

PERCEPÇÃO E PRODUÇÃO DAS PLOSIVAS DO INGLÊS (L2) POR APRENDIZES BRASILEIROS

Bruno Moraes Schwartzhaupt¹
Ana Hemmons Baratz²
Ubiratã Kickhöfel Alves³

1. Introdução

Ao direcionarmos nossa atenção aos estudos voltados à investigação de percepção e produção das consoantes plosivas iniciais do Inglês (L2)⁴ por aprendizes brasileiros, observamos que, nos últimos anos, o número de trabalhos desenvolvidos tem crescido consideravelmente. Dentre muitos outros, podemos mencionar os de Cohen (2004), Alves (2007), Reis e Nobre-Oliveira (2008), Magro (2010), Alves (2011) e França (2011). Além disso, também podemos observar que houve considerável mudança no que diz respeito aos cuidados metodológicos adotados ao se estudarem os padrões de vozeamento dessas consoantes. O recente acesso a importantes ferramentas de análise e coleta de dados, tais como o *software Praat* (BOERSMA; WEENINK, 2011), sem dúvida, fez evoluir a metodologia dos estudos na área.

A investigação desse aspecto fonético-fonológico da Língua Inglesa se faz importante devido ao caráter perceptualmente distintivo que a aspiração tem na produção das plosivas /p/, /t/ e /k/. Sem a devida instrução, o aprendiz tende a não adquirir essa aspiração, característica do Inglês, uma vez que ele associa os segmentos aspirados da L2, [p^h], [t^h] e [k^h], às produções não aspiradas do Português Brasileiro (PB), [p], [t] e [k]. Por não discriminar esses padrões de *Voice Onset Time* (VOT), a produção oral em L2 do aprendiz se dá sem a aspiração nos contextos necessários, acarretando problemas de inteligibilidade em sua comunicação em Inglês (cf. CELCE-MURCIA *et al.*, 2010). Tal problema ocorre como consequência da semelhança existente entre o padrão de VOT das plosivas surdas do PB e as plosivas sonoras da língua-alvo (ambas produzidas com VOT zero, sem aspiração).

Considerados estes pontos, o presente estudo buscou investigar a produção e percepção dos padrões de VOT das consoantes plosivas do Inglês por aprendizes brasileiros em dois diferentes níveis de proficiência. Essa investigação apresenta diversos objetivos, para as quais formulamos as seguintes questões norteadoras, que objetivamos responder:

- 1) Os aprendizes brasileiros são capazes de identificar as plosivas com VOT característico do Inglês, em ambos os níveis de proficiência?
- 2) Os índices de VOT das plosivas produzidas em Inglês pelos aprendizes são semelhantes ao padrão da língua-alvo, em ambos os níveis de proficiência?

2.1) Os índices de VOT produzidos em português por esses aprendizes são significativamente diferentes daqueles produzidos em Inglês?

¹ Acadêmico de Letras – Bolsista de Iniciação Científica da UFRGS. E-mail: bruno.schwartzhaupt@yahoo.com.br

² Acadêmica de Letras – Bolsista PET-Letras da UFRGS. E-mail: anabaratz@hotmail.com

³ Professor Adjunto do Programa de Pós-Graduação em Letras da UFRGS. E-mail: ukalves@gmail.com

⁴ Este trabalho não faz distinção entre os termos Segunda Língua (L2) e Língua Estrangeira (LE).

3) Os aprendizes são capazes de fazer distinção entre plosivas com VOT Negativo, VOT Zero Natural, VOT Zero Artificial e VOT Positivo⁵?

3.1) Há diferenças entre os níveis de proficiência na acuidade em discriminar os padrões de VOT contrastados?

3.2) Há diferenças significativas em termos de acuidade na discriminação entre os contrastes, comparando as condições:

a) VOT Zero Natural e VOT Zero Artificial?

b) VOT Zero Artificial e VOT Positivo?

Neste trabalho, primeiramente apresentamos os padrões de VOT estudados e discutindo os problemas que envolvem sua aquisição em L2. Em seguida, explicitamos a metodologia com que o estudo foi desenvolvido. Na seção seguinte, apresentamos e discutimos os resultados da investigação. Finalmente, mostramos nossas conclusões a respeito do estudo, bem como os próximos passos que pretendemos seguir.

2. Referencial teórico

Começaremos esta seção por explicitar o fato fonético-fonológico em questão, principal objeto deste estudo, o *Voice Onset Time*. Em seguida, pretendemos caracterizar a aquisição fonético-fonológica de L2 como um processo dinâmico, dependente de diversos fatores, em oposição à caracterização binária tradicional. Por fim, serão discutidos os principais problemas que estão envolvidos na aquisição dos padrões de VOT da língua-alvo, à luz de dois tipos de transferência: a fonético-fonológica, explicada claramente pelos modelos *Speech Learning Model* (FLEGE; MUNRO; MACKAY, 1995) e *L2 Perceptual Learning Model* (BEST; TYLER, 2007), e a *Transferência Grafo-Fônico-Fonológica* (ZIMMER; SILVEIRA; ALVES, 2009).

