

A identificação das variedades regionais do português brasileiro através da informação entoacional

The identification of Brazilian Portuguese regional varieties through intonational information

Daniel Oliveira Peres

Universidade de São Paulo – São Paulo – São Paulo – Brasil



Resumo: Este estudo tem como objetivo verificar se falantes do português brasileiro são capazes de reconhecer suas próprias variedades regionais a partir das informações entoacionais. Além disso, também pretende procurar pistas nas variações de F0 que possam justificar o desempenho dos participantes. Os dois experimentos realizados neste trabalho foram feitos com base em três variedades do Português Brasileiro: a variedade de Pelotas (RS), São Paulo (bairro da Mooca) e Senador Pompeu (CE). O experimento 1 foi desenvolvido com estímulos de fala delexicalizada, e.g., com apenas as informações entoacionais; o experimento 2 apresentou estímulos de fala sem variação de F0, e.g., apenas com a curva melódica monotônica. Os resultados obtidos através da *Teoria de Detecção* (MACMILLAN; CREELMAN, 2005) mostraram que os participantes tiveram um desempenho significativo. Os resultados obtidos nos dois experimentos foram diferentes, pois houve mais dificuldade de reconhecimento no experimento com estímulos delexicalizados. A tendência dos alarmes falsos também mostrou que a informação entoacional desempenha um papel importante no reconhecimento de variedades regionais.

Palavras-chave: Sociofonética; Percepção de fala; Entoação; Português do Brasil

Abstract: This study aims to verify if Brazilian Portuguese speakers are able to recognize their own regional varieties only by listening to the intonational information. It also aims to search for cues in F0 variation, which can justify the performance of the participants. The two experiments conducted in this study deal with three varieties of Brazilian Portuguese: the variety of Pelotas (RS), São Paulo (Mooca district) and Senador Pompeu (CE). Experiment 1 was developed with delexicalised speech stimuli, e.g., intonational information only; experiment 2 presented stimuli with a flat melodic curve only, e.g., with no F0 variations. The results based on the *Detection Theory* (MACMILLAN; CREELMAN, 2005) showed that the participants had a performance above chance. The results obtained in both experiments were different as we observed more recognition difficulties in the experiment with delexicalized stimuli. The tendency of false alarms also showed that intonational information plays an important role in the recognition of regional varieties.

Keywords: Sociophonetics; Speech perception; Intonation; Brazilian Portuguese

Introdução

Este trabalho analisará a variação regional do português brasileiro (PB) com foco no viés perceptivo dos estudos da linguagem. A variabilidade é uma característica largamente estudada, seja ela no âmbito dialetal ou idiolectal, além de ser um fato comum às línguas naturais como um todo. Por conta disso, deve-se buscar maneiras de entender como essa variação é processada na percepção da fala (CLOPPER; PISONI, 2004). Para tal, este estudo pretende trazer para os estudos de percepção

da fala, especificamente os de percepção da entoação, os fins encontrados nos estudos sociolinguísticos. Esse encontro tem gerado um vasto número de metodologias aplicadas aos estudos sociolinguísticos que se interessam pela representação mental das variações dialetais.¹ Assim, propõe-se uma pesquisa que integre os vieses perceptivo e sociolinguístico, especificamente, um trabalho que busque nas características acústicas o entendimento de

¹ Como exemplos dessas metodologias, temos: *map-drawing task*, *attitude judgments*, *speech caricatures*, *vowel matching* e *matched-guise technique*. Consultar Clopper e Pisoni (2004) para mais detalhes.

como a fala pode ser percebida e suas implicações no reconhecimento das variedades regionais do português brasileiro.

Neste estudo, a abordagem perceptual está restrita ao papel da entoação como pista para o reconhecimento dialetal. Os trabalhos de Atkinsons (1968), Bonte (1975), Maidment (1976) e Ohala e Gilbert (1978) já atestaram a importância da prosódia como pista no reconhecimento de línguas. A motivação desses trabalhos foi a observação ordinária de que as línguas diferem umas das outras por seu inventário segmental e por suas restrições fonotáticas, além de possuírem critérios de formação de palavras e construção de sentenças muito diferentes entre si, mesmo que essas diferenças sejam mais tênues entre algumas línguas e mais salientes entre outras. Outros estudos abordaram o reconhecimento da prosódia de diferentes variedades do inglês, como no estudo de Bush (1967), sobre as variedades britânica, americana e indiana; e o de Richardson (1973), sobre o *Black* e o *White English*. De acordo com esses estudos, pode-se afirmar que as características contidas nas variações de frequência fundamental, amplitude e duração desempenham o papel de sinalizadores de diversos tipos de variação linguística, sejam elas de caráter diatópico, diafásico ou diastrático.

Com base na hipótese do reconhecimento linguístico e dialetal por meio das informações entoacionais, este trabalho visa a desenvolver um estudo perceptual no qual se possa testar a habilidade dos falantes de português na tarefa de reconhecimento de sua variedade linguística. Além disso, com o propósito de saber quais informações são relevantes para tal reconhecimento, será feita uma análise de produção para avaliar quais informações podem estar interferindo na escolha dos ouvintes. As três variedades estudadas neste trabalho são: a variedade de São Paulo, especificamente a do bairro da Mooca; a variedade do Ceará, da cidade de Senador Pompeu, e a variedade do Rio Grande do Sul, da cidade de Pelotas. Os informantes participaram de dois experimentos de percepção; o primeiro foi desenvolvido com estímulos de fala delexicalizada, o segundo com estímulos de fala monotonal. Os dados foram analisados estatisticamente, além da análise via Teoria da Detecção (MACMILLAN; CREELMAN, 2005), e, por fim, confrontados com a análise de produção.

1 Breve histórico da formação do português brasileiro

Abordar a história de qualquer língua, seja ela vernácula ou não, encerra um trabalho exaustivo que não se pode abarcar de uma única vez. Quando se anuncia, neste capítulo, fazer um breve histórico, tem-se em mente a possibilidade de traçar linhas gerais que possam justificar

as diferenças que, de outros tempos até os dias de hoje, são notadas entre as variedades estudadas neste trabalho. Ademais, essa breve visão geral tem como objetivo justificar a escolha das variedades regionais analisadas neste trabalho.

