

A PERCEPÇÃO DA QUALIDADE DE VOZ POR BRASILEIROS BILÍNGUES

Ana Paula Petriu Ferreira Engelbert^{1*}

¹Universidade Tecnológica Federal do Paraná, Curitiba, Paraná, Brasil

Denise Cristina Kluge^{2**}

²Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil

Resumo

Estudos revelam que falantes bilíngues podem alterar a voz quando falam uma língua estrangeira (LE) em comparação com a voz em sua L1. Resta saber se essas diferenças são percebidas pelos ouvintes. Assim, o presente estudo aborda a percepção da qualidade de voz por ouvintes brasileiros bilíngues em emissões em português brasileiro (PB) e em inglês (IN) produzidas por falantes também brasileiros e bilíngues. Tal objetivo foi alcançado por meio de um teste de discriminação das vozes em PB e em IN do mesmo falante. Os ouvintes julgaram se as vozes eram iguais ou diferentes, descrevendo suas características caso fossem diferentes. Os resultados apontaram para uma certa variabilidade nos julgamentos, mas também mostraram que os ouvintes são capazes de identificar diferenças de *pitch* e intensidade entre as línguas, bem como atribuir características de personalidade e emoção à fala nas duas línguas.

Palavras-chave: Qualidade de Voz; Percepção; Fala Bilíngue.

THE PERCEPTION OF VOICE QUALITY BY BRAZILIAN BILINGUALS

Abstract

Research has shown that bilingual speakers can alter their voice when speaking a foreign language in comparison to their voice in L1. It remains to be seen whether listeners perceive these differences. Thus, the present study explores the perception of voice quality by bilingual Brazilian listeners in speech in Brazilian Portuguese (BP) and English (EN) produced by bilingual Brazilian speakers. To achieve this objective, a

* Professora adjunta da Universidade Tecnológica Federal do Paraná – Câmpus Pato Branco, vinculada ao Programa de Pós-graduação em Letras. Possui graduação em Letras-Inglês pela Universidade Federal do Paraná - UFPR (2003), mestrado (2007) e doutorado (2015) pela mesma universidade. Seus interesses de pesquisa são a percepção e produção da fala em língua materna e estrangeira, ensino-aprendizagem de inglês como língua estrangeira e formação de professores de inglês. Seu e-mail é anapaulapetriu@gmail.com. ORCID: <https://orcid.org/0000-0003-4904-607X>.

** Professora adjunta da Universidade Federal do Rio de Janeiro (UFRJ) vinculada ao Departamento de Letras Anglo-germânicas da Faculdade de Letras e ao Programa de Pós-graduação em Letras da Universidade Federal do Paraná (UFPR). Possui graduação em Letras pela Universidade do Vale do Rio dos Sinos - Unisinos (2000), mestrado em Letras Inglês e Literatura Correspondente pela Universidade Federal de Santa Catarina - UFSC (2004) e doutorado (2009) e pós-doutorado (2016) em Letras Inglês também pela UFSC. Seus interesses de pesquisa são: percepção e produção de sons nativos e não-nativos, efeito de pistas visuais, treinamento perceptual, aquisição/aprendizagem de segunda língua e ensino de pronúncia. Seu e-mail é deniseckluge@gmail.com. ORCID: 0000-0003-4656-7902.



voice discrimination task with semantically similar speech samples from the same speaker was carried out. Listeners judged if voices in LE were the same or different than voices in L1, describing their characteristics in case they were different. Results point to certain variability in judgements, but also reveal that listeners are capable of identifying differences in pitch and intensity between languages, as well as conferring characteristics of personality and emotion to the speech samples in both languages.

Keywords: Voice Quality; Perception; Bilingual Speech.

1. Introdução

Soamos diferentes quando falamos uma língua estrangeira¹ (LE)? Pelo menos de forma impressionística, tal mudança na voz parece ser recorrente em comparação à fala em língua materna (L1). Esta pesquisa buscou investigar se pessoas falando português brasileiro (PB) como L1 soam diferentes quando falam inglês (IN) como LE do ponto de vista perceptual. Quando falamos, nossa voz fornece uma série de informações sobre nossas características tanto físicas quanto psicológicas. Podemos reconhecer quando uma pessoa está cansada ou entediada, distraída ou doente, quando quer contar um segredo ou está entusiasmada. Embora nós, como ouvintes, percebamos todos esses elementos que as vozes nos revelam, nosso julgamento nem sempre é preciso; ainda assim, geralmente sabemos se uma pessoa é um homem ou uma mulher, se está cansada ou irritada, doente ou animada, além de percebermos atitudes do falante, como ironia e sarcasmo, competência e credibilidade.

Segundo Kreiman e Sidtis (2011), definir os termos “voz” e “qualidade de voz” não é tarefa fácil, já que os dois têm sido frequentemente usados de maneira intercambiável. Essas autoras defendem definições baseadas no uso científico dos termos, de modo que a “voz” tenha uma base física e fisiológica que remete ao sinal acústico (fala), e a “qualidade de voz” seja referente à impressão perceptiva resultante do sinal acústico, de maneira análoga à distinção entre frequência (propriedade física da vibração) e *pitch* (sensação do ouvinte). Neste trabalho, adotamos a definição de “voz” como sinônimo de “fala” para contemplar tanto os aspectos laríngeos quanto os supralaríngeos da produção, e utilizamos o termo “qualidade de voz” como a resposta de um ouvinte ao som da fala.