2.1 Voice Onset Time

O *Voice Onset Time* (VOT) pode ser definido como o período de tempo existente entre a explosão de uma consoante plosiva (/p/, /b/, /t/, /d/, /k/, /g/) e o início da vibração de cordas vocais da vogal que segue essa consoante. Lisker e Abramson (1964) apontam três principais categorias de VOT encontradas nas línguas do mundo: (i) o VOT negativo (também chamado *pre-voicing*), padrão de vozeamento no qual a vibração vocálica precede a soltura do segmento plosivo em um tempo que varia de 75 a 125 ms; (ii) o VOT zero, que ocorre quando o início da vibração vocálica se dá quase simultaneamente à soltura da plosiva - isto é, de 0 a 35 ms após essa soltura; e (iii), o VOT positivo (conhecido como *aspiração*), que é o padrão no qual há a soltura do segmento plosivo sem vibração de cordas, e essa vibração só inicia após um intervalo que varia de 35 a 100 ms.

Como já sugerido por Cohen (2004), Reis e Nobre-Oliveira (2008), e Alves (2010), os valores de VOT não são absolutos, e é impossível considerá-lo como uma entidade isolada dentro de qualquer sistema linguístico. Segundo Cohen (2004), o valor do VOT pode ser influenciado por fatores como a prosódia, a tonicidade, a velocidade

⁵ Tais padrões são devidamente discutidos na seção 2.1. O VOT zero artificial, fruto de um experimento de manipulação de VOT, é explicado na seção 3.4.2, referente à metodologia do estudo.

da fala, entre outros aspectos fonético-fonológicos. De acordo com os autores supracitados, os valores do VOT encontrados nas plosivas surdas produzidas em Inglês variam de 30 a 80 ms (VOT positivo): /k/ apresenta aproximadamente 80 ms; /t/, 70 ms; e /p/, 55 ms. As plosivas vozeadas, no Inglês, /b/, /d/ e /g/ tendem a ser produzidas com VOT zero (0 a 35 ms); embora, como aponta Simon (2010), a produção dessas consoantes com VOT negativo também seja possível. Diferentemente do Inglês, no PB, as plosivas surdas tendem a ser produzidas com VOT zero; /k/ apresenta aproximadamente 38 ms de VOT, /t/, 18 ms e /p/, 12 ms (REIS & NOBRE-OLIVEIRA, 2008). Além disso, no PB, as plosivas vozeadas são produzidas com VOT negativo. Assim sendo, Cohen (2004) sugere que, no que concerne ao VOT, a Língua Inglesa e o Português Brasileiro pertencem a dois grupos distintos. Caberá ao aprendiz brasileiro, portanto, adquirir um novo padrão de vozeamento ao produzir as plosivas surdas da língua-alvo.

2.2 Aquisição fonético-fonológica de L2: uma perspectiva dinâmica

O fenômeno de aquisição dos sons de um novo sistema linguístico pode ser explicado por uma série de variáveis que agem em conjunto, de modo a incluir, no conjunto dessas variáveis, detalhes fonéticos que tendem a desempenhar um papel minoritário sob uma perspectiva tradicional. Uma perspectiva binária, caracterizada apenas pela presença ou ausência de aspiração, não consegue retratar, com tamanha precisão, as nuances dos estágios fonético-fonológicos pelos quais passam os aprendizes em sua trajetória desenvolvimental. Nesse sentido, estudos como o de França (2011) mostram que uma descrição dos dados que vá além da caracterização “aspiração *sim vs. não*” é capaz de revelar etapas de aquisição em que o aprendiz já não mais mostra o padrão da L1, ainda que, sob uma perspectiva binária, tal fato pudesse ser omitido sob o rótulo da não-ocorrência da aspiração.

É importante deixar claro, entretanto, que diferentes maneiras de descrição de dados podem revelar diferentes concepções de fonética e fonologia; e, inclusive, distintas visões sobre o processo de aquisição fonético-fonológica (cf. FRANÇA, 2011). Nesse sentido, voltar-se para o nuance fonético e considerá-lo como aspecto fundamental para a explicação das etapas desenvolvimentais dos aprendizes implica uma concepção de aquisição baseada de caráter dinâmico. Uma concepção dinâmica de aquisição de segunda língua (cf. DE BOT *et al.*, 2007) caracteriza-se por apresentar um conjunto de diversas variáveis que interagem entre si ao longo do tempo, de modo que uma grande variedade de fatores, não vistos isoladamente, mas sim integrados, se mostram necessários para dar conta da aquisição de um único e determinado aspecto linguístico. Conforme explicam Zimmer e Alves (2010), ao se caracterizar como um processo de natureza dinâmica, a aquisição do sistema sonoro se revela tanto na percepção quanto na produção dos sons do sistema-alvo.

No que diz respeito à percepção, é necessário que o aprendiz note pistas acústicas que não necessariamente se mostram salientes no seu sistema de língua-materna. Além disso, é necessário, muitas vezes, que múltiplas pistas acústicas sejam notadas em conjunto, para a distinção entre dois ou mais segmentos: as pistas acústicas agem em conjunto, de modo que, de acordo com o sistema linguístico em jogo, uma pista pode exercer um papel prioritário, ou, ainda, de caráter mais coadjuvante.

