



# CONCEITOS DE IDENTIDADE E MÉTODOS PARA SEU ESTUDO NA SOCIOLINGUÍSTICA

---

CONCEPTS OF IDENTITY AND METHODS  
FOR ITS STUDY IN SOCIOLINGUISTICS

Livia Oushiro  
Universidade Estadual de Campinas (oushiro@iel.unicamp.br)

**Resumo:** Este artigo tem o objetivo de revisar algumas definições do conceito de identidade e estudos que dele se valem, com vistas a avaliar sua operacionalização e replicabilidade em pesquisas na área da Sociolinguística Variacionista (LABOV, 2008 [1972]). Os métodos para acessar as identidades dos falantes incluem, entre outros, a observação etnográfica, a formulação de perguntas específicas no roteiro da entrevista, a aplicação de questionários, a elaboração de índices, testes de percepção e a análise de posturas, cada qual com suas vantagens. Em última instância, deve-se reconhecer a natureza fluida e dinâmica das identidades, mas tal fato não impede o pesquisador de buscar métodos menos impressionísticos para o estudo de sua possível influência sobre os padrões de variação linguística.

**Palavras-chave:** Identidades; Metodologias; Variação Linguística.

---

**Abstract:** This article aims at reviewing definitions of the concept of identity as well as studies that explore it in order to evaluate its operationalization and replicability in sociolinguistic studies (LABOV, 2008 [1972]). Methods to access speakers' identities include, among others, ethnographic observations, the formulation of specific questions within the interview schedule, the application of questionnaires, the development of indices, perception tests and the analysis of stance, each of which having their own advantages. Researchers ultimately ought to recognize the fluid and dynamic nature of identities, but this fact should not hinder them from searching and developing more objective methods to study the possible influence of identities on patterns of language variation.

**Keywords:** Identities; Methodologies; Language Variation.

## INTRODUÇÃO

Os estudos sociolinguísticos de vertente variacionista costumam se desenvolver com base em certas categorizações sociais dos falantes: (i) *seu sexo e/ou seu gênero* – categorizações que vêm sendo problematizadas nos últimos anos (ver, p.ex., FREITAG, 2015); (ii) *sua faixa etária*, que os categoriza em diferentes momentos de vida – grosso modo, jovens, pessoas de meia idade e pessoas mais velhas –, para cada qual se esperam certos comportamentos, incluindo-se aí os linguísticos; e, sobretudo no Brasil, (iii) *seu nível de escolaridade*, a partir de classificações em número de anos que frequentou o ensino formal, ou categorias como Analfabeto, Ensino Fundamental, Ensino Médio e Ensino Superior. No Brasil, são vários os bancos de dados que se baseiam nessas categorias para estratificação de suas amostras, como, por exemplo, o PEUL (PAIVA; SCHERRE, 1999), o VARSUL (VARSUL, s/d), o VALPB (HORA, 2004), o Iboruna (GONÇALVES, s/d), o SP2010 (MENDES; OUSHIRO, 2012) e o PEPP (LOPES, 2018),<sup>1</sup> para citar apenas alguns.

Ao mesmo tempo, nos últimos anos, tem crescido o interesse sobre o tema da identidade dos falantes, principalmente em trabalhos que se filiam à vertente que Eckert (2012) chamou de *terceira onda da Sociolinguística*. Desse célebre artigo, cabe retomar algumas das críticas colocadas pelos estudos sociolinguísticos de terceira onda à abordagem mais tradicional (ver também MENDES, 2017). Os estudos caracterizados como “de primeira onda” muitas vezes partem de macrocategorias sociológicas – como o sexo, a faixa etária e o nível de

<sup>1</sup> Respectivamente, Programa de Estudos sobre o Uso da Língua, sediado na UFRJ; Projeto Variação Linguística na Região Sul do Brasil, projeto coordenado entre UFRGS, PUC-RS, UFPR e UFSC; Projeto Variação Linguística no Estado da Paraíba, sediado na UFPB; Iboruna, banco de dados do Projeto Amostra Linguística do Interior Paulista, sediado na Unesp-São José do Rio Preto; Projeto SP2010, sediado na USP; e Programa de Estudos sobre o Português Popular de Salvador, sediado na UFBA.

---

escolaridade dos falantes – e, *depois* de verificar padrões de variação, acabam atribuindo aos falantes certos usos linguísticos como resultado de certas “identidades”, que coincidiriam com as macrocategorias determinadas *a priori* pelo pesquisador. Note-se que essa crítica não é quanto ao emprego dessas categorias em si, mas ao que é possível concluir a partir dessa abordagem.

Por exemplo: não é rara a observação, principalmente em comunidades urbanas, de que as mulheres tendem a empregar a forma considerada de prestígio mais do que os homens da mesma comunidade (CHAMBERS, 1995; PAIVA, 2004), do que resulta uma correlação estatística significativa em análises realizadas em programas como Goldvarb ou R. Uma frequente explicação dada para tal observação empírica é porque elas seriam, supostamente, mais “sensíveis” às formas de prestígio. Esse, talvez, seja um dos pontos mais mal compreendidos da Sociolinguística. Embora Labov (2001, p. 274) empregue o termo *sensitivity* para descrever o comportamento das mulheres a fim de dar conta desse recorrente padrão em estudos de variação, uma melhor tradução para esse termo talvez não seja “sensibilidade”, mas sim “suscetibilidade”; ou seja, as mulheres não seriam mais “sensíveis”, mas sim mais *sujeitas* a pressões normativas do que homens, por conta dos papéis sociais a elas atribuídos na sociedade e das expectativas sobre seu comportamento. Assim, uma crítica da terceira onda sobre o uso de macrocategorias para explicar a variação linguística são interpretações de cunho essencialista, que atribuem a certos grupos de falantes um determinado comportamento como se fosse algo que “naturalmente” decorreria do fato de que aquele indivíduo pertence a determinada categoria social (MENDOZA-DENTON, 2002).

A terceira onda, desse modo, procura ter um olhar mais atento aos próprios indivíduos e a suas práticas cotidianas, pois uma categoria imposta de cima para baixo por um pesquisador pode não ser o modo como a própria pessoa se definiria ou se identificaria. Identidades de gênero são um bom exemplo disso, porque não é possível classificar todas as pessoas dentro de apenas duas caixinhas – homens e mulheres –, que são categorias que se referem ao sexo biológico somente, e não necessariamente ao modo como as pessoas se identificam em seu dia a dia.

