

Línguas diferentes, vozes distintas: evidências da fala de bilíngues em português e inglês

Ana Paula Petriu Ferreira Engelbert^{*}
Universidade Tecnológica Federal do Paraná
Curitiba, PR, BR

Denise Cristina Kluge^{**}
Universidade Federal do Paraná
Curitiba, PR, BR

Adelaide Hercília Pescatori Silva^{***}
Universidade Federal do Paraná
Curitiba, PR, BR

Resumo

A qualidade de voz resulta da modulação dos articuladores do trato vocal e da configuração laríngea, produzindo um efeito de longo-termo na produção e percepção da fala. Bilíngues podem alterar a qualidade de voz quando falam línguas diferentes, conforme apontam estudos na área. O presente artigo apresenta resultados referentes à produção da voz em português brasileiro (PB) e inglês (IN) por falantes bilíngues brasileiros. Para chegar a tais resultados, foi realizado um experimento de produção que consistiu na gravação de emissões em PB e em IN de tarefas de leitura de textos, bem como amostras de fala semiespontânea. Os resultados mostram diferenças estatisticamente significativas entre médias e extensão de f_0 entre as línguas, bem como entre medidas de declínio espectral (LTAS) entre línguas e tarefas.

Palavras-chave: Qualidade de Voz; Fala Bilíngue; Análise Acústica

Different languages, different voices: evidence from Portuguese/English bilinguals' speech

Abstract

Voice quality results from the modulation of vocal tract articulators and laryngeal activity and produces a long-term effect in speech production and perception. Bilingual speakers can alter voice quality when speaking different languages, according to evidence shown in research in the field. The present study shows results from a voice production experiment in Brazilian Portuguese (BP) and English (EN) by Brazilian bilingual speakers. The experiment consisted of the recordings of reading and semi-spontaneous tasks in BP and EN. Results show statistically significant differences between languages in f_0 measures, as well as in spectral characteristics (LTAS) between languages and tasks.

Keywords: Voice Quality; L2 Speech; Acoustic Analysis

^{*} Ana Paula Petriu Ferreira Engelbert é professora adjunta no Departamento de Línguas Estrangeiras Modernas da Universidade Tecnológica Federal do Paraná (UTFPR). Possui doutorado em Linguística pela Universidade Federal do Paraná com período sanduíche na University of California, Los Angeles (2015). Tem interesse em pesquisa sobre a aquisição de L2, especialmente a fonologia do inglês como L2, qualidade de voz e fonética acústica. Email: anapaulapetriu@gmail.com, anapetriu@utfpr.edu.br

^{**} Denise Cristina Kluge é professora adjunta da Universidade Federal do Paraná (UFPR) vinculada ao Departamento de Letras Estrangeiras Modernas (DELEM) e ao Programa de Pós-graduação em Letras. Possui doutorado em Letras Inglês pela Universidade Federal de Santa Catarina (2009). Seus interesses de pesquisa são a percepção e produção de sons, efeito de pistas visuais, treinamento perceptual, aquisição/aprendizagem de segunda língua e ensino de pronúncia. Email: deniseckluge@gmail.com

^{***} Adelaide Hercília Pescatori Silva é professora associada da Universidade Federal do Paraná (UFPR). Possui doutorado em Linguística pela Universidade Estadual de Campinas (2002). Tem experiência na área de Linguística, com ênfase em Teoria e Análise Linguística, atuando principalmente nos seguintes temas: fonética, fonologia, análise acústica, róticos e modelos dinâmicos de produção da fala. Email: adelaidehpsilva@gmail.com



1. Introdução

Pesquisas mostram que há algo de diferente em nossas vozes quando falamos uma língua estrangeira em comparação com quando falamos nossa língua materna (BRUYNINCKX *et al.*, 1991, 1994; HARMEGNIES *et al.*, 1991; TODAKA, 1995; NG *et al.*, 2012; CAMARGO *et al.*, 2013). A presença de tais diferenças na voz pode ser motivada por fatores culturais da comunidade linguística na qual o falante está inserido, ou por questões de personalidade, identidade e aspectos emocionais do falante (ESLING, 2013). As pesquisas supracitadas mostram que bilíngues realizam mudanças na voz pelo uso de diferentes tipos de fonação, bem como média e extensão de frequência fundamental (f_0) distintas. Assim, o principal objetivo do presente estudo foi comparar amostras de fala de brasileiros bilíngues falantes de português (PB) como L1 e inglês (IN) como L2. As comparações realizadas tinham como objetivo verificar diferenças nas amostras entre as línguas e entre as tarefas desempenhadas pelos participantes (leitura de texto e fala semiespontânea). Além disso, foi investigada uma possível correlação entre variáveis de aquisição (idade de aquisição e tempo de experiência em país em que se fala o IN como L1) na produção da fala do IN. Com base nos resultados encontrados por estudos anteriores, a nossa hipótese constava da presença de diferenças entre as amostras de fala no PB e no IN no que concerne às características acústicas espectrais e de f_0 dos falantes brasileiros bilíngues. Esperou-se encontrar diferenças entre as línguas e entre as tarefas desempenhadas, bem como alguma influência das variáveis de aquisição do IN na qualidade de voz dos brasileiros bilíngues.

2. Qualidade de voz

A definição dos termos ‘voz’ e ‘qualidade de voz’ não é consensual, mas, em linhas gerais, estão relacionados a características resultantes da modulação dos articuladores do trato vocal que produzem um efeito de longo-termo na produção da fala e sua percepção pelo ouvinte. Kreiman e Sidtis (2011, p. 6) afirmam que os termos ‘voz’ e ‘qualidade de voz’ são duas faces de uma

mesma moeda: o primeiro está relacionado às características da produção da fala, e o segundo, às da percepção. Por meio da configuração do trato vocal e, consequentemente, da produção de diferentes características vocais, é possível ajustarmos nossa voz a determinadas situações comunicativas. Isso acontece, pois há muitas outras informações na fala para além do conteúdo linguístico que se emprega durante a enunciação. Como falantes, fornecemos ao nosso ouvinte informações sobre quem somos, de onde viemos e o que estamos sentindo. Como ouvintes, também somos capazes de inferir uma série dessas características a partir da voz do falante. A produção da voz também é determinada pela composição fisiológica do trato, já que o formato e posição dos articuladores como os lábios, língua e mandíbula influenciam diretamente na realização dos sons da fala. Neste artigo, focalizamos as características vocais da produção da voz de brasileiros bilíngues, considerando, contudo, a indissociável relação entre produção e percepção.

A qualidade de voz pode ter tanto função linguística,¹ ou seja, que implique contrastes em uma dada língua, como paralinguística, quando fornece informações físicas e psicossociais do falante. A paralingua é responsável por veicular informações relacionadas às características físicas e psicossociais do falante, bem como suas atitudes e emoções na produção do enunciado. Todos esses aspectos fazem parte do que são consideradas as características indiciais² do falante. Segundo Pike (1945), as atitudes e intenções veiculadas na comunicação oral são pelo menos tão importantes quanto às informações linguísticas e, às vezes, têm mais importância que estas.

A literatura convencionou discernir as características indiciais da qualidade de voz em duas partes: as que são “orgânicas” e as que são “adquiridas”. As primeiras são aquelas resultantes da configuração anatômica do falante, e que, consequentemente, são relativamente imutáveis, como média de frequência fundamental (f_0) e frequência de formantes. Por serem motivadas fisiologicamente, as características orgânicas da voz tendem a ser estáveis para falantes adultos, fornecendo bons índices de identidade, sexo e idade (KREIMAN e SIDTIS, 2011, p. 65). As últimas são padrões habituais que o

falante aprende com a experiência linguística, que incluem o sotaque, taxa de elocução, curvas entoacionais, f_0 habitual e qualidade de voz. Essas características são cruciais para sinalizar o vínculo à determinada comunidade linguística e para fornecer atributos pessoais. Portanto, a voz é o produto da composição fisiológica e da origem sociolinguística do falante. No presente trabalho, acredita-se que as características vocais adquiridas sejam as responsáveis pelas diferenças na produção da voz entre L1 e L2, considerando-se a forte influência de fatores como motivação, atitude e emoção no processo de aquisição de uma L2.