No que diz respeito à produção, também conforme explicam Zimmer & Alves (2010), a fala deve ser vista como um processo multimodal, que resulta da orquestração de diversas unidades dinâmicas de ação, a partir das quais será estabelecido, também, o nível abstrato. Sob essa concepção, portanto, nem o físico nem o abstrato apresentam caráter meramente binário, uma vez que produção e percepção são resultados de

múltiplos fatores, que agem em conjunto e que, a partir de tal ação conjunta, dão conta da representação das unidades sonoras. Assim, a cada novo sistema linguístico a ser adquirido, o aprendiz deverá notar e posteriormente produzir, a partir da cadeia sonora de fala, as pistas acústicas que se mostram pertinentes e fundamentais para o estabelecimento dos contrastes no sistema em questão (ZIMMER; ALVES, 2008). A cada sistema a ser adquirido, diferentes pistas são, portanto, ressaltadas, não somente em termos de percepção, mas, também, de produção.

É importante ressaltar, nesse sentido, que: (a) a percepção da cadeia sonora tende a sofrer um viés, em função da língua materna, de modo que o aprendiz tenda a perceber os novos sons a partir das pistas acústicas que são ressaltadas em sua L1; (b) a produção também pode sofrer influências da L1, de modo que muitas pistas acústicas função de o aprendiz reforçar aquelas pistas já trazidas de seu primeiro sistema linguístico, pistas essas que não necessariamente se mostram como determinantes para o estabelecimento das oposições na língua-alvo. Esses dois casos caracterizam a transferência fonético-fonológica da L1 para a L2. Nas seções que seguem, discutiremos, com maiores detalhes, os tipos de transferência enfrentados pelo aprendiz frente à pista acústica de *Voice Onset Time*.

2.3 Transferência dos padrões fonético-fonológicos L1-L2

Uma das dificuldades enfrentadas por aprendizes brasileiros na aquisição de aspectos fonético-fonológicos do Inglês é apontada pelo *Speech Learning Model* (FLEGE; MUNRO; MACKAY, 1995) e pelo *L2 Perceptual Learning Model* (BEST; TYLER, 2007).

Neste trabalho, baseamo-nos, sobretudo, na caracterização de Best e Tyler (2007), uma vez que, no trabalho em questão, defendemos uma concepção dinâmica de aquisição dos sons. No que diz respeito à aquisição dos padrões de VOT, ao verificarmos que Best e Tyler (*op.cit.*) fazem alusão ao gesto fônico como a unidade primitiva de análise, os aprendizes tendem a não produzir a aspiração necessária, devido ao fato de não perceberem a diferença existente entre os padrões de vozeamento do PB e do Inglês.

Uma vez que os elementos fônicos dos sistemas da L1 e da L2 interagem em um espaço fonológico comum, é possível que os aprendizes venham a assimilar os sons da língua a ser adquirida como aqueles de sua língua materna; entende-se por “assimilação”, aqui, um processamento a partir do qual o aprendiz não se mostra capaz de formar novas categorias de sons para a L2, por se basear nas pistas acústicas prioritárias de sua L1. Considerando-se o papel das pistas acústicas ressaltado na seção anterior, bem como o fenômeno investigado neste trabalho, podemos entender claramente o processo de assimilação aqui descrito. De fato, as plosivas surdas iniciais do português e do inglês mostram-se distintas no que diz respeito às pistas acústicas prioritárias de cada língua – na L2, a aspiração é um detalhe majoritário, o que não ocorre no português. Entretanto, apesar de tal diferença entre os sistemas, o aprendiz corre o risco de processar os sons da língua-alvo a partir daquelas pistas ressaltadas na sua L1 – nesse sentido, o VOT longo do inglês é “assimilado”, de modo a não ser notado pelo aprendiz brasileiro, que não percebe, tampouco produz, o diferente papel desta pista em ambos os sistemas.

2.4 Transferência Grafo-Fônico-Fonológica de L1 para L2

No que diz respeito ao processo de aquisição das plosivas surdas do inglês por brasileiros, um dos fatores mais relevantes na transferência entre os conhecimentos de

L1 e L2 é a Transferência Grafo-Fônico-Fonológica de L1 para L2 (ZIMMER; SILVEIRA; ALVES, 2009). De acordo com os autores, embora o PB e a Língua Inglesa façam uso do mesmo sistema alfabético, a relação que existe entre ortografia e escrita é distinta nas duas línguas. Enquanto ortografia e pronúncia do PB têm uma correspondência transparente, essa relação é opaca no Inglês. A tendência, portanto, é de que o aprendiz transfira o padrão do PB para sua produção oral em Língua Inglesa (ALVES, 2010).

Considerando-se a afirmação feita no parágrafo anterior, entende-se, portanto, que o sistema alfabético pode servir, também, como um meio que fortifica o uso e a percepção daquelas pistas acústicas que se mostram prioritárias na língua materna, reforçando o padrão da L1 e, dessa forma, impedindo que o aprendiz passe a ressaltar outras pistas acústicas que, apesar de na L1 não apresentarem relevância, possuem papel fundamental para as diferenças funcionais entre os sons da L2. Como verificado por Zimmer (2004), as produções não-aspiradas de /p/, /t/ e /k/, encontradas na fala em Inglês produzida por aprendizes brasileiros, são resultado da transferência da relação estabelecida entre os grafemas “p”, “t” e “k” e seus sons correspondentes em Português. Em outras palavras, quando os estudantes não são instruídos sobre a pronúncia destes sons, eles tendem a não produzir o padrão do Inglês [p^h], [t^h] e [k^h]; ao invés disso, sua produção é equivalente àquela encontrada no PB, [p], [t] e [k] (MAGRO, 2010).

