pela sociolinguística variacionista, no método de abordagem no contato com a comunidade e nas questões éticas envolvidas nesse contato. Embora reconheçamos o valor dessas críticas, é nossa intenção nesse artigo apresentar soluções que evitem os problemas levantados por Rajagopalan (2003), mantendo, desse modo, ainda válidas as propostas enunciadas por Labov (1982) e Wolfram (1993).

As soluções a serem apresentadas têm como base três ideias distintas, mas que, unidas, podem fornecer saídas às críticas de Rajagopalan (2003). O princípio da indissociabilidade entre Ensino, Pesquisa e Extensão pode servir como esteio para refutar a crítica realizada ao tipo de pesquisa feito pela sociolinguística variacionista. A utilização de meios alternativos aos tradicionais livros, artigos e dissertações, tais como vídeos e internet respondem às questões sobre o contato com a comunidade. Por fim, as bases teóricas da pesquisa-ação trazem discussões a respeito da ética na pesquisa com relação a pesquisador e pesquisado que podem contribuir para a solução do problema ético levantado por Rajagopalan (2003).

Ao fim, argumentaremos que estamos de posse de um cenário em que ações de divulgação científica no âmbito da Sociolinguística contariam com bases de diversas fontes para sua realização bem-sucedida. Desse modo, a preocupação com a comunicação entre linguistas e sociedade, anunciada no início dessa introdução, tende a ser diminuída, uma vez que novos caminhos são apresentados para serem postos à prova e para discussões futuras.

Neste artigo, a preocupação lançada por Saussure (1969) e abordada por Labov (1982) e Wolfram (1993) será o foco da próxima seção. A seção seguinte traz as críticas de Rajagopalan (2003) a essa abordagem. Na terceira seção, apresentamos as três propostas de respostas a essas críticas. A seção que encerra este artigo conta com considerações a respeito da comunicação entre linguistas e sociedade e ilustra essas considerações com exemplos concretos de ações de divulgação científica.

## **Fazer Linguística é comunicar-se**

Embora haja divergências quanto ao papel e ao destaque que se deve dar ao trabalho de Saussure e sua relação com a história da Linguística, é inegável que há uma corrente que o considera como fundador de um modo de pensar e fazer Linguística que prima pelo rigor científico de suas bases teóricas e metodológicas e de seus objetivos. Tomando emprestado por alguns momentos esse ponto de vista que valoriza o trabalho de Saussure, que alguns chegam ao ponto de considerar o “pai da Linguística”, podemos, também, por extensão, considerá-lo como o fundador da preocupação de levar os conhecimentos obtidos por essa disciplina à sociedade em geral. Essa preocupação está explícita logo nas primeiras páginas do *Curso*, como se pode perceber na citação abaixo:

Qual é, enfim, a utilidade da Linguística? [...] não há domínio onde tenha germinado ideias tão absurdas, preconceitos, miragens, ficções. Do ponto de vista psicológico, esses erros não são desprezíveis; a tarefa do linguista, porém, é, antes de tudo denunciá-los e dissipá-los tão completamente quanto possível. (SAUSSURE, 1969, p. 14)

Apesar dos seus quase cem anos, essa citação permanece atual, tendo em vista que ainda hoje há um desconhecimento profundo, por parte da sociedade, do trabalho do linguista e dos resultados de suas pesquisas. A “polêmica do livro didático”, estabelecida nos meses de maio e junho de 2011 a respeito de um livro didático que “ensinaria a falar errado”<sup>2</sup> é um dos exemplos mais marcantes desse desconhecimento. Saussure pouco pôde fazer para remediar esse tipo de situação e cumprir a tarefa do linguista que ele próprio estipulou, uma vez que se ocupou em construir o edifício teórico que serviria de base para o Estruturalismo europeu e teve seus trabalhos interrompidos por sua morte.

A preocupação de Saussure com a comunicação com a sociedade veio se refletir em estudos posteriores de vários linguistas, dentre os quais destacamos o de Labov, que formalizou, em trabalho de 1982, dois princípios que deveriam guiar a relação entre o linguista e a sociedade, a saber, o **principle of error correction** e o **principle of debt incurred**. Recorremos a Catellan (1996) para uma tradução não apenas desses princípios como de suas definições. Respectivamente, temos o **princípio da correção do erro** e o **princípio do débito contraído**. Suas definições são, respectivamente (CATELLAN, 1996, p. 172): “Um cientista que se dá conta de que um ideia corrente ou prática social com importantes consequências é invalidada por dados que tem à disposição é obrigado a trazer estes dados para apreciação da maior audiência pública possível” e “Um investigador que obteve dados linguísticos junto aos membros de determinada comunidade de fala tem a obrigação de tornar aqueles dados acessíveis à comunidade, quando ela precisar disso”.

Esses princípios foram formulados dentro de um contexto específico, qual seja, a discussão do caso ocorrido em Ann Arbor, nos anos sessenta, que culminou no julgamento da situação que envolvia a existência ou não de um Inglês Vernacular Negro. Dado que linguistas possuíam a informação a respeito dessa existência, de acordo com Labov (1982), os mesmos deveriam comunicá-la à comunidade, auxiliando-a no processo. Podemos notar que a participação de linguistas no julgamento obedece aos dois princípios enunciados por Labov (1982): há uma ideia incompatível com os dados dos linguistas (a inexistência do Inglês Vernacular Negro), e, portanto, cabe aos linguistas informar a comunidade desse erro; a posse da informação e sua consequente comunicação são úteis à comunidade, que está precisando delas para seu benefício. Labov (1982), por assim dizer, propõe um *modus operandi* da atuação do linguista em relação às comunidades de fala: desmistificar ideias errôneas, comunicando aquilo que foi encontrado em suas pesquisas e contribuir para situações concretas toda vez que a comunidade precisar.

Labov (1982) toma o cuidado de estabelecer três princípios de objetividade para que o envolvimento do linguista com determinada comunidade não seja tomado de modo subjetivo, o que prejudicaria sua argumentação a respeito do valor dos seus dados frente às ideias incorretas que com ele entram em conflito. O grau de controle sobre os dados, o tempo entre a observação e o registro dos dados e semelhança nas fontes de dados são parâmetros científicos, portanto objetivos e neutros, que podem balizar a argumentação

2 Trata-se da obra *Por uma vida melhor*, de Heloisa Ramos, voltado para EJA (Ensino de Jovens e Adultos).

do linguista. Dados obtidos de forma aleatória, sem ter passado pelo crivo do tempo e obtidos a partir de fontes heterogêneas tendem a não ser creditados. Dados com as características inversas levam a uma robustez maior nas conclusões teóricas empreendidas pelo investigador.

Wolfram (1993) retoma os princípios de relação do linguista com a sociedade propostos por Labov (1982) para argumentar que deve ser acrescentado mais um, que ele nomeou como **princípio da gratuidade linguística**: “Investigadores que obtiveram dados linguísticos a partir de membros de uma comunidade deveriam procurar ativamente modos positivos nos quais eles podem retornar favores linguísticos à comunidade” (traduções nossas). De que modo esse novo princípio se relaciona com os anteriores e em que medida os complementa e estende? Seguiremos a argumentação de Wolfram (1993) para sua defesa, que nota, em primeiro lugar, que não se trata de uma relação de exclusão entre os princípios de Labov (1982) e o por ele proposto. Pelo contrário, trata-se de uma relação de complementação, uma vez que a situação à qual se aplica o novo princípio é diferente das expostas por Labov (1982). Podemos diferenciar os princípios de Labov (1982) e o de Wolfram (1993) por um critério de extensão de atuação: enquanto os princípios labovianos seriam de atuação pontual, o princípio de Wolfram seria de atuação contínua. Assim é que no princípio da correção do erro, deve haver a identificação de uma ideia errônea e sua conseqüente alteração; no princípio do débito contraído, o linguista atua de acordo com as necessidades da comunidade. Ora, se um linguista não perceber nenhuma ideia em conflito com seus dados e se não houver nenhuma situação concreta de necessidade da comunidade, os princípios de Labov (1982) não entrariam em ação. O que Wolfram (1993) propõe é que a ação dos linguistas frente à comunidade aconteça de modo contínuo, independente de haver percepção de alguma ideia errônea ou de uma necessidade premente. No caso do princípio da gratuidade linguística, haveria uma intenção do linguista de, independentemente de situações percebidas ou que surgissem, organizar sua relação com a comunidade de modo colaborativo, integrado e contínuo.