Com relação à voz e à qualidade de voz em LE, Lee e Sidtis (2017, p.11) afirmam que, baseadas em seus resultados, “falantes bilíngues produzem padrões de vozes notadamente distintos entre línguas”, e que “essa diferença implica que a língua seja um fator adquirido que contribui para a manifestação da variabilidade de atributos vocais intra e inter-falantes”.² Vale ainda ressaltar que quando nos referimos a “falantes bilíngues” (ou multilíngües) neste trabalho, consideramos aqueles falantes com diversos graus de conhecimento e uso de duas (ou mais) línguas, definidos por Grosjean (2010, p. 4) como “pessoas que precisam e usam duas ou mais línguas (ou dialetos) nas suas vidas cotidianas”. Partindo do pressuposto da existência de diferenças vocais entre línguas, como também verificado por Engelbert et al. (2016), o principal objetivo deste estudo foi

investigar se ouvintes bilíngues percebem diferenças entre emissões em PB e em IN por meio de um experimento de percepção da qualidade de voz.

2. Qualidade de voz e fala em LE

A qualidade de voz é a matéria-prima da prosódia, sendo responsável por veicular tanto características linguísticas como paralinguísticas na comunicação. Suas funções, especialmente no domínio paralinguístico, envolvem a expressão da emoção e a intenção do falante, bem como fornecem características psicológicas e de personalidade dos falantes. Para Laver (1980), a qualidade de voz é:

... uma abstração cumulativa de um falante em um período de tempo, caracterizando-se como qualidade, que é acumulada a partir das flutuações momentâneas e espasmódicas de articulações de curto-termo, usadas pelo falante para a comunicação linguística e paralinguística. (LAVÉR, 1980, p. 1)³

Laver utiliza a definição que corresponde à percepção do resultado da ação coordenada do sistema respiratório, pregas vocais, língua, mandíbula, lábios e palato mole (KREIMAN et al., 2005). Tal definição, contudo, não especifica a contribuição do ouvinte na interpretação da qualidade de voz que é, por sua vez, um fenômeno perceptual por natureza. Segundo Kreiman et al. (2005), diferentes aspectos da percepção da qualidade de voz dependem do contexto, da atenção, da experiência do ouvinte, sendo também afetados pelo tipo de tarefa perceptiva que se desempenhe.

Segundo Esling (2013), pelo fato de a qualidade de voz ser socialmente indicial,⁴ indivíduos bilíngues ou multilíngues mudam a voz quando falam línguas diferentes ou mesmo combinam ajustes das línguas que falam. Essas mudanças, contudo, podem ter motivações diversas. 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. Os resultados de Chen e Bond (2010), que investigaram a percepção da personalidade e a acomodação cultural em diálogos entre chineses bilíngues e estadunidenses, mostraram que estes são percebidos como tendo personalidade mais extrovertida, emocionalmente mais estável, mais aberta e mais assertiva que os chineses. Portanto, para os falantes bilíngues do presente estudo, possíveis mudanças na voz entre o PB e o IN podem ter origem nas diferenças culturais que existem entre a sociedade brasileira e a estadunidense.⁵

Para nosso conhecimento, não há na literatura um expressivo número de pesquisas sobre a qualidade de voz e a fala bilíngue, não sendo possível, portanto,

estabelecer interpretações inequívocas sobre as motivações que levam um falante a produzir vozes diferentes em línguas diferentes. Por conta da pouca atenção que aspectos prosódicos têm recebido na pesquisa sobre a aprendizagem da LE, pouco se sabe sobre como um aprendiz de LE adquire elementos prosódicos da língua-alvo – ou mesmo se esses elementos são “ensináveis” e “aprendíveis”. Por isso, ainda não está claro para aprendizes e professores onde reside a dificuldade, se na produção dos elementos prosódicos, na percepção dos mesmos, ou se nas duas esferas. Isso se reflete na ausência de materiais didáticos que proponham o aumento da consciência do aprendiz de LE sobre a importância das características prosódicas da LE nos atos de comunicação. Mais especificamente com relação às pesquisas sobre qualidade de voz, o número também reduzido de estudos sobre a voz dos bilíngues ainda conta com a grande variedade de formas de medir ou avaliar o fenômeno, de encaminhamentos metodológicos, e de tratamento e apresentação dos resultados. Ademais, as tarefas desempenhadas pelos participantes das pesquisas variam bastante, dificultando a comparação entre os resultados e possíveis explicações para o fenômeno.

3. Percepção da qualidade de voz em LE

Reconhecer vozes é uma habilidade desenvolvida muito cedo, a partir do momento em que o bebê identifica a voz da mãe e gradualmente expande seu repertório de vozes familiares (KREIMAN; SIDTIS, 2011). Não há limite de número de vozes que compõem esse repertório, que vai sendo construído ao longo da vida de uma pessoa. Entretanto, não somente ouvimos vozes familiares em nosso dia-a-dia; entramos em contato com inúmeras vozes que podem tanto continuar a ser não-familiares, ou podem tornar-se familiares com o tempo. A pesquisa na área da percepção da voz considera que usamos diferentes tipos de processamentos cognitivos para a percepção da voz: (a) o processamento por traços, que pressupõe a decomposição do estímulo em elementos aditivos e relevantes; (b) o processamento por reconhecimento de padrões, que implica a percepção como um todo (KREIMAN; SIDTIS, 2011, p. 158). Na primeira abordagem, considera-se que o processo envolva a percepção e classificação da presença ou ausência de uma série de características previamente estabelecidas. Na segunda abordagem, certas características do estímulo levam à correta interpretação da interação entre as partes e o todo, que por sua vez é determinado pela natureza intrínseca do estímulo e dependem de contexto, familiaridade e memória. Essa diferença entre as duas abordagens faz todo o sentido para a pesquisa em percepção da voz ou da sua qualidade; a tarefa que o ouvinte desempenha em um experimento de percepção pode exigir que ele faça o reconhecimento do padrão da voz como um todo (tarefas de identificação ou reconhecimento) ou atentar aos detalhes característicos da voz (tarefas de discriminação). As tarefas de identificação e de reconhecimento implicam que a voz seja familiar ao ouvinte, enquanto que a tarefa de discriminação pode ser realizada com vozes não-familiares.