3. Medidas acústicas de qualidade de voz

Variações no modo em que as pregas vocais vibram podem produzir mudanças perceptíveis na maneira como nossa voz soa. O tipo de fonação mais frequentemente produzido por nós seres humanos é a voz modal, resultado da ação relativamente periódica das pregas vocais. Ladefoged (1971) propôs um modelo simplificado de representação dos tipos de fonação por um contínuo de constrição glotal. Nesse modelo, o tamanho da glote, que depende da distância entre as pregas, pode variar do ponto em que elas estão em abertura máxima (como em uma fricativa glotal) até o ponto em que estão completamente fechadas (como em uma oclusiva glotal). Em ambos os extremos, não há fonação. Nos estágios intermediários de abertura entre as pregas, momentos em que há fonação, ocorrem a voz soprosa (*breathy*), voz modal (*modal*) e voz crepitante (*creaky*) (FIGURA 1):



Figura 1: Contínuo de fonação.

Fonte: Ladefoged (1971, *apud* Gordon e Ladefoged, 2001, p. 384)

Segundo Keating e Esposito (2007, p. 85), por conta da utilização do contínuo na representação do modelo, é possível conceber a existência de graus não somente

de sopro e crepitação, mas também de graus de voz modal. Como consequência, não há categorias discretas na fonação. Além disso, não há correspondência interlinguística na produção da fonação (o que é considerado voz soprosa em uma língua pode ser modal em outra) e nem entre falantes (o que é considerado voz soprosa para um falante pode ser modal para outro, mesmo dentro da mesma língua).

As medidas acústicas mais utilizadas para caracterizar a produção da voz e a fonação são realizadas no domínio espectral. As mais comuns são resultado da diferença entre a amplitude dos harmônicos, calculadas a partir dos espectros obtidos pela *Fast Fourier Transform* (FFT).³ O *Long Term Average Spectrum* (LTAS)⁴, que é realizado em trechos relativamente longos de fala, fornece informações relativas à distribuição de energia em diferentes faixas de frequência. Mais especificamente, o LTAS é resultado da média da amplitude ou da intensidade em uma determinada faixa de frequência de um trecho de fala contínua. De acordo com Li *et al.* (1969, *apud* Pittam, 1987), o LTAS deve ser obtido em trechos de fala entre 30 e 40 segundos, duração suficiente para que os efeitos de segmentos individuais não sejam significativos. Segundo Pittam (1987), pelo fato de a qualidade de voz ser uma característica de longo-termo da fala, as medidas de LTAS são bastante apropriadas para descrever tal fenômeno por dois motivos: (a) quando o foco é tratar da qualidade de voz como um marcador social ou de personalidade, e (b) quando julgamentos perceptuais são associados às medidas acústicas. As medidas do LTAS têm sido usadas para investigar questões relacionadas à identificação de falantes, comparação entre grupos de culturas diferentes e classificação de estados emocionais. Para Pittam (1987), o LTAS é um meio eficaz de obter principalmente informações relacionadas à fonação, como a sopro e a tensão laríngea.

O declínio espectral, medido pela diferença de energia entre bandas de frequência mais baixa do espectro e as de frequência mais alta, tem relação com a percepção da sopro na voz (HILLENBRAND *et al.*, 1994; HILLENBRAND e HOUDE, 1996). Além disso, Hammarberg *et al.* (1980) afirmam que o declínio espectral entre determinadas faixas de frequência do

espectro está relacionado à percepção de características fonatórias como a soproidade e a crepitação.

A medida de declínio espectral de longo termo mais utilizada é a “razão alfa”, que é a proporção de energia entre a faixa de frequência que vai de 0 a 1000Hz e a que vai de 1 a 5000Hz.⁵ Kitzing (1986) investigou a produção de quatro tipos de qualidade de voz (voz modal em amplitude normal, voz modal em amplitude reduzida, voz soproada e voz áspera⁶) realizada por fonoaudiólogos experientes. Os resultados dessa pesquisa mostraram que o declínio espectral foi o parâmetro mais potente para distinguir as qualidades vocais produzidas. Hammarberg *et al.* (1980) também relatam que a diferença de energia entre as faixas mais baixas e mais altas de frequência (como a razão alfa) podem indicar diferentes características glotais, como por exemplo, a ocorrência da voz soproada quando há uma baixa concentração de energia entre 400 e 4000Hz. Sundberg e Nordenberg (2006) afirmam que a razão alfa está diretamente relacionada com a intensidade do sinal acústico (quanto maior o volume, mais energia na faixa mais alta de frequência). Com o objetivo de cruzar os resultados de avaliações perceptuais com os de produção da voz, Hammarberg *et al.* (1980) utilizaram a diferença entre os picos de amplitude de diferentes faixas de frequência para relacionar aos resultados da análise fatorial dos julgamentos de percepção. Os autores adotaram as diferenças entre os picos de amplitude de 0 a 2000Hz e 2000 a 5000Hz e entre 2000 a 5000Hz e 5000 a 8000Hz. Uma maior diferença entre 0 a 2000Hz e 2000 a 5000Hz (maior declínio espectral) indica a produção da voz soproada; enquanto que uma maior diferença entre 2000 a 5000Hz, e 5000 a 8000Hz, é indicativa de voz crepitante.

Medidas relacionadas à f0 são também amplamente usadas para a caracterização da voz. Sabe-se que línguas e dialetos podem variar no que concerne ao emprego do *pitch*,⁷ apresentando diferentes médias e extensões de f0 (KEATING e KUO, 2012). Segundo essas autoras, se até os falantes de um mesmo grupo social em dada língua podem usar a f0 de maneira distinta, diferenças nesse parâmetro acústico entre línguas ou entre emissões de falantes bilíngues são bastante plausíveis:

A f0 é até certo ponto um aspecto arbitrário da fala, e uma determinada extensão de f0 pode fazer parte da estrutura fonética de uma língua, de maneira que no limite, um falante soaria não-nativo (tendo um sotaque estrangeiro) se usasse uma extensão de f0 diferente. (KEATING; KUO, 2012, p. 1050⁸)

Se a f0 faz parte da estrutura fonética de uma língua, é possível considerar que um bilíngue utiliza diferentes médias ou extensões dessa medida nas suas duas línguas. Por outro lado, ao mesmo tempo em que a f0 habitual fornece características a uma determinada língua ou dialeto, ela também depende dos aspectos fisiológicos e emocionais do falante como indivíduo. Assim, a f0 também depende fortemente de fatores como o sexo, a raça, e características psicossociais do falante.

A maioria dos estudos interlínguas para a caracterização da f0 de falantes bilíngues adota medidas distribucionais de longo termo, ou seja, medidas baseadas na distribuição de valores de f0 dentro de uma amostra de fala (MENNEN *et al.* 2014). As medidas de f0 mais utilizadas para comparar emissões de fala bilíngue são a média, o desvio padrão e a extensão, esta última calculada pela diferença entre os valores máximo e mínimo. Tais medidas são influenciadas tanto pelas configurações fisiológicas quanto pelas características habituais de produção e, portanto, podem revelar se o bilíngue realiza diferenças entre sua língua materna e a L2.

4. Qualidade de voz em L2

A qualidade de voz está diretamente relacionada à composição fisiológica do falante e, portanto, tende a ser a mesma independente da língua que a pessoa fale. Contudo, a qualidade de voz é também um parâmetro indicial e, segundo Esling (2013), bilíngues ou multilíngues podem alterar a qualidade de voz quando falam línguas diferentes, ou combinar qualidades das línguas que falam. Isso acontece, pois, dentro de uma comunidade linguística, os parâmetros de qualidade de voz são responsáveis por fornecer características tanto regionais quanto sociais dos seus falantes. Além das mudanças na qualidade de voz com motivação dialetal, Esling (2013) afirma que há modulações resultantes da personalidade e do estado emotivo do falante, que certamen-

te dependem de padrões culturais de dada comunidade linguística. Da mesma forma, a vontade de modificar dados aspectos da fala em L2 depende da atitude que o falante tem em relação à mesma, além de questões de identidade e autoestima, que são sinalizados pela qualidade vocal.

