

## EFEITOS DE LONGO PRAZO DO TREINAMENTO PERCEPTUAL NA PERCEPÇÃO E PRODUÇÃO DAS PLOSIVAS INICIAIS SURDAS DO INGLÊS POR ESTUDANTES BRASILEIROS: IMPLICAÇÕES PARA O ENSINO DE PRONÚNCIA

Ubiratã Kickhöfel Alves<sup>†</sup>

Felipe Rodrigues Kampff<sup>†\*\*</sup>

<sup>†</sup>Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil

### Resumo

Neste artigo, verificamos os efeitos do treinamento perceptual na percepção e produção dos padrões de VOT de plosivas iniciais do Inglês por aprendizes brasileiros. O experimento consistiu em um pré-teste, um pós-teste imediato e dois pós-testes postergados. O primeiro pós-teste postergado foi realizado um mês após o treinamento, e o último, três anos mais tarde. Os resultados mostraram um aumento tímido nos índices de identificação por parte do Grupo Experimental, não tendo sido verificado um aumento nas durações de VOT entre o pré-teste e o último pós-teste. Através destes dados, chamamos a atenção para a necessidade de uma abordagem de treinamento perceptual conjugada a atividades adicionais de conscientização acerca dos aspectos fonético-fonológicos, além de integrada a uma exposição contínua às formas-alvo.

**Palavras-chave:** Treinamento Perceptual; Plosivas Iniciais; Ensino de Pronúncia

## LONG-TERM EFFECTS OF PERCEPTUAL TRAINING ON THE PERCEPTION AND PRODUCTION OF WORD-INITIAL VOICELESS STOPS BY BRAZILIAN LEARNERS: IMPLICATIONS FOR PRONUNCIATION TEACHING

### Abstract

In this article, we investigate the effects of perceptual training on the perception and production of word-initial VOT patterns in English by Brazilian learners. The experiment consisted of a pre-test, a post-test and two delayed post-tests. The first delayed post-test took place one month after training, while the second one took place three years later. The results show a slight increase in the identification rates by the Experimental Group, and no significant increase between the pre-test and the final post-test in the

<sup>†</sup> Professor do Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul (UFRGS). Bolsista de Produtividade – Nível 2 do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Seu e-mail é ukalves@gmail.com. ORCID: 0000-0001-6694-8476.

<sup>\*\*</sup> Licenciado em Letras – Português/Inglês pela Universidade Federal do Rio Grande do Sul. Durante sua graduação, atuou como Bolsista de Iniciação Científica (com bolsa CNPq) na área de Desenvolvimento Fonético-Fonológico de L2. Seu e-mail é lipekampff@hotmail.com. ORCID: 0000-0002-8796-3998.



production data. Based on these results, we draw attention to the need of associating perceptual training with additional awareness-raising tasks on phonetic-phonological aspects, combined with a continuous exposure to the target forms.

**Keywords:** Perceptual Training; Word-initial Plosives; Pronunciation Teaching

## Introdução

Na Língua Inglesa, o *Voice Onset Time* (VOT), caracterizado como o intervalo de tempo entre a explosão de uma consoante plosiva e o início da pulção glotal (anterior ou posterior à explosão) (LISKER; ABRAMSON, 1964; ABRAMSON; WHALEN, 2017), é a principal pista acústica para o estabelecimento perceptual da oposição funcional entre oclusivas surdas e sonoras iniciais. Plosivas produzidas com um VOT Positivo, o que corresponde à aspiração de tais segmentos, são identificadas como instâncias de segmentos surdos pelos falantes nativos da língua (SCHWARTZHAUPT; ALVES, FONTES, 2015). Por sua vez, oclusivas que apresentam um padrão de VOT Zero (ou seja, sem pré-vozeamento e com um intervalo de VOT sem aspiração, bastante curto), ou um padrão de VOT Negativo<sup>1</sup> (com vibração de pregas vocais antes mesmo da soltura), são identificadas pelos falantes nativos de Inglês como instâncias de segmentos sonoros. Em outras palavras, em Inglês, é a realização ou não de aspiração que leva à identificação, por falantes nativos, do segmento como surdo (/p/, /t/, /k/) ou sonoro(/b/, /d/, /g/), respectivamente.

Entretanto, o VOT não corresponde à pista prioritária para as distinções de sonoridade entre segmentos plosivos em todas as línguas. Isso porque, ainda que tal aspecto acústico exerça papel fundamental em Inglês, em outros sistemas linguísticos tal pista pode exercer um papel menos importante ou secundário para a distinção entre segmentos surdos e sonoros. De fato, as oposições fonológicas entre segmentos podem variar em função das diferentes pistas acústicas tomadas como prioritárias por cada sistema linguístico, de modo que as diferenças fonológicas se estabeleçam a partir dos pesos maiores ou menores atribuídos às pistas acústicas (cf. a proposta de ‘pesos de pistas’, ou *cueweighting*, de Holt; Lotto, 2006), considerando-se um modelo psicoacústico de percepção (como o *Speech Learning Model* de Flege, 1995). No que diz respeito aos segmentos plosivos, uma série de estudos recentes (AHN, 2018; AL-TAMIMI; KHATTAB, 2018; KLEBER, 2018; LADD; SCHMID, 2018 dentre outros) tem evidenciado que diversas outras pistas podem exercer papel fundamental para o estabelecimento de distinções funcionais em diferentes sistemas. Nessas diferentes línguas, considerando-se a aprendizagem de Inglês como L2, são necessárias técnicas instrucionais para levar o aprendiz a tomar a pista de VOT como fundamental na diferenciação entre segmentos surdos e sonoros iniciais.

