

Atlas Linguístico do Paraná – v. II: comentários sobre a Dialetoimetria

(Linguistic Atlas of Paraná – v. II: comments on Dialectometry)

Fabiane Cristina Altino¹

¹Universidade Estadual de Londrina (UEL)

fabiane_altino@uol.com.br

Abstract: This paper aims to describe the preparation of the second volume of the Linguistic Atlas of Paraná - ALPR vol II and discuss the treatment of data according to the Dialectometry, a method which has not been much disseminated in Brazil. Besides providing an instrument to access dialectologic data of rural language spoken in Paraná for students and researchers, the second volume of ALPR analyses the data based on a dialetoimetric perspective in order to find lexical areas related to the settlement of the State.

Keywords: Dialectology; Dialectometry; ALPR v. II.

Resumo: Este trabalho busca descrever a preparação do segundo volume do Atlas Linguístico do Paraná – ALPR vol II e discutir o tratamento de dados à luz da Dialetoimetria, método pouco disseminado no Brasil. Além de oferecer aos estudantes e pesquisadores da língua portuguesa um instrumento de acesso a dados dialetoológicos sobre a linguagem rural paranaense, o segundo volume do ALPR analisou os dados na perspectiva dialetoimétrica na busca de áreas lexicais ligadas ao povoamento do Estado.

Palavras-chave: Dialetoologia; Dialetoimetria; ALPR v. II.

Introdução: um pouco sobre a dialetoologia brasileira e os atlas regionais

A Língua Portuguesa no Brasil – PB – não é uma nem monolítica e apresenta muitas variedades dialetais¹ Essa afirmação parece ser um consenso entres os pesquisadores do PB, que também concordam com a necessidade de descrição desses dialetos, em sua maioria ainda não descritos cientificamente. Não é preciso procurar muito; a literatura popular está repleta de obras – dicionário do gauchês, de baianês, vocabulário nordestino, vocabulário popular paranaense, para citar alguns – que traduzem o falar dos habitantes das diversas regiões do país. Todas essas obras servem para identificar e marcar geográfica e socialmente seus usuários, mas não trazem em si o rigor científico desejado pela academia.

Outro ponto a ser comentado é que, em decorrência dos diversos falares, registram-se várias formas de preconceitos consequentes do valor social que se atribuem aos diferentes modos de fala: é muito comum a observação de que as variedades linguísticas de menor prestígio são inferiores ou erradas. Essa ‘exigência’ linguística ganha força à medida que as comunidades vão entrando em contato com a escola e a evolução científica e tecnológica.

¹ Para o conceito de dialeto e de suas derivações, adotou-se, para este trabalho, a definição de Alvar (1996, p. 13), para quem dialeto é “un sistema de signos desgajado de una lengua común, viva o desaparecida; normalmente, con una concreta limitación geográfica, pero sin una fuerte diferenciación frente a otros de origen común. De modo secundario, pueden llamarse dialectos las estructuras lingüísticas, simultáneas a otras, que no alcanzan la categoría de lengua”. Cf. também Chambers e Trudgill (1994), Dubois (1978) e Camara Jr. (1981).

A cultura, antes regional e particular, isolada e com costumes relativamente uniformes e cristalizados, deu lugar à civilização universalista e urbana, com seus ‘novos’ problemas. As comunidades, agora letradas e em contato constante com o mundo (através da tecnologia e da mídia), iniciaram uma nova era na comunicação humana, mais exigente e formal. A migração, o contato entre as culturas, o menor ou o maior acesso à mídia e à escolaridade, entre outros fatores, contribuem para que a língua seja cada vez mais dinâmica e se transforme continuamente.

Porém essas mudanças não se fazem de forma homogênea. Há comunidades em que, por diversas razões (culturais, geográficas), a mudança se faz mais lentamente, perpetuando ali hábitos linguísticos há muito deixados para trás por outras comunidades de fala.

Buscar as diferenciações regionais, verificando as alterações sócio-históricas que cada comunidade apresenta, e averiguar as diversidades geográficas na busca do retrato linguístico são metas que devem nortear a tarefa de um dialetólogo. Esses objetivos devem estar presentes não só nos atlas nacionais, mas, principalmente, nos regionais, por possibilitarem o estudo da língua por meio de um inventário linguístico verticalizado, no qual se espelhem algumas mudanças sociais, oferecendo, assim, subsídios para a constituição de um quadro sinótico da língua.

Há muito, no Brasil, estudiosos da linguagem vêm pregando a necessidade de estudos dos falares nacionais. Esse propósito já era discutido por Amadeu Amaral (1976, p. 2):

Seria de se desejar que muitos observadores imparciais, pacientes e metódicos se dedicassem a recolher elementos em cada uma dessas regiões, limitando-se estritamente ao terreno conhecido e banindo por completo tudo quanto fosse hipotético, incerto, não verificado pessoalmente. Teríamos assim um grande número de pequenas contribuições, restritas em volume e em pretensão, mas que na sua simplicidade modesta, escorreita e séria prestariam muito maior serviço do que certos trabalhos mais ou menos vastos, que de quando em quando se nos deparam, repositórios incongruentes de fatos recolhidos a todo preço e de generalizações e filiações quase sempre apressadas.

Essa preocupação com os falares brasileiros começa a tomar forma no Congresso Internacional de Linguística, em Haia, em 1928, onde se proclama a necessidade de desenvolver estudos dialetológicos no Brasil, solicitando junto ao governo federal, o patrocínio dos trabalhos necessários ao Atlas Linguístico do Brasil. Essa preocupação também é preconizada por Houaiss (1960), quando relata sobre a criação do Centro Latino-Americano de Pesquisa de Ciências Sociais, no Rio de Janeiro, em 1957. Assinala a necessidade de o recém-fundado centro levar em consideração os aspectos fundamentais do problema linguístico brasileiro, uma vez que

o estudo científico do instrumento de comunicação – no nosso caso concreto, a língua portuguesa feioada às nossas características nacionais – está cada vez mais na dependência de pesquisas de campo, graças às quais se poderá levar a bom termo, num futuro que esperamos não seja remoto, o Atlas Linguístico do Brasil, melhor, o Atlas Dialectológico Brasileiro. (HOUAISS, 1960, p. 40)

A ideia de mapear os falares brasileiros ganha corpo com Cunha e Silva Neto, no II Colóquio Internacional de Estudos Luso-Brasileiros, realizado em Lisboa em 1957, que igualmente afirmavam a urgência de estudos dialetológicos pelo método geolinguístico,

ressaltando a importância da elaboração de atlas regionais que pudessem investigar, mais detidamente as variantes de cada região do Brasil. Soma-se a esses pesquisadores Nascentes, cujas proposições estão registradas na publicação das *Bases para a elaboração do atlas lingüístico do Brasil* (1958; 1961), obra que fornece as diretrizes gerais para o estabelecimento da rede de pontos, o perfil dos informantes e a elaboração do instrumento de coleta de dados. É nessa efervescência que começa a surgir o ‘gosto’ pela dialetologia no Brasil e, a partir desses primeiros estudos, pesquisadores se lançam em busca de mapear os falares regionais.