Jones e Evans (1995) afirmam que falantes de diferentes línguas utilizam modulações de qualidade de voz distintas, o que resulta em uma impressão auditiva de quem fala aquela língua, como por exemplo, a voz labializada e dentalizada do francês ou a voz anasalada e tensa do chinês. Essa impressão é geralmente o primeiro contato consciente com a fonologia de uma L2, que se é percebido quando alguém imita a maneira como uma língua soa antes mesmo de saber como falar nela. Jones e Evans (1995) defendem que um curso de pronúncia em L2 deve começar pela qualidade de voz, pelo fato de ser uma abordagem holística de diferentes aspectos fonético-fonológicos de maneira integrada. Essa abordagem dá ao aprendiz a chance da experiência com a pronúncia da L2 de forma intuitiva, bem como inserida em contextos reais de comunicação.

A maioria dos estudos sobre a produção da voz em bilíngues utiliza medidas acústicas para comparar os dados de fala. Segundo Mennen *et al.* (2010), as medidas acústicas para a análise da fala têm uma metodologia relativamente bem-estabelecida, são não-invasivas e bastante acessíveis. As pesquisas geralmente comparam a produção nas duas línguas utilizando medidas de f0, intensidade, tipos de fonação e LTAS.

Bruyninckx *et al.* (1991; 1994) investigaram a qualidade de voz produzida por participantes bilíngues falantes de catalão e espanhol.⁹ Foram coletados os dados de 24 participantes (12 homens e 12 mulheres), divididos por dominância em uma das línguas, realizando a leitura de um texto foneticamente balanceado em cada língua, em cinco repetições. Medidas de LTAS – índice de dissimilaridade SDDD¹⁰ de 0 a 5000Hz – foram utilizadas verificar possíveis diferenças. Duas análises intrafalantes foram realizadas: interlínguas (amostras de catalão X espanhol do mesmo falante) e intralíngua (amostras de catalão X catalão do mesmo falante). Houve maior variabilidade nos resultados interlínguas do que nos intralíngua. Além disso, houve diferenças

estatisticamente significativas entre dois grupos no que concerne às comparações intralíngua – mulheres com dominância no catalão e homens com dominância no espanhol. Os autores relatam que:

Há uma tendência para uma maior variabilidade na qualidade de voz para um dado falante na sua língua dominante, enquanto que o grau de coerência tende a ser maior na língua não-dominante.¹¹ BRUYNINCKX *et al.* (1991, p. 401)

Harmegnies *et al.* (1991) avaliaram o efeito do tipo de *corpus* na pesquisa sobre diferenças de qualidade de voz em bilíngues catalão/espanhol. Os autores utilizaram quatro tipos de *corpora* para cada uma das línguas (A, B, C e D¹²). Participaram do estudo seis bilíngues de catalão e espanhol, sendo três (dois homens e uma mulher) com dominância no catalão, e os outros três (dois homens e uma mulher) com dominância no espanhol. Foram feitas comparações de LTAS de dados intrafalantes, tanto interlínguas (catalão-espanhol) quando intra-língua (catalão-catalão e espanhol-espanhol), com textos dos corpora A, B e C em um primeiro experimento, e com os textos dos corpora C e D em um segundo experimento. Os resultados do primeiro experimento mostraram que, mesmo quando a similaridade fonética dos corpora aumenta, a variabilidade interlínguas do falante permanece a mesma. O segundo experimento, que comparou dois corpora equivalentes em similaridade interlíngua, mostrou que há um efeito maior nos valores do LTAS da língua utilizada em relação ao efeito do conteúdo dos corpora. Em suma, os autores argumentam que seus resultados vão contra a ideia de que as mudanças nos LTAS em comparação interlínguas sejam motivadas pela variação fonética das línguas.

Todaka (1995) analisou a fala de quatro bilíngues (dois homens e duas mulheres), falantes de IN como L1 e japonês como L2. Os participantes foram submetidos a experimentos para a obtenção de medidas aerodinâmicas, acústicas e eletroglotográficas. O objetivo do estudo era comparar as características laríngeas e respiratórias da produção dos bilíngues, e analisar as características supralaríngeas de acordo com a classificação “tensa” versus “frouxa” de Laver (1980). Os resultados estatisticamente significativos ($p < 0,05$) mostraram que:

(a) as participantes mulheres apresentaram voz mais soprosa em japonês do que em IN; (b) todos os sujeitos apresentaram valores de f_0 mais altos em japonês do que em IN; (c) todos os participantes apresentaram uso mais amplo do espaço vocálico em IN do que em japonês. Esses resultados sugerem, segundo o autor, que as diferenças de qualidade de voz entre as duas línguas pesquisadas têm motivação sociocultural.

Ng *et al.* (2012) examinaram a fala de 40 bilíngues (20 homens e 20 mulheres), falantes nativos de cantonês e proficientes em IN como L2, comparando as características vocais dos participantes nas duas línguas. A tarefa do experimento era a da leitura de dois textos: “*The North Wind and the Sun*” em cantonês e “*The Rainbow Passage*” em IN. Os autores mediram média de f_0 , além de medidas obtidas de espectros de longo termo como o primeiro pico, média de energia e declínio espectrais. Os resultados que foram estatisticamente significativos mostraram as seguintes diferenças: as participantes mulheres tiveram maior média de f_0 em IN do que em cantonês; maior média de energia espectral em cantonês do que em IN para todos os participantes; menor média de declínio espectral em cantonês do que em IN. As medidas espectrais revelaram que os participantes, quando falavam cantonês, faziam-no com maior tensão laríngea e com mais ruído de alta frequência, sendo tal diferença relacionada com a predominância da ressonância posterior característica do cantonês. O declínio espectral também revelou que maior produção de sopro no cantonês em relação ao IN.

Camargo *et al.* (2013) investigaram a produção de um bilíngue, falante de IN como L1, espanhol como L2 e português brasileiro como L3 no que concerne aos ajustes de qualidade vocal e a aspectos de dinâmica vocal. O participante, que tinha 71 anos, foi gravado produzindo fala semiespontânea e lendo textos nas três línguas. A análise perceptivo-auditiva foi realizada por meio do protocolo VPAS-PB (Camargo e Madureira, 2008) e a análise acústica foi feita pelas medidas obtidas pelo script SGExpression Evaluator (Barbosa, 2009). Os resultados mostraram que houve diferença na qualidade e na dinâmica vocais entre as línguas nas seguintes características: supralaríngeas (lábios e língua); laríngeas (fonatória e de tensão); medidas acústi-

cas (f_0 , derivada de f_0 e LTAS); tipo de tarefa (leitura ou fala semiespontânea). Além disso, a medida manual de declínio espectral revelou o agrupamento das emissões por língua.

Em suma, os estudos relacionados à produção da qualidade de voz por bilíngues revelam que há certas variáveis que influenciam os resultados de maneira geral. Primeiramente, parece haver um efeito na qualidade de voz em função da dominância linguística, pois há evidências que o bilíngue apresenta maior variabilidade na sua língua dominante. Outro ponto crucial na pesquisa sobre qualidade de voz em bilíngues é o tipo de tarefa que o participante realiza, já que há diferenças entre dados obtidos por meio de leitura de textos ou amostras de fala espontânea ou semiespontânea. Além disso, a variável sexo indiscutivelmente exerce efeito sobre a produção de diferentes qualidades de voz, principalmente motivadas por diferenças fisiológicas e culturais. Fatores como proficiência, idade de aquisição, ambiente de aprendizagem, tempo de instrução na L2 certamente influenciam a produção dos bilíngues, como já está estabelecido pela literatura em aquisição de L2.

No presente artigo, são apresentados resultados relacionados às diferenças entre as línguas e entre as tarefas (leitura e fala semiespontânea), bem como a correlação entre as variáveis de aquisição do IN como L2 e as medidas de LTAS e f_0 .