A língua portuguesa trazida no século XVI e em subsequentes incursões dos portugueses no Brasil já não era um bloco, um todo uniforme, mas trazia variações que estavam presentes no português europeu. Fernão de Oliveira² (1981 apud SILVA NETO, 1986, p. 42) diz que “os homens de Estremadura são diferentes dos D’antre Douro e Minho: porque assi como os tempos, assi também as terras criam diversas condições e conceitos”. Parece, então, natural que essa língua ganhe aqui, em terras brasileiras, novos nuances. A diversidade linguística encontrada no português brasileiro, do ponto de vista diatópico, é um indício forte de que o português incorporou elementos autóctones. Castilho (2000) sintetiza as duas abordagens acerca da formação do português brasileiro: i) *a hipótese de mudança natural, ou hipótese internalista* – identificação de tensões no interior da estrutura linguística, ou seja, a língua segue a sua deriva natural, sem motivação externa provocada pelos contatos linguísticos do português europeu; ii) *a hipótese crioulística* – a língua possivelmente sofreu influências fonológicas e gramaticais do substrato indígena e de línguas africanas aqui chegadas durante a escravidão. A segunda delas trabalha em conjunto com a demografia histórica, ou seja, os caminhos que os colonizadores percorreram ao ocupar o território nacional e os contatos linguísticos resultantes dessa ocupação. A união da demografia e do resultado linguístico dos contatos seria o ideal para acessar o que de fato ocorreu, mas há muito pouco interesse da historiografia com relação às características linguísticas: “os historiadores, por via de regra, prestam pouquíssima atenção às maneiras de falar dos povos sobre os quais escrevem” (KIENAN, 1993³ apud MATTOS e SILVA, 2001, p. 281).

O que há documentado são núcleos primários (RIBEIRO, 1914) estabelecidos pelos portugueses no litoral brasileiro. Silva Neto (1986, p. 134-135) os enumera da seguinte maneira: 1) Maranhão e Pará, que geram as unidades administrativas do extremo Norte; 2) Pernambuco, que gera os núcleos secundários da Paraíba, Rio Grande do Norte, Ceará e Alagoas; 3) Bahia, que absorve Ilhéus e Porto Seguro e gera Sergipe; 4) Rio de Janeiro, que em 1658 se emancipou da Bahia, englobava a Paraíba do Sul e tinha superior jurisdição sobre o Espírito Santo, São Vicente e Santo Amaro; 5) São Paulo, donde

² OLIVEIRA, F. *Gramática da linguagem portuguesa*. 1. ed. Lisboa: Biblioteca Nacional, 1981. Edição fac-similada.

³ KIENAN, V. Língua e conquistadores. In: BURKE, F.; PORTER, R. (Org.), p. 258-284, 1993.

evoluiu todo o Oeste (Goiás emancipado em 1744, Minas em 1720, Mato Grosso em 1748).

A incursão de São Paulo para o Sul (PR, SC e RS) deu-se a partir do século XVIII. Em uma classificação cronológica, Castilho (2010, p. 174) enumera o povoamento e a implantação do português em oito focos litorâneos: 1) quatro focos no século XVI: São Vicente/São Paulo (1532, 1554); Olinda/Recife (1535); Salvador (1549); Rio de Janeiro (1557); 2) dois focos no século XVII: São Luís do Maranhão (1612) e Belém (1616); 3) dois focos no século XVIII: Florianópolis (1738) e Porto Alegre (1752). De acordo com os trabalhos supramencionados, as variedades propostas para a análise deste trabalho estão em núcleos diferentes, na classificação de Silva Neto (1986), e aparecem em épocas distintas nos focos de ocupação cronologicamente organizados por Castilho (2010), principalmente as variedades paulista e gaúcha.

A ocupação do território brasileiro não foi fruto de uma colonização bem definida, com metas e normas rígidas. Como consequência disso, nasce a possibilidade de pensar o Brasil como um conjunto de regiões culturais (DIEGUES, 1960), pois cada uma delas é o reflexo das relações estabelecidas entre os povoadores, os povos nativos e a terra ocupada. Em consonância com essa característica, Silva Neto (1963, p. 70) diz que “o extenso território brasileiro é, na verdade, um arquipélago constituído de *ilhas* históricas: Maranhão ou Pará, Pernambuco, Bahia, Rio de Janeiro, São Paulo”. É importante ressaltar que essa delimitação não é discreta, ou seja, as ilhas históricas, nos termos de Silva Neto, e as regiões culturais, nos termos de Diegues, não são justapostas, mas são alteradas ao longo da ocupação.

Para Diegues (1960), uma região cultural seria um conjunto ecológico de pessoas, agregadas pela similaridade das relações sociais, pela unidade das relações espaciais da população, somando a essas a estrutura econômica local. Assim, as regiões culturais seriam espaços territoriais delimitados por certas características que unificam as ideias, os sentimentos e o estilo de vida de um grupo populacional (DIEGUES, 1960, p. 7). As três variedades analisadas neste trabalho estão dispostas em regiões culturais distintas: i) Nordeste Mediterrâneo (Ceará); ii) Pastoril do Extremo Sul (Rio Grande do Sul); iii) Café (São Paulo). Sob o ponto de vista linguístico, as regiões culturais inevitavelmente compartilhariam traços linguísticos particulares, pois o papel da interação social exerce grande influência na variação e mudança linguísticas. Labov (1972) considera que não há como tratar das diferenças linguísticas, sejam elas numa abordagem diacrônica – em caso de mudança –, ou sincrônica – em caso de variação – sem levar em conta a vida social da comunidade e as pressões sociais

existentes nela. Parece, então, que o pressuposto para a categorização tanto sociocultural como sociolinguística faz uso das relações e dos contatos externos e internos para compor uma região cultural, na concepção de Diegues, e uma comunidade de fala (CF), na visão de Labov.⁴

A abordagem clássica trata uma CF como uma distribuição linguística dentro de um espaço social ou geográfico (GUMPERZ, 1968). Outra definição mais recente trata uma CF como “uma área geográfica delimitada por critérios não linguísticos como a demografia ou fronteiras sócio-políticas” (KERSWILL, 1994, p. 23). Já para Bloomfield (1933), a partir de um viés linguístico, uma CF seria um grupo de pessoas que interagem por meio da fala. Neste trabalho, a demarcação das áreas estudadas obedece à distância geográfica, à história, ao povoamento e ao conjunto de características linguísticas vistos em Silva Neto (1976), Noll (2008) e Castilho (2010).

2 Abordagens sociofonéticas e categorização dialetal

O termo sociofonética refere-se a um tratamento descritivo das variações fonéticas presentes nas diversas variedades de uma língua, estabelecendo seus estudos nos limites dos estudos sociolinguísticos e fonéticos. Os estudos sociofonéticos levantaram hipóteses e questões sobre as teorias fonéticas e fonológicas, deixando de distinguir, muitas vezes, *fonética* de *fonologia*. Ohala (1990) contesta a expressão “interface entre fonética e fonologia” argumentando que fonética e fonologia são dependentes entre si e não precisam constituir duas disciplinas autônomas.⁵ Os trabalhos na área têm interesse sobre a relação entre as formas fonético-fonológicas e os fatores sociais, em busca das origens e da transmissão de mudanças linguísticas. Esta subárea da fonética e da sociolinguística tem dado contribuições para além da teoria linguística, levando em consideração as características das variações regionais, sociais e contextuais para a síntese de fala. Além disso, tem auxiliado na descrição de como as variações acontecem e quão sistemáticas elas podem ser, tornando possível gerar sistemas de reconhecimento de fala muito mais robustos (FOULKES, 2006).