A importância dos atlas regionais, já discutida por Cunha e Silva Neto, Nascentes (1958 e 1961) e também preconizada por Houaiss (1960), reside na exaustão com que os dados coletados são trabalhados, fornecendo, além de uma visão panorâmica da língua, a possibilidade de fazer seu inventário e documentar o grau de aculturação presente no espaço geográfico, de realizar o estudo da formação da língua e da sua história. Além disso, podem oferecer materiais agrupados em esferas semânticas e mostrar a distribuição geográfica das palavras. Houaiss (1960, p. 62) sinaliza para a importância dos atlas linguísticos:

As razões determinantes da necessidade de que cada comunidade lingüística possa ter sua fisionomia caracterizada exaustivamente são de várias naturezas. A técnica de apreensão dessa fisionomia lingüística é chamada dialetologia, cujos resultados só se tornam válidos se expostos em forma de atlas. O atlas lingüístico de uma comunidade lingüística é, por conseguinte, aquele registro mercê do qual todos os aspectos e particularidades dessa comunidade lingüística possam ser descritos, sistematizados, analisados.

É preciso mencionar, ainda, que esses estudos auxiliam na reconstituição das fases da língua, fornecendo uma visão sincrônica e diacrônica, ao mesmo tempo. Segundo o referido autor, “com respeito especificamente ao português do Brasil, ocorreria lembrar que de há muito se vem fazendo sentir a necessidade imperativa de uma pesquisa exaustiva e exata do nosso domínio dialetológico” (HOUAISS, 1960, p. 44).

Sobre a importância do trabalho dialetológico, Guarner (1953, p. 9) preconiza que

La Ciencia tiene el deber de recopilar antes de su pérdida, estos testimonios vernáculos que reflejan una mentalidad autóctona, una concepción local particular de la vida, elaborados lentamente [...]. El terreno a explorar, cada año que pasa, se hace más estéril y más arduo: de aquí la urgencia de la labor.²

Assim, começaram a ser desenvolvidas pesquisas dialetais que buscavam as realidades regionais, elaborando atlas que demonstrassem de forma concreta as possíveis diferenças linguísticas existentes dentro do território nacional. O primeiro deles foi o *Atlas Prévio dos Falares Baianos* (APFB – 1963), seguido pelo *Esboço de um Atlas Lingüístico de Minas Gerais* (EALMG – 1977), *Atlas Lingüístico da Paraíba* (ALPB – 1984), *Atlas Lingüístico de Sergipe* (ALS – 1987), *Atlas Lingüístico do Paraná* (ALPR – 1994), *Atlas Lingüístico-Etnográfico da Região Sul do Brasil* (ALERS – 2002), *Atlas Geo-sociolingüístico do Pará* (ALISPA – 2004), *Atlas Lingüístico de Sergipe – II* (ALS II – 2005), *Atlas Lingüístico do Mato Grosso do Sul* (ALMS – 2007), *Atlas Lingüístico do Estado do Ceará* (2010) e *Atlas Lingüístico-Etnográfico da Região Sul do Brasil II* (ALERS II – 2011).

² Tradução livre: A Ciência tem o dever de recopilar, antes da sua perda, estes testemunhos vernáculos que reflitam uma mentalidade autóctone, uma concepção local particular da vida, elaborados lentamente [...]. O território a explorar, cada ano que passa, será mais estéril e árduo: aqui a urgência do trabalho.

O Paraná e seus atlas

O *Atlas Lingüístico do Paraná – ALPR*; (AGUILERA, 1994), quinto atlas regional brasileiro, traz o registro de alguns aspectos da fala rural paranaense. Como resultados, apresentam-se 191 cartas geolinguísticas, que contemplam itens lexicais e aspectos fonéticos, ambos apresentando notas explicativas das variantes registradas e isoglossas da linguagem paranaense. A metodologia adotada para esse trabalho – escolha da rede de pontos, definição do instrumento de coleta de dados e o perfil dos informantes – foi norteadas pela Dialetoologia tradicional, postulada por Cunha e Silva Neto (1957) e Nascentes (1958; 1961), que, após algumas adaptações à realidade histórica e regional, ofereceu a homogeneização dos procedimentos de coleta de dados necessária à pesquisa dentro dos moldes científicos.

O *Atlas Linguístico do Paraná*, entretanto, foi publicado apresentando um recorte no *corpus* coletado e somente 37,85% das questões foram cartografadas no primeiro volume. Embora as entrevistas realizadas para este projeto já tenham sido exploradas por inúmeros pesquisadores e acadêmicos das diversas áreas de estudo da linguagem, uma quantidade significativa do *corpus* permanecia ainda inexplorada do ponto de vista geolinguístico. As respostas inéditas, representando 62,15% do questionário aplicado, são o objeto de estudo para o segundo volume do *Atlas Linguístico do Paraná*. A exemplo do primeiro volume, as cartas apresentam informações semântico-lexicais, fonético-fonológicas e notas explicativas para reproduzir as variantes registradas nos 65 pontos de estudo. Além de fornecer subsídios para outras pesquisas, o volume dois do ALPR oferece aos estudantes e pesquisadores da língua portuguesa um instrumento de acesso a dados dialetológicos e geolinguísticos mais amplos sobre a linguagem rural paranaense.