5. Metodologia

Para caracterizar as vozes em PB e em IN como L2 produzidas por bilíngues brasileiros, foram utilizadas as seguintes medidas acústicas espectrais: razão alfa, diferença entre os picos de amplitude de 0 a 2000Hz e de 2000 a 5000Hz (chamada de D0225kHz) e diferença entre os picos de amplitude de 2000 a 5000Hz e de 5000 a 8000Hz (chamada de D2558kHz). Para a caracterização da f_0 , foram utilizadas a média (e desvio padrão) e extensão dessa medida nas amostras de fala.

5.1 Corpus

O corpus desta pesquisa contou com tarefas de leitura e de respostas a perguntas em PB e em IN. Deci-

diu-se trabalhar com a leitura e a fala semiespontânea neste estudo, pois, segundo Kreiman e Sidtis (2011), as habilidades de produção da voz diferem-se substancialmente a depender da tarefa: fala (semi) espontânea, leitura em voz alta, repetição e canto.

Para realizar as medidas acústicas, optou-se por selecionar um texto que fosse foneticamente balanceado, e que produzisse amostras de fala com duração de pelo menos 30 segundos. A escolha do texto – *The Rainbow Passage*, de Fairbanks (1940) – foi baseada na sua ampla utilização por estudos experimentais na área da fonética em IN (PITTAM, 1987; HILLENBRAND e HOUDE, 1996; CLOPPER e SMILJANIC, 2011; KEATING e KUO, 2012). A partir do texto original em IN, foi feita uma versão para o PB. Foi selecionado apenas o primeiro parágrafo do texto para a realização das gravações, já que essa primeira parte já supria as condições de duração da amostra. Além da leitura dos textos, os participantes foram solicitados a responder perguntas, o que gerou as amostras de fala semiespontânea desta pesquisa. Foram apresentadas três perguntas¹³ aos participantes, e eles tinham que escolher duas dessas para responder. Para cada uma das perguntas, havia sugestões de tópicos para as respostas dos participantes, que poderiam ser utilizadas ou não. As mesmas perguntas que eles respondiam em PB, teriam que responder também em IN.

5.2 Participantes

Os participantes do experimento foram recrutados na *University of California Los Angeles* (UCLA) nos meses de fevereiro e março de 2014. Todos eram estudantes de graduação ou pós-graduação nessa universidade. Participaram deste estudo 16 brasileiros bilíngues, dez mulheres e seis homens, com média de idade de 27,5 anos (20 a 39) na época da coleta dos dados, tendo o PB como língua materna e o IN como L2.¹⁴

Os participantes revelaram detalhes sobre a experiência de estudo do inglês, bem como a vivência em países em que esse idioma é falado como língua materna (QUADRO 1).

Participante	Idade (anos)	Idade início aprendizado IN (anos)	Tempo de estudo IN (meses)	Tempo de vivência em país de LI (meses)
Fal1	25	10	96	4
Fal2	26	8	96	10
Fal3	24	14	36	24
Fal4	23	11	120	5
Fal5	39	14	36	144
Fal6	34	22	36	60
Fal7	30	18	24	60
Fal8	30	7	120	30
Fal9	21	7	24	5
Fal10	27	12	24	7
Fal11	20	10	108	84
Fal12	29	9	96	48
Fal13	31	8	60	68
Fal14	25	11	168	168
Fal15	23	7	96	84
Fal16	33	12	48	68

Quadro 1: Detalhes sobre aquisição e experiência do IN dos participantes

O Quadro 1 traz a descrição detalhada dos dados obtidos pela entrevista. Primeiramente, com relação à idade de aquisição do IN, a maioria dos participantes relatou ter começado a aprendizagem quando criança (dos 7 aos 12 anos). Somente alguns relataram ter iniciado a aquisição quando adultos (dos 14 aos 22 anos). Muitos relataram terem estudado inglês formalmente em escola de idiomas no Brasil ou em outros países por vários anos. Quanto ao número de meses que os participantes tinham morado em países de LI, houve uma grande variabilidade no número de meses de permanência. Alguns sujeitos moravam no país há mais de 14 anos, enquanto outros haviam chegado poucos meses antes das gravações acontecerem.¹⁵

5.3 Experimento

O experimento foi realizado no *Voice Perception Laboratory*, que faz parte do *Head and Neck Surgery*

Department da David Geffen School of Medicine, da UCLA. Primeiramente, os participantes leram e assinaram o termo de consentimento.¹⁶ Depois disso, eles responderam às questões da entrevista demográfica, que foi feita em PB. Foram gravadas primeiramente as tarefas em PB e, logo em seguida, as do IN. A duração total da coleta foi de aproximadamente 45 minutos para cada participante, sem pausas.¹⁷ As etapas do experimento ocorreram da seguinte maneira:

1. Leitura do texto em PB. Instruções: ler o texto no volume e velocidade habituais. Foi pedido aos participantes que lessem o texto em silêncio para se familiarizarem com o mesmo. Posteriormente, foram feitas as três gravações da leitura desse texto.
2. Resposta às perguntas propostas. Instruções: ler as perguntas e escolher duas para responder. Foi solicitado aos participantes que falassem por cerca de um minuto em cada resposta.

Após a coleta em PB, foram realizadas as mesmas etapas na coleta dos dados em IN: leitura do texto e resposta às perguntas. Foi solicitado que o participante respondesse em IN às mesmas perguntas que havia respondido em PB. Todas as instruções dessa etapa foram as mesmas utilizadas na coleta dos dados em PB, porém dadas em IN.

A gravação dos dados de fala foi coletada em um computador da marca *Apple*, modelo *MacBookPro*. Foi utilizado um microfone de cabeça, unidirecional, marca *Shure*, modelo SM10A-CN, com resposta de frequência entre 50Hz e 15000Hz. Esse microfone foi conectado a um pré-amplificador, marca *Blue*, modelo *Icicle*, que por sua vez foi conectado ao computador. O microfone foi posicionado a 10 centímetros da boca do participante, em um ângulo aproximando de 45°. O software utilizado para gravação foi o *Audacity 2.0.5*,¹⁸ que é um programa gratuito para edição de áudio. O registro dos dados foi feito em formato *wav*, com frequência de amostragem de 44000Hz e taxa de quantização de 32 bits.

5.4 Medidas de LTAS e f0

As medidas de LTAS e f0 foram realizadas nas tarefas de leitura dos textos e da fala semiespontânea. Os valores foram obtidos pelo *Praat*, seguindo os seguintes critérios:

1. A medida razão alfa foi obtida pelo comando *To Ltas (pitch corrected)*, também com correção de *pitch* (50 a 300Hz para homens e 100 a 500Hz para mulheres), com frequência máxima de 8000Hz, e largura de banda de 160Hz. As diferenças D0225kHz e D2558kHz foram obtidas a partir do *Ltas object* gerado pelo *Praat*, a partir das diferenças entre os picos de amplitude das seguintes bandas de frequência: 0 a 2000Hz, 2000 a 5000Hz e 5000 a 8000Hz (a diferença entre as duas primeiras faixas e as duas últimas). Foram obtidas dez medidas para cada uma das duas diferenças, todas expressas em dB.
2. Para f0, foram usadas as configurações padrão do programa, a não ser por dois ajustes: (a) a correção das frequências de *pitch* mínima (*floor*) e máxima (*ceiling*) que eram originalmente de 75 a 600Hz, passando à configuração de 50 a 300Hz para homens, e de 100 a 500Hz para mulheres; (b) o intervalo de tempo de medição, modificado para 10 milissegundos. Depois da extração das medidas, foram selecionadas a média, a extensão e o desvio padrão de cada amostra. No total, foram obtidas dez medidas por falante, todas as medidas de f0 são expressas em Hz.