O **princípio da gratuidade linguística** é ilustrado no artigo de 1993 com os programas de conscientização sobre língua (“language awareness programs”, no original). Esse tipo de programa torna o contato entre linguistas e comunidade mais frequente e prolongado. Por ser um programa, pode contar com uma estrutura que envolve equipes, objetivos previamente traçados, um cronograma, entre outras características. Dois programas são descritos: um que versa sobre o Inglês Vernacular Negro e outro sobre o dialeto de uma comunidade da Carolina do Norte. Reaser e Adger (2007) relatam outras ações de linguistas que podem ser descritas como atendendo ao princípio da gratuidade linguística, com destaque para o documentário “Do you speak american?”, com sua conseqüente aplicação em escolas de ensino médio e criação de um site para abrigar informações a seu respeito (<http://www.pbs.org/speak>). Charity (2008) relaciona seu trabalho com estudantes universitários, e sua inserção em ambientes escolares de ensino fundamental para observar e intervir em situações que envolvem o Inglês Vernacular Negro, com o princípio proposto por Wolfram (1993). Essas ações de diversos linguistas têm em comum, como exposto acima, a característica de partirem dos próprios linguistas sem esperar uma situação que os motivasse a se comunicar com a comunidade. Além disso, não são ações pontuais, mas constantes, dado o caráter programático de cada uma delas.

Assim como Labov (1982), Wolfram (1993) apresenta as preocupações éticas com esse tipo de atividade. No entanto, não o faz de modo formal, como nos princípios

de objetividade, mas, por outro lado, utiliza situações reais, que aconteceram nos dois programas descritos, para levantar problemas e questionamentos sobre o alcance do trabalho dos linguistas. Incentivar a comunidade a participar demonstrando a importância desse tipo de programa é um desses trabalhos. Isso deve ser feito com limites, pois o linguista não pode exagerar na argumentação sobre seu trabalho sob pena de estar atendendo mais a seus próprios interesses do que os da comunidade. Ouvir a voz da comunidade e assimilá-la em seu trabalho é outro desafio do linguista, uma vez que essa voz pode ser dissonante em relação a suas ideias, provenientes de seus dados.

Apesar dos esforços de Labov (1982), Wolfram (1993) e outros linguistas que seguem essa linha de raciocínio, esse tipo de trabalho, que visa a concretizar o exposto por Saussure, não tem passado incólume a críticas provenientes da própria academia. Rajagopalan (2003), baseado no trabalho de Cameron et al. (1993 apud RAJAGOPALAN, 2003), apresenta algumas críticas que parecem colocar em xeque o *modus operandi* proposto pelos sociolinguistas variacionistas. Na próxima seção, faremos um levantamento dessas críticas.

### **Armadilhas da comunicação...**

Rajagopalan (2003) está interessado em discutir relações entre linguagem e ética, mais especificamente, como linguistas, que lidam diretamente com a linguagem, se posicionam frente a questões éticas envolvidas quando se prestam a colocar seus trabalhos para a comunidade dita não-linguista.

Rajagopalan (2003) se baseia na análise que Cameron et al. (1993 apud RAJAGOPALAN, 2003) fazem do trabalho já citado de Labov (1982), submetendo-o a uma análise crítica que se centra no tipo de pesquisa empreendida por linguistas como Labov (1982), a sociolinguística variacionista. Os autores afirmam que no julgamento Labov (1982) estava defendendo os próprios interesses, já que apresentava uma versão do debate que ia ao encontro dos trabalhos por ele realizados. Com isso, faltava ao sociolinguista um contato mais intenso com a comunidade, que ficou sem voz, uma vez que o linguista a representava. Essa representação seria uma ilusão pois foi decidida pelo linguista e o que ele apresenta não foi discutido pela comunidade. Tal situação é gerada a partir da metodologia de pesquisa variacionista, classificada como “positivista”<sup>3</sup>. Dado que há uma divisão a priori entre pesquisador e pesquisado e somente o primeiro detém o poder na pesquisa, o distanciamento entre sujeito e objeto seria uma consequência natural.

Outra crítica apresentada menciona especificamente o modo pelo qual se deu a abordagem com a comunidade no caso relatado por Labov (1982). Seguindo-se os princípios propostos por esse linguista, foi através da percepção do linguista de que ele poderia ajudar que o contato se deu. Além disso, como observam Cameron *et al.* (apud RAJAGOPALAN, 2003), o linguista não integra a comunidade em questão, é um membro externo. Sendo assim, possui interesses diversos, que podem não ser representativos da comunidade como um todo. Labov (1982) emprestou a sua voz somente a uma parte da comunidade, voz que, uma vez mais se observa, atende a seus interesses. Essa falta de representatividade da comunidade como um coletivo, que debate e decide seus inte-

<sup>3</sup> Tal classificação se refere a uma posição epistemológica que define a separação entre pesquisador e objeto de pesquisa como sendo a base para a produção do conhecimento, o que se opõe a abordagens que consideram que o conhecimento provém das relações entre os envolvidos em uma pesquisa.

resses, não estaria comprometendo o papel da própria comunidade? Podemos pensar em um exercício em que, após o contato do linguista com a comunidade, a mesma rejeite algumas das conclusões de suas pesquisas. Como deveria o linguista agir? Recuaría de sua posição, baseada em descobertas científicas, para dar voz à comunidade, ou silenciaria essa voz, privilegiando aquela que advém de um conhecimento especializado? São questões que são lançadas por Cameron et al. (apud RAJAGOPALAN, 2003) e retomadas por Rajagopalan a cerca desse tipo de comunicação entre linguistas e comunidades. Essas questões levam ao terceiro ponto levantado por Rajagopalan (2003).

O tema da ética é frisado com ênfase na avaliação do trabalho realizado por Labov. O primeiro ponto faz menção ao segundo princípio laboviano, que trata do débito contraído. Seria uma espécie de dever do linguista quitar esse débito com a comunidade, auxiliando-a quando ela precisasse. O que Rajagopalan (2003) questiona é que essa percepção de que há um débito parte do linguista e o leva a uma possível posição assistencialista, de prestador de serviços, pois o mesmo se sente na obrigação de auxiliar a comunidade. Ora, pode-se argumentar que qualquer contribuição que o linguista venha a fornecer para a comunidade deveria ser encarada como natural e espontânea, parte integrante do seu próprio trabalho; do modo que está exposto no segundo princípio de Labov, seria somente através dessa percepção que o linguista acordaria para a necessidade de que seu trabalho seja útil à comunidade. O segundo ponto diz respeito à posição que o linguista assume nesse contato, como sendo portador de um tipo de conhecimento que se diz ser superior ao conhecimento da comunidade. Uma vez que seu trabalho possui bases científicas, é estruturado por pesquisas em um ambiente especializado, o conhecimento adquirido a partir desse cenário tomaria ares de superioridade frente ao conhecimento da comunidade, ao qual se pode referir como senso comum ou leigo ou não-especializado. Rajagopalan (2003) alerta para o perigo de isso levar a uma imposição por parte do linguista do seu conhecimento em detrimento do conhecimento fomentado pela comunidade.

No caso relatado, parece claro para Cameron *et al.* (apud RAJAGOPALAN, 2003) e para Rajagopalan (2003), que essa questão ética não foi observada por Labov. Isso porque o linguista apresentou uma contribuição que defendia seus próprios interesses e podia não refletir os interesses da comunidade como um todo, que não teria sido consultada. Como consequência, o linguista assumiria um papel de detentor do saber, colocando-se em uma posição assimétrica frente à comunidade. Tomado em conjunto, esse cenário colocaria em xeque uma série de questões éticas, tornando problemático, para dizer o mínimo, o tipo de comunicação proposto por Labov. Não é outro se não esse tom crítico que Rajagopalan (2003) adota na análise do que Labov propõe.

As três críticas acima levantadas seriam passíveis de uma réplica consistente, que as respondesse, ponto a ponto? Essas críticas seriam extensivas ao princípio proposto por Wolfram (1993), e em caso positivo, também haveria como abordá-las de modo direto, apresentando alternativas que sejam consideradas como suficientes para sustentá-lo? Esse é o desafio que será abordado na próxima seção.

## **Organizar para comunicar**

Conforme apresentado na Introdução, resumimos as críticas de Rajagopalan em três pontos: o tipo de pesquisa realizado pela sociolinguística variacionista; o modo pelo qual aconteceu o contato com a comunidade; e nas questões éticas envolvidas no trabalho

do linguista. Nesta seção, abordaremos cada crítica em particular, utilizando de argumentos teóricos e analíticos. O procedimento a ser adotado para cada abordagem será o mesmo: iniciamos com a exposição de um conteúdo que, embora não faça menção direta à crítica em questão, acreditamos ser relevante para sua réplica para em seguida realizar uma análise desse conteúdo frente à crítica propriamente dita. O primeiro ponto a ser abordado é a respeito do tipo de pesquisa realizado pela sociolinguística variacionista, classificado como “positivista”.

O isolamento do pesquisador em relação a questões sociais tem sido uma crítica frequente ao modo de se fazer pesquisas na Universidade nos dias de hoje. Apesar de ser uma realidade palpável e inegável, essa situação não está prevista por aquilo que deve reger a vida universitária, a saber, instruções normativas firmadas em documentos oficiais. Sendo assim, recorreremos à Constituição Federal (BRASIL, 1988), em seu artigo 207, para estabelecer um primeiro contato com o tema da pesquisa na Universidade: “As universidades [...] obedecerão ao princípio de indissociabilidade entre ensino, pesquisa e extensão.” A palavra que fornecerá a chave para a compreensão da resposta à primeira crítica de Rajagopalan (2003), sem dúvida, é “indissociabilidade”. Diversos autores, dentre os quais podemos destacar Dias (2009) e Tauchen (2009), trazem importantes discussões a respeito de como esse princípio da indissociabilidade pode ser obedecido através da prática universitária. Parece ser consensual, primeiramente, que se trata de um desafio extremamente complexo de ser resolvida dada a própria relação entre os níveis e as condições de sua concretização.