⁴ Para Labov, uma CF é definida pela participação de seus integrantes num conjunto de normas compartilhadas pautadas pela uniformidade dos padrões abstratos de variação, levando-se em consideração o conceito de língua como um comportamento social (LABOV, 1972). A dimensão areal das CF também é variável, abrangendo também pequenos agrupamentos urbanos, os quais carregam algum tipo de particularidade em sua fala (FEAGIN, 1996). O termo também foi usado para nomear minorias linguísticas como os falantes de francês de Ontário, no Canadá (MOUGEON; BENIAK, 1996).

⁵ Jakobson (1972, p. 17), apesar de autonomizar a fonologia, deixa claro que “quanto mais o fonólogo se voltar para a abundância dos dados da fonética, tanto mais frutuoso será o seu trabalho; quanto mais substância fonética a fonologia experimentar, tanto melhor”.

O primeiro estudo autointitulado como sociofonético foi o trabalho de Deshaies-Lafontaine (1974, apud FOULKES, 2006), o qual abordou o francês falado no Canadá. Os trabalhos nessa área abarcam vários interesses relacionados às línguas e suas variações, mas têm suas publicações, em sua maioria, direcionadas ao estudo de produção de fala. Embora os estudos de percepção sejam largamente desenvolvidos em estudos nas áreas da fonética e da psicologia, ainda não é tão profícuo na área da sociolinguística. Segundo Thomas (2002), uma das razões pela qual isso ocorre é a tradição impressionística que subjaz os estudos de variação, baseada no desenvolvimento do Alfabeto Fonético Internacional e do sistema de Vogais Cardinais por Daniel Jones. O problema desse tipo de abordagem é que, mesmo que se acredite na fidelidade dos dados transcritos, estes não escapam ao crivo perceptual do pesquisador.⁶

Os trabalhos sociofonéticos que tratam de percepção de fala utilizaram, de maneira geral, palavras ou sentenças as quais foram apresentadas a ouvintes leigos para que pudessem julgar de acordo com o questionamento feito pelo pesquisador, tanto em experimentos de discriminação quanto de identificação. Os experimentos desenvolvidos nesses estudos tinham como objetivo: i) identificar as variedades regionais, etnia ou nível socioeconômico; ii) verificar a influência de estereótipos na percepção de sons; iii) analisar a influência das diferenças dialetais na categorização dos fones; iv) avaliar os traços pessoais baseada na voz (THOMAS, 2002).

3 Metodologia

3.1 Separação segmental-suprasegmental

Para se obter os estímulos utilizados nos experimentos deste trabalho, é necessário o isolamento das características segmentais e suprasegmentais da fala. Para avaliar qual informação os ouvintes podem perceber sem que estejam presentes os segmentos originais de fala, vários métodos foram criados para que pudesse haver uma separação entre segmentos e suprasegmentos. É, porém, importante deter-se ao fato de que esses métodos promovem uma “distorção” em um dos níveis que se pretende estudar. Os trabalhos que lidam com esses tipos de técnica experimental têm como objetivo, em sua maioria, avaliar a identidade do falante e o reconhecimento de línguas (MOFTAH; ROACH, 1988).

Nos estudos de Bush (1967) e de Ohala e Gilbert (1978), as pistas segmentais foram modificadas para que os ouvintes não conseguissem ter acesso aos segmentos que foram produzidos pelos falantes. As técnicas utilizadas

para esse tipo de experimento são chamadas, em termos mais amplos, de fala distorcida. De acordo com os estudos que trabalham com experimentos perceptuais com fala “sem palavras”, há alguns métodos distintos que têm como objetivo retirar a informação segmental da cadeia sonora da fala. Um deles é o sinal laringográfico,⁷ que capta a onda diretamente da fonte, isto é, o som produzido pela laringe. A vantagem desse tipo de método é que o som captado não passou pelas ressonâncias do trato vocal e provavelmente não sofreu nenhuma interferência do filtro. A desvantagem do sinal laringográfico é o seu caráter invasivo, pois é coletado por meio de eletrodos colocados no pescoço do informante. Estudos que lidam com fala espontânea provavelmente serão afetados por conta desse método.

Outro método utilizado é a filtragem passa-baixa na qual é utilizado um filtro que permite passar somente as frequências pré-determinadas pelo pesquisador, abaixo de um limiar previamente estabelecido. O método de filtragem muitas vezes não consegue eliminar por completo os resquícios que os segmentos possam deixar no sinal acústico. Por essa razão, optou-se, neste trabalho, pela utilização do método PURR (*Prosody Unveiling Restricted Representation*), desenvolvido por Sonntag e Portele (1998).⁸ Esse método consiste em retirar do sinal acústico da fala toda a informação segmental relevante, tornando-o ininteligível do ponto de vista segmental. Ao mesmo tempo, o resultado desse processo preserva as características prosódicas como variações de F0, amplitude e duração. Os autores demonstraram, em experimento de reconhecimento, que os informantes conseguiram relacionar frases escritas de mesma estrutura sintática com o estímulo gerado por PURR, mesmo que os segmentos fossem diferentes dos representados na escrita. Os resultados validam a hipótese de que a percepção do estímulo PURR é totalmente separada das qualidades segmentais e que sentenças de estruturas sintáticas semelhantes podem ser comparadas apesar das diferenças lexicais (SONNTAG; PORTELE, 1998).

De maneira oposta, para isolar os segmentos, palavras e demais informações morfológicas e sintáticas, sem que haja variação de F0, propõe-se a utilização de uma curva melódica monotônica (CMM). Com ela, todas as informações segmentais são preservadas, exceto as características entoacionais. A CMM foi obtida por meio do tom médio (TM) de cada estímulo extraído pelo *software* ExProsodia.^{9,10} A partir dos valores de TM,

⁷ *Laryngograph signal* (abreviado como Lx).

⁸ A utilização do PURR foi feita por meio de um *script* para Praat (BOERSMA; WEENINK, 1996), criado por Petra Wagner (2004).

⁹ FERREIRA NETTO, W. ExProsodia. *Revista da Propriedade Industrial – RPII*, Rio de Janeiro, v. 2038, n. 167, p. 167, 2010.

