

# PERCEPÇÕES FONÉTICAS DO (-R) EM SÃO PAULO: PRINCIPAIS CORRELAÇÕES

**Larissa Grasiela Mendes Soriano\***  
**Ronald Beline Mendes\*\***

**Resumo:** Em entrevistas sociolinguísticas (MENDES; OUSHIRO, 2012), paulistanos tecem comentários sobre “graus de retroflexão” do (-r) em coda: uma “pronúncia forte” (que corresponderia à retroflexa) e uma “pronúncia fraca” (a aproximante alveolar). Supostamente, a primeira ocorreria na fala de migrantes do interior, enquanto a segunda seria mais comum na fala de certos paulistanos. Essas duas variantes ainda se oporiam ao tepe, mais “prototipicamente paulistano”. O presente artigo discute resultados de um experimento que mediu quão salientes são essas sutilezas acústicas para diferentes grupos de paulistanos.

**Palavras-chave:** Percepção sociofonética. /-r/ em coda. São Paulo.

## INTRODUÇÃO

■ **É** plausível imaginar que todas as pessoas ouçam os sons da mesma maneira, uma vez que, *a priori*, todos os indivíduos que não possuem distúrbios auditivos têm a mesma capacidade de ouvir os sons. No entanto, mais do que um processo físico (em que as ondas sonoras, mecânicas, vibram nos órgãos auditivos do ouvinte), a capacidade de distinguir os sons é também um processo cognitivo e, por isso, está intimamente ligado ao repertório (tanto linguístico quanto social) de cada indivíduo. Essa capacidade está relacionada, portanto, tanto ao sistema fonético-fonológico do dialeto da língua materna (BYBEE, 2001) que o indivíduo adquiriu, como também a ideologias (IRVINE,

\* Universidade de São Paulo (USP) – São Paulo – SP – Brasil. E-mail: larissa.soriano@gmail.com

\*\* Universidade de São Paulo (USP) – São Paulo – SP – Brasil. E-mail: rbeline@usp.br

2001) acerca das variantes linguísticas a que está exposto e aos significados sociais a elas associados.

O intuito deste trabalho foi investigar como indivíduos de diferentes grupos sociais, falantes de uma mesma variedade, percebem as diferenças acústicas entre os *sons*. No experimento de percepções sociolinguísticas (CAMPBELL-KIBLER, 2006, 2007; HAY; DRAGER, 2007) aqui apresentado, pessoas nascidas na cidade de São Paulo atribuíram graus de diferença a pares formados por duas de cinco variantes de (-r) em coda silábica presentes no dialeto paulistano: 1. vibrante com três “batidas” da língua nos alvéolos; 2. vibrante com duas batidas; 3. tepe; 4. aproximante alveolar; e 5. retroflexo (MENDES, 2010; OUSHIRO, 2015; CALLOU et al., 1996, 2002 apud OUSHIRO, 2015).

Os resultados mostram que as percepções dos participantes variam de forma significativa e sistemática de acordo com as características sociais do ouvinte: seu sexo/gênero e sua região de residência na cidade (central ou periférica).

## BASES E JUSTIFICATIVA

No que concerne ao dialeto paulistano, as variantes tepe e retroflexa são as mais produtivas na fala das pessoas nascidas na cidade de São Paulo (MENDES, 2010; OUSHIRO; MENDES, 2013; OUSHIRO, 2015). A pronúncia tepe é a mais frequente em todas as regiões da cidade, em todas as faixas etárias e graus de escolaridade analisados. De fato, o tepe funciona como “estereótipo” (LABOV, 1972) desse dialeto, uma vez que mesmo pessoas não estudosas da língua se referem a essa variante como algo característico desse “sotaque”. No entanto, em uma amostra representativa da cidade, Oushiro (2015, p. 104) mostra que o retroflexo corresponde a 28,3% das 9.226 ocorrências de (-r) em coda efetivamente pronunciadas (desconsiderando-se, portanto, os casos de apagamento). Nesses trabalhos sobre a *produção* linguística dos falantes, não foi feita uma análise acústica das pronúncias. Assim, para esses resultados em São Paulo, não se diferenciaram “tepe” e “vibrante”, pois ambos se oporiam conjuntamente ao retroflexo; assim como também não foram efetivamente analisadas as diferenças entre as pronúncias aproximante alveolar e retroflexa, uma vez que ambas funcionam como “não tepe”. Dados do Projeto Nurc (CASTILHO; PRETI, 1986) indicam que outras variantes de (-r) também eram produzidas no falar culto paulistano da década de 1970 (mesmo que em quantidades menos expressivas), principalmente a vibrante alveolar, que representou 3% dos /-r/ produzidos em posição de coda medial e 7% na coda final (CALLOU et al., 1996, 2002 apud OUSHIRO, 2015). Esses resultados não mostram, contudo, quais grupos sociais privilegiariam o uso dessas vibrantes.

Embora sejam produtivas nessa variedade, não se pode dizer que tepe e retroflexo tenham o mesmo valor social na comunidade. Oushiro (2015) mostra que o retroflexo é favorecido por falantes do sexo masculino, menos escolarizados, moradores das regiões mais periféricas da cidade, com menor mobilidade geográfica dentro de São Paulo e cujos pais não são paulistanos. De forma geral, essa distribuição aponta para o fato de que, além de a fala paulistana estar longe de ser homogênea no que diz respeito à produção de /-r/, a classe social e a região de residência na cidade parecem ser fatores fundamentais para a diferenciação sociolinguística, bem como o prestígio das variantes. Os resultados em

relação à centralidade do bairro possivelmente indicam uma diferenciação de *status* das variantes: uma vez que as classes mais altas se concentram em bairros mais centrais e as mais baixas em bairros mais periféricos, a associação do tepe com os primeiros também indica uma associação com classe social.

