

MODELAÇÃO DE DADOS E PRODUÇÃO DE “REALIDADES” NA ANÁLISE SOCIOLINGUÍSTICA¹

Gredson dos SANTOS²

RESUMO: O texto faz considerações críticas sobre o tratamento quantitativo dos dados no interior da sociolinguística e seu impacto na formulação de “realidades” no que tange ao espaço linguístico brasileiro. O trabalho comenta brevemente como as operações de seleção, tratamento, análise e interpretação de dados, segundo uma metodologia quantitativa, “formatam”, em distintos graus de intensidade, o “real” que está no escopo da análise sociolinguística.

PALAVRAS-CHAVE: Filosofia da Ciência. Filosofia da Linguística. Sociolinguística.

Introdução

A filosofia da linguística é uma área que tem como principal preocupação a relação entre as diversas teorias linguísticas e o seu objeto de estudos. Em outras palavras, uma das tarefas que a filosofia da linguística impõe a si é verificar *criticamente* até que ponto as formulações das teorias produzidas no interior dos estudos linguísticos se aproximam ou se afastam do “real” que pretendem investigar. Esse campo de estudos enquadra-se numa área mais

1 Este texto é fruto de algumas indagações do autor decorrentes das discussões travadas na disciplina Seminários Avançados III, oferecida pelo Programa de Pós-graduação em Letras e Linguística da Universidade Federal da Bahia (UFBA) e coordenada pela professora Dra. Rosa Virgínia Mattos e Silva. As reflexões apresentadas são absolutamente provisórias. Embora o texto resulte das discussões realizadas nas aulas de Mattos e Silva e do professor Dr. Carlos A. Faraco, e embora este último tenha feito algumas sugestões de abordagem na linha da epistemologia da linguística, os equívocos presentes no trabalho são, evidentemente, de minha inteira responsabilidade.

2 Programa de Pós-Graduação em Letras e Linguística, Instituto de Letras, UFBA, Salvador, Bahia, Brasil. Professor Assistente de Linguística do Centro de Formação de Professores, da Universidade Federal do Recôncavo da Bahia (UFRB). gredsons@bol.com.br

ampla do conhecimento, a filosofia da ciência, que se volta para a natureza das teorias e práticas científicas e a sua relação com os resultados que produzem. No âmbito da linguística brasileira, as discussões acerca da natureza das várias teorias que aqui frutificaram ainda são tímidas. No entanto, a reflexão crítica acerca do trabalho científico é um imperativo ético e acadêmico: ético porque diz respeito, necessariamente, ao comprometimento do cientista com os valores da transparência e publicidade de seu trabalho, além do seu compromisso socialmente assumido de contribuir para a compreensão de fatos do mundo; acadêmico porque a revisão crítica do trabalho conduz a resultados mais condizentes com a “verdade” buscada pela sua prática, além de refinar o alcance das formulações feitas no interior de dada teoria. Aí está situada a relevância do trabalho que aqui se propõe, na medida em que este busca produzir uma pequena reflexão sobre um dos modos – dentre tantos outros – de operar com o fato linguístico.

O objetivo geral deste texto é tecer considerações críticas sobre o tratamento quantitativo dos dados no interior da análise sociolinguística e seu impacto na formulação de “realidades” no que tange ao espaço linguístico brasileiro. O trabalho pretende ainda comentar sucintamente como as operações de seleção, tratamento, análise e interpretação de dados, segundo uma metodologia quantitativa, “formatam”, em distintos graus de intensidade, o “real” que está no escopo da análise sociolinguística.

Segundo aponta Borges Neto (2004, p. 9), uma vez que a filosofia da linguística (às vezes chamadas também de epistemologia da linguística) precisa de “boas descrições” das teorias linguísticas, esse campo de saber se vincula estreitamente à história/historiografia da linguística. O presente trabalho, ao propor uma reflexão crítica sobre uma técnica de análise sociolinguística, não é propriamente um trabalho de sociolinguística.

Nosso posicionamento vincula-se a uma formulação produzida no interior da filosofia heideggeriana. Em *Ciência e pensamento do sentido* (2002), o filósofo alemão, desdobrando sua afirmação de que “a ciência é a teoria do real”, faz uma reflexão sobre o sentido das palavras “teoria” e “real”, mostrando com isso que *teoria* pode ser entendida: 1) como observação (contemplação), que prescindiria de qualquer elaboração do real – o que permitiria falar em ciência pura, desinteressada e sem propósito; 2) no sentido de tratar, empenhar-se, trabalhar – sentido mediante o qual “a ciência é uma elaboração do real terrivelmente intervencionista” (HEIDEGGER, 2002, p. 48). Nesse particular, importa

lembrar que o filósofo distingue, em sua reflexão, o modo grego de pensar o fazer científico (contemplar o vigente em sua vigência) e o modo moderno. Assumindo que o segundo conceito é o que marca a ciência moderna, Heidegger postula que “a ciência põe o real”, transformando-o em objetividade – um real capaz de ser submetido a um conjunto de operações e processamentos, isto é, numa sequência de causas aduzidas que se podem prever. Mais adiante em sua reflexão, assegura que qualquer real que tenha sua “objetividade” processável passa a ser o traço fundamental de representação produzida pela ciência moderna – a ciência moderna torna o real objetividade, recorta-o de modo a poder calculá-lo. Mas, um traço característico desse proceder da ciência é justamente que, para que a ciência se produza, ela precisa de um método, que não pode ser questionado pelo cientista em sua atividade como cientista (o cientista deve seguir o método, deve submeter-se a ele) – a não ser que se afaste dessa posição. Por isso é que Heidegger afirma na sequência de sua exposição: “Nenhuma física tem condições de falar da física, como física [...] o mesmo vale para a filologia. Na condição de teoria da língua e de literatura, a filologia nunca poderá ser objeto de um exame filológico. É o que vale para toda a ciência” (2002, p. 55).