6. Resultados

Primeiramente, com o objetivo de verificar se havia diferenças intrassujeito entre as três repetições da leitura do texto e as duas amostras de fala semiespontânea, foram rodados testes estatísticos em cada língua (Friedman para as três repetições da leitura e Wilcoxon para as amostras de fala semiespontânea). Os resultados mostraram não haver diferenças estatisticamente significati-

vas¹⁹ nas medidas de LTAS e f0 entre as três repetições da leitura do texto em cada língua ou entre as duas amostras de fala semiespontânea ($p>0,05$). Portanto, foram calculadas a média e o desvio padrão referentes à cada uma das medidas (LTAS e f0) para cada tarefa (leitura ou fala semiespontânea) para cada língua.

A partir dessas médias, foram realizados testes estatísticos para realizar a comparação entre as línguas, entre as tarefas, e a correlação entre as medidas e as variáveis de aquisição do IN com L2.

6.1 Comparação entre línguas

Primeiramente, foram calculadas a média e o desvio padrão das medidas de LTAS para as amostras de leitura de texto (TX) e fala semiespontânea (SE) em ambas as línguas, conforme ilustrado na Tabela 1:

Tabela 1: Média e desvio padrão das medidas de LTAS, em DB.

Medida	TX PB	TX IN	SE PB	SE IN
LTAS	Média (DP)	Média (DP)	Média (DP)	Média (DP)
Razão alfa	10,24 (0,30)	9,53 (0,40)	8,83 (1,07)	8,67 (1,52)
D0225kHz	19,42 (0,45)	17,18 (0,60)	16,81 (1,28)	16,28 (1,88)
D2558kHz	14,20 (3,88)	17,92 (2,16)	17,26 (1,90)	18,48 (1,86)

Para verificar possíveis diferenças nas características de LTAS na comparação entre a produção do PB e do IN, foi rodado o teste Wilcoxon, cujo resultado é apresentado na Tabela 2:

Tabela 2: Resultados do teste WILCOXON entre PB e in para medidas de LTAS

Medida de LTAS	TX: PB X IN	SE: PB X IN
Razão Alfa	Z=-2,534, p=,011	Z=-1,034 p=,301
D0225kHz	Z=-3,516, p>,001	Z=-1,500, p=,134
D2558kHz	Z=-3,361, p=,001	Z=-2,327, p=,020

Os resultados da Tabela 2 mostram que há diferenças estatisticamente significativas na comparação entre

a produção em IN e em PB na tarefa de leitura de texto. A razão alfa e a D0225kHz tiveram médias mais altas em PB. Segundo Kitzing (1986), a noção psicoacústica de sonoridade na qualidade de voz é refletida em um declínio espectral menos acentuado. Portanto, a partir dos resultados dessas duas medidas, pode-se afirmar que as vozes dos participantes tiveram maior intensidade no IN que no PB. Keating e Garellek (2015) afirmam que a voz crepitante tem harmônicos mais fortes na região entre 3000 e 5000Hz, justamente o que aconteceu na produção em IN descrita pelo declínio espectral menos acentuado nessa língua. Assim, de maneira geral, os brasileiros produziram uma voz mais soprosa no PB que no IN. A medida D2558kHz também revelou essa tendência, pois as médias foram mais altas em IN que em PB, indicando maior presença de voz crepitante nessa língua. Segundo Hammarberg *et al.* (1980), a voz soprosa está relacionada com um maior declínio espectral entre as faixas de 0 a 2000Hz e 2000 a 5000Hz, enquanto que a voz crepitante está relacionada com um maior declínio espectral entre as faixas 2000 a 5000Hz e 5000 a 8000Hz.

Com relação à fala semiespontânea, houve diferença estatisticamente significativa na medida D2558kHz entre o PB e o IN, indicando a presença de mais crepitação na voz no IN.

Os presentes resultados estão em consonância com os de Ng *et al.* (2012), que compararam as características espectrais de bilíngues cantonês/IN na tarefa de leitura de texto. Nesse estudo, as participantes mulheres apresentaram declínio espectral mais baixo que os homens (médias mais altas). Além disso, todos os participantes apresentaram declínio espectral mais baixo em cantonês, sua língua materna, do que em IN. No estudo de Todaka (1995), da mesma forma, as mulheres apresentaram configuração laríngea mais relaxada e constante abertura glotal em japonês que em IN.

Com relação à f0, os valores da média e extensão produzidos na tarefa de leitura de textos e fala semiespontânea são apresentados na Tabela 3:

Tabela 3: Média e extensão de f0 em TX e se, em Hz²⁰.

Medida de f0	TX – PB	TX – IN	SE – PB	SE – IN
Média	161,56 (66,55)	158,00 (66,50)	159,24 (65,54)	155,73 (63,05)
Extensão	298,36 (105,30)	304,71 (107,63)	297,01 (93,93)	301,28 (91,60)

Para verificar se havia diferenças nas medidas de f0 entre as línguas, foram obtidos os resultados a partir do teste Wilcoxon mostrados na Tabela 4:

Tabela 4: Resultados do teste WILCOXON entre PB e IN para medidas de f0 em TX e SE

Medida de f0	TX: PB X IN	SE: PB X IN
Média	Z=-1,965, p=,049	Z=-1,655, p=,098
Extensão	Z=-1,965, p=,049	Z=-0,517, p=,605

Segundo a Tabela 4, na tarefa de leitura de texto, as duas medidas de f0 tiveram diferença estatisticamente significativa entre o PB e o IN. A média de f0 foi mais baixa no IN que no PB, enquanto que a extensão foi mais alta. Isso indica que, por mais que os brasileiros tenham um *pitch* mais baixo na leitura em IN, eles também apresentam maior variação na f0 nessa língua. Por outro lado, não houve diferença estatisticamente significativa entre o PB e o IN na fala semiespontânea.

Outros estudos também verificaram valores de f0 mais altos para as línguas maternas do que para as línguas estrangeiras. Todaka (1995) constatou que suas participantes produziam f0 mais altas em japonês que em inglês; as bilíngues russo/IN de Altenberg e Ferrand (2006) produziram f0 mais alta em russo que em inglês. Por outro lado, Ng et al. (2012), ao comparar as médias de f0 em inglês e cantonês, verificaram que as participantes produziam f0 mais alto em inglês, sua L2.

6.2 Comparação entre tarefas

Com o objetivo de verificar possíveis diferenças entre as tarefas, foi rodado novamente o teste Wilcoxon para comparar as medidas espectrais das produções dos participantes bilíngues: TX-IN versus SE-IN; TX-PB versus SE-PB, conforme a Tabela 5:

Tabela 5: Resultados do teste WILCOXON para medidas de LTAS entre TX e SE

Medida de LTAS	IN: TX x SE	PB: TX x SE
Razão alfa	Z=-2,327, p=,020	Z=-2,947, p=,003
D0225kHz	Z=-1,603, p=,109	Z=-3,206, p=,001
D2558kHz	Z=-0,621, p=,535	Z=-3,154, p=,002

De acordo com a Tabela 5, nas comparações entre tarefas do IN, houve diferença estatisticamente significativa apenas na razão alfa, sendo a média mais baixa na fala semiespontânea que na leitura de texto. Isso indica que os participantes produziram vozes mais soprosas na leitura de texto em relação à fala semiespontânea. Já no PB, houve diferenças estatisticamente significativas nas três medidas na comparação entre os tipos de tarefas. O declínio espectral foi mais alto na leitura de texto em comparação com a fala semiespontânea nas medidas razão alfa e D0225kHz; na medida D2558kHz, o declínio é maior na fala semiespontânea que na leitura do texto. Esses resultados indicam maior soprosidade na leitura de texto no PB e maior crepitação na fala semiespontânea no IN.

Camargo *et al.* (2013) também encontraram diferenças estatisticamente significativas entre a tarefa de leitura de texto em comparação com a fala semiespontânea nas medidas de declínio espectral. Esses autores reportaram valores mais altos de declínio espectral na fala semiespontânea nas três línguas do participante.

Além disso, o fato de no IN haver menos diferenças entre as tarefas está em consonância com as conclusões dos estudos de Bruyninckx *et al.* (1991; 1994), que mostraram haver mais consistência/menos variabilidade na língua não-dominante. Portanto, nossos resultados parecem indicar que os participantes bilíngues têm maior variabilidade quando falam o PB do que quando produzem o IN.