Mendes e Oushiro (2011) e Oushiro (2015) desenvolveram testes de percepção utilizando a técnica *matched-guise* (LAMBERT et al., 1960 apud CAMPBELL-KIBLER, 2006), a fim de desvendar os significados sociais do tepe e do retroflexo na cidade e as reações subjetivas dos paulistanos acerca delas. Nesse sentido, eles selecionaram excertos da fala de quatro informantes paulistanos (dois homens e duas mulheres com cerca de 30 anos, moradores da zona oeste da cidade) e os manipularam digitalmente, de forma a obter dois estímulos para cada excerto: um em que só consta a pronúncia tepe e outro que só contém a retroflexa. Como cada par de estímulos difere apenas no que diz respeito às pronúncias da coda /-r/, infere-se que diferenças de percepção acerca deles correspondem a tais variantes. Cada participante ouviu um trecho por falante (ou seja, um mesmo participante só ouviu um trecho com tepe ou com retroflexo de cada um dos quatro falantes) e, em seguida, preencheu um questionário de acordo com a imagem que construiu a partir do trecho ouvido. Mendes e Oushiro (2011) aplicaram o questionário a 24 paulistanos, 13 homens e 11 mulheres, todos universitários, com 20 a 30 anos (ou seja, um grupo bastante homogêneo). Esses ouvintes ressaltaram diferenças estatisticamente significativas quanto a *paulistanidade*, *sotaque* e *centralidade do bairro*, a depender do estímulo. Os falantes, quando ouvidos na versão retroflexa, foram percebidos como “menos paulistanos”, com “mais sotaque”, e como moradores de regiões “menos centrais”, relativamente aos estímulos com tepe. Oushiro (2015) ampliou o mesmo teste para uma amostra sociolinguisticamente mais diversificada de 185 ouvintes (129 deles paulistanos) moradores da cidade de São Paulo e, mais uma vez, as três escalas que se relacionam a identidades geográficas foram as que apresentaram as diferenças mais significativas a depender da variante de (-r) ouvida.

Já em uma análise do discurso metapragmático de 60 informantes do Projeto SP2010 (MENDES; OUSHIRO, 2012), Soriano (2014) verificou que os paulistanos espontaneamente identificam dois “tipos” diferentes de retroflexo, que se diferenciariam pelo seu “grau de retroflexão” (embora não tenham usado essa metalinguagem): a variante “mais forte” é geralmente associada a um falar típico do interior, enquanto a “mais fraca” seria pronunciada na cidade de São Paulo:

1. S1<sup>1</sup>: [...] eu procuro ler da/ da maneira mais sem sotaque nenhum... mas a gente falando às vezes informalmente aí sai de qualquer jeito né  
 D1: e e... falando sem sotaque que você diz aí era diferente o erre seria diferente?  
 S1: ah sim... é  
 D1: e aí co/ era como?  
 S1: ah o erre assim do do jeito que eu li né ‘a porta [r] (es)tá aberta’ [r] agora se está no corre-corre aí “fecha ali que a ‘porta [ɹ] (es)tá aberta” [ɹ] quer dizer já sai não assim daquele jeito igual Sabrina Sato ‘aberta’ [ɹ] mais né... mas sai mais assim (Rodolfo R., M3CP).

1 “D1” refere-se ao documentador e “S1” ao informante.

Nesse exemplo, o informante cita esses dois tipos de retroflexão, aquele que ele (paulistano) pronunciaria quando estivesse “na correria” – ou seja, quando não estivesse prestando atenção à própria fala – e outro que lhe parece mais “acentuado”, caracterizando-o como o da apresentadora Sabrina Sato, natural do interior do estado de São Paulo e conhecida por manter o seu sotaque “original” proeminente. Esses dois “tipos de retroflexo” ainda se oporiam ao tepe, ao mesmo tempo “mais formal” e “sem sotaque”, que ele usaria quando estivesse monitorando mais a sua fala, ao fazer leituras.

2. D1: tem algum sotaque do Brasil que você goste mais assim?

S1: eu gosto do sotaque paulistano [risos] [...] uma vez eu saía com uma menina que ela que ela falava tudo ela falava bem assim... ela falava até os erres assim “ah eu vou ver” [pronuncia todas as palavras pausadamente e com o tepe no final de “ver”] não sei acha/ porque também ela fazia teatro então ela marcava bem mas ela fazia um sotaque bem paulistano eu achava engraçado... assim paulistano esse né porque tem paulistano também do do in/ da periferia né que é um pouco diferente

D1: como é que é da periferia?

S1: não é uma coisa ma/ eh o erre já é um o erre já é um pouco/ pouco parecido com o erre do interior né aquela coisa assim... aí tinha a minha namora/ a minha ex-namorada ela falava igual São Paulo só que ela fazia às vezes ela puxava o erre pouco um pouco diferente assim um pouco mais... acho que da periferia não sei mas acho sim... acho que era por aí porque ela viajava pro interior também e ela morou também na na Zona Sul não nas quebradas né (Rodrigo R., M1SC).

Em (2), Rodrigo R. associa dois tipos de /-r/ ao sotaque paulistano: o tepe e o outro mais próximo de um falar típico do interior, que ele associa a regiões mais periféricas da cidade. Em ambos os excertos, os informantes se referem a esses dois tipos de /-r/ na cidade de São Paulo, que seriam diferentes do /-r/ típico do interior.