Finalmente, considerando-se uma concepção dinâmica de aquisição de linguagem, é importante, portanto, salientarmos que, no fenômeno aqui analisado, os dois tipos de transferência acima caracterizados não podem ser vistos como processos plenamente dissociáveis. Na verdade, um fenômeno serve para reforçar o outro: se, por um lado, a transferência dos padrões grafados da L1 reforça a produção e a percepção diferente do falar nativo, ao mesmo tempo a não distinção perceptual das pistas acústicas que caracterizam as plosivas da L1 e da L2 estimula a manutenção dos padrões grafo-fônico-fonológicos da língua materna na produção oral em L2. Em outras palavras, ainda que seja pertinente reconhecer, em termos teóricos, uma diferenciação entre os dois tipos de transferência supracitados, associar a dificuldade de produção dos VOTs do inglês por brasileiros a apenas um dos tipos de transferência constitui tarefa impossível, visto que os dois tipos de transferência agem em conjunto, sem necessariamente uma relação de causa e efeito ordenada e predefinida entre si.

3. Metodologia

Nesta seção, apresentamos informações sobre os participantes, as palavras-alvo, os dois testes de percepção e os dois testes de produção que fizeram parte deste estudo.

3.1 Participantes

Foram voluntários no presente estudo dez estudantes de graduação em Letras – Inglês na UFRGS. Todos os participantes nasceram no Brasil e tinham o Português como sua língua materna. Três deles eram do sexo masculino, enquanto sete eram do sexo feminino.

Os participantes foram submetidos à avaliação do *Oxford Placement Test 1* (ALLAN, 2004), que os classificou em dois níveis de proficiência: intermediário e proficiente. Assim sendo, este trabalho investigou a percepção e produção de cinco aprendizes de nível intermediário e cinco de nível proficiente.

3.2 Palavras-alvo

Os instrumentos de percepção e produção desta pesquisa tiveram como alvo, em Inglês, palavras monossilábicas, iniciadas por consoantes plosivas seguidas pelas vogais altas /i/ e /ɪ/. Como exemplo dessas, temos as palavras *pee*, *tick* e *kill*. Em Português, as palavras-alvo do teste de produção eram dissilábicas, e a vogal que seguia a consoante inicial era /i/ quando a plosiva era bilabial ou velar, ou /u/, se essa consoante era alveolar⁶. Exemplos dessas são as palavras *pico*, *tudo* e *quilo*. A escolha pelas vogais altas é justificada pelo fato de que, como afirmam Yavas & Wildermuth (2006) e Yavas (2008), nesse contexto, a aspiração das plosivas se faz mais evidente.

O número total de palavras-alvo utilizadas nos instrumentos foi 18. Considerando cada ponto de articulação (bilabial, alveolar e velar), o número de palavras foi o mesmo. Pares mínimos, a exemplo de *peer* / *beer* e *tip* / *dip*, foram contrastados no teste de discriminação.

Os estímulos utilizados nos testes de percepção consistiam na produção das palavras-alvo, previamente gravadas em estúdio por duas falantes nativas de Inglês adultas – uma nascida nos Estados Unidos e outra no Canadá, que estavam no Brasil há menos de um ano. Foi realizada a verificação acústica das gravações no *software Praat* (BOERSMA; WEENINK, 2011), a qual revelou que as duas falantes partilhavam dos mesmos padrões de vozeamento. Além disso, a verificação acústica mostrou que, para plosivas sonoras, ambas as falantes produziam, variavelmente, dois padrões de VOT: VOT zero e VOT negativo, o que possibilitou o uso desses dois padrões, além do VOT positivo, nos testes de percepção.

3.3 Testes de produção

Foram realizados, neste estudo, dois testes de produção: o primeiro, em Inglês, e o segundo, em Português. Em ambos os testes, os dez participantes tiveram sua leitura oral das palavras-alvo gravadas individualmente, com o auxílio de *headsets*. Foi solicitado aos aprendizes que lessem em voz alta as palavras que eram apresentadas a eles em um *slide show*. As produções foram posteriormente analisadas no *software Praat* (BOERSMA; WEENINK, 2011) para fins de medição de VOT e subsequente comparação entre L1 e L2.

Cada teste contou com a produção de nove palavras (três por ponto de articulação) em cada língua. Em todas as palavras, a consoante inicial era seguida pela vogal /i/ ou /ɪ/. As nove palavras foram produzidas duas vezes pelos aprendizes, resultando em um total de 180 tokens por língua-alvo (360 ao total). A medição de VOT revelou que um número pequeno de dados não poderia ser aproveitado para o estudo, devido a ruídos durante a gravação. Um total de 354 tokens foi, então, aproveitado.

Tendo sido feita a medição de VOT de todos os dados, analisamos os dados em dois diferentes testes estatísticos. Primeiramente, aplicou-se o teste *Two-Way ANOVA*, buscando constatar se havia diferenças significativas entre os dois níveis de proficiência ao produzir as consoantes /p/, /t/ e /k/ em Inglês. Após essa etapa, realizamos o teste de *Repeated Measures ANOVA*, a fim de verificar se havia diferenças significativas entre as produções das consoantes em L1 e L2.