Sobre a composição de um segundo volume com dados já coletados, Cardoso (2005, p. 108) afirma a “existência de um número considerável de dados ainda inéditos e da importância de tê-los analisados com resultados passíveis de comparação com os que se oferecem ou oferecerão para outras áreas”. Acrescenta, ainda, a necessidade de “exploração de corpora organizados e ainda não trabalhados”, tarefa que se torna compromisso acadêmico e social. Essas afirmações são sentidas aqui e justificam a utilização dos *corpora* já constituídos para o ALPR.

Essa é também a opinião de Vieira (1998), ao ressaltar a necessidade de utilização dos dados coletados para a constituição dos atlas linguísticos e que acabam sem tratamento científico. A autora afirma que

a ausência de um tratamento informatizado, adequado aos dados das entrevistas para agilizar seu aproveitamento sem perda de dados, dificulta, ou impossibilita o acesso dos interessados ao material lingüístico para pesquisas de um rico material, que precisa continuar a ser explorado, tanto para novas cartas lingüísticas de outros volumes dos atlas, quanto para trabalhos de pesquisa, pela comunidade universitária, alunos de graduação, pós-graduação e professores. (VIEIRA, 1998, p. 208)

Pelo desafio de ver concretizado o *Atlas Linguístico do Paraná – ALPR II*, retomamos as discussões de caráter metodológico para a apresentação do registro dos dados coletados por Aguilera entre os anos de 1985-1989, descrevendo e analisando os fatos linguísticos, considerando não apenas o aspecto geográfico, como também o histórico e o social, igualmente estudados no ALPR (1994). Assim, tomando por base as proposições

apresentadas por Aguilera (1990), o ALPR II buscou investigar a possibilidade de identificar em que níveis da linguagem – fonético e/ou lexical – se manifestam mais claramente as diferenças linguísticas regionais; e, ainda, a possibilidade de, por meio de uma nova metodologia, delimitar as diferenças ou as semelhanças linguísticas dos falantes rurais paranaenses.

Outro foco do ALPR II foi inserir nos estudos linguísticos brasileiros os estudos dialetométricos. A Dialetometria, método pouco divulgado no Brasil, mas com tradição na Europa, constitui-se em análises estatísticas de dados coletados pela Geolinguística e permite abordar de maneira mais acurada uma quantidade de dados, auxiliando na obtenção de uma visão geral da variação geográfica das línguas. Os dados coletados para o *Atlas Lingüístico do Paraná – ALPR* (1994), na sua íntegra, foram tratados à luz da Dialetometria e seus resultados analisados na busca de áreas lexicais ligadas ao povoamento do Estado.

Este trabalho também pretendeu estabelecer o grau de influência dos diversos grupos étnicos presentes no Estado em relação aos níveis léxicos e fonéticos constitutivos do falar rural. O Paraná tem, nos movimentos de colonização e povoamento de seu território, o elemento norteador para a divisão em regiões: Tradicional e Moderno.³ O Paraná Tradicional, com povoamento iniciado no século XVII e estendido até o século XIX, tem sua economia baseada, inicialmente, na agropecuária e foi percorrido por indivíduos isoladamente, por bandeiras e por portugueses. O Paraná Moderno subdivide-se em duas regiões: o Norte, no início do século XX, foi ocupado por migrantes paulistas e mineiros e imigrantes europeus atraídos pela atividade cafeeira; o Sudoeste e o Oeste, povoados por migrantes gaúchos e catarinenses em meados do século XX, com economia centrada na plantação de cereais e na criação de suínos. Os aspectos apresentados sobre a colonização apontam para a possibilidade da diversidade linguística do Estado.

O ALPR II conta, ainda, com um glossário com base no levantamento das variantes lexicais mapeadas nas cartas léxico-semânticas. O universo vocabular documentado permitiu organizar um glossário com as 935 variantes lexicais registradas nas questões inéditas das entrevistas para o ALPR (1994) em que constam as anotações sobre a consulta nos dicionários de Moraes Silva (1813) e Houaiss (2001). A utilização desses dois dicionários justifica-se pela importância das obras e por sua data de publicação: 1813 e 2001, respectivamente.

O *Atlas Lingüístico do Paraná*, em sua segunda edição, ainda sem publicação, organiza-se em 2 volumes. O primeiro apresenta a fundamentação teórico-metodológica, as considerações sobre o *Atlas Lingüístico do Paraná – ALPR*, a análise dos dados e o glossário. As cartas linguísticas, segundo volume, encontram-se em CD-ROM. São 175 cartas lexicais e fonéticas, além das cartas introdutórias e duas cartas dialetométricas. Foi necessária, ainda, a constituição de um apêndice no qual estão registradas as matrizes dos dados utilizados na análise dialetométrica e dos anexos que apresentam o questionário para a coleta de dados, o histórico dos 65 pontos de inquiridos no Estado e os dados dos informantes, ambos inseridos no CD-ROM.

O método dialetométrico

Descrever e apresentar os dados tratados à luz da Dialetometria foi outro foco deste trabalho. Essaciência, como já mencionado, utiliza-se dos dados fornecidos pela Dialetologia,

3 Divisão feita por Cardoso e Westphalen (1986).

dentro do método geolinguístico, tanto a tradicional quanto a pluridimensional, e vem somar esforços para o mapeamento dos dados coletados. O método dialetométrico constitui-se, assim, de uma análise global do material coletado pela Geolinguística, o que permite uma visão generalizada dos dados.

Para o estudo dialetométrico dos falares paranaenses, em uma primeira etapa, utilizamos o *corpus* integral do ALPR e, analisado nas relações entre cada par de pontos, o interponto (IP) que, segundo Goebel (1981; 1983), constitui-se no ponto médio do segmento que une cada par de pontos. Cada interponto tem duas funções: a função comunicativa, quando se analisam as semelhanças, e a função discriminatória, quando se analisam as diferenças.

A seleção dos mapas de um atlas linguístico, chamados pela Dialetometria de atributos, e a definição de unidades classificatórias (etimológico, morfossintático, lexical e/ou fonético) compreendem as respostas em cada um dos atributos. Para o estudo, foram agrupadas as respostas dadas pelos informantes nos 65 pontos do ALPR, no que se refere ao léxico, haja vista o número de possibilidades lexicais sugeridas no questionário e verificadas nas entrevistas.