¹⁰ Para uma visão mais abrangente do funcionamento e dos resultados do ExProsodia, ver Ferreira Netto (2016).

⁶ Uma abordagem mais profunda sobre a validade da transcrição fonética pode ser encontrada no trabalho de Kerswill e Wright (1990).

gera-se uma senoide de mesmo valor e duração da frase original. O passo seguinte é criar um arquivo no comando *To manipulation* do Praat, (com tempo de 0.01, mínimo e máximo de 75 Hz e 600 Hz) para a senoide e para o som original. Em seguida, extraem-se as camadas tonais do som original e da senoide utilizando o comando *extract pitch tier*. Por fim, a camada tonal da senoide é sobreposta ao arquivo original através do comando *replace pitch tier*.

3.2 Teoria da Detecção

A Teoria da Detecção – TD – (MACMILLAN; CREELMAN, 2005)¹¹ tem como intuito aferir o desempenho de participantes de experimentos de percepção em que o sinal (visual ou auditivo) deve ser detectado na presença de ruído. Em outras palavras, os participantes são avaliados quanto a sua habilidade de perceber um determinado sinal do ruído. No caso deste estudo, a TD será utilizada para mensurar a habilidade dos participantes em reconhecer a sua variedade, mediante estímulos manipulados com a técnica PURR e com a curva melódica monotônica (CMM). A grande vantagem da TD é o uso do índice d' (*d-prime* ou *d-linha*), que é um quantificador da *sensibilidade* do participante com relação às características associadas ao alvo a detectar, ou seja, d' reflete o grau de separação entre o sinal e o ruído. Para mensurar a sensibilidade do participante, a TD não analisa somente os acertos dos participantes (*hits*), mas leva em consideração também os erros cometidos (*false alarms*). Com essas duas medidas, acertos (H) e falsos alarmes (FA), consegue-se obter o valor d' , que é calculado subtraindo-se a inversa normal da distribuição dos FA da inversa normal dos H: $d' = z(H) - z(FA)$. A TD trabalha com quatro tipos de respostas dadas pelos informantes: *hit*, *miss*, *false alarm* e *correct rejection*. Neste estudo, *hit* (acerto) é o reconhecimento de uma variedade pelo participante (paulistano reconhecendo a variedade paulistana etc.); *miss* (erro) é o não reconhecimento da variedade pelo participante (paulistano dizendo *não* para a variedade paulistana); *false alarm* (alarme falso) é o reconhecimento indevido de uma variedade (paulistano reconhecendo a variedade pelotense como sendo a sua); e *correct rejection* (rejeição correta) é o não reconhecimento de uma variedade da qual o participante não faz parte (paulistano não reconhecendo a variedade pelotense).

Tabela 1. Exemplo de organização dos dados em TD (GREEN; SWETS, 1966)

	Sim	Não
Sim	Hit	False Alarm
Não	Miss	Correct Rejection

Quando os valores de H forem superiores aos de FA, tem-se a possibilidade de avaliar a sensibilidade dos participantes ($d' > 0$), mas se $d' = 0$, as distribuições de H e FA estão sobrepostas e indicam que não há distinção entre o sinal e o ruído. Como exemplo, imagine um experimento com 30 estímulos (15 com sinal presente e 15 com sinal ausente), em que um participante julgou 21 estímulos dos 30 disponíveis como sendo portadores do sinal. Das 21 detecções, 11 foram feitas corretamente (H) e 10 incorretamente (FA). Nessa situação, temos $H \approx FA$ e $d' \approx 0$. Considerando a fórmula acima $d' = z(H) - z(FA)$, temos, com as devidas proporções de H e FA: $d' = z(0,73) - z(0,66)$. O valor de d' é de aproximadamente 0,2, ou seja, valor demasiado baixo (próximo de zero), demonstrando que o participante não foi capaz de diferenciar o sinal do ruído. Como exemplo contrário, considere um participante que apresente uma alta taxa de acertos e uma baixa taxa de falsos alarmes; então $H > FA$ e $d' > 0$. Considerando o mesmo experimento acima (com 30 estímulos), o participante reconheceu 15 estímulos como sendo sinal. Dos 15 estímulos reconhecidos, 14 foram corretamente detectados (H), mas um estímulo foi detectado incorretamente como sinal (FA). Aplicando novamente a fórmula $d' = z(H) - z(FA)$, temos: $d' = z(0,93) - z(0,06)$. O valor de d' é de aproximadamente 3,00, valor que indica uma separação satisfatória entre sinal e ruído. Entretanto, diferentes proporções de H e FA podem resultar no mesmo índice de sensibilidade d' . Essas diferenças, então, são atribuídas a diferentes critérios de especificidade. O índice c é uma medida de critério de reconhecimento que determina o enviesamento dos participantes, i.e., se um participante é mais propenso a detectar o sinal (observador *liberal*), há um aumento da proporção de FA, mas a probabilidade de se obter altos índices de H também aumenta.¹² Por outro lado, se o participante é propenso a não detectar o sinal (observador *conservador*), a taxa de FA será baixa, da mesma forma que a taxa de H decairá. Nota-se, pois, que o critério de decisão influencia diretamente as proporções de H e FA. O valor de c é calculado por $c = \frac{1}{2}(zF + zH)$. Quando $c = 0$, não há enviesamento; quando $c < 0$, temos um critério liberal, e quando $c > 0$, temos um critério conservador:

Participante 1

$$d' = z(H) - z(FA)$$

$$d' = z(0,67) - z(0,03)$$

$$d' = 2,35$$

Participante 2

$$d' = z(H) - z(FA)$$

$$d' = z(0,83) - z(0,08)$$

$$d' = 2,35$$

¹¹ A TD é baseada na Teoria da Detecção do Sinal (GREEN; SWETS, 1966).

¹² O índice β é um índice de enviesamento mais popular na literatura, mas muitos autores têm recomendado o c , dentre outras razões, por β ser influenciado por variações de d' .

Critério de decisão: c

Participante 1

$$c = \frac{1}{2}(zF + zH)$$

$$c = \frac{1}{2}[z(0,67) + z(0,03)]$$

$$c = 0,74$$

Participante 2

$$c = \frac{1}{2}(zF + zH)$$

$$c = \frac{1}{2}[z(0,83) + z(0,08)]$$

$$c = 0,21$$

No exemplo acima, o participante 1 teve um critério ligeiramente conservador se comparado ao participante 2, que apresentou um viés moderado (próximo de 0).