Comentários como esses, sobre “diferentes graus de retroflexão”, indicam que certos paulistanos atentam para as diferenças acústicas entre a aproximante alveolar e o retroflexo, e que tal diferença fonético-fonológica (bastante sutil) pode ser acompanhada de significação social. Por outro lado, metacomentários desse tipo não permitem concluir que tais avaliações são sistemáticas na comunidade, ou seja, há que se verificar se diferentes paulistanos percebem essas diferenças e se a elas se correlacionam fatos de natureza social.

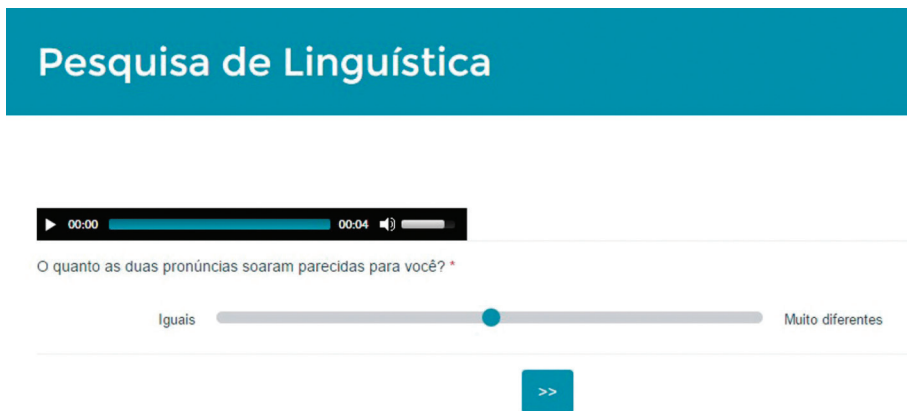
Nesse sentido, realizou-se um experimento com o objetivo de *medir* o quanto os ouvintes percebem como diferentes cinco variantes de (-r). Estudos de percepções sociofonéticas (FOULKES; SCOBIE; WATT, 2010; HAY; DRAGER, 2007; DRAGER, 2010; MENDES, 2014; OUSHIRO, 2015, entre outros) indicam que mesmo diferenças fonéticas mais sutis podem carregar significado social de maneira sistemática e exploram como percepção e produção linguística se relacionam. Alguns trabalhos, como o Ladefoged e Broadbent (1957) e Willis (1972 apud DRAGER, 2010) indicam que as variantes observadas na fala de um indivíduo podem refletir a forma como ele percebe os sons, enquanto outros (como Hay et al., 2006 apud DRAGER, 2010) mostram que os ouvintes podem ser bastante precisos em identificar alguns sons, mesmo que não os distingam em

sua produção. Já Niedzielski (1999) evidencia que não só as variáveis fonéticas influenciam o modo como as características sociais são apreendidas, mas que o fornecimento de certas informações sociais pode influenciar como variantes fonéticas são classificadas pelos ouvintes.

Seguindo esta linha de investigação entre as relações da fonética com o mundo social, o trabalho de percepção que aqui se apresenta averigua quais realizações, dentre as cinco variantes de (-r) incluídas, os informantes paulistanos assinalam como significativamente diferentes, a fim de entender até que ponto as “imagens” que eles constroem acerca dos falantes que pronunciam essas variantes de /-r/ podem afetar o modo como eles percebem as diferenças fonéticas entre elas.

## METODOLOGIA

O experimento consistiu em apresentar aos participantes diferentes pares formados por um mesmo item lexical, nos quais a única diferença é o segmento rótico – uma vibrante com três batidas (daqui em diante, “vibrante 3” ou “3”); uma vibrante com duas batidas (“vibrante 2” ou “2”); tepe (1); aproximante alveolar (A); aproximante retroflexa (R). Para cada par ouvido (p.ex., “pardal-2xA” – ou seja, “pardal” primeiro pronunciado com a vibrante 2 e, em seguida, pronunciado com a aproximante alveolar), os participantes deveriam responder quão diferentes entre si as pronúncias lhes soavam, clicando no ponto escolhido numa escala que vai de “Igual” (0 – zero) a “Muito Diferente” (100 – cem) (Figura 1). A partir desses valores numéricos, realiza-se a análise estatística das respostas dadas pelos participantes para cada par.



**Figura 1** – Tela de respostas do experimento na plataforma online SurveyGizmo

Os participantes não foram informados sobre o foco do teste. Assim, ao indicar que duas pronúncias de uma mesma palavra são iguais, por exemplo, isso significa que a diferença fonética entre as duas variantes de (-r) não lhes foi relevante. Evidentemente, não se pode afirmar que o participante “não ouviu uma diferença acústica”, mas sim que ele não *atribuiu* um grau de diferença a tal par.

Para que os ouvintes não “automatizassem” suas respostas (DRAGER, 2013), foram incluídos no experimento pares distratores formados por itens lexicais que não contivessem /-r/ em posição de coda silábica, mas que também variassem entre si, p.ex., “grande” com e sem a palatalização do /d/, “fazenda” com e sem a ditongação do /e/ nasal etc.

Os itens lexicais usados como palavras-alvo foram definidos a partir de Oushiro (2015): a vogal precedente, o contexto fônico seguinte, a classe da palavra e a tonicidade da sílaba são variáveis estatisticamente relevantes para a variação na produção do /-r/. O Quadro 1 apresenta os itens lexicais escolhidos:

**Quadro 1** – Itens lexicais que formam os estímulos

		t	d
tônica	i	Airton	—
	e	aperto	perda
	ɛ	oferta	cerda
	a	carta	farda
	ɔ	porta	corda
	o	aborto	gordo
	u	curto	surdo
átona	i	virtude	—
	e	certeza	verdade
	a	artéria	pardal
	o	português	mordomo
	u	urtiga	surdez

Fonte: Elaborado pelos autores.

