



Efeitos do abaixamento na frequência do terceiro formante na pronúncia aproximante do /r/ em coda sobre a percepção social de falantes em São Bernardo do Campo

Effects of the decrease in the frequency of the third formant in the approximant pronunciation of coda /r/ on the social perception of speakers in São Bernardo do Campo

Gustavo de Campos Pinheiro da SILVEIRA*

RESUMO: Este estudo apresenta evidências de que diferenças graduais na frequência do terceiro formante da pronúncia aproximante do /r/ em coda podem provocar diferenças igualmente graduais no modo como falantes são socialmente avaliados em São Bernardo do Campo, município da Grande São Paulo. Conduziu-se um experimento de percepção sociolinguística, elaborado de acordo com a técnica *matched-guise* (LAMBERT *et al.*, 1960), com 58 residentes do município, que foram solicitados a escutar uma gravação da fala de uma pessoa lendo uma frase e, em seguida, responder a um questionário acerca do falante. Os estímulos auditivos diferiam unicamente nas ocorrências do /r/ em coda, cada um apresentando apenas um de três graus de abaixamento do terceiro formante. Os resultados indicam correlação entre o grau de abaixamento e a maneira como os falantes foram avaliados. A queda na frequência do

ABSTRACT: This study reports evidences that gradual differences in the frequency of the third formant in the pronunciation of coda /r/ can provoke equally gradual differences in how speakers are socially evaluated in São Bernardo do Campo, a city in Greater São Paulo. A sociolinguistic perception experiment designed according to the *matched-guise* technique (LAMBERT *et al.*, 1960) was conducted with 58 residents of the city, which were asked to listen to a recording of the speech of a person reading a sentence, and then to fill out a questionnaire about the speaker. The auditory stimuli differed only in the tokens of coda /r/, each of them presenting only one of three levels of decrease in the third formant. The results indicate correlation between the decrease and the way in which the speakers were evaluated. The decrease in the frequency of the third formant correlated with the increase in the rate of the responses that

* Mestrando em Linguística pela Universidade Estadual de Campinas (UNICAMP). ORCID: <https://orcid.org/0000-0002-2869-2024>. silveira@tuta.io

terceiro formante se correlacionou com o crescimento na taxa das respostas que avaliaram os falantes como tendo mais sotaque e sendo residentes de São Bernardo do Campo. Esse resultado sugere que significados sociais de formas linguísticas podem variar gradativamente quando veiculados por variáveis fonéticas graduais.

PALAVRAS-CHAVE: /r/ em coda. /r/ retroflexo. Percepção sociolinguística. Matched-guise. Terceiro formante.

evaluated the speakers as having more accent and as being resident of São Bernardo do Campo. This result suggests that social meanings of linguistic forms can vary gradually when indexed by gradual phonetic variables.

KEYWORDS: Coda /r/. Retroflex /r/. Sociolinguistic perception. Matched-guise. Third formant.

1 Introdução

Neste artigo serão discutidos os resultados de um experimento de percepção sociolinguística que, combinando o método *matched-guise* (LAMBERT *et al.*, 1960) com técnicas de manipulação digital dos sinais de fala (CAMPBELL-KIBLER, 2009), teve como objetivo testar a hipótese de que diferenças graduais na queda da frequência do terceiro formante (F3) na pronúncia aproximante do /r/ em coda silábica (p. ex., em palavras como “porta” e “tarde”) podem provocar diferenças igualmente graduais no modo como os falantes são socialmente avaliados em São Bernardo do Campo, município da Região Metropolitana de São Paulo.

Uma das primeiras pesquisas a utilizar o método *matched-guise* para estudar a percepção social do /r/ em coda no português brasileiro foi conduzida por Oushiro (2015, 2019) com o propósito de investigar os significados sociais associados às variantes tepe e aproximante¹ na cidade de São Paulo. A pesquisadora elaborou um experimento de percepção seguindo um protocolo inspirado no estudo de Campbell-

¹ Apesar de a autora, como vários outros pesquisadores, chamar de “retroflexas” as pronúncias aproximantes do /r/ em coda silábica que são popularmente qualificadas como “caipiras”, aqui elas serão designadas apenas por “aproximantes”, por motivos que serão expostos na próxima seção.

Kibler (2009), em que trechos de gravações de fala espontânea são digitalmente manipulados para criar estímulos auditivos controlados, que diferem apenas nas ocorrências da variável linguística de interesse, preservando a naturalidade da fala.

No estudo de Oushiro, os estímulos auditivos foram criados a partir de quatro trechos selecionados respectivamente da fala de quatro paulistanos nativos, com perfis sociais semelhantes. Os trechos foram retirados de gravações de entrevistas sociolinguísticas e cada um apresentava de quatro a sete ocorrências de /r/ em coda silábica. Utilizou-se o software Praat (BOERSMA; WEENINK, 2020) para substituir as ocorrências de /r/ em coda por instâncias controladas, produzidas pelos mesmos falantes, das variantes tepe e aproximante. Para cada trecho, obteve-se um par de estímulos idênticos em todos os aspectos, com exceção de que um dos estímulos apresentava apenas a variante tepe e o outro, apenas a aproximante. O experimento contou com a participação de 185 residentes da cidade de São Paulo e a tarefa solicitada a esses participantes consistia em escutar um único estímulo de cada falante e responder um questionário, avaliando-os num conjunto de características. As respostas obtidas foram analisadas estatisticamente a fim de determinar se as avaliações feitas aos falantes diferiam sistematicamente a depender da variante de /r/ em coda empregada.

Em conformidade com estudo preliminar conduzido por Mendes e Oushiro (2012), os resultados a que Oushiro (2015) chegou indicam que a pronúncia variável do /r/ em coda silábica funciona em São Paulo como índice de sotaque, paulistanidade e centralidade da região de origem. Quando ouvidos nos estímulos com pronúncia aproximante, os falantes foram avaliados como menos paulistanos, como tendo mais sotaque e como residentes de bairros menos centrais. Essas avaliações são evidências de que a pronúncia aproximante do /r/ é percebida na cidade de São Paulo como uma forma mais marcada e associada a bairros mais periféricos.

