UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM SEMIÓTICA E LINGUÍSTICA GERAL

**O falar paulist[ẽ:]no e os significados sociais de (AN):
correlações entre origem do ouvinte e percepção**

Maria Eugênia Martins Barcellos

São Paulo
2020

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM SEMIÓTICA E LINGUÍSTICA GERAL

**O falar paulist[ẽ:]no e os significados sociais de (AN):
correlações entre origem do ouvinte e percepção**

Maria Eugênia Martins Barcellos

Dissertação apresentada ao Programa de Pós-Graduação em Semiótica e Linguística Geral do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, como requisito para obtenção do título de Mestre em Linguística.

Orientador: Prof. Dr. Ronald Beline Mendes

São Paulo

2020

Para minha mãe
que sempre me disse “vai lá, filha”,
mesmo que esse “lá” fosse para longe dela.

Para meu pai,
sempre orgulhoso das minhas conquistas.
Acredito que dessa também se orgulharia.

E para Lúcia,
minha segunda mãe.

Sobre ser “de fora” e ter concluído meu ciclo em São Paulo:

Com o passar do tempo, hibridizamos nosso novo eu com o antigo. Nos tornamos a mistura de várias terras, desfazemos e refazemos lugares, recebemos e deixamos marcas, encontramos super-objetos que nunca antes imaginamos, descobrimos pessoas e lugares, nos deleitamos com a criatividade culinária [cachorro quente com purê!] e, metamorfoseados, atravessamos rotas inacreditáveis.

Nilen Cohen

Agradecimentos

Os agradecimentos são, geralmente, um tanto contraditórios: umas das primeiras palavras que aparecem no trabalho, mas as últimas a serem escritas. Isso porque, pelo menos para mim, a sensação de dever cumprido e ciclo fechado tomou conta assim que finalizei esta dissertação, e as palavras que me restaram foram exatamente “muito obrigada”. E, aqui, outra contradição me ocorre: escrevo esses agradecimentos com uma enorme vontade de abraçar cada uma das pessoas que cito, mas num contexto que nos impõe distanciamento e isolamento social. As palavras que escrevo são, para muitas dessas pessoas, de gratidão e saudade.

Agradeço à minha família. À minha mãe, a quem gosto de chamar carinhosamente de Dona Nilda, por sempre ter me incentivado a alcançar meus sonhos, mesmo que para isso nós tivéssemos que nos distanciar fisicamente. Ao meu pai, Seu Barcellos, quem sempre comprava livros para os filhos e sentava junto à mesa para fazer os deveres de casa. Com certeza, isso se reflete no fato de eu estar me tornando mestra por uma excelente universidade pública, e sei que ele estaria muito orgulhoso disso. Ao meu irmão, Pedro Henrique, por ser um exemplo para a irmã caçula aqui. Aos meus outros irmãos, Mônica, Fernando e Ana Paula, por sempre torcerem por mim. À Lúcia, minha segunda mãe, por ser voluntariamente minha segunda mãe.

Agradeço ao meu orientador, Ronald – para mim, respeitosamente, o *professor*. Agradeço por ter me acolhido tão bem desde nosso primeiro encontro, quando me levou para conhecer os prédios da FFLCH, indo e voltando da Letras até o vão da História. Para mim, capixaba com receio da recepção em São Paulo e por uma universidade que não me conhecia, isso significou o mundo. Agradeço também pela orientação, pelas sugestões que acrescentam muito na minha redação e por me ouvir nos meus não tão bons momentos. Obrigada, professor!

Agradeço ao melhor grupo de estudos da vida in-tei-ra, o GESOL: Amanda Santana (ou Musa Osasquence), Dany Gonçalves, Germano Jorge, Giovanna Indelicato (a Giovããna), Gustavo Gomes, Isabel Pie (ou Bel), Julia França, Larissa Soriano, Livia Oushiro, Luciana Massai, Luigi Parrini, Monique Freitas (ou mana Monique) e Wendel Santos (a senhora do grupo). Agradeço a vocês todos pelas discussões de trabalho, pelas ajudas com esta dissertação, pelas trocas de e-mails e mensagens de WhatsApp para diversos fins, pelas fofocas, pelos cafés com “bólos” e, claro, pelo apoio emocional. Carrego cada um de vocês com muito calorzinho no meu coração.

Agradeço a Livia Oushiro, por quem nutro uma enorme admiração de cunho profissional e pessoal. Agradeço por cada e-mail que respondeu, fosse sobre a minha pesquisa, fosse sobre seu curso de verão do R. Agradeço pelas caronas de ida e volta da Unicamp. Agradeço pelo que você faz pela pesquisa desse país – e isso não só sobre questões estritamente acadêmicas; falo do modo como você acolhe quem pede a sua ajuda. Obrigada, Livia!

Por falar em acolhimento, agradeço à professora Christina Abreu Gomes, a quem eu sempre me refiro carinhosa e respeitosamente como *professora*. Agradeço por ter me acolhido tão bem quando eu era uma graduanda do terceiro período de Letras da UFRJ e depois, como aluna especial. Com a senhora, professora, eu dei meus primeiros passos na pesquisa acadêmica e tomei gosto pela sociolinguística. Agradeço por confiar no meu trabalho e por ter contado comigo na redação e publicação de alguns textos acadêmicos. Obrigada, professora!

Aproveito o “gancho” e agradeço ao Marcelo Alexandre Silva Lopes de Melo por ter me inspirado tanto com sua apresentação sobre o (s) em coda na minha aula de Fonética do segundo período de Letras. Foi por conta dela que tive vontade de fazer iniciação científica na “sócio”.

Agradeço ao professor Paulo Chagas, que gentilmente topou participar da defesa do meu exame de qualificação e agora da minha dissertação pronta.

Por fim, agradeço àqueles que sempre lutaram pela educação pública de qualidade. Este trabalho foi fruto dela e a ela muito deve.

Resumo

Esta dissertação de mestrado investiga os significados sociais associados a /a/ nasalizado (AN) e se a origem dos ouvintes tem efeito na percepção sociolinguística dessa variável. Interessa verificar se ouvintes paulistanos e não paulistanos (do interior e de fora do estado de São Paulo) associam, do mesmo modo, (AN) a noções como inteligência, formalidade, escolaridade, sucesso profissional, classe social, localização do bairro, amigabilidade, masculinidade e feminilidade e outras características sociais. A variante alongada é aqui referida como “tipicamente paulistana”, em contraposição ao seu par, a variante não alongada. Para atingir seus objetivos, esta pesquisa embasou-se nos pressupostos teórico-metodológicos da Sociolinguística Variacionista (Labov, 2006[1966]; 2001; 2008[1972]; Eckert, 2008; 2012). Primeiramente, conduziram-se entrevistas metalinguísticas, no intuito de eliciar reações espontâneas a estímulos com a variante alongada. As análises evidenciaram divergências nas respostas dos 27 ouvintes paulistanos e 23 não paulistanos. Em seguida, aplicou-se um experimento de percepção, com disfarces gravados por quatro falantes paulistanos (dois homens e duas mulheres) que foram organizados de acordo com a técnica *matched-guise* (Lambert *et al.*, 1960; Campbell-Kibler, 2006; 2009). Os disfarces constituem-se de sentenças com duas ocorrências de uma das variantes de (AN) e cujo tema é “infância”. As análises de regressão logística (R Core Team, 2019) mostram que a (AN) não tem efeito na percepção de nenhuma variável quantitativa nas respostas dos 70 ouvintes paulistanos, mas teve efeito na percepção da variável nominal “branca” (etnia). Já nas respostas dos 53 ouvintes não paulistanos, (AN) tem efeito na percepção de 8 das 10 escalas quantitativas: os falantes soam mais inteligentes, amigáveis, paulistanos, bem sucedidos profissionalmente, de classe média alta, residentes de bairros mais centrais da cidade de São Paulo, mais masculinos e menos femininos quando ouvidos em seus disfarces com a variante alongada. Ainda

para esse grupo de ouvintes e para essa variante, os falantes são significativamente associados a um adulto por volta dos 20 anos e conservador. Ao estudar a percepção de uma variável linguística, cuja produção ainda não foi analisada, esta pesquisa mostra que há outras maneiras de se trabalhar o lado social da variação sociolinguística e que a origem do ouvinte pode ser determinante para os estudos de percepção sobre o falar específico de determinada região.

Palavras-chave: Português paulistano. /a/ nasalizado. Origem do ouvinte. Percepção sociolinguística.

Abstract

This master's thesis investigates the social meanings associated to nasal /a/ (AN) and whether the listeners' origin has an effect on the sociolinguistic perception of this variable. It searches for whether listeners from the city of São Paulo (Paulistanos) and non-Paulistanos (listeners from the state countryside and from other states) similarly associate (AN) with notions of intelligence, formality, education, professional success, social class, neighborhood location, friendliness, masculinity and femininity, as well as with other social characteristics. The lengthened variant is referred to as "typically Paulistano", in contrast to its counterpart, the non-lengthened variant. To achieve its goals, this research is based on Variationist Sociolinguistics theoretical and methodological assumptions (Labov, 2006 [1966]; 2001; 2008 [1972]; Eckert, 2008; 2012). At first, open-ended interviews were conducted in order to elicit spontaneous reactions to stimuli with the lengthened variant. Such analysis show divergences in the responses of 27 Paulistano listeners and 23 non-Paulistanos. After that, a perception experiment was carried out, with guises recorded by four Paulistano speakers (two men and two women), each of them talking about their childhood. These guises were organized according to the matched-guise technique (Lambert *et al*, 1960; Campbell-Kibler, 2006; 2009), and consisted of sentences with two occurrences of one of the variants of (AN). Logistic regression analysis (R Core Team, 2019) show that (AN) has no effect on the perception of any quantitative variable in the responses by the 70 Paulistano listeners, but it does have an effect on the perception of the nominal variable 'white' (regarding ethnicity). As for the responses of the 53 non-Paulistanos, (AN) has an effect on the perception of 8 of the 10 quantitative scales: the speakers sounded more intelligent, friendly, Paulistano, professionally successful, upper middle class, like residents of more central neighborhoods of the city of São Paulo, more masculine and less feminine when

heard in their lengthened (AN)-guises. For this group of listeners and for that variant, speakers were more frequently perceived as a conservative adult around 20 years old. When studying the perception of a linguistic variable that has not yet been analyzed in production studies, this research shows that there are other ways of working on the social side of sociolinguistic variation, and that the listener's origin can be determinant to studies of perception regarding the speech of specific region.

Keywords: São Paulo Portuguese. Nasal /a/. Listener's origin. Sociolinguistic perception.

Sumário

Resumo	8
Abstract.....	10
Sumário.....	12
Introdução - “O paulista tem vários tipos de sotaque”	14
Capítulo 1 - “O nasalês de São Paulo”	21
1.1 A nasalidade vocálica no português	21
1.2 Análise espectral e cepstral de (AN)	24
1.2.1 Falante Gisele.....	25
1.2.2 Falante Ricardo.....	28
1.2.3 Falante Bruna.....	31
1.2.4 Falante Rodolfo	33
Capítulo 2 - Estudos de percepção da fala	37
2.1 Da psicologia social à sociolinguística	37
2.2 As três ondas da sociolinguística.....	55
Capítulo 3 - Hipóteses e montagem do experimento.....	64
3.1 Preparação dos estímulos	64
3.2 Entrevistas metalinguísticas	66
3.3 O experimento de percepção.....	75
3.4 Hipóteses e expectativas	77
3.5 Coleta e perfil dos ouvintes	80
Capítulo 4 - Análises: Percepções sociolinguísticas acerca de (AN) por ouvintes paulistanos e não paulistanos.....	84
4.1 CP1 – Prestígio e <i>Status</i> social	91
4.2 CP2 – Gênero e Profissionalismo	100
4.3 CP3 – Amigabilidade e Gênero	105
4.4 Variáveis qualitativas	110

4.5 Caixas de seleção	113
Considerações finais	115
Referências	122
Apêndices	139

Introdução

“O paulista tem vários tipos de sotaque”¹

A presente pesquisa sobre percepção sociolinguística originou-se da observação de dois fatos de natureza empírica: i) o modo “específico” como falavam paulistanos VJs² da antiga emissora *MTV Brasil*³; e ii) o modo como

¹ Marcelo Adnet em entrevista ao “Programa do Jô” em 2009 (<https://www.youtube.com/watch?v=Vp0pMLsTSk0> [acessado em 22/12/2019]).

²Do inglês *videojockey*; apresentador da TV de videoclipes e músicas. O nome surgiu por contraste a DJ, *disco jockey*.

fluminenses frequentemente caçoam dos sotaques paulistas/paulistanos. Em princípio, emergiram as seguintes perguntas: como falam os paulistas e paulistanos? e quais são os elementos linguísticos que caracterizam esses sotaques e levam falantes de outros lugares a “registrá-los” como paulistas ou paulistanos? Marcelo Adnet, um humorista *carioca da gema*, oferece algumas pistas para responder a essas perguntas em sua entrevista no Programa do Jô⁴ de 2009. De suas imitações caricatas, podem-se destacar:

- O paulista italiano: *tem o paulista italiano que fala assim, que é da Mooca, que tem um sotaque mais puxado de Catanduva* (imitando algo como um “sotaque de italiano” e fazendo gesto com a mão, como mostra a **Figura 1**);
- O paulista do interior: *que fala “po[ɿ]ta”, “co[ɿ]te”, “mo[ɿ]te”, “so[ɿ]te” e “fo[ɿ]te”; que é o paulista que deve falar inglês, exatamente é esse, né? “Piracicaba isnea[ɿ] orfa[ɿ]?”*;
- O “nasalês” de São Paulo: *o “nasalês” é aquela menina que fala “Eu sou a Juliana e tenho 25 anos”* (com as vogais destacadas pronunciadas com maior duração, de modo mais “anasalado” (a partir de então, representadas por [ẽ:]) e com caretas, como mostra a **Figura 2**);
- Sotaque da periferia de São Paulo: *que na periferia de São Paulo, a voz das pessoas vai ficando grave (...) “E aí, mano, tudo ce[ɿ]to, mano? Incrusive, mano, quando eu falo assim com essa voz, dá mó vontade de rimar, ce[ɿ]to?”* (com voz grave, repetição de “mano”, rotacismo em “incrusive”, “mó” como advérbio de intensidade para “muito” e (-r) retroflexo em “certo”).

Figura 1 Marcelo Adnet imitando “o paulista italiano” no Programa do Jô

³ Para a pesquisadora, os VJs favam diferentemente dela e de seus conterrâneos capixabas.

⁴ <https://www.youtube.com/watch?v=Vp0pMLsTSk0> [acessado em 22/12/2019].

Figura 2 Marcelo Adnet imitando o “o ‘nasalês’ de São Paulo” no Programa do Jô

É interessante observar como a constatação de Adnet sobre o (-r) em coda é o evidenciado em Oushiro (2015: 111): para o humorista, o paulistano da periferia produz o retroflexo tal qual o paulista do interior e, de fato, a autora mostra que moradores de bairros mais periféricos de São Paulo favorecem fortemente o emprego dessa variante (que é associada à fala de pessoas do interior por certos paulistanos), enquanto moradores de bairros mais centrais tendem a evitá-la, preferindo o tepe. Ademais, o estudo também evidencia que filhos de pais do interior tendem a usar mais o retroflexo relativamente aos

filhos de paulistanos, estrangeiros ou de pais de origem mista – e menos do que filhos de nordestinos e nortistas.

Adnet parece constatar ainda que a tal “menina do nasalês” não é da periferia, pois o jeito de falar “Juliana” ou “anos” é diferente do jeito de falar “mano” que, em princípio, poderiam ser pronunciadas da mesma maneira, visto que apresentam /a/ nasalizado (doravante (AN)) em posição de sílaba tônica. Sendo assim, Adnet parece sugerir que esses “tipos de paulistanos” pertencem a lugares diferentes de São Paulo.

Essa breve análise das representações de paulistas/paulistanos por Adnet levaram a outras questões que inspiraram o presente estudo: que imagens as pessoas têm de uma menina que emprega o “nasalês”?; de que local da capital paulista imaginam que ela seja?; as pessoas concordam com Adnet ao não associar tal menina à periferia?; e a qual classe social ela pertenceria? De modo geral: quais significados sociais associam-se a essa variante que parece ser tipicamente paulistana?

Mais uma vez, na própria entrevista, Marcelo Adnet parece lançar luz sobre tais questões. Ao ser questionado por Jô Soares sobre “onde se encaixa aquele ‘cantar’ de falar assim ‘foi em nov[ẽ]bro, ent[ẽ]de?’” (com /e/ nasal ditongado), Adnet responde, de modo caricato:

Foi em nov[ẽ]bro, h[ẽ:]n... H[ẽ:]n... H[ẽ:]n... [...] Eu acho que esse é o “nasalês”: “Eu sou a Juli[ẽ:]na, tenho 25 [ẽ:]nos. Estou aprend[ẽ]do muito. Eu nasci no mês de set[ẽ]bro”.

A percepção de Adnet, portanto, é de que a mesma menina que fala *Juli[ẽ:]na* fala *set[ẽ]bro*, o que leva ao questionamento de se essas duas variantes são percebidas sociolinguisticamente de modo similar. De fato, /e/ nasal ditongado é associado à paulistanidade, como evidenciou Mendes (2016). Além desse resultado, após expor 44 ouvintes a oito estímulos auditivos contendo uma única ocorrência de /e/ nasal (doravante, (EN)), produzidos por quatro falantes (dois homens e duas mulheres), o estudo mostrou que, de modo geral,

a pronúncia ditongada também é associada à feminilidade e a áreas mais centrais da cidade de São Paulo. Os resultados mostram, no entanto, que tal padrão de percepção muda a depender do falante. Contudo, por questões de interesse de estudo, Mendes não levou em consideração, na análise dos dados, a origem dos ouvintes – se eles eram ou não paulistanos.

Esta variável, contudo, parece ser relevante para estudos de percepção, pois, como mostram as respostas coletadas de entrevistas abertas⁵ por Oushiro (2015) em seu estudo de avaliação⁶, o (EN) ditongado parece ser um estereótipo (Labov, 2008 [1972]) do linguajar paulistano para quem é de “fora” de São Paulo – e não para os nativos da cidade –, visto que apenas dois dos 118 paulistanos entrevistados no estudo desenvolveram algum discurso metalinguístico sobre a variante quando perguntados o que achavam de quem fala “meu, (vo)cê (es)tá entend[êj]do o que eu (es)tou diz[êj]do?”. Ou seja, para os paulistanos, a variante está abaixo do nível de consciência (Labov, 2001). Por outro lado, pessoas de “fora” de São Paulo produzem metacomentários sobre a variante ditongada, como em Bieler da Silva (2015: 108):

S1: *então eu até até hoje eu ainda falo aqueles ‘[êj]nta’ lá aqueles trinta [vogal longa e nasalizada] ainda pego*

D1: *uhum*

S1: *ainda peguei isso das meninas de São Paulo porque não dá para você não pegar*

D1: *é*

S1: *o sotaque de São Paulo ele é muito forte e eu morava com três de São Paulo*

(F1B_cristina)

Diferentemente de (EN), a variável (AN) ainda não foi estudada. Portanto, um dos objetivos deste trabalho é verificar quais significados sociais são associados as suas variantes, aqui representadas por [ẽ:] e [ẽ] (a variante

⁵ Proposta de tradução de Oushiro (2015) para “*open-ended interviews*” (Campbell-Kibler, 2006).

⁶ Em seu estudo, Oushiro (2015: 32) diferencia conceitualmente *avaliação* de *percepção* linguística: o primeiro diz respeito aos discursos metalinguísticos dos falantes acerca das variantes – que, por serem subjetivos e espontâneos, são um objeto de estudo em si –, ao passo que o segundo refere-se às inferências, conscientes ou não, que os indivíduos fazem ao ouvirem um falante e que, portanto, podem não ter sido objeto dos metacomentários da avaliação.

“tipicamente paulistana” e a “comum”, respectivamente). Ademais, como se observou na discussão acima, a origem dos ouvintes parece ser uma variável relevante para os estudos de percepção. Portanto, também se averigua se ela tem efeito na percepção sociolinguística de (AN). Para tanto, analisaram-se separadamente as respostas de ouvintes paulistanos e não paulistanos a formulários de um experimento de percepção. Aqui, entendem-se como não paulistanos tanto pessoas do interior do estado como de “fora” de São Paulo, mas que residem na capital paulista. Desse modo, tal como fizeram Mendes (2016), Oushiro (2015) e Santos (2020), este trabalho visa a contribuir para os estudos sociolinguísticos acerca da percepção do português paulistano. Todos esses (e outros) trabalhos são discutidos no **Capítulo 2** desta dissertação.

Para alcançar os objetivos propostos, esta pesquisa dividiu-se em dois grandes momentos, como se abordou no **Capítulo 3**. No primeiro, empregaram-se entrevistas metalinguísticas⁷ (Campbell-Kibler, 2006; Oushiro, 2015) com o objetivo de buscar as reações naturais e espontâneas de 50 ouvintes (27 paulistanos e 23 não paulistanos) acerca de [ẽ:], a partir da escuta de áudios de dois paulistanos (um homem e uma mulher) da amostra São Paulo 2010⁸ e a partir da visualização do vídeo do Marcelo Adnet – apenas com a parte do “nasalês”. Aqui, fez-se uma análise qualitativa dos discursos metalinguísticos apresentados. As respostas coletadas nesse primeiro momento já mostraram divergências entre os grupos de ouvintes, visto que, muitas vezes, os paulistanos disseram não identificar nada “de mais” nos áudios, ao passo que não paulistanos chegaram a reconhecer a variante em questão e associá-la a certos tipos de pessoas (“paulistana”, “blogueira”, entre outras).

⁷ Proposta de tradução do termo “*open-ended interviews*” (Campbell-Kibler, 2006).

⁸ O Projeto SP2010 reúne amostras da fala paulistana coletadas, entre os anos 2010 e 2012, pelo Grupo de Estudos e Pesquisa em Sociolinguística (GESOL) da Universidade de São Paulo (USP), na forma de gravações e transcrições, com o intuito de contribuir para a disseminação de trabalhos sociolinguísticos sobre essa variedade de fala. Para acessar ao site do projeto, basta ir em <http://projetosp2010.fflch.usp.br/>

Com base nessas respostas, elaboraram-se formulários para o segundo momento da pesquisa: um experimento de percepção. Isso significa que algumas noções apresentadas nos discursos metalinguísticos apareceram como variáveis quantitativas ou nominais no formulário de percepção. Para o experimento, empregou-se a técnica *matched-guise*⁹ (Lambert *et al.*, 1960; Campbell-Kibler, 2009; Mendes, 2018) com disfarces produzidos por quatro falantes paulistanos (duas mulheres e dois homens), contendo, cada disfarce, duas ocorrências de uma das variantes de (AN). A técnica tem por objetivo eliciar atitudes subjetivas e encobertas acerca das variantes. As respostas dos 123 ouvintes (70 paulistanos e 53 não paulistanos) foram coletadas na plataforma online Google Forms e os dados foram analisados no *software* de computação estatística R (R Core Team, 2019).

Como se pode observar, o **Capítulo 3** apresenta as metodologias empregadas para a pesquisa, discorrendo sobre as entrevistas metalinguísticas e a montagem e execução do experimento de percepção com a técnica *matched-guise*. O capítulo aborda, ainda, as hipóteses da pesquisa. Em seguida, o **Capítulo 4** apresenta os resultados do experimento de percepção em gráficos e tabelas, e suas interpretações. Por fim, nas **Considerações Finais**, são apresentadas as contribuições desta pesquisa para os estudos (sócio)linguísticos.

Ao fim da discussão sobre os “sotaques paulistas” feita nesta introdução e antes de começar a revisão epistemológica sobre trabalhos de percepção da fala (no **Capítulo 2**), é necessário entender o que é, afinal, esse “nasalês” de São Paulo. No **Capítulo 1**, a seguir, discutem-se as diferenças na configuração acústica de [ẽ:] e [ẽ], a fim de melhor entender o objeto de estudo desta dissertação.

⁹ Traduzido como “técnica de estímulos pareados” por Oushiro (2015).

Capítulo 1

“O nasalês de São Paulo”¹⁰

O objetivo deste capítulo é entender minimamente como a variante “tipicamente paulistana” se configura acusticamente e se diferencia do seu par. Isto posto, não se propõe uma análise pormenorizada e extensa de ambas as pronúncias – objetivo das teorias de fonética acústica e fonologia

¹⁰Marcelo Adnet em entrevista ao “Programa do Jô” em 2009.
(<https://www.youtube.com/watch?v=Vp0pMLsTSk0> [acessado em 22/12/2019])

experimental¹¹. De todo modo, uma descrição das propriedades articulatórias e acústicas das variantes de (AN) é importante na medida em que as diferenças fonéticas entre elas implicam em diferentes percepções sociolinguísticas, como mostra este estudo mais a seguir. No entanto, antes dessa descrição, é necessário apresentar a discussão da literatura acerca da nasalidade vocálica no português.

1.1 A NASALIDADE VOCÁLICA NO PORTUGUÊS

A literatura admite que a nasalidade vocálica no português é um fenômeno complexo e difícil de ser explicado por divergências conceituais e por razões articulatório-acústicas (Raposo de Medeiros, 2007: 166; Moraes, 2013: 103; Barbosa e Madureira, 2015: 450-1). Ela é, portanto, um dos “[pontos] mais controversos da fonologia do português” (Moraes e Wetzels, 1992: 153).

A começar, há diversas abordagens de diferentes correntes teóricas que explicam o fenômeno (Cf. Moraes & Wetzels, 1992). Algumas abordagens estruturalistas e gerativistas consideram a vogal como nasal propriamente dita, de forma que ela se encontra numa estrutura *monofonêmica* (algo como /lã/ para “lã”). Por outro lado, há aquelas que defendem uma estrutura *bifonêmica*, sendo a sílaba composta por uma vogal oral que sofre assimilação regressiva do arquifonema nasal seguinte (/laN/), numa configuração, portanto, tautossilábica¹².

Estruturalistas e gerativistas convergem quanto à metodologia: enquanto aqueles buscam respostas para o fenômeno numa análise diacrônica da morfologia do português (Câmara Jr, 1953, 1970), estes pautam-se em processos de derivação fonológica a partir de estruturas subjacentes (Wetzels, 1991; Bisol, 2013). De todo modo, tanto abordagens monofonêmicas quanto bifonêmicas são exemplos de *nasalidade contrastiva* (*oufonêmica*), pois a nasalidade é traço

¹¹ Ver Moraes e Wetzels (1992), Raposo de Medeiros (2007), Moraes (2013), Bisol (2013), Barbosa e Madureira (2015), *inter alia*.