Essa complexidade talvez explique porque a pesquisa se tornou isolada dos demais níveis, levando a uma postura de distanciamento em relação ao objeto de pesquisa e de superioridade no conhecimento por ela obtido. Entretanto, tal situação de isolamento é nociva à vida universitária, algo que o artigo 207 externaliza ao enunciar que o princípio da indissociabilidade deve ser obedecido. Está colocado, portanto, o desafio de realizar uma atividade que, ao mesmo tempo, se relacione com duas outras. Como apontado no parágrafo anterior, não apenas isso aumenta a carga de trabalho como exige habilidades múltiplas do docente específicas dos três níveis em questão, ensino, pesquisa e extensão. Não apenas não há formação voltada para o desenvolvimento dessas habilidades como o próprio ambiente universitário se viciou em torná-las independentes causando dificuldades para aqueles que anseiam tornar suas atividades conectadas entre os três níveis.

Apesar desse cenário adverso, os estudos realizados sobre o tema da indissociabilidade também parecem ser unânimes quanto à extrema urgência de que esse princípio seja obedecido o quanto antes no fazer universitário. Assume-se que somente essa integração entre os níveis pode resolver os problemas inerentes a cada um deles. Em outras palavras, por mais paradoxal que possa parecer, os problemas enfrentados pela atividade do pesquisador podem ser resolvidos quando se leva em conta as atividades de ensino e de extensão. Nesse ponto, há propostas específicas como a levantada no Plano Nacional de Extensão Universitária (BRASIL, 2001), de que a atividade de pesquisa deveria ser levada a cabo pensando em temas da educação básica (e daí o ponto de contato com o ensino), de modo a contribuir para a solução de problemas específicos dessa área (levando a um contato com a extensão). Tal proposta teria como efeito imediato repensar todo o quadro que justifica uma pesquisa e o embasamento para sustentar sua existência.

Nada disso parece acontecer de modo ingênuo e torna-se evidente que o aspecto social está ganhando relevo na consideração do princípio da indissociabilidade e sua concretização. Isso acontece porque se privilegia o nível extensionista como sendo o balizador dos demais níveis: pesquisa e ensino devem ter uma contrapartida social inerente, de tal modo que pensar a separação é que deveria ser estranho e complexo e não a indissociabilidade, como é a prática corrente. Desse modo, a percepção de que a pesquisa deve englobar de forma natural e caracterizadora contribuições para a solução de problemas de ensino em situações concretas pode reverter o quadro desfavorável que é construído pelos críticos da postura epistemológica da sociolinguística variacionista.

O segundo ponto abordado por Rajagopalan (2003), com base em Cameron et al. (1993 RAJAGOPALAN, 2003), a respeito dos princípios propostos por Labov diz respeito ao modo como se deu o contato com a comunidade. De acordo com os críticos da postura laboviana, esse contato se deu numa via de mão única, pois partia do pesquisador para a comunidade, e de modo parcial, uma vez que apenas parte da comunidade foi abordada e consultada, não havendo uma aproximação mais profunda com essa parte e nem uma que tivesse atingido uma parcela maior dessa comunidade. Nesse ponto, passamos a apresentar alternativas que possam ser vistas como capazes de solucionar os problemas levantados nesse parágrafo.

A preocupação levantada por Rajagopalan (2003) é vista em outros contextos em que se estimula o trabalho universitário a construir contratos com a sociedade. O já referido Plano Nacional de Extensão Universitária (BRASIL, 2001) menciona várias vezes que é imprescindível haver esses contratos. Entretanto, poucas vezes deixa claro como esses contratos devem ser realizados. Em algumas passagens, são citados os movimentos sociais e a sociedade organizada como portadores da voz com a qual os docentes universitários devem dialogar. Ainda que reconheçamos a legalidade dessas instituições, nos questionamos sobre sua representatividade, tendo em vista o baixo grau de interesse das pessoas em participar ou se informar sobre essas instituições, as pressões que são exercidas por interesses alheios aos formadores de uma sociedade (como em associações de moradores, por exemplo) e as ligações muitas vezes evidentes com partidos políticos, o que compromete a isenção desses grupos. A partir daí, imaginar que um pesquisador ou mesmo um grupo pode realizar a tarefa de estabelecer um contato com uma comunidade inteira, ou parte significativa, pode ser apenas um ideal, mas longe de ser concretizado.

Como, então, seria possível resolver esse dilema? Uma proposta razoável parece ser seguir o caminho daqueles instrumentos que conseguem atingir parcelas significativas da população: os meios de comunicação de massa. No Brasil, esses são representados pelos canais abertos de televisão. Seria possível atingi-los? Antes de mais nada, precisamos confirmar se desejamos isso. Pois tal não é a postura de Baronas (2010), que critica revistas de divulgação como a “Língua Portuguesa”, vendida em bancas de jornais e disponível, em parte, na internet. Sua crítica se centra na simplificação da linguagem e das argumentações disponibilizadas. Desse modo, propõe que a divulgação científica de linguística não deveria seguir esse caminho. Tal não é a proposta aqui, onde justamente se argumenta pela massificação da informação, com o que nos apoiamos em Nicoletis (2010), sendo esse o primeiro ponto que ele destaca na necessidade de divulgar a ciência no Brasil. Um exemplo clássico de como informações linguísticas foram parar nas residências de milhões de telespectadores foram os comerciais, veiculados na década

de noventa do século passado, por uma rede de *fast food*, em que se explicava por que as crianças falam “eu sabo”: por analogia com formas regulares (“beber” – “eu bebo”); ainda se completava a explicação a informação, linguisticamente correta, de que a criança aprenderia a forma irregular sem esforço. Note-se que tal comercial está hoje disponível na internet, sendo ainda possível sua circulação e divulgação.<sup>4</sup>

Conscientes das dificuldades de inserção direta em um canal de televisão aberta, podemos sugerir o caminho inverso ao que aconteceu com o comercial acima: inserir o material em um meio de comunicação como a internet para posteriormente tornar possível sua inserção em um meio mais amplo. Tal rota, da internet para a televisão, acontece com frequência, o que pode se notar em programas humorísticos, novelas, etc. Com isso, a internet, por si só, conta com um poder de circulação muito maior do que os números concretos revelam, uma vez que um vídeo pode ser assistido por um número *x* de pessoas, mas comentado por um número mais expressivo ainda. Nossa proposta insiste em apostar nesse caminho, que está sendo muito pouco explorado pelos linguistas. A aposta se confirma uma vez que na polêmica do livro didático, ocorrida nos meses de maio e junho de 2011, parcela da sociedade demonstrou extremo desconhecimento de décadas de trabalho linguístico na universidade. Talvez nossa falta de habilidade ou (e?) interesse em divulgar nossos trabalhos explique esse fato.

Note-se que acima utilizamos a expressão “parcela da sociedade”, embora muitos linguistas que tenham se pronunciado, ainda que de modo informal, tenham utilizado o termo “sociedade”. Escolhemos isso porque muitas das demonstrações de desconhecimento vieram da mídia impressa e de comentários de leitores em fóruns na internet. Como saber se a parcela da sociedade não representada pela mídia impressa e por esses leitores corrobora essa visão de desconhecimento? Ao que parece, é uma resposta difícil (se não impossível) de se obter. Estando, portanto, todos nós no “escuro” frente ao que signifique “estabelecer contato com a comunidade/sociedade”, nossa proposta, ao focalizar a divulgação em massa, pretende atingir o maior número possível de pessoas, numa tentativa de responder a esse desafio do contrato, sem, no entanto, dar a certeza de que se poderá resolvê-lo.

A terceira crítica levantada por Rajagopalan (2003) ressalta as dificuldades éticas do trabalho do linguista quando em contato com a sociedade. Deve esse contato ser intrínseco ao seu trabalho? Em caso positivo, que postura adotar frente a situações como quando há conflito de interesses entre pesquisador e sujeitos envolvidos na pesquisa? Como o pesquisador deve colocar seus interesses em jogo em uma relação que não seja assimétrica? Ao que parece, são questões muito profundas que a abordagem laboviana não teria como dar conta, já que apenas anuncia que contatos devem ser estabelecidos mas não se detém nessas questões. É certo que Labov (1982) apresenta preocupações com a questão ética, dados seus princípios de objetividade acima descritos. Entretanto, trata-se de outro tipo de questão, que, no caso de Labov (1982), focaliza apenas a objetividade do pesquisador, para que ele não caia na armadilha da subjetividade. Ora, esse é apenas um dos pontos levantados por Rajagopalan (2003) e, ainda assim, não parece ser o mais crucial.