No experimento de Oushiro (2015), adotou-se uma perspectiva categórica binária da pronúncia variável do /r/ em coda, mas, como nota Soriano (2016, p. 45), um dos resultados fornece indícios de que os ouvintes paulistanos também são sensíveis a diferenças graduais no uso dessa variável. Dos quatro pares de estímulos auditivos, apenas um (i.e., o par de um dos quatros falantes) não apresentou efeitos significativos sobre as respostas dos participantes. Para Oushiro, variáveis intervenientes, como a ordem não aleatorizada de apresentação dos estímulos e as diferenças no conteúdo temático dos trechos, podem ter neutralizado os efeitos das variantes tepe e aproximante no caso específico desse par. Por outro lado, Soriano sugere outra explicação, apoiando-se na constatação feita por Oushiro (2015, p. 291) de que nos estímulos do falante em questão a pronúncia da variante aproximante soa como mais “fraca”, não apresentando tão claramente a qualidade comumente associada ao “r caipira”, quando comparada à pronúncia dos demais falantes. Para Soriano, é possível que a associação de significados sociais às pronúncias do /r/ em coda em São Paulo seja sensível a diferenças fonéticas mais sutis, como diferenças de grau de retroflexão. Alternativas a essa expressão “grau de retroflexão”, menos comprometidas com a articulação do segmento rótico, são as utilizadas por Ladefoged (2003, p. 149): grau de roticidade (*degree of rhoticity*) ou tonalidade de r (*r-colouring*).

Se essa hipótese estiver correta, pode-se cogitar que os paulistanos discriminem ao menos duas variantes aproximantes, uma com menor grau de roticidade, que seria socialmente menos marcada, e outra com grau mais acentuado, remetendo-se a regiões mais periféricas e ao “sotaque caipira”. A maior proximidade dos significados sociais da forma menos marcada com relação ao tepe explicaria a ausência de efeitos significativos daquele par de estímulos.

A partir de análises qualitativas dos comentários metalinguísticos feitos por paulistanos nas entrevistas sociolinguísticas do Projeto SP2010 (MENDES; OUSHIRO, 2012), Soriano (2016) constatou que alguns participantes mencionaram a diferença

entre dois tipos de aproximante, uma “forte”, associada a um falar mais interiorano, e uma “fraca”, mais típica da fala paulistana. Essas evidências motivaram a pesquisadora a conduzir um experimento de percepção com o intuito de examinar como as características sociais dos ouvintes que residem na cidade de São Paulo se correlacionam com a capacidade de discriminação de variantes fonéticas do /r/ em coda.

No teste de percepção conduzido por Soriano (2016) em plataforma *online*, os participantes tinham que escutar diferentes pares de estímulos auditivos. Cada par era formado por duas pronúncias de um mesmo item lexical, sendo este item diferente para cada par. As pronúncias se distinguiam apenas na variante empregada do /r/ em coda, que podia ser uma de cinco possibilidades: vibrante alveolar com três, com duas e com uma “batida” (esta última, equivalente ao tepe), e duas aproximantes, chamadas pela pesquisadora de “alveolar” e “retroflexa” – aqui elas serão referidas, respectivamente, por “fraca” e “forte”. Para cada par, os ouvintes eram instruídos a avaliar a diferença entre as duas pronúncias usando uma barra deslizante que ia de “Igual” a “Muito diferente”.

A análise quantitativa das 4.796 respostas de 109 moradores do município de São Paulo indicou que a capacidade de discriminação das pronúncias do /r/ em coda varia a depender das variantes em oposição. A maior diferença foi atribuída ao par formado pela aproximante forte e a vibrante alveolar de três “batidas”, e a menor, ao par formado pelas vibrantes de duas e de três “batidas”. Os ouvintes avaliaram o tepe como diferindo mais da aproximante forte do que da fraca, mas a aproximante fraca foi vista como mais semelhante à variante forte do que ao tepe.

Uma das principais contribuições do experimento de Soriano foi mostrar que a capacidade de discriminação das variantes do /r/ em coda pode variar a depender de certas características sociais dos ouvintes. A pesquisadora constatou que os participantes nascidos em regiões interioranas em que as pronúncias aproximantes são

bastante produtivas atribuíram um grau de diferença entre as formas fraca e forte significativamente menor do que os participantes nascidos na cidade de São Paulo, onde o tepe é mais frequente. Para Soriano, esse resultado parece indicar que nascer e crescer em regiões onde a exposição às variantes aproximantes é maior faz com que as diferenças entre as formas fraca e forte sejam menos salientes para o ouvinte.

Outro fator que se mostrou relevante para explicar as diferenças nas respostas obtidas diz respeito ao reconhecimento da variante na própria fala por parte do ouvinte. Os participantes que acreditavam que não faziam uso das formas aproximantes do /r/ em coda, mas apenas das vibrantes, como o tepe, atribuíram diferenças significativamente maiores aos pares de estímulos formados por uma vibrante e uma aproximante do que os participantes que acreditavam que faziam uso não só das vibrantes, mas também das aproximantes.

Os estudos de Oushiro (2015, 2019) e de Soriano (2016) deixam claro que a percepção social da pronúncia variável do /r/ em coda em São Paulo é um fenômeno altamente complexo. Por um lado, as variantes socialmente significativas não parecem se reduzir à oposição binária entre tepe e aproximante, tendo em vista que há evidências de que os ouvintes também são capazes de identificar mais de um tipo de aproximante. Por outro, os estudos conduzidos até então ainda não são suficientes para determinar, nas diferentes produções do /r/ em coda, os limites dos contrastes fonéticos capazes de provocar nos ouvintes associações significativas de natureza social.

Dando prosseguimento a essa linha de investigação, o presente estudo analisou as percepções sociolinguísticas da pronúncia variável do /r/ em coda silábica em São Bernardo do Campo. Dois motivos levaram à escolha desse município para a realização da pesquisa. Em primeiro lugar, os estudos sobre o /r/ em coda no estado de São Paulo até então se concentraram nas duas populações dessa região a que mais se associam os significados sociais dessa variável: os nativos da cidade de São Paulo,

para os quais o tepe é a norma (MENDES; OUSHIRO, 2011), e os nativos do interior, sobretudo das zonas rurais, para os quais prevalece a pronúncia aproximante, comumente chamada de “r retroflexo” ou de “r caipira” (AMARAL, 2020; BRANDÃO, 2007; LEITE, 2004, 2010). Os habitantes das demais cidades que se aglomeram ao redor da capital paulista, entre os quais se encontram os são-bernardenses, não se encaixam nesses dois grupos, pois, a rigor, eles não são nem paulistanos, nem interioranos. Investigar um dos municípios da Grande São Paulo contribuirá para determinar em que medida os padrões de percepção sociolinguística se propagam para além da capital, na direção das cidades periféricas da metrópole paulista.