¹² Que estão na mesma sílaba.

distintivo capaz de opor significado (/lã/ *versus* /lá/). Aqui, diz-se que a vogal é nasalizada fonologicamente, o que, segundo a Linguística portuguesa, confere a ela o título de *vogal nasal* (Cf. Barbosa e Madureira, 2015: 451).

Pelo viés estruturalista (Câmara Jr, 1970), há também a *nasalidade alofônica*(ou *fonética*), incapaz de opor pares mínimos e de natureza heterossílábica, determinada pela presença de consoante nasal em posição de ataque na sílaba seguinte à da vogal nasalizada. Ou seja, tem-se uma consoante nasal intervocálica, de onde o traço de nasalidade se espalha regressivamente para a vogal na sílaba anterior, como nos exemplos de Adnet “Juliana” e “anos”. Nestes casos, a implementação da nasalidade se dá foneticamente e, portanto, tem-se uma *vogal nasalizada* (Cf. Barbosa e Madureira, 2015: 451). Essa assimilação regressiva é “o resultado da aplicação de uma regra variável, sensível a fatores como localização do acento vocábulo, natureza da consoante nasal subsequente, natureza da vogal nasalizada e o dialeto” (Moraes, 2013: 98).

Não cabe aqui discutir o mérito de uma ou outra abordagem¹³. Para esta dissertação, é relevante saber que elas fazem uma distinção entre os tipos de nasalidade. Seguindo os exemplos dados na própria imitação caricata de Adnet, os disfarces do experimento de percepção desta pesquisa apresentaram somente o que a tradição brasileira reconhece como vogais nasalizadas de (AN), em palavras como “Fabiana”, “Juliana”, “Mariana”, “Adriana”, “pano”, “cano”, “cigano” e “fulano” (Cf. **Capítulo 3** desta dissertação). Ainda seguindo os exemplos de Adnet, tomou-se cuidado para que todas as ocorrências de /a/ nasalizado estivessem em sílaba tônica e apresentassem uma consoante nasal dental na sílaba subsequente. Seguiram-se tais exemplos pois ainda não há estudos sobre a produção de /a/ nasalizado tipicamente paulistano e, portanto, não se sabe se a variação ocorre em outros contextos de tonicidade silábica e de consoante seguinte.

¹³ Para outras abordagens, ver Moraes e Wetzels (1992) e Moraes (2013).

Como se mencionou, a nasalidade vocálica no português é complexa também por razões articulatório-acústicas. Na realidade, o fenômeno da nasalidade em geral, seja de consoantes ou vogais, tem sua análise espectral¹⁴ prejudicada devido ao próprio modo de articulação das consoantes nasais.

Na produção de um som nasal, há o acomplamento das cavidades (ou tratos) oral e nasal, de forma que há ressonâncias nessas duas regiões. A grande questão é que, pelo fato de haver alguma obstrução na passagem de ar pela cavidade oral no momento de produção da nasal, os efeitos da ressonância não conseguem ser exteriorizados pelos lábios e, portanto, o som não é plenamente captado pelo microfone (Barbosa e Madureira, 2015: 106). Esse som “lesado” recebe tecnicamente o nome de *antiformante*. Se os *formantes* são “zonas de frequência em que há maior concentração de energia acústica” (Callou, Moraes e Leite, 2013: 79), os antiformantes são exatamente esses picos de energia vibratória aprisionados que, por consequência, têm a amplitude de suas faixas de frequência minimizada ao serem capturadas pelo microfone (Barbosa e Madureira, 2015: 106). Dessa forma, como se verá nas figuras mais adiante, se os formantes são representados por picos nas análises acústicas, os antiformantes são representados por vales.

Além disso, diferentemente de outros sons, o espectro final de uma nasal é a combinação entre os espectros da cavidade oral e da nasal. Dessa forma:

Se houver um formante no mesmo lugar de um antiformante, a depender da amplitude de cada um, pode-se ter uma anulação dos dois efeitos e a amplitude na região em que isso ocorreu fica semelhante à amplitude da mesma região do espectro da glote. [...] [O] grande problema com a composição de formantes e antiformantes em lugares próximos ou idênticos é a impossibilidade de observação dos dois efeitos em separado, uma vez que o microfone à boca capta o efeito final (Barbosa e Madureira, 2015: 119).

Com base em tal problema, Barbosa e Madureira (2015: 460) afirmam que somente a *análise de Fourier* dos sons nasais é insuficiente para observar as

¹⁴ “Análise dos componentes em frequência de um som” (Barbosa e Madureira, 2015: 140).

regiões formânticas e antiformânticas, apesar de ela ser a mais comumente usada em estudos de fonética acústica para descobrir as amplitudes e fases de ondas periódicas¹⁵. Para uma observação satisfatória de tais regiões, é necessário comparar o espectro de Fourier ao espectro da *análisecepstral*, cujos cálculos permitem “obter uma representação dos sinais da fala com a variação de amplitude na ordenada e de uma variável na abscissa chamada de quefrequência” (Barbosa e Madureira, 2015: 163)¹⁶.

Por vezes neste estudo, adotaram-se os termos técnicos da fonética acústica para identificar o espectro de cada análise: o espectro da análise de Fourier é chamado *espectro*, ao passo que o da cepstral é chamado *cepstro*.

Para este estudo, uma ocorrência de [ẽ:] de cada falante foi analisada por meio de um espectro e um cepstro, e o resultado foi comparado ao espectro e ao cepstro da outra variante para a mesma palavra. Antes, no entanto, traçaram-se os espectrogramas dessas mesmas ocorrências para melhor constatar o tempo de duração de produção das variantes de (AN).

1.2 ANÁLISE ESPECTRAL E CEPSTRAL DE (AN)

Analísaram-se as variantes de (AN) da palavra “Mariana” para a falante Gisele, da palavra “Juliana” para Ricardo, “Fabiana” para Bruna e “Adriana” para Rodolfo. Os falantes apresentam-se sob pseudônimos para que sua identidade permaneça desconhecida por questões éticas. Serão apresentadas as análises espectrais e cepstrais para cada falante nessa ordem – a mesma do experimento, como se vê no **Capítulo 3** –, a fim de diferenciar a configuração acústica das variantes.

1.2.1 FALANTE GISELE

¹⁵ Para mais sobre a análise de Fourier, ver Barbosa e Madureira (2015).

¹⁶ Para mais sobre a análise cepstral, ver Barbosa e Madureira (2015).

Gisele é uma paulistana que gravou a seguinte sentença para o experimento de percepção: “Eu era muito amiga do Paulo, da Mariana... A gente fingia ser príncipe, princesa, herói, cigano, pirata... A gente criava muita história!”. Na **Figura 3** e **4** abaixo, têm-se os espectrogramas de banda larga (janela de 5 ms) da palavra “Mariana” com [ẽ] e [ẽ:], respectivamente. É relevante dizer que a representação das análises de Fourier e cepstrais deste estudo seguiram o modelo proposto por Barbosa e Madureira (2015) para as vogais nasalizadas. Além disso, o símbolo fonético de (AN) “não tipicamente paulistano” ([ẽ]) também está de acordo com os autores.

Figura 3 Espectrograma de banda larga (janela de 5 ms) da palavra “Mariana” com a variante [ẽ] produzida pela falante Gisele

Figura 4 Espectrograma de banda larga (janela de 5 ms) da palavra “Mariana” com a variante [ẽ:] produzida pela falante Gisele

Na **Figura 3**, destacam-se duas fases da produção da vogal (AN): [ẽ] e /N/. Na **Figura 4**, as fases são [ẽ:] e /N/. As duas fases de produção de cada variante estão presentes em todas as ocorrências de todos os falantes apresentadas nessa subseção. Para as vogais em si, observa-se uma diferença no tempo de duração: enquanto [ẽ] dura 137 milissegundos, [ẽ:] prolonga-se por 211 – uma diferença de 74 milissegundos. No que diz respeito à segunda fase de produção de (AN), isto é, /N/, a diferença entre as variantes é de somente 2 milissegundos. Por estes motivos e por ainda não haver na literatura estudos sobre o (AN) “tipicamente paulistano”, representou-se foneticamente esta variante pela junção de [ẽ] com o diacrítico que representa um som longo (:), formando, assim, o símbolo [ẽ:]. Com base nessas análises, também se refere à [ẽ] como “variante não alongada” e à [ẽ:] como “variante alongada” neste estudo.

Uma análise mais fidedigna dos formantes e antiformantes é feita, como se discutiu, por meio da análise de Fourier e cepstral das pronúncias, como nas **Figuras 5 e 6** abaixo para a variante não alongada e a alongada, respectivamente.

Figura 5 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ] da falante Gisele na palavra “Mariana”

Figura 6 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ:] da falante Gisele na palavra “Mariana”

Como se discutiu anteriormente, na leitura dessas figuras, deve-se ter em mente que as regiões de concentração de picos representam os formantes, ao passo que os vales indicam os antiformantes. O ideal é que o espectro e o

cepstro desenhem juntos o mesmo traçado acústico dos sons nasais, apesar de usarem cálculos diferentes. Ademais, entende-se, nas figuras, que o primeiro pico de energia refere-se à frequência fundamental (f_0) ou *pitch* (Barbosa e Madureira, 2015), e não ao primeiro formante em si.

Na **Figura 5**, para a variante não alongada, têm-se um primeiro formante por volta de 1600 Hz e outro em 2750 Hz, com a energia da vogal nasalizada caindo consideravelmente após este. Também são dois os antiformantes: um em torno de 1250 Hz e outro em 2250 Hz. Na **Figura 6**, para a variante alongada, têm-se um primeiro formante em 1650 Hz, outro em 2550 Hz e um terceiro em 3800 Hz. Também identificam-se três antiformantes em torno de 1100 Hz, 2000 Hz e 3250 Hz.

Nota-se uma diferença na quantidade de formantes e antiformantes entre as pronúncias de Gisele, o que evidencia que a variante alongada demanda mais energia para ser produzida e, ao mesmo tempo, é a que mais “sofre” efeito da obstrução da passagem de ar pela cavidade oral. Uma explicação possível para essas evidências pode estar no fato de tal variante ter, como se viu, maior tempo de duração na produção. Além disso, o valor do segundo formante de [ẽ] é 200 Hz maior do que o de [ẽ:] – evidencia de uma anteriorização da língua¹⁷. Por fim, na produção de [ẽ:], a língua está ligeiramente mais abaixada do que na de [ẽ], dado o valor maior do primeiro formante¹⁸.

1.2.2 FALANTE RICARDO

Ricardo leu a seguinte passagem para o experimento: “Lá onde eu morava era uma garotada... Eu, o Pedro, a Juliana... E, um dia, sem querer, a gente furou um cano e foi uma aguaceira.” As **Figura 7** e **8** abaixo apresentam

¹⁷ O segundo formante relaciona-se ao eixo horizontal da língua: quão maior seu valor, mais anterior está a língua, ao passo que, quão menor, mais posterior.

¹⁸ O primeiro formante relaciona-se ao eixo vertical da língua: quão maior seu valor, mais baixa está a língua, ao passo que, quão menor, mais alta.

os espectrogramas de banda larga (janela de 5 ms) da palavra “Juliana” com [ẽ] e [ẽ:], respectivamente.

Figura 7 Espectrograma de banda larga (janela de 5 ms) da palavra “Juliana” com a variante [ẽ] produzida pelo falante Ricardo

Figura 8 Espectrograma de banda larga (janela de 5 ms) da palavra “Juliana” com a variante [ẽ:] produzida pelo falante Ricardo

Aqui também há diferença no tempo de produção das vogais em si, com a alongada durando 75 milissegundos a mais do que [ẽ] – valor bem próximo à Gisele (74 ms). O segmento nasal de [ẽ:] é também maior do que de [ẽ], em 19 ms. Portanto, os termos “variante não alongada” e “variante alongada” são coerentes. Passa-se, agora, às análises dos espectros e cepstros.

Figura 9 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ] do falante Ricardo na palavra “Juliana”

Figura 10 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ:] do falante Ricardo na palavra “Juliana”

Para a variante não alongada, como mostra a **Figura 9** tem-se o primeiro formante em torno de 600 Hz, o segundo em 1500 Hz e o terceiro em 2800 Hz. O

primeiro antiformante encontra-se por volta de 400 Hz, o segundo em 1100 Hz e o terceiro em 2450 Hz. Na alongada, representada pela **Figura 10**, o primeiro formante encontra-se em torno de 1500 Hz, o segundo em 3200 Hz e o terceiro em 3800 Hz. Os antiformantes estão em 1050 Hz, 1800 Hz e 2550 Hz. Os picos menores em 1950 Hz e 2400 Hz podem ser *formantes nasais*, que compensam a redução da intensidade dos formantes da vogal pelo acoplamento dos tratos (Souza e Pacheco, 2012: 404).

Para Ricardo, a língua estava bem mais baixa e anterior na produção da variante alongada (relativamente a sua par), visto os valores elevados de seus primeiro e segundo formantes, respectivamente. O fato de [ẽ:] apresentar formantes nasais corrobora o fato de que essa variante sofre maiores efeitos do acoplamento dos tratos e obstrução do som na cavidade oral.

1.2.3 FALANTE BRUNA

A paulistana Bruna leu o trecho “Eu brincava muito com o Lucas, com a Fabiana... A gente tinha uma brincadeira assim de amarrar um pano no olho e pegar o outro... Era muito legal!” As **Figura 11** e **12** abaixo apresentam os espectrogramas de banda larga (janela de 5 ms) da palavra “Fabiana” com [ẽ] e [ẽ:], respectivamente.

Figura 11 Espectrograma de banda larga (janela de 5 ms) da palavra “Fabiana” com a variante [ẽ] produzida pela falante Bruna

Figura 12 Espectrograma de banda larga (janela de 5 ms) da palavra “Fabiana” com a variante [ẽ:] produzida pela falante Bruna

Mais uma vez, evidencia-se a diferença no tempo de produção das vogais em si: [ẽ:] dura 69 milissegundos a mais do que sua par. Aqui, no entanto, /N/ da variante não alongada dura 21 ms a mais do que o segmento nasal de sua par – o que, contudo, não anula o fato de [ẽ:] durar mais tempo. Observam-se, a seguir, os traçados de seus parâmetros acústicos.

Figura 13 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ] da falante Bruna na palavra “Fabiana”

Figura 14 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ:] da falante Bruna na palavra “Fabiana”

A **Figura 13** mostra o primeiro formante da variante não alongada em torno de 1850 Hz, e o segundo em 2800 Hz. Os antiformantes estão perto de 1300 Hz e 2500 Hz. Para a alongada, na **Figura 14**, vê-se que o primeiro formante está em 1900 Hz, o segundo em 2650 Hz. Os antiformantes estão em torno de 1200 Hz, outro em 2250 Hz e 3500 Hz. O pico menor em 3750 Hz é, possivelmente, um formante nasal.

Dessa forma, sabe-se que, para Bruna, a língua estava mais baixa e mais posterior na produção da alongada, relativamente à não alongada. Além disso, novamente se percebe, pela quantidade maior de antiformantes de [ẽ:], que tal variante sofre mais efeito do modo de articulação das nasais.

1.2.4 FALANTE RODOLFO

O último falante apresentado no experimento foi Rodolfo, que leu a seguinte passagem: “Eu brincava com o Caio, com a Adriana... E, assim, na hora de montar time pra jogar bola, sempre tinha um fulano que discutia”. A seguir,

estão as **Figura15** e **16** com os espectrogramas da palavra “Adriana” com [ẽ] e [ẽ:], respectivamente.

Figura 15 Espectrograma de banda larga (janela de 5 ms) da palavra “Adriana” com a variante [ẽ] produzida pelo falante Rodolfo

Figura 16 Espectrograma de banda larga (janela de 5 ms) da palavra “Adriana” com a variante [ẽ:] produzida pelo falante Rodolfo

Aqui, a vogal e o segmento nasal da variante alongada duram 119 ms e 6 ms a mais do que a não alongada, respectivamente. Observam-se as diferenças entre seus formantes e antiformantes a seguir.

Figura 17 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ] do falante Rodolfo na palavra “Adriana”

Figura 18 Espectro (janela de 30 ms) em vermelho e cepstro (filtragem de 350 Hz) em azul da região central da vogal [ẽ:] do falante Rodolfo na palavra “Adriana”

Na **Figura 17**, para a variante não alongada, têm-se o primeiro formante por volta de 850 Hz, o segundo em 1480 Hz e o terceiro em 2750 Hz. Também são dois os anti-formantes: um em torno de 700 Hz, outro em 950 Hz e um

terceiro em 2000 Hz. A variante alongada de Rodolfo é a que mais se diferencia daquela produzida pelos outros falantes. Sua energia de produção é baixa, pois, como se observa na **Figura 18**, os picos não são altos a partir de 470 Hz. Isso dificulta a análise de seus parâmetros acústicos. Seu primeiro formante parece estar em torno de 800 Hz e o segundo em 1500 Hz. A partir de então, a energia, que já era baixa, cai consideravelmente, com o aparecimento de vários formantes nasais. O primeiro deles parece estar em 1250 Hz e outro em 2200 Hz. Os antiformantes estão em 750 Hz, outro em 1100 Hz e um terceiro em 2000 Hz. A baixa energia, com forte ação de formantes nasais, evidencia mais uma vez que a variante alongada é mais “lesada” por conta do acoplamento dos tratos e obstrução da passagem de ar pela boca. Aqui também a língua está mais baixa e anterior na produção da variante alongada comparativamente à não alongada.

Foi objetivo deste capítulo entender minimamente como as variantes de (AN) se configuram acusticamente e se diferenciam uma da outra. De modo geral, viu-se que, relativamente à variante não alongada, a alongada é produzida: i) com maior tempo de duração; e ii) com a língua mais abaixada. Para Gisele e Bruna, no eixo horizontal, a língua encontra-se mais anterior, enquanto para Ricardo e Rodolfo, ela está mais posterior. Isso, contudo, não teve efeito na percepção sociolinguística, como mostra o **Capítulo 4**.

Antes de seguir com as hipóteses e o experimento deste trabalho, é importante discutir os estudos que o antecederam e o ajudaram tanto na formulação das questões da pesquisa quanto na formulação do próprio experimento.

Capítulo 2

Estudos de percepção da fala

Neste capítulo, discutem-se trabalhos sobre avaliações sociolinguísticas, desde seu início, na psicologia social, até chegar aos estudos sociolinguísticos, com mais detalhamento no tratamento dado ao lado social da variação em trabalhos mais recentes da sociolinguística, com as chamadas “ondas” (Eckert 2008, 2012). Essa discussão é pertinente na medida em que tais trabalhos dão suporte a essa pesquisa, tanto no tocante à elaboração dos seus objetivos quanto às metodologias de coleta e análise dos dados.

2.1 DA PSICOLOGIA SOCIAL À SOCIOLINGÜÍSTICA

Pesquisas empíricas sobre avaliações sociolinguísticas começaram na Psicologia Social com o trabalho de Pear (1931, *apud* Giles e Billings, 2004; Campbell-Kibler, 2006) sobre os efeitos de diferentes variedades do inglês britânico e do convívio social na avaliação de perfis de locutores de rádio da BBC. Com base na escuta de estímulos produzidos por diferentes falantes (elemento constitutivo da técnica *verbal guise*), os ouvintes preencheram formulários acerca da idade, sexo, ocupação, local de nascimento do falante, entre outras informações. Apesar dos resultados das análises dos dados dos ouvintes que não conviviam com os locutores diferirem significativamente daqueles que os conheciam, havia certa homogeneidade em suas respostas – uma evidência de que compartilhavam noções estereotípicas para cada falante. Em outras palavras, o estudo conseguiu estabelecer que os ouvintes não são muito habilidosos em ler a personalidade através fala, mas tendem a concordar fortemente uns com os outros em suas avaliações (Campbell-Kibler, 2006: 12).

Em um dos trabalhos pioneiros na Linguística, Lambert *et al.* (1960) estudaram atitudes inter-étnicas, a fim de entender como franco e anglo-canadenses percebem-se uns aos outros e a si mesmos na comunidade bilíngue

de Montreal. Com o objetivo de amenizar os efeitos das características fisiológicas, como tamanho e forma do trato vocal, na percepção sociolinguística, os autores gravaram o mesmo falante produzindo estímulos diferentes – mas que variam em algum domínio –, inaugurando a técnica *matched-guise* (ou *técnica de estímulos pareados* na proposta de tradução de Oushiro, 2015), utilizada até hoje nos estudos de percepção – com algumas adaptações (Campbell-Kibler, 2009; Levon, 2014; Mendes, 2018; *inter alia*). A técnica visa a eliciar atitudes subjetivas e encobertas acerca de diferentes variedades ou variantes linguísticas, ou mesmo línguas, pelos ouvintes.

Nesse estudo, Lambert *et al.* (1960) gravaram falantes bilíngues lendo a mesma passagem filosófica em inglês e em francês, e essas gravações serviram como estímulos para o experimento. Tomou-se cuidado para que soassem naturais e autênticas, no sentido de não parecerem uma leitura forçada. Com base nesses estímulos, os 130 ouvintes preencheram um formulário contendo: i) 14 escalas de diferenciais semânticos de seis pontos (indo de menos/pouco X a mais/muito X); ii) ocupação do falante; iii) três escalas para medir suas atitudes linguísticas e preconceitos; e iv) informações pessoais. Dos muitos resultados obtidos, destacam-se: a) anglo-canadenses avaliaram falantes do seu próprio grupo étnico mais positivamente em metade das 14 escalas; e b) franco-canadenses não apenas foram na mesma direção, como avaliaram o outro grupo étnico mais positivamente em 10 das 14 escalas. Tais resultados mostram como falantes podem ter atitudes negativas em relação à sua própria comunidade e como o código ou variedade tem efeito na percepção dos falantes, o que permite discussões acerca da percepção da língua em determinado contexto sociopolítico.

Para Labov (2006 [1966]: 102), este estudo de Lambert *et al.* (1960) é uma engenhosidade, na medida em que investiga sistematicamente as reações subjetivas dos ouvintes – o que o autor considera ser “a questão mais obscura e difícil” dos estudos linguísticos (idem.: 265). Interessado nessa grande questão,

Labov estudou as reações de falantes novaiorquinos quanto a cinco variáveis fonológicas: (r), (oh), (æh), (th) e (dh). Esse interesse, no entanto, surgiu a partir dos resultados de estudos prévios sobre produção.

Nesses primeiros estudos, Labov (2006 [1966]) evidenciou, por exemplo, que o uso das variantes de (r)¹⁹ era condicionado por dois fatores extralinguísticos que interagem entre si: classe social e estilo (aqui, referente à atenção prestada à fala). O estudo verificou que todas as classes sociais tendiam a produzir mais a variante padrão em contextos de fala mais monitorados. Contudo, o mais interessante é que, ainda nesses contextos, a taxa de produção da classe média alta superava a da classe alta para essa variante. Tal resultado, segundo o autor, é consequência da “hipercorreção” (no sentido social), visto que os falantes da classe média alta querem soar como pertencentes a uma classe ainda mais alta. Nesse sentido, interessou a Labov entender como os falantes se sentiam ou pensavam acerca das variantes linguísticas: elas indiciam um *status* social?, algumas são “melhores”, mais “bonitas”, mais “corretas” do que outras?.

Para responder a tais questões, o autor aplicou testes de reações subjetivas. Neles, Labov pediu a 122 novaiorquinos que se imaginassem na situação de gerente de uma grande empresa que estava prestes a contratar funcionárias. Eles ouviriam as candidatas às vagas (cinco mulheres gravadas para a pesquisa) em ordem aleatória e, num formulário, deveriam selecionar o cargo adequado para cada uma: personalidade televisiva, secretária executiva, recepcionista, operadora de telefone, vendedora, operária de fábrica ou nenhuma das anteriores. Além de sentenças-teste, havia as *zero sentence* “sentenças zero” (Labov, 2006 [1966]: 282) que não continham nenhuma das variáveis em observação e serviam apenas para controlar o teste.

¹⁹ Discute-se mais profundamente a produção variável de (r) em Nova Iorque na subseção seguinte “As três ondas da sociolinguística”.

No geral, os resultados evidenciaram que as variantes de maior prestígio social, produzidas frequentemente pelas classes socioeconômicas mais altas, tendiam a receber avaliações maiores nas escalas do formulário, relativamente às variantes de menor prestígio, favorecidas pelos falantes de classe social mais baixa.

Observa-se, como exemplo, os resultados para (r) por um ouvinte, na **Figura 19** abaixo. Os números distribuídos no topo da tabela (na figura, de 12 a 19), indicam as diferentes sentenças gravadas que foram tocadas. Cada barra circulada indica a resposta de determinado ouvinte. As sentenças 14 e 18 foram gravadas por uma falante, e as 15 e 19, por outra. A diferença entre esses pares de sentenças está na consistência da pronúncia rótica, ou seja, na quantidade de vezes em que (r) foi produzido. 14 e 15 tiveram produções mais consistentes, relativamente a 18 e 19. Neste caso, nota-se que, para as mesmas falantes, as sentenças mais consistentes são mais valoradas do que as sentenças menos consistentes. Dessa forma, é confirmada a hipótese laboviana de que a pronúncia rótica é a variante prestigiada socialmente.

Figura 19 Resposta ao formulário de reações subjetivas para (r) de um ouvinte (parte da Figura 11.4 de Labov 2006[1966]: 284)

A partir desse trabalho de Labov, ampliou-se o leque de estudos sociolinguísticos interessados em acessar os significados sociais das variantes (Campbell-Kibler, 2006; Mendes, 2017). A maioria desses estudos adaptou a técnica *matched-guise*, a fim de responder a questões das mais diversas naturezas teóricas. Esta dissertação, por exemplo, emprega tal metodologia a

fim de verificar quais significados sociais são associados à (AN), bem como para averiguar se a origem do ouvinte tem efeito na percepção das variantes dessa variável.

Em seu conhecido estudo de percepção sociolinguística, Campbell-Kibler (2011) analisou o efeito da interseção de *pitch*, sibilância e (ING) na percepção de masculinidade, sexualidade²⁰ e competência, e sua relação com estereótipos. Com base em quatro clipes de áudio de quatro falantes masculinos, 175 ouvintes responderam a um formulário que continha nove escalas graduadas em seis pontos, a saber: inteligência, grau de instrução, masculinidade, *gayness*²¹, amigabilidade, descontração, interioranidade²², educação (em termos de estudo/educação formal) e confiança.