Assumindo como pressuposto válido as formulações presentes em Heidegger (2002 e 2008), Borges Neto (2004), Freire-Maia (2008), entre outros, proceder-se-á aqui a uma discussão bem geral em torno de certos aspectos da análise sociolinguística, mormente no que diz respeito a noções como as de “fatores sociais”, “vernáculo” e “norma culta” e a acomodação das mesmas em análises estatísticas.

Breves considerações em torno do conceito de “verdade” e de sua relação com a ciência³

Uma consulta a dicionários etimológicos associa o significado da palavra latina *veritas* ao sentido que costumamos atribuir à palavra *verdade*, que diz da qualidade pela qual as coisas se apresentam tais como são. Além disso, *verdade* expressa uma conformidade do que se diz com o que é. No âmbito da

3 As considerações que se fazem aqui não são uma formulação do autor deste texto, que assume aqui sua limitação em produzir uma reflexão mais profunda no campo em que ora se aventura. As palavras escritas nesta seção são antes anotações de leituras feitas de obras que tratam do assunto, mormente Heidegger (2002 e 2008), Freire-Maia (2008) e Borges Neto (2004). Saliente-se ainda que as noções expostas a seguir importam ao trabalho na medida em que o autor tentará relacioná-las a alguns aspectos do modelo sociolinguístico de análise linguística.

filosofia em geral e da filosofia da ciência em particular, o conceito de verdade não é algo de fácil definição, justamente porque é um conceito associado à transcendência, é uma noção primária e intuitiva, aut-evidente e ao mesmo tempo misteriosa, assim como as noções do *belo*, do *bem* e do *ser*.

É Freire-Maia (2008, p. 27-28) que nos aponta que, embora derivada de *veritas*, a palavra *verdade* pode ser empregada em três dimensões distintas:

1) a que está ligada à própria origem latina: uma narração fiel ao que aconteceu. *Veritas*. Neste caso, se dizemos “É verdade que Paulo viajou”, isso realmente ocorreu; Paulo não está mais entre nós;

2) outro sentido provém da palavra correspondente em grego: *aleteia*. Neste caso, a verdade resulta de um descobrimento. A verdade é como a revelação de algo que estava encoberto. “Este livro é de linguística...” significa que se trata de um livro que aborda um assunto em particular e não qualquer um; algo que foi constatado no momento em que se examinou o livro;

3) o outro sentido está relacionado com a palavra hebraica *emunah*. Algo verdadeiro e confiável. “Fulano não é meu amigo de verdade” significa que, se posto à prova, não se mostra amigo, não é confiável.

Ainda segundo Freire-Maia, esses três aspectos da *verdade* podem ser assim tratados. Se nos referimos a algo passado e sua descrição nos permite prever o que virá a acontecer no futuro, está-se usando o *emunah*, que diz daquilo em que se pode confiar; de algo cujo comportamento esperado será confirmado. Se a descrição está relacionada a algo que exatamente aconteceu, então é o *veritas* latino que aparece. Quando, por outro lado, se conhece algo que não se conhecia antes, ou quando algo que não era visto passa a ser, então ocorre a verdade *aleteia*.

Há ainda duas outras maneiras de se tentar precisar o conceito de *verdade*. Neste particular, pode-se afirmar que há dois tipos de *verdade*: as lógicas e as factuais. Uma verdade lógica independe de fatos. Se se diz que $2 = 2$, está-se enunciando uma verdade lógica. Se o termo “variantes” é empregado para se referir a duas coisas diversas, mas de mesma natureza, como duas formas iguais de se dizer a mesma coisa, sejam elas quais forem, pode-se assumir isso como uma verdade lógica, tal como $(2+4) = (3+3)$. As verdades factuais são verdades empíricas. Sua validade advém da verificação do que se diz e de sua correspondência com o que é ou com o que aconteceu. São exemplos de verdades factuais: “o gavião voa”; “mamíferos não põem ovos”; “a língua portuguesa deriva da língua latina”.

Heidegger (2008), em sua *Introdução à filosofia*, especialmente na primeira seção do livro, em que discute as relações entre filosofia e ciência, atribui a essência da verdade e o seu pertencimento ao *ser-aí*. O ser-aí é, em linhas bem gerais, aquele que tem como essência a qualidade de poder, na acepção do autor, “ser-um-com-o-outro”: o homem, que diferente dos objetos sem vida (classificados pelo filósofo como entes por si subsistentes), é um ser cuja natureza implica o compartilhamento da verdade, o desvelamento do ente. Em sua constituição ôntica, o ser-aí é um ser descobridor, ao passo que a verdade do ente por si subsistente é o ter sido descoberto (HEIDEGGER, 2008, p. 127-8). Note-se que, nesta preleção, o filósofo levanta a discussão acerca da essência da verdade em virtude de pretender abordar a essência do conhecimento científico.

Heidegger aponta ainda que o conceito escolástico de verdade – caracterizada como adequação do intelecto às coisas – está associado à posição que encara a verdade como uma verdade proposicional. De acordo com esse entender, a verdade é uma verdade do juízo, do enunciado, na medida em que este tem o caráter de mostrar as coisas. Cabe lembrar que é esse conceito de verdade que é tradicionalmente associado à prática científica. Disso implica que se a verdade reside mesmo na proposição e a ciência pode ser definida como um conhecimento que busca a verdade, então ela, a ciência, é uma conexão de proposições verdadeiras (HEIDEGGER, 2008, p. 50).