Em segundo lugar, há evidências de que São Bernardo do Campo é avaliado por paulistanos como uma região caracterizada pela pronúncia “caipira” do /r/ em coda. Ao solicitar a paulistanos que opinassem a respeito da sentença “a porta tá aberta”, com pronúncia “caipira” do /r/, Oushiro (2015) constatou que os respondentes associaram essa forma não apenas a municípios do interior paulista, como Sorocaba e Campinas, mas também à cidade de São Bernardo do Campo. Essa constatação chama atenção ao se considerar que São Bernardo do Campo faz divisa e se integra com o sudeste da cidade de São Paulo, formando com ela um contínuo urbano, além de se encontrar a não mais do que 30 quilômetros do litoral, estando longe, portanto, de ser uma cidade interiorana. Esse discurso metalinguístico, que aparentemente destoa da estereotípica associação da variante aproximante com o interior, motivou a condução do experimento em São Bernardo do Campo, no intuito de investigar como os próprios são-bernardenses percebem e avaliam socialmente a aproximante rótica.

Na seção a seguir, serão descritos os procedimentos adotados na elaboração e execução do experimento. Em seguida, serão apresentados e discutidos os resultados principais. Será mostrado que os são-bernardenses são capazes não apenas de discriminar três variantes da pronúncia aproximante do /r/ em coda, diferenciadas na

frequência do F3, como também de associar essas variantes a informações sociais a respeito do falante.

2 Metodologia

O desenho do experimento foi elaborado de acordo com a técnica *matched-guise* (LAMBERT *et al.*, 1960). Seu objetivo foi verificar se o modo como o falante é avaliado socialmente pelos são-bernardenses se altera em função de mudanças no que se costuma chamar de “grau de retroflexão” do /r/ em coda. Neste artigo esse termo não será utilizado por não ser adequado do ponto de vista fonético.

As aproximantes róticas, das quais o /r/ retroflexo faz parte, se manifestam acusticamente sobretudo na frequência do F3 (DELATTRE; FREEMAN, 1968; LADEFOGED; MADDIESON, 2009; LINDAU, 1985). Quanto mais acentuada for a queda nessa frequência, maior é a sensação no ouvinte de “roticidade”, para usar a expressão de Ladefoged (2003). No entanto, esse padrão acústico apresenta uma correspondência ambígua com movimentos articulatorios. A teoria acústica dos sons da fala prevê queda nos valores do F3 quando há constrição na região inferior da cavidade faríngea ou na região pós-alveolar da cavidade oral, e o arredondamento dos lábios também pode contribuir para a queda na frequência de ambos F2 e F3 (LINDAU, 1985, p. 163). Conseqüentemente, o mesmo efeito acústico do que se costuma chamar de “r retroflexo” pode ser alcançado por mais de uma configuração articulatoria, sem que envolva necessariamente a retroflexão do ápice da língua. Por exemplo, a retração da raiz da língua, ao diminuir o espaço faríngeo, pode levar ao abaixamento do F3 e causar no ouvinte uma sensação de “r retroflexo”, mesmo que não tenha ocorrido retroflexão do ápice lingual. Isso significa que, com base apenas em critérios auditivos ou acústicos, não é possível determinar se houve ou não retroflexão da língua numa ocorrência aproximante do /r/ em coda, sendo para isso

necessário dispor de registros articulatórios da produção do segmento, como os obtidos por Delattre e Freeman (1968) via cineradiografia.

Apesar de as informações acústicas não serem confiáveis para determinar graus de retroflexão, ainda assim elas podem ser utilizadas para investigar as percepções sociais das aproximantes róticas. Neste experimento, a frequência do F3 foi utilizada como critério para determinar uma gradação de pronúncias aproximantes do /r/ em coda em termos da sensação de “roticidade” que elas produzem no ouvinte, assumindo a correspondência entre o grau do abaixamento na frequência do F3 e o aumento da sensação de “roticidade” (LADEFOGED; MADDIESON, 2009; LADEFOGED, 2003).

Com o objetivo de produzir estímulos auditivos diferenciados unicamente no grau de abaixamento na frequência do F3 na pronúncia aproximante do /r/ em coda silábica, dois falantes paulistanos com faixa etária de 20 a 25 anos, uma mulher e um homem, foram gravados lendo uma frase curta com duas ocorrências de /r/ em coda. A falante mulher foi gravada lendo a sentença (1) e o falante homem, lendo a sentença (2), sendo que as palavras grifadas consistem nas palavras-alvo do experimento (i.e., aquelas em que há ocorrência de /r/ em coda):

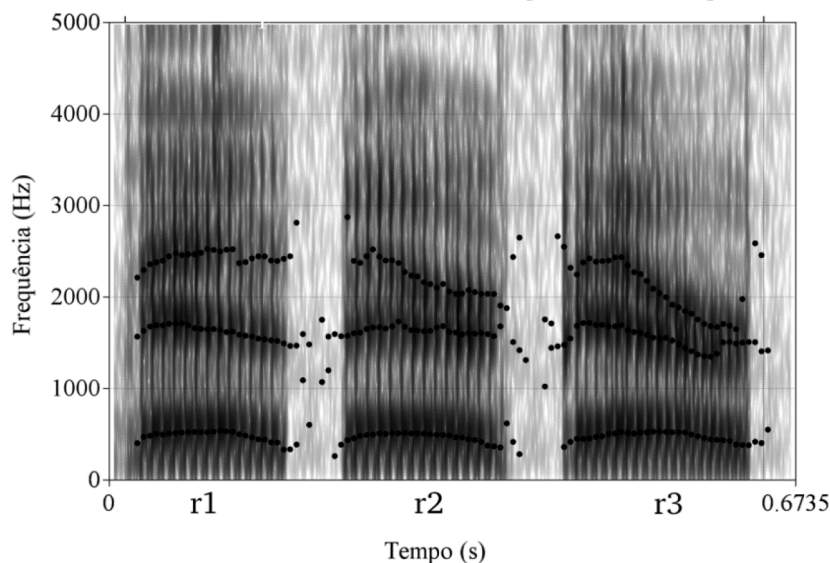
1. Agora eu vou *jantar*, a gente se fala mais *tarde*;
2. Tem uma loja na rua do lado que fica *aberta* o dia todo, fica *perto* do ponto de ônibus.

Na primeira etapa de gravação, os falantes foram orientados a ler sua respectiva sentença de modo confortável, sem impor modificações na pronúncia. O objetivo nesse momento era obter uma leitura mais natural. Na etapa seguinte, os falantes foram gravados repetidas vezes lendo a mesma frase, mas, dessa vez, instruídos a pronunciar as duas ocorrências de /r/ em coda com três graus de “r caipira”. A cada gravação,

carregou-se o áudio no Praat e analisaram-se os formantes no espectrograma a fim de verificar se a frequência do F3, obtida com auxílio da técnica LPC (ATAL; HANAUER, 1971), nas pronúncias do /r/ em coda tinha se aproximado suficientemente do nível almejado.