3.3 Desenvolvimento e aplicação dos experimentos

3.3.1 Experimento 1 com técnica PURR

Para o experimento 1 – enunciados delexicalizados com a técnica PURR – foram utilizadas amostras de fala adulta do sexo masculino divididas entre nove informantes (distribuídas igualmente entre as três variedades escolhidas [SP, RS e CE]). Os informantes gravados nasceram e residem em suas respectivas cidades, sendo muito breves e raras as suas ausências dos locais de nascimento. Os detalhes de cada sujeito estão expostos no Quadro 1, a seguir.

Quadro 1. Relação de características por informante.

Informantes	Idade	Estado	Escolaridade
1	32	SP	2º Grau completo
2	34	SP	Superior completo
3	62	SP	Superior completo
4	31	RS	2º Grau completo
5	34	RS	Superior completo
6	43	RS	2º Grau completo
7	30	CE	2º Grau completo
8	58	CE	2º Grau incompleto
9	59	CE	2º Grau incompleto

Ao todo, foram selecionados seis trechos fluentes de fala espontânea de cada sujeito (três trechos curtos [TC] e três trechos longos [TL]). A duração dos trechos curtos escolhidos para o teste foi de até 10 segundos, e a dos trechos longos entre 15 e 25 segundos. As gravações de São Paulo (Mooca) e do Ceará (Senador Pompeu) foram feitas com gravador digital ZOOM H2, com taxa de 44.1 kHz. As gravações do Rio Grande do Sul (Pelotas) fazem parte do banco de dados VARX (AMARAL; BORGES, 2001), também com taxa de 44.1 kHz.

Houve um breve treino em que os ouvintes se familiarizaram com o que é entoação. Todas as instruções para a feitura do experimento foram gravadas para que os informantes ouvissem sempre a mesma explicação, evitando, dessa maneira, algum tipo de enviesamento na aplicação dos experimentos. Os informantes também puderam acompanhar as explicações por meio de uma versão escrita. Para a fase de familiarização, uma frase de cada variedade foi apresentada em sua versão não manipulada e manipulada com PURR, com o intuito de fazer o ouvinte entender o procedimento. Em seguida, 13 ouvintes de cada uma das localidades ouviram duas sequências de estímulos, a primeira constituída de 27 trechos curtos (TC) das 3 variedades em questão, a segunda constituída de 27 trechos longos (TL) das 3 variedades, ambos dispostos aleatoriamente. Os participantes receberam um formulário para ser preenchido durante o experimento. A cada estímulo, tiveram de que marcar com um 'x' a resposta *sim*, caso reconhecessem o estímulo como sendo a sua variedade linguística, ou *não*, caso não o reconhecessem. O teste foi aplicado individualmente, em lugar silencioso e com a utilização de fone de ouvido Sennheizer 220HD. No total, foram 2106 respostas. Os ouvintes que participaram do experimento possuem idade entre 20 e 55 anos, de ambos os sexos, sem problemas auditivos e alfabetizados.

3.3.2 Experimento 2 com curva melódica monotônica (CMM)

No experimento 2 – enunciados com curva melódica monotônica (CMM) – foi utilizada parte das mesmas amostras de fala do experimento 1. Os ouvintes que participaram do experimento 2 são os mesmos que participaram do experimento 1, obedecendo à mesma metodologia descrita no experimento com PURR. Para a fase de familiarização, uma frase de cada variedade foi apresentada em sua versão não manipulada e monotônica. Após a familiarização, 13 ouvintes de cada uma das localidades ouviram uma sequência aleatória de 18 estímulos constituída de 6 trechos de cada uma das 3 variedades em questão.

4 Análise de dados

4.1 Análise com a Teoria da Detecção

A Tabela 2 apresenta os resultados de cada informante para cada variedade no experimento (PURR), envolvendo trechos curtos e longos.

Os resultados demonstram que há uma certa variabilidade nos valores d' obtidos. Alguns informantes apresentaram maior sensibilidade para a discriminação da própria variedade, bem como a negação das variedades

Tabela 2. Valores de d' e c para cada informante de cada variedade no experimento 1. Os trechos longos e curtos foram abreviados como TL e TC, respectivamente.

CE					SP					RS				
Inf.	TC		TL		Inf.	TC		TL		Inf.	TC		TL	
	(d')	(c)	(d')	(c)		(d')	(c)	(d')	(c)		(d')	(c)	(d')	(c)
1	1.72	0.095	1.02	0.075	1	1.02	0.079	1.36	0.54	1	0.71	-0.074	0.57	0.146
2	0.76	-0.382	0.57	0.146	2	1.11	0.414	0.87	0.014	2	1.11	0.414	1.4	0.268
3	1.66	0.393	1.53	0.004	3	0.71	-0.074	0.9	0.312	3	0.86	-0.009	0.42	0.071
4	1.61	0.035	0.71	-0.074	4	1.36	-0.095	0.45	0.365	4	1.2	0.167	1.08	0.689
5	0.57	0.144	0.028	0.154	5	0.57	0.144	0.72	-0.079	5	0.51	-0.509	0.14	-0.07
6	1.2	-0.166	0.9	-0.311	6	1.2	0.167	1.12	0.408	6	0.86	-0.009	1.05	-0.241
7	1.35	-0.238	0.57	0.146	7	1.53	0.021	1.2	-0.167	7	0.57	0.15	1.2	-0.167
8	1.05	-0.246	1.42	-0.571	8	0.9	0.316	0.86	0.018	8	1.4	0.262	0.86	0.033
9	0.9	0.311	0.42	0.064	9	1.05	-0.241	0.28	0.044	9	0.58	0.15	0.28	0.006
10	1.65	-0.393	0.57	-0.146	10	1.4	0.268	1.5	-0.469	10	1.2	0.167	1.5	-0.469
11	0,04	0.126	-0.43	0.215	11	1.36	-0.54	1.02	0.079	11	1.36	-0.54	1.02	0.076
12	1.05	-0.246	0.42	0.071	12	0.57	0.146	0.71	-0.076	12	1.73	0.095	0.71	-0.074
13	0.57	-0.144	1.29	-0.12	13	0.83	0.554	1.03	0.076	13	0.71	-0.074	0.73	0.225
Média	1.08	-0.055	0.69	-0.027	Média	1.05	0.089	0.92	0.082	Média	0.98	0.014	0.84	0.038
D.P.	0.52	0.246	0.55	0.211	D.P.	0.32	0.276	0.34	0.26	D.P.	0.38	0.264	0.42	0.267

alheias. Quanto maior o valor d' , melhor foi a tarefa de reconhecimento. Os informantes senadorense (CE) 1, 3, 4, 7, 8, 10 e 12, em trechos curtos, e 1, 3, 6, 8, e 13, em trechos longos, tiveram um desempenho satisfatório.