Do modo como a autora organizou o formulário, os resultados mostram que: i) masculinidade e sexualidade correlacionam-se negativamente (de modo que quão mais masculino o falante é percebido, menos gay ele soa; e o contrário é verdadeiro); ii) masculinidade e competência correlacionam-se positivamente (tal que quão mais masculino o falante é percebido, mais competente ele soa; e o contrário é verdadeiro); e iii) sexualidade e competência não se correlacionam. No entanto, o segundo resultado não se realiza em duas situações. Primeiro, quando os ouvintes associam /s/ posterior a um homem heterossexual. Para estes ouvintes, os falantes soam masculinos sem necessariamente parecerem competentes. Segundo, quando /s/ anterior e *-ing* velar são associados a um homem gay. Nesse caso, o falante é percebido como muito competente, embora não soe masculino. Para este resultado, a autora argumenta que os ouvintes desconsideram a correlação mais geral entre masculinidade e competência para, ao invés disso, confiar apenas nas representações estereotipadas de tipos de

²⁰ No sentido de “quão gay” soava determinado falante.

²¹ Em termos de “quão gay” soava determinado falante.

²² Proposta de tradução da autora para “country”, no sentido de saber “quão interiorano/caipira” soa o falante.

homem: o heterossexual masculino e sem instrução de um lado e o homossexual efeminado e educado de outro.

Ainda sobre estereótipos, Levon (2014) analisa sua influência na percepção de três diferentes índices linguísticos, a saber: *pitch*, sibilância e TH anterior. Mais especificamente, o autor investigou até que ponto as atitudes e crenças estereotipadas sobre gênero (*masculino/não masculino*), sexualidade (*gay/não gay*) e classe social (*classe trabalhadora/classe não trabalhadora*) servem para permitir a associação de características linguísticas a significados sociais específicos, ao mesmo tempo em que bloqueia outras associações.

A partir da escuta de clipes de áudio de três falantes masculinos, 189 ouvintes responderam a um formulário online que continha oito escalas graduadas em seis pontos. Apesar de não ser confirmada a hipótese de que um estereótipo bloqueia o efeito de outro, foi possível verificar diferentes reações quando as categorias de sexualidade, gênero e classe social foram trabalhadas isoladamente (pelas variáveis sibilância, *pitch* e TH anterior, respectivamente) de quando foram combinadas entre si. Com relação à sibilância, /s/ anterior é associado a um falante menos masculino e mais gay, de forma que as duas escalas se correlacionam negativamente. No que concerne ao *pitch*, os falantes eram percebidos como menos competentes quando ouvidos no *pitch* médio mais alto relativamente a quando ouvidos com *pitch* médio mais baixo (Levon, 2014: 551). Por fim, observou-se há uma correlação positiva entre TH anterior e os julgamentos de simpatia (idem.: 556) – resultado que converge com o discurso popular de que a classe trabalhadora é mais atrativa socialmente (idem.: 555).

Contudo, o interessante é que, mesmo nesse caso, em que um efeito estereotipado poderia potencialmente emergir nas escalas de “competência” e “simpatia”²³, a sibilância neutralizou o potencial indicial de *pitch* e de TH

²³ Visto que os resultados do estudo mostraram que: (i) os falantes eram percebidos como menos competentes quando ouvidos no *pitch* médio mais alto em relação a quando ouvidos com *pitch* médio

anterior para todos os ouvintes, independentemente de suas atitudes. O autor propõe, por um viés cognitivo, que os ouvintes atendem seletivamente às variáveis mais perceptivelmente proeminentes nos estímulos. Por outro lado, foi confirmada a hipótese de que atitudes individuais moderam efeitos de bloqueio dos estereótipos e podem influenciar diretamente a percepção. Através de uma pesquisa de atitudes sobre papéis masculinos, observou-se que ouvintes mais conservadores/tradicionais associam *pitch* elevado e sibilação anterior a “não masculinidade” e a “homossexualidade”, ao passo que, para aqueles que rejeitam tais estereótipos, *pitch* e sibilação não têm esse efeito. A partir desse estudo, Levon (2014) sugere que a percepção sociolinguística dá-se pela soma de fatores cognitivos e atitudinais.

No Brasil, ainda são recentes os trabalhos de percepção, principalmente com a técnica *matched-guise*. Dentre eles, estão os estudos sobre o falar do Rio de Janeiro (Melo, 2017; Melo e Gomes, 2019; Carvalho, 2019) e de São Paulo (Mendes 2016, 2017, 2018; Oushiro, 2015; Santos, 2020) – nos quais este trabalho se insere.

Sobre a comunidade de fala carioca, Melo e Gomes (2019) investigaram se a variante glotal de coda (s), como em *me[h]mo*, é socialmente estigmatizada (segundo Labov, 2008[1972]) relativamente a sua par pós-alveolar, como em *me[ʃ]mo*, bem como se os falantes dessa comunidade compartilham as mesmas avaliações sobre tais variantes.

Para tanto, 36 ouvintes participaram de um teste de percepção com a técnica *matched-guise*, sendo 12 ouvintes universitários de classe média média e média baixa (grupo UFRJ²⁴), 12 adolescentes moradores de favelas, participantes de um curso para monitor de uma instituição de ensino e pesquisa e com escolarização regular (grupo Fiocruz) e 12 adolescentes também oriundos de favelas e que cumpriam medida socioeducativa de internação numa

mais baixo (Levon, 2014: 551); e (ii) há uma correlação positiva entre TH anterior e os julgamentos de "simpatia" (idem: 556).

²⁴ Em referência aos alunos da Universidade Federal do Rio de Janeiro.

instituição do estado (grupo EJLA²⁵) – os dois últimos grupos de classe média baixa.

Para o teste de percepção, os participantes ouviram 18 sentenças (12 sentenças com uma ou duas variantes da coda (s) e seis sentenças distratoras) produzidas por uma mesma falante carioca e, após ouvir cada uma, responderam qual das seguintes profissões eles achavam que a tinha dito: médica, auxiliar de enfermagem e faxineira. Segundo os autores, devido à obrigatoriedade de um curso superior, a profissão de médica tem maior prestígio social; a de faxineira, por não necessitar de escolaridade ou formação específica, tem menor prestígio; por fim, auxiliar de enfermagem fica no meio do *continuum*, por exigir uma formação específica – não necessariamente ensino superior – e ser geralmente exercida por pessoas de camadas sociais mais populares. Diferentes valores foram atribuídos a cada uma dessas ocupações: o valor 1 foi atribuído à médica, o 2 à profissão de auxiliar de enfermagem e o valor 3 à faxineira. A soma dos valores atribuídos às respostas de todos os ouvintes indica o grau de estigma ou prestígio das variantes de coda (s), de modo que valores menores indicam maior prestígio/menor estigma, e valores maiores, menor prestígio/menor estigma.

Além da variável linguística coda (s) (fricativa glotal *versus* fricativa pós-alveolar) e do grupo social do ouvinte (UFRJ, Fiocruz ou EJLA), controlou-se também a frequência de variantes por sentença (com uma ou duas ocorrências na mesma sentença).

Os resultados dos 432 dados analisados mostram que: i) não há efeito da frequência de variantes por sentença; ii) a nível global, a variante glotal tem um valor maior (315) do que a pós-alveolar (209) – ou seja, tem menor prestígio; iii) a nível do grupo, a diferença entre os valores das variantes só é significativa para UFRJ e Fiocruz, na mesma direção do observado a nível global – para EJLA, no entanto, a diferença é pequena e não significativa; e iv) se é possível

²⁵ Em referência à amostra EJLA (Escola João Luiz Alves).

verificar a associação da glotal ao estigma social, o contrário (associar a variante pós-alveolar a prestígio) não é verdadeiro para nenhum grupo.

A partir desses resultados, os autores concluem que: i) para coda (s), a frequência de variantes por sentença pode, de fato, não ter efeito na percepção de prestígio, ou talvez a metodologia empregada não tenha conseguido captar tal efeito (diferentemente da metodologia de Labov, 2011); ii) os grupos UFRJ e Fiocruz pertenceriam a mesma comunidade de fala, ao passo que o EJLA pertenceria a outra; iii) apesar de pertencerem a mesma classe social e de serem oriundos de favelas, as práticas e redes sociais dos grupos Fiocruz e EJLA parecem distanciá-los um do outro e aproximar os falantes da Fiocruz dos do grupo UFRJ (ambos têm maior contato com uma escolarização regular, diferentemente dos falantes do EJLA que vivem em situação de exclusão social); e iv) a variante pós-alveolar pode funcionar como característica regional da variedade carioca. Por fim, Melo e Gomes (2019: 144) defendem que o conceito de sistema adaptativo complexo (Beckner *et al.*, 2009 *apud* Melo e Gomes, 2019) é o que melhor explica os resultados encontrados, por entender que a língua não se desenvolve de maneira ordenada e uniforme, ao contrário, “é dependente das interações específicas dos indivíduos (Mufwene, 2008, p. 62, 2013, p. 207-208), que estão encaixadas na estrutura da sociedade, uma vez que a experiência do indivíduo é também social e histórica”.

Ainda sobre a comunidade de fala do Rio de Janeiro, Carvalho (2019) estudou a percepção sociolinguística dos pronomes *tu* e *você* na posição de sujeito. A hipótese central dessa dissertação é que a forma *você* é percebida pelos ouvintes cariocas como adequada a diversos contextos sociointeracionais, ao passo que *tu* associa-se a contextos mais específicos, a depender da relação estabelecida entre os falantes, de modo que é considerada mais adequada em relações simétricas (entre amigos, por exemplo) e menos adequada em relações assimétricas (entre um diretor de uma escola e um aluno, por exemplo). Além

disso, a forma *tu* indicaria mais informalidade, intimidade, *carioquicee* juventude, enquanto *você* soaria mais “neutra” na variedade carioca.

Para atestar suas hipóteses, a autora utilizou dois experimentos: um julgamento de adequação sociolinguística e um questionário de avaliação subjetiva (com base na técnica *matched-guise*). Para o primeiro experimento, adaptado da Psicolinguística, 40 ouvintes universitários assistiram a cenas de interação entre dois ou mais personagens, a fim de julgar a aceitabilidade e a naturalidade das sentenças com *tu* e *você* de acordo com o tipo de relação entre os falantes da cena (simétrica ou assimétrica) e ao tipo de frase (com ou sem verbo).

Para o segundo experimento, dois homens e duas mulheres cariocas foram gravados lendo o mesmo texto, com duas ocorrências da variável linguística, e quatro disfarces foram criados para cada falante: um disfarce com *tu* com verbo conjugado em 3ª pessoa do singular, outro com *você* e a mesma conjugação, um terceiro disfarce com um *você* e um *tu*, e o último com *tu* com verbo na 2ª pessoa do singular (doravante, *tu-2SG*). Após ouvir cada disfarce, os 50 participantes preencheram um questionário com 12 características, graduadas em diferenciais semânticos de cinco pontos, relacionadas à identidade nacional e regional (carioca, suburbano e favelado), à solidariedade (honestidade, intimidade e formalidade), a *status* (rico, trabalhador, letrado) e a atrativo pessoal (elegante, cortês, velho).

Dentre os vários resultados do primeiro experimento, destaca-se que os disfarces com *você* são bem aceitos em todos os casos, ao passo que *tu* é significativamente mais aceito em relações simétricas e em sentenças com verbo (mesmo que na conjugação de 3ª pessoa). A autora afirma que, por esse resultado, a percepção de *tu* pode não estar relacionada à falta de concordância canônica. As variáveis sociais sexo e região de residência não têm efeito na percepção de nenhuma forma em nenhum contexto.

Também se destacam alguns resultados obtidos no segundo experimento: i) para a dimensão da identidade local, *tu*-2SG é avaliado como pouco carioca, ao passo que *tu*, *você* e *você/tu* são muito cariocas; ii) na dimensão da solidariedade, *tu*-2SG esta no pólo mais formal, ao passo que *tu* e *você/tu* são mais informais e *você* é “neutro”; iii) na dimensão do *status*, os falantes foram percebidos como mais ricos ao usarem a variante *tu*-2SG, relativamente às outras; e iv) na dimensão do atrativo pessoal, essa mesma variante soa como pertencente a um falante mais cortês, elegante e mais velho.

Uma das conclusões atestadas é que *você* comporta-se como um *indicador* (Labov, 2008 [1972]), enquanto *tu* comporta-se como um *marcador*, isto é, está sujeito à variação estilística e sua percepção sofre efeito do tipo de interação entre os falantes.

Há também estudos no Brasil que investigam a percepção da “fala gay” (Mendes, 2014, 2017, 2018; Gonçalves, 2020). Mendes (2014) analisou os significados sociais da concordância nominal (doravante, (CN)) na fala de quatro homens (sob os pseudônimos de Carlos, Robson, Jaime e Lucas). Segundo o autor, os dois primeiros soam particularmente masculinos e dão a entender, em suas entrevistas metalinguísticas, que são heterossexuais, ao passo que os dois últimos falantes soam efeminados e identificam-se como gays.

No total, 100 participantes responderam a um formulário de percepção online nos moldes da técnica *matched-guise*. Os resultados evidenciam que a variante não padrão é associada a homens menos escolarizados, menos efeminados, menos formais, menos inteligentes e de classe social mais baixa. Ademais, o sexo do ouvinte tem efeito na percepção sociolinguística das variantes: apesar de ouvintes do sexo masculino e feminino considerarem a escolaridade como fator mais relevante na percepção da forma linguística (com (CN) padrão relacionada a mais escolaridade e (CN) não padrão, menos escolaridade), para participantes homens, o segundo fator mais relevante foi a masculinidade e, para mulheres, a inteligência e a classe social.

Com base nos dados do trabalho supracitado, Mendes (2017) estudou o efeito de (CN) na percepção de masculinidade dos quatro falantes homens a partir da análise do componente principal Efeminidade – composto pela escala “efeminidade”²⁶ – e da caixa de seleção “gay”²⁷.

De modo geral, os falantes soam “mais efeminados” quando ouvidos em seus disfarces com a variante padrão. Apesar de Lucas e Jayme soarem “mais efeminados” em ambos os seus disfarces (como já se esperava), tal percepção é acentuada quando ouvidos com a variante padrão. Carlos e Robson, por outro lado, tendem a soar “menos efeminados” em seus disfarces com a variante não padrão e “mais efeminados” com a padrão.

Quanto à caixa de seleção “gay”, o estudo evidencia que ela foi mais frequentemente assinalada: (i) para Lucas (50) e Jaime (42) do que para Carlos (15) e Robson (14); (ii) por ouvintes homens relativamente às mulheres; e (iii) nos disfarces com a variante padrão do que a não padrão – para ambos os gêneros/sexos. Todas as diferenças são significativas do ponto de vista estatístico. Dessa forma, o estudo de Mendes (2017) mostra que (CN) indicia significados baseados em gênero e sexualidade e que a variante não padrão pode auxiliar na projeção de *personae* masculinas estereotipadas.

Mendes, agora em seu estudo de 2016 sobre /e/ nasal na cidade de São Paulo, mostra que a variante ditongada leva a percepções de feminilidade, paulistanidade e pertencimento a áreas mais centrais da cidade, apesar de haver diferenças perceptuais a depender do falante. Como se mencionou na **Introdução** desta dissertação, o estudo analisou disfarces de quatro falantes: Lucas (o falante que soa mais efeminado), Robson (aquele que soa mais

²⁶ Os componentes principais são gerados a partir da Análise de Componentes Principais (ACP) – procedimento matemático cujo objetivo é averiguar se as respostas em cada uma das escalas se deram de maneira independente entre si, ou seja, verificar se há algum padrão de correlação nas respostas coletadas. Dessa forma, uma ou mais escalas podem integrar um mesmo componente. Em Mendes (2017), apenas uma escala (“efeminidade”) integrou o componente principal Efeminidade. O **Capítulo 4** desta dissertação aprofunda a discussão sobre a ACP.

²⁷ Característica não escalar que pode (ou não) ser atribuída a determinado falante. Cf. **Capítulo 3** desta dissertação.

masculino), Ariane e Janaína. A partir das respostas coletadas nesse estudo, o autor fez um novo estudo em 2018; desta vez, interessado nos efeitos de (EN) na percepção de gênero e sexualidade, novamente pela análise do componente Gênero – integrado pelas escalas de masculinidade e feminilidade em correlação negativa – e pela caixa de seleção “gay/lésbica” – como fez para (CN).

A variável não teve efeito na percepção de quão feminina soa Ariane e de quão masculino soa Robson, ao passo que, em seus disfarces com a variante ditongada, Janaína é percebida como mais feminina e Lucas, como menos masculino. Como Lucas “sempre” soa mais efeminado e não se tem resultados somente para as falantes mulheres, o autor sugere que, metodologicamente, talvez seja melhor não mesclar vozes de homens e mulheres em um mesmo experimento de percepção. Além disso, a variável não teve efeito sob a característica “gay/lésbica”, talvez pela quantidade de ouvintes.

Após Mendes (2016, 2017 *videsupra*) verificarem que (CN) e (EN) isoladamente têm efeito na percepção de gênero e sexualidade, o autor combinou as variáveis em um terceiro estudo, a fim de verificar se o efeito é potencializado. Após 208 ouvintes participarem de um experimento de percepção nos moldes da técnica *matched-guise*, Mendes (2018) evidenciou que as variáveis funcionam de maneira independente – ou seja, do ponto de vista estatístico, seus efeitos não interagem. O autor sugere que esse resultado tem base no fato das variáveis pertencerem a diferentes níveis linguísticos: (EN) é de natureza fonética, ao passo que (CN) pode ser entendida como de natureza morfossintática. Com isso, o autor questiona a colocação laboviana (1993, 2011) de que estruturas linguísticas de “nível mais profundo” são menos perceptíveis ao “monitorador sociolinguístico”²⁸ (proposta de tradução do autor para

²⁸ De acordo com Labov *et al.* (2011: 434), “O monitor pode ser concebido como um módulo de processamento e armazenamento separado ou como a capacidade de fazer um cálculo imediato a qualquer momento por uma inspeção de ocorrências lembradas” (tradução da autora). Os autores se

“sociolinguistic monitor”)²⁹ dos falantes por não se atribuírem significados sociais a elas.

Santos (2020) propõe o mesmo questionamento ao analisar os efeitos das morfologias verbais de subjuntivo e indicativo tanto na percepção de significados sociais de falantes ludovicenses e paulistanos, quanto na percepção de paulistanidade dos falantes paulistanos por ouvintes de São Luís e de São Paulo. Para fazê-lo, em um primeiro momento, o autor coletou dados de entrevistas metalinguísticas (ou *entrevistas abertas*, segundo o autor com base na tradução proposta por Oushiro, 2015) que serviram como base para o experimento de percepção, na medida em que algumas noções das entrevistas foram transformadas em escalas no formulário.

Num segundo momento, o autor aplicou um experimento de percepção, seguindo a técnica *matched-guise*, com disfarces produzidos por quatro falantes (dois ludovicenses e dois paulistanos), contendo orações subordinadas no modo indicativo ou no subjuntivo, sendo elas: i) adverbiais com *embora* ou *talvez* (“*embora estivesse/estava* aqui, a oportunidade foi dada a outro candidato” e “*talvez estivesse/estava* no tempo de perguntar aos funcionários sobre sua satisfação no trabalho”); e ii) substantivas introduzidas por *querer* ou *acreditar* (“o chefe *quer* que a secretária *permaneça/permanece* na reunião” e “ele *acredita* que todos *façam/fazem* a sua parte”). No total, 501 ouvintes responderam a um formulário contendo: i) oito variáveis quantitativas (amigabilidade, antipatia, educação [no sentido de polidez], escolaridade, formalidade, inteligência, paulistanidade e seriedade); ii) três variáveis nominais/escalas (nível de escolaridade, faixa etária e classe social); e iii) uma seção para informações pessoais do ouvinte (sexo, escolaridade, idade etc.).

mantêm neutros quanto a essa questão (idem.: 435). De qualquer forma, assumem que o monitor é responsável por rastrear, armazenar e processar informações sobre variação linguística.

²⁹ Mendes (2018: 32) afirma que o termo “monitorador” é mais eloquente do que “monitor” por sinalizar que “ocorre uma ‘ação’ (ainda que inconsciente), por parte do falante/ouvinte, de monitorar a enunciação (própria ou de outrem)”.

Os resultados mostram que as oito escalas reduziram-se a quatro componentes principais que se diferiram a depender do grupo de ouvintes. Do modo como se configurava o formulário, têm-se para os ludovicenses: i) Competência (formada pela correlação positiva de escolaridade e inteligência); ii) Antipatia (amigabilidade e antipatia correlacionando-se negativamente); iii) Paulistanidade (composto somente pela escala paulistanidade); e iv) Formalidade (correlação positiva de formalidade e seriedade). Para os ouvintes paulistanos, os componentes são: i) Antipatia (em que antipatia se correlaciona negativamente com educação e amigabilidade que, por sua vez, relacionam-se positivamente); ii) Competência (este e o próximo componente formados pelas mesmas escalas dos ludovicenses); iii) Paulistanidade; e iv) Seriedade (unicamente composto pela escala seriedade).

De modo geral, os falantes soam mais competentes e antipáticos quando ouvidos na forma subjuntiva, por ambos os grupos de ouvintes. No entanto, falantes ludovicenses são percebidos como mais sérios e formais do que os paulistanos. Apesar da paulistanidade para falantes paulistanos, há divergência nas respostas dos ouvintes: paulistanos soam mais paulistanos em seus disfarces no subjuntivo para ouvintes ludovicenses, ao passo que, para ouvintes paulistanos, a percepção de paulistanidade é mais acentuada nos disfarces no indicativo. O autor sugere que esse resultado se dá porque: i) ludovicenses acreditam que paulistanos estão na “roda econômica” do país e estão mais cercados de bens culturais do que pessoas de outras partes do Brasil; e ii) ouvintes paulistanos identificam-se com o indicativo, mesmo não apresentando metacomentários sobre seu uso. Outra divergência está nas respostas para a variável “nível de escolaridade”: ouvintes paulistanos percebem todos os falantes como de nível superior, independentemente do disfarce, enquanto ouvintes ludovicenses associam os disfarces no indicativo a pessoas com escolaridade mediana. Segundo Santos (2020: vi), tais resultados

“sugerem que a percepção dessas formas é socialmente estratificada em São Luís, mas não em São Paulo”.

Para além de evidenciar que variáveis morfosintáticas também estão sujeitas a avaliações subjetivas – objetando Labov (1993, 2011) –, Santos (2020) mostra a importância de se considerar a origem dos ouvintes nos estudos de percepção sociolinguística, tal qual faz esta dissertação.

Ainda sobre a percepção do falar paulistano, Oushiro (2015) investigou os significados sociais de (-r) em coda silábica por meio de duas distintas metodologias de estudo. Num primeiro momento, a autora aplicou *entrevistas abertas*³⁰ (proposta de tradução para “*open-ended interviews*” de Campbell-Kibler 2006, 2009) com 30 ouvintes, a fim de coletar discursos metalinguísticos que surgiam espontaneamente a partir da escuta de diferentes disfarces (gravados por duas falantes mulheres e dois homens paulistanos), contendo de 4 a 7 ocorrências somente de tepe ou somente de retroflexo. Esses disfarces eram idênticos em todos os aspectos, exceto pela pronúncia da coda (-r), que foi digitalmente controlada no Praat (Boersma & Weenink, 2012). As respostas coletadas ajudaram na formulação do experimento de percepção com base na técnica *matched-guise*, o segundo momento da pesquisa.

Nesse experimento, 185 ouvintes (dos quais 120 eram paulistanos) escutaram clipes de áudio de quatro falantes e preencheram um formulário contendo: i) 10 variáveis quantitativas, graduadas em diferenciais semânticos de cinco pontos (extroversão, escolarização, feminilidade/masculinidade, inteligência, formalidade, amigabilidade, paulistanidade, ter ou não sotaque, ter ou não amigos, centralidade do bairro); ii) quatro variáveis qualitativas (zona de residência, faixa etária, escolaridade e classe social); e iii) 30 caixas de seleção numa lista de características (alta, branca, tímida, descolada, entre outros). Coletaram-se também os dados pessoais dos ouvintes.

³⁰ As entrevistas são “abertas” no sentido de não terem um roteiro pré-definido e porque perguntas bem gerais são feitas aos ouvintes (p.ex., “Como você imagina que seja essa pessoa que está falando?”, “Como ela é fisicamente?”, “E a personalidade dela?”, entre outras) (cf. Oushiro, 2015: 275).

Os resultados do experimento de percepção mostram que o retroflexo associa-se a falantes paulistanos que soam como pertencentes à classe e nível de escolaridade baixos, menos formais, articulados, sofisticados e de regiões periféricas da cidade; porém, muito trabalhadores, simples, sinceros e solidários. Além disso, o estudo revela que a região de residência e a origem dos ouvintes influenciam na percepção de paulistanidade e de sotaque das variantes: moradores de bairros periféricos de São Paulo “aceitam” mais o retroflexo como paulistano, enquanto migrantes de outros estados não percebem uma diferença significativa entre as variantes para paulistanidade. Contudo, pessoas da região metropolitana e do interior, moradoras de áreas mais periféricas, amenizam o retroflexo como um “sotaque”, enquanto migrantes de outros estados e moradores de áreas mais centrais o consideram uma forma mais marcada do tepe. Dessa forma, o trabalho de Oushiro possibilita a discussão sobre a relevância de se analisar o local de residência dos ouvintes, bem como sua origem, nos estudos de percepção – ainda mais acerca de variantes “registradas” (Agha, 2007) como pertencentes a determinado grupo.

Ainda sobre o (-r) em coda, Soriano (2016) interessou-se em estudar o que se entende como “rcapira” e “r paulistano”. De modo mais técnico, a autora investigou como as características sociais dos ouvintes, bem como os significados sociais das variantes, têm efeito sobre a maneira como são atribuídas diferenças acústicas a pares formados entre cinco variantes de (-r) em coda silábica: i) vibrante com três batidas, ii) vibrante com duas batidas, iii) tepe (vibrante simples), iv) aproximante alveolar, e v) aproximante retroflexa.

Para tanto, 109 paulistanos participaram de um experimento de percepção sociofonética (Campbell-Kibler, 2006, 2007, 2009; Hay&Drager, 2007 *apud* Soriano, 2016). Um único falante homem foi gravado, com o objetivo de controlar quaisquer interferências que desviassem a atenção do (-r). Após escutar esse falante, os ouvintes tinham que indicar, em uma escala numérica

(cujos valores eram invisíveis aos participantes) indo de “Igual” (0 – zero) a “Muito diferente” (100 – cem), a percepção que tinham acerca das pronúncias da mesma palavra pronunciada duas vezes seguidas – sendo o segmento rótico a única diferença entre elas. Dado o objetivo do experimento, não se mencionou tal diferença aos ouvintes.