Embora o trabalho de seleção dos dados a serem submetidos ao tratamento dialetométrico tenha tido início com os 325 conceitos investigados pelo instrumento de pesquisa, foi necessária uma triagem. Foram suprimidas as questões que contemplavam respostas abertas – lendas, superstições e narrativas pessoais –, além das questões que continham respostas não verificáveis, isto é, questões do instrumento de coleta de dados que não permitiam a observação imediata e irrefutável do referente.⁴ Ao final dessa primeira seleção, foram escolhidas selecionadas 209 questões que foram tabuladas. Para cada referente investigado adotou-se um código numérico para as respostas obtidas (taxonomia numérica) em que estão anotadas as variantes registradas para cada conceito.

Após essa etapa foi preciso fazer a quantificação das semelhanças, designadas coidentidades, da qual se retirou o índice de semelhança existente entre dois pontos, denominado índice relativo de identidade. As diferenças e codiferenças perfizeram o índice relativo de distância. Desse modo, um ponto linguístico estudado teve seus registros lexicais confrontados com os das demais localidades, utilizando um sistema binário.

A partir dessa segunda tabulação, obteve-se uma observação geral dos conceitos com o número elevado de não respostas, alvo da segunda seleção. O índice estabelecido de não resposta teve como parâmetro o trabalho realizado por Goebel (2003), em que são assinalados 6% para a taxa de não resposta. Segundo ele,

Nous n'avons taxé que des planches – ALF aussi complètes que possible tout en écartant celles qui, sur le total des 638 réponses théoriquement possibles, disposent de moins de 600 entrées transcrites et pourtant utilisables: ceci pour minimiser le danger de distorsions numériques qui résulteraient inévitablement de la prise en compte taxométrique de cartes d'atlas par trop lacunaires.⁵ (GOEBL, 2003, p. 65)

4 Por exemplo: *Quais os tipos de terreno que o(a) Sr.(a) conhece?* e *E quais os tipos de milho que o(a) sr(a) conhece?* (16 questões); e as com respostas com denominação única para o referente, como na questão 156 para as denominações para Cabeça: *Como se chama essa parte aqui? (Apontar)* (93 questões).

5 Tradução livre: Para a análise, foram apenas escolhidos os mapas que apresentavam um número aceitável de respostas no seu todo. Assim, foram eliminados todos aqueles que, num total de 638 respostas, teoricamente possíveis, não atingiram 600 entradas transcritas. Isto para minimizar o perigo de distorções numéricas que, inevitavelmente, resultariam, no processo taxométrico, da inclusão de mapas com demasiadas lacunas.

Para o trabalho com os dados coletados para o ALPR foi estabelecida a taxa de 4,5% para a não resposta e, assim, o número de questões selecionadas chegou a 148 atributos.

Para a etapa seguinte foram selecionados os interpontos, ou seja, as ‘fronteiras’ entre os pontos, seguindo o critério de contiguidade, sem se cruzarem dentro do mapa. Segundo Saramago (1986, p. 2),⁶ “o critério de união é o da menor distância geográfica, portanto os que estão mais próximos entre si”, obtendo, assim, a relação dos interpontos, base para encontrar os índices relativos. É a partir desses dados que se constitui a matriz simétrica e se constroem os mapas dialetométricos. Os Índices Relativos de Diferenças (IRD) e os Índices Relativos de Identidade (IRI) para os pontos são calculados a partir do número de semelhanças e diferenças encontradas.

Identificados os interpontos, a fase subsequente consistiu-se em obter uma média geral do IRD e, a partir dessa média, obter uma mediana do IRI. Segue-se a distribuição dos valores denominada de intervalação. Para a distribuição dos valores do IRI e IRD nas cartas, Goebel (1983)⁷ emprega o algoritmo MINMWMAX (média) para o IRD e o algoritmo MED, MEDMW (mediana) para o IRI. Os resultados dessas operações foram divididos por três, como recomenda Goebel:

[...] les écarts numériques entre le minimum et la moyenne arithmétique et entre le maximum et la moyenne arithmétique sont respectivement divisés par trois. Les valeurs ainsi déterminées constituent les tailles des intervalles situés au-dessous et au-dessus de la moyenne arithmétique.⁸ (1981, p. 362)

Dessa forma, expõem-se o número de interpontos em cada intervalo e o símbolo utilizado para a cartografia do Índice Relativo de Distância nas 65 localidades do Paraná. Da mesma forma, os valores da intervalação para o estudo das identidades são encontrados por meio do valor mediano das porcentagens e seguem a sistematização feita para os índices de distância e submetidos à divisão por três, como propõe Goebel (2003).

Para a cartografia dos dados numéricos obtidos utilizamos dois suportes visuais. O primeiro suporte, designado por Goebel *parquet triangulaire*, é obtido a partir da união linear entre os pontos da rede em que são visualizados os valores de IRI. O segundo suporte, designado pelo mesmo autor como *parquet polygonal*, é construído sobre o primeiro, através da aplicação do método dos polígonos de Thiessen: para cada linha do *parquet triangulaire* é traçada a respectiva mediatriz. Os pontos de encontro entre cada par de mediatrizes são os ângulos de união de cada um dos lados do polígono onde será visualizada a relação interpontual entre um determinado ponto e aqueles que lhe são vizinhos. Para este estudo, os dois suportes mencionados foram desenhados pelo professor Dr. Xulio Sousa Fernández, do Instituto de Língua Galega da Universidade de Santiago de Compostela – Espanha.

6 Artigo cedido pelo autor.

7 Cf. também Goebel (1983; 1987) e Saramago (1986).

8 Tradução livre: As distâncias numéricas entre o mínimo e a média aritmética e entre o máximo e a média aritmética são, respectivamente, divididas por três. Os valores assim determinados constituem recortes dos intervalos situados abaixo e acima da média aritmética.

Olhando para os dados

Seguindo os pressupostos da Dialetoimetria, os dados, para o IRD – Índice Relativo de Distância, e o IRI – Índice Relativo de Identidade, estão sistematizados nas duas cartas que seguem. A carta da figura, reproduzida a seguir, registra o IRD, e a da Figura 2, o IRI.