Para ilustrar esse fenômeno, a Figura 1 apresenta o espectrograma de banda larga dos três graus de abaixamento do F3 no segmento / ϵ r/ na palavra “aberta”, produzidos pelo falante do sexo masculino. Nas três realizações, o F3 começa aproximadamente na mesma região, em torno de 2.500 Hz, correspondendo aos valores da pronúncia da vogal anterior média-baixa [ϵ]. Depois de aproximadamente 50 milissegundos de cada segmento, os terceiros formantes começam a se diferenciar devido à realização da aproximante rótica. Enquanto o primeiro apresenta queda muito sutil, em torno de 2.400 Hz, o segundo manifesta um nítido abaixamento, chegando aos 2.000 Hz. Com os menores valores, a frequência do F3 no terceiro segmento se aproxima dos 1.700 Hz, provocando uma queda também no F2.

Figura 1 – Espectrograma de banda larga exemplificando três graus de abaixamento do terceiro formante na pronúncia aproximante do /r/ em coda. Os segmentos r1, r2 e r3 consistem em três pronúncias do segmento / ϵ r/ na palavra “aberta” produzidas pelo falante de sexo masculino. Eles foram recortados dos áudios originais e colocados lado a lado a fim de tornar mais clara a diferença entre eles nos padrões formânticos. As linhas pontilhadas indicam os valores estimados via LPC das frequências dos três primeiros formantes.



Fonte: elaborada pelo autor.

Esse processo de gravação e análise acústica foi repetido até se obter, para cada ocorrência de /r/ em coda nas duas sentenças, três exemplares cujas frequências do F3 estivessem suficientemente diferenciadas e, em conjunto, apresentassem uma gradação consistente. A Tabela 1 apresenta as frequências médias do terceiro formante dos exemplares do /r/ em coda, nas duas palavras-alvo da sentença de cada falante, que foram selecionados para a criação dos estímulos auditivos. Nota-se que em todas as palavras há diminuição consistente na frequência do terceiro formante, com queda de aproximadamente 400 Hz a cada grau.

Tabela 1 – Frequências médias estimadas em Hertz via LPC do terceiro formante nas instâncias de /r/ em coda silábica nas palavras-alvo usadas na elaboração dos estímulos auditivos.

		r1	r2	r3
Falante mulher	jantar	2413	2060	1698
	tarde	2412	2015	1717
Falante homem	aberta	2790	2341	2036
	perto	2737	2362	2061

Fonte: elaborada pelo autor.

Seguindo a estratégia de Campbell-Kibler (2009), também adotada por Oushiro (2015), os estímulos auditivos foram criados a partir da edição digital dos sinais de fala. A gravação da leitura de cada sentença feita na primeira etapa, em que as sentenças foram lidas de modo confortável e mais natural, serviu de base para a criação dos estímulos. Elas foram modificadas unicamente nas realizações do /r/ em coda, incluindo a vogal imediatamente precedente. Com uso do Praat, foram extraídas das gravações feitas na segunda etapa as três realizações selecionadas (as mesmas a que se refere a Tabela 1), correspondentes aos três graus de abaixamento na frequência do F3 dos segmentos /ar/ nas palavras “jantar” e “tarde” e dos segmentos /ɛr/ nas palavras “aberta” e “perto”. Esses segmentos extraídos substituíram, num processo do tipo “recortar e colar”, as ocorrências correspondentes de /r/ em coda nas gravações mais

naturais². Como resultado, para cada sentença, obtiveram-se três cópias quase idênticas da gravação mais confortável, sendo a única diferença a pronúncia do /r/ em coda: cada cópia do áudio apresentava nas respectivas palavras-alvo apenas uma das três pronúncias aproximantes.

Esses procedimentos apresentam três vantagens. Em primeiro lugar, ao usar como base a leitura mais confortável, sem modificações articulatórias controladas, obtiveram-se estímulos mais naturais da fala. Segundo, substituindo apenas os segmentos de interesse, garantiu-se um controle completo entre estímulos: nada difere entre os estímulos além das pronúncias de /r/ em coda. Terceiro, ao invés de substituir todas ocorrências de /r/ em coda nas diferentes palavras-alvo por uma única e mesma instância controlada de cada grau de abaixamento do F3, preferiu-se substituir cada ocorrência por um exemplar controlado produzido no mesmo item lexical dentro da mesma frase. Garantiu-se, assim, que a substituição de cada ocorrência se desse por uma versão sujeita às mesmas modificações de natureza coarticulatória e prosódica. O resultado foi a obtenção de estímulos em que a edição digital é imperceptível ao ouvinte. A fim de tornar este experimento mais reprodutível, e em acordo com os recentes esforços na comunidade científica de promover a ciência aberta, os áudios desses estímulos se encontram disponíveis no seguinte repositório: <https://github.com/silveira7/r-coda-sbc>.

O experimento foi aplicado em uma plataforma *online* desenvolvida exclusivamente para esse propósito³ e divulgado em grupos de moradores de São Bernardo do Campo no Facebook e no WhatsApp. Ao acessar o *website*, o participante

² Para evitar um salto de amplitude na gravação editada, os segmentos selecionados foram inseridos de modo que a sua forma de onda se conectasse com a da gravação exatamente no eixo de amplitude zero.

³ O *website* foi desenvolvido pelo autor deste texto com uso das linguagens de programação HTML5, CSS e Javascript, e hospedado na plataforma GitHub. O seu código-fonte encontra-se disponível *online* no mesmo repositório mencionado anteriormente. Os dados fornecidos pelos participantes eram automaticamente enviados e armazenados na plataforma gratuita Simple Form.

era instruído sobre o experimento e solicitado a escutar um estímulo auditivo, aleatoriamente selecionado por um algoritmo da plataforma. Em outras palavras, cada participante escutava apenas um falante (ou a falante mulher ou o falante homem, mas não ambos) em apenas um dos três graus de abaixamento do F3 na pronúncia do /r/ em coda. Em seguida, o participante respondia, com base nas suas impressões acerca da fala que escutara, cinco perguntas (ver Quadro 1), quatro delas com escalas semânticas diferenciais, a respeito do falante e de si mesmo. No final do questionário, o participante era solicitado a informar o seu gênero (feminino ou masculino), sua faixa etária (até 18 anos, 19 a 30 anos, 31 a 60 anos ou mais de 60 anos), se era ou não residente de São Bernardo do Campo e, caso fosse, há quanto tempo residia no município (menos de 1 ano, entre 1 e 5 anos, 6 e 15 anos, mais de 15 anos ou a vida inteira).