O estudo evidenciou que as percepções acústicas das variantes de (-r) em coda não se correlacionam à Tonicidade da Sílabas, Vogal Precedente nem à Consoante Seguinte, mas as respostas dos ouvintes variam conforme seu Sexo/Gênero, Local de Residência, Região de Nascimento na Cidade e Variante que Reconhece na Própria Fala: i) no geral, as mulheres atribuíram as maiores diferenças aos pares ouvidos, assim como os moradores de regiões mais centrais; ii) pessoas que nasceram na “região do retroflexo” (Goiás, Paraná, interior de São Paulo, sul de Minas Gerais) são os que menos diferenciaram as aproximantes; e iii) o grupo que disse reconhecer vibrantes e aproximantes na sua fala foi o que menos atribuiu diferenças a essas variantes. Dessa forma, confirmou-se a hipótese de que fatores extralinguísticos têm efeito até mesmo na percepção fonética – um fenômeno que, *a priori*, segundo a autora, variaria de forma menos sistemática do que a produção sociolinguística.

É interessante observar como os estudos mais recentes de percepção sociolinguística fomentam a abordagem proposta por Eckert (2008, 2012) sobre como os significados sociais não são estáveis nem se dão com a variante isolada, mas emergem do contexto (sócio)interacional, podendo sofrer efeito de diversas variáveis sociolinguísticas simultaneamente. Essa exploração das interações entre variáveis é possível, talvez, não somente pelas discussões de *Zeitgeist*³¹ pós-modernista (Agha, 2005, 2007; Podesva, 2006, 2007a, 2007b, 2011; Eckert 2008, 2012; *inter alia*), mas pelo desenvolvimento de novas técnicas e plataformas de análises estatísticas. A discussão feita por Eckert tem sua gênese nas críticas ao tratamento dado aos significados sociais por uma onda mais

³¹ Termo alemão cuja tradução significa espírito da época, espírito do tempo ou sinal dos tempos.

tradicional da sociolinguística. Aprofundar tal discussão é importante na medida em que esta dissertação investiga os significados sociais de (AN), a partir de diferentes recursos de diferentes ondas, ou seja, porque hibridiza artifícios de ondas distintas.

2.2 AS TRÊS ONDAS DA SOCIOLINGUÍSTICA

Por meio de um debate epistemológico, Eckert (2012) propõe que a sociolinguística divide-se em três ondas de estudo que se diferenciam quanto ao enfoque e à metodologia de análise. No entanto, a autora não sugere que uma onda substitua a outra, visto os interesses e relevâncias de cada uma.

A primeira onda começa com o trabalho de Labov (1966 [2006]) sobre a produção linguística variável na cidade de Nova Iorque, como se discutiu na subseção anterior. O autor estudou a frequência e a distribuição do uso de diversas variantes de diferentes variáveis fonológicas, a fim de identificar um padrão linguístico. Tal padrão foi, então, correlacionado a diferenças paralelas na estrutura social (Sexo, Classe Social, Idade e Estilo – esta última, entendida como grau de atenção à fala) com a finalidade de isolar os fatores extralinguísticos condicionantes da variação.

Para o estudo do (r), por exemplo, Labov evidenciou que a variante não rótica, marca da fala local, tende a perder espaço para a variante rótica na medida em que a classe socioeconômica aumenta. Em outras palavras, a frequência de ocorrência da não rótica é maior para falantes de classe mais baixa, ao passo que, para os falantes de classe mais alta, a frequência maior é da variante rótica. No entanto, quando Classe Socioeconômica é cruzada com Estilo, observa-se um resultado interessante, conforme mostra a **Figura 20**: em contextos de maior monitoramento e atenção à fala, todas as classes tendem a produzir a variante rótica; inclusive, aqueles da classe média alta (6 a 8) tendem a superar as taxas de produção dos falantes da classe mais alta na leitura de

lista de palavras e de pares mínimos. Como se discutiu anteriormente, segundo Labov, isto ocorre porque os falantes de classe média alta querem soar como pertencentes a uma classe social ainda mais alta, num fenômeno cunhado pelo autor de “hipercorreção (social)”.

Figura 20 Frequência de ocorrência da variante rótica, de acordo com classe socioeconômica e estilo, no inglês de Nova Iorque (adaptado de Labov 2006 [1966] *apud* Mendes, 2018)

Dados os objetivos de Labov (1966 [2006]), Eckert (2012) afirma que o foco dos estudos de primeira onda estaria em desvendar padrões gerais de usos de uma comunidade de fala, por meio de análises quantitativas, buscando fatores linguísticos e extralinguísticos/sociais que explicariam os processos de variação e mudança.

Em Labov (2006 [1966]), Classe Socioeconômica, bem como o cruzamento desta variável com Estilo, são fatores sociais condicionantes da produção variável de (r). Eckert (2012) pontua, contudo, que, por essa perspectiva, o “lado social” da variação restringe-se a macrocategorias sociais (sexo, faixa etária, classe social, escolaridade etc.) às quais os falantes pertencem e que se correlacionam à tendência do uso de determinada variante. Dessa forma, a autora critica o tratamento secundário dado aos significados sociais da variação

nessa “onda” de estudos. A linguista argumenta, ainda, que esses trabalhos desconsideram:

- (i) a agentividade social³² do falante enquanto ser ativo com intenções e práticas comunicativas realizadas também por meio dos significados sociais das variantes linguísticas. Segundo a autora, essa onda de estudos posiciona o falante como um ser passivo às tendências observadas para sua macrocategoria social ou limitam a agentividade à autocorreção – fruto da atenção à fala –, por considerarem que os indivíduos, sensíveis ao *status* relativo das variedades de classes, afastam-se do vernáculo à medida que adotam formas mais padronizadas em sua fala mais cuidadosa (Eckert, 2012: 89); e
- (ii) as microcategorias sociais (ou categorias locais) ligadas a identidades e crenças compartilhadas por um grupo de indivíduos que se reafirma por meio das práticas linguísticas (Campbell-Kibler, 2011).

Por consequência de (i) e (ii), há o risco de se cair num falso essencialismo ao impor padrões de cima para baixo no intuito de explicar os fenômenos linguísticos variáveis.

No intuito de evitar que isso ocorra, a segunda onda dos estudos sociolinguísticos, ainda que de caráter quantitativo, emprega métodos etnográficos para focar em aspectos locais a partir de grupos menores de indivíduos. Nessa perspectiva, entende-se o vernáculo como uma expressão de identidade local ou de classe e a ele atribui-se a agentividade social do indivíduo, visto que, apesar do seu possível valor negativo a nível global, o vernáculo pode ser positivamente valorado a nível local (Eckert, 2012: 91). É o que evidencia Milroy (1987) em seu estudo sobre a variação fonológica em redes sociais de Belfast, capital da Irlanda do Norte. A autora observou a

³² Do inglês *social agency* (Eckert, 2012).

relação entre a interação de falantes da classe trabalhadora e a realização variável da fricativa interdental sonora do inglês em contexto intervocálico, como em *brother* (irmão), e evidenciou o efeito da “força da rede”: redes sociais mais densas, em que há forte interação entre os indivíduos (são colegas de trabalho e vizinhos, por exemplo) valorizam positivamente o vernáculo (apagamento de (th)), enquanto indivíduos que participam de redes dispersas e apresentam orientação para fora do grupo tendem a preferir a variante padrão. Dessa forma, a autora concluiu que o comportamento linguístico dos falantes é influenciado pela força de suas redes sociais e que práticas locais podem funcionar como resistência a normas supralocais.

Em seu próprio estudo sobre dois grupos de alunos do colégio *BeltenHigh*, em Detroit (Michigan, EUA), Eckert (2000) também evidenciou que a valoração do vernáculo deve ser entendida a partir de microcategorias sociais (ou categorias locais). Os *jocks*, da classe média, são atletas ligados à vida extracurricular da escola (esportes, principalmente), formam uma hierarquia rígida e competitiva, e mantêm relações de cooperação com o corpo docente e administrativo. Por sua vez, os *burnouts*, da classe trabalhadora, focam no currículo profissional e mantêm suas redes, identidades e vidas sociais atreladas à vizinhança e à aglomeração mais ampla, rejeitando o papel social da escola.

O estudo mostrou que os grupos apresentavam comportamento linguístico diferente (acerca da posteriorização de vogais médias e baixas, alçamento de (ay) e formas sintáticas para negação) a fim de constituírem seu estilo e, dessa forma, afastarem-se do estilo do outro grupo. É importante dizer que, na segunda onda, entende-se estilo como “algo essencial na construção de categorias sociais e não tem a ver apenas com a atenção direcionada à fala, mas sobretudo com práticas (tais como participar de certos ‘rolês’, em princípio contra-indicados a *jocks*, sobretudo meninas *jocks*)” (Mendes, 2018: 33).

A pesquisa de Eckert evidenciou ainda que, apesar de existir alguma correlação entre o comportamento linguístico dos grupos e a categoria social dos pais, os padrões de variação continuaram se desenvolvendo na adolescência, mostrando claramente que tais padrões não são definidos na infância, mas servem como recursos na construção da identidade mais tardiamente na vida. Essa constatação indica que “correlações mais amplas de classe não são simplesmente consequências da educação, ocupação e renda, mas refletem a dinâmica local enraizada em práticas e ideologias que moldam e, por sua vez, são moldadas por classe” (Eckert, 2012: 92). Contudo, a autora constata que, tal qual os estudos da primeira onda, os estudos da segunda onda enfocam em categorias sociais aparentemente estáticas (mesmo que microcategorias) e equiparam a identidade do falante às categorias nas quais ele se encontra. De certa forma, portanto, ainda há imposição de padrões sociolinguísticos de cima para baixo.

Finalmente, na terceira onda, a variação deixa de ser entendida como reflexo das identidades e de (macro ou micro)categorias sociais para ser interpretada enquanto *prácticalinguística* na qual os falantes posicionam-se estilisticamente no cenário social (Eckert, 2012: 93-4). Aqui, a noção de estilo é crucial para diferir a terceira onda das demais: é o *locus* privilegiado em que, pela (re)combinação de variantes em um processo contínuo de *bricolage*, atualizam-se os significados sociais e vão-se criando outros (Irvine, 2001; Eckert, 2001, 2012). Se as duas primeiras ondas entendem os significados da variação enquanto consequência incidental do ambiente social, a terceira onda os vê como uma característica essencial da linguagem (idem.), de forma que o falante é visto como ser ativo – e não passivo às categorias sociais nas quais se encontra. Por essa perspectiva, o usuário da língua se mune de formas linguísticas e não linguísticas relacionadas a significados sociais diversos – que emergem durante o processo (sócio)interacional – para projetar uma *persona* (ou

personae) de si mesmo, bem como para criar *personae* do seu interlocutor. Em outras palavras, para a terceira onda, o falante performa diferentes estilos.

Podesva (2006, 2007a, 2007b, 2011) inaugura os estudos dessa onda na medida em que evidenciam que: (i) as variantes e seus significados funcionam diferentemente quando analisadas em interação entre si do que quando analisadas isoladamente; e (ii) o falante é agente, porque atua sobre os diferentes significados sociais que emergem em distintas situações (sócio)interacionais, mudando de estilo a depender dos objetivos do seu engajamento na conversa. Para tanto, o autor estuda a pronúncia variável de /-t/ e /-d/ em final de palavras do inglês (como *fat* e *bad*), a variação na forma e nas propriedades fonéticas de sentenças declarativas e, por fim, o falsete na fala de um gay (sob o pseudônimo de “Heath”) em duas situações: em seu consultório em que atua como médico e num churrasco com amigos.

Figura 21 Diagramas composicionais das *personae* “doutor zeloso” e “diva gay” performadas por Heath em duas diferentes situações: consultório médico e churrasco com amigos (adaptado de Podesva 2007b com tradução proposta por Mendes 2018)

Em seu local de trabalho, Heath frequentemente não apaga as oclusivas coronais, um índice de formalidade (Campbell-Kibler, 2006; Eckert, 2008),

tendendo à hiperarticularização, um índice de competência no inglês (Eckert, 2003; Podesva *et al.*, 2006). Nessa situação, o falante também produz com frequência sentenças declarativas com entonação ascendente, fato que o auxilia a assumir uma postura acolhedora e não ameaçadora na comunicação com seu paciente (Podesva, 2006, 2007a). Por fim, relativamente ao falsete, sua realização fonética mais fraca (menor duração e menor frequência fundamental³³) e sua distribuição restrita (ocorrendo quase exclusivamente nos marcadores de discurso, geralmente precedendo comandos), indiciam expressividade, mas de modo comedido. Dessa forma, Heath só performa uma *persona* médica atenciosa e zelosa se as variantes e seus significados sociais estiverem em interação na conversa. Em outras palavras, esse estilo é marcado não apenas por formalidade, competência, comportamento acolhedor e não ameaçador ou expressividade comedida, mas pela confluência desses significados sociais e de seus índices linguísticos, como mostra a **Figura 21**.

No churrasco com os amigos, o estilo de Heath é outro. Nessa situação, a frequente não realização de /-t/ e /-d/ indicia informalidade, ao passo que sua realização forte, juntamente à hiperarticularização das vogais (não anteriorizadas quando pré-coronais), indicia delicadeza, sugerindo ainda mais precisão a sua fala. Nessa situação de descontração entre amigos, os contornos entonacionais decrescentes de Heath indiciam animação. Finalmente, seu falsete, como discutido anteriormente, o ajuda a criar uma *persona* expressiva – neste caso, no entanto, não há moderação, mas um uso intenso e frequente desse índice linguístico. Como antes, esses significados podem não ser tão nítidos quando isolados, mas ganham clareza quando considerados em conjunto. Os significados sociais como informalidade, delicadeza, animação e expressividade permitem a Heath projetar sua *persona* de gay diva, um estilo caracterizado não apenas por ousadia, expressividade ou informalidade, mas por todas elas juntas, ou seja, de uma só vez (Podesva, 2007b).

³³ A mais baixa e a mais forte frequência componente da série harmônica de um som.

Como constata o autor (Podesva, 2007b), estilos ganham significado quando em contraste com outros estilos: o estilo gay diva de Heath só adquire tal noção ao se opor ao seu estilo médico atencioso e zeloso. Portanto, entende-se que a variação estilística é um sistema de contrastes, oposições.

Também interessado em estudar a interseção de variáveis na composição de diferentes estilos – além de observar se há padrões nos usos de tais variáveis para os falantes paulistanos –, Mendes (2018) analisou a produção de (CN) e (EN) em trechos de entrevista de uma hora de quatro falantes do sexo masculino. Tais objetivos surgiram a partir dos resultados de estudos anteriores sobre a percepção de cada variável isoladamente e em interação (Cf. subseção anterior deste capítulo).

O estudo evidencia que: (i) é muito baixa a frequência da variante não padrão de (CN) para todos os falantes, o que pode ser consequência do efeito intimidador da própria entrevista sociolinguística (entrevistador desconhecido e gravador diante do entrevistado – Cf. Labov, 1966); e (ii) o registro de (EN) é tão inconsistente (Agha, 2007) que os quatro falantes variam equilibradamente na produção das variantes ao longo de suas entrevistas. Desse modo, o autor afirma que analisar performances estilísticas com base em entrevistas sociolinguísticas assemelha-se a “assistir a uma peça de teatro a partir da coxia, através da fechadura de uma porta que a separa do palco” (idem.: 187). De toda forma, autor declara que tal metodologia é eloquente na medida em que, como evidenciou o estudo, os homens que se preocupam em soar mais masculinos tendem a produzir a variante não padrão mais frequentemente do que tendem a produzir a padrão aqueles que preferem não performar um estereótipo de masculinidade.

Após essa discussão das ondas da sociolinguística, é importante dizer que o presente estudo toma como foco e metodologia, respectivamente, o interesse em acessar significados sociais e a técnica *matched-guise* – característicos da terceira onda –, no intuito de buscar padrões sociolinguísticos

– objetivo da primeira onda –, ainda que para a percepção. Dessa forma, há certa hibridização da abordagem em “ondas” proposta por Eckert (2008, 2012). Quanto a isso, há certo consenso na área de que as ondas “não devem ser vistas como sobrepostas numa cronologia linear, mas como partes de um todo” (Camacho, 2010: 158). Entende-se, assim, que cada onda contribui a seu modo para estudos sociolinguísticos para, no final, ter-se um todo completo, visto que a primeira onda caracteriza-se por seu viés macrosociológico, a segunda pelo etnográfico e a terceira, pelo estilístico (Mendes, 2017).

Capítulo 3

Hipóteses e montagem do experimento

Este capítulo discute as metodologias aplicadas com base em Campbell-Kibler (2006), a saber: i) preparação dos estímulos usados no experimento de percepção; ii) realização de entrevistas metalinguísticas; e iii) formulação e aplicação do experimento de percepção com a técnica *matched-guise*.

3.1 PREPARAÇÃO DOS ESTÍMULOS

Para preparar os pares de disfarces do experimento de percepção, duas mulheres e dois homens – cujos pseudônimos são Bruna, Gisele, Ricardo e Rodolfo – foram gravados com o equipamento Tascam DR-100 MK II – com exceção de Ricardo, gravado com iPhone XR, pois o gravador profissional parecia captar uns ruídos estranhos. Bruna foi gravada na sala dos professores de uma escola. Gisele, em sua casa. Ricardo foi gravado na sala dos professores de uma universidade. Rodolfo, na sala vazia de um cinema. A pesquisadora gravou os falantes lendo as passagens abaixo, ora com [ẽ:] ora com [ẽ], quantas vezes necessárias até que soassem naturais e espontâneas – no sentido de não parecem uma leitura forçada.

Bruna:*Eu brincava muito com o Lucas, com a **Fabiana**... A gente tinha uma brincadeira assim de amarrar um **pano** no olho e pegar o outro... Era muito legal!*

Gisele:*Eu era muito amiga do Paulo, da **Mariana**... A gente fingia ser príncipe, princesa, herói, **cigano**, pirata... A gente criava muita história!*

Ricardo:*Lá onde eu morava era uma garotada... Eu, o Pedro, a **Juliana**... E, um dia, sem querer, a gente furou um **cano** e foi uma aguaceira!*

Rodolfo:*Eu brincava com o Caio, com a **Adriana**... E, assim, na hora de montar time pra jogar bola, sempre tinha um **fulano** que discutia!*

Cada sentença apresenta duas ocorrências de (AN) – em destaquena transcrição acima. Manteve-se o mesmo assunto (infância) na elaboração dos disfarces, na intenção de evitar possíveis efeitos de uma diferença de tópicos nas respostas dos ouvintes. As primeiras ocorrências de todos os falantes apresentam somente (AN) na sílaba, precedida de *i* com *-na* como sílaba final. Para Gisele e Rômulo, a segunda ocorrência está em palavras trissílabas com *-no* no final, ao passo que, na segunda ocorrência de Bianca e Ricardo, apesar de a sílaba final também ser *-no*, a palavra é bissílaba.

Os disfarces originais são aqueles com [ẽ:], de modo que as palavras com essa pronúncia foram recortadas e substituídas pelas mesmas palavras com a outra variante. Esse procedimento foi feito com o auxílio doPraat (Boersma e Weenink, 2012) e é necessário para que os pares de disfarces fiquem idênticos, salvo a variante produzida – elemento constitutivo da técnica *matched-guise*.

A qualidade dos estímulos foi testada com a ajuda de quatro professores da rede básica de ensino, ministrantes das áreas de exatas e ciências da natureza – ou seja, de fora da linguística. Inicialmente, apresentaram-se os clipes de áudios como parte de uma pesquisa de mestrado, que eles simplesmente teriam que ouvir e julgar quanto à qualidade (boa ou ruim, sem maiores detalhes). Todos disseram que os áudios estavam bons. No segundo momento, informou-se a eles que os áudios haviam sido manipulados e que as sentenças foram lidas. Em seguida, perguntou-se se isso era perceptível nos estímulos, ao que eles disseram que não se percebiam os cortes e os áudios pareciam fala espontânea, com exceção de Bruna.

Após esses retornos, seguiu-se com a organização dos disfarces, conforme mostra a **Tabela 1**. Os conjuntos **A1** e **A2** apresentam os disfarces com os falantes na mesma ordem, intercalando o gênero. Cada ouvinte ouviu apenas um dos dois conjuntos, ou seja, ouviu cada falante em apenas um de seus disfarces.

Tabela 1 Conjunto de estímulos	
A1	A2
Gisele -[ẽ:]	Gisele-[ẽ]
Ricardo-[ẽ]	Ricardo -[ẽ:]
Bianca-[ẽ:]	Bianca-[ẽ]
Rômulo-[ẽ]	Rômulo -[ẽ:]

Depois da preparação e organização dos estímulos, montou-se o formulário de percepção. No sentido de decidir quais características ou significados sociais integrariam o formulário, foram feitas, anteriormente, entrevistas metalinguísticas, com paulistanos e não paulistanos, cujos resultados se apresentam na subseção a seguir.

3.2 ENTREVISTAS METALINGUÍSTICAS

As entrevistas metalinguísticas (Campbell-Kibler 2009, 2011) foram feitas com 27 ouvintes paulistanos e 23 não paulistanos (do interior e de fora do estado) residentes em São Paulo ou que vão à cidade diariamente para estudar e/ou trabalhar. Todos estavam cientes de que respondiam a uma pesquisa de mestrado em Linguística; no entanto, nada se mencionou sobre “sotaque” ou algo semelhante.

Nas entrevistas, os ouvintes escutaram clipes de áudio de dois falantes da Amostra São Paulo 2010 (sob os pseudônimos de CarlosJ e VivianeC – sempre nessa ordem) e assistiram ao trecho de vídeo em que Adnet imita a menina do “nasalês”. Os clipes de áudio duravam, em média, 1 minuto e 14 segundos e apresentavam relatos dos falantes sobre sua vida pessoal no que

concerne à família e aos estudos, conforme os excertos a seguir³⁴ (as palavras em destaque foram pronunciadas com a variante alongada):

CarlosJ (CJ)

CJ *a minha mãe tem mais o hábito de vê-los né? Eu confesso que vejo mais nas nas festas de final de ano*

D1 *uhn*

CJ *e eventualmente alguns aniversários tal durante o ano a gente não... Não consegue ter um/ antes (havia)/ aí/ antes a gente/ nós quando a minha vó a mãe do meu pai era viva... Então todos todos os domingos nós íamos visitar então a família se reunia lá*

D1 *uhum*

CJ *agora... Depois que ela faleceu então acabou... Dispersando um pouco né? Então eu os encontro mesmo mais em aniversários e nas festas de final de ano Natal. A gente costuma pa/ passamos todos juntos Réveillon não (mas Natal nós) passamos juntos e tal quando... Consigo reencontrar todos assim*

D1 *aham. Quando você era criança você encontrava bastante?*

CJ *sim por causa da minha avó né? Então eheh eu convivi bastante com meus primos na infância por causa disso*

D1 *uhum*

CJ *que é todos todos os domingos à tarde né a gente passava lá*

D1 *tá... Você tem irmãos?*

CJ *é tenho uma irmã uma irmã mais nova um ano mais nova. Tenho trinta e um e ela tem trinta*

D1 *ah legal vocês... Se dão bem assim?*

CJ *simsim. Agora chega né? Já passou da época de brigar a gente... [risos-CJ-D1] Eh não não não na na infância normal né briga de irmão né eh ciuminho um quer atenção o outro não dá [risos-D1] Eh normal. Acho que são brigas de qualquer irmão mas agora hoje/*

VivianeC (VC)

VC *é bom... é bom. Mas estou com vontade de sair da minha casa.*

D1 *é mesmo? você mora com quem?*

VC *moro com a minha mãe. Morava eu minha mãe e minha irmã e aí a minha irmã foi morar em Portugal faz cinco anos já tsc então só eu e minha mãe agora só que já tenho vinte e seis anos né? [risos-D1]*

³⁴ Áudios e transcrições disponíveis no site do Projeto São Paulo 2010: <http://projetosp2010.fflch.usp.br/> [acessado em 22/01]

D1 *ai você quer sair da casa da mãe?*
 VC *então... Ah eu tenho vontade de sair não que eu tenha não tenho nenhum problema com ela a gente se dá super bem mas eu tenho vontade de sair. Só que é muito cara né? A vida em São Paulo.*
 D1 *é mesmo.*
 VC *tsc então. Por enquanto está fora de cogitação. Eu estava com a bolsa FAPESP aí a bolsa acabou tipo eu devia ter saído enquanto eu estava com a bolsa apesar de ser baixa sabe? Tem certos desafios que você tem que assumir e... E fazer. Agora a bolsa acabou eu ti/ então eu não aproveitei a oportunidade. Agora vai saber quando que eu vou conseguir... Trabalho porque né? eu te falei que meu mestrado acaba só em junho então*
 D1 *uhum*
 VC *até lá eu não vou trabalhar provavelmente*
 D1 *uhum... E você... Gosta do bairro?*
 VC *muito gosto muito. Esse é outro problema né? Porque como eu moro lá dos morel então estou morando lá desde os treze até agora que eu tenho vinte e seis. Tipo não só um período grande como de formação praticamente né? então quando você começa a andar sozinha na rua!*

Após escutar cada falante, o ouvinte respondeu livremente a perguntas bem gerais de percepção, como “O que você acha dessa pessoa?”, “Como você acha que ela é?”, “Que imagem dela vem a sua cabeça?”. Quando os ouvintes não diziam de onde eles achavam que o falante era, a pesquisadora os perguntava. O objetivo, nesse momento, foi verificar as reações que surgiram espontaneamente ao se escutaram os falantes.

Para o trecho de vídeo de Marcelo Adnet, as mesmas perguntas foram feitas. Tal trecho foi inserido na entrevista, porque se esperava que a maioria dos entrevistados prestaria atenção à variante “paulistana” e teria mais facilidade em detectá-la, dado o modo caricato como foi pronunciada.

As entrevistas foram aplicadas individualmente ou em grupos de até sete pessoas, permitindo que os ouvintes contribuíssem uns com os outros em suas respostas, fossem em termos de concordância ou discordância, como mostram, a seguir, os trechos (1) sobre VivianeC e (2) sobre CarlosJ:

(1) Kleber³⁵: *É, ou então esse povo blogueiro. Será que não é isso?*

³⁵Todos os ouvintes são representados por pseudônimos.