O Português e o Espanhol podem ser enquadrados no caso acima. Nestas duas línguas, ainda que não saibamos exatamente quais são as pistas acústicas prioritárias para a distinção entre segmentos surdos e sonoros (uma vez que a distinção pode estar sendo estabelecida, inclusive, pela ação conjunta das pistas, em um cenário de acordo com o qual a vibração das pregas vocais durante a oclusão possa ser uma pista importante a ser empregada de forma conjunta com tais pistas adicionais), diversos estudos realizados por nosso grupo de pesquisa (Português: ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; Espanhol: ALVES; LUCHINI, 2016, 2017) mostram que,

diferentemente de falantes norte-americanos, a presença-ausência de aspiração não é a característica principal para que os falantes nativos dessas duas línguas venham a identificar as plosivas iniciais como surdas e sonoras. Tal conclusão foi atingida ao contarmos, nos estudos supracitados, com um padrão de VOT manipulado, que denominamos de VOT Zero Artificial. Tal padrão manipulado era obtido ao retirarmos, através do software *Praat*, o período de VOT longo (aspiração) das plosivas iniciais surdas que serviam como estímulo em nossos experimentos perceptuais, de modo a obtermos um padrão de VOT Zero. Com esta manipulação, o segmento resultante apresentaria o padrão de VOT semelhante ao encontrado em plosivas sonoras na língua, ao mesmo tempo em que apresentaria todas as demais pistas acústicas remanescentes de um segmento originalmente surdo, de modo a torná-lo um segmento acusticamente híbrido.

A identificação de tal segmento como surdo ou sonoro, bem como a comparação das respostas a tal padrão com as referentes ao VOT Zero não-manipulado (que apresentava, portanto, todas as pistas acústicas, não somente a de VOT, como caracterizadoras de uma oclusiva sonora do Inglês), foi-nos reveladora do *status* do VOT como pista majoritária ou não entre ouvintes norte-americanos, brasileiros e argentinos.<sup>2</sup> Em Schwartzhaupt; Alves; Fontes (2015), evidenciamos que tanto o padrão VOT Zero Natural quanto o padrão VOT Zero Artificial eram identificados por falantes nativos monolíngues de Inglês como sonoro, o que indica que, independentemente das demais pistas acústicas, a realização ou não da aspiração é determinante para a atribuição das categorias surdo/sonoro, de modo a caracterizar o *status* fundamental da pista de VOT entre os participantes norte-americanos. Por outro lado, entre aprendizes brasileiros e argentinos de Inglês, mesmo com alto nível de proficiência na L2,<sup>3</sup> o resultado não foi o mesmo: em função de seu grau de proficiência, tais participantes já identificavam como sonoras as plosivas com VOT Zero Natural (assim como faziam os participantes norte-americanos); entretanto, quando expostos ao padrão híbrido, tanto os participantes brasileiros quanto os argentinos tendiam a identificá-lo como surdo. Em outras palavras, tais participantes priorizavam outras pistas acústicas, que não o VOT, para a identificação das oclusivas como surdas ou sonoras.

Ao não exercer papel fundamental na percepção, o VOT tampouco assumirá um *status* fundamental na produção de tais aprendizes. Trata-se de um fato importante, pois a não produção da aspiração em palavras tais como *pit*, *tip* e *kill* pode afetar a inteligibilidade da fala de tais aprendizes quando ouvida por ouvintes cuja L1 é o inglês, conforme apontam os dados de Schwartzhaupt (2015), Schwartzhaupt; Alves; Fontes (2015) e Motta-Ávila (2017).<sup>4</sup> Frente a esse quadro, sentimos a necessidade de que os estudantes brasileiros e argentinos sejam instruídos, implícita ou explicitamente, a prestarem atenção ao *status* fundamental do *Voice Onset Time*. Para isso, abordagens como a instrução explícita dos aspectos fonético-fonológicos em sala de aula,<sup>5</sup> ou estratégias laboratoriais de treinamento perceptual,<sup>6</sup> podem exercer caráter fundamental.

Com base nessas premissas, em Alves; Luchini (2017) realizamos um estudo de treinamento perceptual em que aprendizes argentinos de Inglês (com ní-

veis C1 e C2 de proficiência) eram treinados a identificar o padrão VOT Positivo como instância indicadora de consoantes surdas, e o VOT Zero Artificial, como indicador de consoantes sonoras.<sup>7</sup> Em outras palavras, os aprendizes em questão eram levados, através de tal treinamento, a prestar atenção à pista acústica do VOT como prioritária (uma vez que as demais pistas eram todas iguais, dado o fato de que, conforme já expresso, o padrão VOT Zero Artificial era manipulado a partir de oclusivas que exibiam o padrão VOT Positivo). Ao buscar verificar os efeitos de tal treinamento tanto em termos de percepção (com a aplicação de uma tarefa de identificação) e de produção (através de um instrumento de leitura controlada de palavras-alvo), o instrumento contou com (i) um pré-teste; (ii) um período de uma semana com três sessões de treinamento, de aproximadamente 30 minutos cada; (iii) um pós-teste imediato, realizado maximamente três dias após o término das sessões de treinamento perceptual; (iv) um pós-teste postergado, realizado exatamente um mês após o término do treinamento.

Como resultados, Alves; Luchini (2017) demonstraram que, diferentemente do verificado no grupo controle, o grupo que recebeu treinamento,<sup>8</sup> no que diz respeito à identificação dos padrões VOT Zero Natural e Zero Artificial como sonoros, apresentou uma melhora significativa entre o pré-teste e o pós-teste imediato, sendo que tais resultados se mostraram duradouros no pós-teste postergado. Entretanto, no que diz respeito aos índices de produção de VOT, não foram encontradas diferenças significativas entre o pré-teste e o pós-teste imediato, ou entre o pré-teste e o pós-teste postergado. Como uma possível explicação, Alves; Luchini (2017) levantam a possibilidade de que o período de apenas um mês para a realização do pós-teste postergado possa ter sido demasiadamente imediato. Os autores, dessa forma, consideram a possibilidade de que, ao longo de um intervalo de tempo maior, a contínua exposição à língua em uso poderia contribuir para o desenvolvimento de maiores índices de produção. Em outras palavras, o trabalho de treinamento perceptual contribuiria para que os aprendizes passassem a prestar atenção (cf. LEOW, 2015; LEOW; ADRADA-RAFAEL, 2018) ao VOT no *input* e a tornassem uma pista de *status* prioritário (HOLT; LOTTO 2006). Com isso, paulatinamente, os aprendizes viriam a produzir maiores índices de VOT ao longo do tempo.