Quadro 1 – Questionário respondido pelos participantes durante o experimento. As respostas eram dadas com base nas impressões que tinham acerca do estímulo auditivo.

(1) Você acha que essa pessoa tem...				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Nenhum sotaque	Pouco sotaque	Sotaque moderado	Bastante sotaque	
(2) Você acha que ela é...				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nada simpática	Pouco simpática	Moderadamente simpática	Bastante simpática	Extremamente simpática
(3) Você acha que ela é...				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nada bem-sucedida	Pouco bem-sucedida	Moderadamente bem-sucedida	Bastante bem-sucedida	Extremamente bem-sucedida
(4) Você acha que ela mora...				
<input type="checkbox"/>	Longe de São Bernardo do Campo			
<input type="checkbox"/>	Perto de São Bernardo do Campo			
<input type="checkbox"/>	Em São Bernardo do Campo			
(5) Você fala parecido com essa pessoa?				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nada parecido	Pouco parecido	Moderadamente parecido	Bastante parecido	Extremamente parecido

Fonte: elaborado pelo autor.

O Quadro 1 mostra, na ordem em que foram apresentadas aos participantes, as cinco perguntas do questionário. Com exceção da primeira e da quarta questão, todas as demais foram respondidas numa escala de cinco pontos, de “Nada” a “Extremamente”⁴. A primeira foi elaborada com o objetivo de investigar em que medidas as três variantes aproximantes soam como socialmente marcadas para os ouvintes. A quarta e a quinta visaram aferir em que medida os são-bernardenses se identificam com essas variantes e as reconhecem como característica do município. Por fim, a segunda e a terceira foram criadas para investigar como os são-bernardenses avaliam socialmente essas três variantes.

De acordo com Watson e Clark (2015), as avaliações sociolinguísticas podem ser agrupadas em duas dimensões axiológicas básicas, a da solidariedade (i.e., a dimensão que envolve qualidades como empatia, amigabilidade, simpatia, etc.) e a do prestígio (i.e., que abarca qualidades como sucesso profissional, educação, elegância, etc.). À luz dessa distinção, a segunda e a terceira pergunta buscaram investigar em que medida as avaliações dos falantes nessas duas dimensões se modifica a depender do grau de abaixamento do F3 no /r/ em coda.

3 Resultados

O conjunto de dados coletados foi analisado estatisticamente com uso da linguagem de programação R (R CORE TEAM, 2018). Ele pode ser acessado em sua forma bruta (em arquivo de texto com formatação CSV), em conjunto com o *script* da análise estatística (arquivo de texto escrito em linguagem R), no mesmo repositório *online* em que se encontram os estímulos auditivos.

⁴ Na primeira questão, a dificuldade de encontrar um descritor que soasse gramatical para o quinto ponto levou à adoção de uma escala de quatro pontos.

Foram realizadas algumas transformações a fim de eliminar níveis com poucos dados e tornar a análise estatística mais robusta. No total, 67 pessoas participaram do experimento, mas as respostas de 9 delas foram desconsideradas da análise, pois não faziam parte da população alvo, isto é, não eram residentes de São Bernardo do Campo. A Tabela 2 mostra como ficou a distribuição dos estímulos auditivos, agrupados de acordo com os três graus de abaixamento na frequência do F3, entre os 58 participantes são-bernardenses.

Tabela 2 – Distribuição das frequências absolutas e das proporções de participantes por estímulo auditivo segundo grau de abaixamento na frequência do F3 na pronúncia do /r/ em coda.

Grau de abaixamento na frequência do F3 no estímulo auditivo	Número de participantes
Pouco abaixamento (r1)	18 (31%)
Abaixamento moderado (r2)	18 (31%)
Bastante abaixamento (r3)	22 (38%)

Fonte: elaborada pelo autor.

A Tabela 3 traz a distribuição de frequências do cruzamento das variáveis referentes ao perfil demográfico dos participantes. Predominaram na amostra participantes do sexo feminino e com mais de 30 anos. Os participantes que moram há mais de 15 anos em São Bernardo do Campo compõem 85% da amostra, sendo que 49% deles são nativos, o que indica que os resultados obtidos podem ser considerados representativos das avaliações sociolinguísticas de quem vive em São Bernardo do Campo se não desde nascença, ao menos há vários anos.

Tabela 3 – Distribuição das frequências absolutas e das proporções das variáveis referentes às características demográficas dos participantes.