Para cada item lexical foram formados 10 pares, resultantes de todas as combinações possíveis entre as variantes, por exemplo, “urtiga” pronunciado com a vibrante 3 seguido de “urtiga” pronunciado com a vibrante 2 (urtiga-3x2); “urtiga” com a vibrante 3 seguido de “urtiga” pronunciado com o tepe (urtiga-3x1) e assim por diante, de modo que as cinco variantes fossem contrastadas entre si em cada item lexical, o que leva a um total de 220 pares de estímulos.

Todas as características acústicas dos estímulos foram controladas (DRAGER, 2013; CAMPBELL-KIBLER, 2006, 2009), uma vez que, ao focalizar o detalhe fonético, duas variantes não devem ser contrastadas apenas de acordo com o que se percebe auditivamente (já que aqui mesmo se supõe que essas percepções variam). Como não é possível saber quais características do falante podem suscitar diferentes percepções, sejam elas em relação ao perfil social ou de qualquer particularidade da fala (CAMPBELL-KIBLER, 2006), decidiu-se por usar somente uma “voz” para o experimento e manter controladas quaisquer características que pudessem desviar a atenção dos ouvintes das pronúncias do /-r/. Assim, para cada

item lexical, um mesmo falante gravou<sup>2</sup> as cinco variantes de (-r), que posteriormente foram analisadas e editadas nos programas Praat e Audacity. Para cada item lexical foi escolhida uma pronúncia como base, sobre a qual foram coladas as outras quatro pronúncias do /-r/. Por exemplo, na palavra “português”, foi escolhida como base aquela que tinha sido produzida com a vibrante 2. A vibrante foi recortada da produção original e, em seu lugar, foi colada uma vibrante 3 que havia sido produzida em outra versão da mesma palavra, dando origem a um novo estímulo. Procedeu-se da mesma forma com o tepe, a aproximante alveolar e o retroflexo. Como para cada item lexical manteve-se uma base comum, as características acústicas (como entonação e silabação) permaneceram idênticas em cada par, exceto pela variante de (-r), o que permite dizer que as diferenças indicadas pelos ouvintes referem-se a esse segmento.

O teste foi aplicado online através da plataforma *SurveyGizmo* e divulgado pela internet, de modo que pessoas de diferentes grupos sociais pudessem participar. Foram coletadas 109 respostas de moradores da cidade de São Paulo, distribuídos por diferentes grupos sociais.

## RESULTADOS

As análises objetivam verificar se as diferenças nas percepções por parte dos ouvintes foram aleatórias ou sistemáticas dentro da amostra coletada e se alguma (e qual) característica social ou linguística as influenciam. Desse modo, realizaram-se testes estatísticos no programa R (R CORE TEAM, 2013), a fim de verificar se havia correlações entre, de um lado, variáveis sociais ou linguísticas e, de outro, a variação nas respostas.

### Quadro 2 – Distribuição dos participantes de acordo com suas características sociais

		Feminino (65)		Masculino (44)	
		Centro (36)	Periferia (29)	Centro (25)	Periferia (19)
<b>Região de Nascimento</b>					
SP Capital	(67)	18	20	13	16
Interior de SP, PR, MG, GO	(24)	12	3	7	2
Grande SP	(9)	2	4	2	1
NE, Brasília, RJ, parte de MG	(9)	4	2	3	0

Fonte: Elaborado pelos autores.

O Quadro 2 indica que foram coletadas também respostas de pessoas nascidas em outros lugares do Brasil, no intuito de verificar se sua Região de Nasci-

<sup>2</sup> Os estímulos foram gravados em uma cabine acústica na Universidade de São Paulo (USP). Utilizou-se um gravador Tascam modelo DR-100mkl, com microfone Sennheiser HMD 26.

mento<sup>3</sup> (e, indiretamente, o dialeto que adquiriu quando criança) influencia o grau de diferença atribuído às variantes. No interesse do espaço, contudo, não se discutem aqui os resultados para essa variável; a discussão delimita-se às respostas dos participantes paulistanos.

### Quadro 3 – Variáveis analisadas

<b>Variável dependente</b>	Grau de diferença atribuído	0 a 100
<b>Variáveis linguísticas</b>	Tonicidade	Átona, tônica
	Vogal precedente	[i, e, ε, a, ɔ, o, u]
	Consoante seguinte	[t,d]
	Par de variantes de (-r)	3x2, 2x1, 1xA, AxR, 3x1, 2xA, 1xR, 3xA, 2xR, 3xR
<b>Variáveis sociais</b>	Sexo/Gênero	Masculino, feminino
	Local de residência na cidade	Centro, periferia
<b>Variáveis aleatórias</b>	Participante	
	Estímulo	

Fonte: Elaborado pelos autores.