A respeito das medidas de f0 na comparação das tarefas desempenhadas pelos participantes, obteve-se os seguintes resultados (TABELA 6):

Tabela 6: Resultados do teste WILCOXON para medidas de f0 do entre TX e SE

Medida de f0	IN: TX x SE	PB: TX x SE
Média	Z=-1,344, p=,179	Z=-1,603, p=,109
Extensão	Z=-0,827, p=,408	Z=-0,078, p=,938

A Tabela 6 mostra que os participantes não apresentam diferença estatisticamente significativa entre as medidas de f0 produzidas em PB e as do IN na comparação entre tarefas. Embora não haja diferença estatisticamente significativa entre as tarefas, pode-se constatar

pelos medidas que as médias produzidas na fala semiespontânea foram mais baixas que as na leitura de texto.

6.3 Correlação entre variáveis de aquisição do IN e medidas acústicas

Para verificar a correlação entre as variáveis de aquisição do IN e as medidas acústicas, foram realizados testes Pearson. Primeiramente, reportamos os resultados do teste entre o tempo de experiência em países de LI e as medidas acústicas obtidas a partir da fala em IN (TABELA 7):

Tabela 7: Resultados do teste PEARSON entre tempo de experiência em países de LI e medidas acústicas do IN

		Média f0 TX	Média f0 SE	Ext. f0 TX	Ext. f0 SE	Razão A. TX	Razão A. SE	D0225k TX	D0225k SE	D2558k TX	D2558k SE
Tempo experiência país IN	Corr. Pearson	,339	,334	,336	,488	-,086	-,094	-,052	,034	-,139	,014
	Sig. (2 extr.)	,199	,206	,204	,055	,750	,730	,850	,899	,608	,960

Como se pode verificar na Tabela 7, não houve correlação entre o tempo de experiência em países de língua inglesa e a qualidade de voz em IN dos participantes desta pesquisa. Por outro lado, houve correlação

entre a idade de aquisição do IN e todas as medidas de f0 (TABELA 8):

Tabela 8: Resultados do teste PEARSON entre a idade de aquisição e as medidas de F0

		Média f0 TX	Média f0 SE	Ext. f0 TX	Ext. f0 SE	Razão A. TX	Razão A. SE	D0225k TX	D2558k TX	D0225k SE	D2558k SE
Idade Aquisição IN	Corr. Pearson	,622	,562	,625**	,603*	-,260	-,325	-,216	-,432	-,380	-,155
	Sig. (2 extr.)	,010	,023	,010	,013	,331	,219	,421	,095	,147	,565

A correlação se deu na medida em que quanto mais cedo tivesse ocorrido o início da aquisição do IN pelo participante, mais baixos eram os valores de f0. Essa relação é ilustrada na Figura 2:

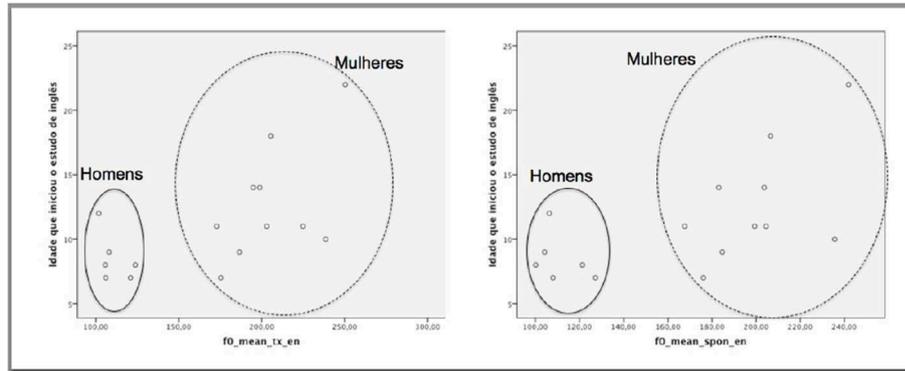


Figura 2: Gráfico de correlação entre médias de f0 em IN e idade de aquisição

É pertinente observar a correlação entre idade de aquisição e média de f0, tanto na leitura de texto, quanto na fala semiespontânea, ilustrado na Figura 2. Os gráficos de dispersão mostram que, especialmente para as mulheres, quanto mais cedo tivesse ocorrido a aquisição do IN, mais baixa era a média de f0.

6.4 Discussão dos resultados

O objetivo geral deste trabalho foi comparar a fala de bilíngues brasileiros em emissões no PB e no IN no que concerne à qualidade de voz. Nossa hipótese, a de que haveria diferenças entre as línguas e entre as tarefas, bem como correlação entre as variáveis de aquisição e as medidas acústicas, foi parcialmente confirmada. Segundo os resultados reportados, houve diferença estatisticamente significativa entre as línguas em todas as medidas na tarefa de leitura de texto. Houve também diferenças significativas entre as tarefas nas medidas espectrais do PB, bem como na razão alfa do IN. A respeito da relação entre variáveis de aquisição e medidas acústicas de qualidade de voz, foi verificada correlação entre idade de aquisição do IN e média e extensão de f0. Nossa hipótese foi refutada nas comparações entre línguas da fala semiespontânea, bem como nas medidas espectrais e de f0 do IN entre as tarefas; tampouco não houve correlação entre o tempo de experiência em países de língua inglesa e as medidas acústicas.

Ficou evidente, pelos resultados apresentados, que os brasileiros bilíngues que participaram deste estudo produziram vozes em IN com algumas características distintas das vozes produzidas em PB. É bastante im-

portante ressaltar que quando se reporta diferenças entre as línguas, estamos utilizando uma como referência da outra. Portanto, quando apontamos uma tendência de voz mais soprosa em uma língua, é porque a outra apresenta menos características acústicas de voz soprosa, por exemplo.

Na tarefa de leitura do texto, os brasileiros produziram vozes mais soprosas em PB e vozes mais crepitantes em IN, assim como indicaram as três medidas espectrais de longo termo. Já na fala semiespontânea, houve diferença apenas na medida de declínio espectral que indica a crepitação na voz – a D2558kHz. Os falantes produziram vozes mais crepitantes em IN do que em PB. Esses resultados estão em consonância com os de Ng *et al.* (2012), que analisaram a qualidade de voz em Cantonês e IN por chineses bilíngues proficientes no que concerne à tarefa de leitura de texto. Seus resultados mostraram que os chineses produziram vozes mais soprosas em cantonês que em IN. Da mesma forma, Harmegnies e Landercy (1985), comparando bilíngues de francês e holandês, Harmegnies *et al.* (1989) e Bruyninckx *et al.* (1994), comparando bilíngues de catalão e espanhol, reportaram diferenças na qualidade de voz por medidas de espectro de longo termo entre as línguas. Contudo, esses estudos não contaram com a análise de amostras de fala semiespontânea, utilizando somente a tarefa de leitura de texto em seus experimentos. Como aponta Byrne *et al.* (1994), medidas espectrais de longo termo são certamente influenciadas pelo tipo de material de fala analisada, mas esses autores discutem apenas as diferenças entre listas de palavras ou conjuntos de frases e pequenos textos. Os nossos resultados, que trazem diferenças estatisticamente significativas em todas as medidas espectrais de longo termo na tarefa de leitura, parecem indicar que os bilín-

gues produzem vozes distintas entre as línguas, como se houvesse um “*reading mode*” específico para cada uma.

Essa tendência de produzir vozes menos soprosas, e muitas vezes mais crepitantes em IN tanto para os homens, mas principalmente para as mulheres, vai ao encontro com pesquisas recentes relacionadas aos tipos de fonação e sexo. Podesva (2013) é um exemplo de estudo que mostra que as mulheres estadunidenses têm produzido altos índices de voz crepitante em sua fala. Esse autor cita uma série de outras pesquisas que têm indicado a mesma tendência, inclusive apontando as possíveis motivações para tal fenômeno. Dentre as impressões que uma voz mais crepitante pode causar estão: assumir postura mais autoritária, mas sem soar agressiva; estar em posição social ascendente, ser profissional, urbana; ser capaz de competir à altura com os homens. Segundo Podesva (2013), “*de acordo com essa interpretação, portanto, o significado social da voz crepitante está enraizada na sua semelhança com a voz dos homens*”²¹ Ele ainda sugere que as altas taxas de voz crepitantes encontradas na fala feminina podem ser consequência do abaixamento do *pitch*.