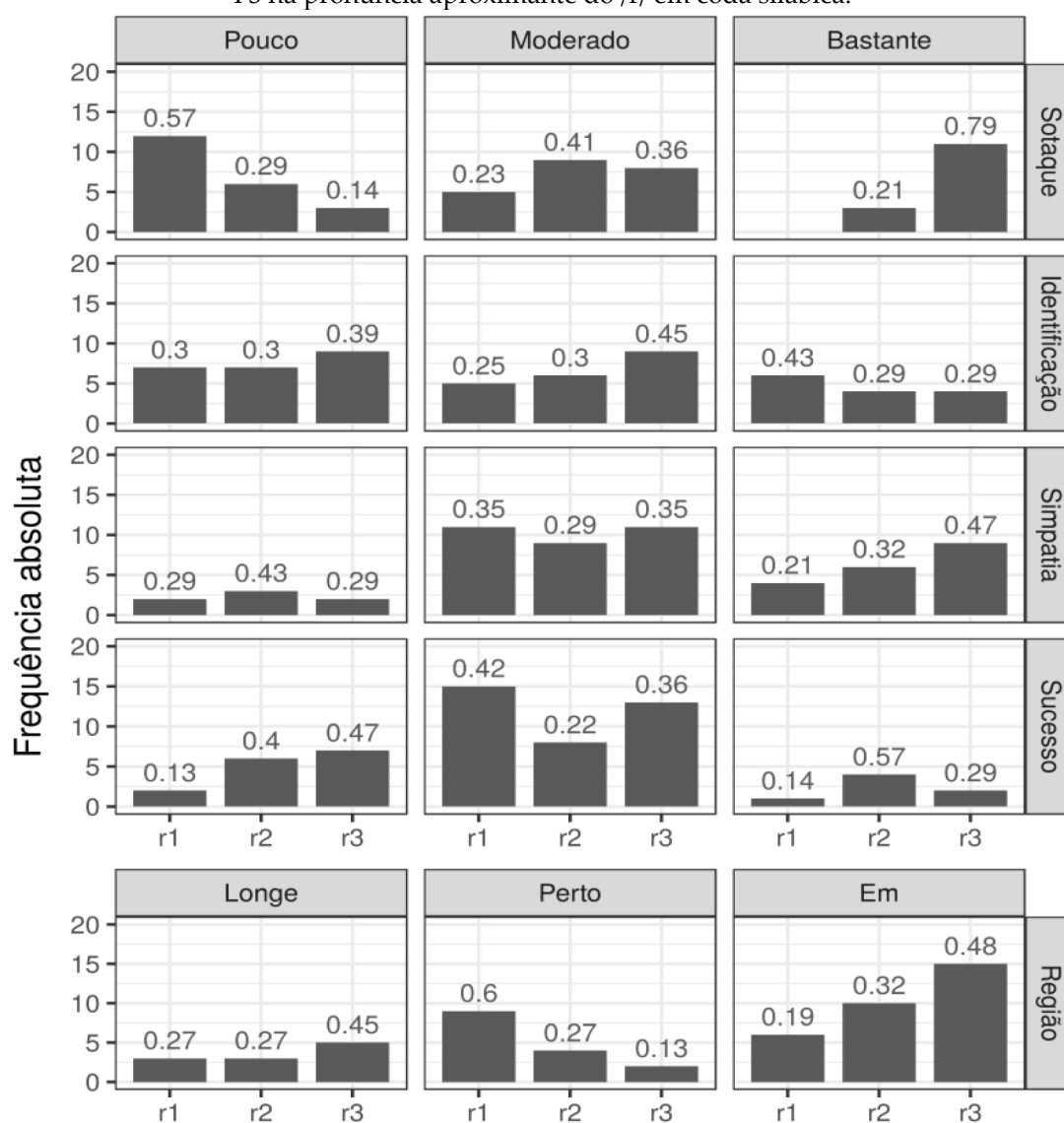
Faixa etária	Tempo residindo em São Bernardo do Campo			N
		Até 15 anos	Mais de 15 anos	
Até 30 anos				
	Feminino	3 (5%)	6 (10%)	9 (15%)
	Masculino	3 (5%)	8 (14%)	11 (19%)
	N	6 (10%)	14 (24%)	20 (34%)
Mais de 30 anos				
	Feminino	1 (2%)	27 (47%)	28 (49%)
	Masculino	2 (3%)	8 (14%)	10 (17%)
	N	3 (5%)	35 (61%)	38 (66%)

Fonte: elaborada pelo autor.

Das quatro variáveis correspondentes às perguntas respondidas em escala semântica diferencial, os níveis extremos (i.e., os pontos “Nada” e “Extremamente”) apresentaram poucas ocorrências em comparação com os valores intermediário (i.e., os pontos “Pouco”, “Moderado” e “Bastante”) e, por isso, foram amalgamados com o respectivo nível adjacente, levando a uma redução dos níveis a três: “Pouco”, “Moderado” e “Bastante”. No caso da primeira questão, na ausência de um quinto ponto, amalgamaram-se apenas os pontos “Nenhum” e “Pouco”. A análise diferenciou as respostas apenas entre os graus de abaixamento do F3 dos estímulos, sem distingui-los em termos de quem os produziu (se a falante mulher ou o falante homem).

A Figura 2 apresenta a distribuição das respostas às cinco perguntas do questionário. Os gráficos em barra exibem a frequência absoluta das respostas em cada nível da escala (i.e., “Pouco”, “Moderado” e “Bastante”) em função dos estímulos auditivos (r1, r2 e r3), e as respectivas proporções se encontram no topo de cada barra.

Figura 2 – Distribuição da frequência absoluta, nas ordenadas, e da proporção, no topo de cada barra, dos níveis das escalas de cada pergunta do questionário em função dos três graus de abaixamento do F3 na pronúncia aproximante do /r/ em coda silábica.



Grau de abaixamento do F3 na pronúncia

Fonte: elaborada pelo autor.

As respostas que mais se diferenciam em função dos estímulos auditivos ocorreram nas escalas de Sotaque e de Região – estas foram as únicas variáveis que apresentaram diferenças no nível estatístico inferencial. No caso de Sotaque, nota-se um padrão consistente: à medida que aumenta a queda na frequência do F3, a frequência de “Pouco” diminui e a de “Bastante” cresce gradativamente. No nível

“Pouco” da escala de Sotaque, constatou-se diferença significativa entre r1 e r3 ($\chi^2 = 5,4$ (1), $p = 0,02$). No nível “Bastante”, além da diferença entre r1 e r3 ($\chi^2 = 11$ (1), $p = 0,0009$), também se verificou diferença entre r2 e r3 ($\chi^2 = 4,5$ (1), $p = 0,03$), mas não entre r1 e r2. Não se observou padrão significativo na distribuição do nível “Moderado”.

No caso de Região, verifica-se aumento consistente na avaliação do falante como sendo morador de São Bernardo do Campo (i.e., no nível “Em”) conforme diminui a frequência do F3 no /r/ em coda. Em sentido contrário, os falantes são progressivamente menos avaliados como residentes de regiões próximas de São Bernardo do Campo (i.e., no nível “Perto”) à medida que a frequência de F3 aumenta. Do ponto de vista inferencial, constataram-se diferenças significativas apenas entre as variantes r1 e r3 em ambos os scores “Perto” ($\chi^2 = 4,45$ (1), $p = 0,03$) e “Em” ($\chi^2 = 3,8$ (1), $p = 0,05$), mas não entre elas e r2. Os estímulos não tiveram efeitos sobre a avaliação dos falantes no caso do nível “Longe”.

4 Discussão

Os resultados das variáveis Sotaque e Região fornecem evidências de que os ouvintes foram capazes de distinguir as três pronúncias aproximantes do /r/ em coda silábica, diferenciadas pelo grau de abaixamento do F3, e de associar informações sociais a essa gradação fonética. Ao menos nessas duas variáveis, foi confirmada a hipótese experimental de que diferenças graduais no abaixamento do F3 na pronúncia aproximante do /r/ em coda silábica podem provocar diferenças igualmente graduais no modo como os falantes são socialmente avaliados em São Bernardo do Campo. Esses resultados têm implicações importantes para a pesquisa variacionista, uma vez que a grande parte dos estudos da variação fonético-fonológica tem trabalhado com uma noção de variável sociolinguística entendida como uma variável categórica não ordenada. Ao definir as variantes aproximantes do /r/ em coda silábica com referência

a uma escala de três graus de abaixamento da frequência de F3, o atual experimento obteve evidências de que diferenças fonéticas mais sutis de natureza escalar também podem ter efeitos significativos na percepção social que os ouvintes têm dos falantes.