As expressões utilizadas por Labov (1982) (“correção do erro” e “débito contrário”) são, para Rajagopalan (2003), reveladoras de uma postura do pesquisador frente

---

<sup>4</sup> Esse comercial está acessível em <<http://www.youtube.com/watch?v=lz5u3i-FJTM>>. Acesso em: 27 ago.2012.

aos sujeitos de sua pesquisa. A ação do pesquisador viria de uma pressão causada por seu trabalho ter sido possível somente com a colaboração de uma comunidade e daí a necessidade de fornecer algum retorno a essa comunidade. Pode-se dizer que essa percepção já revela um afastamento do pesquisador com relação aos sujeitos da pesquisa, posto que ele deve se sentir obrigado a contribuir. Para Rajagopalan (2003), esse sentimento deveria ser inerente ao trabalho do linguista e sua contribuição sendo tão natural e intrínseca a seu trabalho que esse termo “contribuição” deveria perder sua razão de ser, visto que tudo o que deve ser feito pelo linguista deve ser feito com essa intenção, sem precisar ser motivado por alguma pressão. Labov (1982) também não deixa claro como deveria o pesquisador se comportar em caso de conflito entre sua voz e a da comunidade. Rajagopalan (2003) cita o caso da Lei dos estrangeirismos, proposta no final do século passado no Brasil<sup>5</sup>. Enquanto os representantes do povo e parcelas da sociedade apoiavam o projeto que visava a punir a utilização exacerbada de palavras que não fossem de origem portuguesa, os linguistas se posicionavam diametralmente opostos, apontando que o projeto não faria sentido e que não haveria prejuízo algum com esse uso exacerbado. Criou-se, então, uma situação de conflito entre as vozes em jogo, que os linguistas, na época, foram incapazes de solucionar. Na polêmica do livro didático, em 2011, algumas posições de linguistas foram incisivas ao classificar a mídia e muitas manifestações contrárias ao livro em debate como fruto de ignorância, má fé, entre outros termos pouco acadêmicos. O que o uso desses termos revelaria, seguindo a lógica de Rajagopalan (2003), seria uma assimetria entre a voz da academia, superior por ser especializada, e a voz que não é da academia, inferior considerada como senso comum. Os princípios de objetividade levantados por Labov (1982) apenas serviriam para corroborar essa assimetria, pois garantiriam à voz do linguista um peso que a voz da comunidade não teria.

Um tipo de pesquisa que tornasse fluida a relação entre pesquisador e pesquisado, em que propositalmente se emprestasse voz ao sujeito da pesquisa, tornando-o também um pesquisador, participante ativo nos rumos da pesquisa, é proposto pelo que é conhecido como Pesquisa-Ação (TRIPP, 2005). Ao eleger um problema específico, localizado, como alvo de uma pesquisa, a Pesquisa-Ação é capaz de travar um contato mais próximo com o que seria o objeto de sua pesquisa, uma turma de 6º ano do ensino fundamental, por exemplo. Desse modo, a aproximação entre pesquisador e “pesquisado” se torna facilitada, o que leva também a uma maior possibilidade de que o “pesquisado” possa interferir na pesquisa, algo que é desejável de acordo com os princípios da Pesquisa-Ação. Assim, a “contribuição” do pesquisador não se daria de modo externo, mas sim intrínseco a seu trabalho, como aponta Rajagopalan (2003). A transformação de pesquisado em pesquisador pode resolver o problema do conflito entre as vozes, se houver uma interação entre os dois lados a tal ponto que ambos possam ter seu posicionamento respeitado. Como base para isso, está a ideologia de que o pesquisador apenas porta informações a respeito de como conduzir uma pesquisa, quais são os passos a serem percorridos, os problemas a serem enfrentados. Já é pressuposto para ele que ele não domina um conhecimento superior, que vai ser aplicado ou testado em uma situação, mas um conhecimento que vai ser transformado por essa situação, alimentado pelo conhecimento dos outrora pesquisados, agora pesquisadores também.

Que efeitos a adoção de aspectos metodológicos da pesquisa-ação teriam na prática de pesquisa sociolinguística? Podemos citar dois. O primeiro seria uma guinada na

---

5 Trata-se do Projeto de Lei n. 1676/2001, de autoria do então senador Aldo Rebelo (PC do B – SP).

visão a respeito do resultado que se procura obter com a pesquisa. Enquanto uma pesquisa sociolinguística tradicional, por exemplo, pretende responder questões teóricas, tais como “qual variante do /r/ é utilizada em determinada localidade?”, uma pesquisa sociolinguística que incorporasse o conceito de “ação” poderia mudar sua questão básica para “a partir do uso de uma variante do /r/ em determinada localidade, como seus falantes são considerados a partir do ponto de vista da utilização da variante de prestígio?”. A resposta a essa questão contribuiria de modo direto para as relações sociais estabelecidas entre esses falantes e os falantes das variantes de prestígio. Outro feito poderia ser no método de obtenção dos dados de uma comunidade. É uma prática comum na pesquisa sociolinguística tradicional a realização de entrevistas. No entanto, nesse tipo de obtenção de dados, o papel do pesquisador ainda é realçado, distanciando-o do pesquisado, o que abre espaço para possíveis artificialidades na fala do entrevistado. A inserção de modo natural e espontâneo do pesquisador na comunidade poderia reduzir essas artificialidades, tornando-o mais próximo de encontrar dados que reflitam o uso real de determinada variante linguística.

Argumentamos, nesta seção, que a confluência entre a indissociabilidade de Ensino, Pesquisa e Extensão; a utilização de mídias alternativas e as bases da Pesquisa-Ação podem servir de esteio para uma resposta balizada às críticas de Rajagopalan (2003) aos princípios labovianos. Entendemos que o princípio da gratuidade linguística, proposto por Wolfram (1993), por ser complementar aos de Labov (1982), também se encaixa nas críticas levantadas por Rajagopalan (2003) e, por sua vez, também se beneficia das respostas aqui trazidas. Acreditamos que esses três princípios podem ser mantidos e uma vez cobertos pelas propostas apresentadas nesta seção podem produzir trabalhos que venham a encarar de forma concreta os desafios da comunicação entre linguistas e a sociedade. Algumas considerações a esse respeito, acrescentadas de ilustrações desse tipo de trabalho, formam o conteúdo da seção final desse artigo.

## **Comunicar para agir**

Nas seções anteriores, empreendemos algumas discussões sobre o ato de comunicar e sua relação com o trabalho do linguista. Vimos que desde, pelo menos, Saussure, existe uma preocupação explícita em tornar públicos os resultados das pesquisas e estabelecer um canal de comunicação com a sociedade. O trabalho de Labov (1982) apresenta propostas concretas para o estabelecimento desse canal, o que é complementado pela proposta de Wolfram (1993). Rajagopalan (2003) apresentou críticas a essas propostas e na seção anterior, trouxemos elementos que, se não podem talvez refutar essas críticas, se constituem como base para uma resposta concreta a elas.

Acreditamos que os novos olhares que foram trazidos podem trazer benefícios para o ato de comunicar que, por sua vez, como indicado no título da presente seção, deve também ter como consequência uma ação. Em um levantamento bibliográfico realizado, constatamos que os trabalhos de Wolfram (1993) e seu princípio, bem como os de Labov (1982), não são citados em artigos, dissertações, teses, etc. brasileiras. A nosso ver, trata-se de uma lacuna na discussão de como o linguista pode atuar na comunicação de suas pesquisas. O princípio da gratuidade linguística é um passo a mais na formalização dessa comunicação, posto que incita os linguistas a fazerem de sua comunicação algo ativo, positivo e estruturado. Não seriam mais ações pontuais, movidas por um senso de obriga-

ção, mas sim ações contínuas, movidas por um senso de justiça global, dado que torna o ato de comunicar mais amplo apoiando-se no conceito de favores linguísticos, que podem ser úteis à sociedade.

Não poderíamos deixar de finalizar essa discussão sem a apresentação de uma ilustração de como poderia ser uma comunicação com a sociedade que prezasse pelos princípios analisados aliando-se a eles as propostas de solução dos problemas levantados. A elaboração de uma cartilha digital contra o preconceito linguístico,<sup>6</sup> fruto de um projeto de iniciação científica, de um material didático voltado para o ensino de regras da norma padrão baseado nas regras da norma não-padrão, fruto de um projeto de iniciação à docência, e de uma atividade lúdica, em forma de labirinto, a respeito do preconceito linguístico voltado para alunos do 6º ano do ensino fundamental, fruto de um projeto de extensão, pode servir como ilustração da discussão aqui levantada. Estamos conscientes de que se trata de trabalhos em fase inicial de elaboração, ainda sem publicações ou produtos que permitam uma análise crítica mais aprofundada de seus resultados. No entanto, a menção dessas atividades tem como propósito apenas apontar sua existência e revelar que há formas de tornar concretas as propostas de comunicação dos linguistas com a sociedade. Entendemos, portanto, que somente a prática pode levar a uma realimentação da discussão empreendida pelo lado que defende princípios como os de Labov (1982) e Wolfram (1993) e pelo lado que os ataca. Uma vez que estejam disponibilizadas essas situações, será possível construir críticas e alternativas a essas próprias situações, gerando um contínuo de ações, colocando, por fim, em prática, as ideias veiculadas pelos linguistas que se preocupam de modo central da ação de comunicar, cientes que estão de que comunicar também leva a outras ações.