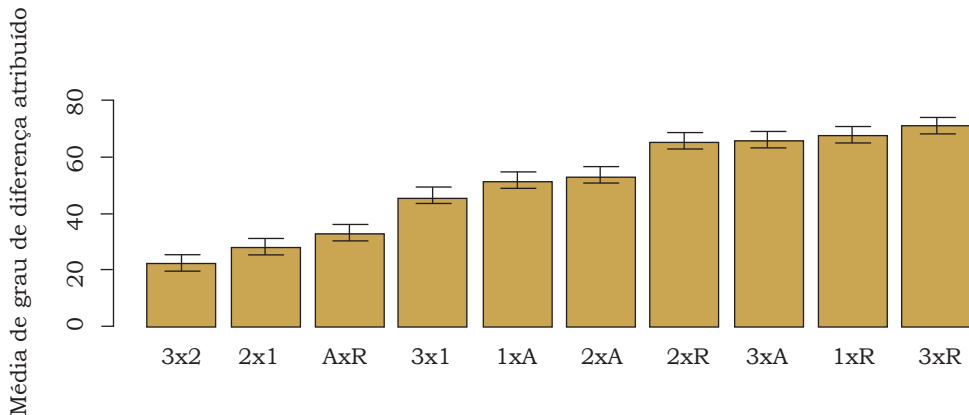
As variáveis linguísticas no Quadro 3 incluem aquelas usadas como critérios de escolha dos estímulos, enquanto as variáveis sociais dizem respeito às características dos ouvintes. Foram realizadas análises multivariadas (BAAYEN, 2008) que incluem dois tipos de variáveis independentes: as fixas e as aleatórias. As primeiras dizem respeito às variáveis que são representativas da população, como o Sexo/Gênero e a Tonicidade, e que podem ser replicadas em estudos futuros. Já as variáveis aleatórias são aquelas que dizem respeito apenas à amostra coletada, ou seja, os Participantes e os Estímulos apresentados (OUSHIRO, 2015).

A Tonicidade da Sílabas, o Contexto Fônico Seguinte e a Vogal Precedente não se mostraram correlacionados com o grau de diferença atribuído, diferentemente da variável Par ( $p < 0,001$ ), justamente a que diz respeito à comparação entre as variantes de (-r). Esses resultados mostram que os contextos linguísticos não influenciaram o modo como os participantes ouviram as variantes; mas que, por outro lado, eles atribuíram graus de diferença maiores ou menores a depender do par de variantes que estavam sendo contrastadas (se vibrante 3 *versus* vibrante 2 (3x2), vibrante 2 *versus* retroflexo (2xR), aproximante alveolar *versus* retroflexo (AxR) e assim por diante).

O Gráfico 1 mostra as médias dos graus de diferença atribuídos a cada par com os respectivos intervalos de confiança (ou seja, a variabilidade prevista para cada um deles). Quando os intervalos de confiança se sobrepõem, há maior probabilidade de que os resultados sejam estatisticamente equivalentes.

3 Agruparam-se os participantes nascidos no interior de São Paulo, Paraná, Goiás e sul de Minas Gerais porque, nessas regiões, predomina o uso da variante retroflexa do (-r). Por outro lado, agruparam-se os participantes que nasceram em estados do Nordeste, Brasília, Rio de Janeiro e norte de Minas Gerais porque nessas regiões predominam as variantes aspiradas.





**Gráfico 1** – Médias de grau de diferença atribuído para cada par

Fonte: Elaborado pelos autores.

No Gráfico 1, os pares estão dispostos em ordem crescente de acordo com a média das diferenças atribuídas. Os pares que contrastam as vibrantes 3 e 2 (3x2) e 2 e tepe (2x1), ou seja, nos quais as variantes se distinguem por uma batida da língua nos alvéolos, receberam as menores notas. Já a diferença entre a vibrante três e o tepe (3x1), ou seja, diferença de duas batidas, foi mais saliente.

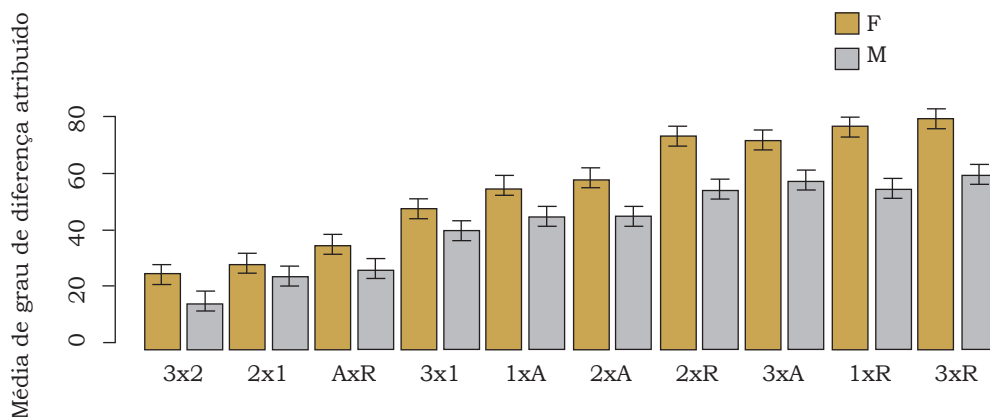
Em seus metacomentários, alguns informantes do Projeto SP2010 haviam dito perceber uma diferença acústica entre as aproximantes alveolar e retroflexa; no presente experimento, contudo, ao par que contrasta essas duas aproximantes (AxR) foi atribuída uma das menores diferenças, com média de 32,6 (dos 100 possíveis). Por outro lado, quando contrastadas ao tepe (a variante mais tipicamente paulistana), as duas aproximantes receberam notas distintas. A média dos graus de diferença atribuídos ao par tepe *versus* aproximante alveolar (1xA) foi 51,2, enquanto a média para par tepe *versus* retroflexo (1xR) é 67,9 ( $p < 0,001$ )<sup>4</sup>. Esse resultado demonstra como os correlatos sociais das variantes podem influenciar o modo como são percebidas acusticamente, já que o par que não contrastaria dois “sotaques” (1xA) foi percebido como menos diferente do que aquele que remeteria a identidades sociais distintas (1xR).