Uma segunda crítica dos estudos de terceira onda sobre as análises mais tradicionais na Sociolinguística diz respeito aos significados sociais de variantes linguísticas *para além* do eixo estigma-prestígio. Com efeito, o olhar sobre os indivíduos em suas práticas sociais demonstra que as pessoas podem se

---

identificar com diversos outros grupos para além de suas classes sociais ou níveis de escolaridade, e preocupar-se não somente com “falar corretamente” de acordo com padrões normativos. Existem patricinhas, manos, youtubers, bolsominions, coxinhas, petralhas... existem identidades regionais: soteropolitanos, paulistas, gaúchos... existem identidades raciais: brancos, negros, asiáticos, indígenas... existem identidades futebolísticas: corinthianos, flamenguistas, cruzeirenses, colorados... e diversas outras identidades e significados que surgem localmente, durante uma interação específica. Essas categorias, ademais, se cruzam: um indivíduo não é apenas “paulista ou baiano”, ou apenas “mulher ou homem”, “gay ou hétero”, ou apenas “de direita ou de esquerda”, mas várias delas ao mesmo tempo – ou nenhuma delas, a depender da interação. Dentro de determinada perspectiva da terceira onda, as pessoas não “são” essas categorias, pois as próprias identidades são construídas por meio da variação linguística. Isso implica assumir que os falantes não falam do modo como falam como *reflexo* inevitável de categorias a que “pertencem”, mas que tais categorias identitárias são *elaboradas*, de forma mais ou menos consciente, pelos falantes durante suas interações sociais. Também implica assumir que as identidades são sempre múltiplas, plurais, e que um falante não pode ser reduzido a uma única dimensão (“ser mulher”, “ser nordestino” etc.).

Nesse sentido, o objetivo deste artigo é, primeiramente, o de revisar alguns conceitos de identidade em estudos sociolinguísticos (Seção 1) e, em um segundo momento, avaliar diferentes métodos que buscam depreendê-las em sua multiplicidade (Seção 2). Se se tem interesse em estudar *identidades sociolinguísticas*, não basta simplesmente um olhar “de cima para baixo” do pesquisador a partir de macrocategorias sociais pré-determinadas. Vai se argumentar, afinal, que é justamente pelo fato de as identidades sociolinguísticas não serem fixas nem singulares que se faz necessário aplicar métodos variados, objetivos e replicáveis, a fim de compreender seu papel sobre a variação linguística.

## 1 CONCEITOS DE IDENTIDADE EM SOCIOLINGÜÍSTICA

Definições de “identidade” costumam ressaltar o caráter relacional desse conceito. Mendoza-Denton (2002, p. 475) o define como “a negociação ativa da relação de um indivíduo com construtos sociais mais amplos, na medida em que essa negociação é sinalizada através de meios linguísticos e outros meios

---

semióticos”<sup>2</sup>. Dessa definição, cabe destacar que identidade é, antes de mais nada, uma *negociação*, ou seja, um indivíduo não tem poder de definir para si uma identidade totalmente nova, que não tenha sido elaborada coletivamente e que não seja aceita por outras pessoas.

A negociação feita pelos indivíduos com construtos sociais mais amplos é bem exemplificada por Battisti (2014), numa reflexão sobre redes sociais e identidades, em seu estudo sobre a palatalização de [ti] e [di] (como em *tia* e *dia*) em Antônio Prado-RS. A autora afirma que “[a] pesar de as identidades serem experimentadas, vivenciadas pelos sujeitos e, nas investigações, serem consideradas pelo exame das práticas sociais individuais, elas são em parte construtos sociais” (BATTISTI, 2014, p. 80). Battisti se baseia no conceito de *habitus* de Bordieu (1977, *apud* BATTISTI, 2014) para relacionar as identidades individuais com práticas que se adquirem no processo de socialização dos indivíduos, da vida em sociedade; essas práticas, por sua vez, vão definindo identidades que, ao mesmo tempo, são passíveis de mudança ao longo do tempo a partir de nossa participação em diferentes comunidades. Battisti (2014, p. 81) considera, assim, que “a construção de identidade consiste em negociar os significados de nossa experiência de pertença a diferentes grupos sociais”.

Kiesling (2013, p. 450), de modo semelhante, ressalta a natureza relacional das identidades, definindo o conceito como “um estado ou processo de relação entre o ‘eu’ e o ‘outro’; a identidade é como os indivíduos definem, criam, ou pensam sobre si em termos de sua relação com outros indivíduos e grupos, sejam eles reais ou imaginários”<sup>3</sup>. Kiesling (2013) também chama a atenção para a necessidade de, ao pensar identidades em seu contexto social, levar em conta a escala em que o conceito é operacionalizado. O autor propõe distinções em pelo menos três grandes níveis: (i) amplos grupos “censitários”, como sexo, classe social, etnia; (ii) papéis institucionais, como “mãe”, “policial”; e (iii) tomadas de postura (*stance*) durante a interação. Note-se, novamente, que não há um abandono das macrocategorias sociológicas, como sexo e classe social, mas sim o entendimento de que elas devem ser pensadas dentro de sua própria dimensão.

<sup>2</sup> Texto original: “the active negotiation of an individual’s relationship with larger social constructs, in so far as this negotiation is signaled through language and other semiotic means.”

<sup>3</sup> Texto original: “Identity is a state or process of relationship between self and other; identity is how individuals define, create, or think of themselves in terms of their relationship with other individuals and groups, whether these others are real or imagined.”

---

Todas essas definições ressaltam o papel do indivíduo que negocia suas identidades com outros indivíduos, grupos e construtos sociais. Assim, não se deve superestimar o papel agentivo do indivíduo na construção de sua própria identidade. É pouco provável que um indivíduo possa monitorar e manipular conscientemente cada traço linguístico de sua fala a todos os momentos; além disso, uma série de construtos sociais que servem de parâmetros identitários preexiste e ultrapassa o domínio da ação individual. Gumperz (1971, p. 152-153) sintetiza essa questão do seguinte modo:

em última instância, é o indivíduo que toma a decisão, mas a liberdade de escolher está sempre sujeita a restrições tanto gramaticais quanto sociais. As primeiras se relacionam com a inteligibilidade das sentenças; as segundas com sua aceitabilidade. [...] O poder de seleção, portanto, é limitado por convenções que servem para categorizar formas linguísticas como informais, técnicas, vulgares, literárias, humorísticas etc.<sup>4</sup>

Desse modo, diferentes definições de identidade sempre ressaltam que um indivíduo não é “x” ou “y”, mas que essas categorizações devem necessariamente estar relacionadas com *outros* indivíduos, uma vez que um olhar sobre a identidade não pode perder de vista as relações sociais. A identidade, portanto, não é um atributo pessoal, muito menos uma posse; ela é um processo de criação de sentidos que deve ser ao mesmo tempo individual e coletivo. A construção de sentidos se dá sempre dentro de uma matriz cultural e ideológica, sobre a qual o indivíduo não exerce total controle.

O arcabouço teórico-metodológico da Sociolinguística Variacionista, por intermédio de análises quantitativas e multivariadas, oferece ferramentas particularmente úteis para o desembaraço da complexa rede de influências entre fatores sociais e linguísticos na formação de identidades sociais (MENDOZA-DENTON, 2002). A próxima seção exemplifica e discute alguns desses métodos.