Na discriminação de vozes não-familiares, a análise de elementos característicos e a comparação entre eles têm papel fundamental. O ouvinte extrai características acústicas básicas como *pitch*, qualidade de voz e taxa de elocução, além de atributos de outra natureza como “inteligência” ou “masculinidade”. Kreiman (1987) afirma que o desempenho de ouvintes normais na tarefa discriminatória sugere que ambas as características gerais e as específicas da voz são utilizadas para comparar vozes não-familiares.

No caso da discriminação de vozes em língua estrangeira, se o ouvinte não conhece o inventário fonético da LE, muito provavelmente terá dificuldades em identificar o que é característica específica do falante do que é característica da língua. Segundo Kreiman e Sidtis (2011, p. 241), há quatro possibilidades para a identificação e discriminação de falantes envolvendo LEs: (1) o falante e o ouvinte têm a mesma língua materna; (2) o falante e o ouvinte falam a mesma língua, porém um como língua materna e o outro como LE; (3) o falante e o ouvinte têm uma LE em comum; (4) o falante usa uma LE que o ouvinte não conhece. A princípio, a familiaridade com a língua parece contribuir para um melhor desempenho na discriminação de vozes ou de falantes.

Estudos sobre a discriminação de vozes de falantes bilíngues geralmente visam à identificação de evidências que separam as características indiciais do falante das características linguísticas da L1 e da LE. Goggin et al. (1991) testaram o papel da familiaridade com a língua na identificação da voz. Esses autores realizaram experimentos de identificação com estímulos de dois grupos de bilíngues: IN/alemão e IN/espanhol. Os resultados mostraram que: (a) ouvintes monolíngues, falantes de IN, identificaram as vozes dos bilíngues com mais eficácia quando as amostras eram no IN do que quando eram no alemão; (b) ouvintes monolíngues alemães também tiveram melhor desempenho na identificação do falante nos estímulos em alemão; (c) ouvintes monolíngues, falantes de IN, identificaram melhor as vozes dos bilíngues nas amostras em IN do que nas em espanhol; (d) ouvintes bilíngues IN/espanhol não mostraram padrão de identificação definido, mas a identificação dos falantes que apresentavam sotaque em IN foi menor do que quando a amostra não tinha sotaque. Em conclusão, os autores afirmam que a confiança na identificação das vozes dobra quando o ouvinte compartilha da língua apresentada no estímulo em comparação com quando ouve uma língua estrangeira.

Winters et al. (2008) investigaram até que ponto a familiaridade com a língua afeta a percepção de propriedades indiciais da fala. Para tanto, os autores realizaram testes de identificação e discriminação da fala da produção de bilíngues (IN/alemão). Em um experimento, os ouvintes foram treinados a identificar os falantes bilíngues em apenas uma das línguas, e testados na sua habilidade de identificar os mesmos falantes produzindo na outra língua. Em outro experimento, os ouvintes discriminavam os falantes bilíngues nas duas línguas por um teste do paradigma AX. Os resultados dessa pesquisa mostraram que há informação indicial que independe da língua (*language-independent*) em quantidade suficiente na fala, de modo que ouvintes conseguem discriminar os

falantes bilíngues, não importando qual língua falem. Aparentemente, os ouvintes processam a informação indicial pelas características *language-dependent* quando eles têm familiaridade com a mesma; por outro lado, realizam tarefas de discriminação baseados nas características *language-independent* quando não conhecem a língua. Assim, os autores concluíram que a identificação do falante pelo ouvinte é feita com base tanto em características dependentes da língua quanto em independentes da língua.

O estudo de Wester (2012) verificou a identificação de falantes bilíngues em três grupos diferentes – IN/alemão, IN/finlandês e IN/mandarim. Os ouvintes, 14 para cada par de línguas, eram estadunidenses monolíngues e realizaram tarefa de discriminação de amostras de fala nas duas línguas, decidindo se as frases eram faladas pela mesma pessoa ou não. Os ouvintes apresentaram altos índices de identificação nos julgamentos IN-IN, bem como em julgamentos LE-LE (e.g. alemão-finlandês), desempenhando de maneira menos acurada quando os estímulos eram IN-LE. Os autores argumentam que, no caso dos julgamentos LE-LE, os ouvintes não eram “distráídos” pelas características linguísticas do sinal da fala, ao contrário do que acontecia nos pares IN-LE.

A percepção e a discriminação de falantes bilíngues também interessam a pesquisadores na área de desenvolvimento de tecnologia de reconhecimento e síntese de fala. Wester et al. (2010), pesquisadores participantes de um projeto internacional para o desenvolvimento de um tradutor de fala que mantivesse as características vocais do falante na versão em LE,⁶ relataram a necessidade da realização de pesquisas sobre a avaliação da similaridade da voz da fala bilíngue. As questões que precisam ser investigadas são se o falante soa a mesma pessoa em L1 e LE e se o ouvinte reconhece a similaridade do falante entre as línguas. Embora os autores admitam que o falante possa soar diferente quando fale línguas distintas, eles acreditam que seja muito provável identificá-lo nas duas línguas em comparação com outros falantes. Já a questão da percepção da similaridade pelo ouvinte parece ser mais complicada de responder. Os fatores que influenciam a discriminação e identificação das vozes dos bilíngues, já reportados pela pesquisa na área, são: (a) a familiaridade do ouvinte com a LE; (b) a influência dos fatores linguísticos no julgamento da voz; (c) a dificuldade na separação das pistas *language-dependent* das *language-independent* utilizadas pelos ouvintes. Em suma, os autores afirmam que é de extrema importância que se leve em conta as línguas que o ouvinte conhece e as línguas produzidas pelo falante.