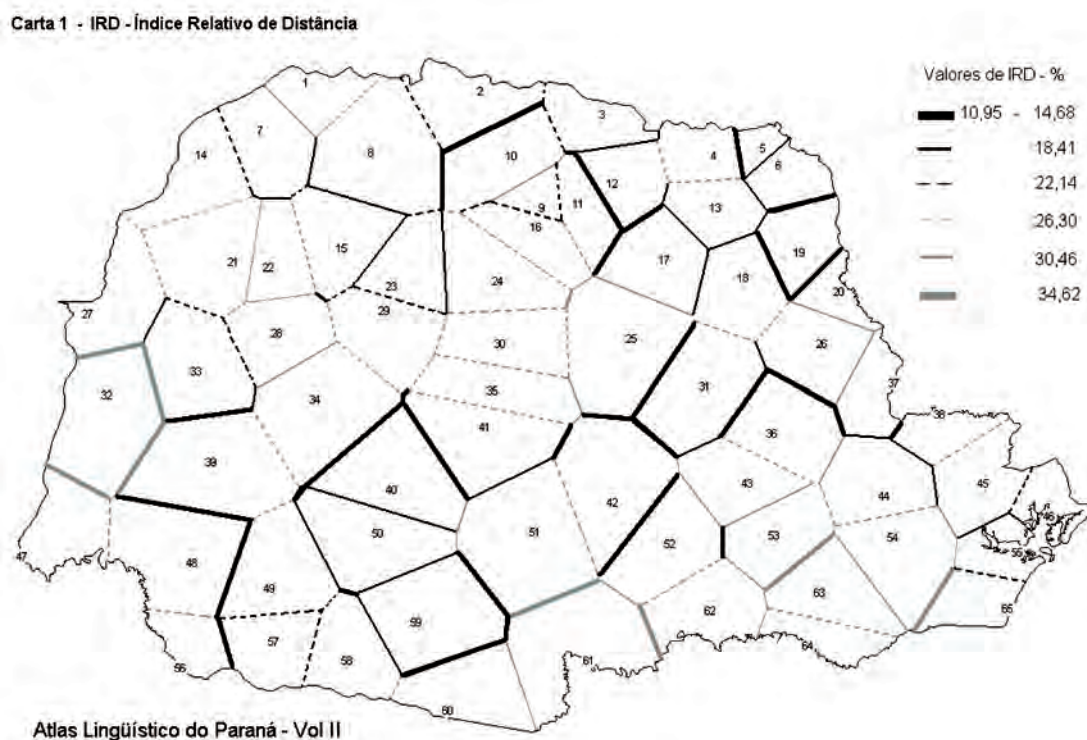


Figura 1 – IRD – Atlas Lingüístico do Paraná – ALPR II

Os dois mapas dialetométricos expostos representam as sínteses dos dados para os 148 atributos selecionados e a leitura deve ser feita seguindo a legenda em que estão expostos do menor ao maior índice de diferença. Quando transportadas para a carta, essas informações demarcam, por exemplo, as áreas em que há menor contato linguístico, ou seja, um maior índice de diferenças entre os pontos estudados.

Quantificando os dados da figura 1 – IRD –, obtiveram-se 72 interpontos (47%) acima da média e 81 (53%) abaixo da média. Dentre os valores acima da média, destacam-se as fronteiras das localidades de São Pedro do Ivaí (ponto 24), Ortigueira (25), União da Vitória (61), Guarapuava (51), São Mateus do Sul (62), Lapa (63) e Palmeira (53). Sobreleva-se dos demais pontos em que há maior índice de distância, a localidade de Marechal Cândido Rondon (32), que tem na maioria de suas fronteiras – com os municípios de Guaíra (27), Assis Chateaubriand (33), Cascavel (39) e Foz do Iguaçu (47) – índices elevados de distância. Ainda sobre a quantificação dos dados na carta 367, são encontrados índices intermediários de distância entre os pontos, em sua maioria, da região noroeste. Os índices de menor distância encontram-se, também em sua maioria, entre os pontos do Norte Pioneiro. Ressalta-se, entre os pontos com índices de menor distância, a cidade de Siqueira Campos (19) com Jacarezinho (6), Ibaiti (18) e São José da Boa Vista (20).

Para a leitura dos dados constantes da Figura 2 – IRI, veem-se expostos do menor ao maior índice de identidade. Observa-se, assim, uma afinidade lexical no noroeste e no

centro-oeste paranaense. Salientam-se os índices de identidade entre o município Francisco Beltrão (57) e seus circunvizinhos, Manguerinha (59), Dois Vizinhos (49) e Pato Branco (58).

Carta 2 - IRI - Índice Relativo de Identidade

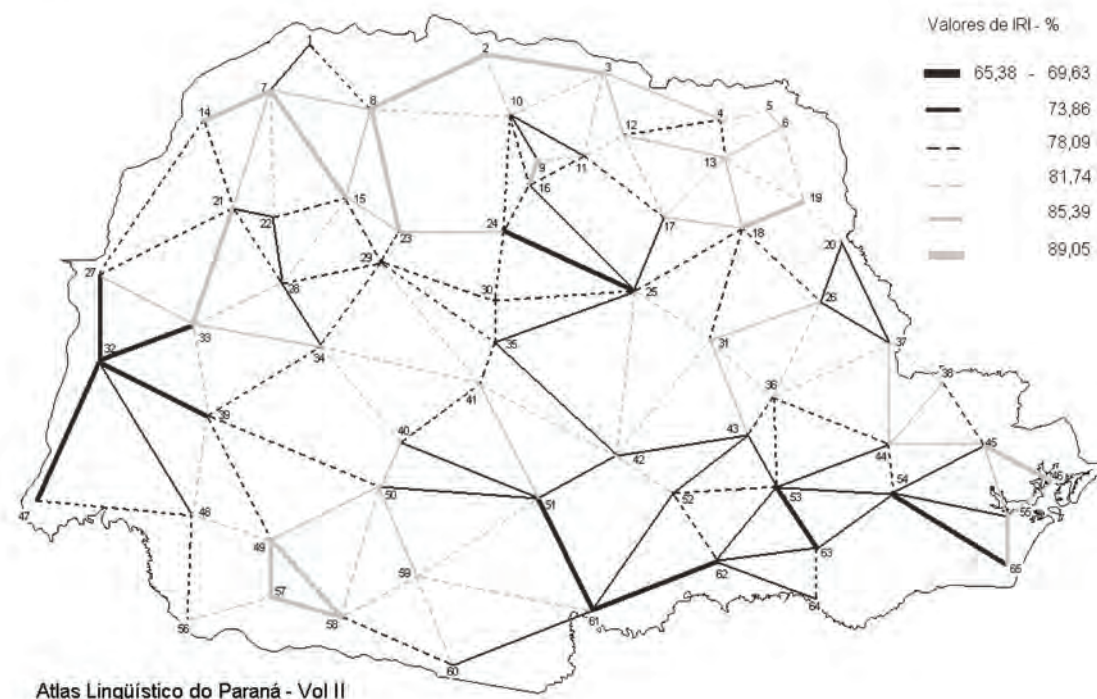


Figura 2 – IRI – Atlas Lingüístico do Paraná – ALPR II

Na constituição da população de cada localidade refletida no perfil dos informantes pode residir a distância encontrada na linguagem. Dessa forma, é preciso uma observação mais detida sobre os dados registrados nas cartas.