## 2 MÉTODOS PARA ABORDAR IDENTIDADES

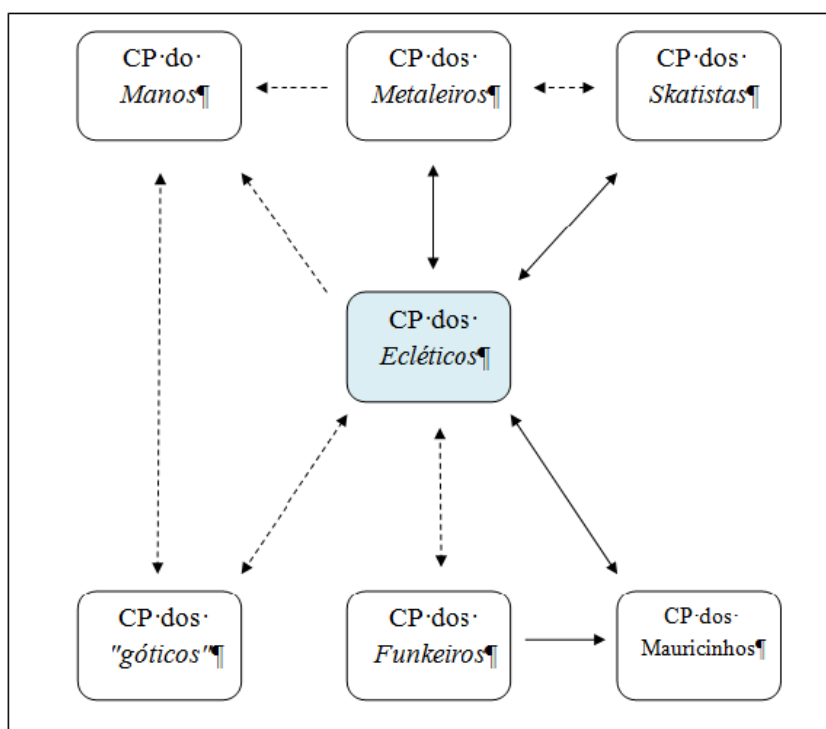
São vários os métodos já aplicados para analisar diferentes identidades dos falantes, e esta seção não tem o objetivo de esgotá-los. Uma discussão mais detalhada pode ser encontrada em Schilling (2013).

<sup>4</sup> Texto original: “ultimately it is the individual who makes the decision, but his freedom to select is always subject both to grammatical and social restraints. Grammatical restraints relate to the intelligibility of sentences; social restraints relate to their acceptability. [...] The power of selection is therefore limited by commonly agreed-on conventions which serve to categorize speech forms as informal, technical, vulgar, literary, humorous etc.”

A *observação etnográfica*, tomada da Antropologia Linguística (DURANTI, 1997), foi adotada por Eckert (1989) em seu estudo sobre *jocks e burnouts* em um colégio em Detroit nos Estados Unidos. O principal intuito dessa metodologia de coleta de dados é permitir que o pesquisador se afaste de macrocategorias sociais e depreenda categorias locais relevantes para membros de uma comunidade.

Um exemplo desse tipo de abordagem nos estudos brasileiros é a tese de Salomão-Conchalo (2015), que realizou um estudo etnográfico em uma escola pública de São José do Rio Preto, na qual identificou diferentes comunidades de práticas. A Figura 1 mostra esses grupos (manos, metaleiros, ecléticos, funkeiros etc.) e as relações entre eles: as linhas tracejadas indicam um contato reduzido e as linhas contínuas sinalizam relações diretas; ao mesmo tempo, as relações podem ter um direcionamento único, como a flecha dos funkeiros em relação aos mauricinhos, ou de mão dupla, como a flecha entre metaleiros e ecléticos.

Figura 1: Relações entre comunidades de práticas em um colégio em São José do Rio Preto-SP, no estudo de Salomão-Conchalo (2015)

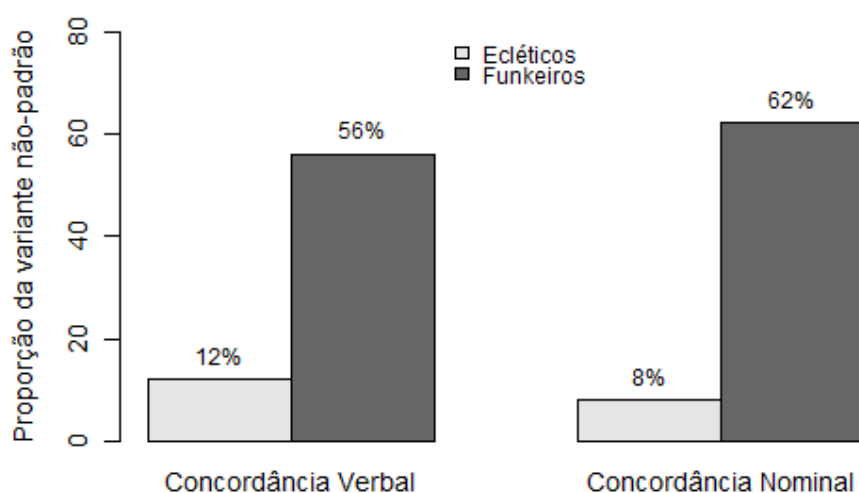


Fonte: Salomão-Conchalo (2015, p. 185)

Depreender esses grupos e saber quem interagiu com quem só foi possível pela observação participante da pesquisadora ao longo de dois anos. Salomão-Conchalo (2015) analisou os padrões de variação de duas variáveis bem

conhecidas na Sociolinguística brasileira: a concordância nominal e a concordância verbal de terceira pessoa do plural. Dentre vários aspectos que podem ser destacados desse trabalho, vale a pena chamar a atenção para as diferenças nos comportamentos linguísticos por parte de dois desses grupos, os funkeiros e os ecléticos. O primeiro grupo era formado por adolescentes avessos à instituição escolar e que se identificavam ideologicamente com a cultura *hip hop*, *funk* e *rap*. Os membros do segundo grupo, autodenominado “ecléticos” porque circulavam em diferentes grupos da escola, gostavam de frequentá-la e de participar de atividades extracurriculares. A Figura 2 compara as proporções de uso das variantes não padrão de concordância nominal (à esquerda) e verbal (à direita) entre esses dois grupos.

Figura 2: Proporções de uso das variantes não-padrão de concordância verbal e nominal nas comunidades de práticas dos ecléticos e dos funkeiros, no estudo de Salomão-Conchalo (2015)



Fonte: Adaptado de Salomão-Conchalo (2015, p. 243)

Enquanto os ecléticos têm baixas taxas de concordância não padrão – 12% para CV e 8% para CN –, os funkeiros apresentam 56% de CV não padrão e 62% de CN não padrão. Cabe notar, aqui, que um estudo sociolinguístico que se baseasse em macrocategorias sociais, definidas *a priori*, provavelmente acabaria classificando todos esses indivíduos dentro de uma mesma categoria: adolescentes, estudantes de escola pública. Esses dados, no entanto, mostram



---

quão diferentes podem ser as práticas linguísticas de pessoas dentro uma mesma macrocategoria social.

Outro método para depreender identidades são *questionários* especificamente voltados para isso. Hoffman & Walker (2010), em um grande projeto de coleta de dados da fala de Toronto, no Canadá, estavam interessados principalmente em diferentes grupos de imigrantes e seus descendentes que compõem a população da cidade. Toronto é uma cidade amplamente diversificada étnica, cultural e linguisticamente, com a presença de chineses, italianos, alemães, portugueses, poloneses, gregos, dentre outros grupos. A fim de investigar se diferentes grupos de imigrantes apresentavam diferentes tendências linguísticas, e se tais diferenças se relacionam com as identidades étnico-culturais dos indivíduos, os autores desenvolveram um questionário de orientação étnica que incluía perguntas sobre como os falantes se identificam etnicamente, sua rede social e usos do inglês e da língua de herança, como exemplificadas em (1). As respostas receberam pontuação entre 1 e 3, para indicar menor ou maior orientação à identidade étnica, e a média dos valores para 35 questões do questionário representou o índice de orientação étnica de cada indivíduo, que foi então utilizado como variável independente no estudo de diversas variáveis sociolinguísticas.