De maneira análoga, os informantes paulistanos (SP) 1, 2, 4, 6, 7, 9 e 11, em trechos curtos, e 1, 6, 11 e 13, em trechos longos, tiveram um bom desempenho no reconhecimento de sua variedade. O mesmo aconteceu com os informantes pelotenses (RS) 2, 4, 10, 11 e 12, em trechos curtos, e 2, 4, 6, 10 e 11, em trechos longos. Por outro lado, pode-se observar que alguns informantes não tiveram um desempenho satisfatório quanto ao reconhecimento, seja pelo não reconhecimento da própria variedade ou pelos alarmes falsos na aceitação de variedades alheias. Nesses casos, o valor d' ficou próximo de 0 ou negativo quando a taxa de alarmes falsos superou a taxa de acertos. Exemplos de baixo desempenho podem ser vistos nos informantes senadorense 5 e 11, nos informantes paulistanos 5 e 12 e nos informantes pelotenses 5 e 9, em trechos curtos e longos.

A Tabela 3, com os resultados de cada participante para o experimento (CMM).

Os resultados demonstram que não houve a mesma variabilidade encontrada no experimento 1 (PURR). Os valores d' obtidos mantiveram-se altos com relação aos índices do experimento 1. Parece que a tarefa de reconhecimento das variedades linguísticas torna-se mais difícil quando os segmentos estão ausentes.

Ao analisarmos os valores do critério c , nota-se que, de maneira geral, os valores foram próximos de zero, indicando um viés quase neutro. Isso indica que as probabilidades de sinal e de ruído são equivalentes.

Entende-se por isso que a probabilidade das respostas dadas pelos participantes não apresenta, de forma geral, enviesamento para respostas *sim* ou *não*. Em outras palavras, os participantes não estão propensos a dizer *sim* (critério liberal) toda vez que um estímulo lhes é apresentado. O oposto também é verdadeiro, ou seja, tampouco estão inclinados a dizer *não* (critério conservador) quando avaliam um estímulo.

Tabela 3. Valores de d' e c para cada informante de cada variedade no experimento 2.

CE			SP			RS		
Inf.	(d')	(c)	Inf.	(d')	(c)	Inf.	(d')	(c)
1	1,64	-0.140	1	2,35	0.208	1	2,35	0.208
2	1,27	0.205	2	2,35	0.187	2	2,35	0.189
3	2,54	-1.058	3	1,93	0.003	3	2,35	0.206
4	3,29	-0.679	4	2,76	0.948	4	2,76	0.948
5	1,18	-0.378	5	0,97	0.497	5	0,97	0.484
6	1,4	0.268	6	1,93	0.035	6	1,93	-0.014
7	3,29	0.679	7	2,76	0.948	7	1,93	-0.02
8	1,4	-0.268	8	2,35	0.208	8	2,76	0.948
9	0,43	-0.215	9	1,81	0.476	9	1,81	0.476
10	0,86	-0.002	10	0,43	0.215	10	0,43	0.215
11	1,71	0.425	11	1,93	0.003	11	1,93	-0.026
12	1,93	-0.001	12	1,93	0.014	12	1,93	-0.005
13	1,93	0.014	13	1,81	0.476	13	1,81	0.477
Média	1,76	-0.088	Média	1,95	0.324	Média	1,95	0.314
D.P.	0,86	0.44	D.P.	0,65	0.316	D.P.	0,65	0.325

Apesar dessa visão geral, os valores ligeiramente conservadores e moderados de c apontaram, muitas vezes, para valores mais elevados de d' . Como exemplo, temos o informante senadorense 3 avaliando trechos curtos delexicalizados ($d'=1,66$ e $c=0,393$). Nesse

caso, o informante apresenta uma ligeira tendência em responder *não* aos estímulos. Há também um exemplo diametralmente oposto como o informante senadorense 10 ($d'=1,65$ e $c=-0,393$), o qual apresentou um valor de c ligeiramente liberal, ou seja, com uma tendência maior a dar respostas positivas. No entanto, ambos apresentaram valores d' quase idênticos. O mesmo pode ser visto nos dados de SP e RS. No experimento 2, essa tendência fica mais evidente, apesar de haver exceções. O informante paulistano 7 e o informante pelotense 4 apresentaram um critério c conservador e um valor d' elevado ($d'=2,76$ e $c=0,948$ para ambos). O informante senadorense 7 apresentou resultado semelhante ($d'=3,29$ e $c=0,679$). Há, como exemplo contrário, o informante paulistano 3 ($d'=1,93$ e $c=0,003$) e o informante pelotense 12 ($d'=1,73$ e $c=0,095$), os quais apresentaram valores elevados de d' , mas praticamente nenhum viés de decisão.

4.2 Análise geral dos dados do experimento 1 – fala delexicalizada

O desempenho dos três grupos de participantes foi o seguinte: i) os informantes de Senador Pompeu (CE) obtiveram 443 (63%) respostas certas contra 259 erradas; ii) os informantes de São Paulo (Mooca) obtiveram 483 (69%) respostas certas contra 219 erradas; iii) os informantes de Pelotas (RS) obtiveram 468 (67%) respostas certas contra 234 erradas. Para uma análise probabilística dos dados, foi feito um teste ANOVA de um fator, envolvendo os dados brutos de acertos e a duração dos estímulos do experimento 1 (PURR). O resultado não foi significativo: $F(1, 72)=0,63$, $p>0,05$. Com isso, o fator duração de estímulo parece não exercer influência no desempenho dos informantes. Um outro teste ANOVA de 1 fator foi feito, agora tendo como fator a variedade dos informantes. O resultado foi significativo: $F(2, 75)=4,04$, $p<0,05$. Para constatar qual era a diferença significativa entre as variedades, foi feito um teste *post hoc* de Tukey. O resultado apontou uma diferença significativa entre as variedades de SP e CE, e entre SP e RS ($p<0,05$).

Num segundo momento da análise do experimento 1 (PURR), foi feito um teste ANOVA de 1 fator, mas com os valores d' apresentados anteriormente (Tab. 2). O fator foi a duração de estímulos e o resultado obtido foi $F(1,76)=5,14$, $p<0,05$. Os índices d' demonstraram que o desempenho dos informantes pode ter sido influenciado pela duração dos estímulos, ao contrário do resultado encontrado nos dados brutos. Este resultado pode ser atribuído à utilização do índice d' , pois fornece uma informação mais detalhada do desempenho de cada informante por levar em consideração também os erros cometidos.