Na carta 367 (IRD) chamam a atenção os dados sobre o município de Marechal Cândido Rondon (32) e seus circunvizinhos Guaíra (27), Assis Chateaubriand (33), Cascavel (39) e Foz do Iguaçu (47). Provavelmente, a ilha de isolamento linguístico que se formou no ponto 32 seja decorrente da sua formação étnica. É preciso historiar a formação da população e o perfil do informante em cada município.

Marechal Cândido Rondon tem, na formação de seu povo, imigrantes de Santa Catarina e do Rio Grande do Sul, descendentes em sua maioria de alemães, como é o caso dos informantes deste atlas (nascidos em Santa Rosa – RS). O outro município, Guaíra, já pertenceu ao Paraguai (século XVI) sob domínio espanhol e teve como primeiros habitantes nativos de várias tribos. No início do século XVII, a Província Del Guayrá passa ao domínio português. Cruzaram por essas terras bandeirantes e padres jesuítas estabelecidos nas reduções, posteriormente abandonadas e destruídas. Já Assis Chateaubriand teve seu território considerado como terras devolutas, ficando entre os anos de 1950 e 1958 em poder da justiça. Após esse período, a Sociedade Colonizadora Norte do Paraná S/A reinicia a colonização, atraindo agricultores de vários pontos do país, principalmente de Minas Gerais. Cascavel teve como primeiros colonos os imigrantes vindos de Guarapuava e, posteriormente, muitos outros sulistas de origem alemã e italiana, atraídos pela agricultura, que lá iniciaram

o cultivo de cereais e a criação de rebanhos de suínos. Por fim, Foz do Iguaçu, cidade fronteiriça, teve a primeira instalação militar em 1889, quando era habitada por paraguaios, brasileiros e argentinos.

Sobre os dados da carta 367 para o menor índice de diferença, cabe ressaltar o município de Siqueira Campos (19) com Ibaiti (18), São José da Boa Vista (20) e Jacarezinho (6), que foram colonizados por paulistas e mineiros que ali se estabeleceram, principalmente com a plantação do café e que tiveram, para este atlas, informantes nascido na localidade.⁹ Ressaltamos, também, sobre os índices de identidade encontrados na carta 368, os municípios Francisco Beltrão e seus circunvizinhos, já mencionados anteriormente. Essas cidades foram colonizadas, em sua maioria, por gaúchos e catarinenses.

Os dados das cartas sintéticas (367 e 368) acima reportam à história da colonização paranaense. A constituição do povo paranaense, que foi inicialmente formado de portugueses, negros e índios concentrados no litoral do Estado, passou, segundo Cardoso e Westphalen (1986), por três ondas povoadoras. Na primeira, durante os séculos XVI e XVII, portugueses e espanhóis, por meio dos bandeirantes paulistas sob ordem do governo português e das missões jesuíticas espanholas, buscaram a posse do território paranaense, dando início a inúmeros núcleos de povoamento ao longo dos principais rios.

Por volta de 1640, tinha sequência o povoamento do litoral, provavelmente iniciado no século anterior, por aventureiros em busca de jazidas de ouro. Nessa mesma época, chegava a Paranaguá o capitão de infantaria Gabriel de Lara, com a missão de resguardar a costa paranaense dos ataques estrangeiros. A descoberta de ouro, embora escasso, atraía moradores para Curitiba, os quais, ao lado da mineração, desenvolviam a agricultura e a criação de gado. A expansão da pecuária determinou a posse das terras circundantes em que se formaram núcleos de povoamento nas regiões de Antonina, Campo Largo, Palmeira e Castro.

Em 1711, por ocasião da venda à Coroa Portuguesa das terras que constituíam a capitania de Paranaguá, de propriedade do Marquês de Cascais, o Paraná passou a integrar a Capitania de São Paulo como a Quinta comarca. Nesse ínterim, por iniciativa particular, continuavam as apropriações de terras dos Campos Gerais, chegando a Ponta Grossa, Tibagi e Jaguariaíva, por volta de 1719.

Em 1727, teve início a abertura da estrada do Viamão-RS, “objetivando a introdução das cavalgadas do sul em direção a São Paulo” (CARDOSO; WESTPHALEN, 1986, p. 44). Tal evento vai determinar o surgimento de povoações em torno das fazendas de pouso, no sudeste paranaense, na direção sul-norte. Com o Tratado de Madri (1750), que reconhecia o direito de ocupação e posse das terras, expedições militares aventuram-se em busca da conquista do interior do estado paranaense.

Em 1829, chegaram a Rio Negro os primeiros colonos alemães e, em 1839, teve início a ocupação e povoamento dos Campos de Palmas. Vieram somar-se, segundo Rodrigues (2007, p. 89), aos imigrantes alemães, franceses, suíços, ingleses e italianos. Por iniciativa governamental, novos grupos de colonos foram fixados próximos às localidades já povoadas, rumo ao interior do Paraná. Essa ação propiciou uma alteração no quadro demográfico, com grupos de imigrantes de alemães de origem russa no Planalto Central paranaense.

⁹ Apenas a informante de Siqueira Campos nasceu na cidade vizinha à cidade pesquisada.

Entre 1840 e 1860, verificou-se o “apogeu do comércio de tropas de mulas conduzidas sobretudo por tropeiros paranaenses, que passam e invernam nos campos do Paraná, com destino às feiras de Sorocaba” (CARDOSO; WESTPHALEN, 1986, p. 56). O Estado se emancipa de São Paulo e a cidade de Curitiba é escolhida como capital de Província do Paraná (1853). Entre 1860 e 1900, multiplicaram-se as colônias formadas por imigrantes poloneses, italianos, alemães, russos alemães, ucranianos e, por último, sírios e libaneses. É nessa mesma época que se iniciaram os núcleos da Colônia Mineira ou Tomazina do Norte Pioneiro, com fazendeiros paulistas e mineiros dedicados à plantação de café.