(1) Identificação étnica

- a. Você se considera italiano, canadense ou ítalo-canadense?
- b. A maioria de seus amigos é italiano?
- c. As pessoas em seu bairro são italianas?
- d. As pessoas com quem você trabalha são italianas?
- e. Quando você era criança, os alunos na sua escola eram italianos? Vocês eram amigos? E as crianças no seu bairro?

Língua

- a. Você fala italiano? Você fala bem? Com que frequência você fala italiano?  
Se não: Você consegue entender italiano?
- b. Onde você aprendeu italiano? Em casa? Na escola?
- c. Você prefere falar italiano ou inglês?
- d. Você prefere ler e escrever em italiano ou em inglês? Você lê jornais em italiano? Quais?
- e. Você prefere ouvir rádio ou ver TV em italiano ou em inglês?

---

Inspirado nesse estudo, um questionário semelhante foi usado no Projeto Processos de Acomodação Dialetal na fala de nordestinos residentes em São Paulo (OUSHIRO, 2016). Nesse projeto, analisou-se a fala de migrantes paraibanos e alagoanos que residem nas cidades de São Paulo e Campinas, em comparação com a fala de paulistas e de paraibanos e alagoanos não migrantes, a partir de amostras dos Projetos SP2010 (MENDES; OUSHIRO, 2012), VALPB (HORA, 2004) e PORTAL (OLIVEIRA, 2017). O objetivo foi o de investigar em que medida os migrantes adquiriram novas formas linguísticas, como o /r/ tepe ou retroflexo em posição de coda silábica (como em *porta*) e a realização africada de /t/ e /d/ antes de [i] (como em *tia* e *dia*), variantes, em princípio, não prototípicas ou menos frequentes em seus locais de origem (ver CARDOSO et al, 2014). Se os migrantes têm usado essas formas com mais frequência do que os conterrâneos que não migraram, quais fatores podem explicar a aquisição ou aumento na produção dessas formas?

Além de características sociodemográficas dos falantes (seu sexo, idade e motivo de migração, tempo de residência etc.), uma hipótese diz respeito ao grau de identidade desses falantes com o local de origem ou com a comunidade anfitriã. Desse modo, foi incluído no roteiro de entrevista sociolinguística um questionário de perguntas sobre a rede social (2), os hábitos (3) e o grau de identificação do falante com cada região (4):

(2) Rede social

- a. Sem contar as pessoas que moram com você, quem são as pessoas com quem você mais convive?
- b. A maioria dos seus amigos hoje é de AL/PB?
- c. Você tem vizinhos que são de AL/PB?

(3) Hábitos

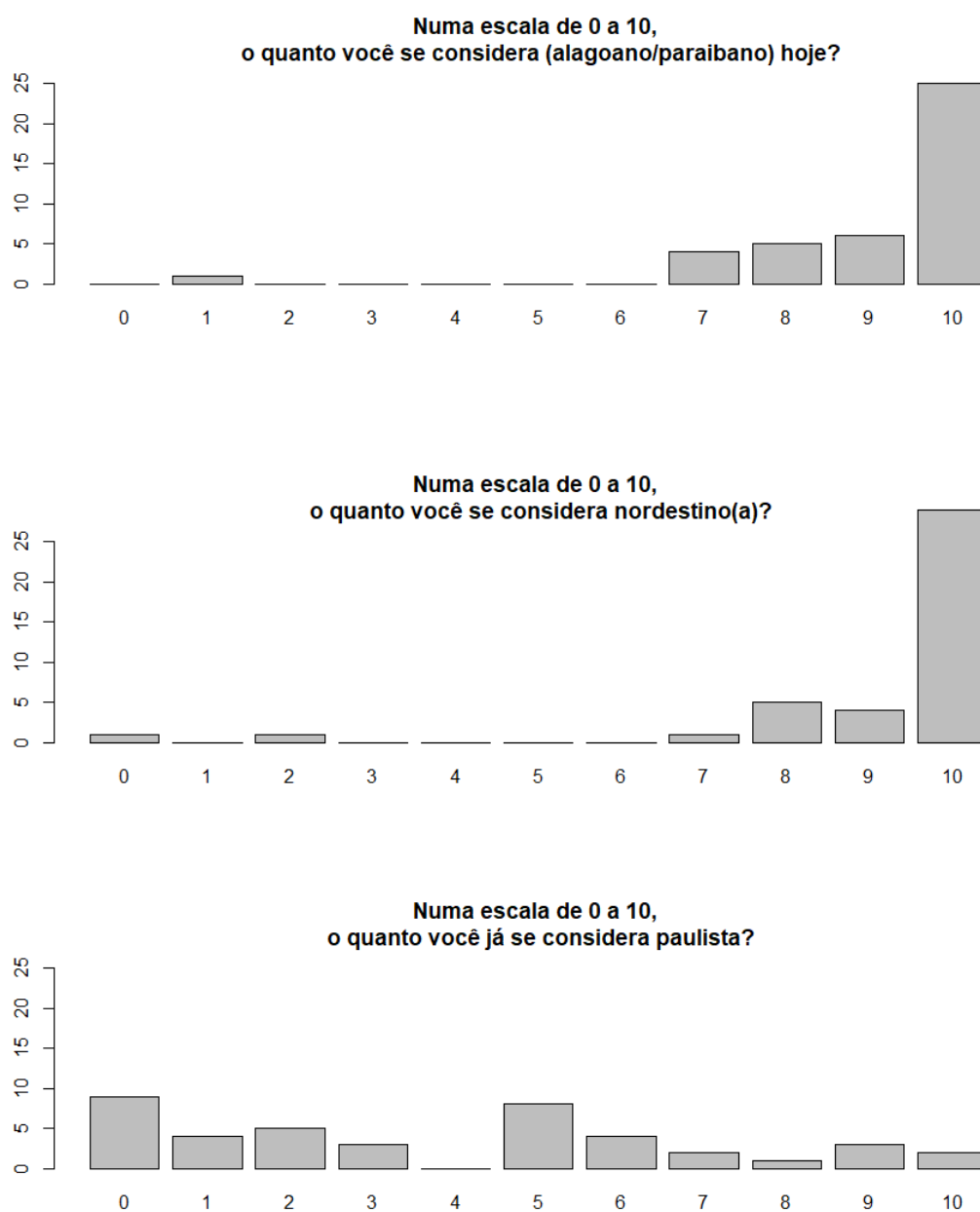
- a. Hoje em dia, com que frequência você come comida nordestina?
- b. Hoje em dia, com que frequência você ouve música nordestina?
- c. Com que frequência você fala com parentes e amigos que não migraram?
- d. Com que frequência você volta para sua cidade natal?

(4) Identidade

- a. Numa escala de 0 a 10, o quanto você se considera alagoano/paraibano hoje?
- b. Numa escala de 0 a 10, o quanto você se considera nordestino?
- c. Numa escala de 0 a 10, o quanto você já se considera paulista?