O trabalho aqui proposto parte justamente da possibilidade levantada no artigo supracitado. No presente estudo, replicaremos o estudo de Alves; Luchini (2017) com aprendizes brasileiros de Inglês (pois, conforme apontado pela literatura da área, tais aprendizes apresentam as mesmas dificuldades dos alunos hispânicos na identificação e produção das distinções de vozeamento das plosivas iniciais). O presente trabalho, portanto, tem o objetivo geral de verificar efeitos de longo prazo do treinamento perceptual na percepção e produção de padrões de VOT iniciais do Inglês por aprendizes brasileiros, acadêmicos de um curso de Letras com formação em Língua Inglesa<sup>9</sup> de uma Universidade Federal do sul do país.

Para atingirmos tal objetivo, o estudo aqui proposto conta com duas etapas distintas. Na primeira etapa, realizada ao final do primeiro semestre de 2015

com estudantes de primeiro ano do curso de Letras, reaplicamos o experimento de Alves; Luchini (2017), de modo a contarmos com tarefas de percepção e produção em um pré-teste, um pós-teste imediato ao treinamento e um pós-teste postergado, um mês após o término da referida intervenção pedagógica. Como segundo momento deste estudo, acrescentamos um novo pós-teste, realizado com os mesmos participantes ao final do primeiro semestre do ano de 2018, quando tais aprendizes cursavam seu último ano de formação em Língua Inglesa no curso de Letras. A partir desta nova verificação, torna-se a verificar se, no grupo experimental, os índices maiores de percepção se mostram igualmente altos ao final do curso, bem como se houve uma melhora nos índices de produção, possibilitada, possivelmente, pelo treinamento realizado no início do curso, aliado à experiência com a Língua Inglesa que o aluno teve ao longo de toda a sua formação. Ao também verificarmos os dados do grupo controle, podemos apontar se possíveis mudanças podem ser realmente atribuídas ao treinamento perceptual, ou se, mesmo sem tal intervenção, os participantes desse grupo vêm a atingir padrões semelhantes aos encontrados no grupo experimental ao final do curso de Letras.

Esperamos, com o presente trabalho, possibilitar uma discussão acerca de possíveis efeitos de longo prazo do treinamento perceptual, bem como discutir as implicações de tais resultados para o ensino de pronúncia da Língua Inglesa.

## **Metodologia**

### ***Participantes***

Na primeira etapa do estudo, alunos de três turmas do primeiro semestre de graduação em Letras de uma Universidade do Sul do Brasil foram convidados a participar do experimento, tendo sido divididos em dois grupos: Experimental (GE) e Controle (GC).<sup>10</sup> Para que suas participações no estudo fossem consideradas, os alunos precisavam estar presentes em todas as etapas instrucionais e de coleta. Com isso, ao final da terceira etapa (Pós-teste Postergado 1), apenas 10 participantes permaneceram em cada um dos grupos, totalizando 20 alunos. Em 2018, na realização do Pós-teste Postergado 2, cinco deles haviam desistido do curso ou trocado de ênfase em língua estrangeira, o que os impediu de continuarem participando do experimento. Na última etapa do experimento, então, o GE contou com sete participantes, enquanto o GC contou com oito, totalizando 15 participantes ao longo das quatro etapas de coleta.<sup>11</sup>

### ***Sessões de Treinamento Perceptual***

Considerando-se o VOT como a pista acústica prioritária para a distinção entre os segmentos surdos e sonoros iniciais por parte de ouvintes nativos do inglês (SCHWARTZHAUPT, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; MOTTA-ÁVILA, 2017), conforme já afirmado na Introdução, o trabalho

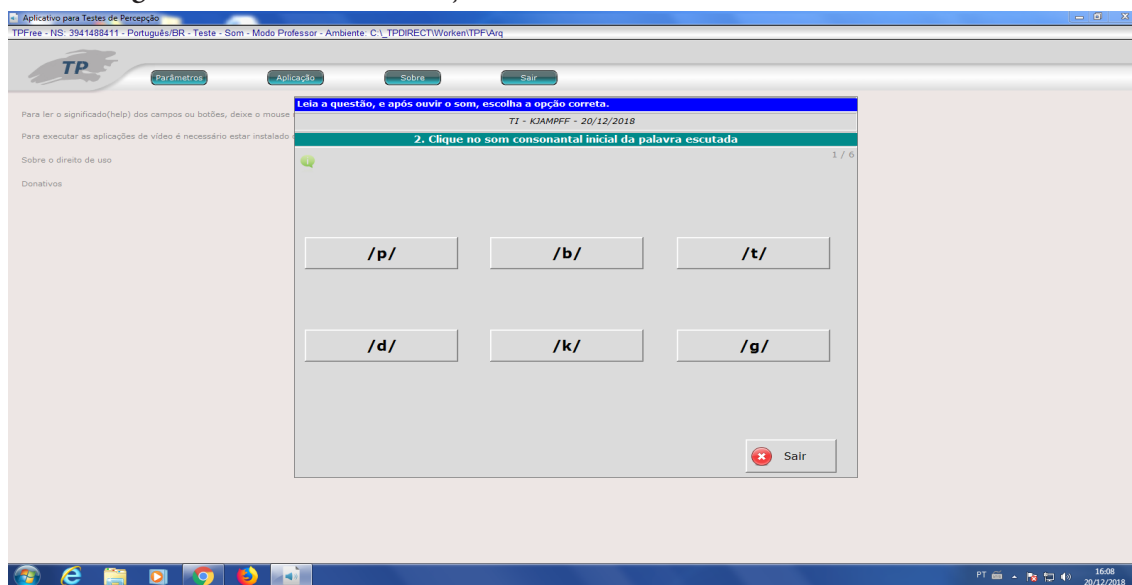
de treinamento perceptual foi idealizado com a expectativa de realocar atenção (cf. LEOW, 2015; LEOW; ADRADA-RAFAEL, 2018), por parte do aprendiz brasileiro, a tal pista no *input*. Dessa forma, tínhamos a intenção de torná-la uma pista de *status* prioritário (LORI; HOLT, 2006) também para estes aprendizes. Tal trabalho contribuiria para que, com a exposição contínua ao insumo mesmo após o treinamento, o estabelecimento das categorias fonológicas de ‘surdo’ e ‘sonoro’ levasse em consideração tal pista acústica como aspecto primordial.