Nessa mesma época, Curitiba ganhava grande movimentação com a extração e comercialização da erva-mate, e passava por transformações de natureza vária, tanto na cultura, como na política, no comércio e na indústria. Em 1888, a fundação da colônia militar em Foz do Iguaçu “registra o ponto de partida para a ocupação e colonização daquela região” (CIGOLINI; MELLO; LOPES, 2001, p. 72). Por volta de 1900, segundo Cardoso e Westphalen (1986, p. 60), o governo imperial concedeu à Estrada de Ferro São Paulo – Rio Grande do Sul, subsidiária da Brazilian Railways Company, terras devolutas nacionais, no sudoeste, oeste e norte do Paraná.

De 1900 a 1940, novas concessões de terras foram realizadas nas várias regiões paranaenses, determinando grandes repercussões na indústria da ocupação do Paraná moderno. As terras, divididas em lotes coloniais de cerca de 15 alqueires e lotes urbanos, eram vendidas também a prestação, medida que atraiu principalmente migrantes paulistas, mineiros e nordestinos, além de imigrantes de variada origem e procedência.

A Paraná Plantation Limited, com sede em Londres, sucedida pela Companhia de Terras do Norte do Paraná, hoje Companhia Melhoramentos do Norte do Paraná, foi uma das concessionárias que recebeu 515 mil alqueires das terras do norte paranaense formando novas fazendas, colônias e cidades tanto de imigrantes quanto de brasileiros. Os interesses econômicos estavam diretamente ligados à cafeicultura. Foi muito significativa, nessa época, no norte paranaense, a onda da imigração japonesa.

Nas décadas de 20 a 40 do século passado, teve início em grande escala a corrente povoadora de agricultores de origem alemã e italiana vinda do sul do país, que se estabeleceu nas matas do atual território localizado entre Pato Branco e Cascavel.

Outro fator importante é a abertura de estradas de rodagem que vieram proporcionar, pouco a pouco, a integração do norte com o sul do Estado, sobretudo com o escoamento da produção cafeeira pelo porto de Paranaguá.

A última fase de ocupação territorial do Paraná deu-se entre 1940 e 1960, nas regiões noroeste, ou Norte Novíssimo, e oeste, com a concessão de terras devolutas, pelo governo e por companhias privadas de colonização. É possível fazer a comparação com os mapas 367 e 368 com essa fase de povoação do Estado e detectar a afinidade lexical entre os pontos nessa área. Assinala-se, ainda, a entrada de colonos do Rio Grande do Sul e de Santa Catarina que, partindo de Pato Branco, rumaram para os vales dos rios Chopim, Piquiri e Paraná. Esses reemigrantes (alemães e italianos, principalmente) que se instalaram no oeste paranaense conferem, em especial ao município de Marechal Cândido Rodon (32), um possível isolamento linguístico.

Ainda sobre a mesma época, segundo Cigolini, Mello e Lopes (2001, p. 74), “colonos europeus continuaram a estabelecer-se nas proximidades de Curitiba e no centro-sul do

estado, formando novos núcleos colonizadores: Irati, Mallet, Prudentópolis e São Mateus do Sul”. Esses novos núcleos¹⁰ apresentados nos mapas de Índices Relativos de Distância – IRD (367) e Índice Relativo de Identidade – IRI (368) refletem a formação étnica de seu povo, registrando índices acima da média para as diferenças.

A partir da história da ocupação do Paraná, em que se evidencia a presença de três ondas povoadoras, é possível encontrar, sob conjunturas diversas e com motivações distintas, ressaltada constituição étnica do Estado como um dos prováveis responsáveis pelas similaridades e diferenças dos falares paranaenses.

À guisa de uma conclusão

O ALPR II teve como objetivo geral a cartografia dos dados inéditos coletados para o *Atlas Lingüístico do Paraná* (1994), com vistas ao registro e à disseminação do acervo linguístico coletado. Por meio das cartas constantes do Atlas, muitos estudiosos poderão se servir dos dados para suas pesquisas, permitindo o entendimento dos usos linguísticos dentro do contexto social da comunidade estudada.

Dessa forma, reafirma-se a importância dos estudos dialetológicos para a preservação, a compreensão e a disseminação dos fatos linguísticos, sejam eles registrados pela Dialectologia tradicional, sejam documentados pela dialetologia pluridimensional, que observa também diversos outros aspectos (diastrático, diageracional, diassexual...) influenciadores da linguagem.

O ALPR II, com suas 175 cartas – 125 lexicais e 50 fonéticas –, é o inventário de algumas formas linguísticas da modalidade oral da língua. A utilização dos dados coletados por meio das entrevistas realizadas para o ALPR (1994) alcança, hoje, 94,5% dos dados. É importante ressaltar que os 5,9% ainda não cartografados (pela natureza aberta das questões) estão servindo de fonte para inúmeros trabalhos monográficos. A retomada, 18 anos depois dos dados coletados entre 1985 e 1989, não inviabiliza a sua cartografia. Ao contrário, aponta para o rigor científico da coleta dos dados que reafirma a possibilidade de análise da documentação. Muitos estudos ainda poderão ser desenvolvidos a partir dos dados já cartografados (ALPR e ALPR II) ou inéditos, seja pela aplicação de novas metodologias, seja pela incorporação de novas coletas de dados às já feitas, demonstrando, segundo Cardoso (2002, p. 164), a necessidade para o “conhecimento efetivo de uma dada realidade lingüística.”

Observa-se, nas cartas do ALPR II, o registro de uma grande produtividade lexical para a denominação do mesmo conceito. Trata-se de um elenco complexo que abarca desde formas inovadoras, seja da linguagem padrão seja de criação popular recente, até formas antigas presas ao português europeu, passando pelos regionalismos, as ocorrências únicas (Hapax legomena) e as formas populares disseminadas pelo português oral de outras regiões do país. Nos dados analisados tem-se, ainda, a manutenção da linguagem popular arcaica, as variantes motivadas pelo fator diatópico, ou seja, são encontradas em uma determinada região dentro do Estado. Há, também, cartas nas quais a variação não ocorre e que expressam, dessa forma, o caráter homogêneo da fala no Paraná.