3.4 Testes de percepção

Neste estudo, os aprendizes tiveram sua percepção dos padrões de VOT da sua L2 testados através de dois instrumentos distintos: primeiramente, um teste de

⁶ Tal escolha se deve ao fato de que, como verificado, no dialeto falado pelos aprendizes, quando a consoante plosiva alveolar é seguida pela vogal /ɪ/, essa plosiva sofre palatalização, resultando na produção de uma consoante africada, cuja análise não é um objetivo deste trabalho.

identificação, de caráter piloto; posteriormente, foi aplicado um teste de discriminação (tarefa AxB). Esses testes são descritos a seguir.

3.4.1 Teste de identificação piloto

O primeiro teste de identificação tinha como objetivo determinar se os participantes eram capazes de identificar produções de consoantes plosivas com os padrões de VOT característicos da língua-alvo. Os estímulos auditivos consistiam na produção de uma única palavra, a exemplo de *pit*, apresentada isoladamente. Foram utilizados fones de ouvido e uma folha-resposta, na qual os aprendizes deveriam indicar a consoante inicial da palavra que ouviram (/p/, /t/, /k/, /b/, /d/ ou /g/).

Apenas produções com VOT positivo e VOT zero, considerados padrão do Inglês, foram utilizadas. Era considerada a hipótese de que os aprendizes teriam dificuldade em identificar as produções com VOT zero como produções de plosivas sonoras, dada a sua semelhança, em termos de VOT, com as plosivas surdas do Português Brasileiro. Assim sendo, dentre as nove palavras que foram apresentadas aos aprendizes, seis iniciavam com uma plosiva sonora, enquanto apenas três iniciavam com uma plosiva surda. Dez estudantes participaram do teste, resultando na identificação de 99 tokens, sendo 66 para segmentos sonoros e 33 para surdos.

3.4.2 Teste de discriminação (tarefa AxB)

Como discutiremos na seção referente aos resultados deste trabalho, a análise dos dados do teste de identificação piloto apontou a necessidade de novos experimentos de percepção. Foi conduzido, portanto, um teste de discriminação, que contrastou os três padrões naturais de VOT (VOT negativo, VOT zero e VOT positivo). Além desses, produções naturais de segmentos surdos com VOT positivo foram manipuladas no *software Praat* (BOERSMA; WEENINK, 2011), visando à obtenção de plosivas com *VOT zero artificial*. Neste procedimento, o VOT das plosivas surdas foi cortado, para que o segmento tivesse o mesmo VOT e soasse como uma plosiva sonora. As consoantes com o VOT zero artificial foram, então, contrastadas, na tarefa AxB, com as produções de VOT positivo e negativo. Segmentos com VOT positivo e negativo foram, também, contrastados entre si.

No teste de discriminação, solicitou-se aos aprendizes que ouvissem sequências de três palavras; sua tarefa seria julgar se a consoante inicial da segunda palavra na sequência era similar à produzida na primeira palavra (AAB), ou à produzida na terceira palavra (ABB). Era, ainda, possível que os participantes afirmassem que a mesma consoante iniciava as três palavras da sequência (AAA). Possíveis estímulos, tomando como exemplo os segmentos bilabiais, eram: (/p/ x /p/ x /b/), (/p/ x /b/ x /b/), (/b/ x /b/ x /p/) e (/b/ x /p/ x /p/) podendo /b/ ser formado por um VOT negativo, zero natural ou zero artificial. O mesmo padrão de estímulo era seguido para questões envolvendo plosivas alveolares e velares.

O teste de discriminação contou com a participação de dez aprendizes, que responderam a seis questões para cada um dos quatro contrastes de VOT aplicados. Sendo assim, o número total de questões do teste foi 24, resultando na produção de 240 tokens. O *Quadro 1* ilustra a distribuição dos tokens no teste de discriminação:

Quadro 1 – Número de tokens no teste de discriminação AxB.

	Número de questões por aprendiz	Nível Intermediário (5 aprendizes)	Nível Proficiente (5 aprendizes)	Total de tokens por condição
<i>Negativo x Zero</i>	06	30	30	60

<i>Natural</i>				
<i>Negativo x Zero Artificial</i>	06	30	30	60
<i>Zero Natural x Zero Artificial</i>	06	30	30	60
<i>Negativo x Positivo</i>	06	30	30	60
Total de tokens por grupo	24	120	120	240

Após a coleta e verificação desses dados, realizamos dois testes estatísticos. O primeiro deles, um *Teste T de Amostras Independentes*, com o objetivo de apontar se havia diferença significativa entre os dois níveis de proficiência na acuidade em discriminar os padrões de VOT contrastados. Após isso, foram feitas *Pairwise Comparisons*, aplicando-se a correção de *Bonferroni*, a fim de verificar se havia diferenças significativas na acuidade em discriminar os padrões de VOT, comparando as condições VOT negativo x VOT zero artificial e VOT negativo x VOT positivo.

4. Resultados e discussão

Nesta seção, apresentamos e discutimos os resultados obtidos no presente estudo. A ordem em que os resultados são mostrados se faz importante para o entendimento das conclusões a que chegou este trabalho: o último experimento conduzido foi o de discriminação, que foi elaborado levando-se em consideração os resultados obtidos nos experimentos de identificação e produção.

4.1 Teste de identificação piloto

A análise dos dados obtidos no teste de identificação piloto aponta que ambos os níveis de proficiência obtiveram o mesmo índice de acuidade: 97,77% (para um total de 45 tokens por grupo). Em resposta à Questão Norteadora 1, concluímos que, em ambos os níveis de proficiência, os aprendizes foram altamente capazes de identificar as plosivas do Inglês.