Portanto, o presente trabalho propôs um teste de discriminação da voz de brasileiros bilíngues falando português e IN como LE, realizada por ouvintes que compartilham da língua materna e da LE dos falantes. Nosso objetivo foi que os ouvintes atentassem para a qualidade de voz dos falantes, na tentativa de detectar possíveis mudanças entre as duas línguas. Houve familiaridade com as línguas para a execução da tarefa, e os ouvintes possivelmente utilizaram tanto pistas *language-dependent* quanto *language-independent* para realizar o teste. Os fatores linguísticos têm influência tanto na L1 como na LE, já que os ouvintes são nativos na primeira e proficientes na segunda, respectivamente.

4. Metodologia

O principal objetivo deste estudo foi verificar se ouvintes bilíngues percebem diferenças de qualidade de voz, quando estas ocorrem, na fala também de bilíngues. Para acessar esse julgamento, um experimento de percepção foi desenhado.

O teste de percepção da qualidade de voz de falantes bilíngues foi realizado pelo software livre TP – Teste de Percepção (RAUBER et al., 2012), na versão 3.172. Foi montada uma tarefa de discriminação, na qual os ouvintes julgaram as vozes dos bilíngues como “iguais” ou “diferentes”.

4.1 Os ouvintes

Como ouvintes do experimento, foram recrutados brasileiros bilíngues que atuavam como professores de IN e que tivessem pelo menos um ano de experiência no ensino dessa língua.⁷ As condições para que o sujeito participasse do experimento foram: (a) ser brasileiro e falar o PB como língua materna; (b) ter a mesma faixa etária que os participantes do experimento de produção⁸ que gerou os estímulos para o teste de percepção (entre 20 e 39 anos); falar IN como LE com proficiência; (c) declarar não possuir deficiência auditiva. Por meio de um questionário, foram levantadas as seguintes informações sobre os participantes: sexo (sete homens e sete mulheres); origem (12 participantes de Curitiba e região metropolitana, 1 de Santos e 1 de Diadema, no estado de SP) e idade (média de 25,6 anos). Os ouvintes também relataram sua experiência (em meses) em países de IN e fizeram autoavaliação de sua proficiência em porcentagem (sendo 0% não-proficiente e 100%, totalmente proficiente), como mostra a TABELA 1:

Tabela 1: experiência em país de IN autoavaliação de proficiência e sotaque estrangeiro dos ouvintes

Ouvinte	Experiência em país de IN (meses)	Autoavaliação Proficiência (%)
Ouv1	12	90
Ouv2	0	70
Ouv3	2	90
Ouv4	1	60
Ouv5	1	60
Ouv6	0	100
Ouv7	0	70
Ouv8	0	85
Ouv9	0	60
Ouv10	12	90
Ouv11	1	70
Ouv12	1	80

Ouv13	0	80
Ouv14	0	70

4.2 Os estímulos

Os estímulos do teste, baseados nos quais foram feitos os julgamentos, são compostos de trechos da fala semiespontânea⁹ produzida por 16 brasileiros bilíngues. Foram selecionadas amostras similares semanticamente no PB e no IN – isso foi possível, pois foi instruído ao participante no experimento de produção que respondesse à mesma pergunta em PB e em IN, preferencialmente mantendo conteúdo. Cada *trial* foi composto por um arquivo de áudio em que a amostra em PB foi concatenada à amostra similar em IN, amostras essas do mesmo falante, e apresentadas nessa ordem (PB e depois, IN). O mesmo *trial* foi apresentado duas vezes, totalizando 32 *trials* (duas vezes cada *trial* dos 16 participantes).

4.3 As instruções e a tarefa

Primeiramente, os ouvintes responderam ao questionário de informações pessoais. Depois, foram instruídos quanto ao uso do *software* e quanto à tarefa que teriam que desempenhar. Foi feita a familiarização dos participantes com o software e a tarefa por meio de um *trial* piloto. Esse *trial* piloto foi constituído de uma amostra de fala em PB e em IN de uma pessoa que não participou do experimento de produção, para servir de exemplo aos ouvintes. A tarefa foi explicada aos ouvintes da seguinte forma¹⁰:

1. Você ouvirá estímulos (*trials*) contendo amostras de fala em PB e em IN, nessa ordem, de uma mesma pessoa. Ao todo, são 16 pessoas diferentes, brasileiros bilíngues falantes de IN como LE. Sua tarefa é analisar e comparar a voz da pessoa quando fala em IN com a voz em PB: não leve em conta desvios sintáticos ou de pronúncia; preste atenção apenas na voz. Você poderá ouvir o mesmo estímulo até cinco vezes antes de se decidir e então deve clicar no botão “igual” ou no botão “diferente” (FIGURA 1):

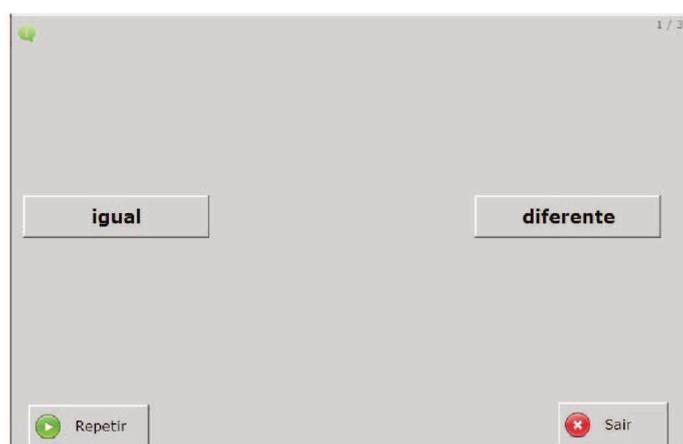


Figura 1: Tela do TP para o julgamento das vozes em “igual” ou “diferente”

2. Após clicar em um dos botões “igual” ou diferente”, você verá uma escala deslizante que vai de “pouco” a “muito”. Se você considerou a voz do falante “igual” nas duas línguas, desconsidere a escala deslizante; apenas utilize-a para quantificar a diferença caso tenha clicado em “diferente”. Você deve quantificar a diferença entre “pouco” e “muito” (FIGURA 2).