Mas Heidegger, mesmo reconhecendo que essa concepção de verdade goza de prestígio no interior da história da filosofia, lembra que nem todo juízo mostra e levanta a seguinte ponderação:

Mas, se a ciência deve ser o encontro e a determinação da verdade, então se torna simultaneamente questionável se o conceito basilar de verdade como verdade proposicional é suficiente para esclarecer a essência da verdade científica. Talvez resida até mesmo na caracterização da verdade como verdade proposicional e na determinação da ciência a partir de seu resultado um único e mesmo erro fundamental. Por meio de uma apreensão mais radical da essência da verdade, precisamos nos colocar em condições de, desde o começo, ver também a essência da ciência de modo mais originário. Precisamos chegar até o ponto em que evitemos desde o principio tomar a ciência como um conteúdo proposicional. (HEIDEGGER, 2008, p. 51-2)

Como se vê, o autor associa o entendimento da essência da ciência à própria busca da essência da verdade, na medida em que se admite que o conhecimento científico busca atingir a verdade. Para isso, o autor elabora uma crítica ao conceito tradicional de verdade como uma proposição, como um juízo. Ele

não nega que a verdade exista. De maneira bem direta, afirmará que “verdade significa desvelamento do ente” (HEIDEGGER, 2008, p. 82). Entretanto, como há entes de diversos modos de ser, existem correspondentemente diversas “modulações da verdade” (HEIDEGGER, 2008, p.112). Com tal raciocínio, o autor conclui que uma determinação essencial da verdade é o fato de que ela não reside na proposição, mas no ser-aí (HEIDEGGER, 2008, p. 114).

Evidentemente que a noção de verdade, especialmente quando correlacionada à ideia de ciência, implica também pensar o que vem a ser o “real”. Se pensarmos à maneira platônica, o real são as ideias. Esse modo de pensar implica rejeitar a realidade de todas as outras coisas.

Analisando o problema ontológico na filosofia da ciência e como ele pode ser aplicado à Linguística, Borges Neto (2004, p. 42) destaca a posição de Popper (1963), para quem é possível pensar no real a partir de três dimensões. O primeiro mundo que goza de realidade é o dos objetos físicos (livro, lápis, campo gravitacional etc.); o segundo mundo real é o dos processos mentais (saúde, crença, desejos etc.); o terceiro compreende as entidades e relações teóricas (teorias, conceitos, argumentos etc.). Segundo mostra Borges Neto, Popper entende, diferentemente de Platão, que os objetos do terceiro mundo são criações da psicologia humana que passam a existir, de fato, a partir do momento em que são criados, ganhando realidade autônoma.

Outra via em que se pode pensar o que vem a ser o real é a partir da dicotomia já clássica entre realidade subjetiva e realidade objetiva. Os objetos pertencentes à realidade subjetiva seriam aqueles cuja existência não pode ser definida a partir de parâmetros únicos universalmente válidos. Um exemplo: num estádio de futebol uma torcida vê seu time perder. Naquela multidão, muitos podem ficar extremamente tristes e até com raiva de um ou de todos os jogadores. Isso não impede, entretanto, que um grupo saia satisfeito com a atuação do time, mesmo em face da derrota, e não tão triste assim. O exemplo nos diz o seguinte: os sentimentos de raiva, satisfação e tristeza são coisas que existem, mas obviamente não podem ser tocadas, cheiradas e nem são sentidas com mesma intensidade em todas as pessoas. Isso é basicamente o que caracteriza o real subjetivo.

Os objetos da realidade objetiva são aqueles que podem ser percebidos por nossos órgãos do sentido ou por equipamentos que ampliam essas limitadas capacidades (como telescópio, microfone, quimógrafo etc.) ou podem ser definidos por meio de uma relação matemática única. Outro exemplo: se menciono

“triângulo equilátero”, todos, independentemente do tamanho que atribuam ao objeto, terão em suas mentes uma figura de três lados A, B e C em que $A=B=C$. Se falo “cadeira”, as pessoas que conhecem o objeto pensarão em algo que é feito para sentar, independentemente do formato, tamanho, cor etc.

Se, considerando essas coisas, voltarmos a pensar no conceito de verdade e no que é a ciência, podemos formular provisoriamente que a ciência se preocupa fundamentalmente com o real que pode ser apreendido de modo objetivo, sobretudo em função do modo como os próprios cientistas operam e dos instrumentos de que lançam mão para apreender e representar esse real.

É novamente Heidegger (2008, p. 169) que coloca a questão: a ciência é, então, um tipo de verdade? O filósofo responde afirmativamente:

a ciência é um tipo de verdade. A verdade, porém, pertence essencialmente ao ser-aí. Esse existe na verdade; a verdade é existente. A ciência como possibilidade de existência do ser-aí é uma possibilidade do ser na verdade. (2008, p. 169)

Convém ressaltar que o filósofo adverte que não se deve tomar a ciência como uma atividade absoluta na lida com a verdade. Ela não implica necessariamente nenhuma elevação do ser-aí humano, mas “na medida em que é uma possibilidade do ser aí, pode ser dessa ou daquela maneira, pode atuar dessa ou daquela forma” (HEIDEGGER, 2008, p. 172). Também diz o autor que a verdade científica não é a única espécie de verdade nem a mais elevada, mas é certamente “uma postura fundamental possível da existência humana” (HEIDEGGER, 2008, p. 178).

A partir desse ponto, podemos começar a pensar como essas considerações podem se relacionar com a linguística em geral e com a sociolinguística em particular. Em outros termos, uma questão que pode ser levantada diz respeito a como as análises linguísticas no interior de uma disciplina como a sociolinguística podem ser entendidas como uma postulação do real em termos de representação no âmbito do exercício teórico de uma prática científica.