Para que sejam possíveis comparações dos resultados aqui apresentados com os verificados em outros trabalhos, será assumido que as variantes aproximantes chamadas por outros pesquisadores de alveolar, pós-alveolar e retroflexa correspondem aproximadamente às três variantes r1, r2 e r3, definidas neste experimento com referência ao grau de abaixamento na frequência do F3.

Os padrões observados na escala de Sotaque se aproximam dos resultados constatados por Mendes e Oushiro (2011) e Oushiro (2015). Semelhante ao que ocorre na cidade de São Paulo, o abaixamento na frequência do F3 na pronúncia da aproximante rótica funciona em São Bernardo do Campo como índice de sotaque. O aumento consistente na avaliação dos falantes como tendo mais sotaque à medida que aumenta na fala deles o grau de abaixamento do F3 consiste em importante evidência de que os significados sociais associados ao /r/ em coda são suscetíveis a diferenças fonéticas graduais envolvendo um mesmo modo de articulação.

Assumindo que a noção de sotaque funcione como um modo de qualificar a saliência social das formas linguísticas, e com base na distribuição da escala de Sotaque, pode-se concluir que, como os paulistanos, os são-bernardenses avaliam as pronúncias mais “róticas” (aquelas com maior queda de F3) do /r/ em coda como socialmente mais marcadas. No entanto, ao contrário dos paulistanos, que, de acordo com os resultados de Oushiro (2019) apresentados anteriormente, não avaliam essa pronúncia comumente chamada de “caipira” como característica da fala paulistana, os são-bernardenses a reconhecem como marca do seu próprio município.

Num primeiro olhar, os padrões observados nas escalas de Sotaque e Região, quando considerados em conjunto, parecem ir na direção contrária à hipótese de Soriano (2016), de acordo com a qual nascer e crescer em regiões onde a exposição às

variantes aproximantes é maior faz com que a diferença entre as formas alveolar e retroflexa seja menos saliente para o ouvinte. Se a variante com maior abaixamento de F3 for, de fato, a forma predominante do /r/ em coda em São Bernardo do Campo, então, seguindo essa hipótese, seria de se esperar que os padrões observados na escala de Sotaque e no nível “Em” da escala de Região fossem inversos: formas avaliadas como mais próprias do município seriam vistas como menos marcadas (i.e., com menos sotaque).

É um fato bastante consolidado nas pesquisas sociolinguísticas que não há necessariamente uma correspondência entre o modo como as pessoas se comportam linguisticamente e o modo como elas avaliam o seu próprio comportamento e o dos demais. Estudos sociolinguísticos observaram casos em que os ouvintes eram pouco acurados na avaliação da sua própria fala (LABOV, 2008; PRESTON, 2013). É possível que, apesar de os são-bernardenses compartilharem com os paulistanos o discurso metalinguístico de que o seu município é uma região com maior frequência do “r retroflexo” do que as regiões próximas, como sugerem os resultados da escala de Região, essa pronúncia não seja, de fato, a mais produtiva na fala deles. Em outras palavras, não se pode descartar a possibilidade de que essa crença seja incorreta e não corresponda aos reais padrões de comportamento linguístico do município⁵.

Com base nessas premissas, é possível explicar a aparente contradição dos resultados com a hipótese de Soriano (2016). Se for assumido que a variante r3 não seja tão produtiva na fala dos são-bernardenses quanto eles acreditam, então, em conformidade com a hipótese de Soriano, se espera que ela seja percebida como mais saliente, o que explica o fato de que 79% das respostas indicando “Bastante sotaque” se concentram no estímulo com r3. Por outro lado, sob influência do discurso

⁵ Até o momento não há estudos da produção do /r/ em coda em São Bernardo do Campo que permitam resolver essa questão.

metalinguísticos contrário, isto é, de que essa forma é recorrente na região, os são-bernardenses percebem os falantes mais frequentemente como moradores de São Bernardo quando ouvidos com pronúncia com bastante abaixamento na frequência do F3. Se essa explicação estiver na direção correta, as respostas dos participantes nessas duas escalas seriam, portanto, resultado da influência de duas dimensões diferentes, a da produção (i.e., os próprios comportamentos linguísticos dos ouvintes), no caso da escala de Sotaque, e a da avaliação (i.e., os valores e os discursos associados aos comportamentos linguísticos), no caso da escala de Região.

Outro resultado que chama atenção refere-se à distribuição de “Perto” na escala de Região. Os residentes de São Bernardo do Campo não apenas avaliam as pronúncias com maior abaixamento na frequência do F3 como uma marca do município, mas também as veem como um traço que o distingue das cidades vizinhas. É possível que esse resultado esteja associado à crença corrente entre paulistas de que falantes da cidade de São Paulo realizam o /r/ em coda como tepe e de que isso os diferencia dos demais paulistas, em especial daqueles que moram no interior. Em outros termos, os são-bernardenses, com base nessa crença, e tendo como referência a vizinha cidade de São Paulo, podem ter avaliado o seu município como um reduto de pronúncias mais “caipiras” no entorno dos tepes dos paulistanos.

Em relação às variáveis Simpatia e Sucesso, as diferenças no grau de abaixamento na frequência do F3 não tiveram efeitos significativos sobre as avaliações dos ouvintes. Aqui deve-se notar que no experimento de Oushiro (2019) também não se observou efeitos da pronúncia variável do /r/ sobre a percepção do falante em termos de amigabilidade, extroversão e inteligência, que são qualidades que se aproximam das noções de simpatia e sucesso. No entanto, é possível identificar, no nível descritivo, dois padrões que estão em conformidade com padrões recorrentes na literatura sobre percepção sociolinguística. Essa literatura tem constatado que, de modo geral, as variantes padrão tendem a ser mais bem avaliadas em termos de

prestígio, enquanto as variantes não padrão, em termos de solidariedade (PRESTON, 2013). As distribuições das respostas no nível “Bastante” da variável Simpatia e no nível “Pouco” da variável Sucesso apontam justamente nesse sentido. A maior parte das respostas que consideraram os falantes como bastante simpáticos se concentra nos estímulos com a pronúncia r3, que pode ser vista como a pronúncia mais não padrão das três variantes. Se, por um lado, os falantes foram avaliados como mais simpáticos quando escutados com variante r3, por outro, essa variante foi a que mais os levou a serem percebidos como pouco bem-sucedidos.