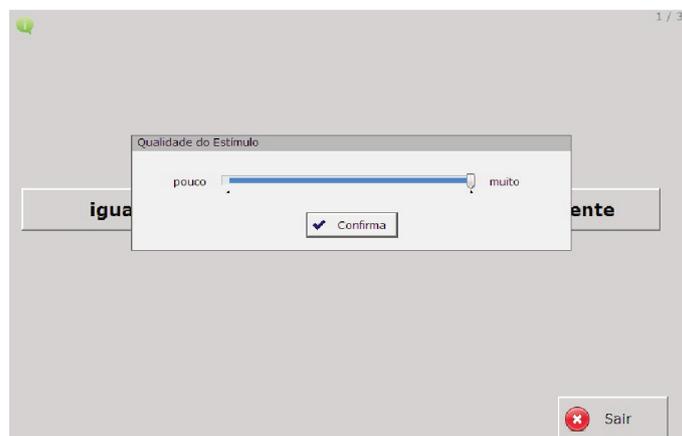


Figura 2: Tela do TP com a escala deslizante para a quantificação da diferença entre as vozes.

3. Caso você julgue que a voz seja diferente no IN em relação ao PB, você terá que dizer qual é essa diferença: você poderá descrever a diferença usando a palavra ou expressão que desejar.¹¹

A partir dos julgamentos, foram geradas planilhas com os resultados da discriminação entre “igual” e “diferente” e com a quantificação das diferenças. Se o ouvinte considerou que o falante soava igual em IN e PB, o valor atribuído ao julgamento foi de zero. Se o ouvinte considerou que o falante soava diferente em IN em comparação ao PB e quantificou essa diferença na escala deslizante, foram gerados valores de 1 a 100, sendo 1 considerado muito pouca diferença e 100, muita diferença.

Como análise inicial foi reportada a estatística descritiva dos dados: média e desvio-padrão dos julgamentos dos 14 ouvintes em relação à produção dos 16 falantes. Além disso, foi rodado o teste Pearson para verificar a correlação entre os julgamentos a fim de verificar a concordância entre os ouvintes. Baseado nos resultados desse teste, foi criado um subgrupo dos dez ouvintes que obtiveram maior número correlação nos julgamentos uns com os outros. Ademais, os falantes foram agrupados (grupo 1 e grupo 2) conforme o julgamento desse subgrupo de ouvintes.

5. Resultados

A partir do experimento de percepção, foram geradas médias dos julgamentos igual/diferente feitos pelos ouvintes com relação às vozes dos falantes em PB e

em IN. Como reportado na seção sobre metodologia, os valores gerados pelo programa TP vão de 0 (vozes iguais em PB e em IN) a 100 (vozes muito diferentes entre o PB e o IN).

De maneira geral, houve certa variabilidade nos julgamentos feitos pelos ouvintes. Essa variabilidade pode ser verificada a partir do desvio-padrão das médias obtidas pelos julgamentos (TABELA 2):

Tabela 2: média e DP dos julgamentos de diferenças entre vozes em PB e em IN pelos ouvintes

	Média	DP
Fal1	58,0	26,7
Fal2	13,4	27,2
Fal3	27,7	25,9
Fal4	19,5	27,4
Fal5	26,7	31,3
Fal6	15,4	24,5
Fal7	76,0	20,8
Fal8	13,2	24,4
Fal9	57,0	28,6
Fal10	23,6	28,9
Fal11	58,8	26,5
Fal12	17,1	26,5
Fal13	19,8	27,3
Fal14	14,1	27,2
Fal15	39,1	31,1
Fal16	28,8	30,2

A partir dessas médias apresentadas na Tabela 2, verificou-se que os falantes que demonstraram maiores diferenças entre as vozes do PB e do IN foram os falantes 1, 7, 9, e 11 (médias entre 57 e 76). Os falantes 3, 5, 10, 15 e 16 apresentaram diferenças um pouco menores (entre 23,6 e 39,1), enquanto que os falantes 2, 4, 6, 8, 12, 13 e 14 tiveram as menores médias (entre 13,2 e 19,8). Os valores do DP foram em geral bastante altos, sendo os maiores valores atribuídos aos falantes 5, 15 e 16 (acima de 30) e o mais baixo, ao falante 7, que também apresentou a maior diferença entre as vozes do PB e do IN. A variabilidade revelada pelos valores de DP reitera o fato de que, segundo Kreiman e Gerratt (1998), ouvintes individuais são razoavelmente consistentes no julgamento de determinadas qualidades de voz, mas a variância dos julgamentos entre os ouvintes pode ser altíssima, devido a outros fatores como estratégia perceptual e atenção a diferentes aspectos do estímulo.

Para verificar a consistência entre os ouvintes no julgamento das diferenças nas vozes em PB e em IN, foi rodado o teste Pearson com o intuito de conferir a correlação entre os julgamentos. Os ouvintes que obtiveram um maior número

de correlações (de oito a onze) uns com os outros foram os de número 1, 2, 6, 7, 8, 10, 11, 12, 13 e 14 ($p < 0,05$). Os outros ouvintes tiveram poucas correlações uns com os outros – de zero a sete. Dessa forma, foram calculadas as médias e DP dos julgamentos feitos por esses dez ouvintes que apresentaram maior número de correlações (TABELA 3):

Tabela 3: média e DP dos julgamentos de diferenças entre vozes em PB e IN pelos ouvintes que mais obtiveram correlações

	Média	DP
Fal1	54,3	27,7
Fal2	8,3	16,6
Fal3	23,4	21,0
Fal4	9,9	19,3
Fal5	32,8	34,1
Fal6	10,0	20,5
Fal7	76,9	17,3
Fal8	5,1	14,6
Fal9	56,9	24,9
Fal10	19,2	27,7
Fal11	62,3	23,8
Fal12	7,1	12,8
Fal13	12,0	21,6
Fal14	5,6	11,6
Fal15	42,3	29,7
Fal16	18,4	20,4

A maioria das médias e DP apresentadas na Tabela 3 foram menores para esse subgrupo de ouvintes. Apenas as médias dos falantes 5, 7, 11 e 15 foram mais altas. Os valores de DP foram também mais baixos a não ser pelos falantes 1 e 5. Assim, para as análises seguintes, serão considerados apenas esses dez ouvintes que apresentaram maior número de correlações entre si.