Tal resultado contrariou uma das hipóteses iniciais do estudo, a qual sugeria que as plosivas sonoras produzidas em Inglês com VOT zero deveriam ser identificadas pelos aprendizes como segmentos surdos, devido à semelhança entre os padrões de VOT na L1 e L2. Levando em conta os resultados desse teste e dos testes de produção, apresentados a seguir, verificamos a necessidade de elaboração de um novo experimento de percepção.

4.2 Testes de produção

Ao respondermos à Questão Norteadora 2, a verificação acústica dos dados coletados no teste de produção em Inglês revelou que, embora próximos, os índices de VOT produzidos pelos participantes ainda não igualam o padrão da língua-alvo. O teste *Two-Way ANOVA* apontou não haver diferença significativa entre as produções de /p/, /t/ e /k/ dos 2 níveis de proficiência ($p < 0,05$); isto é, nesse grupo de 10 participantes, pode-se sugerir que o nível de proficiência não exerce influências no índice de VOT que o indivíduo produz. Destaca-se o fato de que a consoante com o VOT mais próximo ao padrão da L2 é /k/; tais índices poderiam até mesmo ser comparados aos produzidos por um falante nativo de Inglês. Os resultados do teste de produção em Inglês estão ilustrados a seguir, na *Tabela 1*:

Tabela 1 – Resultado do teste de produção em inglês: índice de VOT nas produções dos aprendizes (ms). 3 palavras (x2) por consoante. *DP*: desvio padrão.

Consoante	Intermediários		Proficientes	
	<i>Tokens</i>	<i>Média (DP)</i>	<i>Tokens</i>	<i>Média (DP)</i>
/p/	30	23,53 (14,09)	29	39,76 (30,70)
/t/	29	59,76 (23,28)	30	58,27 (22,87)
/k/	30	76,10 (26,73)	30	79,77 (25,44)

Os altos índices de VOT na produção da plosiva /k/, similares à produção nativa, podem ser justificados devido ao seu ponto de articulação. Como afirmam Cho e Ladefoged (1999), plosivas posteriores caracterizam-se por um VOT com maior duração. Além disso, a duração média de VOT encontrada nessa consoante no próprio PB, como apresentaremos a seguir, pode ser considerada semelhante à da L2.

O teste de produção em português nos permite responder à questão 2.1: verificados os dados, realizamos o teste *Repeated Measures ANOVA*, o qual apontou haver diferença significativa entre as produções em L1 e L2: $F(5, 45) = 38.564$; $p = 0,000$. *Pairwise Comparisons* de *Bonferroni* revelaram que tal diferença significativa existe apenas entre as produções da consoante /t/ ($p = 0,022$); isto é, a produção de /t/ em Inglês tem VOT significativamente maior do que em PB. Os resultados estão ilustrados abaixo, na *Tabela 2*, que compara os índices de VOT produzidos pelos aprendizes em PB e na língua-alvo:

Tabela 2 – Resultado dos testes de produção: índice de VOT nas produções dos aprendizes em L1 e L2 (ms). 3 palavras (x2) por consoante. *DP*: desvio padrão.

Consoante	Português		Inglês	
	<i>Tokens</i>	<i>Média (DP)</i>	<i>Tokens</i>	<i>Média (DP)</i>
/p/	57	21,75 (10,06)	59	30,91 (24,92)
/t/	60	34,21 (11,19)	59	58,40 (22,89)
/k/	59	70,98 (20,53)	60	77,29 (25,93)

Os índices de VOT em português verificados no presente trabalho são mais altos do que aqueles tradicionalmente encontrados na literatura da área, que aponta valores médios de 12, 18 e 38 ms, para /p/, /t/ e /k/, respectivamente (REIS & NOBRE-OLIVEIRA, 2008). Outros trabalhos, no entanto, também registram índices mais elevados, tais como o de França (2011), que verificou, na produção de aprendizes do Sul do estado do Rio Grande do Sul, tempo de aspiração médio de 19,56 ms para /p/, 21,66 para /t/ e 47,20 para /k/.

Esses resultados, em adição à capacidade de identificação das plosivas verificada no teste de identificação piloto, apontaram a necessidade de elaboração de novos experimentos de percepção. Considerando que os aprendizes são capazes de fazer a identificação das plosivas do Inglês, mas não as produzem, uma das possibilidades de explicação para tal fato seria a de que outras pistas acústicas presentes nestas consoantes, que não o VOT, estariam sendo notadas pelos participantes na tarefa de identificação.

4.3 Teste de discriminação

Ao considerarmos a hipótese aventada na seção anterior, conduzimos o experimento adicional de discriminação (tarefa AxB), que fez uso de estímulos com VOT zero artificial, com o objetivo de investigar a existência de outras pistas acústicas

utilizadas pelos aprendizes para distinguir plosivas surdas e sonoras. Os resultados nos permitem responder à Questão Norteadora 3 da seguinte forma: (a) como esperado, os aprendizes são altamente capazes de discriminar plosivas com VOT negativo daquelas com VOT positivo; (b) os aprendizes não são capazes de fazer distinção entre plosivas com VOT negativo e VOT zero natural; e (c) os aprendizes são capazes de discriminar produções com VOT negativo ou VOT zero natural de produções com VOT zero artificial.