A Figura 3 mostra a distribuição das respostas para as três escalas de identidade para os 40 migrantes que foram entrevistados em Campinas. Para a escala de quão paraibano ou alagoano os falantes se consideram (topo), vê-se que a grande maioria se atribuiu uma nota alta da escala, entre 7 e 10. O mesmo é observado para a escala de quão nordestino (centro) eles se consideram.

Figura 3: Notas autoatribuídas por migrantes alagoanos e paraibanos residentes em São Paulo para graus de identidade com estado de origem (topo), região Nordeste (centro) e estado de São Paulo (abaixo), no estudo de Oushiro (2019)



Fonte: Oushiro (2019)

---

A autoatribuição de notas altas para as identidades de origem é, de certo modo, esperada. Sabe-se que “negar as raízes” é uma atitude socialmente mal vista e, pelo menos discursivamente, a grande maioria dos falantes demonstram valorizar sua herança cultural. Isso, contudo, é um indicativo de que uma explicação para o maior ou menor uso de formas paulistas por parte desses migrantes não pode ser simplesmente “sua identidade alagoana/paraibana/nordestina”, pois praticamente todos eles se identificam com o próprio estado e com o Nordeste – enquanto sua produção linguística, por outro lado, é bem menos homogênea (ver OUSHIRO, 2019).

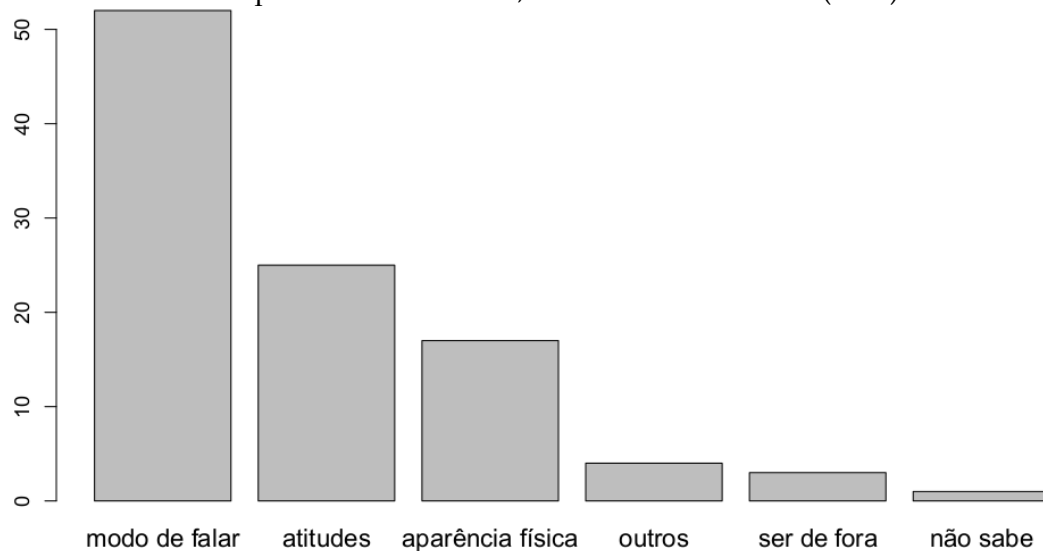
Em contraste, quando se analisa a escala de notas atribuídas para quão “paulista” eles já se consideram, percebe-se que as notas não se concentram no ponto baixo da escala (como se poderia plausivelmente esperar), mas se distribuem de forma mais ou menos equilibrada entre 0 e 10. Desse modo, é possível que a escala de “paulistidade” explique mais dos padrões de uso de formas paulistas do que as escalas de “alagoanidade/paraibanidade” e “nordestinidade”.

Esses resultados mostram que é necessário ter cautela para interpretar as manifestações identitárias dos falantes durante as gravações. Não basta constatar que um falante que mantém os traços linguísticos do local de origem revela orgulho de suas raízes; também é necessário examinar se falantes que mais assimilaram os traços linguísticos da nova comunidade se identificam relativamente mais com a comunidade anfitriã.

Outra questão de interesse ao sociolinguista é não apenas os modos como os falantes se identificam, mas também como eles identificam outros falantes e de que modo eles o fazem. Se é por meio de traços linguísticos, é de interesse saber quais são esses traços mais salientes, para diferentes grupos. No Projeto SP2010 (MENDES; OUSHIRO, 2012), foram incluídas *perguntas específicas* no roteiro de entrevista sociolinguística para descobrir quais são esses traços. Depois de ter perguntado a quais lugares o falante já havia viajado, o documentador perguntava: “Você disse que já foi para X, para Y, para Z. Quando você foi para esses lugares, as pessoas percebiam que você era paulistano?” A essa pergunta, 92% dos participantes disseram que sim. Quando a resposta era sim, perguntava-se na sequência: “Como elas percebiam?” As pessoas deram diversas respostas, como a aparência física (p.ex., a brancura da pele), atitudes (como estar sempre com pressa e ficar perguntando “se vai dar tempo”), e, principalmente, o modo de falar. As respostas foram categorizadas nos grupos apresentados na Figura 4,

que também apresenta a frequência com que foram mencionadas pelos participantes do Projeto SP2010.

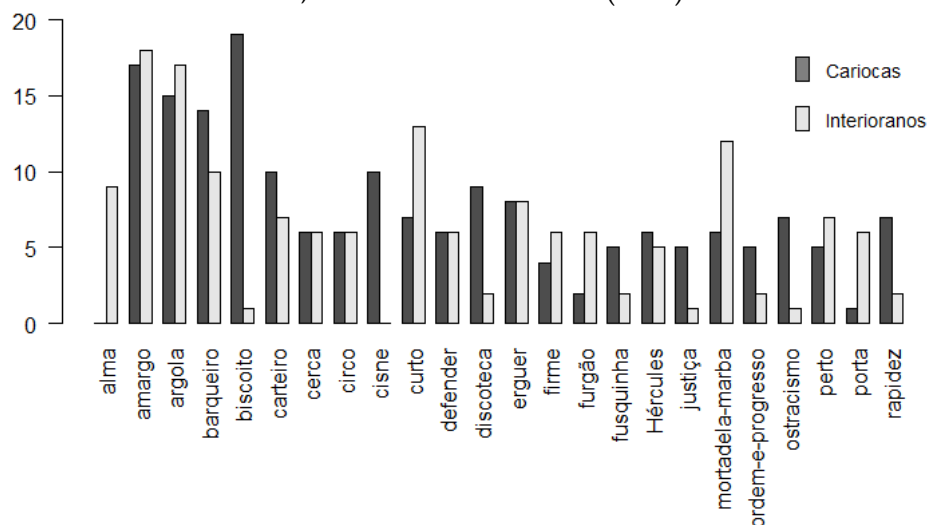
Figura 4: Respostas à pergunta: “Quando você estava em X, as pessoas te reconheciam como paulistano? Como?”, do estudo de Oushiro (2015)



Fonte: Oushiro (2015, p. 26)

No roteiro de entrevista, após a leitura de uma lista de palavras, os documentadores também perguntavam como o participante achava que um carioca e uma pessoa do interior de São Paulo falaria algumas dessas palavras. A expectativa é que a própria seleção de palavras revela quais traços linguísticos são mais salientes para os falantes. A Figura 5 mostra os 20 itens mais mencionados da lista; nela, percebe-se que a grande maioria são palavras que contêm o /r/ em coda silábica – para identificar cariocas e interioranos – e o /s/ em coda para os cariocas.