Seguindo-se Alves; Luchini (2017), o treinamento consistiu em um teste de identificação com *feedback* imediato, através do *software* TP (RAUBER *et al.*, 2013), sendo realizado em três sessões ao longo de uma semana após o Pré-teste. Esse treinamento contava apenas com os padrões VOT Positivo (esperando-se, como resultado do treinamento, que os aprendizes identificassem tal padrão como surdo) e VOT Zero Artificial (que deveria ser identificado como sonoro, indicando que os aprendizes estariam tomando o VOT como pista acústica fundamental para a distinção de sonoridade, uma vez que estariam priorizando o VOT sobre as demais pistas acústicas desse segmento híbrido). Dessa forma, a função do treinamento era levar o aprendiz a tomar o VOT como pista acústica principal na distinção entre oclusivas surdas e sonoras.

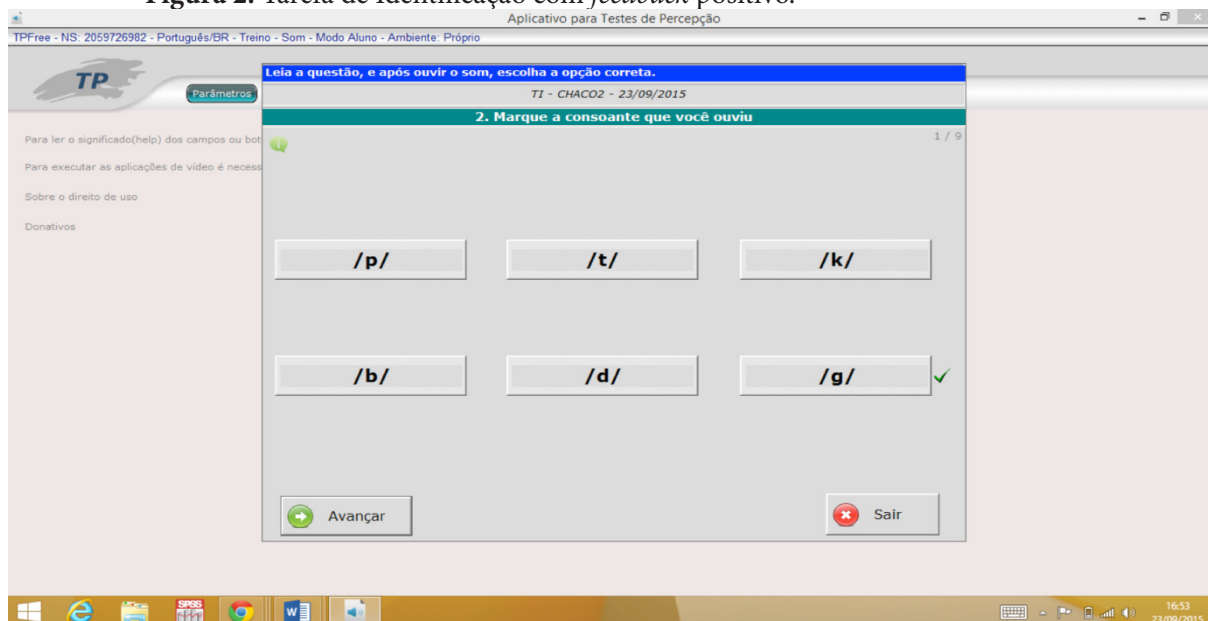
Os estímulos haviam sido gravados em um estúdio profissional particular da cidade de Porto Alegre/RS por seis falantes nativos da variedade norteamericana de Inglês, residentes no Brasil há menos de três meses antes da coleta dos dados (três homens e três mulheres). Seguindo-se Alves; Luchini (2017), os itens lexicais usados como estímulos no treinamento foram ‘*pee*’, ‘*tip*’ e ‘*kit*’. Tal escolha lexical se justifica porque, de acordo com Yavas; Wildermuth (2006) e Schwartzhaupt (2012), as plosivas surdas apresentam um intervalo de VOT maior quando seguidas de uma vogal alta, o que facilitaria a percepção.<sup>12</sup> Cada locutor gravou diversas vezes os itens lexicais supra citados. Ao final das gravações, foram selecionados os que apresentassem os padrões mais próximos da língua nativa, resultando em 18 arquivos de áudio, sendo três itens lexicais produzidos por cada um dos seis falantes; garantimos, assim, que cada participante do estudo tivesse o mesmo número de exposições aos estímulos de cada um dos falantes nativos. Metade desses itens tiveram sua aspiração cortada para que obtivéssemos o padrão VOT Zero Artificial. Na elaboração do instrumento de treinamento, todos os 18 itens foram repetidos 20 vezes em ordem aleatória, o que resulta em 360 *tokens* por sessão, que durava de 20 a 30 minutos para ser concluída. Pausas opcionais eram fornecidas aos participantes a cada 90 *tokens*.