- Barbara: *É! Isso!*
 Denise: *Porque elas falam meio assim, né?*
 Barbara: *Pode ser também. Meio patricinha, meio blogueira de maquiagem, sabe? Alguma coisa assim.*
 Natália: *[imitando uma blogueira] “Oi, meninas! No dia de hoje...”*
 Kleber: *Eu não sou muito de assistir vídeo de blogueira, mas, eu percebo que o início dos vídeos tem esse dialeto...*
- (2) Larissa: *Eu acho que ele é branco. No começo, eu achei que não. Aí depois foi vindo uma imagem de uma pessoa branca, alta.*
 Luíza: *Eu não sei... Pior que eu não sei. Eu acho que não. Não, pra mim ele não é branco. Não sei por que, mas ele não é.*

Interessa notar, sobre o trecho (1), que os ouvintes estavam tentando descobrir com que caricatura “popular” aquela falante se parecia. Era como se a resposta estivesse “na ponta da língua”, mas não saísse. Quando a noção “blogueira” finalmente foi dita, muitos concordaram com certo “alívio”. Em relação ao trecho (2), Luíza discorda de Larissa com relação à etnia/cor branca de CarlosJ; no entanto, ela não o classifica como nenhuma das outras categorias étnicas.

Com base nas respostas dos ouvintese nos termos que eles empregaram para se referir aos falantes que ouviram, criaram-se *nuvens de palavras* no programa Wordle³⁶, que representam visualmente a frequência relativa com que as noções aparecem nas respostas, de forma que as mais frequentes estão em fonte maior e as menos frequentes, fonte menor. Como é de interesse dessa pesquisa averiguar se a origem dos ouvintese tem efeito na percepção de [ẽ:], criou-se uma nuvem de palavra por grupo (paulistanos e não paulistanos), para cada um dos três falantes (CarlosJ, VivianeC e Adnet).

A **Figura 22** mostra os termos atribuídos a CarlosJ pelos paulistanos. Muitas vezes, o falante é relacionado a um “paulistano” por ser “sem-sotaque”.

³⁶ Cf. <http://www.wordle.net/>.

Também apareceram noções como “coxinha”, “certinho”, “arrumadinho”, “cabelo ajeitadinho” e “homossexual”. Sabe-se que: i) há um discurso metalinguístico sobre gays “usarem diminutivo demais” (cf. Mendes, 2018: 4) e que uma pessoa pode usar esse índice linguístico para referir-se a um gay (cf. idem); ii) há os construtos sociais de que mulheres cuidam mais da beleza e de que gays querem “se parecer” com mulheres (Gaudio, 1994) e, por esse viés ideológico, homens heterossexuais que cuidam de sua aparência são vistos como gays; e iii) estudos sociolinguísticos mostram suficientemente que sexo/gênero e sexualidade correlacionam-se negativamente: aquilo que soa menos masculino/mais feminino soa mais gay (e vice-versa) (Gaudio, 1994; Levon, 2007, 2014; Campbell-Kibler, 2011; Mendes, 2018). Com essas questões em mente, variáveis de gênero foram incluídas no formulário do experimento de percepção na forma de: i) uma escala de *feminilidade* e outra de *masculinidade* com diferenciais semânticos de seis pontos (sendo, portanto, uma variável quantitativa); e ii) uma caixa de seleção representada pelo termo *Gay/Lésbica* (uma variável qualitativa) (ver subseção 3.3).

Figura 22 Nuvem de palavras associadas ao falante CarlosJ por ouvintes paulistanos

Ainda sobre CarlosJ, tem-se a **Figura 23** para as noções apresentadas por ouvintes não paulistanos. Nessa nuvem de palavras, “paulistano” é bem mais frequente do que na nuvem dos paulistanos, inclusive com a identificação de

[ẽ:] (representada por “ã”³⁷). Dessa forma, há indícios de que a variante seja um índice de paulistanidade para quem não é originalmente de São Paulo e que a variante parece estar abaixo do nível da consciência para paulistanos. Por isso, *paulistanidade* também apareceu como variável quantitativa no experimento de percepção reportado na subseção seguinte.

De modo geral, nas duas nuvens de palavras de CarlosJ, aparecem noções como “escolarizado” e “professor”. Dessa forma, *escolaridade*, *inteligência* e *formalidade* apareceram no formulário do experimento de percepção como variáveis quantitativas. Também se inseriu a variável *classesocial*, visto que noções como “cozinha”, “branco”, “classe média”, “família estruturada” podem ser associadas a falantes de classe social mais alta.

Figura 23 Nuvem de palavras associadas ao falante CarlosJ por ouvintes não paulistanos

Interessa observar a frequência com que CarlosJ foi associado a uma pessoa branca, inclusive por pessoas negras que, de alguma forma, identificaram-se com sua história: “Apesar de eu ter me identificado, eu imaginei ele um cara, sei lá...Mais alto, branco” (ouvinte sob pseudônimo de Murilo²⁵). Mais interessante ainda é o fato de CarlosJ ser negro. Com base nessas observações, colocou-se a etnia/cor *branca* como uma opção nas caixas de seleção, de forma que pudesse ser selecionada caso o ouvinte associasse a

³⁷ O programa Wordle não lê essa configuração da vogal (de “cabeça para baixo”) nem o diacrítico “:”.

variante a tal etnia/cor. Com base nos dados do IBGE sobre a população “parda” e “preta” do país se encontrar nos estratos socioeconômicos mais baixos ao passo que a população “branca” encontra-se nos mais altos³⁸, buscou-se verificar, no experimento de percepção, se há correlação entre etnia/cor e a classe social do falante.

Quanto à falante VivianeC, a **Figura 24** mostra as reações naturais que surgiram para paulistanos. Ela é associada a uma “paulistana”, da “capital”. Talvez o que tenha gerado esse tipo de metacommentário foi o fato de, no áudio, VivianeC dizer que ganhava bolsa de mestrado da Fapesp³⁹. Quanto à etnia, as poucas noções que aparecem são, ainda assim, “branco” e “branquinha”.

Figura 24 Nuvem de palavras associadas à falante VivianeC por ouvintes paulistanos

Na **Figura 25**, vemos que a noção de “paulistana” aparece bem mais nas respostas de não paulistanos para VivianeC, por conta, muitas vezes, do seu “sotaque” – o que quer que isso signifique para os ouvintes. Além disso, ela é muito mais frequentemente associada a uma pessoa branca. Apenas uma vez, disseram que ela era negra. A classe social a qual ela pertence não é bem

³⁸ IBGE. Pesquisa Nacional por Amostra de Domicílios Contínua, 2018, consolidado de primeiras entrevistas (<https://www.ibge.gov.br/estatisticas/sociais/populacao/25844-desigualdades-sociais-por-cor-ou-raca.html?=&t=resultados> [acessado em 18/01/2020]).

³⁹ Sigla para Fundação de Amparo à Pesquisa do Estado de São Paulo. É uma das principais agências de fomento à pesquisa científica e tecnológica do país (<http://www.fapesp.br/sobre/> [acessado em 18/01/2020]).

evidenciar que a variante é um índice de paulistanidade principalmente para quem não é originalmente da capital paulista.

Figura 26 Nuvem de palavras associadas à imitação de Adnet por ouvintes paulistanos

Figura 27 Nuvem de palavras associadas à imitação de Adnet por ouvintes não paulistanos

A fim de saber se tais noções podem ser atribuídas especificamente à variante alongada e se são percepções distintas para cada grupo (paulistanos e não paulistanos) ou são idiossincrasias dos ouvintes a depender de outras variáveis, é necessária a quantificação desses dados por meio de uma metodologia mais rígida.

3.3 O EXPERIMENTO DE PERCEPÇÃO

As noções discutidas na subseção anterior foram sistematizadas em formulários de percepção, contendo: i) dez escalas de variáveis quantitativas com graus de diferenciais semânticos de seis pontos (inteligência, amigabilidade, feminilidade, formalidade, paulistanidade, escolaridade, sucesso profissional, classe social, localização do bairro, masculinidade); ii) duas variáveis qualitativas (zona e idade do ouvinte); e iii) 23 caixas de seleção com características pessoais do falante – única parte de resposta opcional, podendo todas ou até mesmo nenhuma característica ser selecionada.

i) Dez variáveis quantitativas

Na sua opinião, essa pessoa parece:		
123456		
Nada inteligente	☐ ☐ ☐ ☐ ☐ ☐	Muito inteligente
Nada amigável	☐ ☐ ☐ ☐ ☐ ☐	Muito amigável
Nada feminina	☐ ☐ ☐ ☐ ☐ ☐	Muito feminina
Nada formal	☐ ☐ ☐ ☐ ☐ ☐	Muito formal
Nada paulistana	☐ ☐ ☐ ☐ ☐ ☐	Muito paulistana
Nada escolarizada	☐ ☐ ☐ ☐ ☐ ☐	Muito escolarizada
Mal sucedida na profissão	☐ ☐ ☐ ☐ ☐ ☐	Bem sucedida na profissão
De classe social mais baixa	☐ ☐ ☐ ☐ ☐ ☐	De classe social mais alta
De um bairro mais periférico de São Paulo	☐ ☐ ☐ ☐ ☐ ☐	De um bairro mais central de São Paulo
Nada masculina	☐ ☐ ☐ ☐ ☐ ☐	Muito masculina

ii) Duas variáveis qualitativas

1. Em que zona de São Paulo você acha que ela mora?

- ☐ Norte
- ☐ Sul
- ☐ Leste
- ☐ Oeste
- ☐ Centro

2. Você acha que essa pessoa é...

- ☐ Adolescente
- ☐ Adulta de 20 e poucos anos
- ☐ Adulta por volta dos 30 anos
- ☐ Adulta por volta dos 40 anos
- ☐ Adulta por volta dos 50 anos

iii) 23 caixas de seleção

Você acha que essa pessoa deve ser (marque quantas opções quiser):

- ☐ Alta
- ☐ Branca
- ☐ Gay/Lésbica
- ☐ Tímida
- ☐ Descolada
- ☐ Religiosa
- ☐ Articulada
- ☐ Irritante
- ☐ Metida
- ☐ Mimada
- ☐ Culta
- ☐ Tranquila
- ☐ Conservadora
- ☐ Sofisticada
- ☐ Independente
- ☐ Mauricinho/Patricinha
- ☐ Riquinha
- ☐ Confiável
- ☐ Simples
- ☐ Trabalhadora
- ☐ Caipira
- ☐ Engraçada
- ☐ Calma
- ☐ Outros: _____

Por fim, o ouvinte compartilhou as seguintes informações pessoais, interessantes para as análises estatísticas das suas respostas:

iv) Informações pessoais

Agora, para terminar, me fale um pouco sobre você.

1. Idade: _____

2. Sexo:

- ☐ Feminino
- ☐ Masculino
- ☐ Outro: _____

3. Profissão/ocupação: _____

4. Escolaridade:

- ☐ Ensino fundamental (completo ou incompleto)
- ☐ Ensino médio (completo ou incompleto)
- ☐ Ensino superior (completo ou incompleto)

5. Bairro onde mora: _____

6. Zona de São Paulo onde mora:

- ☐ Sul
- ☐ Norte
- ☐ Oeste
- ☐ Leste
- ☐ Centro

7. Local onde nasceu:

- ☐ São Paulo, capital.
- ☐ São Paulo, interior.
- ☐ Outro estado.

8. Se NÃO nasceu na CIDADE de São Paulo, em que CIDADE e ESTADO nasceu?

9. Se NÃO nasceu na CIDADE de São Paulo, há quanto tempo mora nela, aproximadamente? _____

3.4 HIPÓTESES E EXPECTATIVAS

A expectativa é que certas variáveis quantitativas e qualitativas trabalhem de modo dependente, formando um só conjunto de significados. No entanto, esses conjuntos podem mudar a depender da origem do ouvinte. Há, contudo, conjuntos comuns entre paulistanos e não paulistanos⁴⁰, tal qual mostra a **Tabela 2**, apesar de comportamentos diferentes.

Na **Hipótese 1** para o conjunto **Prestígio**, espera-se que as escalas de inteligência, formalidade, escolaridade, sucesso profissional, classe social e localização do bairro se correlacionem na forma de grandezas proporcionais. O nome “**Prestígio**” é empregado aqui no sentido de “valoração social”: respostas positivamente valoradas parecerem exprimir valoração social positiva (são socialmente “bem vistas”), ao passo que, respostas negativamente valoradas, exprimem valoração social negativa (são socialmente “mal vistas”). Além disso, a hipótese é que valores maiores para **Prestígio** sejam associados à zona oeste de

⁴⁰ Também para este experimento de percepção, *nãopaulistanos* são aqueles nascidos no interior de São Paulo e fora do estado.

São Paulo e menores, à zona sul (seu extremo), que concentram os bairros com altos e baixos Índices de Desenvolvimento Humano, respectivamente (Estadão, 2016). Por fim, relativamente à variável linguística, espera-se que a variante alongada (comparativamente à não alongada) receba valores mais positivos para este conjunto de respostas, ao menos para ouvintes não paulistanos.

Tabela 2 Hipóteses sobre os conjuntos de significados comuns aos dois grupos de ouvintes

	Hipótese 1	Hipótese 2	Hipótese 3
	Prestígio	Amigabilidade	Gênero
Variáveis Quantitativas	Inteligência Formalidade Escolaridade Sucesso profissional Classe social Localização do bairro	Amigabilidade	Feminilidade Masculinidade
Variáveis qualitativas	Zona	-	-
Caixas de seleção	Alta Branca Articulada Metida Mimada Culta Conservadora Sofisticada Maurinha/patricinha Riquinha Simples Trabalhadora	Articulada Metida Mimada Engraçada	Gay/Lésbica Mauricinho/Patricinha

Para a **Hipótese 2**, o conjunto **Amigabilidade** deve conter exclusivamente a escala de mesmo nome como variável quantitativa, mas espera-se que, quando o falante soar mais amigável, ele soe também articulado e engraçado, ao passo que, quando ele soar menos amigável, ele soe metido e mimado. Para ouvintes paulistanos, essa hipótese independe da variante linguística. No entanto, para não paulistanos, espera-se que os disfarces com a

variante alongada sejam percebidos como menos amigáveis – e, portanto, metidos e mimados – relativamente à não alongada.

O conjunto **Gênero** é composto pelas escalas feminilidade e masculinidade e as caixas de seleção gay/lésbica e maurinho/patricinha. Seguindo a tendência de estudos anteriores, como já se mencionou, espera-se que ambas as escalas se correlacionem negativamente, de forma que quão mais feminino o falante soar, menos masculino ele o seja – e o contrário é verdadeiro. Além disso, a hipótese é que os falantes homens que soem mais femininos sejam percebidos como gays e mauricinhos, enquanto as falantes mulheres que soam mais masculinas sejam percebidas como lésbicas (no mesmo sentido de trabalhos como de Levon (2007; 2014) e Mendes (2018)). Isso se aplica a ambos os grupos de ouvintes. Com relação às variantes de (AN), a partir das respostas às entrevistas metalinguísticas, espera-se que, para ouvintes não paulistanos, falantes homens sejam percebidos como mais femininos, menos masculinos e gays nos disfarces com [ẽ:], relativamente aos com [ẽ:]. Essa interpretação também é possível a partir da análise da imitação de Adnet (cf. **Introdução**). O humorista performa uma menina ao imitar o “nasalês”. Assumindo que nenhum discurso é arbitrário, mas apresenta sempre viés ideológico socialmente construído (MacDonell, 1986), pode-se dizer que há algo de feminino (e, por consequência, gay) na variante alongada. Para falantes mulheres nas respostas de ambos os grupos de ouvintes, espera-se que (AN) não tenha efeito na percepção de **Gênero**: elas devem soar mais femininas e menos masculinas (sem a opção lésbica marcada) independentemente da variante linguística.

Dado que esta pesquisa visa a contribuir para os estudos de percepção do falar paulistano, a escala de paulistanidade merece atenção especial. Com base nos comentários metalinguísticos (subseção 3.2), espera-se que os dois grupos de ouvintes se difiram também quanto à percepção dessa escala. Por

essa perspectiva, (AN) não teria efeito na percepção de quão paulistano soa o falante para ouvintes paulistanos, enquanto, para não paulistanos, os falantes seriam percebidos como mais paulistanos em seus disfarces com [ẽ:], relativamente àqueles com [ẽ] – formando, sozinha, o conjunto **Paulistanidade**, como mostra a coluna **Hipótese 4** da **Tabela 3**. É possível que, para não paulistanos, esse conjunto seja, na realidade, formado pelas escalas paulistanidade e localização do bairro, e pela variável qualitativa zona, de modo que os falantes soem mais paulistanos, de bairros mais centrais e do centro ou zona oeste quando ouvidos em [ẽ:] – como se vê na coluna **Hipótese 5** da mesma tabela. A zona oeste de São Paulo é conhecida por concentrar bares e, por isso, ser quase um “ponto turístico” da cidade, o que pode sugerir ser “mais paulistano” para quem é de fora. Por esse viés, a variante teria traço regional mais fortemente marcado – sem diminuir a “força” de seu traço de *status* social.

Tabela 3 Hipóteses sobre os conjuntos de significados para paulistanidade para ouvintes não paulistanos

	Hipótese 4	Hipótese 5
	Paulistanidade	Paulistanidade
Variáveis quantitativas	Paulistanidade	Paulistanidade Localização do Bairro
Variáveis qualitativas	-	Zona
Caixas de seleção	-	-

3.5 COLETA E PERFIL DOS OUVINTES

No total, 123 ouvintes residentes em São Paulo responderam aos formulários *online* da plataforma de pesquisas Google Forms. Usaram-se *chats* de WhatsApp, grupos de Facebook e e-mails para recrutar ouvintes por meio de um *link* que direcionava os voluntários à página do Google Forms em que se encontrava o formulário. Colegas da autora gentilmente compartilharam o *link* com conhecidos.

O formulário do Google Forms dividia-se em 5 seções: as 4 primeiras continham os clipes de áudio dos falantes com os significados sociais a serem relacionados às variantes (ver subseção 3.3) e a última, perguntas de cunho pessoal sobre o ouvinte. O respondente só poderia passar para a seção seguinte após ter respondido todas as questões obrigatórias (i.e., as escalas de variáveis quantitativas e as variáveis qualitativas). No topo da página da primeira seção, havia um pequeno texto com as devidas informações necessárias para que o ouvinte respondesse ao formulário:

Olá! Tudo bem?

Obrigado por participar dessa pesquisa de mestrado da USP sobre a cidade de São Paulo. Para participar, basta ser morador da cidade (não precisa ter nascido nela).

Por gentileza, use fone de ouvido!

Você vai ouvir quatro pessoas (uma por vez) nos clipes de áudio a seguir. Você pode ouvi-las quantas vezes quiser.

Em seguida, por favor, preencha o formulário que segue cada um dos clipes, de acordo com a sua percepção. Não há respostas certas ou erradas.

Vamos lá?

É importante dizer que o Google Forms não permite o uso de áudios, mas sim de vídeos da plataforma YouTube. Por isso, os áudios colocados no formulário apareceram, na realidade, em um vídeo com fundo preto. Ao final de cada fala, o vídeo continuava por um curto tempo (sem som nenhum) e aparecia, na tela, uma mensagem dizendo “Para ouvir novamente, clique aqui” e uma seta apontando para o botão “reiniciar”. Isso foi feito porque, no fim do vídeo, o YouTube sugere ao ouvinte outros vídeos daquele mesmo canal. Tal sugestão de vídeo aparece no meio da tela, de forma que, ao clicar ali, uma nova aba com outro vídeo do mesmo canal do YouTube aparece. Portanto, para evitar tal transtorno, tomou-se a medida de apresentar a mensagem ao fim de cada vídeo.

De modo geral, pode-se dizer que não houve dificuldade na execução da tarefa pela maioria dos ouvintes, apesar de certas hesitações. Alguns colegas da pesquisadora reportaram dificuldade em julgar determinados significados,

como o sucesso profissional do falante, com base num curto trecho de áudio. Além disso, alguns disseram (ou deram a entender) ser “complicado” e até “preconceituoso” julgar se uma pessoa soa feminina/masculina e gay/lésbica – possivelmente porque reconhecem que a comunidade LGBTQIA+⁴¹ é uma minoria já socialmente injustiçada e julgada pela sociedade. Já outros ouvintes pareciam dizer que essas características são “ruins” e, por isso, o cuidado na hora de atribuí-las a alguém.

A partir de outros comentários, observou-se que os ouvintes não tiveram dificuldades técnicas (onde clicar, como prosseguir etc.) nem de nomenclatura (entendimento da semântica dos termos) para realizarem a tarefa.

Com todas essas questões em mente, é possível afirmar que os ouvintes compreenderam bem que o objetivo da atividade não consistia em “acertar” como os falantes verdadeiramente deveriam ser no “mundo real”, mas sim em expressar suas percepções subjetivas sobre eles.

Os 123 ouvintes dividem-se em diferentes perfis sociais, conforme mostra a **Tabela 4**. Aquelas categorias que não apresentaram nenhum ouvinte não foram inseridas na tabela. Nota-se certo desequilíbrio quanto a(o): i) número de respondentes por grupo de ouvintes (73 paulistanos e 50 não paulistanos); ii) Gênero dos ouvintes (há 21 paulistanas a mais do que homens paulistanos, mas, para não paulistanos, há certo equilíbrio, com apenas 2 homens a mais do que mulheres); iii) Faixa Etária⁴² (há 52 ouvintes paulistanos na primeira faixa etária, 20 na segunda e apenas 1 na terceira; no caso de não paulistanos, há 44 ouvintes na primeira faixa etária, 6 na segunda e nenhum na terceira – ou seja, no geral, a maioria dos ouvintes concentra-se na primeira faixa etária); e iv) Nível de Escolaridade⁴³ (apenas 1 respondente com ensino fundamental compõe os

⁴¹ Sigla para Lésbicas, Gays, Bissexuais, Travestis, Transexuais, Transgêneros, Queers, Intersexos e Assexuados (cf. <http://prceu.usp.br/uspdiversidade/lgbtqia/o-que-e-lgbtqia/> [acessado em 13/04/2019])

⁴² Seguiram-se os critérios de divisão da Amostra SP2010: 1ª faixa (de 19 a 34 anos), 2ª faixa etária (de 35 a 59 anos) e 3ª faixa etária (60 anos ou mais) (cf. Oushiro, 2015).

⁴³ Os níveis são ensino fundamental (incompleto ou completo), ensino médio (incompleto ou completo) e ensino superior (incompleto ou completo).

grupos de ouvintes; a maioria tem ensino superior, graças ao método de recrutamento: divulgação em grupos de Facebook de cursos da USP e por meio da divulgação do *link* por colegas do grupo de estudo da pesquisadora). A maior parte dos ouvintes tem/está no ensino superior e na 1ª faixa etária. Sendo assim, quaisquer diferenças de percepção quanto a características sociais dos ouvintes podem mais adequadamente ser explicadas apenas pelas variáveis origem do ouvinte e/ou gênero.

Tabela 4 Distribuição dos perfis sociais dos ouvintes (N = 123)

		Paulistanos (N = 70)					Não Paulistanos (N = 53)			
		Feminino (N = 47)		Masculino (N = 26)			Feminino (N = 24)		Masculino (N = 26)	
		1ª (N = 34)	2ª (N = 13)	1ª (N = 18)	2ª (N = 7)	3ª (N = 1)	1ª (N = 19)	2ª (N = 5)	1ª (N = 25)	2ª (N = 1)
Centro (N = 14)	Ens. Superior (N = 14)	-	2	1	1	-	4	1	4	1
Leste (N = 20)	Ens. Superior (N = 20)	6	3	10	-	-	1	-	-	-
Norte (N = 14)	Ens. Superior (N = 14)	6	-	1	2	1	1	-	3	-
Oeste (N = 49)	Ens. Médio (N = 5)	1	-	1	-	-	1	1	1	-
	Ens. Superior (N = 44)	6	5	4	3	-	11	1	14	-
Sul (N = 26)	Ens. Médio (N = 2)	2	-	-	-	-	-	-	-	-
	Ens. Superior (24)	13	3	1	1	-	1	2	3	-

A variação etária dos participantes também está representada na **Figura 28**. A linha vertical vermelha indica a idade mediana para cada grupo de ouvintes: 30 anos para paulistanos e 26 para não paulistano. Ou seja, a tendência central para as distribuições numéricas das idades de ambos os grupos de ouvintes concentra-se na primeira faixa etária, como já constatado na tabela anterior.

Figura 28 Histogramas da variação etária dos ouvintes paulistanos e não paulistanos

A **Tabela 5** mostra como os ouvintes dividem-se quanto a sua origem. Porque a quantidade de ouvintes do interior de São Paulo (28) e de ouvintes de outro estado (22) não é significativamente grande, resolveu-se amalgamar os dois grupos, formando, assim, o grupo de ouvintes não paulistanos.

Tabela 5 Distribuição dos perfis sociais dos ouvintes por origem

Origem	N
São Paulo - capital	73
São Paulo - interior	28
Outro estado	22
Total	123

O próximo capítulo discute alguns resultados para as respostas dos dois grupos de ouvintes, para quais se aplicaram diversos testes de regressão por meio da plataforma de análise estatística R (R Core Team, 2019) com vistas a verificar se as hipóteses levantadas neste capítulo são corroboradas.

Capítulo 4

Análises: Percepções sociolinguísticas acerca de (AN) por ouvintes paulistanos e não paulistanos

Como se viu no capítulo anterior, 123 ouvintes participaram do experimento de percepção desenvolvido com a técnica *matched-guise*. Cada ouvinte preencheu um formulário para cada um dos quatro falantes, o que resultou num total de 492 respostas. O objetivo deste capítulo é reportar as análises dessas respostas, com destaque para os efeitos de (AN) nas percepções sociolinguísticas dos ouvintes paulistanos e não paulistanos sobre como soam os quatro falantes cujas vozes foram utilizadas no experimento.

Para tanto, realizaram-se testes estatísticos na plataforma R (R Core Team, 2019). Inicialmente, desenvolveram-se as Análises de Componentes Principais (ACP)⁴⁴ para as respostas dadas nas escalas (variáveis quantitativas), por meio da função “*principal*” do pacote “*psych*” (Revelle, 2015). Mediante cálculos de transformação ortogonal, essa função analisa as respostas para cada escala e identifica correlações entre elas. Aquelas escalas cujas respostas são interdependentes entre si, ou seja, se correlacionam, constituem um só componente (doravante CP) no conjunto total de respostas. Dessa forma, segundo Levshina (2015: 361), pode-se “indicar o mínimo possível de componentes ortogonais (não correlacionados), maximizando a variância total explicada”⁴⁵. As ACPs servem, portanto, para reduzir a dimensionalidade das variáveis.