Essa mesma relação é observada nos pares que contrastam a vibrante 2 e as aproximantes (2xA: média 53,2 e 2xR: média 65,6;  $p < 0,001$ ); ou seja, a aproximante alveolar também foi considerada mais parecida com a vibrante 2 relativamente ao retroflexo. Uma vez que as diferenças entre os pares que comparam o tepe e a vibrante 2 à aproximante alveolar (1xA: 51,2 e 2xA: 53,2,  $p = 0,99$ ) e ao retroflexo (1xR: 67,9 e 2xR: 65,6,  $p = 0,99$ ) não foram significativas, é possível inferir que em comparação com as aproximantes, uma ou duas batidas da língua nos alvéolos não é tão saliente para os ouvintes, e que a vibrante 2 tem “valor” de tepe. Já a vibrante 3, quando oposta às aproximantes, não leva a percepções significativamente diferentes (3xA: 65,8; 3xR: 71,5;  $p = 0,31$ ). De fato, as médias das respostas aos pares 1xR, 2xR, 3xA e 3xR são estatisticamente equivalentes, o que reforça o argumento de que a distinção entre o tepe e as aproximantes é a mais saliente e importante.

<sup>4</sup> Resultados do teste de significância Tukey para o modelo: (GRAU DE DIFERENÇA ATRIBUÍDO ~ PAR\*REGIÃO DE RESIDÊNCIA+(1|ESTIMULO)+(1|PARTICIPANTE)).

Em relação ao perfil social dos ouvintes, foram selecionadas como significativas no modelo de efeitos mistos o sexo/gênero ( $p=0,001$ )<sup>5</sup> e sua interação com o Par ( $p<0,001$ ), bem como a interação entre o local de residência na cidade e o par ( $p<0,001$ ).

O Gráfico 2 mostra as respostas de homens e mulheres para cada par. As barras mais escuras indicam as médias das respostas das mulheres e as mais claras, as dos homens. De maneira geral, as mulheres atribuíram maiores graus de diferença a todos os pares, embora as respostas delas e deles nem sempre sejam estatisticamente diferentes.



**Gráfico 2** – Médias dos graus de diferença atribuídos a cada par por homens e mulheres

Fonte: Elaborado pelos autores.

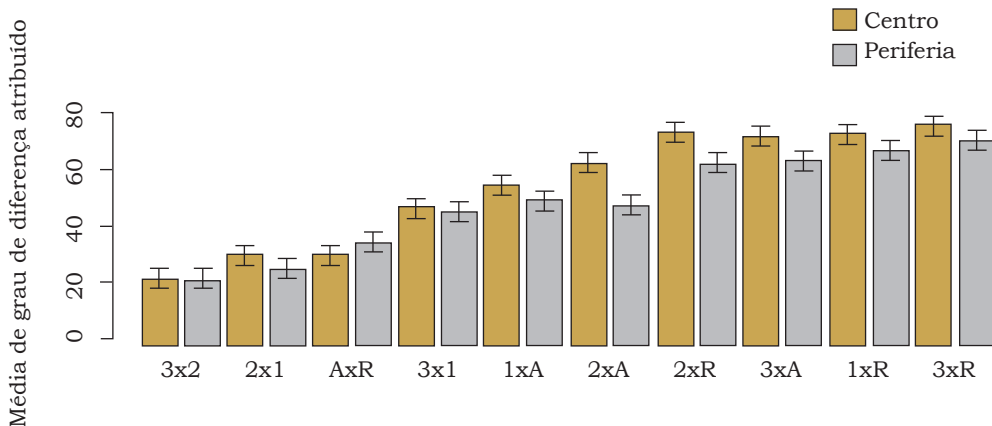
De fato, a principal diferença entre as percepções desses subgrupos reside justamente na comparação das aproximantes com as vibrantes, o que reforça a indicação de que o contraste das vibrantes entre si não é tão saliente na comunidade. Os homens também atribuíram o mesmo grau de diferença às aproximantes quando em contraste com o tepe ( $1xA_M: 46,2$  e  $1xR_M: 55,8$ ;  $p=0,41$ )<sup>6</sup>, enquanto as mulheres consideraram a diferença entre tepe e retroflexo maior do que entre tepe e aproximante alveolar ( $1xR_F: 76,9$ ;  $1xA_F: 55,0$ ;  $p<0,001$ ). O mesmo acontece com os pares formados pela vibrante 2 e as aproximantes ( $2xA$  e  $2xR$ ): enquanto as mulheres consideraram a vibrante 2 mais diferente do retroflexo do que da aproximante alveolar ( $2xA_F: 58,4$ ,  $2xR_F: 73,7$ ;  $p<0,001$ ), os homens não diferenciaram as aproximantes nesse contexto ( $2xA_M: 46,5$  e  $2xR_M: 55,0$ ;  $p=0,65$ ). Esse resultado sobre percepção se equipara ao de inúmeros trabalhos sobre produção, em diferentes comunidades, em que se observa que as mulheres tendem a ser mais “atentas” ao prestígio social das variantes, relativamente aos homens.

5 Resultados do teste ANOVA para o modelo (GRAU DE DIFERENÇA ATRIBUÍDO~LOCAL DE RESIDÊNCIA\*PAR+SEXO/GÊNERO\*PAR+VOGAL PRECEDENTE\*PAR+CONTEXTO FÔNICO SEGUINTE\* PAR+(1|PARTICIPANTE) + (1|ESTÍMULO) / Valor de referência: Par 1xA, Centro, Feminino, /e/, /d/.