## REFERÊNCIAS

BARONAS, R. Da necessidade premente de se cometer uma política de divulgação científica qualificada dos trabalhos da lingüística do Brasil. *Revista da Anpoll*, v. 29, n. 1, p. 235-258. 2010.

BRASIL. Plano Nacional de Extensão Universitária 1999-2001. 2001. Disponível em <<http://www20.fcm.unicamp.br/extensao/arquivos/pne.pdf>>. Acesso em: 28 ago. 2012.

\_\_\_\_\_. Constituição da República Federativa do Brasil de 1988. 1988. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/constituicao/constitui%C3%A7ao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constitui%C3%A7ao.htm)>. Acesso em: 28 ago. 2012.

CATELLAN, J. *O gênero discursivo religioso católico: uma materialização discursiva previsível*. 1996. 230 f. (Mestrado em Linguística) – Universidade Federal do Paraná, Curitiba. Disponível em <<http://dspace.c3sl.ufpr.br/dspace/bitstream/handle/1884/24416/D%20D20CATELLAN,%20JOAO%20CARLOS.pdf?sequence=1>>. Acesso em 28/08/12.

CHARITY, A. Linguists as Agents for Social Change. *Language and Linguistics Compass*, 2/5: p. 923-939, 2008.

<sup>6</sup> Disponível em <http://www.cartilhacontraopreconceitolingustico.blogspot.com.br>

DIAS, A. Discutindo caminhos para a indissociabilidade entre ensino, pesquisa e extensão. *Revista Brasileira de Docência, Ensino e Pesquisa em Educação Física*, v. 1, n. 1, p. 37-52, ago. 2009.

LABOV, W. Objectivity and commitment in linguistic science. *Language in Society*, 11, 165-201, 1982.

NICOLELIS, M. *Manifesto da Ciência Tropical: um novo paradigma para o uso democrático da ciência como agente efetivo de transformação social e econômica no Brasil*. 2010. Disponível em <<http://www.viomundo.com.br/entrevistas/nicolelis-lanca-manifesto-da-ciencia-tropical-vai-ditar-a-agenda-mundial-do-seculo-xxi.html>>. Acesso em: 28 ago. 2012.

RAJAGOPALAN, K. Linguagem e ética: algumas considerações gerais. In: \_\_\_\_\_. *Por uma linguística crítica: linguagem, identidade e a questão ética*. São Paulo: Parábola, 2003. Cap. 1, p.15-23.

REASER, J., ADGER, C. Developing language awareness materials for nonlinguists: lessons learned from the *do you speak american?* Curriculum development project. *Language and Linguistics Compass*, 1/3, p. 155-167, 2007.

SAUSSURE, F. *Curso de linguística geral*. Cultrix: São Paulo, 1969. 216 f.

TAUCHEN, G. *O princípio da indissociabilidade universitária: um olhar transdisciplinar nas atividades de ensino, pesquisa e extensão*. 2009. 147 f. (Doutorado em Educação) – Faculdade de Educação, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.

TRIPP, D. Pesquisa-ação: uma introdução metodológica. *Educação e Pesquisa*, São Paulo, v. 31, n. 3, p. 443-466, set./dez. 2005.

WOLFRAM, W. Ethical considerations in language awareness programs. *Issues in Applied Linguistics*, 4, p. 225-255, 1993.

# O “r caipira” em Mato Grosso do Sul – estudo baseado em dados do *ALMS, Atlas linguístico do Mato Grosso do Sul*

(The “caipira r” in Mato Grosso do Sul - Brazil: study based on data from *ALMS, Atlas linguístico do Mato Grosso do Sul*)

Vandersí Sant’Ana Castro

Instituto de Estudos da Linguagem - Universidade Estadual de Campinas (Unicamp)

vandersi@iel.unicamp.br

**Abstract:** In this study we verify the occurrence of the “caipira r” in the Brazilian state of Mato Grosso do Sul, based on data from *ALMS, Atlas linguístico do Mato Grosso do Sul* (OLIVEIRA, 2007). The context considered is the syllabic coda, within or at the end of the word. The results show the predominance of the “caipira r” over other variants of /r/ in this context. It is an evidence of the influence of “caipira dialect” (a dialect in the state of São Paulo) in this area, which is historically explained by the presence of the *paulistas* in the settlement of this region.

**Keywords:** dialectology; dialect geography; linguistic atlas; Mato Grosso do Sul – Brazil; “caipira r”.

**Resumo:** Neste estudo, procuramos verificar a ocorrência e a frequência do “r caipira” no Mato Grosso do Sul, com base em dados do *ALMS, Atlas linguístico do Mato Grosso do Sul* (OLIVEIRA, 2007). O contexto examinado é a coda silábica, no interior e no final de palavra. Os resultados mostram uma indiscutível predominância dessa variante entre as realizações róticas atestadas no contexto em questão, sinalizando a influência do dialeto caipira nesse território, historicamente explicada pela atuação dos bandeirantes paulistas no povoamento dessa região.

**Palavras-chave:** dialetologia; geolinguística; atlas linguístico; Mato Grosso do Sul; “r caipira”.

## Introdução

O presente estudo, como o título indica, trata do “r caipira”, tema que implica uma referência obrigatória a Amadeu Amaral. Em 1920, Amaral (1982 [1920]) publica uma descrição do dialeto caipira, variedade do português popular brasileiro que teria sido falada pela grande maioria da população no território da antiga província de São Paulo até por volta do final do século XIX. O autor descreve o dialeto de forma abrangente e detalhada, apresentando e exemplificando suas características fonéticas, morfológicas, sintáticas e lexicais. Uma das variantes mais marcantes do dialeto é o r retroflexo, também denominado “r caipira” desde o trabalho de Amaral. Nos registros de Amaral, a variante é atestada em final de sílaba (ex.: *carta*) e em posição intervocálica (ex.: *arara*). Embora Amaral não seja preciso quanto às áreas em que coletou seus dados, sabemos, através de Duarte (1976, p. 91), que suas investigações se circunscreveram “às zonas de Capivari, Piracicaba, Tietê, Itu, Sorocaba e São Carlos, onde fez suas observações e de onde se originaram muitos dos seus poucos informantes”.

Muito tempo depois da publicação do trabalho de Amaral, a permanência dessa variante em território paulista tem sido atestada por diversos estudiosos. Rodrigues (1974), por exemplo, constatou a vitalidade do dialeto caipira como um todo na zona rural da região de Piracicaba, no início da década de 1970. No que se refere particularmente ao “r caipira”, a autora registrou, na área investigada, a ocorrência da variante nos dois contextos

apontados por Amaral (final de sílaba e posição intervocálica) e também como membro de encontro consonantal antes da vogal da sílaba (ex.: *grilo*). Na mesma época, Head (1973) documentou a ocorrência da variante em posição pós-vocálica (final de palavra e final de sílaba interna), em Franca, interior de São Paulo. Mais recentemente, realizações retroflexas em final de sílaba interna também foram atestadas na fala espontânea de informantes de São José do Rio Preto-SP, por Leite (2004), que identificou a aproximante retroflexa como a pronúncia típica da cidade. (Embora não tenha feito sua pesquisa *in loco*, Leite coletou seus dados junto a estudantes de São José do Rio Preto, que, na época, faziam graduação na Universidade Estadual de Campinas, em Campinas-SP.) Ampliando a investigação sobre a presença dessa variante no Estado de São Paulo, e considerando especificamente o contexto final de sílaba (interna e externa), exploramos, em trabalhos anteriores, os dados coletados pelo Projeto do Atlas linguístico do Brasil (ALiB)<sup>1</sup> referentes a 23 cidade paulistas – 10 delas situadas em uma larga faixa ao longo do Rio Tietê (CASTRO, 2009a); seis, no Vale do Paraíba e no litoral (CASTRO, 2010); seis, no norte do Estado (CASTRO, 2011); e também examinamos os dados referentes à capital (CASTRO, 2009b). À exceção da capital São Paulo e da cidade de Santos, em todas as localidades investigadas constatamos a presença marcante do “r caipira”, que ocorre como a variante indiscutivelmente predominante, quando não exclusiva. Nessas localidades, a principal alternativa ao retroflexo, quando ele não ocorre, é o tepe. Na cidade de São Paulo e em Santos, no entanto, o “r caipira” tem presença discreta. Entre os paulistanos, o tepe é a realização predominante, e a fricativa velar tem registro insignificante, figurando como uma variante dissonante. Entre os santistas, também é o tepe a realização rótica predominante, mas a fricativa velar, de ocorrência esporádica na Capital e nas outras localidades investigadas, tem em Santos uma presença mais sistemática, ocupando o segundo lugar entre as variantes róticas, vindo o retroflexo na terceira posição.