As sessões de treinamento, que consistiam em uma tarefa de Identificação, foram realizadas no laboratório de informática da universidade, com todos os participantes de um mesmo grupo testados ao mesmo tempo, um em cada computador, e usando fones de ouvido. Os aprendizes deveriam clicar na opção referente à consoante inicial da palavra ouvida, conforme mostrado na Figura 1, a seguir.

Figura 1: Tarefa de Identificação no TP.



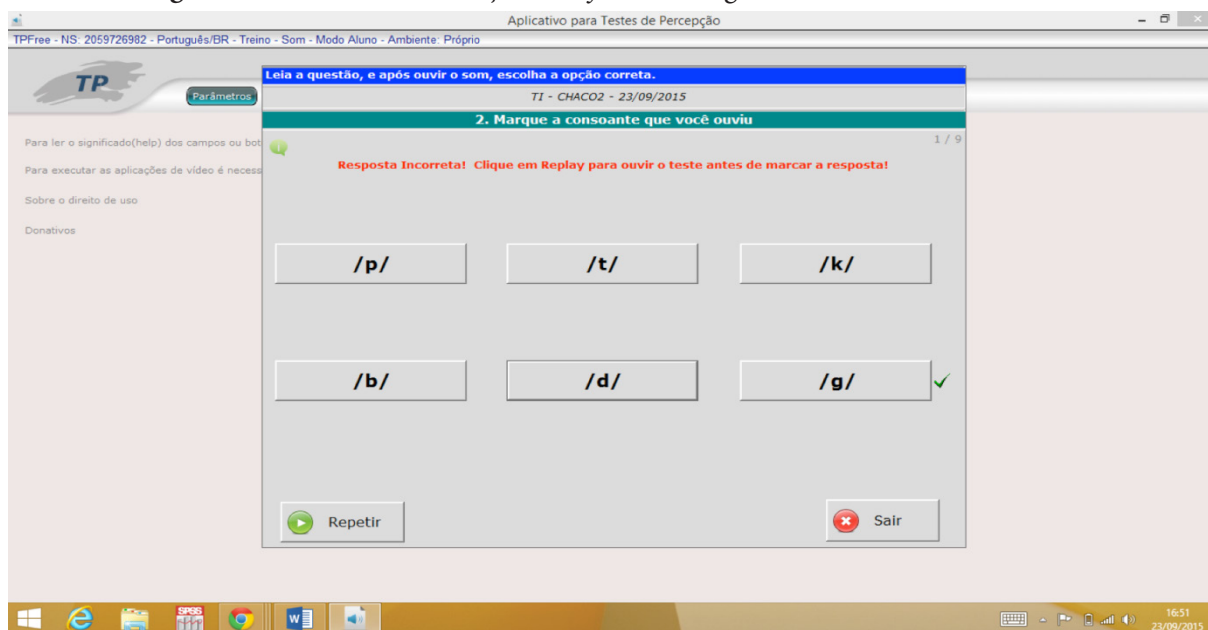
O *feedback* era fornecido imediatamente após cada uma das respostas escolhidas pelos aprendizes. Conforme já explicado, estímulos com o VOT Zero Artificial eram considerados corretos se os aprendizes selecionassem a plosiva inicial do estímulo como sonora, além do acerto em relação ao ponto de articulação, como mostra a Figura 2, a seguir. Ademais, os estímulos com VOT Positivo eram considerados corretos se, além do acerto referente ao ponto de articulação, tais segmentos fossem identificados como surdos. Dessa forma, esperávamos que os aprendizes fossem treinados a não prestarem atenção a outras pistas acústicas senão o VOT, já que a presença/ausência de aspiração era decisiva para as suas respostas.

Figura 2: Tarefa de Identificação com *feedback* positivo.



Quando as respostas desviavam do esperado, tanto no que diz respeito à escolha da sonoridade esperada do segmento ouvido ou no que concerne ao ponto de articulação da consoante, os aprendizes eram imediatamente informados sobre a resposta correta, tendo de ouvir o estímulo outra vez antes que pudessem ir adiante na tarefa, como mostra a Figura 3.

**Figura 3:** Tarefa de Identificação com *feedback* negativo.



### *Instrumentos de coleta de dados (Pré e pós-testes).*

Nesta seção, serão descritas as tarefas de percepção (identificação) e produção utilizadas no pré-teste e nos três pós-testes.

### *Tarefa de Identificação*

A Tarefa de Identificação, também realizada através do TP (RAUBER *et al.*, 2013), segue o mesmo padrão empregado em Alves; Motta (2014), Alves; Zimmer (2015), Schwartzhaupt; Alves; Fontes (2015), Alves; Luchini (2016, 2017).

Na Tarefa de Identificação aplicada no pré-teste e nos três pós-testes, os aprendizes ouviam estímulos individuais e escolhiam a consoante inicial da palavra ouvida, sem qualquer tipo de *feedback* após a escolha da resposta. Nesta tarefa, os seguintes estímulos, apresentados em quatro padrões de VOT (Negativo, Zero, Positivo e Artificial), foram incluídos e apresentados em ordem aleatória: 'pit'/bit', 'tick'/dick', 'kill'/gill'.<sup>13</sup> Além disso, a tarefa contava com 48 estímulos para serem identificados, com cada um dos quatro padrões de VOT sendo apresentados em 12 *tokens* (quatro para cada ponto de articulação, com a mesma palavra produzida por um falante diferente)<sup>14</sup>.