4.3 Análise geral dos dados do experimento 2

O desempenho dos três grupos de participantes no experimento 2 foi o seguinte: i) os informantes de Senador Pompeu (CE) obtiveram 179 (76,5%) respostas certas contra 55 erradas; ii) os informantes de São Paulo (Mooca) obtiveram 182 (78%) respostas certas contra 52 erradas; iii) os informantes de Pelotas (RS) obtiveram 194 (83%) respostas certas contra 40 erradas. Seguindo a mesma metodologia de análise do experimento 1, foi analisado o papel da variável “variedade regional” sobre o nível de acertos e índices d' . A análise prévia de resíduos, via teste de normalidade de Shapiro-Wilk, demonstrou que a distribuição dos acertos não era normal ($p<0,05$). A análise de homocedasticidade, via teste de Fligner-Killeen, apontou para uma distribuição não homogênea das variâncias ($p<0,05$). Por conta disso, foi feito um teste não paramétrico de Kruskal-Wallis. O resultado não foi significativo, ou seja, os acertos parecem não ter um desempenho diferenciado de acordo com as variedades. O modelo de análise com índice d' também não cumpriu os pressupostos da ANOVA. Assim, foi feito um teste de Kruskal-Wallis, que apresentou resultado não significativo ($p>0,05$).

4.4 Análise de produção

A análise de produção feita neste trabalho partiu dos dados fornecidos via *ExProsodia*. Três características foram abordadas: i) F0 final dos estímulos, ou finalização (Hz); ii) tom médio dos estímulos (Hz); iii) dispersão das variações de F0 (Hz). No primeiro passo da análise, valores de F0 final dos estímulos, o teste de Kruskal-Wallis forneceu um resultado não significativo ($p>0,05$). No segundo passo, tom médio entre as variedades, um teste ANOVA de 1 fator apresentou resultado significativo: $F(2,51)=17,09$, $p<0,05$. O resultado apontou uma diferença significativa entre os tons médios das variedades regionais. O teste de Tukey apontou para uma diferença significativa entre as variedades paulistana e senadorense, e entre as variedades pelotense e paulistana ($p<0,05$). No terceiro passo, dispersão de F0, um teste ANOVA de um fator apresentou resultado significativo: $F(2, 51)=18,04$, $p<0,05$. O teste de Tukey apresentou novamente diferença significativa entre as variedades paulistana e senadorense, de um lado, e pelotense e paulistana, de outro ($p<0,05$). Em resumo, das três possíveis características que poderiam estar influenciando a escolha dos informantes, o tom médio dos estímulos e a dispersão de F0 apresentaram resultado significativo. O Quadro 2 resume os resultados dos testes estatísticos feitos até então.

Quadro 2. Resumo das análises.

	<i>Testes</i>	<i>Resultados</i>
<i>Experimento 1 (PURR)</i>	Anova – 1 fator – v. dependente - acertos; v. independente - duração dos estímulos (TL e TC)	$F(1,72) = 0,63, p > 0,05$
	Anova – 1 fator – v. dependente - acertos; v. independente – variedades regionais	$F(2,75) = 4,04, p < 0,05$
	Tukey – acertos e variedades regionais	Diferenças significativas entre as variedades SP e CE, SP e RS, $p < 0,05$
	Anova – 1 fator – v. dependente – σ^2 ; v. independente – duração dos estímulos (TL e TC)	$F(1,76) = 5,14, p < 0,05$
<i>Experimento 2 (CMM)</i>	Kruskal-Wallis – v. dependente - acertos; v. independente – variedades regionais	$p > 0,05$
	Kruskal-Wallis – v. dependente – σ^2 ; v. independente – variedades regionais	$p > 0,05$
<i>Experimentos 1 e 2</i>	Kruskal-Wallis – v. dependente – acertos; v. independente – tipo de experimento (PURR e CMM)	$p < 0,05$
<i>Análise de Produção</i>	Kruskal-Wallis – v. dependente – F0 final; v. independente – variedades regionais	$p > 0,05$
	Anova – 1 fator – v. dependente – F0 médio; v. independente – variedades regionais	$F(2,51) = 17,09, p < 0,05$
	Tukey – F0 médio e variedades regionais	Diferenças significativas entre as variedades SP e CE, SP e RS, $p < 0,05$
	Anova – 1 fator – v. dependente – dispersão de F0; v. independente – variedades regionais	$F(2,51) = 18,04, p < 0,05$
	Tukey – dispersão de F0 e variedades regionais	Diferenças significativas entre as variedades SP e CE, SP e RS, $p < 0,05$

5 Discussão

A primeira questão deste trabalho, sobre a capacidade dos ouvintes em reconhecer suas próprias variedades em meio às demais e de forma atípica (PURR e CMM), foi respondida de maneira positiva. Assim, os resultados deste estudo sobre o português brasileiro corroboram os estudos anteriores. Entretanto, responder a essa questão não traz à luz o que ocasionaria o reconhecimento das variedades regionais. A ANOVA (acertos e variedades regionais) apresentou diferenças significativas de reconhecimento, especificadas pelo teste de Tukey entre as variedades de São Paulo e do Ceará, e entre as do Rio Grande do Sul e de São Paulo. Essa primeira divisão no desempenho já é um indício de que os informantes de São Paulo se comportaram de maneira diferente dos demais. Na análise de produção, os valores de tom médio parecem se organizar de maneira análoga à performance dos informantes. Outra vez, a variedade de São Paulo ficou isolada, diferenciando-se da do Ceará e do Rio Grande do Sul. Os resultados da análise de dispersão de F0 também apresentaram semelhança à análise de tom médio,

diferenciando a dispersão de F0 da variedade de São Paulo das demais. O que parece haver de sistemático, quando os resultados são cruzados, é uma distribuição espelhada entre percepção e produção, isolando a variedade senadoreense num extremo, a variedade pelotense numa posição central, mas ainda assim um pouco mais próxima da senadoreense, e a variedade paulistana noutro extremo.

Se os senadoreenses são distintos dos paulistanos, mas nem tanto dos pelotenses, como seriam os alarmes falsos deles? Da mesma forma, qual seria a tendência dos alarmes falsos dos demais? Verificar os alarmes falsos de cada variedade é uma maneira de corroborar a análise acima ou de colocar um ponto de interrogação na influência das características tidas, até então, como prováveis pistas utilizadas na escolha das variedades. A hipótese seria de um aumento no número de alarmes falsos quando o informante se deparasse com uma variedade semelhante à sua. Os alarmes falsos dos informantes pelotenses foram 31% direcionados à variedade paulistana e 69% à senadoreense. Os alarmes falsos dos paulistanos foram 34% direcionados à variedade senadoreense e 66% à pelotense. Por fim, os senadoreenses apresentaram 48%

dos alarmes falsos direcionados à variedade paulistana e 52% à pelotense. Um teste ANOVA para os alarmes falsos dos informantes pelotenses teve como resultado $F(1,34)=12,98$, $p<0,05$. Um teste de Kruskal-Wallis para os alarmes falsos cometidos pelos informantes paulistanos teve como resultado $p<0,05$.