Nosso objetivo no presente trabalho é levar a investigação para o território de Mato Grosso do Sul, tarefa facilitada pela publicação do atlas linguístico desse Estado (OLIVEIRA, 2007). Há justificativas para se investigar essa área, uma vez que os bandeirantes tiveram função importante no povoamento do território correspondente ao antigo Estado do Mato Grosso, agora dividido em dois Estados (Mato Grosso e Mato Grosso do Sul). Com a Guerra dos Emboabas em território mineiro, os bandeirantes paulistas, perdedores, são pressionados pela situação a procurarem outras áreas, no caso, Mato Grosso e Goiás, onde as descobertas de minérios se renovam. Nessas incursões, utilizam sobretudo o caminho dos rios, no movimento conhecido como monções. Com esse patente vínculo dos primeiros paulistas com o território de Mato Grosso, é compreensível encontrar nessa região traços do dialeto caipira. Daí nosso interesse em verificar a presença do “r caipira” no Mato Grosso do Sul.

## O “r caipira”

Parece-nos oportuno retomar os termos de Amaral (1982 [1920], p. 47-48) na descrição do “r característico” do dialeto caipira, atestado em duas posições, como já foi mencionado: final de sílaba (ex.: *carta*) e entre vogais (ex.: *arara*). A descrição de Amaral é minuciosa:

---

<sup>1</sup> Uma descrição resumida do Projeto ALiB encontra-se em Cardoso (1998).

[...] r inter e post-vocálico (carta, arara), possui um valor peculiar: é línguo-palatal e guturalizado. Na sua prolação, em vez de projetar a ponta da língua contra a arcada superior, movimento este que produz a modalidade portuguesa, a língua leva os bordos laterais mais ou menos até os pequenos molares da arcada superior e vira a extremidade para cima, sem tocá-la na abóbada palatal. Não há quase nenhuma vibração tremulante. Para o ouvido, este r caipira assemelha-se bastante ao inglês post-vocálico. (AMARAL, 1982 [1920], p. 47-48)

Head (1973) observa que a descrição de Amaral (“a língua... vira a extremidade para cima, sem tocá-la na abóbada palatal”) permite classificar esse “r” como retroflexo, articulação cujo correlato acústico é rebaixado ou bemolizado. Observa, ainda, que o levantamento do dorso da língua (uma articulação posterior ou velarizada) produz o mesmo efeito acústico – bemolizado. Essa articulação posterior explicaria o termo “guturalizado”, usado por Amaral (1982 [1920]), expressão que seria, então, entendida como uma articulação com o dorso da língua.

Outro aspecto da descrição de Amaral que merece ser destacado é sua afirmação de que “não há quase nenhuma vibração tremulante” na produção do “r caipira” (1982 [1920], p. 47/48). Head (1987) demonstrou a incompatibilidade articulatória da *vibração* com a *retroflexão*, confirmando a característica apontada na descrição de Amaral, e desautorizando a classificação dessa variante como vibrante retroflexa. Head (1987, p. 25) pondera, no entanto, que, dado que é convencional essa classificação entre nossos estudiosos, deve-se aceitar essa terminologia, “tecnicamente contraditória”, tendo em mente que “vibrante”, no caso da retroflexa, corresponde a uma “consoante líquida não lateral”, ou seja, não indica vibração.

## O Atlas linguístico do Mato Grosso do Sul – ALMS

Um dos últimos atlas linguísticos regionais brasileiros publicados é o *ALMS, Atlas linguístico do Mato Grosso do Sul* (OLIVEIRA, 2007). Embora tenha sido publicado há quase cinco anos, acreditamos que é oportuno apresentar algumas informações sobre esse trabalho, considerando que esse tipo de material, infelizmente, não tem muita circulação entre nós. Evidentemente, essas informações são limitadas e não substituem a consulta ao atlas. Muito ao contrário, pretendem ser um estímulo a essa consulta.

**Rede de pontos** – A rede de pontos do ALMS é constituída por trinta e duas localidades, “distribuídas por cinco setores, cujos municípios principais são Três Lagoas, Corumbá, Aquidauana, Campo Grande e Dourados”. A escolha das localidades se fez com base em critérios “demográficos, históricos e sociais”, como consta na introdução do atlas (OLIVEIRA, 2007, p. 21).

**Informantes** – Foram inquiridos 128 informantes, quatro por localidade, distribuídos equitativamente nos dois sexos, e em duas faixas etárias – não há no texto introdutório do atlas uma explicitação dos limites dessas faixas, mas consultando a apresentação dos informantes, pode-se constatar que na primeira faixa os entrevistados variaram de 18 a 36 anos, e, na segunda faixa etária, de 42 a 83 anos.<sup>2</sup> Quanto à escolaridade, os informantes

<sup>2</sup> Na descrição dos Procedimentos metodológicos, seção que consta do texto introdutório do atlas, é apresentada a lista dos informantes, por localidade, identificando-se cada um pelo estado civil, profissão, idade, grau de escolaridade e naturalidade, dados antecedidos das iniciais da pessoa. Na ordem de

variam de nenhuma instrução (analfabetos) ao primário completo (4ª série do ensino fundamental) – pudemos constatar que a maioria (52 entrevistados) tem o primário incompleto; 37 informantes concluíram o primário; e 33 são analfabetos.

**Questionário** – Nos inquéritos foi utilizado um questionário de 557 perguntas, subdividido em dois módulos: um Questionário Fonético-Fonológico (QFF), voltado para aspectos fonéticos, e um Questionário Semântico-lexical (QSL), constituído de questões lexicais. O levantamento de aspectos morfossintáticos, também cartografados, foi feito a partir das narrativas dos informantes sobre fatos marcantes em sua experiência de vida.

**Cartas** – O atlas é composto por 231 cartas linguísticas: 57 fonéticas, 167 semântico-lexical e 7 morfossintáticas.<sup>3</sup> Como nome da carta figura o número da pergunta correspondente: por exemplo, a primeira carta do atlas é nomeada QFF 01.a – FAMÍLIA (logo abaixo se repete o item lexical, destacando-se o segmento cartografado – famí(lia)).<sup>4</sup> As variantes registradas são apresentadas na legenda e associadas a ícones correspondentes. Ao lado de cada uma indica-se a respectiva porcentagem de ocorrência atestada. No mapa, as variantes são representadas pelos respectivos ícones dispostos em cruz, que permitem identificar a realização correspondente a cada informante. (Além da disposição em cruz, os ícones recebem a cor específica da categoria de informante que produziu a variante, o que seria dispensável, pois a disposição em cruz seria suficiente para a identificação do informante. Uma das cores escolhidas é o amarelo, que dificulta muitas vezes a identificação do ícone que representa a variante.) Essa representação permite uma rápida visualização, mas tem o inconveniente de não permitir o registro de mais de uma variante por informante, se for o caso.

A título de exemplificação, reproduzimos, no ANEXO I, uma carta fonética do ALMS. Trata-se da Carta QFF 13.a – CALOR, que registra os dados relativos à pronúncia da última sílaba da palavra.

Após essa visão geral do atlas, podemos descrever o material especificamente selecionado para nossa análise.

---

apresentação, são alistados inicialmente os dois informantes da primeira faixa etária, o homem e em seguida a mulher, e depois os entrevistados da segunda faixa etária, o homem e em seguida a mulher. Em relação à localidade de Camapuã, os dados estão incompletos - só consta um informante, do sexo masculino, havendo um equívoco de digitação que impossibilita a identificação de sua idade. Em relação a Porto Murтинho, faltam dados sobre a escolaridade e a naturalidade da informante da segunda faixa etária; também falta a indicação da escolaridade do informante da segunda faixa etária de Paiaguás e da informante dessa mesma faixa etária de Porto Esperança.

3 Essa contagem não coincide com a que Oliveira (2007, p. 267) apresenta na Conclusão do atlas: pelo critério aqui adotado foram computadas todas as cartas que integram o atlas, inclusive as que correspondem a desdobramentos de uma mesma pergunta, diferentemente do autor, que considerou o número de perguntas que deram origem às cartas e que as nomeiam. Por exemplo: à questão QFF 01 (família) correspondem, efetivamente, duas cartas – QFF 01.a famí(lia) e QFF 01.b (fa)mília –, que focalizam, respectivamente, o último e o primeiro segmento do item lexical (o segmento em foco é colocado entre parênteses). Para Oliveira, conta-se aí uma carta apenas.

4 Para destacar o segmento cartografado, curiosamente adota-se o critério de colocá-lo entre parênteses, recurso geralmente usado com o efeito contrário. O mais esperado seria colocar entre parênteses o segmento desconsiderado na cartografiação, recurso adotado, por exemplo, pelo Atlas linguístico-etnográfico da região Sul do Brasil – ALERS (KOCH *et alii*, 2002).