Para este estudo, o primeiro passo no desenvolvimento das ACPs foi determinar quantos componentes mais explicam a variância nas respostas dadas pelos ouvintes. Para isso, foram analisados os *eigenvalues* gerados a partir da função “*principal*” – que são responsáveis por mostrar exatamente quanto da

⁴⁴ Do inglês *Principal Component Analysis* (PCA).

⁴⁵ Do original “The main purpose of PCAs is to find as few orthogonal (uncorrelated) components as possible while maximizing the total explained variance”.

variação total é explicada por cada CP (cf. Levshina, 2015: 354). Segundo o critério de Kaiser (idem.: 355), adotado neste trabalho, apenas CPs com *eigenvalue* acima de 1 devem ser considerados.

A **Tabela 6** a seguir mostra os CPs da primeira análise deste estudo e seus respectivos *eigenvalues*. Nota-se que nem todos os CPs serão considerados, visto que somente os três primeiros têm valor acima de 1 tanto para ouvintes paulistanos (4,23, 1,79 e 1,28) quanto para os não paulistanos (3,76, 1,85 e 1,24) – para CP1, CP2 e CP3, respectivamente.

Tabela 6 *Eigenvalues* dos Componentes Principais gerados das respostas dadas por ouvintes paulistanos e não paulistanos nas 10 escalas – sem rotação *Promax*

	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8	CP9	CP10
<i>Eigenvalues</i>										
Paulistanos	4,23	1,79	1,28	0,80	0,60	0,41	0,38	0,22	0,19	0,10
<i>Eigenvalues</i>										
Não Paulistanos	3,76	1,85	1,24	0,94	0,72	0,44	0,37	0,35	0,18	0,15

Pode-se também inspecionar esses resultados visualmente, por meio dos gráficos⁴⁶ (Cattell, 1966) da **Figura 29** abaixo. O eixo *x* corresponde ao número de componentes a serem extraídos⁴⁷ e o eixo *y*, aos *eigenvalues* gerados. O objetivo do gráfico é encontrar o ponto (comumente chamado de *cotovelo*) em que, posteriormente a ele, a adição de novos componentes não explica significativamente a variância das respostas (cf. Levshina, 2015: 319). Dessa forma, a partir do cotovelo, a variância apresenta uma tendência descendente linear (Reise et al, 2000). Como se vê na figura, para ambos os conjuntos de dados (respostas de ouvintes paulistanos e não paulistanos), o cotovelo está no ponto 3, pois, a partir dele, os *eigenvalues* são menores que 1.

⁴⁶ Em inglês *Cattell's screeplots*.

⁴⁷ É relevante dizer que não se espera que a quantidade de CPs extraídos seja a mesma quantidade de escalas analisadas. Isso não faria sentido, pois significaria dizer que cada escala funciona de modo independente e, como se viu no **Capítulo 3** desta dissertação, espera-se que haja correlação entre respostas dadas a escalas diferentes. Esta medida apenas segue a tendência de estudos anteriores para começar as ACPs (Mendes, 2019; Santos, 2020).

Figura 29 Gráficos de Cattell das Análises de Componentes Principais das respostas dadas por paulistanos e não paulistanos nas 10 escalas

Os resultados obtidos nesse primeiro passo das ACPs evidenciam que as dez escalas deste estudo (inteligência, amigabilidade, feminilidade, formalidade, paulistanidade, escolaridade, sucesso profissional, classe social, localização do bairro e masculinidade) podem, na realidade, ser reduzidas a três componentes, tanto para as respostas de ouvintes paulistanos quanto de não paulistanos.

Após determinar a quantidade de CPs, o próximo passo trata de verificar o grau de correlação entre as escalas que compõem cada um deles, bem como definir, com mais acurácia, quais explicam mais a variância nas respostas⁴⁸. Isso é feito por meio da rotação "*Promax*" (cf. Levshina, 2015: 363) como argumento na função "*principal*", e o resultado está representado na **Tabela 7** a seguir. Para entender a tabela, considera-se, neste estudo, $|0,60|$ ⁴⁹ como índice-limite para considerar que certa escala integre um CP. Além disso, o primeiro CP é aquele

⁴⁸ A função "*principal*" já indica quais CPs mais explicam a variância nas respostas; no entanto, segundo Levshina (2015), a rotação "*Promax*" o faz com mais acurácia mediante seus cálculos.

⁴⁹ Lê-se: zero vírgula sessenta absoluto. Ou seja, o valor numérico independente do seu sinal (negativo ou positivo). Dessa forma, escalas com valores menores que -0,60 ou maiores que 0,60 foram considerados como constituintes de um CP.

que mais explica a variância nas respostas (verificar *eigenvalues*), seguido do segundo e, por fim, do terceiro.

Tabela 7 Componentes Principais gerados pelas correlações entre as respostas nas 10 escalas para ouvintes paulistanos e não paulistanos – com rotação *Promax*

	Paulistanos			Não Paulistanos		
	Prestígio CP1	Gênero CP2	Amigabilidade CP3	Status social CP1	Profissionalismo CP2	Gênero CP3
Inteligência	0,51	0,03	0,43	0,00	0,81	-0,11
Amigabilidade	-0,24	0,08	0,95	-0,40	0,79	0,24
Feminilidade	0,02	0,96	0,12	0,06	0,13	-0,84
Formalidade	0,77	0,11	-0,23	0,33	0,17	-0,38
Paulistanidade	0,15	-0,10	0,59	0,49	0,11	0,67
Escolaridade	0,83	-0,41	0,17	0,44	0,52	-0,04
Sucesso profissional	0,74	-0,07	0,21	0,11	0,77	-0,12
Classe social	0,91	0,02	-0,11	0,92	-0,02	0,00
Localização do bairro	0,91	-0,03	-0,24	0,91	-0,11	-0,01
Masculinidade	0,01	-0,97	0,03	-0,06	0,01	0,89
<i>Eigenvalue</i>	3,80	1,90	1,61	2,42	2,25	2,19
%Variância	38	19	16	24	23	22
%Cumulativa	38	57	73	24	47	69

Na **Tabela 7**, logo se vê que, apesar da quantidade de CPs ser a mesma para os dois grupos de ouvintes, as escalas que compõem cada um são diferentes. Por exemplo, o CP1 para ouvintes paulistanos (que explica 38% da variância nas respostas) é composto pelas escalas formalidade (índice = 0,77), escolaridade (0,83), sucesso profissional (0,74), classe social (0,91) e localização do bairro (0,91), positivamente correlacionadas – inteligência participa marginalmente (0,51). Por outro lado, para ouvintes não paulistanos, o CP1 (que explica 24% da variância nas respostas) é integrado apenas pelas escalas classe social (0,92) e localização do bairro (0,91), também correlacionadas positivamente. De fato, o interesse em realizar ACPs para paulistanos e não paulistanos separadamente está em verificar a correlação (ou a falta dela) nas respostas às escalas e, por consequência, se há semelhanças ou diferenças nos CPs de cada grupo.

As escalas que compõem o CP1 para ouvintes paulistanos e o modo como se correlacionam podem ser entendidas como componentes do conceito social de Prestígio, no sentido de “valoração social”: respostas positivamente valoradas parecerem exprimir valoração social positiva, ao passo que, respostas negativamente valoradas, exprimem valoração social negativa. Ou seja, parece “bom”, do ponto de vista da valoração social, ser mais formal, mais escolarizado, mais bem sucedido profissionalmente, de classe social mais alta e de um bairro mais central da cidade de São Paulo. Ao mesmo tempo em que o oposto disso é “ruim”. Por outro lado, para o CP1 dos não paulistanos, a correlação positiva entre classe social e localização do bairro parece exprimir a ideia de *Status* social, no sentido de condição socioeconômica mais ou menos favorecida. Pessoas parecem ter uma condição socioeconômica mais favorável se soam como pertencentes a uma classe social mais alta e como moradores de uma área mais central da cidade, e menos favorável se soam como pessoas que pertencem a uma classe social mais baixa e como moradores de uma área mais periférica.

O resultado para o CP2 nas respostas dos paulistanos converge com resultados de estudos anteriores (Levon 2007, 2014; Mendes, 2019) sobre as escalas de feminilidade (neste estudo, índice = 0,96) e masculinidade (-0,97), isto é, elas se correlacionam negativamente. Esse CP é, aqui também, chamado Gênero, e explica 19% da variância nas respostas obtidas neste estudo. Para não paulistanos, o CP2, que explica 23% da variância, é constituído pelas escalas inteligência (0,81), amigabilidade (0,79) e sucesso profissional (0,77) em correlação positiva – escolaridade participa marginalmente (0,52). Seguindo estudos anteriores sobre essa correlação entre tais escalas (Campbell-Kibler, 2007, 2009; Levon & Fox, 2014; *inter alia*), deu-se a esse CP o nome de Profissionalismo.

O terceiro componente nas repostas dos paulistanos é formado exclusivamente pela escala amigabilidade (0,95), que dá nome ao CP –

paulistanidade participa periféricamente (0,59). Sozinho, este CP é responsável por explicar 16% da variância nas respostas dos paulistanos e, conjuntamente aos CPs anteriores, explica 73%. Para os ouvintes não paulistanos, o terceiro CP, que explica 22% da variância nas respostas, é constituído pelas escalas feminilidade (-0,84), paulistanidade (0,67) e masculinidade (0,89), sendo que as duas últimas se correlacionam positivamente entre si e negativamente com a primeira. Juntos, os três CPs dos não paulistanos explicam 69% da variância.

Na comparação entre os CPs dos dois grupos de ouvintes, percebe-se que, além de diferentes escalas constituírem diferentes componentes, paulistanidade (foco deste estudo) integra um dos CPs dos não paulistanos, ao passo que, para paulistanos, essa escala somente participa marginalmente do terceiro componente. Além disso, formalidade e escolaridade são escalas que integram o CP1 de paulistanos, mas não participam de nenhum dos CPs dos não paulistanos. Já inteligência, que participa de CP2 dos não paulistanos, não integra nenhum CP dos paulistanos.

Com base nessas diferenças entre os CPs identificados nas respostas de paulistanos e de não paulistanos, pode-se afirmar que os grupos de ouvintes percebem (AN) de maneira diferente, visto que todos ouviram os mesmos disfarces e que somente a variante linguística diferenciava um disfarce do seu par. As questões a serem exploradas, a partir de agora, são: i) (AN) tem efeito nas respostas dos ouvintes (nos três CPs)?; ii) se sim, como ela funciona para os dois grupos de ouvintes?; iii) se não, quais outras variáveis explicam os CPs?; e iv) há efeito interativo da variável linguística com outras variáveis nas respostas dadas pelos ouvintes?

Para responder a essas questões, realizaram-se modelos de regressão no programa R. Tais modelos têm por objetivo averiguar as possíveis relações entre uma variável dependente e variáveis independentes. Como Análises de Componentes Principais consideram as correlações entre escalas, a variável dependente de cada modelo não é uma escala em si, mas um Componente

Principal. Portanto, é nos CPs que, hipoteticamente, têm efeito as demais variáveis (as independentes) do experimento: variável linguística (AN), falante, o sexo do ouvinte, sua escolaridade, idade e zona da cidade de São Paulo onde mora.

A seguir, apresentam-se as análises de regressão feitas para todos os três Componentes Principais identificados nas respostas de cada grupo de ouvintes. Em seguida, apresentam-se os resultados para as variáveis qualitativas e, por fim, os das caixas de seleção – também para cada grupo de ouvintes separadamente.

4.1 CP1 – PRESTÍGIO E *STATUS* SOCIAL

Esta subseção apresenta a análise do primeiro componente principal tanto para ouvintes paulistanos quanto para não paulistanos. Como se viu anteriormente, para paulistanos, o CP1 é chamado Prestígio e é composto pelas respostas nas escalas formalidade, escolaridade, sucesso profissional, classe social e localização do bairro, positivamente correlacionadas. Para não paulistanos, o CP1, intitulado *Status* social, é formado apenas pelas escalas classe social e localização do bairro, também em correlação positiva.

A **Figura 30** apresenta os gráficos de dispersão⁵⁰ das respostas dos dois grupos de ouvintes para seus respectivos CP1 por Variante. Note-se que ouvintes paulistanos e não paulistanos perceberam os falantes (em conjunto) como mais prestigiosos e de maior *status* social, respectivamente, quando os ouviram com a variante alongada (caixa azulada), relativamente à não alongada (caixa alaranjada). Observe-se que os valores no eixo y são *scores* gerados pela ACP para os componentes – e não os valores 1 a 6 das escalas originais.

Interessa agora saber se essa diferença entre a percepção de [ẽ] e [ẽ:] no CP1 é significativa para os dois grupos de ouvintes. Pela visualização dos

⁵⁰ Este gráfico de dispersão em específico é tecnicamente chamado de *boxplot* ou, em tradução própria, *gráfico de caixas*.

gráficos, espera-se que, para os ouvintes paulistanos, a diferença não seja significativa, pois a concentração de respostas (nesse caso, a *mediana* que é representada pelo traço mais escuro dentro das caixas) para ambas as variantes estão bem próximas. Para não paulistanos, no entanto, vê-se que a mediana para respostas baseadas em [ẽ:] está bem acima da mediana de [ẽ] e, por isso, espera-se um possível efeito da Variante.

Figura 30 Distribuição das respostas dos ouvintes paulistanos para o CP Prestígio e dos ouvintes não paulistanos para o CP *Status* social, por Variante

Para checar tais expectativas, fez-se um modelo de regressão linear simples de efeitos mistos. É *simples* por testar o efeito de apenas uma variável independente (neste caso, a Variante) sobre uma dependente⁵¹, e de *efeitos mistos* por considerar o Ouvinte como efeito aleatório, ou seja, porque caso novos ouvintes fossem recrutados aleatoriamente, provavelmente eles não seriam os mesmos da amostra anterior (cf. Oushiro, 2017: 174). A **Tabela 8** a seguir mostra os resultados desse modelo de regressão. Para entender essa tabela e outras semelhantes que seguem, é importante saber quais são e como funcionam suas cinco colunas: Efeitos fixos, Estimativa, Erro padrão, Valor-*t* e Valor-*p*.

⁵¹Em contraposição a um modelo de regressão linear *múltiplo*, que testa o efeito de múltiplas variáveis independentes sobre uma variável dependente.

Tabela 8 Resumo dos resultados do modelo de regressão linear simples para o CP1 de paulistanos (Prestígio) e não paulistanos (*Status* social) (Efeitos Mistos)

Ouvintes	Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Paulistanos	<i>Intercept</i>	-0,058	0,094	-0,615	0,540
	[ẽ:]	0,116	0,097	1,191	0,235
Não paulistanos	<i>Intercept</i>	-0,207	0,099	-2,088	0,038 *
	[ẽ:]	0,414	0,136	3,028	0,002 **

Fórmula: CP1 ~ Variante + (1|Ouvinte), data=DATA

Intercept= [ẽ]

Em Efeitos fixos, têm-se as variantes das variáveis previsoras (ou independentes) incluídas no modelo. O *Intercept* é o conjunto de variantes tomado como referência e é definido, seguindo o *default*⁵² no R, por ordem alfabética e ordem numérica crescente. É importante dizer que o R tem dificuldades em ler diacríticos. Dessa forma, para este estudo, ao invés de representar as variantes como [ẽ] e [ẽ:] no programa, optou-se por representá-las como A e AN, respectivamente. Para facilitar a leitura dos gráficos e tabelas, no entanto, retomou-se à representação original de tais variantes.

Considerando-se tal mudança de representação no R, sabe-se que, na ordem alfabética, A vem antes de AN. Dessa forma o *Intercept* é a variante A (ou [ẽ]). O valor de sua Estimativa na tabela é o *valor de referência*. Isso significa que o outro valor da coluna Estimativa (referente à outra variante, i.e., à variante AN/[ẽ:]) deve ser comparado a ele. Mais precisamente, os valores da Estimativa indicam a diferença entre o valor do *Intercepte* o da outra variante. Portanto, para ouvintes paulistanos, a diferença entre o *Intercept*([ẽ]) e [ẽ:] é de 0,116. Do ponto de vista estatístico, essa diferença não é significativa, como mostra o valor da coluna “Valor-p”.

Também conhecido como “valor de significância”, o Valor-p indica a probabilidade de se obter os mesmos resultados (em termos de distribuição das respostas às variantes das variáveis em foco) caso a hipótese nula fosse verdadeira (ou seja, caso não houvesse correlação entre cada uma das variáveis independentes e a variável dependente). Nas Ciências Humanas,

⁵² O pesquisador pode não usar o *default* e definir outra configuração para o *Intercept* a depender dos seus interesses.

convencionou-se limitar tal valor a 5% (em valores decimais, como nas tabelas: 0,05). O asterisco (*) simboliza os níveis de significância, de forma que p entre 0,01 e 0,05 é representado por um *, p entre 0,001 e 0,01 por dois **, e $p < 0,001$ por três ***. O ponto final simboliza um valor pouco acima de 0,05.⁵³

Como o Valor- p de [ẽ:] é 0,235 (ou 23,5%), é alta a probabilidade da hipótese nula ser verdadeira, ou seja, de não haver correlação entre a variante e CP1 para as respostas de paulistanos. Em outras palavras, o efeito da variável linguística (AN) não é significativo na percepção de Prestígio para os ouvintes paulistanos. Para os não paulistanos, por outro lado, a diferença entre a Estimativa do *Intercept* e de [ẽ:] é de 0,414 e ela é significativa do ponto de vista estatístico, pois $p < 0,05$ (mais precisamente, p está entre 0,001 e 0,01). Em outras palavras, para não paulistanos, os falantes (em conjunto) são percebidos como significativamente de maior *status* social quando ouvidos no seu disfarce com [ẽ:], relativamente à [ẽ]. Aqui, já se confirma, do ponto de vista da percepção, os comentários feitos pelos paulistanos e não paulistanos nas entrevistas metalinguísticas (cf. **Capítulo 3** desta dissertação): enquanto aqueles, apesar de tecerem comentários relacionados à situação socioeconômica dos falantes, não discorrem acerca da variante alongada, estes não só tecem ainda mais comentários nesse sentido, como mencionam a pronúncia [ẽ:], especificamente.

Nas tabelas, há também a coluna de Erro Padrão que, segundo Oushiro (2017: 135), é responsável por medir a precisão das previsões, de modo que quanto menor seu valor, mais preciso é o modelo. Na quarta coluna, tem-se o Valor- t , correspondente à divisão entre estimativa e erro padrão que servirá de base para calcular o resultado do Valor- p .

Após verificar somente o efeito da Variante, apurou-se se outras variáveis (sociais) poderiam ter efeito nas respostas que compõem o CP1. Modelos de regressão linear múltipla de efeitos mistos foram testados,

⁵³ Cf. Oushiro, 2017: 136.

tomando-se como variáveis independentes: Variante, Falante, Gênero do ouvinte, sua Idade, Escolaridade e Zona⁵⁴ onde mora.

A **Tabela 9** mostra as respostas para ouvintes paulistanos. O *Intercept* é a falante Bruna com a variante [ẽ] para as respostas de ouvintes mulheres, de idades mais baixas, com ensino médio e moradoras do centro da cidade de São Paulo. Somente as variáveis Falante e Zona têm efeito significativo na percepção de Prestígio. A diferença entre o valor de referência (Estimativa = -0,299) e Gisele (-0,961), e Ricardo (-0,487), e Rodolfo (-0,990) é significativa, pois $p < 0,001$ em todos os casos. Os resultados mostram que, em relação a Bruna, todos os falantes são percebidos como de menor prestígio. Para Zona, a diferença para ouvintes da zona sul (0,952) é significativa ($p < 0,05$). Isso significa que ouvintes da zona sul tendem a perceber Bruna como de maior prestígio do que ouvintes do centro (que estão no *Intercept*). Os resultados para as duas variáveis selecionadas podem ser visualizados, respectivamente, nas **Figura 31** e **32** a seguir. Por fim, Variante (tal como no modelo anterior), Gênero, Idade e Escolaridade não têm efeito significativo na percepção de Prestígio para paulistanos, pois, em todos esses casos, $p > 0,05$.

Tabela 9 Resumo dos resultados do modelo de regressão linear múltipla para o CP1 de paulistanos (Prestígio)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Intercept	-0,299	0,611	-0,489	0,626
Gisele	-0,961	0,115	-8,347	<0,001 ***
Ricardo	-0,487	0,115	-4,222	<0,001 ***
Rodolfo	-0,990	0,115	-8,577	<0,001 ***
[ẽ:]	0,091	0,081	1,120	0,264
Masculino	-0,170	0,178	-0,957	0,342
Mais velhos	0,016	0,010	1,658	0,102
Leste	0,789	0,412	1,916	0,059 .
Norte	0,501	0,430	1,166	0,247
Oeste	0,735	0,393	1,869	0,066 .
Sul	0,952	0,417	2,285	0,025 *
Ensino superior	-0,319	0,372	-0,858	0,394

Fórmula: $\text{Imer}(\text{CP1} \sim \text{Falante} + \text{Variante} + \text{Sexo} + \text{Idade} + \text{Zona} + \text{Escolaridade} + (1 | \text{Ouvinte}), \text{data}=\text{DATA})$
Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

⁵⁴ Nas tabelas, as variantes de Variante são [ẽ] e [ẽ:], de Falante são Bruna, Gisele, Ricardo e Rodolfo, de Sexo são Feminino e Masculino, de Idade são Menor e Maior, de Zona são Centro, Leste, Norte, Oeste e Sul e, finalmente, de Escolaridade são Ensino médio e Ensino superior.

Figura 31 Distribuição das respostas dos ouvintes paulistanos para o CP1 (Prestígio) por Falante

Figura 32 Distribuição das respostas dos ouvintes paulistanos para o CP1 (Prestígio) por Zona

Após verificar o efeito das variáveis isoladamente, averiguou-se se há efeito interativo entre as variáveis sociais e a linguística (que é o maior foco deste estudo). Os resultados mostram que não há efeito interativo entre Variante e nenhuma outra variável - é relevante dizer que este e outros resultados não apresentados no corpo do capítulo encontram-se nos **Apêndices** desta dissertação.

As respostas dos ouvintes não paulistanos nos modelos de regressão linear múltipla estão na **Tabela 10**. Note-se que Falante, Variante (tal como na **Tabela 8**) e Zona têm efeito na percepção de *Status* social. Todos os ouvintes são significativamente percebidos como de menor *status* social relativamente a Bruna. Ouvintes da zona leste percebem Bruna como de maior *status* social relativamente aos ouvintes do centro.

Tabela 10 Resumo dos resultados do modelo de regressão linear múltipla para o CP1 de não paulistanos (*Status* social)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Intercept	0,493	0,488	1,009	0,317
Gisele	-1,055	0,159	-6,599	<0,001 ***
Ricardo	-1,057	0,159	-6,612	<0,001 ***
Rodolfo	-1,200	0,159	-7,504	<0,001 ***
[ẽ:]	0,414	0,113	3,668	<0,001 ***
Masculino	0,210	0,141	1,491	0,143
Mais velhos	-0,004	0,010	-0,423	0,674
Leste	1,520	0,515	2,949	0,005 **
Norte	0,282	0,290	0,972	0,336
Oeste	0,091	0,181	0,505	0,616
Sul	0,284	0,253	1,125	0,267
Ensino superior	-0,004	0,299	-0,014	0,989

Fórmula: lmer(CP1 ~ Falante + Variante + Sexo + Idade + Zona + Escolaridade + (1|Ouvinte), data=DATA)

Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

No modelo que inclui interações⁵⁵, como mostra a **Tabela 11** abaixo, apenas Falante interage com Variante: Gisele[ẽ:] é significativamente percebida como de maior *status* social que Bruna[ẽ:]. No entanto, Zona não interage com Variante. Ademais, é interessante notar como a interação entre Gênero do ouvinte e Variante (na tabela, “[ẽ:]Masculino”) é marginalmente significativa, de forma que ouvintes homens tendem a perceber Bruna como de maior *status* social quando ouvida com [ẽ:] relativamente à [ẽ] e a ouvintes mulheres. Por fim, nesse modelo de interação, que é mais complexo, Variante não se mostrou significativa para a percepção de *Status* social – o que não anula o fato de ela o ser nos modelos anteriores, menos complexos.

⁵⁵ A interação é simbolizada pelo asterisco (*) nas fórmulas e pelos dois-pontos (:) nas tabelas.

Tabela 11 Resumo dos resultados do modelo de interação
para o CP1 de não paulistanos (*Status social*)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Intercept	0.362	0.628	0.576	0.565
Gisele	-1.412	0.222	-6.345	<0,001 ***
Ricardo	-1.355	0.260	-5.211	<0,001 ***
Rodolfo	-1.459	0.260	-5.613	<0,001 ***
[ẽ:]	0.753	0.826	0.911	0.363
Masculino	-0.015	0.183	-0.084	0.933
Mais velhos	0.005	0.013	0.380	0.704
Leste	1.739	0.692	2.512	0.013 *
Norte	0.213	0.378	0.565	0.573
Oeste	0.246	0.253	0.974	0.332
Sul	0.145	0.339	0.430	0.668
Ensino superior	0.150	0.389	0.386	0.700
[ẽ:]Gisele	0.714	0.314	2.270	0.024 *
[ẽ:]Ricardo	0.414	0.392	1.057	0.293
[ẽ:]Rodolfo	0.338	0.392	0.863	0.390
[ẽ:]Masculino	0.451	0.229	1.970	0.050 .
[ẽ:] Mais velhos	-0.018	0.017	-1.098	0.274
[ẽ:]Leste	-0.420	0.866	-0.486	0.628
[ẽ:]Norte	0.135	0.473	0.286	0.775
[ẽ:]Oeste	-0.301	0.316	-0.952	0.343
[ẽ:]Sul	0.286	0.424	0.673	0.501
[ẽ:]Ensino superior	-0.308	0.487	-0.632	0.528

Fórmula: CP1 ~ Falante * Variante + Sexo * Variante + Idade * Variante + Zona * Variante + Escolaridade * Variante + (1|Ouvinte), data=DATA

Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

O resultado para Variante:Falante pode ser visualizado na **Figura 33**. É relevante frisar que o modelo com interações revela que as variáveis (neste caso, Variante e Falante) funcionam de forma interdependente. De fato, as medianas nas respostas baseadas nas duas variantes de Bruna estão muito próximas uma da outra, o que explica Variante não ter efeito na percepção de CP1 para ela. Para Gisele, as medianas estão distantes e essa distância (diferença) é significativa, como se viu na **Tabela11**. Para Ricardo e Rodolfo, apesar de certa distância entre as medianas de suas variantes, a diferença não é significativa do ponto de vista estatístico.