Antes de abordarmos mais de perto essa questão, será reproduzida aqui uma longa citação de Heidegger, que servirá de base para as considerações a serem traçadas na próxima seção do trabalho. Eis as palavras do filósofo:

[...] Somente a ciência traz consigo a autêntica verdade sobre o ente. Um simples exemplo pode deixar claro esse estado de fato: tomemos o modo como é descoberto o Sol. Para os gregos dos primeiros tempos, em seus primórdios, ele era o deus

Hélios, o deus que atravessava o espaço celeste em sua carruagem de fogo e submergia no oceano. Mais tarde, essa interpretação foi perdendo força, e o Sol tornou-se um disco que percorre sua trajetória. Logo o disco se mostrou como uma esfera de fogo, como uma bola que se movimenta em torno da Terra. Depois disso, foi a Terra que se tornou uma esfera que se movimenta em torno do Sol, que é o centro. E por fim, esse sistema solar apareceu como um sistema entre muitos outros. Nosso Sol foi então amplamente investigado por meio do espectro solar.

Onde está então a verdade? Será que a física e a astronomia atuais conseguem afirmar que elas descobriram o sol tal como ele é? Onde se acha o critério para que a concepção atual do sistema solar possa se mostrar como a única concepção verdadeira, para que ela possa ser assumida ao mesmo tempo como mais verdadeira do que as concepções mais antigas e até mesmo mais verdadeira do que a concepção mítica? No entanto, não continuamos, ainda assim, a falar do pôr-do-sol? Seria esse apenas um modo de falar? Realmente não o vemos se pôr, e esse desvelamento do sol não domina nosso ser-aí, cotidiano? (2008, p. 176)

Essas palavras não são de modo algum uma tentativa de usurpar a credibilidade e consistência das formulações que faz a ciência. Com elas quer o autor mostrar que o conhecimento científico goza de uma especificidade em relação ao ente, que resulta na definição mesma de ciência e de verdade: deixar justamente o ente ser como ele é. A ciência é, assim, “ser no desvelamento do ente em virtude do desvelamento” (HEIDEGGER, 2008, p. 192). Isso envolve a contemplação do ente, que, na teorização científica, é impelido a “responder” aos questionamentos da ciência. Assim, quando um cientista se debruça sobre, por exemplo, como se dá a aquisição por uma criança de um segmento silábico com *onset* ramificado em língua portuguesa, ele opera com procedimentos que “forçam” o surgimento daquilo que era um todo, um *continuum* fônico indistinto (para os pais da criança, por exemplo); o cientista “provoca”, então, a emergência dessa categoria outrora invisível (ou inexistente mesmo) e o surgimento de um padrão de regularidade e hierarquia que torna possível ao pesquisador postular a tal entidade e colocar como problema a sua aquisição.

Estas últimas observações apontam para um fato que marca a ciência moderna: a intervenção no real. Uma vez que uma ciência só é ciência na medida em que circunscreve o seu real, o seu *positum*, passa a fazer parte de sua essência a modulação do ente a ser desvelado por ela – alguns exemplos da linguística podem mesmo nos mostrar isso.

A “realidade” da linguagem na linguística

As teorias linguísticas são uma maneira que os pesquisadores encontraram de produzir uma interpretação de um real difícil de ser definido e captado em sua integralidade. Se perguntarmos a um homem comum o que é uma língua talvez ele tenha uma dificuldade de fornecer uma definição precisa ou talvez ele a defina tal como ouviu na escola. Uma coisa é certa, porém, qualquer que seja a alternativa: todos nós, especialistas ou não, nos percebemos como seres cuja existência bio-psico-social se dá na e pela língua. Ela está tão em nós que submetê-la aos rigores da investigação científica não é uma tarefa fácil. Uma das razões se deve ao fato de que a língua(gem) enquanto objeto de investigação não pode ser estudada sem que esse mesmo objeto se constitua instrumento mesmo de análise do pesquisador. Em outras palavras: a botânica se dedica à análise de algo que é exterior ao ser do cientista. Nessa ciência, o objeto será estudado com o recurso a instrumentos que não são em sua natureza “botânicos”. São sempre instrumentos exteriores à constituição vegetal. Tal não é de modo algum o caso da linguagem. Se, na ciência em geral, a terminologia é um problema importante na medida em que participa da definição mesma do campo das diversas teorias elaboradas, na linguística esse problema, por exemplo, é ainda mais complicado: o significado de um termo pode ser objeto de análise (basta pensarmos no problema clássico da Semântica) e ao mesmo tempo meio sem o qual dada análise não se sustenta (pensemos na rigorosa definição de competência em Chomsky ou nos problemas em torno do conceito de palavra em Morfologia). Note-se que a Terminologia é uma disciplina no interior da linguística. E não é novidade a frustração de Saussure com a imprecisão dos termos empregados em linguística (cf. BENVENISTE, 1995, p. 44).

No que tange ao modo como o real buscado pela linguística é formulado pelos pesquisadores da área, pode-se dizer que esse real (que, por sinal é “manipulado” de modos bem diversos em função do modo como é contemplado teoricamente⁴) nem sempre encontra correspondência com o modo como os usuários não-linguistas o percebem. Basta pensarmos no seguinte:

4 Um exemplo dessa “manipulação”: se no âmbito do modelo Padrão, a teoria chomskyana era basicamente um “sistema de regras”, a partir dos anos 80 a gramática passa a ser vista pelos gerativistas como um “sistema de princípios”, em que a subespecificação de parâmetros variáveis é a explicação dada para a diversidade das línguas – note-se que postular “princípios” e “parâmetros” é uma operação que dá existência, no nível teórico, a “coisas” que não podem ser captadas diretamente da realidade.