Para um maior entendimento da próxima tabela, que apresenta os resultados do teste, é necessário esclarecer: (i) *Acuidade* corresponde à porcentagem de vezes em que os aprendizes foram capazes de discriminar corretamente os padrões de VOT, acertando qual das palavras iniciava com uma consoante diferente na sequência do estímulo; e (ii) *Igualdade* corresponde à porcentagem de vezes em que os aprendizes não foram capazes de distinguir os padrões de VOT contrastados, assinalando que as três palavras na sequência iniciavam pela mesma consoante. A seguir, a *Tabela 3* apresenta os resultados do teste de discriminação:

Tabela 3 – Resultado do teste de discriminação AxB. 6 questões por contraste.

Padrão de VOT	Intermediários		Proficientes	
	<i>Acuidade</i>	<i>Igualdade</i>	<i>Acuidade</i>	<i>Igualdade</i>
<i>Negativo x Zero Natural</i>	13,33% (4/30)	86,66% (26/30)	0% (0/30)	100% (30/30)
<i>Negativo x Zero Artificial</i>	73,33% (22/30)	20% (6/30)	83,33% (25/30)	13,33% (4/30)
<i>Zero Natural x Zero Artificial</i>	66,66% (20/30)	26,66% (8/30)	50% (15/30)	23,33% (7/30)
<i>Negativo x Positivo</i>	90% (27/30)	0% (0/30)	96,66% (29/30)	0% (0/30)

Objetivando responder à Questão Norteadora 3.1, que indaga se há diferenças entre os níveis de proficiência na acuidade em discriminar os padrões de VOT contrastados, recorreremos ao Teste-T de Amostras Independentes. O teste apontou não haver diferença significativa entre os dois níveis de proficiência na acuidade em discriminar os padrões de VOT: (a) resultado para o contraste VOT negativo x zero natural: $t(08) = -2,138$; $p = 0,650$; (b) resultado para a condição VOT negativo x zero artificial: $t(08) = 0,647$; $p = 0,536$; (c) resultado ao contrastar VOT zero natural e zero artificial: $t(08) = -0,953$; $p = 0,368$; (d) resultado para o contraste VOT negativo x positivo: $t(08) = 1,265$; $p = 0,242$. Novamente, há a sugestão de que, para esse grupo de 10 aprendizes, o nível de proficiência não exerce diferenças significativas na sua capacidade de percepção dos padrões de VOT da L2.

Também buscamos responder à Questão Norteadora 3.2, a qual questionava a existência de diferenças significativas em termos de acuidade na discriminação entre os contrastes, comparando as condições (a) VOT Zero Natural e VOT Zero Artificial e (b) VOT Zero Artificial e VOT Positivo. Nesse sentido, *post-hoc Pairwise Comparisons*, aplicando-se a correção de *Bonferroni*, sugerem que: (i) ao compararmos as condições *VOT Negativo x Zero Natural* e *VOT Negativo x Zero Artificial*, há diferença significativa entre seus índices de acuidade ($p = 0,000$); (ii) ao compararmos as condições *VOT Negativo x Zero Artificial* e *VOT Negativo x Positivo*, não há diferença significativa entre seus índices de acuidade ($p = 0,568$). A partir do que foi verificado na

realização dessas *Pairwise Comparisons*, sugerimos que, para os aprendizes, o VOT zero artificial é perceptualmente diferente do VOT zero natural e, por outro lado, semelhante ao VOT Positivo.

A tarefa de discriminação AxB nos permite fazer três considerações a respeito da discriminação de padrões de VOT: perceptualmente, para esses participantes, *VOT negativo* e *VOT zero natural* são similares, *VOT zero natural* e *VOT zero artificial* são diferentes, e *VOT zero artificial* e *VOT positivo* são similares. Tais considerações nos levam a sugerir que, para esses aprendizes, o *Voice Onset Time* parece não ser a única pista acústica determinante na distinção entre plosivas surdas e sonoras. Isso porque, se o VOT fosse a única pista, supomos que os aprendizes não fariam distinção entre os padrões *zero natural* e *zero artificial*, que têm virtualmente os mesmos índices de VOT. Além disso, os padrões *zero artificial* e *positivo*, que têm índices de VOT tão distintos, não deveriam ser considerados como produções da mesma consoante. Outras pistas acústicas possivelmente consideradas pelos aprendizes nesse contraste podem ser, por exemplo, a intensidade da explosão e o tempo de clausura dessas consoantes.

5. Considerações finais

Este trabalho apresentou os primeiros resultados de um estudo de caráter piloto. Tais resultados contribuirão para a elaboração de novos instrumentos de percepção e produção, que nos permitirão apresentar dados mais consistentes a respeito da aquisição das plosivas do Inglês por brasileiros. Para atingir esse objetivo, pretendemos aumentar o número de participantes do estudo, analisando também a percepção e produção de falantes de nível iniciante. Além disso, objetivamos verificar os efeitos da instrução formal na aquisição desse aspecto fonético-fonológico da L2.

Levando em consideração o que foi verificado na realização da tarefa de discriminação AxB, pretendemos investigar, através de ferramentas como o *software Praat* (BOERSMA; WEENINK, 2011), que outras possíveis pistas acústicas, além do VOT, determinam a distinção surdo x sonoro na produção de plosivas do Inglês. Outra meta a perseguirmos seria a de aplicar testes de discriminação entre diferentes padrões de VOT a falantes nativos de Inglês, a fim de verificar se essas mesmas pistas acústicas são percebidas por esses falantes. A partir de tais estudos futuros, julgamos que conseguiremos demonstrar, mais uma vez, a complexidade e a dinamicidade do processo de aquisição fonético-fonológica de L2, processo esse em que múltiplas pistas acústicas agem em conjunto, com diferentes graus de prioridade, para a distinção dos sons em cada sistema.