Na comparação entre as tarefas, houve diferença estatisticamente significativa em algumas medidas espectrais em cada língua. A tarefa de leitura de texto apresentou diferença apenas na razão alfa no IN em comparação à fala semiespontânea, com a presença de maior sopro na leitura de texto. No PB, contudo, as três medidas apresentaram diferenças entre as tarefas: mais sopro na leitura de texto e mais crepitação na fala semiespontânea. O fato de haver mais diferenças nas vozes entre as tarefas em PB que em IN converge com os resultados de Bruyninckx *et al.* (1994), cujos resultados mostraram que há uma maior coerência na língua não dominante e uma maior variabilidade na língua dominante.

Além da dominância linguística e do tipo de tarefa provocarem diferenças nas vozes dos bilíngues entre as emissões em PB e em IN, podem também estar envolvidas no fenômeno as questões da personalidade e da emoção. Sabe-se que a voz é altamente condicionada tanto pelos estados emocionais transientes quanto pelas características identitárias mais estáveis (KREIMAN e SIDTIS, 2011). Os atributos vocais do falante podem

revelar características físicas (altura, beleza, idade), psicológicas (inteligência, extroversão, dominância), emocionais, bem como relacionadas à masculinidade ou feminilidade. Some-se a isso, a qualidade de voz também tem motivação cultural, podendo variar entre as línguas por conta de estereótipos vocais existentes em diferentes países. Todas essas explicações podem dar conta das diferenças produzidas pelos bilíngues da presente pesquisa.

Como explica Moyer (2004), há outros fatores para além da idade de aquisição e tempo de experiência que influenciam a aprendizagem da L2. Ela relata que a intensidade da motivação em termos do alcance de uma boa proficiência ou em termos profissionais são fatores que influenciam fortemente a aprendizagem da L2. Além desses citados, as oportunidades de contato naquela L2, a atitude em relação à cultura do país em questão e a intenção em permanecer nele têm peso na progressão durante o processo de aprendizagem. Pode ser que os participantes do presente estudo tenham sido influenciados de maneira mais contundente por esses fatores sociais e afetivos envolvidos na aquisição da L2, principalmente pela qualidade de voz ser crucial para veicular atributos paralinguísticos na comunicação. Contudo, como não é do escopo de nossa pesquisa, não foram coletados dados de motivação e atitude para com o IN e, por isso, não se pode afirmar inequivocadamente que a qualidade de voz seja realmente afetada por tais fatores sociais e afetivos.

7. Considerações finais

A partir dos resultados apresentados e da discussão em torno deles, ficou evidenciado que os participantes desta pesquisa produzem algumas diferenças entre as vozes no PB e no IN. Dentre as características que mais se distinguem entre as línguas estão as medidas de LTAS, especialmente a razão alfa e a D0255kHz, além da média de f0. As diferenças detectadas entre o PB e o IN pelas medidas acústicas variaram principalmente em função da tarefa desempenhada, do sexo, e da idade de aquisição do IN.

A qualidade de voz é a matéria-prima da prosódia, sendo responsável por veicular tanto características lin-

guísticas como paralinguísticas na comunicação. Suas funções, especialmente no domínio paralinguístico – o mais importante para este trabalho – envolvem a expressão da emoção e a intenção do falante, bem como fornecem características psicológicas e de personalidade dos falantes. Em pesquisas sobre os efeitos da língua na expressão da personalidade em falantes bilíngues, há evidências que, a depender da língua que utilize, o falante muda suas características de personalidade principalmente para adequar-se aos padrões culturais adotados por determinada sociedade. Chen e Bond (2010) afirmam que características de personalidade expressas pela fala podem ter origem em diferentes papéis sociais que o falante desempenha entre as línguas. Segundo esses autores, as línguas que os bilíngues falam ativam diferentes comportamentos culturais específicos de cada língua. Outro fator que parece estar associado às mudanças de personalidade em função da língua é a competência linguística, de forma que quanto mais fluente o bilíngue seja, menos diferenças de personalidade ele produza. Assim, é possível especular que, para os falantes bilíngues do presente estudo, as mudanças na voz entre o PB e o IN podem ter origem nas diferenças culturais que existem entre a sociedade brasileira e estadunidense. Nesse raciocínio, a motivação para as mudanças de qualidade de voz dos brasileiros pode ser uma adaptação às características de personalidade atribuídas aos estadunidenses.

Em conclusão, a presente pesquisa vem a contribuir com a área dos estudos relacionados à produção da voz no que concerne às diferenças que bilíngues apresentam em sua fala na utilização de uma ou de outra língua. Parece que a impressão de vozes distintas, comumente percebidas por aprendizes de L2 ou bilíngues no dia a dia, realmente existe para algumas pessoas.

Notas

1. Não são encontrados, no PB ou no IN, contrastes linguísticos veiculados por diferentes tipos de qualidades de voz.
2. Do inglês “*indexical*”.
3. A Transformada de Fourier é um método de transformação de um sinal digital que converte o formato de uma onda acústica em um espectro. (JOHNSON, 2012).
4. Em português, “espectro de longo termo”.
5. Frokjaer-Jensen and Prytz (1976) nomearam essa proporção “*alpha ratio*”.
6. “*Leaky voice*” foi traduzida como voz soprosa e “*strained voice*”, como voz áspera.
7. *Pitch* é o correlato perceptual da f_0 que, por sua vez, é a medida acústica relacionada à taxa de repetição do ciclo de uma onda sonora em um segundo.
8. A tradução das citações cujos originais estão em inglês são de responsabilidade das autoras deste artigo. Os originais aparecerão em itálico em nota de rodapé. *Speaking F0 is to some extent an arbitrary aspect of speech, and a particular F0 range may be part of the phonetic structure of a language, such that in the limit, a speaker would sound non-native (have a foreign accent) using a different F0 range.*
9. A variedade do espanhol considerada no artigo é o castelhano (*Castilian*).
10. *Standard Deviation of the Differences Distribution* (Harmegnies, 1988).
11. *Thus, there is a tendency towards greater voice quality variability for a given speaker in his or her dominant language, while the degree of voice coherence tends to be higher in the non-dominant language.*
12. Os corpora do tipo A eram constituídos por textos foneticamente balanceados, compostos por uma distribuição equilibrada dos elementos fonéticos de cada língua. Os corpora B, por sua vez, de frases contendo elementos fonéticos comuns ao catalão e ao espanhol. Os corpora C contêm o mesmo formato dos corpora B, porém com a maior similaridade fonética entre os enunciados nas duas línguas que foi possível conceber. Por fim, os corpora D, já que não era possível criar dois enunciados equivalentes foneticamente em línguas distintas, foram compostas frases tão diferentes das dos corpora C quanto os corpora C eram distintos uns dos outros.
13. 1. Onde você nasceu e cresceu? Fale um pouco sobre sua cidade natal. 2. Do que mais você sente falta do Brasil? 3. Do que você não gosta nos Estados Unidos/ Los Angeles?
14. Os participantes tinham como estado de origem SC (5), SP (5), RS (1), PR (1), RJ (1), MG (1), BA (1), DF (1).
15. Alguns participantes relataram conhecimento em outras línguas estrangeiras, todas aprendidas depois do IN: quatro sujeitos relataram conhecimento em espanhol; um em alemão; um em francês; um em japonês; um em mandarim; e um em coreano. Os outros participantes reportaram não ter conhecimento em outras línguas estrangeiras.