No caso dos paulistanos, houve um enviesamento nas respostas, quando incorretas, para uma variedade regional mais próxima do ponto de vista geográfico (a dos pelotenses). Somente os alarmes falsos cometidos pelos senadorenses não condizem com essa tendência ($p>0,05$). Em outras palavras, a hipótese da tendência dos alarmes falsos foi mantida no desempenho das variedades paulistana e pelotense. No caso da variedade senadorenses, parece não haver direção para os equívocos. A relação que se pôde estabelecer entre tom médio, dispersão de F0 e a tendência dos alarmes falsos foi significativa, mas este trabalho tem a ciência de que os resultados obtidos aqui necessitam de um estudo mais detalhado, sobretudo na produção dos estímulos, abarcando os demais parâmetros acústicos da entoação. Um outro refinamento possível da análise seria a manipulação dos parâmetros de F0 estudados neste trabalho e a replicação do experimento 1, ou seja, a manipulação dos tons médios e da dispersão de F0 e compará-los com os resultados deste trabalho. Tal procedimento reafirmaria o papel do tom médio e da dispersão de F0 na percepção das variedades regionais em questão, uma vez que a mudança dessas características deveria, em tese, ocasionar uma mudança na percepção. Com relação ao ritmo, poder-se-ia aplicar um teste com CMM e PURR combinados de maneira que somente as durações pudessem ser percebidas. Para os fins deste trabalho, a entoação pode ser encarada como parte determinante no reconhecimento das variedades regionais aqui estudadas. A análise de produção conseguiu evidenciar pistas relevantes nas características de F0, as quais podem ser consideradas, nos limites experimentais deste trabalho, como responsáveis pelo reconhecimento das variedades regionais do PB por falantes brasileiros.

Referências

- AMARAL, Luís; BORGES, Paulo. *Banco de dados sociolinguísticos variáveis por classe social – VarX*. Pelotas: UFPel, 2001.
- BUSH, Clara. Some acoustic parameters of speech and their relationships to the perception of dialect difference. *TESOL Quarterly*, v. 1, n. 3, p. 20-30, 1967.
- ATKINSON, Kay. Language identification from nonsegmental cues. *Working Papers in Phonetics* (UCLA), 10, p. 85-89, 1968.
- BOERSMA, Paul; WEENINK, David. Praat, a system for doing phonetics by computer, version 5.1.10. *Institute of Phonetic sciences of the University of Amsterdam*, Report, p. 132-182, 1996.
- CASTILHO, Ataliba T. de. *Nova Gramática do Português Brasileiro*. São Paulo: Contexto, 2010.
- CLOPPER, Cynthia G.; PISONI, David B. Perception of Dialect Variation. In: PISONI, David B.; REMEZ, Robert E. (Ed.). *The Handbook of perception*. Oxford: Blackwell, 2004. p. 314-337.
- DIEGUES JÚNIOR, Manuel. *Etnias e culturas no Brasil*. Rio de Janeiro: Biblioteca do Exército, 1960.
- FEAGIN, Crawford. Peaks and Glides in Southern States Short-a. In: GUY, Gregory R.; FEAGIN, Crawford; SCHIFFRIN, Deborah; BAUGH, John (Ed.). *Towards a Social Science of Language: Papers in Honor of William Labov*. Vol. 1. Variation and Change in Language and Society. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996.
- FERREIRA NETTO, Waldemar. *ExProsodia*. Resultados Preliminares. São Paulo: Paulistana, 2016.
- FOULKES, Paul. Sociophonetics. In: BROWN, Keith (Ed.). *Encyclopedia of Language and Linguistics*. 2. ed. Amsterdam: Elsevier Press, 2005. p. 495-500.
- GREEN, David. M.; SWETS, John A. *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.
- GUMPERZ, John. The speech community. In: *International Encyclopedia of the Social Sciences*. London: Macmillan, 1968. p. 381-386.
- KERSWILL, Paul. *Dialects converging: Rural speech in urban Norway*. Oxford: Oxford University Press, 1994.
- KERSWILL, Paul; WRIGHT, Susan. On the limits of auditory transcription: a sociophonetic perspective. *Language Variation and Change*, v. 2, p. 255-75, 1990.
- JAKOBSON, Roman. *Fonema e fonologia*. Rio de Janeiro: Livraria Acadêmica, 1972.
- LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- MACMILLAN, Neil A.; CREELMAN, C. Douglas. *Detection Theory: a user's guide*. New Jersey: LEA, 2005.
- MAIDMENT, John. Voice fundamental frequency characteristics as language differentiators. *Speech and Hearing Working in Progress*, London, University College London, 1976. p. 74-93.
- MATTOS e SILVA, Rosa Virgínia. *Para a história do português brasileiro*. Vol. 2. Primeiros Estudos. São Paulo: Humanitas/Fapesp, 2001.
- MOUGEON, Raymond; BENIAK, Édouard. Social class and language variation in bilingual speech communities. In: GUY, Gregory R.; FEAGIN, Crawford; SCHIFFRIN, Deborah; BAUGH, John. (Eds.). *Towards a Social Science of Language: Papers in Honor of William Labov*. Vol. 1. Variation and Change in Language and Society. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996.
- NETO, Serafim da Silva. *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro: Presença, 1986.
- NOLL, Volker. *O português brasileiro: formação e contrastes*. São Paulo: Globo, 2008.
- OHALA, John. J.; GILBERT, Judy B. Listener's ability to identify languages from their prosody. *Report of the Phonology*

Laboratory. Berkeley: University of California, v.2, p. 126-132, 1978.

OHALA, John J. There is no interface between phonetics and phonology: a personal view. *Journal of Phonetics*, v. 18, p. 153-171, 1990.

RIBEIRO, João Batista. *História do Brasil*. Curso Superior. Rio de Janeiro: Editora Francisco Alves, 1914.

RICHARDSON, Joyce A. C. *The Identification by Voice of Speakers Belonging to Two Ethnic Groups*. 1973. Dissertação (Mestrado) – Ohio State University, Columbus, 1973.

SONNTAG, Gerit P.; PORTELE, Thomas. PURR – a method for prosody evaluation and investigation. *Computer Speech and Language, Special Issue on Evaluation*, v. 12, n. 3, p. 437-451, 1998.

THOMAS, Erik R. Sociophonetics applications of speech perception experiments. *American Speech*, v. 77, n. 2, p. 115-146, 2002.

Recebido: 30 de setembro de 2016

Aprovado: 07 de março de 2017

Contato: danielperes@usp.br