é provável que quando pense na expressão “língua portuguesa”, um linguista logo pense em um mosaico de coisas que compõem essa entidade. Há quem pense nos parâmetros que a especificam frente aos princípios presentes nas demais línguas. Alguns talvez pensem como uma realidade estruturada em níveis como a fonologia, a sintaxe e a morfologia. Há ainda os que podem considerar a expressão inadequada se se pensa no modo como se fala em Portugal e no Brasil. É possível que um sociolinguista imagine logo um conjunto de variáveis morfossintáticas e fonéticas que compõem essa entidade...

Certamente, isso tudo, todos esses conceitos, todas essas entidades do terceiro mundo de Popper (1963), não são reais para um falante comum. Isso explica o fato de muitos acharem mesmo absurda a ideia de que, para certos indivíduos (os linguistas), não exista o modo certo de falar. Para os falantes comuns esse mundo povoado de “coisas” como fonemas, morfemas, tópico, foco, anáfora, sujeito nulo, ergativos e tudo o mais que a imaginação dos pesquisadores “inventa” não existe, e a expressão “língua portuguesa” é mesmo referente a algo usado com as mesmas características do norte ao sul do Brasil.

A verdade que a linguística persegue, ao implementar suas investigações, mesmo se considerarmos o âmbito restrito de atuação dos linguistas, não é mesmo de fácil demonstração. Se, como afirmou Heidegger (2008), a ciência desvela o ente e esse desvelar se faz múltiplo em função da diversidade do ser do ente, podemos dizer que o ente que a linguística procura desvelar tem uma constituição ôntica tal que seu desvelamento está ligado ao próprio modo como o ser-aí (neste caso o linguista) se coloca junto a esse ente. Em resumo: o ente se revela múltiplo e, além disso, o ser-aí que o desvela ocupa posições diversas.

Borges Neto (2004, p. 34), tentando encontrar uma resposta para o problema da multiplicidade de teorias linguísticas, lançou mão da noção de objeto observacional e objeto teórico. Nos termos do autor, o objeto observacional é o conjunto dos fenômenos estudados iguais a todas as teorias. Entretanto, no nível da formulação teórica, esse objeto é notado de diversas maneiras. Isso explicaria a diversidade de teorias linguísticas: o objeto observacional é uno, mas o objeto teórico é diverso.

Uma interpretação talvez nem tanto diversa do problema pode ser a seguinte: o ente linguagem (objeto observacional, nas palavras de Borges Neto) não é de modo algum igual se é observado por um linguista ou por um falante comum. Aquilo que os linguistas veem como linguagem certamente difere do que os demais humanos veem, embora a percepção natural inicial possa

ser idêntica. Além disso, esse ente, mesmo entre os linguistas, é de diversas maneiras, isto é, ele se mostra de diversas formas – o que talvez nos permita dizer que ele não é *um* objeto observacional, mas *vários* mesmo. Além disso, o modo como o ser-aí (linguista ou não) que o investiga se coloca junto a ele é bem multifacetado – isso se alinha bem à noção de objeto teórico colocada por Borges Neto. Em outras palavras: talvez o objeto observacional uno, tal como proposto por Borges Neto, só possa ser postulado como um uno muito etereamente.

Assim é que a língua no interior da doutrina saussuriana é interpretada como uma entidade dupla-face que é, ao mesmo tempo, individual e social. Mas o interessante é que a noção de fato social em Saussure, herdada das formulações durkheimianas, não é múltipla. O social da linguagem, no âmbito deste raciocínio, é que ela é compartilhada (sempre da mesma forma) por um grupo de indivíduos, visto também como corpo coletivo uno que compartilha algo que está nesse corpo, mas é fora dele. Tanto assim que as chamadas variações foram atribuídas, no interior da proposta saussuriana, às idiossincrasias do indivíduo (e não como uma propriedade intrínseca à natureza do ente língua), não merecendo maior atenção por parte do linguista, que tinha à mão um *estado* de língua para estudar.

Uma questão que se coloca é: até que ponto tal formulação é verdadeira ou falsa? Em outros termos, qual o grau de realismo da proposta saussuriana? A resposta tem um caráter epocal. É evidente que quando essas ideias começaram a circular elas gozavam de aceitação por parte de uma ampla comunidade de linguistas que, conforme o seu senso de realidade, viam nessa formulação uma tradução do ser da língua. Entretanto, quando começou a ser testada, quando começou a ser colocada à prova, as formulações estruturalistas começaram a dar sinais – para certo grupo de linguistas – de que seu grau de realismo não era tão forte quando se olhava mais amplamente para o ente linguagem. É Lucchesi (2004, p. 57) que mostra o dilema: “a questão da mudança converteu-se então num anátema no âmbito da linguística saussuriana ortodoxa, uma ameaça permanente à sua concepção da língua”. Com isso, na visão daqueles que propunham uma alternativa à análise estrutural, a teoria estruturalista não deixava, para usar uma terminologia heideggeriana, o ente ser como ele é. Nesse sentido, a intervenção saussuriana nesse real

era tão forte que tentava impor uma adequação da natureza do ente à contemplação do ser-aí e não o contrário.

Na tentativa de solução desse problema, a interpretação sociolinguística propõe uma mudança nas regras desse jogo teórico. Resolve integrar a variabilidade à natureza mesma do ente linguagem e assume uma interpretação do ser social da linguagem diversa: o social não é mais o social homogêneo de Durkheim. Agora, esse “social” implica necessariamente uma multiplicidade de individualidades, de identidades mesmo. E isso se transfere para a língua: se ela é algo social, ela é também diversa. Com isso, a teoria tentava melhor acomodar a mudança em seu interior – note-se que isso ilustra bem o fato de qualquer modelo teórico opera uma intervenção no real; os modelos nunca são apenas uma contemplação do ente como ele é.