## Tarefa de Produção

A tarefa de produção também segue o padrão de estudos anteriores, como o de Alves; Zimmer (2015) e Alves; Luchini (2017). Em uma sala silenciosa, individualmente, os aprendizes liam um arquivo de .ppt com palavras em ordem aleatória, dentre as quais se encontravam as palavras-alvo ‘peer’, ‘pit’, ‘pee’, ‘team’, ‘tick’, ‘tip’, ‘kit’, ‘keel’ e ‘kill’,<sup>15</sup> com cada uma delas sendo produzida duas vezes, o que resulta em seis *tokens* por consoante por aprendiz. Nas três primeiras etapas, os participantes foram gravados em um *Macbook White late 2009*, com o microfone embutido do próprio computador, através do *software Audacity*, versão 2.010. Na quarta e última etapa, foi usado um *Macbook Pro 2017*, também com microfone embutido, através do mesmo *software*, na versão 2.3.0.

### *Análises acústica e estatística de dados*

Após as coletas, os dados de produção foram analisados acusticamente no *software Praat* (BOERSMA; WEENIK, 2015).<sup>16</sup> As estatísticas descritivas e inferenciais foram rodadas no programa SPSS - versão 21 para Windows (SPSS Inc.; Chicago, IL, USA).

## Descrição dos resultados

### *Tarefa de Identificação*

Na Tabela a seguir, apresentamos os índices de acertos<sup>17</sup> nas Tarefas de Identificação (Pré-teste, Pós-Teste, Pós-teste Postergado 1 e Pós-teste Postergado 2) e dos resultados dos testes de Friedman<sup>18</sup> para cada um dos dois grupos.

**Tabela 1:** Índices de acerto de cada um dos padrões de VOT (percentagem de acuidade na primeira linha, média e desvio-padrão na segunda) e resultados dos testes de Friedman referentes à Tarefa de Identificação.

<i>Grupo 1 - Experimental</i>					
	Pré-teste	Pós-teste	Pós-teste Postergado 1	Pós-teste Postergado 2	X2 (df)
<i>VOT Negativo</i>	<b>97,62%</b> 11,71 (0,79) 82/84	<b>100%</b> 12,00 (0,00) 84/84	<b>100%</b> 12,00 (0,00) 84/84	<b>100%</b> 12 (0,00) 84/84	9,00 (3) (* )
<i>VOT Positivo</i>	<b>96,43%</b> 11,57 (1,13) 81/84	<b>94,05%</b> 11,29 (1,15) 79/84	<b>96,43%</b> 11,57 (0,76) 81/84	<b>100%</b> 12,00 (0,00) 84/84	7,65 (3) (*?)
<i>VOT Zero Natural</i>	<b>70,24%</b> 8,42 (1,25) 59/84	<b>100%</b> 12 (0,00) 84/84	<b>96,43%</b> 11,57 (0,49) 81/84	<b>97,62%</b> 11,71 (0,76) 82/84	18,42 (3) (***)

<i>VOT Zero Artificial</i>	<b>35,71%</b> 4,29 (3,29) 30/84	<b>59,52%</b> 7,14 (2,24) 50/84	<b>59,52%</b> 7,14 (2,24) 50/84	<b>51,19%</b> 6,14 (2,12) 43/84	13,74 (3) (**)
<b>Grupo 2 - Controle</b>					
	Pré-teste	Pós-teste	Pós-teste Postergado 1	Pós-teste Postergado 2	X2 (df)
<i>VOT Negativo</i>	<b>100%</b> 12,00 (0,00) 96/96	<b>100%</b> 12,00 (0,00) 96/96	<b>100%</b> 12,00 (0,00) 96/96	<b>100%</b> 12,00 (0,00) 96/96	0,000 (3)
<i>VOT Positivo</i>	<b>100%</b> 12,00 (0,00) 96/96	<b>100%</b> 12,00 (0,00) 96/96	<b>98,96%</b> 11,88 (0,35) 95/96	<b>97,92%</b> 11,75 (0,71) 94/96	2,00 (3)
<i>VOT Zero Natural</i>	<b>79,17%</b> 9,50 (2,00) 76/96	<b>81,25%</b> 9,75 (1,91) 78/96	<b>77,08%</b> 9,25 (2,12) 74/96	<b>83,33%</b> 10,00 (2,00) 80/96	1,88 (3)
<i>VOT Zero Artificial</i>	<b>21,88%</b> 2,63 (1,19) 21/96	<b>20,83%</b> 2,50 (2,88) 20/96	<b>25%</b> 3,00 (2,00) 24/96	<b>28,13%</b> 3,38 (1,69) 27/96	2,00 (3)

**Nota:** \*? p < 0,10 (marginalmente significativo); \*p < 0,05; \*\* p < 0,01; \*\*\*p < 0,001.

Vemos que três dos padrões apresentam diferenças inferenciais significativas no Teste de Friedman, sendo eles o VOT Negativo, o Zero Natural e o Zero Artificial, com o padrão VOT Positivo apresentando uma diferença marginalmente significativa. Por sua vez, o GC não apresenta diferença significativa em nenhum dos padrões. A partir disso, poderíamos à primeira vista afirmar que o treinamento perceptual possibilita resultados que podem ser considerados benéficos.

Ao olharmos os índices descritivos da Tabela, verificamos que os participantes do GE aumentaram o número de acertos referentes ao padrão VOT Negativo. Tal diferença significativa não era esperada, uma vez que os índices do Pré-teste já eram altos. Merecem atenção, também, os resultados referentes aos índices de acuidade do GE que aumentaram entre as etapas, inclusive entre as duas últimas, como os casos dos padrões de VOT Positivo e Zero Natural.

Cabe mencionar que, de acordo com os resultados encontrados em estudos anteriores (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015), aprendizes brasileiros tendem a identificar o padrão Zero Artificial como surdo. Na Tabela 1, através da descrição dos índices percentuais, vemos que 35,71% dos dados dos participantes já eram identificados como sonoros pelo GE. Ao aumentar o índice de acuidade para 59,52% no Pós-teste e mantê-lo no Pós-teste Postergado 1, notamos que, nesta etapa do experimento, o treinamento contribuiu positivamente para a identificação de tal padrão como sonoro, o que sugere a tomada do VOT como pista acústica prioritária. Cabe mencionar, entretanto, que o índice máximo de acertos, de 59,52%, denota grande variabilidade nas escolhas dos participantes, apesar do aumento significativo verificado.