6 Resultados do teste de significância Tukey a partir do modelo (GRAU DE DIFERENÇA ATRIBUÍDO ~ PAR \* SEXO/GÊNERO+(1|PARTICIPANTE)+(1|ESTÍMULO)).

O modelo de efeitos mistos ainda mostrou uma interação entre o local de residência e o par; ou seja, ser morador da região central ou periférica de São Paulo também influencia o modo como o ouvinte avalia certos pares de variantes. De maneira geral, as respostas desses dois subgrupos de paulistanos foram semelhantes estatisticamente, mas, mais uma vez, se diferenciam na comparação entre uma vibrante e uma aproximante.

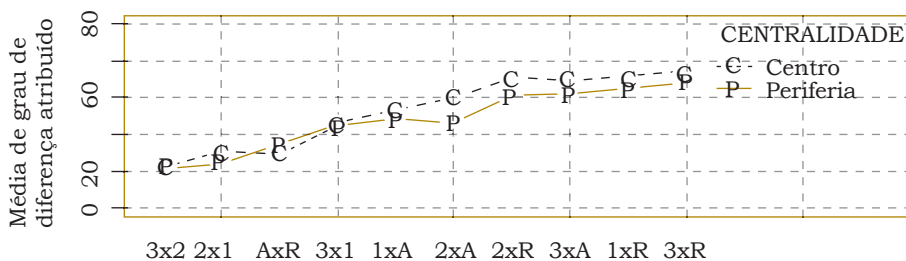
O Gráfico 3 mostra as médias atribuídas para cada par por esses dois subgrupos de paulistanos e revela que a única diferença significativa se observa no par vibrante 2 *versus* aproximante alveolar (2xA), para o qual os moradores de regiões periféricas atribuíram uma diferença menor do que aqueles que moram em regiões centrais.



**Gráfico 3** – Médias dos graus de diferença atribuídos a cada par por moradores de regiões centrais e periféricas

Fonte: Elaborado pelos autores.

Já no Gráfico 4, nota-se que as linhas que indicam as respostas dos moradores do centro (pontilhada) e da periferia (contínua) estão sempre muito próximas, mas mostram padrões diferentes a depender do par. A maior distância entre as linhas aparece efetivamente no par 2xA e as respostas se sobrepõem para os pares 3x2 e 3x1. Embora haja uma aparente “inversão” para o par AxR, ao qual os moradores da periferia atribuíram maior grau de diferença do que os do centro, essa diferença também não é significativa.



**Gráfico 4** – Interação entre a variável linguística par e a social local de residência

Fonte: Elaborado pelos autores.

A principal diferença entre esses subgrupos parece residir exatamente na avaliação que eles fazem a respeito da vibrante 2: enquanto os moradores da periferia a consideraram mais diferente do retroflexo do que da aproximante alveolar ( $2x_{A_p}$ : 46,8 e  $2x_{R_p}$ : 61,0;  $p=0,001$ ), os moradores do centro não diferenciaram esses dois pares ( $2x_{A_c}$ : 60,9,  $2x_{R_c}$ : 70,8;  $p=0,32$ ). Ao mesmo tempo, os moradores de regiões periféricas também atribuíram uma diferença maior para o par  $3xA$  do que para  $2xA$  ( $3x_{A_p}$ : 62,2,  $2x_{A_p}$ : 46,8;  $p=0,0003$ ), enquanto para os moradores de regiões centrais não houve diferença entre as vibrantes 2 e 3 em comparação com a aproximante ( $3x_{A_c}$ : 69,9;  $2x_{A_c}$ : 60,9;  $p=0,51$ ). Os padrões das linhas sugerem que, para os moradores da periferia, a vibrante 2 tem um “valor” mais próximo do tepe, enquanto para os moradores do centro, ela adquiriria um “valor de vibrante” quando em contraste com as aproximantes. Da mesma forma, esses resultados sugerem que a vibrante 2 e a aproximante alveolar têm o mesmo “valor de paulistanidade” para esses moradores da periferia (o que motivaria uma atribuição de diferença baixa), enquanto a vibrante 3 se afastaria de uma paulistanidade prototípica. As diferenças entre as percepções desses dois subgrupos, no entanto, carecem ser mais exploradas, principalmente no que tange a uma possível interação entre o local de residência e outras variáveis sociais que não foram abordadas na presente pesquisa.

### CONSIDERAÇÕES FINAIS

Os resultados aqui relatados acerca das percepções fonéticas de grupos de paulistanos demonstram que, em São Paulo, o “grau de paulistanidade” das variantes de (-r) é de grande importância para a atribuição de diferença fonética entre elas. A aproximante alveolar e o retroflexo, cujas diferenças fonéticas são bastante sutis, foram avaliados de forma diferente pelas mulheres quando em comparação com o tepe:  $1xA$  (que seriam “mais paulistanas”) foram consideradas mais parecidas entre si, do que  $1xR$ . Os homens, por outro lado, não fizeram tal distinção. Do mesmo modo, a diferença entre o tepe e a vibrante 2 não parece saliente para nenhum dos grupos de ouvintes, enquanto a vibrante 3 se destaca das outras duas vibrantes.