Nesse caso, pode-se mesmo falar numa adequação das principais formulações estruturalistas à nova teoria, que não é de modo algum um rompimento qualitativo face à concepção anteriormente vigente. Ou seja: a sociolinguística – sem uma revisão crítica dos fundamentos – integra em seu mundo teórico uma série de entidades que povoavam outro mundo (cf. LUCCHESI, 2004, p. 206). Dois, dentre os vários problemas de cunho epistemológico que isso implica, são estes: a língua é então um reflexo (total ou parcial) das contradições sociais? O desvelar-se do ente da teoria sociolinguística em função da multiplicidade de identidades a que ele se prende revelar-se-á menos problemático, ou seja, o ser-aí sociolinguista estará mais próximo da verdade do ente?

O tratamento da linguagem na teoria sociolinguística e a modelação do real

Em *Ciência e pensamento do sentido*, Heidegger faz a seguinte observação:

Como a arte, a ciência tampouco é, apenas, um desempenho cultural do homem. É um modo decisivo de se apresentar tudo que é e está sendo.

Por isso devemos dizer: o que se chama de ciência ocidental europeia determina também em seus traços fundamentais e em proporção crescente, a realidade na qual o homem de hoje se move e tenta sustentar-se. (2002, p. 39)

Essa reflexão nos conduz a entender que o fazer científico não é uma atividade afastada do real que investiga. Ela, de fato, determina esse real. No dizer de Heidegger, a ciência vai ao mundo tomar-lhe as contas. Para isso, precisa recortá-lo de modo a poder mensurá-lo (pense-se já aqui na análise quantitativa da variação linguística). Assim, a verdade da ciência é uma verdade derivada dessa alteração no real, ainda que possamos considerar que tal atividade tente priorizar o deixar ser do ente como ele é.

No âmbito da teoria sociolinguística, a questão do tratamento do real é uma constante. O fato de a teoria ter integrado em seu interior os conceitos estruturalistas de sistema, estrutura e regularidade, colocando-os junto às noções de variação e mudança, impõe aos pesquisadores uma postura de busca de adequação dos procedimentos a esse novo real que lhes afronta. Um desses procedimentos inclui a análise estatística rigorosa, que é uma tentativa que os sociolinguistas empreendem de dar maior objetividade ao ente que investigam e fazer emergir os padrões de regularidade que a teoria atribui à variação linguística.

Em *Sociolinguística quantitativa: instrumental de análise*, Guy e Zilles oferecem ao leitor detalhes dos procedimentos da metodologia quantitativa em sociolinguística. Numa seção em que esclarecem o que é e por que usar o VARBRUL, os autores declaram:

de modo geral, o uso do VARBRUL facilita a construção de um modelo quantificado de processos linguísticos (e dentro de certos limites [...] dos processos sociais) que controlam e produzem os padrões regulares da variação linguística (2007, p. 105).

Note-se, nesta observação, a palavra “construção”. Seu emprego evidencia a intervenção do método no real a ser investigado. Uma vez que esse real não salta aos olhos do linguista de forma estruturada e justamente pelo fato de que esse real não pode dizer como quer ser abordado, o linguista precisa fazer uma intervenção. No caso da teoria sociolinguística, isso implica *construir* os padrões de regularidade das formas linguísticas. É interessante notar ainda que os próprios autores admitem a limitação do método no que tange à modelação/construções dos padrões sociais ligados à análise linguística. Note-se ainda que isso se relaciona com o problema epistemológico que colocamos acima, na medida em que o modelo capta melhor os padrões de regularidade linguística do que propriamente sociolinguística.

Evidentemente que essas observações não têm nenhuma pretensão de discutir falhas do modelo ou mesmo de levantar dúvidas sobre sua validade. O objetivo básico é tentar mostrar como a prática dominante na sociolinguística quantitativa produz um real cuja natureza pode se distanciar mais ou menos do ente investigado.

Um problema interessante na análise sociolinguística é a definição da amostra e seu tratamento. O uso do programa coloca questões que vão desde a extensão da amostra bem como sua representatividade face à comunidade em que ela é coletada até a melhor forma de executar seu tratamento no programa. Isso por si já mostra que, mesmo coletada do modo mais rigoroso possível, a amostra é já um recorte do real que pode apresentar maior ou menos discrepância em relação à realidade tal como ela é. Evidentemente que esse é um problema presente em toda a ciência, já que ela não é (nem pretende ser) a reprodução do real que investiga.

Mas é Heidegger novamente que nos diz:

toda objetivação do real é um cálculo, quer corra atrás dos efeitos e suas causas, numa explicação causal [o que não parece ser o caso da sociolinguística], quer, enfim, assegure em seus fundamentos, um sistema de relações e ordenamentos. (2002, p. 50)

Nesse caso, isso implica necessariamente o emprego de um método que deve ser aceito pelo cientista como válido e como o mais apropriado para sua análise. Assim, o que se evidencia é que, em função desse imperativo, o cientista – no nosso caso, o linguista – não consegue tocar a essência de sua prática. Nas palavras de Heidegger (2002, p. 56), não assiste ao cientista “a possibilidade de acesso ao incontornável de sua essência”. É por isso que o autor afirma que nenhuma física é capaz de falar da própria física.

De modo mais direto, o que se quer com isso é afirmar que, embora consciente de sua operação e de que ela é uma dentre outras possíveis, o sociolinguista no exercício de seu método tem que assumi-lo como o mais apropriado para tratar do real. Com isso escapa-lhe, entretanto, a consciência de que esse real apropriado ao método é *um real particular*, definido em seus termos no exercício teórico de uma disciplina ou doutrina científica.