Figura 33 Distribuição das respostas dos ouvintes não paulistanos para o CP1 (*Status* social) por Falante e Variante

De modo geral, viu-se que os primeiros componentes nas respostas dadas pelos dois grupos de ouvintes funcionam de maneira diferente. Para paulistanos, a Variante não tem efeito na percepção de Prestígio, ao passo que, para não paulistanos, ela tem efeito na percepção de *Status* social. Além disso, para este último grupo, a variável linguística interage com Falante, de forma que Gisele é significativamente percebida como de maior *status* social no seu disfarce com [ẽ:], diferentemente de Bruna. Ademais, Zona tem efeito na percepção de CP1 para não paulistanos, apesar de não interagir com Variante.

Dado o objetivo deste trabalho, destaca-se a evidência de que a origem dos ouvintes tem efeito na percepção sociolinguística de (AN) para os componentes aqui analisados (com suas respectivas escalas). Espera-se que o mesmo seja evidenciado para os resultados das respostas do segundo componente principal, apresentados na subseção seguinte. Novamente, a expectativa é que a variável linguística tenha efeito nas respostas de ouvintes não paulistanos, mas não nas de ouvintes paulistanos.

4.2 CP2 – GÊNERO E PROFISSIONALISMO

Nesta subseção, apresentam-se as análises do segundo componente principal para ouvintes paulistanos e não paulistanos. Para os ouvintes não paulistanos, o CP2 é chamado Profissionalismo e é composto pelas escalas inteligência, amigabilidade e sucesso profissional em correlação positiva. Para os paulistanos, por outro lado, o CP2 é intitulado Gênero e é formado pelas escalas feminilidade e masculinidade em correlação negativa.

Da mesma forma que para o CP1, apresentam-se primeiramente, na **Figura 34**, os gráficos de dispersão das respostas dos dois grupos de ouvintes para seus CP2 por Variante. Para os paulistanos, a dispersão nas respostas baseadas em ambas as variantes para a percepção de Gênero parece mais equilibrada. Já para os não paulistanos, os falantes parecem ser percebidos como mais profissionais quando ouvidos com a variante alongada, relativamente à não alongada. A fim de verificar se essa diferença é significativa do ponto de vista estatístico, fez-se um modelo de regressão linear simples de efeitos mistos, cujos resultados são resumidos na **Tabela 12**.

Figura 34 Distribuição das respostas dos ouvintes paulistanos para o CP Gênero e dos ouvintes não paulistanos para o CP Profissionalismo, por Variante

Tabela 12 Resumo dos resultados do modelo de regressão linear simples para o CP2 de paulistanos (Gênero) e não paulistanos (Profissionalismo) (Efeitos Mistos)

Ouvintes	Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Paulistanos	<i>Intercept</i>	-0,064	0,082	-0,775	0,439
	Variante[ẽ:]	0,128	0,117	1,097	0,274
Não paulistanos	<i>Intercept</i>	-0,157	0,113	-1,382	0,170
	Variante[ẽ:]	0,314	0,115	2,718	0,007**

Fórmula: CP2 ~ Variante + (1|Ouvinte), data=DATA

Intercept= Variante[ẽ]

Mais uma vez, a Variante não tem efeito na percepção de Gênero nas respostas dos ouvintes paulistanos, ao passo que tem efeito na percepção de Profissionalismo nas respostas dos não paulistanos. O conjunto de falantes é percebido como significativamente mais profissional quando ouvidos com a variante alongada, em comparação à não alongada.

Os possíveis efeitos das outras variáveis (Falante, Gênero, Idade, Zona e Escolaridade do ouvinte) foram testados e os resultados para paulistanos são apresentados na **Tabela13** a seguir.

Tabela 13 Resumo dos resultados do modelo de regressão linear múltipla para o CP2 de paulistanos (Gênero)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
<i>Intercept</i>	1,130	0,208	5,422	<0,001 ***
Gisele	-0,513	0,074	-6,858	<0,001 ***
Ricardo	-1,847	0,074	-24,634	<0,001 ***
Rodolfo	-2,142	0,074	-28,578	<0,001 ***
[ẽ:]	-0,038	0,053	-0,722	0,470
Masculino	0,118	0,059	1,993	0,047 *
Mais velhos	0,000	0,003	0,293	0,769
Leste	-0,025	0,136	-0,186	0,852
Norte	-0,012	0,142	-0,857	0,392
Oeste	-0,007	0,130	-0,609	0,542
Sul	-0,002	0,138	-0,157	0,875
Ensino superior	-0,004	0,123	-0,038	0,970

Fórmula: lmer(CP2 ~ Falante + Variante + Sexo + Idade + Zona + Escolaridade + (1|Ouvinte), data=DATA)

Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

Para os ouvintes paulistanos, a variável Falante tem efeito na percepção de Gênero, de modo que Gisele, Ricardo e Rodolfo soam menos femininos (e mais masculinos) do que Bruna. O Gênero do ouvinte também tem efeito na percepção: ouvintes do sexo masculino tendem a perceber a perceber os falantes

como mais femininos do que as ouvintes mulheres. Os resultados para ambas as variáveis podem ser visualizados nas **Figura 35** e **36** a seguir, respectivamente. As outras variáveis não têm efeito significativo.

Figura 35 Distribuição das respostas dos ouvintes paulistanos para o CP2 (Gênero) por Falante

Figura 36 Distribuição das respostas dos ouvintes paulistanos para o CP2 (Gênero) por Gênero

O passo seguinte foi verificar se as variáveis selecionadas interagem com a Variante, e os resultados mostram que não há interação.

Também foram analisadas as respostas do CP2 (Profissionalismo) para os ouvintes não paulistanos no modelo de regressão linear múltipla, como mostra a **Tabela14**. Nota-se que a única variável que não tem efeito nas respostas é Gênero do ouvinte. Para Falante, Gisele, Ricardo e Rafael são percebidos como significativamente menos profissionais do que Bruna. No tocante à Variante, [ẽ:] soa mais profissional do que [ẽ]. Quanto à Idade, ouvintes mais velhos percebem os falantes como menos profissionais do que os mais novos. Para Zona, ouvintes da zona leste, norte e sul tendem a perceber os falantes como mais profissionais do que os ouvintes do centro. Por fim, ouvintes com ensino superior em geral percebem os falantes como menos profissionais do que os ouvintes que têm nível médio de escolaridade. Os efeitos dessas variáveis nas respostas dos ouvintes podem ser visualizados na **Figura37**. No modelo de interação, verificou-se que nenhuma variável interage com Variante.

Tabela 14 Resumo dos resultados do modelo de regressão linear múltipla para o CP2 de não paulistanos (Profissionalismo)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Intercept	1,770	0,581	3,044	0,003 **
Gisele	-0,422	0,155	-2,712	0,007 **
Ricardo	-0,627	0,155	-4,030	<0,001 ***
Rodolfo	-0,518	0,155	-3,327	0,001 **
[ẽ:]	0,314	0,110	2,858	0,004 **
Masculino	-0,131	0,169	-0,774	0,443
Mais velhos	-0,031	0,012	-2,584	0,013 *
Leste	1,336	0,619	2,157	0,036 *
Norte	0,965	0,349	2,762	0,008 **
Oeste	0,214	0,218	0,981	0,332
Sul	0,749	0,304	2,463	0,017 *
Ensino superior	-0,974	0,360	-2,705	0,009 **

Fórmula: lmer(CP2 ~ Falante + Variante + Sexo + Idade + Zona + Escolaridade + (1 | Ouvinte), data=DATA)

Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

Figura 37 Distribuição das respostas dos ouvintes não paulistanos para o CP2 (Profissionalismo) por Falante, Variante, Gênero, Idade, Zona e Escolaridade

Como se viu até agora, os resultados das análises dos CPs evidenciam que a origem dos ouvintes (se são ou não da cidade de São Paulo) tem correlação com sua percepção acerca dos significados sociais de (AN). A variável linguística propriamente tem efeito nas respostas dos não paulistanos, mas não nas dos paulistanos. Desse modo, evidencia-se que a variante [ẽ:] está abaixo do nível de consciência (Labov, 2001) dos ouvintes paulistanos, mas acima do nível de consciência para não paulistanos.

Para o terceiro CP, a expectativa é que esse resultado se mantenha, principalmente porque a escala de paulistanidade compõe CP3 dos não paulistanos e não o faz na dos paulistanos. O objetivo da subseção seguinte é verificar se a variante alongada pode também ser um índice de paulistanidade para quem é de fora da capital paulista, ou ainda, em termos labovianos (Labov, 2008 [1972]) se seria um *estereótipo* da fala paulistana para quem é de fora.

4.3 CP3 – AMIGABILIDADE E GÊNERO

Para os paulistanos, o CP3 foi intitulado Amigabilidade, já que é constituído unicamente pela escala de mesmo nome. Para não paulistanos, o terceiro CP foi chamado “Gênero”, formado pelas escalas feminilidade, paulistanidade e masculinidade, sendo que as duas últimas se correlacionam positivamente entre si e negativamente com a primeira.

Seguindo os mesmos passos anteriores, apresentam-se inicialmente, na **Figura38**, os gráficos de dispersão das respostas dos dois grupos de ouvintes para seus respectivos CP3, de acordo com a Variante. As respostas dos paulistanos para Amigabilidade não parecem ter correlação com a variável linguística (AN), visto que as concentrações de respostas (as medianas) baseadas em ambas as variantes estão bem próximas uma da outra. Para não paulistanos, por sua vez, a distância entre as medianas é maior, o que parece evidenciar que há efeito da Variante na percepção de Gênero para as respostas desse grupo. Para verificar esse prognóstico, fizeram-se testes de regressão linear simples, cujos resultados são resumidos na **Tabela15**.

Figura 38 Distribuição das respostas dos ouvintes paulistanos para o CP Amigabilidade e dos ouvintes não paulistanos para o CP Gênero, por Variante

Tabela 15 Resumo dos resultados do modelo de regressão linear simples para o CP3 de paulistanos (Amigabilidade) e não paulistanos (Gênero) (Efeitos Mistos)

Ouvintes	Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Paulistanos	<i>Intercept</i>	0,013	0,098	0,134	0,894
	[ẽ:]	-0,026	0,091	-0,287	0,774
Não paulistanos	<i>Intercept</i>	-0,183	0,098	-1,858	0,064 .
	[ẽ:]	0,366	0,139	2,628	0,009 **

Fórmula: CP3 ~ Variante + (1|Ouvinte), data=DATA

Intercept= [ẽ]

Mais uma vez, (AN) não tem efeito nas respostas dos paulistanos (Amigabilidade) – como já se esperava, com base nos discursos metalinguísticos dos ouvintes paulistanos das entrevistas abertas (cf. **Capítulo 3**) e nas análises dos CPs anteriores. Por outro lado, para os não paulistanos, quando ouvidos no disfarce com a variante alongada, o conjunto de falantes é percebido como mais masculino/paulistano e menos feminino. Como se viu no **Capítulo 3**, esperava-se que a variável linguística tivesse efeito na percepção de todos os componentes do grupo de ouvintes não paulistanos, mas era inesperado o fato da escala de paulistanidade se correlacionar às escalas de feminidade e masculinidade.

Uma possível interpretação para essa correlação é que, para não paulistanos, ser paulistano pode estar relacionado a performar um homem heterossexual, entendedor do mercado de trabalho, que trabalha em ambiente corporativo, trajando roupas sociais. Essa interpretação pauta-se na análise do vídeo “TED TALK” do grupo de humor Porta dos Fundos, disponível na plataforma de vídeos YouTube⁵⁶. Nele, o humorista e ator carioca Gregório Duvivier representa um personagem que é o criador de “5 das 7 maiores *startups* de *LifeCoaching* do Brasil”. Nesse vídeo, o personagem, Peu Fraccaroli, está dando uma palestra sobre como resolver os problemas econômicos do Brasil. Sobre sua fala, destacam-se os usos de tepe em coda, de “meu”, da variante alongada em palavras como “abolição” e “tanto” e das palavras “top”

⁵⁶ Disponível em <https://www.youtube.com/watch?v=9Z-MgUDvD6g>. Acesso em 14 de mai. de 2020.

e “toppíssimo”. O (-r) tepe é um índice de paulistanidade – ao menos, para quem é de fora da cidade de São Paulo – (Oushiro, 2015; Soriano, 2016) e o “meu” parece funcionar da mesma forma (Oushiro, 2015), o que pode sugerir que a variante alongada é mais um índice linguístico que ajuda a compor um personagem homem paulistano do mundo corporativo. As palavras “top” e “toppíssimo”, por sua vez, são popularmente associadas a um grupo específico de homens heterossexuais, os “héterotopzera”⁵⁷. Considerando-se que o roteiro é também de autoria do carioca Gregório Duvivier, parece razoável dizer que, para certos ouvintes não paulistanos, a variante alongada correlaciona-se a um homem heterossexual paulistano que trabalha num ambiente corporativo.

Para a percepção de Amigabilidade para ouvintes paulistanos, testaram-se os possíveis efeitos das outras variáveis, e os resultados estão na **Tabela 16** a seguir.

Tabela 16 Resumo dos resultados do modelo de regressão linear múltipla para o CP3 de paulistanos (Amigabilidade)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Intercept	-1,017	0,605	-1,680	0,097 .
Gisele	0,488	0,124	3,927	<0,001 ***
Ricardo	0,048	0,124	0,387	0,698
Rodolfo	0,297	0,124	2,386	0,017 *
[ẽ:]	-0,033	0,088	-0,374	0,708
Masculino	-0,351	0,176	-1,989	0,050 .
Mais velhos	0,014	0,009	1,432	0,156
Leste	0,955	0,407	2,346	0,022 *
Norte	0,533	0,424	1,256	0,213
Oeste	0,975	0,389	2,507	0,014 *
Sul	1,245	0,412	3,022	0,003 **
Ensino superior	-0,419	0,367	-1,142	0,257

Fórmula: lmer(CP3 ~ Falante + Variante + Sexo + Idade + Zona + Escolaridade + (1|Ouvinte), data=DATA)

Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

As variáveis que têm correlação com as respostas dos paulistanos no componente Amigabilidade são Falante e Zona. Gisele e Rodolfo são significativamente percebidos como mais amigáveis do que Bruna. Os ouvintes das zonas leste, oeste e sul percebem as vozes como mais amigáveis,

⁵⁷ Cf. <https://pt.quora.com/Como-se-comporta-o-hetero-topzera> e <https://www.buzzfeed.com/br/florapaul/teste-hetero-top>. Acesso em 14 de mai. de 2020.

relativamente aos ouvintes do centro de São Paulo. Esses resultados também podem ser visualizados nas **Figura 39** e **40** a seguir. Gênero do ouvinte tem efeito marginalmente significativo nas respostas dadas pelos paulistanos ($p=0,050885$): ouvintes masculinos perceberam Bruna como menos amigável do que ouvintes mulheres. Por fim, no modelo de interação, verificou-se que nenhuma variável interage com Variante.

Figura 39 Distribuição das respostas dos ouvintes paulistanos para o CP3 (Amigabilidade) por Falante

Figura 40 Distribuição das respostas dos ouvintes paulistanos para o CP3 (Amigabilidade) por Zona

Os resultados da análise das respostas dos não paulistanos estão resumidos na **Tabela 17** abaixo. Além de Variante, Falante também tem efeito na percepção de Gênero: Gisele, Ricardo e Rodolfo soam significativamente mais masculinos/paulistanos do que Bruna – o que também pode ser evidenciado na **Figura 41**. No modelo com interações, verificou-se que nenhuma variável interage com Variante.

Tabela 17 Resumo dos resultados do modelo de regressão linear múltipla para o CP3 de não paulistanos (Gênero)(Efeitos Mistos)

Efeitos fixos	Estimativa	Erro padrão	Valor-t	Valor-p
Intercept	-1,334	0,262	-5,087	<0,001 ***
Gisele	0,859	0,102	8,364	<0,001 ***
Ricardo	1,808	0,102	17,588	<0,001 ***
Rodolfo	2,152	0,102	20,940	<0,001 ***
[ẽ:]	0,366	0,072	5,038	<0,001 ***
Masculino	-0,000	0,074	-0,013	0,990
Mais velhos	-0,007	0,005	-1,315	0,190
Leste	0,257	0,273	0,941	0,348
Norte	0,070	0,154	0,456	0,649
Oeste	0,079	0,096	0,828	0,408
Sul	0,081	0,134	0,611	0,542
Ensino superior	0,079	0,159	0,499	0,618

Fórmula: lmer(CP3 ~ Falante + Variante + Sexo + Idade + Zona + Escolaridade + (1|Ouvinte), data=DATA)

Intercept= Bruna, [ẽ], Feminino, Mais novos, Centro, Ensino médio

Figura 41 Distribuição das respostas dos ouvintes não paulistanos para o CP3 (Gênero) por Falante

Para o terceiro componente das respostas dadas pelos dois grupos de ouvintes, notou-se que a variante alongada está abaixo do nível de consciência dos ouvintes paulistanos, ao passo que funciona como um *estereótipo* (Labov, 2008 [1972]) para ouvintes não paulistanos por ser um índice de identidade sociodemográfica (paulistanidade) para esse grupo.

4.4 VARIÁVEIS QUALITATIVAS

Nesta subseção, são discutidos os resultados para as duas variáveis qualitativas, a saber: Zona de Residência e Faixa Etária. Foram feitos aqui testes de qui-quadrado, cuja função é verificar se há correlação entre duas variáveis nominais/qualitativas, sendo uma dependente (Zona ou Faixa Etária) e outra independente (a variável linguística). Os resultados para Zona por grupo de ouvintes podem ser visualizados na **Figura 42** a seguir.

Figura 42 Distribuição das respostas para Zona dos ouvintes paulistanos e dos não paulistanos
 $+ = p > 0,05$; $* = p < 0,05$; $** = p < 0,01$; $*** = p < 0,001$

De modo geral, para ambos os grupos de ouvintes, as variantes linguísticas não têm efeito na percepção da Zona de Residência do falante (esse resultado é representado pelo símbolo “+” na figura). Apesar de (AN) associar-se à localização do bairro (em termos de periférico-central) para ouvintes não paulistanos, como se viu para o CP1 *Status* social, o mesmo não se aplica para as cinco zonas. Uma possível explicação para esse resultado pode estar no fato de

que, por serem de fora da capital paulista, esses ouvintes não sabem direito como funciona a divisão da cidade em termos de zona e as características de cada uma. Certa ouvinte, inclusive, comentou algo semelhante com a autora após preencher o formulário. Com relação aos ouvintes paulistanos, esse já era um resultado esperado dados os comentários metalinguísticos do primeiro momento desta pesquisa (cf. **Capítulo 3**).

A variável Faixa Etária só se correlaciona com (AN) nas repostas de ouvintes não paulistanos, como se vê na **Figura43** abaixo (resultado representado pelo símbolo “*” na figura). Para este grupo, os falantes são percebidos mais frequentemente como adultos por volta dos 20 anos quando ouvidos com [ẽ:], relativamente [ẽ] ($p < 0,05$). Essa foi também a faixa etária mais assinalada para esse grupo de ouvintes. Ou seja, a variante alongada “tipicamente paulistana” é mais frequentemente associada a pessoas mais jovens, o que vai ao encontro do que se esperava, com base nos discursos metalinguísticos dos não paulistanos e na imitação feita pelo carioca Marcelo Adnet, que diz que a *Juli[ẽ:]na* tem 25 [ẽ:]nos – cf. **Introdução**. Já para os paulistanos, também como esperado com base em seus discursos metalinguísticos e nas análises dos CPs, (AN) não tem efeito na percepção da Faixa Etária do falante. Para esses ouvintes, a maioria das respostas concentra-se em duas faixas etárias: adulto por volta dos 20 anos e dos 30 anos. Para ambos os grupos, “Adolescente” e “Por volta dos 50” foram pouco selecionadas e a opção “Adulto por volta dos 40 anos” não foi selecionada nenhuma vez e, por isso, não consta nos gráficos.

Figura 43 Distribuição das respostas para Faixa Etária dos ouvintes paulistanos e dos ouvintes não paulistanos
 $+ = p > 0,05$; $* = p < 0,05$; $** = p < 0,01$; $*** = p < 0,001$

Nesta subseção, verificou-se que as variáveis qualitativas/nominais também revelam a importância das dimensões sociais das variantes de (AN) a depender da origem do ouvinte. O fato de que as respostas dos ouvintes não paulistanos para Zona e localização do bairro não terem caminhado na mesma direção pode ser uma evidência de que ser de fora da capital paulista implica em conhecer menos a divisão da cidade. Por essa perspectiva, talvez, *periferia* e *centro* têm o mesmo sentido para esses ouvintes, independentemente de onde eles sejam: *periferia* é onde se localizam famílias de menores rendas e em situação de exclusão social (com todos os prejuízos que isso acarreta), ao passo que os *centros* concentram melhores condições de vida – relativamente à periferia. Tanto é que, para este grupo de ouvintes, localização do bairro compõe o mesmo CP que classe social (cf. subseção 4.1). No entanto, a divisão de cidades em zonas (norte, sul, leste, oeste e centro) não é a realidade de todas as cidades brasileiras, e, naquelas em que é, nem sempre as características de suas zonas são as mesmas das da capital paulista. Portanto, a divisão da cidade de São Paulo em zonas pode ser algo complexo para não paulistanos. Esse grupo, ainda, percebe diferentemente as variantes de (AN) para Faixa Etária, associando a variante alongada a falantes mais novos. Por sua vez, para

paulistanos, a variável linguística não tem efeito na percepção de nenhuma das variáveis qualitativas aqui analisadas.

4.5 CAIXAS DE SELEÇÃO

Esta subseção reporta os resultados das análises das respostas para as 23 características (alta, branca, gay/lésbica etc.) apresentadas em forma de caixas de seleção no formulário de percepção. Assim como nas outras subseções, o objetivo aqui é verificar se a variável linguística tem efeito na percepção de tais características a depender da origem dos ouvintes. Para paulistanos, espera-se que (AN) não tenha efeito, de forma que eventuais diferenças significativas sejam por razão de outra variável. Para os não paulistanos, a expectativa é que a variante alongada seja significativamente relacionada às noções de “branca”, “articulada”, “metida”, “mimada”, “cult”, “conservadora”, “sofisticada”, “mauricinha/patricinha” e “riquinha” para falantes homens e mulheres, com base no que se discutiu no **Capítulo 3**. No caso somente de falantes homens, ainda para esse grupo de ouvintes e essa variante, espera-se também que a noção de “gay” seja frequentemente assinalada.

Nesta parte também foram realizados testes de qui-quadrado, já que se está lidando com duas variáveis nominais. Como se vê na **Tabela 18**, para ambos os grupos de ouvintes, apenas uma característica correlaciona-se significativamente com a variável (AN): “branca” para paulistanos e “conservadora” para não paulistanos. A tabela apresenta a frequência e a proporção com que tais características foram assinaladas (✓) e não assinaladas (-) de acordo com a variante ouvida, bem como os valores de qui-quadrado e seus respectivos valores-*p*. Os valores em negrito destacam a variante que favoreceu o assinalamento da caixa de seleção.

Para ouvintes paulistanos, a variante alongada ([ẽ:]) é aquela associada a uma pessoa branca, o que remete a uma questão tanto de etnia quanto de *status*, considerando-se que brancos usufruem de privilégios sociais e são a maioria com melhores condições socioeconômicas no Brasil. Por outro lado, para ouvintes não paulistanos, a variante não alongada ([ẽ]) é associada a uma pessoa conservadora, o que relaciona-se à esfera de posicionamento político-ideológico do indivíduo. Correlacionando-se esse resultado aos encontrados para as respostas dos CPs desse mesmo grupo de ouvintes e para a mesma variante, é plausível afirmar que pessoas de maior *status* social (CP1), mais profissionais (CP2), mais masculinas, paulistanas e menos femininas (CP3) são percebidas também como mais conservadoras.

Tabela 18 Características relacionadas à (AN)
para ouvintes paulistanos e não paulistanos

Ouvintes	Característica	[ẽ]	[ẽ] %	[ẽ:]	[ẽ:] %	Total	χ^2	P
Paulistanos	Branca							
	-	73	50	54	36,9	127		
	✓	73	50	92	63	165	4,51(1)	0,0336*
Não paulistanos	Conservadora							
	-	76	46	89	53,9	165		
	✓	24	68,5	11	31,4	35	4,98(1)	0,0255*

✓ = Assinalada; - = Não assinalada.

Por fim, o ouvinte tinha ainda a opção de assinalar a caixa de seleção “outro” no formulário e informar uma característica que não estivesse listada, de acordo com a sua percepção. Apenas quatro ouvintes paulistanos assinalaram tal opção e cada um escreveu uma das seguintes características: “negra”, “pedante”, “um ser humano” e “sonhadora” – as três primeiras nos disfarces com [ẽ], e a última, com [ẽ:]. Desse modo, essas características não puderam ser analisadas.

Conclusão

A observação de dois fatos de natureza empírica motivou esta dissertação de mestrado: i) o modo “peculiar” como falavam os VJs da antiga emissora *MTVBrasil*; e ii) as imitações caricatas de “sotaques paulistas/paulistanos” feitas por falantes fluminenses. Nessas imitações, destacam-se: i) “o paulista do interior” e “o sotaque da periferia de São Paulo”, com frequente uso de (-r) retroflexo; ii) “o paulista italiano” com uma pronúncia “italianada” – o que quer que isso signifique; e iii) “o ‘nasalês’ de São Paulo”, com frequente uso de (EN) ditongado e (AN) alongado. Sobre a percepção do (-r) em coda na cidade de São Paulo, têm-se os trabalhos de Oushiro (2015) e Soriano (2016). Por sua vez, os significados sociais correlacionados ao (EN) na capital paulista foram estudados por Oushiro (2015) e Mendes (2016). Por fim, o “paulista italiano” é estudado por Carmo (2018). Não havia trabalhos sobre os significados sociais associados à (AN) na cidade de São Paulo. Esta dissertação propôs-se a investigar tais significados, a fim de contribuir para os estudos de percepção sociolinguística do falar paulistano.

Além disso, os estudos de Oushiro (2015) e Santos (2020) evidenciam que a origem dos ouvintes é uma variável importante para estudos de percepção sociolinguística, na medida em que ouvintes de diferentes origens têm percepções distintas acerca das mesmas variantes de uma variável linguística. Desse modo, esta pesquisa de mestrado teve como objetivos, além de averiguar os significados sociais correlacionados às variantes de (AN), verificar se a origem do ouvinte tem efeito na percepção dessa variável linguística.