A fim de agrupar os falantes conforme as médias de diferença entre as vozes em PB e em IN atribuídas pelos ouvintes, foi rodada a análise aglomerativa de k-médias, conforme a Tabela 4:

Tabela 4: resultados da análise aglomerativa de K-médias para o agrupamento dos falantes

Número do Falante	Grupo (Cluster)	Distância
1	1	91,83
2	2	32,23

3	2	80,50
4	2	61,28
5	2	128,47
6	2	61,29
7	1	0
8	2	44,37
9	1	98,23
10	2	73,30
11	1	80,52
12	2	40,30
13	2	54,37
14	2	0
15	1	136,58
16	2	77,34

Os resultados apresentados na Tabela 4 mostram que os falantes foram separados em dois grupos: no grupo 1, estão os falantes 1, 7, 9, 11 e 15, e no grupo 2, os falantes 2, 3, 4, 5, 6, 8, 10, 12, 13, 14, e 16. O grupo 1 apresentou as maiores diferenças entre as vozes em PB e em IN, enquanto que o grupo 2 apresentou as diferenças mais sutis entre essas duas línguas.¹² A Tabela 5 traz os resultados do teste Anova gerado pela análise aglomerativa:

Tabela 5: resultados do teste ANOVA sobre a análise aglomerativa de K-médias

ANOVA						
	Cluster		Erro		Z	Sig.
	Quadrado Médio	Df	Quadrado Médio	Df	Quadrado Médio	df
Julg_Ouv1	8665,78	1	488,48	14	17,74	,001
Julg_Ouv2	5557,65	1	193,91	14	28,66	,000
Julg_Ouv6	10857,14	1	351,99	14	30,84	,000
Julg_Ouv7	10902,84	1	566,80	14	19,23	,001
Julg_Ouv8	4921,09	1	260,70	14	18,87	,001
Julg_Ouv10	8744,40	1	459,99	14	19,01	,001
Julg_Ouv11	5871,05	1	287,42	14	20,42	,000
Julg_Ouv12	5863,309	1	372,486	14	15,741	,001
Julg_Ouv13	6169,205	1	188,289	14	32,765	,000
Julg_Ouv14	3371,598	1	163,474	14	20,625	,000

Como mostra a Tabela 5, os valores do quadrado médio do cluster revelam que os ouvintes que apresentaram maior poder de discriminação entre os grupos foram os de número 1, 6, 7 e 10. Por outro lado, os ouvintes 2, 13 e 14 tiveram maior poder explicativo para a constituição dos dois grupos de falantes (menores valores do quadrado médio do erro). Os resultados de Z, dado pelo

quociente entre o quadrado médio do cluster e o quadrado médio do erro, indicam que quem mais contribuiu para a definição dos dois grupos foram os ouvintes 2, 6 e 13.

Além de os ouvintes quantificarem as diferenças entre as vozes em PB e em IN, eles tinham também que descrevê-las, sempre caracterizando as vozes em IN em relação ao PB. Assim, os dez ouvintes descreveram as vozes em IN dos cinco falantes do grupo 1, os que apresentaram maior grau de diferenças. No Quadro 1, ao lado das palavras utilizadas pelos ouvintes, há o número de ocorrências dos termos (entre parênteses), considerando-se os 32 julgamentos (duas vezes cada falante):

Quadro 1: descrições das diferenças entre as vozes (julgamento da voz em IN em relação à voz em PB)

Julgamento dos ouvintes
Fal 1 (Homem): mais grave (9), mais baixa (9), mais lento (6), mais inseguro, contido (4), menos variação de tom (4), mais fraca (2).
Fal 7 (Mulher): mais baixa (14), mais suave (6), mais sussurrada (5), mais aguda (2), mais contida (2), menos confiante (2).
Fal 9 (Homem): mais alta (18), mais assertiva (4), mais aguda (3), mais confiante (2), mais forte (2), mais expressiva (1).
Fal 11 (Mulher): mais alta (18), mais aguda (8), mais confiante (4), maior variação de tom (3), mais rápida (2) mais empolgada (1), mais anasalada (1).
Fal 15 (Mulher): mais grave (9), mais baixa (9), menos expressiva (3), mais crepitante (2), menos confiante (2), mais rouca (2), menos variação de tom.

A descrição feita pelos ouvintes seguiu o critério de caracterizar a voz em IN em relação à voz em PB da mesma pessoa. As descrições revelaram que aspectos relativos ao *pitch* e à intensidade são bastante salientes na percepção das vozes. Houve também a descrição de aspectos de personalidade e de emoção.

A voz do falante 1, por exemplo, foi caracterizada como sendo mais grave e baixa em IN, o que possivelmente culminou em uma interpretação de que seria mais contido e inseguro nessa LE. A falta de confiança também esteve presente na descrição da voz das falantes 7 e 15. Por outro lado, os falantes 9 e 11 apresentaram acréscimo de intensidade em IN, o que resultou na interpretação das vozes como mais confiantes, assertivas e expressivas.

6. Discussão

Os resultados do experimento mostraram que houve relativa variabilidade no julgamento das vozes produzidas pelos falantes bilíngues feito pelos ouvintes leigos também bilíngues. Segundo Kreiman et al. (2005), de forma geral, diferentes aspectos da percepção da qualidade de voz dependem do contexto, da atenção, da experiência do ouvinte, sendo também afetados pelo tipo de tarefa perceptiva que se desempenhe. Por isso, os julgamentos perceptivo-auditivos de uma voz geralmente não são constantes entre ouvintes distintos.