Para ilustrar: uma questão mal resolvida em muitos trabalhos sociolinguísticos diz respeito ao tratamento de categorias que pertencem

à chamada variável explanatória, mais especificamente dos fatores chamados de extralinguísticos. É comum encontrarmos trabalhos que quantificam e confrontam ocorrências de formas entre homens e mulheres, mas não elaboram uma reflexão das representações de gênero no interior da comunidade em que a amostra foi colhida. Muitos dos trabalhos limitam-se ao estabelecimento de categorias tais como escolaridade, classe social, gênero etc., sem uma reflexão do quanto tal categorização se aproxima e importa, de fato, para a compreensão de como a utilização de certos fatos de língua está vinculada às práticas sociais/culturais/identitárias do grupo estudado. É muito comum a ideia de que basta correlacionar frequências de certas formas linguísticas a categorias como escolaridade e faixa etária para se ter um retrato sociolinguístico completo da comunidade de fala.

A noção de *vernáculo*, muito importante no âmbito da teoria, guia a coleta da amostra – tanto em termos de extensão da mesma quanto no que tange ao modo de coleta dos dados. Entretanto, a definição seca de *vernáculo* como o falar “espontâneo” do falante em situações “naturais” de interação é muitas vezes assumida sem uma reflexão crítica da própria noção de identidade: o que seria “natural” em termos de interação? A depender do perfil da comunidade investigada, não seria “natural” falar em situações formais? A identidade fragmentária do falante e mesmo da comunidade, muitas vezes é tratada como uma questão de estilo, apenas, não como algo que é da própria constituição ôntica do indivíduo.

Qual a implicação mais séria disso? Esse modo acrítico de exame pode realmente “forjar” um real ou mesmo deformá-lo, produzindo cenários que, de fato, não guardam a proximidade imaginada com a realidade. Guy e Zilles fazem uma advertência similar:

Uma prática bastante comum nos estudos sociolinguísticos no Brasil tem sido a de organizar amostras estratificadas, com igual número, de por exemplo, homens e mulheres, jovens e velhos etc. [...]. a prática de constituir amostras estratificadas faz sentido. No entanto, as categorias utilizadas – sexo, idade, escolaridade, etnia, classe social etc. – são muitas vezes estabelecidas de modo apriorístico ou mesmo pré-teórico. O avanço na pesquisa está apontando para a necessidade de repensar tais categorias. (2007, p. 111)

Outro exemplo: uma categoria que se revela extremamente problemática no âmbito da sociolinguística brasileira é o que se chama geralmente de “norma culta”. Por vários motivos essa noção é problemática: pela própria noção de norma (herdada da tradição coseriana); pelas implicações semânticas do adjetivo que compõem a expressão; pelo problema de delimitar a extensão do que vem a ser isso, já que ela apresenta muitas coincidências com outras normas.

Até mesmo um critério como o de que falantes com curso superior completo são os falantes da norma culta é extremamente problemático. E isso levanta uma questão séria que não é enfrentada (e é mal resolvida) no geral dos trabalhos de sociolinguística: a norma culta é um vernáculo? Se a resposta é afirmativa, há que se considerar que o Brasil da época do NURC mudou bastante e que há muitos brasileiros (e conheço de perto vários casos) que hoje têm curso superior e cujos pais nunca foram à escola – daí: se o vernáculo é adquirido na convivência com a família, como considerar que um falante de nível superior cujos pais não tenham sido nem alfabetizados tenha como vernáculo a norma culta? Como lidar com essa discrepância entre método e conceito?

Como se vê, a entidade norma culta tem existência complicada e já é de difícil aceitação hoje. Faraco (2008), que faz um exame aprofundado do conceito de norma culta e de outros que lhe estão associados, chega mesmo à seguinte conclusão:

Por tudo que afirmamos no texto, talvez melhor faríamos se abandonássemos a denominação norma culta. De um lado, nos livraríamos de sua carga de injustificável elitismo. Por outro lado, estaríamos nos aproximando de uma análise mais precisa da realidade linguística brasileira na medida em que não há, pelo menos no plano da fala, diferenças substanciais entre o que se poderia chamar de norma culta e a linguagem urbana comum. (2008, p. 64)

Como se vê, a modelação dos dados no interior da análise sociolinguística intervém num real que, em função do método analítico a que é submetido, pode se distanciar muito do ente a ser desvelado pelo ser-aí pesquisador. Problema similar já foi discutido por Odete Menon (1993), durante o Seminário Nacional sobre a Diversidade Linguística e o ensino de Língua Materna. Naquela ocasião, a autora discutia *A questão das afirmações generalizantes sobre o português do Brasil e os dados do projeto NURC*. Sintetizo a seguir alguns pontos da argumentação dela.

Menon cita o trabalho de Moreira da Silva (1983, p. 4, apud MENON, 1996, p. 208), que afirma que o pronome *se* não ocorre na língua oral. Destaca também que Pontes (1992, p. 21) faz uma afirmação semelhante, ao dizer que no PB oral, o *se* desaparece em construções existenciais. Logo a seguir, Menon faz a seguinte afirmação:

Coincidentemente, ambos são mineiros e, no seu dialeto, existe uma tendência, constatada já há algum tempo, a de apagar o pronome *se* reflexivo de alguns verbos pronominais. Mas daí a estender esse fato a nível de PB e a todos os *se*, inclusive o que exerce função de sujeito, vai uma certa extrapolação. É perfeitamente plausível que o fenômeno em tela se estenda a outras regiões do Brasil. No entanto, é nesse momento que se apresenta uma séria dificuldade: a falta de trabalhos descritivos dos vários dialetos do PB para se contrapor a esse tipo de afirmação. (MENON, 1996, p. 208)

Ao longo de seu texto, a autora segue apontando outros exemplos de afirmações generalizantes sobre o PB com base no Projeto NURC, ao passo que destaca alguns problemas metodológicos do Projeto e enfatiza a necessidade de trabalhos mais pontuais acerca do português falado por indivíduos escolarizados e não-escolarizados.