É importante novamente destacar que o GC não apresentou melhora significativa em nenhum dos padrões. Conforme já verificado em estudos anteriores (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015), os padrões VOT Negativo e VOT Positivo não se mostram perceptualmente difíceis para os aprendizes, o que explica os altos índices de acuidade na identificação de tais padrões desde o Pré-teste. Apesar disso, cabe mencionar que há oscilações, em termos de estatística descritiva, nos índices de acuidade deste grupo, como podemos ver nos padrões Zero Natural e Artificial.

Na próxima tabela, trazemos os resultados dos testes *post-hoc* de Wilcoxon (com correção de Bonferroni) apenas do GE (por ser o único a apresentar diferenças significativas no teste de Friedman).

**Tabela 2:** Resultados do teste *post-hoc* de Wilcoxon (com correção de Bonferroni) para a Tarefa de Identificação.

<i>Grupo 1 - Experimental</i>				
	VOT Negativo	VOT Positivo	VOT Zero Natural	VOT Zero Artificial
Pré-teste vs. Pós-teste	n.s.	n.s.	p = 0,17 *?	n.s.
Pré-teste vs. Pós-teste Postergado 1	n.s.	n.s.	p = 0,17 *?	n.s.
Pré-teste vs. Pós-teste Postergado 2	n.s.	n.s.	p = 0,17 *?	p = 0,18 *?
Pós-teste vs. Pós-teste Postergado 1	n.s.	n.s.	n.s.	n.s.
Pós-teste vs. Pós-teste Postergado 2	n.s.	n.s.	n.s.	n.s.
Pós-teste Postergado 1 vs. Pós-teste Postergado 2	n.s.	n.s.	n.s.	n.s.

**Nota:** n.s.: não significativo, \*? = marginalmente significativo ( $p < 0,02$ ), \*p  $< 0,017$ .

Em relação ao padrão VOT Zero Natural, as etapas mostraram uma diferença marginalmente significativa entre Pré-teste e Pós-teste, Pré-teste e Pós-teste Postergado 1, e Pré-teste e Pós-teste Postergado 2 ( $p = 0,017$ , considerada marginalmente significativa em função da Correção de Bonferroni realizada manualmente). Ao analisarmos tais resultados em termos descritivos, o alto índice de acurácia dos participantes acontece a partir da primeira etapa, mantendo-se ao longo de todo o experimento, e, inclusive, no Pós-teste Postergado 2, realizado cerca de três anos após o Pós-teste Postergado 1.

Conforme os dados da Tabela 1, ainda que o teste de Friedman tenha mostrado uma diferença significativa para o VOT Negativo e uma diferença marginalmente significativa para o Positivo, os testes *post-hocs* não apresentaram diferenças significativas nestes dois padrões. É provável que isso se dê porque os valores descritivos referentes a esses dois padrões mostram-se bastante altos desde o Pré-teste, não evidenciando, portanto, uma diferença estatisticamente significativa entre etapas de coleta em específico. É possível, também, que o baixo número de participantes seja responsável por tal resultado.

Em relação ao VOT Zero Artificial, ao contrário do que poderíamos esperar a partir da verificação dos índices apresentados na Tabela 1, que mostram um acréscimo entre Pré-teste e Pós-teste, os resultados inferenciais não apresentam diferença significativa em nenhuma etapa, trazendo apenas uma diferença marginal ( $p= 1,8$ ) entre as etapas de Pré-teste e Pós-teste Postergado 2.

Passemos às análises inter-grupos, verificadas a cada um dos momentos temporais tomados individualmente para análise. A seguir, apresentamos os resultados do teste de Mann-Whitney.

**Tabela 3:** Resultados do teste Mann-Whitney para a Tarefa de Identificação.

	Pré-teste U	Pós-teste U	Pós-teste Postergado 1 U	Pós-teste Postergado 2 U
VOT Negativo	16,000	28,000	28,000	56,000
VOT Positivo	24,000	12,000 (*?) $p = 0,072$	15,000	24,500
VOT Zero Natural	18,000	<b>7,000(*)</b>	<b>10,000 (*)</b>	<b>10,500 (*)</b>
VOT Zero Artificial	27,000	<b>4,000 (*)</b>	<b>1,000 (*)</b>	<b>7,500 (*)</b>

**Nota:** \*? = marginalmente significativo ( $p < 0,10$ ), \*  $p < 0,05$ .

Podemos ver que diferenças significativas entre os grupos, nos índices de acuidade referentes à identificação, se deram no Pós-teste, no Pós-teste Postergado 1 e no Pós-teste Postergado 2 para os padrões VOT Zero Natural e Zero Artificial. Isso indica que ambos os grupos se apresentaram estatisticamente equivalentes no Pré-teste, mas, a partir das seguintes etapas, já com o treinamento perceptual, a diferença entre eles aumentou. É válido mencionar, também, que diferenças entre os dois grupos não foram encontradas nos padrões VOT Negativo e VOT Positivo em nenhuma das etapas, dado que a identificação destes dois padrões como representantes de consoantes sonora e surda, respectivamente, não se mostra difícil para nenhum dos dois grupos desde o Pré-teste, conforme já discutido.