Referências

- ALLAN, Dave. *Oxford Placement Test 1*. Oxford University Press, 2004.
- ALVES, Mariane Antero. *Production of English and Portuguese Voiceless stops by Brazilian EFL Speakers*. Dissertação (Mestrado em Letras) - Universidade Federal de Santa Catarina, Florianópolis, SC, 2011.
- ALVES, Ubiratã Kickhöfel. *Uma discussão conexionista sobre a explicitação de aspectos fonético-fonológicos da L2: dados de percepção e produção da plosiva labial aspirada do inglês*. In: POERSCH, José Marcelino; ROSSA, Adriana Angelim (org.). *Processamento da Linguagem e Conexionismo*. Santa Cruz do Sul: Editora da UNISC, 2007, p. 155-185.

_____. *Efeitos da Instrução Formal na Aquisição de Aspectos Fonético Fonológicos do Inglês (L2) por Brasileiros*. Projeto de pesquisa. Porto Alegre: Universidade Federal do Rio Grande do Sul (UFRGS), 2010.

BEST, C; TYLER, M. D. *Nonnative and second-language speech perception: commonalities and complementarities*. In: BOHN, O. S.; MUNRO, M. J. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2007, p. 13-34.

CELCE-MURCIA, Marianne; BRINTON, Donna M.; GOODWIN, Janet M; GRINER, Barry. *Teaching pronunciation: a coursebook and reference guide*. Cambridge University Press, 2010.

CHO, Taehong; LADEFOGED, Peter. *Variation and universals in VOT: evidence from 18 languages*. *Journal of Phonetics*, 27, p. 207-229, 1999.

COHEN, Gustavo Vargas. *The VOT Dimension: a bi-directional experiment with English Brazilian Portuguese stops*. Dissertação – Mestrado em Letras. UFSC, Florianópolis: 2004.

DE BOT, Kes; LOWIE, Wander; VERSPOOR, Marjolijn A *Dynamic Systems Theory approach to second language acquisition*. *Bilingualism: Language and Cognition*, v. 10, n. 1, p. 7-21, 2007.

FLEGE, J. E; MUNRO, M. J.; MacKAY, I. R. A. *Factors affecting strength of perceived foreign accent in a second language*. *Journal of the Acoustical Society of America*, v. 97, n.5, p. 3125-3134, 1995.

FRANÇA, Karoline. *A aquisição da aspiração das plosivas surdas do inglês por falantes do português brasileiro: implicações teóricas decorrentes de duas diferentes formas de descrição dos dados*. Dissertação (Mestrado em Letras) - Universidade Católica de Pelotas, Pelotas, RS, 2011.

LISKER L.; ABRAMSON, A. *A Cross-language study of voicing in initial stops: Acoustical measurements*. *Word*, New York, United States, v.3, n.20, p. 384-422, 1964.

MAGRO, Vivian. *Effects of different types of instruction on the acquisition of aspiration in word-initial /p/ by Brazilian Portuguese speakers*. Monografia de conclusão (Especialização em Ensino de Língua Inglesa) - Centro Universitário Ritter dos Reis, Porto Alegre, RS, 2010.

REIS, Mara; NOBRE-OLIVEIRA, Denize. *Effects of perceptual training on the identification and production of English voiceless plosives aspiration by Brazilian EFL learners*. *Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. Florianópolis: UFSC, 2008. p. 372-381.

SIMON, Ellen. *Voicing in Contrast: Acquiring a Second Language Laryngeal System*. Ghent: Academia Press, 2010.

YAVAS, Mehmet. *Factors influencing the VOT of English long lag stops and interlanguage phonology*. In: RAUBER, Andrea S.; WATKINS, Michael A.; BAPTISTA, Barbara O. (eds.). *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. Florianópolis: Universidade Federal de Santa Catarina, 2008, p. 492-498.

YAVAS, M; WILDERMUTH, Renée. *The effects of place of articulation and vowel height in the acquisition of English aspirated stops by Spanish speakers*. *IRAL*, v. 44, p. 251-263, 2006.

ZIMMER, Márcia Cristina. *A transferência do conhecimento fonético-fonológico do português brasileiro (L1) para o inglês (L2) na recodificação leitora: uma abordagem conexionista*. Tese (Doutorado em Letras) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, 2004.

_____; SILVEIRA, Rosane; ALVES, Ubiratã Kickhöfel. *Pronunciation instruction for Brazilians: bringing theory and practice together*. Newcastle upon Tyne: Cambridge Scholars Publishing, 2009.

_____; ALVES, Ubiratã Kickhöfel. On the Status of Terminal Devoicing as an Interlanguage Process among Brazilian learners of English. *Ilha do Desterro*, v. 55, p. 41-62, 2008.

_____; ALVES, Ubiratã Kickhöfel. Learning to orchestrate time: Voicing patterns and gestural drift in L2 speech production. *Livro de Resumos do São Paulo School of Advanced Studies in Speech Dynamics*. São Paulo, 2010, p. 47-48.