É preciso assinalar, ainda, a importância do glossário organizado com base no levantamento, feito a partir das cartas léxico-semânticas, das variantes lexicais registradas.

¹⁰ Dos municípios citados, fazem parte da rede de pontos do ALPR: Irati (ponto 52), Prudentópolis (42) e São Mateus do Sul (62).

Sobre a utilização dos métodos, é preciso ressaltar que tradicionalmente, no Brasil, o método geolinguístico vem sendo utilizado para a apresentação dos dados registrados nas entrevistas *in loco*. Essa distribuição espacial das variantes lexicais ou fonéticas tem prestado uma valiosa contribuição para o conhecimento da língua falada em diversas regiões do país.

O método dialetométrico constitui-se, por sua vez, de uma análise global dos dados coletados pela Geolinguística, o que permite uma visão generalizada desses dados, as contribuições dessa metodologia são as de fornecer um quadro sintético do comportamento linguístico no Estado do Paraná.

A aplicação de tais métodos – geolinguístico e dialetométrico – ao *corpus* coletado para o ALPR (1994) é complementar, uma vez que cada uma das metodologias adotadas consagra-se por sua abrangência: o método geolinguístico, por permitir a visualização das particularidades de cada ponto de inquérito e de cada informante; e o método dialetométrico, por considerar todo o volume de dados e sintetizá-los por meio das cartas que demonstram os índices de identidade e de distância entre os pontos de um atlas. As cartas dialetométricas, constituídas de análises estatísticas de dados coletados pela geolinguística, evidenciam a diversidade linguística e, neste trabalho, a sua relação com a história da colonização paranaense, em que se estabelece o grau de influência dos diversos grupos colonizadores do Paraná em relação ao nível léxico no falar rural.

Enfim, pretendeu-se contribuir para o conhecimento da língua portuguesa falada no Paraná, no que se refere à sua constituição e às pistas que podem levar ao conhecimento de sua história.

REFERÊNCIAS

AGUILERA, V. de A. *Atlas Linguístico do Paraná – ALPR*. Curitiba: Imprensa Oficial do Estado, 1994.

_____. *Atlas linguístico do Paraná*. 1990. 426p. Tese (Doutorado em Letras) – Universidade Estadual Paulista Júlio de Mesquita Filho, Assis. 1990. 2 v.

ALVAR, M. (Dir.) *Manual de dialectologia hispânica: el español de España*. Barcelona: Ariel, 1996.

AMARAL, A. *O dialeto caipira*. 2. ed. São Paulo: HUCITEC/ Secretaria da Cultura, 1976.

CAMARA JR, J. M. *Dicionário de lingüística e gramática*. Petrópolis: Vozes, 1981.

CARDOSO, J. A.; WESTPHALEN, C. M. *Atlas histórico do Paraná*. 2. ed. rev. e ampl. Curitiba: Livraria Chain, 1986.

CARDOSO, S. A. M. *Atlas lingüístico de Sergipe II*. Salvador: EDUFBA, 2005.

- _____. La dialetologie au Brasil aperçu historique et bilan actuel. In: CONTINI, M. (Org.). *La géolinguistique en Amérique latine*. Hors série n. 2 de Géolinguistique. Grenoble - FR: Université Stendhal, Grenoble 3, 2002. p. 197-229.
- CHAMBERS, J. K.; TRUDGILL, P. *La Dialectología*. Madri: Visor Libros, 1994. p. 19-32.
- CIGOLINI, A.; MELLO, L. de; LOPES, N. *Paraná: quadro natural, transformações territoriais e economia*. 2. ed. rev. e ampl. São Paulo: Saraiva, 2001.
- CUNHA, S.; SILVA NETO, S. da. *Guia para estudos dialectológicos*. 2. ed. Belém: Conselho Nacional de Pesquisa/Instituto Nacional de Pesquisas da Amazônia, 1957.
- DUBOIS, J. et al. *Dicionário de lingüística*. São Paulo: Cultrix, 1978.
- GOEBL, H. Regards dialectométriques sur les données de l'Atlas Linguistique de la France (ALF): relations quantitatives et structures de profondeur. *Estudis Romànics*, Barcelona, Institut D'Estudis Catalans, v. 25, p. 59-96, 2003.
- _____. Parquet polygonal et treillis triangulaire: les deux versants de la dialectométrie interponctuelle. *Revue de Linguistique Romane*, Strasbourg, tome 47, p. 353-412, 1983.
- _____. Eléments d'analyse dialectométrique avec application à l'AIS. *Revue de Linguistique Romane*, Strasbourg, tome 44, p. 349-420, 1981.
- GUARNER, M. S. *La cartografía lingüística en la actualidad y el atlas de la Península Ibérica*. Madrid: Instituto Miguel de Cervantes, 1953.
- HOUAISS, A. *Dicionário Eletrônico Houaiss da Língua Portuguesa*. Rio de Janeiro: Objetiva, 2001.
- _____. *Sugestões para uma política da língua*. Rio de Janeiro: Ministério da Educação e Cultura/Instituto Nacional do Livro, 1960.
- MORAES SILVA, A. de. *Diccionario da lingua portugueza*. Lisboa: Typographia Lacérdina, 1813. (edição digitalizada)
- NASCENTES, A. *Bases para a elaboração do Atlas Lingüístico do Brasil: questionário*. Rio de Janeiro: MEC/Casa de Rui Barbosa, 1961.
- _____. *Bases para a elaboração do Atlas Lingüístico do Brasil*. Rio de Janeiro: MEC/Casa de Rui Barbosa, 1958.
- RODRIGUES, R. E. de S. B. *Em busca de uma história para o léxico rural paranaense*. 2007. 378p. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina. 2 v.

SARAMAGO, J. A. das P. Diferenciation lexicale: um essai dialectometrique applique aux materiaux portugais de l'A.L.E. *Bulletin du Centre de Dialectologie*, Grenoble, v. 2, p. 1-31, 1986. [cópia cedida pelo autor]

VIEIRA, H. G. Fundamentos para organizar, implementar e manter um banco de dados geolingüístico. In: AGUILERA, V. de A. (Org). *A geolingüística no Brasil: caminhos e perspectivas*. Londrina: Eduel, 1998. p. 207-224.