O que tudo isso sugere é que a verdade do ente buscada pela sociolinguística é uma dentre outras e que, nem sempre, as imagens que ela cria desse real permitem que o ente seja mesmo como ele é. Como o desvelamento desse ente se sujeita também ao modo como o ser-aí se coloca junto a ele, só o olhar afastado, só a contemplação descerradora do ente pode aproximar as formulações teóricas do próprio ente contemplado. O problema é que o exercício do método, como salienta Heidegger (2002), pressupõe sua aceitação de certa forma incondicional – e é aí que se revela, como nos aponta o filósofo, esse caráter terrivelmente intervencionista da ciência enquanto teoria do real. É aí que se revela a construção de uma realidade operada pelo tratamento do linguista – é, portanto, essa situação própria da ciência que justifica que se desenvolva uma ‘filosofia da linguística’.

Considerações finais

Este texto pretendeu fazer uma breve reflexão sobre o modo como a sociolinguística trata os dados que coleta e como isso implica uma produção do real que pode ser mais o menos condizente com o ente que a teoria procura

desvelar. Durante as considerações destacou-se que, em alguns trabalhos sociolinguísticos, as noções de “sociedade”, “grupos sociais”, “comunidade” são empregadas de modo apriorístico ou mesmo pré-teórico, o que pode comprometer o alcance das formulações de tais estudos. Além disso, apontou-se que tal situação se relaciona também ao fato de que ao submeter os dados a um tratamento estatístico computacional sem uma reflexão crítica sobre o método em si, o sociolinguista “forja” uma “realidade” da qual ele nem mesmo tem consciência. Note-se que não se está negando aqui a validade do modelo laboviano; o que interessa aqui é dizer que, como qualquer outro, é um modelo que intervém no real de um certo modo. Assim, quando se faz sociolinguística (ou gerativismo ou linguística histórica) há que se seguir um método, mas é fundamental ter clareza de que o método não nos dá o ente tal como ele é.

As conclusões sinalizam para os seguintes aspectos: a) urgente se faz a produção de trabalhos no âmbito da filosofia/historiografia/epistemologia da linguística brasileira – o que permitiria uma revisão crítica dos trabalhos que têm sido produzidos. Tal revisão poderia contribuir para uma melhor avaliação do impacto que as pesquisas linguísticas em geral e as sociolinguísticas em particular têm na compreensão dos diversos fenômenos que circunscrevem a realidade linguística do Brasil. Além disso, tal revisão poderia contribuir ainda para uma aplicação mais efetiva da pesquisa sociolinguística no espaço escolar brasileiro; b) apenas o exercício do método não fornece ao sociolinguista elementos para que ele pense sua prática – daí porque é importante um exame dos fundamentos filosóficos do seu método e de sua ciência.

Agradecimentos: Ao professor Carlos Alberto Faraco e à professora Rosa Virgínia Mattos e Silva.

SANTOS, Gredson dos. Data modelling and the production of “realities” in sociolinguistic analysis. *Revista do Gel*. São Paulo, v. 6, n. 2, p. 85-103, 2009.

ABSTRACT: *This text makes critical considerations on the quantitative data treatment in sociolinguistics and its impact on the formulation of “realities” with regards to the Brazilian linguistic space. Furthermore, this paper holds brief comments on how the data selection, treatment, analysis, and interpretation operations, following a quantitative methodology, “format”, in different intensity degrees, the “real” which is in the scope of the sociolinguistic analysis.*

KEYWORDS: *Science Philosophy. Linguistic Philosophy. Sociolinguistics.*

Referências

BENVENISTE, Émile. Saussure após meio século. In: _____. **Problemas de linguística geral I**. Tradução de Maria da Glória Novack e Maria Luisa Neri. Campinas, SP: Pontes, 1995. p. 34-49.

BORGES NETO, José. **Ensaio de filosofia da linguística**. São Paulo: Parábola, 2004.

FARACO, Carlos Alberto. **Norma culta brasileira**: desatando alguns nós. São Paulo: Parábola, 2008.

FREIRE-MAIA, Newton. **Verdades da ciência e outras verdades**: a visão de um cientista. São Paulo: EdUNESP; Ribeirão Preto, SP: SBG, 2008.

GUY, Gregory R.; ZILLES, Ana. **Sociolinguística quantitativa**: instrumental de análise. São Paulo: Parábola, 2007.

HEIDEGGER, Martin. Ciência e pensamento do sentido. In: _____. **Ensaio e conferências**. Tradução de Emmanuel Carneiro Leão, Gilvan Fogel e Márcia Sá Cavalcante Schuback. Petrópolis, RJ: Vozes, 2002. p. 39-60.

_____. **Introdução à filosofia**. Tradução de Marco Antonio Casanova. São Paulo: Martins Fontes, 2008.

LUCCHESI, Dante. **Sistema, mudança e linguagem**: um percurso na história da linguística moderna. São Paulo: Parábola, 2004.

MENON, Odete Pereira da Silva. A questão das afirmações generalizantes sobre o português do Brasil e os dados do NURC. In: CARDOSO, Suzana Alice Marcelino (Org.). **Diversidade linguística e ensino**. Salvador: EdUFBA, 1996. p. 206-214.

PONTES, Eunice. **Espaço e tempo na língua portuguesa**. Campinas: Fontes, 1992.

POPPER, Karl. **Conjecturas e refutações**. Brasília: Editora da UNB, 1963.