A fim de verificar minimamente como as variantes deste estudo se diferenciam acusticamente, o **Capítulo 1** discutiu suas propriedades articulatórias por meio de análises espectrais e cepstrais de uma ocorrência de cada variante na mesma palavra para cada um dos quatro falantes deste estudo. De modo geral, as análises mostram que a variante “tipicamente paulistana” é

produzida com maior tempo de duração (por isso sua representação fonética [ẽ:] e o nome “variante alongada”) e com a língua mais abaixada, relativamente à [ê] (variante não alongada).

O **Capítulo 2** apresentou uma revisão dos estudos de avaliação da fala, desde seu início na Psicologia Social, até chegar à Linguística, com ênfase na discussão de Eckert (2008, 2012) sobre o tratamento dado ao lado social da variação nas diferentes abordagens da Sociolinguística – as chamadas “ondas” de estudo sociolinguístico. Essa discussão foi pertinente na medida em que esta pesquisa de mestrado tomou como foco e metodologia, respectivamente, o interesse em acessar significados sociais e a técnica *matched-guise* – característicos da terceira onda –, no intuito de buscar padrões sociolinguísticos – objetivo da primeira onda –, ainda que para a percepção.

Em seguida, o **Capítulo 3** apresentou as metodologias acionadas no desenvolvimento do trabalho. Como se viu, o primeiro passo desta pesquisa foi fazer entrevistas metalinguísticas, a fim de obter metacomentários espontâneos dos ouvintes para a variante alongada, por meio da audição de trechos da fala de dois paulistanos (um homem e uma mulher) e de uma imitação caricata do “nasalês de São Paulo”, feita pelo humorista carioca Marcelo Adnet. As respostas dos 50 ouvintes (27 paulistanos e 23 não paulistanos – do interior e de fora do estado) foram analisadas separadamente e os resultados, já nesse primeiro momento, mostraram divergências entre os grupos de ouvintes: muitas vezes, os paulistanos disseram não identificar nada “de mais” nos estímulos, ao passo que não paulistanos chegaram a reconhecer a variante em questão e associá-la a certos tipos de pessoas (“paulistana”, “blogueira”, entre outras).

Com base nas respostas às entrevistas metalinguísticas, algumas hipóteses foram aventadas, apesar de melhor formalizadas mais adiante no **Capítulo 3**. Essas respostas serviram, ainda, como base para a escolha das

variáveis quantitativas e nominais que comporiam o formulário do teste de percepção com base na técnica *matched-guise*.

O teste de percepção foi respondido, via Google Forms, por 70 ouvintes paulistanos e 53 não paulistanos. No **Capítulo 4**, por meio da Análise de Componentes Principais (ACPs), constatou-se que, para os dois grupos de ouvintes, as 10 variáveis quantitativas podem, na realidade, reduzir-se a três componentes. No entanto, as escalas que compõem os CPs de cada grupo são diferentes – já uma evidência de que a origem dos ouvintes tem efeito na percepção de (AN), visto que todos os ouvintes foram expostos aos mesmos disfarces e que somente a variável linguística em foco os diferenciava. As análises seguintes foram realizadas no sentido de verificar se (AN) tinha efeito nas respostas que integravam todos os CPs identificados nas respostas de ambos grupos de ouvintes, além de verificar se variáveis sociais correlacionavam-se a elas.

O primeiro componente principal nas respostas dos ouvintes paulistanos é Prestígio, composto pelas escalas formalidade, escolaridade, sucesso profissional, classe social e localização do bairro, positivamente correlacionadas, e não se correlaciona com a variável linguística. Por outro lado, o CP1 identificado nas respostas dos ouvintes não paulistanos, *Status* social, é integrado pelas escalas classe social e localização do bairro, em correlação positiva, e correlaciona-se com (AN), de modo que os disfarces com a variante alongada são significativamente percebidos como produzidos por falantes de maior *status* social. Desse modo, para não paulistanos, [ẽ:] correlaciona-se à condição socioeconômica do falante.

A variável linguística comporta-se de modo diferente também para o segundo CP: para paulistanos, ela não tem efeito na percepção de Gênero (integrado pelas escalas de feminilidade e masculinidade em correlação negativa), ao passo que, para não paulistanos, (AN) tem efeito na percepção de

Profissionalismo (composto pelas escalas de inteligência, amigabilidade e sucesso profissional, positivamente correlacionadas).

Para as respostas que integram o CP3 das respostas dos ouvintes paulistanos (Amigabilidade, formado exclusivamente pela escala de mesmo nome), novamente não se verificou efeito significativo de (AN). Contudo, para não paulistanos, a variável linguística teve efeito na percepção de Gênero (composto pelas escalas de masculinidade, paulistanidade e feminilidade – as duas primeiras em correlação positiva e em correlação negativa com a terceira). Os resultados dos três CPs evidenciam que a variante alongada está abaixo do nível de consciência dos ouvintes paulistanos, enquanto funciona como um *estereótipo* (Labov, 2008 [1972]) para ouvintes não paulistanos por ser um índice de identidade sociodemográfica (paulistanidade) para esse grupo – tal qual (EN) ditongado em Oushiro (2015).

O fato de a variante alongada ser positivamente valorada com mais frequência nas respostas aos três CPs dos não paulistanos e o modo como esses CPs se configuram podem ser uma evidencia de que [ẽ:] correlaciona-se, como se discutiu no decorrer do **Capítulo 4**, com uma *persona* tal qual aquela acionada pela personagem Peu Fraccaroli do vídeo “TED TALKS” do grupo de humor Porta dos Fundos: um homem branco criador de 5 das 7 maiores *startups* de *LifeCoaching* do Brasil. Uma pessoa dessas certamente pertence à classe média alta e reside em bairros nobres (escalas do CP1 dos ouvintes não paulistanos), além de ser inteligente e bem sucedida profissionalmente (duas das três escalas de CP2 do mesmo grupo). No tocante à paulistanidade desse personagem, podem-se observar os seguintes comentários feitos sobre o vídeo: “toda vez que o gregorio [sic] faz um idiota, ele faz um paulistano” e “21 mil fundadores de startups de lifecoaching paulistanos não curtiram...”. Esses comentários justificam o fato de a variante alongada correlacionar-se à paulistanidade (uma das escalas de CP3 dos não paulistanos), mas não justifica a correlação com amigabilidade (uma das escalas de CP2 desse grupo de

ouvintes), visto que uma pessoa “idiota” dificilmente seria considerada amigável. A questão aqui é diferença entre o tema do vídeo e o dos disfarces: Peú Fraccaroli fala sobre a volta da escravidão como resolução para a questão econômica do Brasil, ao passo que os falantes desta pesquisa leram trechos que remetem à infância. Portanto, o tema dos disfarces pode afetar a percepção das variáveis linguísticas.

Outra importante observação sobre o vídeo “TED TALKS” é que ele foi produzido por um humorista carioca (Gregório Duvivier), o que é mais uma evidência de que os ouvintes de fora de São Paulo percebem a variante alongada como índice do falar paulistano. Não é possível dizer ao certo de onde são as pessoas que teceram os comentários sobre o vídeo, mas pouco provavelmente o primeiro é paulistano, dado o modo um tanto ofensivo com que correlaciona um “paulistano” a um “idiota” – muito provavelmente, o autor do comentário não se incluiria nessa correlação.

Para as variáveis qualitativas, apenas Faixa Etária correlaciona-se à (AN) nas respostas dos não paulistanos: a variante alongada é associada a adultos por volta dos 20 anos. Esse resultado era esperado, visto que, em sua imitação caricata, o carioca Marcelo Adnet diz que *Juli[ẽ:]na* tem 25 [ẽ:]nos. As respostas desse grupo de ouvintes para Zona não mostram correlação com (AN) (diferentemente do que se esperava), um indício de que, aparentemente, os não paulistanos não sabem direito como funciona a divisão da cidade de São Paulo em termos de zona e as características de cada uma. Com isso, a depender dos interesses do seu estudo, um pesquisador deve ater-se a essas variáveis que parecem ter sentido distinto a depender da origem dos participantes. Nas respostas dos ouvintes paulistanos, (AN) não teve qualquer efeito na percepção dessas variáveis sociais.

Por fim, para as respostas às caixas de seleção, ouvintes paulistanos associam a variante alongada ([ẽ:]) a uma pessoa branca, o que remete a uma questão tanto de etnia quanto de *status*, considerando-se que brancos usufruem

de privilégios sociais e são a maioria com melhores condições socioeconômicas no Brasil. Por outro lado, para ouvintes não paulistanos, a variante não alongada ([ẽ]) é associada a uma pessoa conservadora, o que relaciona-se à esfera de posicionamento político-ideológico do indivíduo. É plausível afirmar que o mesmo personagem Peu Fracaroli, que trabalha em ambiente corporativo e é criador de *startups*, pode ser associado a alguém que compactua com uma ideologia mais neoliberal/capitalista e, portanto, é mais conservador. Talvez essa seja a associação feita pelos ouvintes não paulistanos que se voluntariaram a participar do experimento.

Esta pesquisa de mestrado estudou os significados sociais correlacionados a uma variável linguística nunca antes estudada, diferentemente de trabalhos prévios, que partem do estudo de *produção* de uma variável para depois realizar estudos de *percepção*. Com isso, este trabalho mostra que é possível estudar o lado social da variação por outro viés que independe dos resultados para as categorias sociais (macro ou micro) aos quais os falantes pertencem e que condicionam o uso de determinada variante. No entanto, sabe-se que os estudos de percepção e produção da mesma variável contribuem substancialmente para o debate sobre a variação sociolinguística e, nesse sentido, este trabalho espera contribuir com estudos vindouros acerca da produção de (AN).

Ainda sobre os estudos de produção, acredita-se que eles sejam feitos anteriormente aos estudos de percepção para que seja constatado o uso condicionado (por fatores linguísticos e sociais) de determinada variante e, dessa forma, correlacionar as categorias sociais dos falantes aos significados das variantes por eles produzidas. No entanto, este trabalho mostrou que partir de análises de natureza empírica, como a observação de vídeos de humor, para suscitar questões de estudo e analisar determinados resultados é uma metodologia que funciona para os estudos de percepção. Entende-se que vídeos de humor, como o de Marcelo Adnet e Gregório Duvivier apresentados neste

trabalho, utilizam-se de exageros e caricaturas, a partir da constatação fatos cotidianos que saltam aos ouvidos, para gerar humor.

Finalmente, esta dissertação contribui para os estudos de percepção da fala paulistana, levando-se em consideração quem de fato diz que aquela fala é paulistana (indivíduos da cidade de São Paulo ou de fora dela?). Por meio da padronização das percepções sociolinguísticas de (AN), e revisitando trabalhos prévios sobre variantes associadas à paulistanidade, abrem-se novos caminhos para os estudos dos significados sociais das variáveis linguísticas: como os diferentes índices de paulistanidade se recombinariam na projeção de diferentes *personae*? Trata-se, portanto, de propor novos olhares para o lado social da variação.

Referências

AGHA, A. 'Voice, Footing, Enregisterment'. In: **Journal Linguistic Anthropology**, Vol. 15, p. 38-59, 2005.

_____. **Language and social relations**. Cambridge: Cambridge University Press. 2007.

BARBOSA, P. A.; MADUREIRA, S. **Manual de Fonética Acústica Experimental. Aplicações a dados do português**. 1. ed. São Paulo: Cortez, 2015. 591p.

BIELER, M. E.; MENDES, R. B. **Entre duas metrópoles: (-r) em Itanhandu**. 2015. Universidade de São Paulo, São Paulo, 2015. Disponível em: <http://www.teses.usp.br/teses/disponiveis/8/8139/tde-25112015-111636/>

BISOL, L. Fonologia da nasalização. In: ABAURRE, M. B. M. (Org.). **Gramática do português culto falado no Brasil, Volume VII: A construção fonológica da palavra**. 1ªed. São Paulo: Contexto, 2013, v. VII, p. 113-140.

BOERSMA, P. & WEENINK, D. **Praat: doing phonetics by computer**. URL: <http://www.fon.hum.uva.nl/praat/>, 2012. Acessado em: 16 de fev. de 2020.

CALLOU, D.; MORAES, J. A.; LEITE, Y. As vogais orais: um estudo acústico-variacionista. In: ABAURRE, M. B. M. (Org.). **Gramática do português culto falado no Brasil, Volume VII: A construção fonológica da palavra**. 1ed. São Paulo: Contexto, 2013, v. VII, p. 75-93.

CAMACHO, R. G. 'Uma Reflexão Crítica sobre a teoria sociolinguística'. In: **D.E.L.T.A. Documentação de Estudos em Linguística Teórica e Aplicada**, nº 26, vol. 1, p. 141-162, 2010.

CAMARA JR. J. M. **Para um estudo da fonêmica portuguesa**. Rio de Janeiro: Organização Simões, 1953.

_____. **Estrutura da Língua Portuguesa**. Petrópolis: Vozes, 2009[1970]).

CAMPBELL-KIBLER, K. **Listener perceptions of sociolinguistic variables: the case of (ing)**. Tese de Doutorado. Stanford University, 2006. 282f.

_____. 'Accent, (ING), and the social logic of listener perceptions'. In: **American Speech**, vol. 82, p. 32-61, 2007.

_____. 'The nature of sociolinguistic perception'. In: **Language Variation and Change**, vol. 21, p. 135-156, 2009.

_____. 'Intersecting variables and perceived sexual orientation in men'. In: **American Speech**, vol. 86, nº 1, p. 52-68, 2011.

CARMO, L. M. **O português falado na Mooca: Variação linguística na fala de moradores do bairro Mooca em São Paulo**. Projeto de Doutorado apresentado ao Programa de Pós-Graduação em Semiótica e Linguística Geral do Departamento de Linguística-FFLCH/USP, 2018. Dr.

CARVALHO, B. B. A. **“O que você acha do uso de tu?”: a percepção da variação dos pronomes de 2SG no dialeto carioca.** Dissertação de Mestrado. Rio de Janeiro: Faculdade de Letras/UFRJ. 2019.

CATTELL, R. B. The Scree Plot Test for the Number of Factors. In: **Multivariate Behavioral Research**, 1, p.140-161. 1966.
http://dx.doi.org/10.1207/s15327906mbr0102_10

ECKERT, P. **Linguistic Variation as Social Practice.** Blackwell Publishers, 2000.

_____. ‘Style and Social Meaning’. In: ECKERT, P.; RICKFORD, J. R. **Style and Sociolinguistic Variation.** p. 119-126. Cambridge: Cambridge University Press, 2001.

_____. ‘Variation and the indexical field’. In: **Journal of Sociolinguistics**, nº 12, vol. 4, p. 453-476, 2008.

_____. ‘Three Waves of Variation Study: the emergence of Meaning in the Study of Sociolinguistic Variation’. In: **Annual Review of Anthropology**, vol. 41. p. 87-100, 2012.

_____. ‘Variation, meaning and social change’. In: COUPLAND, N. **Sociolinguistics: Theoretical debates.**Cambridge: Cambridge University Press, p. 68-85, 2016.

IDH, **Os 20 melhores e os 20 piores distritos de São Paulo.** Estadão, São Paulo. 24 de abr. de 2016. Disponível em <https://fotos.estadao.com.br/galerias/cidades,idh-os-20-melhores-e-os-20-piores-distritos-de-sao->

paulo,24925?utm_medium=website&utm_source=archdaily.com.br Acessado em 10 de jan. de 2020.

GAUDIO, R. P. Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men. In: **American Speech**, nº 69, p. 30–57, 1994.

GILES, H.; BILLINGS, A. C. 'Assessing Language Attitudes: speaker evaluation studies'. In: DAVIES, A. & ELDER, C. (Eds.). **The Handbook of Applied Linguistics**. Malden, Ma: Blackwell Publishing, p. 187-209, 2004.

GONÇALVES, D. T. **Social Indexation of Sexuality with Brazilian Speakers of Rio de Janeiro**. 2020. (Apresentação de Trabalho).

IRVINE, J. T. Style as distinctiveness: the culture and ideology of linguistic differentiation. In: Eckert, P. & Rickford, J. (eds.) **Style and sociolinguistic variation**. Cambridge: Cambridge University Press, p. 21-43, 2001.

LABOV, W. 'The unobservability of structure and its linguistic consequences'. **Paper presented at the 22nd New Ways in Analyzing Variation conference**, University of Ottawa, October 22–25, 1993.

_____. **Principles of Linguistic Change: external factors**. Oxford & Cambridge: Blackwell, 2001.

_____. **The Social Stratification of English in New York City**. São Paulo: Cambridge University Press, 2006[1966].

_____. **Padrões Sociolinguísticos**. São Paulo: Parábola, 2008[1972].

_____; ASH, Sharon; RAVINDRANATH, Maya; WELDON, Tracey; BARANOWSKI, Maciej; NAGY, Naomi. 'Properties of the sociolinguistic monitor'. In: **Journal of Sociolinguistics**, vol. 15 (4): Blackwell Publishing, pp. 431-463, 2011.

LAMBERT, W. E., HODSON, R. C., GARDNER, R. C. & FILLENBAUM, S. 'Evaluational reactions to spoken languages'. In: **Journal of Abnormal and Social Psychology**, vol. 60(1), 44-51, 1960.

LEVON, E. 'Categories, stereotypes, and the linguistic perception of sexuality'. In: **Language in Society**, vol. 43, p. 539-566, 2014.

_____; FOX, Sue. 'Social Salience and the Sociolinguistic Monitor: a Case Study of ING and TH-fronting in Britain'. In: **Journal of English Linguistics**, nº 3, vol. 42, p. 185-218, 2014.

LEVSHINA, N. **How to do Linguistics with R**. Amsterdam/Philadelphia: John Benjamins, 2015.

MACDONELL, D. **Theories of Discourse: An Introduction**. Oxford: BasilBlackwell, 1986

MELO, M. A. S. L. **Direcionalidade da mudança sonora: o papel do item lexical e da avaliação social**. Rio de Janeiro, UFRJ, Faculdade de Letras, 2012. Tese de Doutorado em Linguística.

_____. GOMES, C. A. Percepção da variação da coda (s) na comunidade de fala do Rio de Janeiro: acessando o significado social da variante fricativa

posterior. In: VIEIRA, M. S. M.; WIEDEMER, M. L. (Org.). **Dimensões e Experiências em Sociolinguística**. 1ed. São Paulo: Blucher, v. 1, p. 129-148. 2019.

MENDES, R. B. Gendered Perceptions of Noun Agreement in Brazilian Portuguese. In: **Revista Internacional de Linguística Iberoamericana**, v. XVII, p. 93-108, 2014.

_____. 'Nonstandard Plural Noun Phrase Agreement as an Index of Masculinity'. In: LEVON, E.; MENDES, R. B. **Language, Sexuality, and Power: Studies in intersectional sociolinguistics**. New York: Oxford University Press, 2016.

_____. 'A Terceira Onda da Sociolinguística'. In: FIORIN, José Luiz. **Novos Caminhos da Linguística**. São Paulo: Contexto, p. 103-123, 2017.

_____. **Percepção e Performance de Masculinidades: Efeitos da concordância nominal de número e da pronúncia de /e/ nasal**. Tese (Livre Docência) – FFLCH-USP. São Paulo, 225f, 2018.

MILROY, L. **Language and social networks**. Oxford: Basil Blackwell, 2 ed., 1987 [1980].

MORAES, J. A. Produção e percepção das vogais nasais. In: ABAURRE, M.B.M. (Org.). **Gramática do português culto falado no Brasil, Volume VII: A construção fonológica da palavra**. 1ed. São Paulo: Contexto, v. VII, p. 95-112. 2013.

_____.; WETZELS, L. . Sobre a duração dos segmentos nasais e nasalizados em português. Um exercício de fonologia experimental. In: **Cadernos de Estudos Linguísticos** (UNICAMP), CAMPINAS, v. 23, p. 153-166, 1992.

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

_____. **Introdução à Estatística para Linguistas**, v.0.3 (mai/2017). Cópia disponível com autora: oushiro@iel.unicamp.br, 2017.

_____. A computational approach for modeling the indexical field / Uma abordagem computacional para a modelagem de campos indexicais. **Revista De Estudos da Linguagem**, v. 27, p. 1737-1786, 2019.

PODESVA, R. **Phonetic detail in sociolinguistic variation: Its linguistic significance and role in the construction of social meaning**. PhD dissertation. Stanford University, 2006.

_____. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. In: **Journal of Sociolinguistics**, 11/4, 478-504. 2007a.

_____. **Three sources of stylistic meaning**. Washington D.C.: Georgetown University. 2007b.

_____. Salience and the social meaning of declarative contours: Three case studies of gay professionals. In: **Journal of English Linguistics** 39(3), p. 233-264, 2011.

R TEAM, Development Core. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL: <http://www.R-project.org/>.

RAPOSO de MEDEIROS, B. Vogais nasais do português brasileiro: reflexões preliminares. In: **Revista de Letras** (Curitiba. 1996), v. 72, p. 165-188, 2007.

REISE, S. P.; WALLER, N. G.; COMREY, A. L. 'Factor Analysis and Scale Revision'. In: **Psychological Assessment**, vol. 12, nº 3, p. 287–297, 2000.

REVELLE, W. **Procedures for Personality and Psychological Research**. R package version. 2015.

SANTOS, W. **Percepções sociolinguísticas acerca da variação subjuntivo/indicativo em São Luís e São Paulo**. 235f. Tese (Doutorado em Semiótica e Linguística Geral) – FFLCH, USP, São Paulo, 2020.

SORIANO, L. G. M. **Percepções sociofonéticas do (-r) em São Paulo**. 137f. Dissertação (Mestrado em Semiótica e Linguística Geral). FFLCH, USP, São Paulo, 2016.

SOUZA, L. C. S.; PACHECO, V. Uma análise acústica das vogais orais, nasais e nasalizadas no dialeto de Vitória da Conquista-BA. **Signum. Estudos de Linguagem**, v. 2, p. 401-431, 2013.

WETZELS, W.L.M. Harmonia Vocálica, Truncamento, Abaixamento e Neutralização no Sistema Verbal do Português: Uma Análise Autosegmental. In: **Cadernos de Estudos Linguísticos** (UNICAMP), v. 1991, p. 25-58, 1991.

Apêndices

Apêndice A – Formulário de percepção *online*

Pesquisa

Olá! Tudo bem?

Obrigado por participar dessa pesquisa de mestrado da USP sobre a cidade de São Paulo. Para participar, basta ser morador da cidade (não precisa ter nascido nela). Por gentileza, use fone de ouvido!

Você vai ouvir quatro pessoas (uma por vez) nos clipes de áudio a seguir. Você pode ouvi-las quantas vezes quiser.

Em seguida, por favor, preencha o formulário que segue cada um dos clipes, de acordo com a sua percepção. Não há respostas certas ou erradas.

Vamos lá?

***Obrigatório**

Vamos à primeira dessas quatro pessoas.

1

Na sua opinião, essa pessoa parece... *

	1	2	3	4	5	6	
Nada inteligente	☐	☐	☐	☐	☐	☐	Muito inteligente

*

	1	2	3	4	5	6	
Nada amigável	☐	☐	☐	☐	☐	☐	Muito amigável

*

	1	2	3	4	5	6	
Nada feminina	☐	☐	☐	☐	☐	☐	Muito feminina

*

	1	2	3	4	5	6	
Nada formal	☐	☐	☐	☐	☐	☐	Muito formal

*

	1	2	3	4	5	6	
Nada paulistana	☐	☐	☐	☐	☐	☐	Muito paulistana

*

Nada escolarizada 1 2 3 4 5 6 Muito escolarizada

☐ ☐ ☐ ☐ ☐ ☐

*

Mal sucedida na profissão 1 2 3 4 5 6 Bem sucedida na profissão

☐ ☐ ☐ ☐ ☐ ☐

*

De classe social mais baixa 1 2 3 4 5 6 De classe social mais alta

☐ ☐ ☐ ☐ ☐ ☐

*

De um bairro mais periférico de SP 1 2 3 4 5 6 De um bairro mais central de SP

☐ ☐ ☐ ☐ ☐ ☐

*

	1	2	3	4	5	6	
Nada masculina	☐	☐	☐	☐	☐	☐	Muito masculina

Em que zona de São Paulo você acha que essa pessoa mora? *

- ☐ Norte
- ☐ Sul
- ☐ Leste
- ☐ Oeste
- ☐ Centro

Você acha que essa pessoa é: *

- ☐ Adolescente
- ☐ Adulta de 20 e poucos anos
- ☐ Adulta por volta dos 30 anos
- ☐ Adulta por volta dos 40 anos
- ☐ Adulta por volta dos 50 anos

Você acha que essa pessoa deve ser (marque quantas opções quiser):

- ☐ Alta
- ☐ Branca
- ☐ Gay/Lésbica
- ☐ Tímida
- ☐ Descolada
- ☐ Religiosa
- ☐ Articulada
- ☐ Irritante
- ☐ Metida
- ☐ Mimada
- ☐ Culta
- ☐ Tranquila
- ☐ Conservadora
- ☐ Sofisticada
- ☐ Independente
- ☐ Mauricinho/Patricinha
- ☐ Riquinha
- ☐ Confiável
- ☐ Simples
- ☐ Trabalhadora
- ☐ Caipira
- ☐ Engraçada
- ☐ Calma
- ☐ Outro: _____

Agora, para terminar, me fale um pouco sobre você.

Idade: *

Sua resposta

Sexo: *

☐ Feminino.

☐ Masculino.

☐ Outro: _____

Profissão/Ocupação: *

Sua resposta

Escolaridade: *

☐ Ensino fundamental (completo ou incompleto)

☐ Ensino médio (completo ou incompleto)

☐ Ensino superior (completo ou incompleto)

Bairro onde mora: *

Sua resposta

Zona de São Paulo onde mora: *

- ☐ Sul
- ☐ Norte
- ☐ Oeste
- ☐ Leste
- ☐ Centro

Local onde nasceu: *

- ☐ São Paulo, capital.
- ☐ São Paulo, interior.
- ☐ Outro estado.

Se NÃO nasceu na CIDADE de São Paulo, em que CIDADE e ESTADO nasceu?

Sua resposta

Se NÃO nasceu na CIDADE de São Paulo, há quanto tempo mora nela, aproximadamente?

Sua resposta

Basta clicar no "enviar" e é isso! Muito obrigada pela sua participação!

Voltar

Enviar