Figura 5: Palavras mais frequentemente mencionadas pelos participantes para diferenciação entre o português paulistano e o português de cariocas/interioranos de SP, do estudo de Oushiro (2015)



Fonte: Oushiro (2015, p. 95)

A própria situação de entrevista sociolinguística, então, quando bem montada e aplicada sistematicamente, pode fornecer bases sólidas para que o pesquisador possa afirmar que determinado traço linguístico é mais saliente e quais identidades se associam a esse traço, para além de suas próprias impressões.

Por outro lado, os falantes podem identificar ou categorizar outros falantes a partir de reações inconscientes, sem saber por que o fizeram. Ao lado de comentários explícitos dos participantes, também é possível obter reações subjetivas a certos usos linguísticos, por meio de métodos indiretos. A *técnica de estímulos pareados* (*matched-guise technique*), desenvolvida por Lambert et al (1960), é um dos métodos que permite depreender diferentes significados sociais que se associam a diferentes línguas, variedades ou variantes linguísticas. A técnica consiste em criar pares de estímulos, geralmente sonoros, em que o pesquisador controla o aspecto que deseja analisar.

A pesquisa de Oushiro (2015), por exemplo, buscou eliciar reações subjetivas às variantes tepe e retroflexa de /r/ em coda silábica por parte de moradores da cidade de São Paulo. O objetivo era o de identificar diferentes significados sociais dessas variantes, e se esses significados se relacionam com os padrões de produção verificados na comunidade (a saber, uma mudança em

tempo aparente na direção da variante retroflexa nas regiões periféricas da cidade).

Para tanto, foram criados pares de estímulos, idênticos em todos os aspectos, exceto pela pronúncia do /r/ em coda. Inicialmente, foram gravados quatro falantes paulistanos de semelhantes perfis sociais (dois homens e duas mulheres, moradores da Zona Oeste da cidade, com nível superior de escolaridade e com cerca de 30 anos), em uma situação de conversa informal. Das gravações, foram extraídos excertos curtos, de 15 a 20 segundos, e que continham uma certa concentração de ocorrências de /r/ em coda (4 a 7). Os falantes foram então recontactados a fim de que produzissem realizações controladas das variantes de /r/ em coda nas mesmas sentenças como tepe ou como retroflexo. Tais produções foram inseridas nas gravações originais por meio do programa Praat (BOERSMA; WEENINK, 2014), a fim de produzir dois estímulos para cada falante, um que continha apenas tepes e outro que continha apenas retroflexos.

Posteriormente, esses estímulos foram separados em dois grupos (Quadro 1), de modo que os participantes ouvintes escutassem cada falante uma única vez – o estímulo com tepe ou o estímulo com retroflexo.

Quadro 1: Separação dos estímulos para experimento de percepções no estudo de Oushiro (2015)

	<b>Grupo A</b>	<b>Grupo B</b>
Falante 1 (M)	tepe (1.t)	retroflexo (1.r)
Falante 2 (F)	retroflexo (2.r)	tepe (2.t)
Falante 3 (M)	retroflexo (3.r)	tepe (3.t)
Falante 4 (F)	tepe (4.t)	retroflexo (4.r)

Fonte: Oushiro (2015, p. 276)

Após ouvir cada estímulo, os ouvintes preenchem um questionário sobre como imaginavam ser aquela pessoa. Uma parte do questionário está transcrita em (5):

(5)

Para você, essa pessoa parece...

(Escolha uma opção em cada linha)

	Pouco					Bastante
Extrovertida	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Escolarizada	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Feminina <sup>5</sup>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Inteligente	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Formal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Amigável	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Paulistana	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Ter sotaque	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Ter amigos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Essa pessoa deve morar num...  
(Escolha uma opção)

Bairro mais periférico

Bairro mais central

Fonte: Oushiro (2015, p. 279)

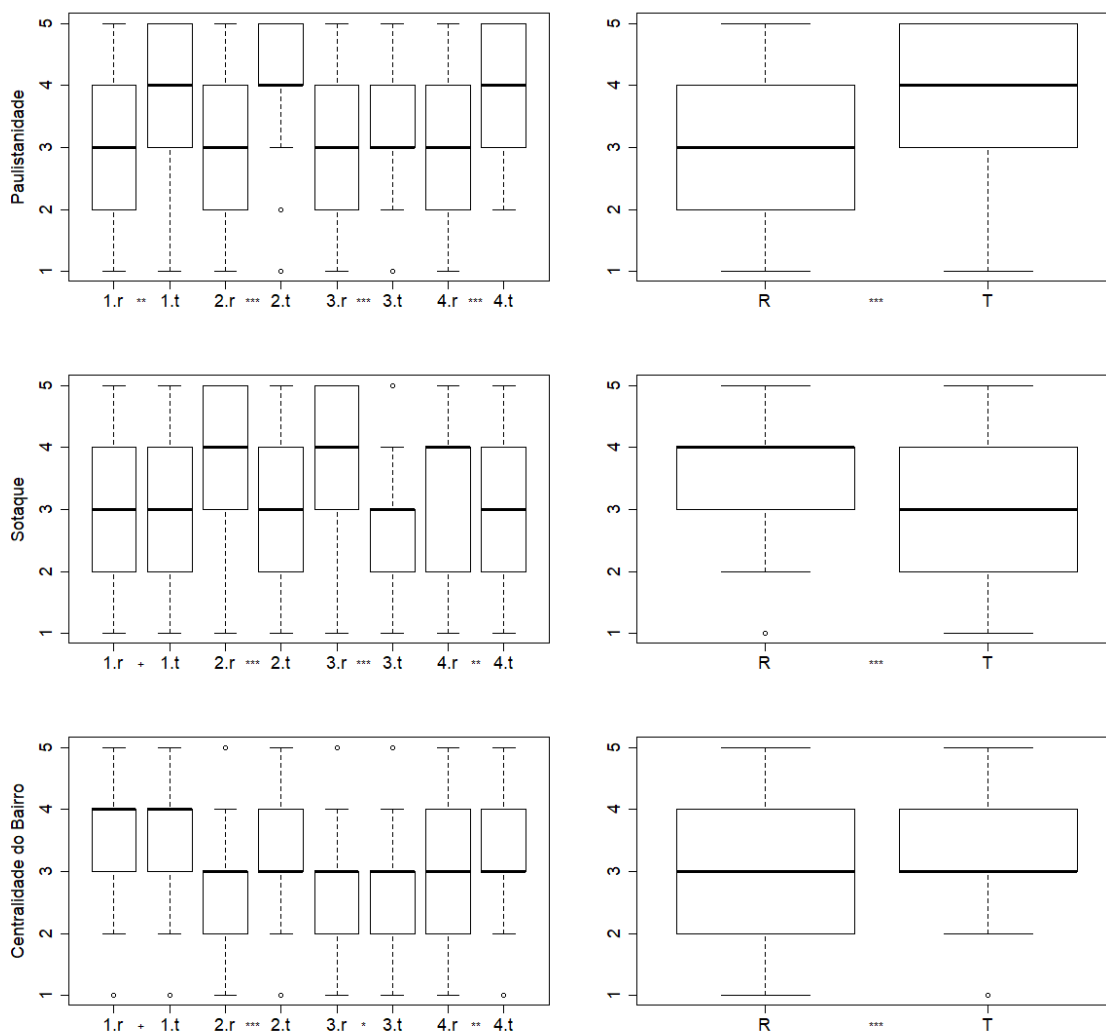
A análise das respostas ao questionário, preenchido no total por 185 participantes, buscou verificar se um mesmo falante, ouvido com uma variante ou com outra, recebia diferentes julgamentos a depender de qual foi a variante ouvida. Tendo em mente que cada par de estímulos é idêntico quanto à qualidade de voz, conteúdo da mensagem etc., se as respostas para uma característica atribuída ao falante são significativamente diferentes para os estímulos com tepe e com retroflexo, isso significa que os ouvintes, conscientemente ou não, associam cada variante linguística com diferentes grupos de falantes.