## O corpus

Das 57 cartas fonéticas do ALMS, 12 são relevantes para se verificar a ocorrência do “r caipira” em final de sílaba, incluindo-se aí os casos de rotacismo, como *purmão* por *pulmão*, e os casos de metátese, como *porcissão* por *procissão*. Os itens em questão são os seguintes:<sup>5</sup> *revólver* (por rotacismo – carta QFF 08.a); *revólver* (carta QFF 08.b); *calor* (carta QFF 13.a); *calção* (por rotacismo – carta QFF 14.a); *fervendo* (QFF 15.b); *gordura* (QFF 17. a); *procissão* (por metátese – carta QFF 23.a); *pulmão* (por rotacismo – carta QFF 36.a); *colher* (QFF 44.a); *garfo* (QFF 45.a); *fósforo* (por metátese – carta QFF 46.a); *porta* (QFF 47.a).

Considerando que são 32 localidades, com quatro informantes cada, e 12 itens lexicais relevantes para o contexto em exame, a expectativa era de analisarmos 1536 dados. Todavia, constatamos 200 casos de ausência de dados, situação indicada nas cartas pelas siglas RP (respostas prejudicadas, em nossa interpretação) ou RNP (questões não formuladas pelo inquiridor).<sup>6</sup> Uma das localidades, Nhecolândia, foi excluída da análise por ausência total de dados. Os casos de ausência de respostas reduziram para 1336 o número de dados examinados. Desses dados, são numerosos os que não apresentam o contexto relevante para a análise. São os casos em que os processos de rotacismo (em *revólver*, *pulmão* e *calção*) ou de metátese (em *procissão* e *fósforo*) não se efetivaram; são ainda os casos em que o rótico se deslocou da posição de coda para o início de uma nova sílaba (ex.: *revólver* = *virri*) ou para antes da vogal da sílaba (ex.: *fervendo* = *fre*). Também se enquadram na mesma situação as ocorrências em que, por hipercorreção, o rótico da coda silábica é realizado como semivogal (foi frequente, por exemplo, a realização *gaw* para *garfo*); ou em que o rótico é vocalizado, como ocorre na linguagem infantil (ex.: *gaj* para *garfo*). O conjunto de casos dessa natureza corresponde a 449 ocorrências, o que significa que se reduzem a 887 os dados efetivamente pertinentes para a análise.

## Resultados da análise

Considerando as localidades em conjunto, mesmo porque os dados não variam muito de uma para outra, constatamos, nesses 887 dados pertinentes, a ocorrência de quatro realizações róticas - retroflexo, tepe, fricativa velar e fricativa glotal – e a realização zero, ou seja, a queda do rótico, nos índices apontados na **Tabela 1**, apresentada abaixo.

5 Destacamos em negrito o segmento em foco.

6 As siglas RP e RNP não são explicadas, mas pode-se deduzir que correspondem, respectivamente, a “resposta prejudicada” e “resposta (sic) não perguntada” – siglas semelhantes são usadas no ALERS, uma possível referência para o ALMS, com significados semelhantes: RP = resposta prejudicada, e NP = não perguntado.

**Tabela 1. Realizações do rótico em coda silábica – Mato Grosso do Sul (cf. dados do ALMS)**

variante fonética	ocorrências	
	número	porcentagem
retroflexo	711	80,15%
zero	138	15,55%
fricativa velar	22	2,48%
tepe	14	1,57%
fricativa glotal	2	0,22%
total	887	99,97% <sup>7</sup>

Como se pode observar, os números apontam o predomínio indiscutível do retroflexo entre as realizações atestadas.

Excluindo os casos de zero e comparando somente as ocorrências de róticos, já que nosso interesse principal é ver a opção dos informantes entre os róticos, obtemos os índices apresentados na **Tabela 2**, apresentada a seguir, em que fica mais clara a distância entre a frequência do “r caipira” e a das outras realizações róticas no desempenho dos informantes.

**Tabela 2. Variantes róticas em coda silábica – Mato Grosso do Sul (cf. dados do ALMS)**

variante fonética	ocorrências	
	número	porcentagem
retroflexo	711	94,92%
fricativa velar	22	2,93%
tepe	14	1,86%
fricativa glotal	2	0,26%
total	749	99,97% <sup>8</sup>

Reforçando o que os números acima indicam, salientamos alguns aspectos dos dados analisados, que também apontam o “r caipira” como a pronúncia característica do Mato Grosso do Sul:

- a) – o “r caipira” foi documentado em todas as localidades investigadas;
- b) – é a única variante atestada em 21 das 31 localidades investigadas,<sup>9</sup> inclusive na Capital, que, em geral, fornece o modelo para o uso linguístico nas demais localidades;
- c) – essas 21 localidades em que só foi atestado o “r caipira” no contexto em questão se distribuem por praticamente todo o Estado;
- d) – nas localidades em que se registram outros róticos (o tepe, a fricativa velar e a fricativa glotal), o retroflexo é sempre a realização predominante, ocorrendo em índices que variam de 68% a 97%, enquanto as outras variantes ocorrem em índices

7 O total não é de 100% porque só consideramos até a ordem das dezenas após a vírgula e não fizemos aproximação da porcentagem.

8 O total não é de 100% porque só consideramos até a ordem das dezenas após a vírgula e não fizemos aproximação da porcentagem.

9 Lembramos que em Nhecolândia houve ausência total de dados, ficando a localidade excluída de nossa análise.

muito inferiores, que vão de 3% a 18% (mais precisamente: a fricativa velar, de 7% a 17%; o tepe, de 3% a 18%; a fricativa glotal, em 10%).

Reconhecido o predomínio do retroflexo, cabe verificar se as outras variantes se registram de forma consistente em alguma área específica do Mato Grosso do Sul ou revelam algum tipo de regularidade em sua ocorrência.

Começando pela fricativa glotal, que só teve duas ocorrências em um total de 749 dados, observamos que foi registrada em uma única localidade, Água Clara, e no desempenho de um único informante. O fato de ser usada por um único informante (homem, 70 anos) leva a pensar que o fator geográfico não estaria envolvido, tendo talvez pesado alguma particularidade da história pessoal do falante, não transparente nos dados registrados na ficha que consta no atlas. Entre quatro dados relevantes, o informante realizou dois retroflexos e duas fricativas glotais – com dados mais numerosos seria mais seguro verificar a consistência do desempenho do falante.

O tepe, que somou 14 registros, foi atestado em 6 pontos (Paiaguás – 5 ocorrências; Nabileque – 3 ocorrências; Rio Negro e Rochedo – 2 ocorrências; Iguatemi e Eldorado – 1 ocorrência). Essas localidades, como se pode ver no mapa do ANEXO II, se situam a noroeste, no centro-norte e no extremo sul do Estado, mostrando alguma concentração geográfica. Todavia, como os registros são pouco numerosos em cada localidade e sobretudo pouco significativos em comparação com a presença predominante e muito mais volumosa do retroflexo, acreditamos que é prematuro fazer qualquer afirmação mais contundente sobre o peso do fator geográfico na questão. Deve-se observar, ainda, que as ocorrências se distribuem entre as quatro categorias de informantes, não revelando um peso maior dos fatores sexo ou idade do falante.

Situação semelhante se observa quanto às ocorrências da fricativa velar. Foram registradas em 7 localidades (Paiaguás, Rochedo, Nabileque e Eldorado – 4 ocorrências; Sete Quedas – 3 ocorrências; Fátima do Sul – 2 ocorrências; Ponta Porá – 1). Essas localidades, como se pode ver no mapa do ANEXO II, também se situam no noroeste, no centro-norte e no sul do Estado, mostrando certa concentração geográfica. Todavia, também no caso desta variante, os registros são pouco numerosos, sobretudo em relação à presença muito mais significativa do retroflexo, predominante, o que sugere cautela na identificação do fator geográfico como influente. Deve-se observar, ainda, que os dados não apontaram influência dos fatores idade e sexo do falante, já que as ocorrências se distribuíram nas 4 categorias de informantes. Um fato a destacar é que a variante fricativa velar foi registrada unicamente na realização do item *fervendo*.

Duas observações, todavia, podem ser feitas com relação à distribuição diatópica do tepe e da fricativa velar: a) as áreas no noroeste, no centro-norte e no sul do Estado, em que se concentra a ocorrência dessas variantes, poderiam ser vistas como áreas de certa instabilidade na realização do rótico no contexto em questão, apresentando, inclusive, em várias localidades, três variantes diferentes; e b) muitas das localidades em questão se situam em região próxima à fronteira com países vizinhos, o que propicia a situação de contato linguístico, que pode ter alguma atuação na alternância observada.

## Considerações finais

Chegamos, assim, às considerações finais deste estudo. Os dados do ALMS analisados mostraram, de forma consistente, que o “r caipira” tem presença marcante no Mato Grosso do Sul. Foi atestado em todas as localidades investigadas, como variante exclusiva ou nitidamente predominante. É um sinal evidente da influência do dialeto caipira na região, historicamente explicada pela atuação dos bandeirantes paulistas no povoamento desse território.

Em algumas localidades, foi constatada a alternância do retroflexo com a fricativa velar, o tepe e a fricativa glotal. A presença dessas três últimas variantes, no entanto, não se mostrou consistente a ponto de permitir identificá-las como a pronúncia típica de alguma área específica. A ocorrência dessas variantes permanece, todavia, como uma questão a ser estudada mais profundamente. O exame de dados mais numerosos e a consideração de aspectos históricos relativos ao povoamento das localidades em questão poderiam ser caminhos adotados na investigação do tema.