Desta forma, para que fosse possível obter generalizações sobre as diferenças e diminuir tal variabilidade entre os julgamentos, foi necessário selecionar os ouvintes mais consistentes entre si. Assim, foi verificado que dez dos 14 ouvintes produziram julgamentos compatíveis a respeito das diferenças de qualidade de voz produzidas pelos falantes entre o PB e o IN. Dos quatro ouvintes que foram menos consistentes, três autoavaliaram sua proficiência em IN em 60%, mostrando ser possível que o grau de proficiência tenha um impacto nos julgamentos.

Acredita-se que a tarefa realizada no experimento tenha contribuído para a confiabilidade dos julgamentos entre esses dez ouvintes, pois consistia em comparar emissões em IN em relação às em PB do mesmo falante. Segundo Kreiman et al. (2007), quando é fornecido ao ouvinte um estímulo de referência para o seu julgamento, a dependência de padrões internos de qualidade de voz é diminuída e a concordância entre os ouvintes aumenta. Além disso, o fato de os ouvintes terem familiaridade com ambas as línguas dos falantes também contribuiu para a confiabilidade dos julgamentos. Assim, a partir dos julgamentos feitos pelos dez ouvintes mais consistentes entre si, os 16 falantes foram separados em dois grupos estabelecidos pela quantidade de diferenças entre as vozes em PB e em IN: o grupo que apresentava maior distinção na voz em IN em comparação à voz em PB (grupo 1), e o grupo que apresentou menor diferença entre as línguas (grupo 2).

Na descrição das diferenças feita pelos ouvintes para os falantes do grupo 1 sobre as características acústicas, foi verificada uma relativa consistência entre os julgamentos. Na maioria dos casos, os ouvintes apontaram diferenças de *pitch* entre as vozes. Foram utilizados termos como “grave” e “aguda” para designar as distinções entre PB e IN, além das descrições relativas à variação do tom da voz. Para quatro dos cinco falantes, as descrições de *pitch* realizadas pelos ouvintes corresponderam às diferenças médias de f_0 entre o PB e o IN presentes no resultados de Engelbert et al. (2016). Em três dos cinco falantes, as diferenças de intensidade da voz foram adequadamente captadas pelos ouvintes. Contudo, três dos cinco falantes produziram vozes mais agudas em IN que em PB, contrariando a tendência de vozes com *pitch* mais baixo em IN que em PB. Os ouvintes também empregaram os termos “alta” e “baixa” para diferenças de intensidade e “mais suave”, “mais sussurrada” e “crepitante” para diferenças gerais da voz.

Houve ainda a utilização de termos relacionados às características emocionais e de personalidade dos falantes na descrição do julgamento dos ouvintes. Vocábulos associados à personalidade como assertividade, confiança e segurança na fala dos bilíngues foram recorrentes. Houve também descrição da emoção dos falantes, como o uso dos termos “mais empolgada”. 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.

7. Considerações finais

Esta pesquisa trouxe resultados concernentes à percepção da qualidade de voz entre línguas por ouvintes leigos. Ficou evidenciado que ouvintes brasileiros bilíngues percebem diferenças entre as vozes em PB e IN de falantes brasileiros também bilíngues. Embora haja grande variabilidade na percepção de tais diferenças, este estudo mostrou que, por meio de tarefa de discriminação e da familiaridade com as línguas, foi possível detectar certa consistência entre os julgamentos de dez dos 14 ouvintes. As diferenças mais notáveis entre o PB e o IN nas vozes foram relacionadas ao *pitch* do falante e ao volume do sinal. Segundo Kreiman e Sidtis (2011), nossa audição responde a medidas de *pitch* e intensidade de maneira bastante estável, inclusive entre indivíduos.

Mais estudos relacionados ao assunto são necessários para que sejam possíveis maiores generalizações acerca da percepção da qualidade de voz de bilíngues por bilíngues. Há questões relacionadas ao tipo de fala (fala lida, roteirizada, semiespontânea e espontânea) e ao tipo de tarefa que o ouvinte precisa realizar (tarefa de discriminação, descrição e quantificação) que se mostram fatores influentes nos resultados. Além disso, há de se averiguar se fatores como experiência linguística e grau de sotaque estrangeiro têm efeito na percepção da qualidade de voz entre línguas.

Em conclusão, esta pesquisa corrobora a impressão de que muitos falantes bilíngues alteram suas características vocais quando falam uma LE em relação à sua L1. Sabe-se também que essas diferenças podem ter motivações distintas – questões de personalidade, emocionais e socioculturais da produção da fala em cada língua. Como implicação pedagógica, sugere-se aumentar a consciência dos estudantes de LE em relação ao impacto que as diferentes qualidades de voz têm na comunicação como um todo, tanto em L1 como em LE.

Notas

1. Utilizamos o termo “língua estrangeira”, pois os participantes da presente pesquisa eram brasileiros que aprenderam inglês como língua estrangeira no Brasil. Para maiores informações a respeito da terminologia L2/LE, ver Eckert e Frosi, 2015.
2. “*bilingual speakers produced notably different voice patterns in their two spoken languages...*”. “*This difference implies that language alone is an acquired factor that contributes to the manifestation of within and between-speaker variability of vocal attributes*” (Traduções de responsabilidade das autoras).
3. “*a cumulative abstraction over a period of time of a speaker characterizing quality, which is gathered from the momentary and spasmodic fluctuations of short-term articulations used by the speaker for linguistic and paralinguistic communication.*”
4. Do inglês *indexical*; o que se associa ou identifica um falante (“*Indexical*”. Def. 2b. Merriam-Webster.com. Acessado em 13/04/2018 em <https://www.merriam-webster.com/dictionary/indexical>
5. As amostras de fala utilizadas no presente estudo foram gravadas nos Estados Unidos, país onde residiam os brasileiros bilíngues que participaram do experimento de produção descrito em Englebert et al (2016). Assim, justifica-se a escolha do povo em questão nesta comparação.