A Figura 6 mostra a distribuição das respostas para três das escalas do questionário: quão paulistano soa o falante (topo), seu grau de sotaque (centro) e o grau de centralidade de seu bairro de residência (abaixo). À esquerda, contrastam-se as respostas para os estímulos com retroflexo e tepe por falante (1.r vs 1.t, 2.r vs 2.t etc.), e à direita a distribuição geral de toda a amostra.

<sup>5</sup> Ou “Masculina”, para vozes masculinas.



Figura 6: Correlações entre variante de /r/ ouvida e percepções sobre Paulistanidade (topo), Sotaque (centro) e Centralidade do Bairro (abaixo) do falante no estudo de Oushiro (2015)

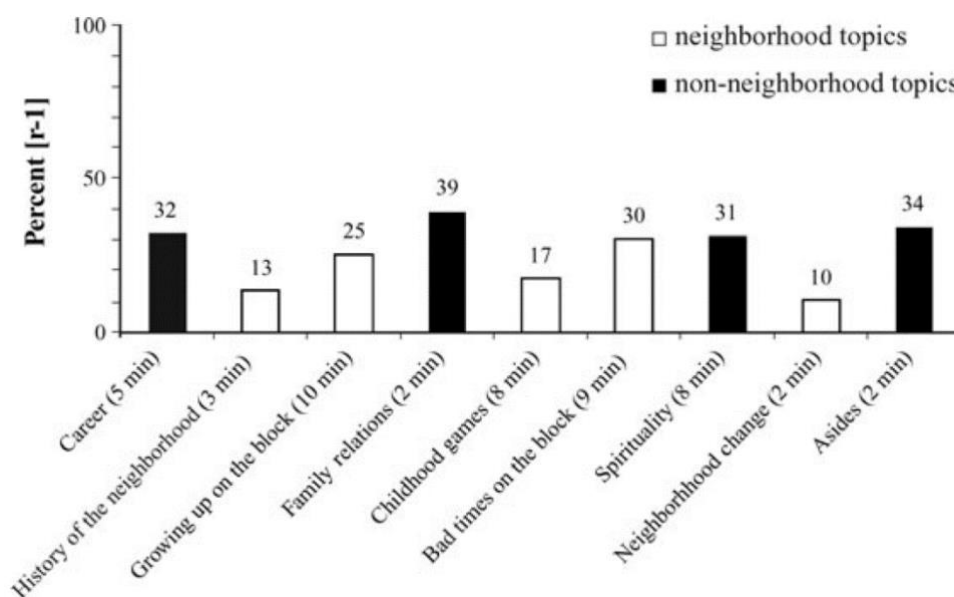


Fonte: Oushiro (2015, p. 290)

Resumidamente, a figura mostra que os falantes foram julgados como mais paulistanos quando ouvidos com o tepe, como tendo mais sotaque quando ouvidos com o retroflexo, e como residentes de bairros mais centrais quando ouvidos com o tepe. Nota-se, contudo, que os falantes não necessariamente foram avaliados como “não paulistanos” quando ouvidos com o retroflexo, como “não tendo sotaque” com o tepe ou como “residentes de bairros periféricos” quando ouvidos com o retroflexo; nesses casos, as respostas se concentram num ponto “neutro” (o valor 3 da escala), que mostra que as categorizações sociais dessas variantes não se organizam como pares de opostos.

Por fim, outro método para o estudo de identidades é uma análise da variação estilística de acordo com os assuntos tratados e diferentes posturas do falante (JAFFE, 2009) ao longo da entrevista sociolinguística. Um exemplo desse tipo de análise é a pesquisa de Becker (2009). A autora realizou um estudo em Nova Iorque que, dentre outros objetivos, replicou a pesquisa de Labov (2006 [1966]) no bairro de Lower East Side em Nova Iorque, a fim de verificar se as mudanças que o autor havia observado na década de 1960 continuavam na mesma direção. Uma das variáveis analisadas por Becker foi a realização do /r/ pós-vocálico do inglês, que em Nova Iorque pode ser apagado ou realizado. Becker (2009) percebeu que, na fala de diferentes indivíduos, as taxas de realização do /r/ mudavam a depender do assunto. A Figura 7 mostra a taxa de realização de /r/ em diferentes momentos da entrevista de um de seus participantes, Michael.

Figura 7: Taxas de realização de /r/ para um falante de acordo com tópicos da entrevista sociolinguística, no estudo de Becker (2009)



Fonte: Becker (2009, p. 651)

Na Figura 7, as barras escuras são os momentos em que o falante mais pronuncia o segmento, que é a variante de prestígio nos Estados Unidos: quando Michael fala sobre a carreira, relações familiares e espiritualidade, a realização de /r/ ultrapassa 30%; quando ele fala sobre a história do bairro, como foi crescer no Lower East Side, brincadeiras de infância e as mudanças que vêm ocorrendo no bairro, sua taxa de realização cai para até 10%. Becker interpretou que isso tem a

---

ver não só com o assunto sobre o qual falava, mas também com a identidade do falante: o apagamento do /r/, apesar de ser estigmatizado nos Estados Unidos, é um traço típico de Nova Iorque e do Lower East Side. Esse falante, assim, deixa de realizar os /r/ pós-vocálicos como forma de sinalizar sua identificação com o bairro. Essa análise de posturas mostra como os falantes podem manipular (de forma consciente ou não) o uso de variáveis linguísticas para sinalizar filiação a certos grupos e como a variação linguística não é reflexo, mas também a matéria prima básica com a qual os falantes constroem identidades e expressam diversos significados sociais.

## CONSIDERAÇÕES FINAIS

Este artigo teve o objetivo de apresentar algumas definições de identidade em estudos sociolinguísticos, bem como discutir sua aplicação em estudos de variação, de modo a revelar e interpretar os padrões sociolinguísticos que emergem dos dados. As diferentes definições de identidade ressaltam o caráter relacional do conceito, que deve passar, necessariamente, pela relação entre um indivíduo e construtos sociais mais amplos. Diferentes métodos podem acessar diferentes categorizações feitas pelos falantes, de modo direto ou indireto, por meio de suas práticas, discursos e percepções. A análise de agrupamentos menores e de indivíduos também permite mais bem entender diferentes significados sociais de variantes linguísticas, cujos usos são dinâmicos.