## REFERÊNCIAS

AMARAL, A. *O dialeto caipira: gramática, vocabulário*. 4. ed. São Paulo: Hucitec/INL-MEC, 1982. (ed. fac.simil. da 2. ed. 1955).

CARDOSO, S. A. M. O Atlas linguístico do Brasil – um projeto nacional. In: AGUILERA, V. A. (Org.). *A geolinguística no Brasil: caminhos e perspectivas*. Londrina: Ed. UEL, 1998. p. 165-176.

CASTRO, V. S. *O “r caipira” em localidades do norte do Estado de São Paulo – estudo com base em dados do Atlas linguístico do Brasil (ALiB)*. Apresentação de trabalho. GEL, jul. 2011.

\_\_\_\_\_. *O “r caipira” em localidades paulistas do Vale do Paraíba e do litoral – estudo com base em dados do Atlas linguístico do Brasil (ALiB)*. Apresentação de trabalho. GEL, jul. 2010.

\_\_\_\_\_. *O “r caipira” no Estado de São Paulo: estudo com base em dados do Atlas linguístico do Brasil (ALiB)*. Apresentação de trabalho. GEL, jul. 2009a.

\_\_\_\_\_. *O “r caipira” em São Paulo: estudo com base em dados do Atlas linguístico do Brasil (ALiB)*. Apresentação de trabalho. Uberlândia: SILEL, nov. 2009b.

DUARTE, P. *Amadeu Amaral*. São Paulo: Hucitec, Secretaria de Cultura e Tecnologia, 1976.

HEAD, B. F. Propriedades fonéticas e generalidades de processos fonológicos: o caso do *r caipira*. *Cadernos de estudos linguísticos*, Campinas, n. 13, p. 5-39, 1987.

\_\_\_\_\_. O estudo do ‘r caipira’ no contexto social. *Revista de cultura Vozes*, [s.l.], n. 67, v. 8, p. 43-49, 1973.

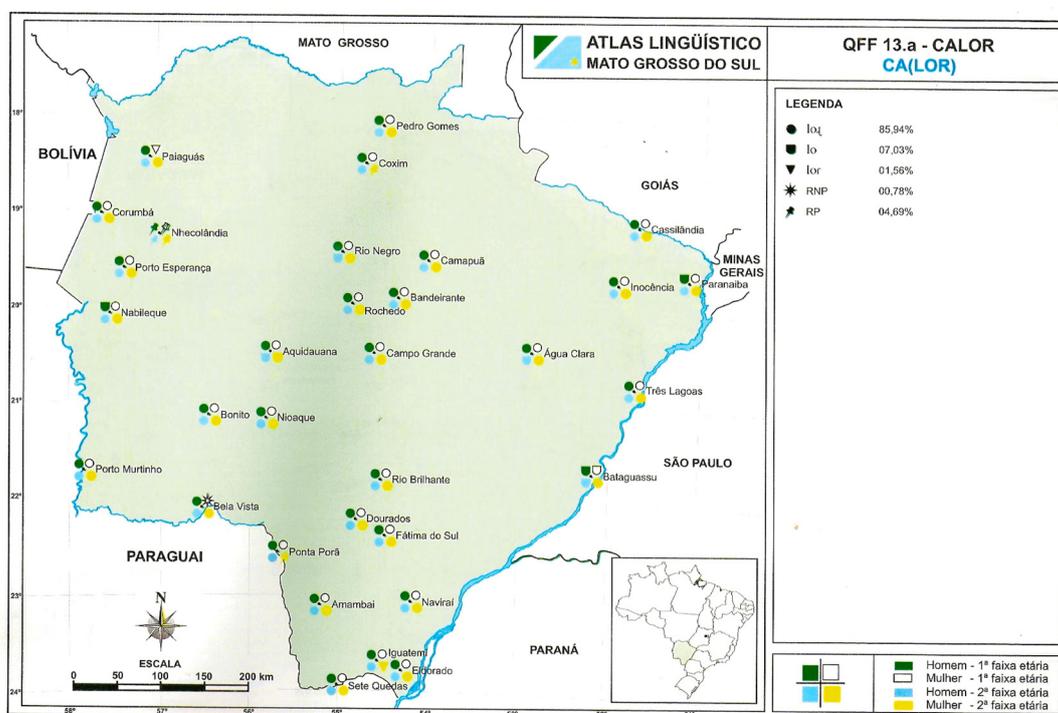
KOCH, W., KLASSMANN, M. S. e ALTENHOFEN, C. V. (Org.). *Atlas linguístico-etnográfico da região Sul do Brasil (ALERS)*. Porto Alegre/Florianópolis/Curitiba: Ed. UFRGS/Ed. UFSC/Ed. UFPR, 2002. v. 2 [Cartas fonéticas e morfossintáticas].

LEITE, C. M. B. *Atitudes linguísticas: a variante retroflexa em foco*. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Unicamp, 2004.

OLIVEIRA, D. P. (Org.). *ALMS, Atlas linguístico do Mato Grosso do Sul*. Campo Grande, MS: Ed. UFMS, 2007.

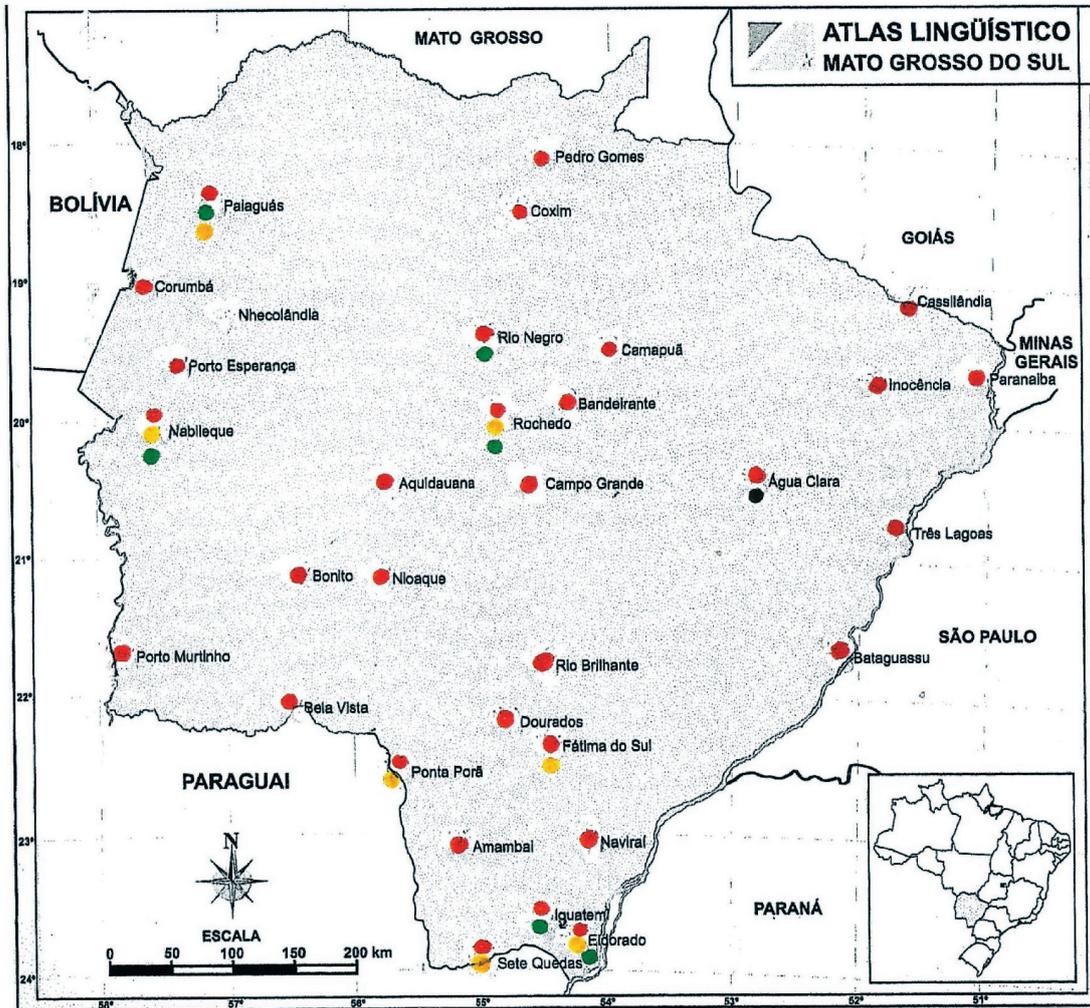
RODRIGUES, A. N. *O dialeto caipira na região de Piracicaba*. São Paulo: Ática, 1974.

## ANEXO I



## ANEXO II

Róticos em coda silábica – Mato Grosso do Sul (cf. dados do ALMS)



- Legenda:
- retroflexo
  - fricativa velar
  - tepe
  - fricativa glotal