6. O projeto foi chamado de EMIME (*Effective Multilingual Interaction in Mobile Environments*). Para mais informações, ver www.emime.org
7. A justificativa para que o participante fosse professor de IN era uma maior garantia do grau de proficiência do mesmo, já que não foi realizado teste de nivelamento para tal fim.
8. O experimento de produção está descrito detalhadamente em Engelbert et al. (2016). Resumidamente, o experimento contou com a participação de 16 brasileiros bilíngues (dez mulheres e seis homens), com média de idade de 27,5 anos, que estavam morando nos Estados Unidos à época da coleta de dados. Os participantes haviam começado a aprender IN no Brasil (como LE) e continuavam a aprender essa língua como L2 nos EUA. A origem dos participantes era majoritariamente dos estados do sul do Brasil e de São Paulo. Não foram incluídos no estudo os dados de fala de uma participante que havia aprendido PB e IN simultaneamente quando criança. As amostras de fala foram gravadas na University of California-Los Angeles, entre janeiro e março de 2014.
9. Respostas às seguintes perguntas (em PB e em IN): 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? Não houve critério fonético para a seleção das amostras, já que o objetivo do experimento era obter julgamentos acerca da voz dos falantes, que figura no domínio paralinguístico da prosódia.
10. O experimento foi realizado nos meses de outubro e novembro de 2014 na UFPR.
11. Não foram sugeridos termos que descrevessem a voz do falante. Esperava-se que o ouvinte utilizasse adjetivos que descrevessem tanto a voz propriamente dita, como características emocionais ou de personalidade do falante. Foi pedido ao ouvinte que falasse o descritor, e a pesquisadora anotava essa palavra ou expressão no papel, já que não havia a possibilidade de escrevê-la no software.
12. A análise aglomerativa tem por objetivo separar os elementos de um grande grupo em função de sua dissimilaridade. Nesse caso, o grupo de 16 falantes foi separado em dois grupos; o falante 7 é o centro do grupo 1 e o falante 14 é o centro do grupo 2. Os outros falantes se relacionam com o centro de cada grupo através do valor expresso pela distância. Quanto maior for o valor da distância, maior a dissimilaridade do elemento em relação ao centro.

Referências

- 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.
- ECKERT, K.; FROSI, V. M. Aquisição e aprendizagem de línguas estrangeiras: princípios teóricos e conceitos-chave. *Domínios de Lingu@Gem*, v. 9, p. 198-216, 2015.
- ENGELBERT, A. P. P. F.; KLUGE, D. C.; SILVA, A. H. P. Línguas diferentes, vozes distintas: evidências da fala de bilíngues em português e inglês. *Ilha do Desterro*, v. 69, nº1, p. 033-048, 2016.
- ESLING, J. Voice Quality. *The Encyclopedia of Applied Linguistics*. Wiley Online Library, p.1-7, 2013.
- GOGGIN, J., THOMPSON, C., STTRUBE, G., SIMENTAL, L. The role of language familiarity in voice identification. *Memory Cognition*, 19 (5), 448-458, 1991.
- GROSJEAN, François (2010). *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press
- KREIMAN, J. *Human Memory for Unfamiliar Voices*. Tese de Doutorado não-publicada. Chicago-IL, University of Chicago, 1987.

- KREIMAN, J.; GERRATT, B. R. Validity of rating scale measures of voice quality. *Journal of the Acoustic Society of America*, 104 (3), 1589-1608, 1998.
- KREIMAN, J.; VANLACKER-SIDTIS, D.; GERRATT, B. R. Perception of voice quality. IN PISONI, D.B.; REMEZ, R.E (eds) *Handbook of Speech Perception*. Blackwell, Oxford, pp. 338–362, 2005.
- KREIMAN, J.; GERRATT, B. R.; ANTONANZAS-BARROSO, N. Measures of glottal source spectrum. *Journal of Speech and Hearing Research*, 50, 595–610, 2007.
- KREIMAN, J.; SIDTIS, D. *Foundations of Voice Studies: An Inter-disciplinary Approach to Voice Production and Perception*. Wiley-Blackwell, Walden, MA, 2011.
- LAVER, J. **The phonetic description of voice quality**. Cambridge: Cambridge University Press, 1980.
- LEE, B.; SIDTIS, D. V. L. The bilingual voice: vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, vol. 20 (3), 174-185, 2017.
- RAUBER, A., RATO, A., KLUGE, D.C.; SANTOS, G. R.; FIGUEIREDO, M. TP – Teste de Percepção. Versão. 3.1, 2012.
- WESTER, M., DINES, J., GIBSON, M., LIANG, H., WU, Y.-J., SAHEER, L., KING, S., OURA, K., GARNER, P., BYRNE, W., GUAN, Y., HIRSIMA, T., KARHILA, R., KURIMO, M., SHANNON, M., SHIOTA, S., TIAN, J., TOKUDA, K., YAMAGISHI, J. Speaker adaptation and the evaluation of speaker similarity in the EMIME speech-to-speech translation project. In *Proc. 7th ISCA Speech Synthesis Workshop*, Kyoto, Japão, Setembro de 2010.
- WESTER, M. Talker Discrimination Across Languages. *Speech Communication*, 54, p. 781-790, 2012.
- WINTERS, S., LEVI, S., PISONI, D. Identification and discrimination of bilingual talkers across languages. *Journal of the Acoustical Society of America*, 123 (6), 4524-4538, 2008.

Recebido em: 14/11/2017

Aceito em: 13/04/2018