De maneira geral, os resultados apresentados neste artigo demonstram que mesmo a percepção fonética, fenômeno que, *a priori*, poderia ser considerado “objetivo”, está longe de ser um processo trivial (bem como de ser desvendado). Este trabalho argumenta, em consonância com estudos recentes (FOULKES; SCOBIE; WATT, 2010; HAY; DRAGER, 2007; DRAGER, 2010), que assim como a produção linguística, a percepção fonética dos indivíduos não varia aleatoriamente, mas é estruturada na comunidade de acordo com fenômenos de natureza social. Tais fenômenos envolvem tanto as características sociais do ouvinte até os significados sociais que os falantes/ouvintes atribuem às variantes (como de “paulistanidade” e “sotaque”) até a noções de prestígio. Mais do que concluir a respeito do que afeta as diferenças de percepção fonética entre os indivíduos, espera-se que este trabalho estimule indagações e curiosidades a respeito de um fenômeno ainda pouco explorado na sociolinguística brasileira.

**PHONETIC PERCEPTIONS ABOUT POST-VOCALIC (-R) IN SÃO PAULO**

**Abstract:** In sociolinguistic interviews, *paulistanos* (people born and raised in the city of São Paulo) comment on two “degrees of retroflexion” of post-vocalic (-r): a “strong” and a “weak” pronunciation, which respectively correspond to the retroflex and the alveolar approximant. The former is generally believed to be employed by people who live in the state’s countryside, while the latter is considered to be present in Paulistano speech. These two variants are in opposition to the tap, the prototypical Paulistano variant. This paper discusses results of a sociophonetic experiment that measures how differently these acoustic subtleties are perceived by different groups of Paulistanos. The main goal is to verify if the speakers’ social categories have an influence on whether the variants are perceived as more different or more similar.

**Keywords:** Phonetic perception. Post-vocalic (-r). São Paulo Portuguese.

**REFERÊNCIAS**

- BAAYEN, R. *Analysing linguistic data: a practical introduction to statistics*. Cambridge: Cambridge University Press, 2008.
- BYBEE, J. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.
- CAMPBELL-KIBLER, K. *Listener perceptions of sociolinguistic variables: the case of (ing)*. 2006. 282 f. Tese (Doutorado em Linguística)–Stanford University, Stanford, 2006.
- CAMPBELL-KIBLER, K. Accent, (ING), and the social logic of listener perceptions. *American Speech*, Durham, v. 82, n. 1, p. 32-61, Spring 2007.
- CAMPBELL-KIBLER, K. The nature of sociolinguistic perception. *Language Variation and Change*, Cambridge, n. 21, p. 135-156, 2009.
- CASTILHO, A.; PRETI, D. (Ed.). *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*. São Paulo: T. A. Queiroz, 1986. v. I – Elocuções Formais.
- DRAGER, K. Sociophonetic variation in speech perception. *Language and Linguistics Compass*, v. 4, n. 7, p. 473-480, July 2010.
- DRAGER, K. Experimental methods in sociolinguistics: matched guise and identification tasks. In: HOLMES, J.; KIRK, H. (Ed.). *Research methods in sociolinguistics: a practical guide*. Hoboken: Wiley-Blackwell, 2013.
- FOULKES, P.; SCOBIE, J.; WATT, D. Sociophonetics. In: HARDCASTLE, W.; LAVER, J.; GIBBON, F. (Ed.). *Handbook of phonetic sciences*. Oxford: Blackwell, 2010. p. 703-754.
- HAY, J.; DRAGER, K. Sociophonetics. *Annual Review Anthropology*, v. 36, p. 89-103, Oct. 2007.
- IRVINE, J. Style as distinctiveness: the culture and ideology of linguistic differentiation. In: ECKERT, P.; RICKFORD, J. (Ed.). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press, 2001. p. 21-43.

- LABOV, W. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- LADEFOGED, P.; BROADBENT, D. E. Information conveyed by vowels. *The journal of the Acoustical Society of America*, v. 29, n. 1, p. 98-104, Jan. 1957.
- MENDES, R. B. Sounding paulistano: variation and correlation in São Paulo. Trabalho apresentado no NWAV39, San Antonio, Texas, 2010.
- MENDES, R. B. Gendered perceptions of noun agreement in Brazilian Portuguese. *Revista Internacional de Linguística Iberoamericana (RILI)*, Madrid, v. 23, n. 1, p. 93-108, 2014.
- MENDES, R. B.; OUSHIRO, L. R. Percepções sociolinguísticas sobre as variantes tepe e retroflexa na cidade de São Paulo. In: HORA, D.; NEGRÃO, E. V. (Ed.). *Estudos da Linguagem*. Casamento entre temas e perspectivas. João Pessoa: Ideia, 2011. p. 229-245.
- MENDES, R. B.; OUSHIRO, L. O paulistano no mapa sociolinguístico brasileiro. *Alfa*, Araraquara, v. 56, n. 3, p. 973-1001, 2012.
- NIEDZIELSKI, N. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, v. 18, n. 1, p. 62-85, Mar. 1999.
- OUSHIRO, L. *Identidade na pluralidade*. Avaliação, produção e percepção linguística na cidade de São Paulo. 2015. 394 f. Tese (Doutorado em Semiótica e Linguística Geral)–Universidade de São Paulo, São Paulo, 2015.
- OUSHIRO, L.; MENDES, R. B. A pronúncia do (-r) em coda silábica no português paulistano. *Revista do GEL*, São Paulo, v. 8, n. 2, p. 66-95, 2013 [2011].
- R CORE TEAM. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. Disponível em: <<http://www.R-project.org/>>. Acesso em: 16 jul. 2016
- SORIANO, L. G. M. Percepções sociolinguísticas: o /-r/ em São Paulo, 2014. Trabalho apresentado no 3°CIDS, Londrina.

Recebido em maio de 2016.  
Aprovado em julho de 2016.