16. O experimento de produção teve aprovação do Comitê de Ética da UCLA (*Institutional Review Board - IRB*), sob o protocolo IRB#13-001788, a partir da data de 08/01/2014. A coleta de dados foi iniciada dia 13/02/2014.
17. A coleta de dados foi realizada pela primeira autora deste artigo.
18. Disponível em <http://audacity.sourceforge.net>
19. Foram considerados estatisticamente significativos os resultados cujo valor de p foi menor ou igual a 0,05.
20. Os valores do desvio padrão, que estão entre parênteses, foram altos por conta da grande diferença entre valores de f_0 para homens e mulheres.
21. *According to this interpretation, then, the social meaning of creaky voice is rooted in its resemblance to men's voices.*

Referências

- ALTENBERG, E. P.; FERRAND, C. T. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. **Journal of Voice**, 20(1), 89–96, 2006.
- BARBOSA, P.A. Detecting changes in speech expressiveness in participants of a radio program. In: **Proceedings of Interspeech 2009 - Speech and Intelligence**. Londres: Causal Productions, 2155-2158, 2009.
- BOERSMA, P.; WEENIK, D. **Praat: doing phonetics by computer**. Version 5.3.82, 2014.
- BRUYNINCKX, M.; HARMEGNIES, B.; LLISTERRI, J.; e POCH, D. Effects of language change on voice quality. An experimental study of catalan-castilian bilinguals. In ICPhS 1991. **Actes du 12 ème Congrès International de Sciences Phonétiques**. Vol. 2. (pp. 398-401). Aix-en-Provence: Université de Provence, Service des Publications, 1991.
- BRUYNINCKX, M.; HARMEGNIES, B.; LLISTERRI, J.; e POCH, D. Language-induced voice quality variability in bilinguals. **Journal of Phonetics**, 22, 19-31, 1994.
- BYRNE, D.; DILLON, H.; TRAN, K.; ARLINGER, S.; WILBRAHAN, K; COX, R.; HAGER-MAN, B.; HETU, R.; KEI, L.; LUI, C.; KIESSLING, J.; KOTBY, N. M.; NASSER, N. H. A.; WAFAA, A. H. El Kholy; NAKANISHI, Y.; OYER, H.; POWELL, R.; STEPHENS, D.; MEREDITH, R.; SIRIMANNA, T.; TAVARTKILADZE, G.; FROLENKOV, G. I.; WESTERMAN, S.; LUDVIGSEN, C. "An international comparison of long-term average speech spectra," **Journal of the Acoustical Society of America**, 96, 2108–2120. 1994.
- CAMARGO, Z.; MADUREIRA, S.; SCHMITZ, J.R. Qualidade de voz e produção de fala em três línguas: um estudo de caso. **Revista Intercâmbio**, v.27, 110-140, 2013.
- CHEN, S. X., BOND, M. H. Two languages, two personalities? Examining language effects on the expression of personality in a bilingual context. **Personality and Social Psychology Bulletin**, 36, 1514-1528, 2010.
- CLOPPER, C.G., SMILJANIC, R. Effects of gender and regional dialect on prosodic patterns in American English. **Journal of Phonetics**, 39(2): 237–245, 2011.
- COHEN, H.; DOUAIRE, J.; ELSABBAGH, M. The role of prosody in discourse processing. **Brain and Cognition**, 46, p. 73-82, 2001.
- ESLING, J.H. Voice Quality. **The Encyclopedia of Applied Linguistics**. Wiley Online Library, p.1-7, 2013.
- FAIRBANKS, G. **Voice and Articulation Drill Book**. Harper, New York: p. 168, 1940.
- FROKJAER-JENSEN, B.; PRYTZ, S. Registration of voice quality. *Bruel and Kjaer Technical Review*, 3, 1976, 3-17.
- GORDON, M.; LADEFOGED, P. Phonation types: a cross-linguistic overview. **Journal of Phonetics**, 29:383-406, 2001.
- HAMMARBERG, B., FRITZELL, B., GAUFFIN, J., SUNDBERG, J.; WEDIN, L. "Perceptual and acoustic correlates of ab- normal voice qualities," **Acta Otolaryngol** 90, 441–451, 1980.
- HARMEGNIES, B.; LANDERCY, A. Language features in the long-term average spectrum. **Revue de Phonétique Appliquée**, 73-75: 69-79, 1985.
- HARMEGNIES, B.; BRUYNINCKX, M.; LLISTERRI, J.; e POCH, D. Effects of language change on voice quality: An experimental contribution to the study of the Catalan-Castilian case. In **Eurospeech 1989: Proceedings of the 1st European Conference on Speech Communication and Technology**, 489-492, 1989.
- _____. Effects of language change in voice quality in bilingual speakers: Corpus content effects. In **Eurospeech 1991. Proceedings of the 2nd European conference on speech communication and technology**, Vol. 1, p.165-8. Genova, Italy, 1991.
- HILLENBRAND, J.M.; CLEVELAND, R.A.; and ERICKSON, R.L. "Acoustic correlates of breathy vocal quality," **Journal of Speech and Hearing Research**, 37, 769-778, 1994.
- HILLENBRAND, J.; HOUDE, R.A. Acoustic correlates of breathy voice quality: dysphonic voices and continuous

- speech. **Journal of Speech, Language, and Hearing**, 39, 298-310, 1996.
- JOHNSON, K. **Acoustic and Auditory Phonetics**. 2nd Edition. Oxford: Blackwell, 2012.
- JONES, R.H.; EVANS, S. Teaching pronunciation through voice quality. **LTAS Journal** 49: 3 244- 251, 1995.
- KEATING, P.A.; ESPOSITO E, C. Linguistic Voice Quality. **UCLA Working Papers in Phonetics**, nº 105, p.85-91, 2007.
- KEATING, P. A.; KUO, G. Comparison of speaking fundamental frequency in English and Mandarin. **Journal of the Acoustic Society of America**, 132 (2), P. 1050 – 1060, 2012.
- KEATING, P. ; GARELLEK, M. Acoustic analysis of creaky voice. Poster apresentado em sessão especial sobre voz crepitante no Encontro Anual da Linguistic Society of America em Portland (OR), 2015.
- KITZING, P. LTAS criteria pertinent to the measurement of voice quality. **Journal of Phonetics**, v. 14, p. 477-482, 1986.
- KREIMAN, J.; SIDTIS, D. **Foundations of Voice Studies: An Inter-disciplinary Approach to Voice Production and Perception**. Wiley-Black-well, Walden, MA, pp. 1–518, 2011.
- LADEFOGED, P. **Preliminaries to linguistic phonetics**. Chicago: University of Chicago Press, 1971.
- LAVER, J. **The phonetic description of voice quality**. Cambridge: Cambridge University Press, 1980.
- MENNEN, I.; LEEUW, E. Beyond Segments. **Studies in Second Language Acquisition**, 36, p. 183-194, 2014.
- MOYER, A. Age, accent and experience in second language acquisition. Clevedon: **Multilingual Matters**, 2004.
- NG, M. L.; CHEN, Y.; CHAN, E. Y. K. Differences in Vocal Characteristics Between Cantonese and English Produced by Proficient Cantonese-English Bilingual Speakers—A Long-Term Average Spectral Analysis. **Journal of Voice**, Vol. 26, No. 4, p. e171-e176, 2012.
- OHTA, A. S. *SLA Processes in the Classroom: Learning Japanese*, Lawrence Erlbaum Associates, 2001.
- PIKE, K. **The Intonation of American English**. Ann Arbor, MI: University of Michigan Press, 1945.
- PITTAM, J. The long term spectral measurement of voice quality as a social and personality marker : a review. In **Tadjen**, Kris, 1-12,1987.
- PODESVA, R.J. Gender and the social meaning of non-modal phonation types. **Proceedings of the Berkeley Linguistics Society**, v.37, 427-448, 2013.
- SUNDBERG, J.; NORDENBERG, M., “Effects of vocal loud- ness variation on spectrum balance as reflected by the al- pha measure of long-term-average spectra of speech” *J. Acoust. Soc. Am.*, **120**, 453–457 (2006).
- TODAKA, Y. A preliminary study of voice quality differences between Japanese and American English: Some pedagogical suggestions. **JALT journal**, 17(2), 261-268, 1995.

Recebido em: 14/09/2015

Aceito em: 09/11/2015