A distribuição da escala de Identificação não apresentou diferenças significativas entre os estímulos auditivos. Há pelo menos duas maneiras de interpretar esses resultados. Uma possibilidade é vê-los como indicação de que, de modo geral, os participantes se identificam igualmente com as três variantes da pronúncia aproximante. Outra é interpretá-los como indício de que essa escala não foi capaz de eliciar com acurácia o grau de identificação do ouvinte com essas variantes. Se for assumida a primeira opção, a questão mais premente a ser respondida é o porquê de a escala de Identificação não ter se correlacionado com a escala de Região. É natural supor que, se os participantes avaliam o aumento do grau de abaixamento na frequência do F3 como indício de que os falantes moram em São Bernardo do Campo, os participantes, sendo eles mesmos são-bernardenses, também se identificariam mais com a pronúncia com mais abaixamento. A hipótese que aqui se faz, para que estudos posteriores possam testá-la, é de que, apesar de verem a variante r3, dita “retroflexa”, como marca do município, os são-bernardenses não se identificam com ela, talvez em decorrência do estigma social que acompanha o /r/ “caipira” no Brasil.

5 Considerações finais

O experimento discutido neste artigo teve o objetivo de começar a explorar no contexto do português brasileiro uma abordagem mais dinâmica da percepção sociolinguística, tentando abarcar diferenças graduais de um traço fonético, o abaixamento na frequência do F3 na pronúncia aproximante do /r/ em coda silábica. Os resultados mostraram que os ouvintes são capazes de perceber diferenças graduais no abaixamento do F3 e de atribuir significados sociais a essa gradação fonética. Em conformidade com os resultados de estudos anteriores sobre as percepções sociais do /r/ na cidade de São Paulo, os significados sociais atribuídos pelos ouvintes são bernardenses a essa variável se referem principalmente ao grau de sotaque do falante e à sua região de origem. No entanto, diferente dos paulistanos, os são bernardenses reconhecem as pronúncias mais “caipiras” como típicas do seu próprio município, embora resistam em se identificar com elas.

Referências

AMARAL, A. **O dialeto caipira**. São Paulo: Parábola, 2020.

ATAL, B. S.; HANAUER, S. L. Speech analysis and synthesis by linear prediction of the speech wave. **The Journal of the Acoustical Society of America**, v. 50, n. 2B, p. 637–655, 1971. DOI: <https://doi.org/10.1121/1.1912679>

BOERSMA, P.; WEENINK, D. **Praat**: doing phonetics by computer [Computer program]. Disponível em: <http://www.praat.org/>.

BRANDÃO, S. F. Nas trilhas do -r retroflexo. **SIGNUM: Estud. Ling**, 2007. DOI: <https://doi.org/10.5433/2237-4876.2007v10n2p265>

CAMPBELL-KIBLER, K. The nature of sociolinguistic perception. **Language Variation and Change**, v. 21, n. 1, p. 135–156, 2009. DOI: <https://doi.org/10.1017/S0954394509000052>

DELATTRE, P.; FREEMAN, D. C. A dialect study of American r's by x-ray motion picture. *Linguistics*, v. 44, p. 29–68, 1968. DOI: <https://doi.org/10.1515/ling.1968.6.44.29>

LABOV, W. **Padrões Sociolinguísticos**. Tradução: Marcos Bagno, Maria Marta Pereira Scherre, Caroline R. Cardoso. São Paulo: Editora Parábola, 2008.

LADEFOGED, P. **Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques**. Malden: Blackwell Publishing, 2003.

LADEFOGED, P.; MADDIESON, I. **The Sounds of the World's Languages**. Oxford: Blackwell Publishers, 2009.

LAMBERT, W. E. *et al.* Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, v. 60, n. 1, p. 44–51, 1960. DOI: <https://doi.org/10.1037/h0044430>

LEITE, C. M. B. **Atitudes lingüísticas: a variante retroflexa em foco**. Dissertação de mestrado—Campinas: Universidade de Campinas, 2004. Disponível em: <http://repositorio.unicamp.br/jspui/handle/REPOSIP/268969>

LEITE, C. M. B. **O /r/ em posição de coda silábica no falar campineiro**. Tese de doutorado—Campinas, SP: Universidade de Campinas, 2010. Disponível em: <http://repositorio.unicamp.br/jspui/handle/REPOSIP/268968>

LINDAU, M. The story of r. *In*: FROMKIN, V. A. (ed.). **Phonetic Linguistics**. Orlando: Academic Press, 1985.

MENDES, R. B.; OUSHIRO, L. Percepções sociolinguísticas sobre as variantes tepe e retroflexa na cidade de São Paulo. *In*: DA HORA, D.; NEGRÃO, E. V. (ed.). **Estudos da linguagem: casamento entre temas e perspectivas**. João Pessoa: Ideias, 2011. p. 229–245.

MENDES, R. B.; OUSHIRO, L. O paulistano no mapa sociolinguístico brasileiro. **ALFA: Revista de Linguística**, v. 56, n. 3, p. 973–1001, 2012. DOI: <https://doi.org/10.1590/S1981-57942012000300011>

OUSHIRO, L. **Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo**. Tese de doutorado—São Paulo: Universidade de São Paulo, 2015. DOI:

OUSHIRO, L. A computational approach for modeling the indexical field. **Revista de Estudos da Linguagem**, v. 27, n. 4, p. 1737–1786, 2019. DOI: <https://doi.org/10.17851/2237-2083.0.0.1737-1786>

PRESTON, D. R. Language with an Attitude. *In*: CHAMBERS, J. K.; SCHILLING, N. (ed.). **The Handbook of Language Variation and Change**. Oxford, UK: John Wiley & Sons, 2013. p. 157–182. DOI: <https://doi.org/10.1002/9781118335598.ch7>

R CORE TEAM. **R: A language and environment for statistical computing**. [s.l.] R Foundation for Statistical Computing, 2018. Disponível em: <https://www.r-project.org/>.

SORIANO, L. G. M. **Percepções sociofonéticas do (-r) em São Paulo**. Dissertação de mestrado—São Paulo: Universidade de São Paulo, 2016. DOI:

WATSON, K.; CLARK, L. Exploring listeners' real-time reactions to regional accents. **Language Awareness**, v. 24, n. 1, p. 38–59, 2015. DOI: <https://doi.org/10.1080/09658416.2014.882346>

Artigo recebido em: 11.03.2021

Artigo aprovado em: 22.10.2021