Há que se ter em mente, ademais, que tais significados são múltiplos, e que a identidade declarada pelos falantes em seu discurso não pode ser tomada *prima facie* como única explicação para determinados usos linguísticos. Ao assumir que identidades sociolinguísticas são múltiplas e dinâmicas, faz-se necessário aplicar métodos variados, objetivos e replicáveis para validação das correlações entre a variação estilística, percepções e identidades. Alguns desses métodos foram aqui revisitados: a observação etnográfica, aplicação de questionários, perguntas específicas no roteiro de entrevistas sociolinguísticas, testes de percepção e análises de posturas. Os estudos sociolinguísticos só têm a ganhar com a ampliação e o desenvolvimento de novas técnicas de coleta, análise e interpretação dos usos sociais da variação.

---

## REFERÊNCIAS

- BATTISTI, Elisa. Redes sociais, identidade e variação linguística. In: FREITAG, Raquel Meister Ko (ed.), *Metodologia de coleta e manipulação de dados em Sociolinguística*, São Paulo: Blucher, 2014, p. 79–98.
- BECKER, Kara. /r/ and the construction of place identity on New York City's Lower East Side. *Journal of Sociolinguistics*, vol. 13, n. 5, p. 634–658, 2009.
- BOERSMA, Paul; WEENINK, David. Praat: doing phonetics by computer, 2014. Disponível em [<http://www.fon.hum.uva.nl/praat/>]. Último acesso em 29 jan. 2019.
- CARDOSO, S. A. M. da S.; MOTA, J. A.; AGUILERA, V. de A.; ARAGÃO, M. do S. S. de; ISQUERDO, A. N.; RAZKY, A.; MARGOTTI, F. W.; ALTENHOFEN, C. V. *Atlas Linguístico do Brasil*, vol. 1. Londrina: EDUEL, 2014.
- CHAMBERS, Jack K. *Sociolinguistic theory: language variation and its social significance*. Oxford: Blackwell, 1995.
- DURANTI, Alessandro. *Linguistic Anthropology*. Cambridge: Cambridge University Press, 1997.
- ECKERT, Penelope. *Jocks and burnouts: social categories and identities in the High School*. New York: Teachers College University, 1989.
- ECKERT, Penelope. Three waves of variation study: the emergence of meaning in the study of variation. *Annual Review of Anthropology*, vol. 41, p. 87–100, 2012.
- FREITAG, Raquel Meister Ko. (Re)discutindo sexo/gênero na sociolinguística. In: FREITAG, R.M.K.; SEVERO, Cristine Gorki (orgs.). *Mulheres, linguagem e poder*. Estudos de gênero na sociolinguística brasileira. São Paulo: Blucher, p. 17–73, 2015.
- GONÇALVES, Sebastião Carlos Leite Gonçalves. Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista. Disponível em [<http://www.iboruna.ibilce.unesp.br/index.php>]. Último acesso em 23 jan. 2019.
- GUMPERZ, John J. Linguistic and social interaction in two communities. In: DIL, Anwar S. (ed.), *Language in social groups*. Essays by John J. Gumperz, p. 151–176. Stanford: Stanford University Press, 1971.
- HOFFMAN, Michol; WALKER, James A. Ethnolects in the city: ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, vol. 22, p. 37–67, 2010.
- HORA, Dermeval da. *Estudos sociolinguísticos: perfil de uma comunidade*. Santa Maria: Palotti, 2004.
- JAFFE, Alexandra. *Stance*. Oxford: Oxford University Press, 2009.
- KIESLING, Scott F. Constructing identity. In: CHAMBERS, J. K.; SCHILLING, Natalie (eds.), *The handbook of language variation and change*. 2 ed. Malden, MA: Wiley-Blackwell, 2013, p. 448–467.

---

LAMBERT, W. E.; HODSON, R. C.; GARDNER, R. C.; FILLENBAUM, S. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, vol. 60, n. 1, p. 44–51, 1960.

LABOV, William. *Principle of linguistic change: social factors*. Cambridge: Blackwell, 2001.

LABOV, William. *The social stratification of English in New York City*. São Paulo: Cambridge University Press, 2006 [1966].

LABOV, William. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline R. Cardoso. São Paulo: Parábola, 2008 [1972].

LOPES, Norma da Silva. O PEPP e os estudos sobre o português de Salvador. *A Cor das Letras*, Feira de Santana, vol. 19, n. Especial, p. 23–39, 2018.

MENDES, Ronald Beline. A terceira onda da sociolinguística. In: FIORIN, José Luiz (org.) *Novos Caminhos da Linguística*. São Paulo: Contexto, p. 103–123, 2017.

MENDES, Ronald Beline; OUSHIRO, Livia. O paulistano no mapa sociolinguístico brasileiro. *Alfa*, vol. 56, n. 3, p. 973–1001, 2012. Disponível em [<http://seer.fclar.unesp.br/alfa/article/view/4942/4369>]. Último acesso em 22 jan. 2019.

MENDOZA-DENTON, Norma. Language and identity. In: CHAMBERS, J.K.; TRUDGILL, P.; SCHILLING-ESTES, N. (eds.), *The handbook of language variation and change*. 1 ed. Malden, MA: Blackwell, p. 475–499, 2002.

OLIVEIRA, Alan Jardel. Projeto PORTAL: variação linguística no português alagoano, 2017. Disponível em [<http://www.portuguesalagoano.com.br/>]. Último acesso em 22 jan. 2019.

OUSHIRO, Livia. *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*, 2015, 390f, Tese (Doutorado em Linguística). FFLCH, USP, São Paulo.

OUSHIRO, Livia. Projeto Processos de Acomodação Dialetal na Fala de Nordestinos Residentes em São Paulo. Projeto Regular FAPESP (Processo 2016/04960-7), 2016.

OUSHIRO, Livia. Contato dialetal e identidade(s). Trabalho apresentado no 67o. Seminário do GEL, 2019. Disponível em [<https://www.letraria.net/67-seminario-do-gel/>]. Último acesso em 13 set. 2019.

PAIVA, Maria da Conceição. A variável gênero/sexo. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (orgs.) *Introdução à Sociolinguística: o tratamento da variação*. São Paulo, Contexto, 2004.

PAIVA, Maria da Conceição; SCHERRE, Maria Marta Pereira. Retrospectiva sociolinguística: contribuições do PEUL. *DELTA*, vol. 15, n. Especial, p. 201–232, 1999.

SALOMÃO-CONCHALO, Mircia. A variação estilística na concordância nominal e verbal como construção de identidade social, 2015, 313f, Tese (Doutorado em Estudos Linguísticos). IBILCE, UNESP, São José do Rio Preto.

SCHILLING, Natalie. *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press, 2013.

---

VARFUL. Projeto Variação Linguística na Região Sul do Brasil. Disponível em [http://www.varsul.org.br/]. Último acesso em 23 jan. 2019.

Nota do editor:  
Artigo submetido para avaliação em: 28/02/2019.  
Aprovado em sistema duplo cego em: 07/06/2019.