Em suma, ao considerarmos os resultados dos testes inferenciais referentes à tarefa de percepção, vemos que o treinamento perceptual apresenta um impacto positivo no que se refere à identificação, considerando que todos os padrões, para o

GE, se mostraram pelo menos marginalmente significativos no teste de Friedman (intra-grupos), além do fato de, nos testes inter-grupos de Mann-Whitney, os testes de VOT Zero Natural e Artificial apresentarem diferenças significativas após o treinamento. Cabe mencionar que, apesar de tais diferenças significativas, o índice de identificação do padrão VOT Zero Artificial como sonoro nunca chegou a ser superior a 60%, o que evidencia uma instabilidade na tomada do VOT como pista acústica prioritária. Ainda no que diz respeito à comparação intra-grupos, cabe a ressalva de que, nas comparações *post-hoc*, os resultados não se mostraram muito claros (de modo a apresentarem, unicamente, diferenças marginalmente significativas), o que pode ser decorrente do baixo número de participantes.

### Tarefa de Produção

Na Tabela 4, apresentamos os valores de VOT e os resultados do teste de Friedman para cada ponto de articulação.

**Tabela 4:** Resultados do Teste de Produção (média (em ms) na primeira linha de cada plosiva, desvio padrão na segunda e mediana na terceira) e resultados do Teste de Friedman.

<b>Grupo 1 - Experimental</b>					
	Pré-teste	Pós-teste	Pós-teste Postergado 1	Pós-teste Postergado 2	X2 (df)
/p/	<b>25,39</b> (6,05) Mdn: 29,17	<b>24,96</b> (9,69) Mdn: 23,27	<b>38,48</b> (15,77) Mdn: 24,17	<b>18,09</b> (6,07) Mdn: 18,50	9,0 (3) (* )
/t/	<b>61,09</b> (22,22) Mdn: 54,64	<b>64,21</b> (25,07) Mdn: 60,93	<b>68,99</b> (28,45) Mdn: 60,76	<b>51,23</b> (22,14) Mdn: 51,76	3,85 (3)
/k/	<b>79,7</b> (21,00) Mdn: 66,46	<b>81,31</b> (21,00) Mdn: 74,92	<b>84,24</b> (19,70) Mdn: 84,94	<b>77,85</b> (14,60) Mdn: 76,02	1,80 (3)
<b>Grupo 2 - Controle</b>					
	Pré-teste	Pós-teste	Pós-teste Postergado 1	Pós-teste Postergado 2	X2 (df)
/p/	<b>17,56</b> (7,12) Mdn: 15,42	<b>18,85</b> (7,19) Mdn: 18,32	<b>23,92</b> (13,62) Mdn: 23,06	<b>19,93</b> (7,63) Mdn: 16,47	4,82 (3)
/t/	<b>67,71</b> (17,78) Mdn: 65,95	<b>66,04</b> (12,68) Mdn: 63,42	<b>71,79</b> (18,95) Mdn: 72,08	<b>72,37</b> (37,36) Mdn: 51,23	1,50 (3)
/k/	<b>77,10</b> (11,06) Mdn: 77,49	<b>84,10</b> (17,89) Mdn: 85,73	<b>89,76</b> (24,30) Mdn: 91,72	<b>88,74</b> (21,03) Mdn: 89,50	6,30 (3) (*?)

**Nota:** Mdn = mediana; DPs = apresentados entre parênteses; \*? = marginalmente significativo ( $p < 0,10$ ), \*  $p < 0,05$ .

Semelhantemente ao verificado no estudo com treinamento perceptual de Alves; Luchini (2017), com participantes argentinos, parece não haver uma generalização clara do treinamento perceptual à produção dos segmentos plosivos iniciais. São encontrados resultados significativos para /p/ no GE, e marginalmente significativos para /k/ no GC.

No GE, apenas /p/ apresenta uma diferença significativa ( $p=0,029$ ) nas médias de produção entre os valores do Pré-teste e os do Pós-teste Postergado 1. Por sua vez, também é importante frisar que os valores referentes à quarta etapa são descritivamente mais baixos. Analisaremos este fato mais a fundo após apresentarmos os resultados dos testes *post-hoc* de Wilcoxon, a seguir.

Ainda sobre o GE, cabe-nos destacar as médias trazidas em /t/ e /k/. Apesar de os resultados inferenciais não se mostrarem significativos, em termos descritivos, podemos ver que houve aumento nas médias das duas plosivas. Para ambas as consoantes, verificamos um aumento nos valores descritivos entre as três primeiras etapas. Ambas as consoantes, três anos depois, apresentam médias menores que as do Pré-teste, em termos descritivos.

De forma surpreendente, conforme já apontado, o GC apresentou uma diferença marginalmente significativa ( $p=.098$ ) para os valores de /k/. Entretanto, conforme mostrarão os testes *post-hoc*, as diferenças entre as etapas não serão significativas, de modo que possamos atribuir tal fato surpreendente a alguma diferença individual de um aprendiz que, no Pré-teste, ainda apresentava valores de VOT mais baixos do que os demais, mas que, com o tempo, “alcançou” os demais colegas. Na próxima tabela, trazemos os resultados do Teste de Wilcoxon (com correção de Bonferroni).

**Tabela 5:** Resultados do teste *post-hoc* de Wilcoxon (com correção de Bonferroni) para a Tarefa de Produção.

	<b>Grupo 1 - Experimental</b>		
	/p/	/t/	/k/
<i>Pré-teste vs. Pós-teste</i>	n.s.	-----	-----
<i>Pré-teste vs. Pós-teste Postergado 1</i>	n.s.	-----	-----
<i>Pré-teste vs. Pós-teste Postergado 2</i>	$p = 0,18^*?$	-----	-----
<i>Pós-teste vs. Pós-teste Postergado 1</i>	n.s.	-----	-----
<i>Pós-teste vs. Pós-teste Postergado 2</i>	n.s.	-----	-----
<i>Pós-teste Postergado 1 vs. Pós-teste Postergado 2</i>	$p = 0,18^*?$	-----	-----

	<b>Grupo 2 - Controle</b>		
	/p/	/t/	/k/
<i>Pré-teste vs. Pós-teste</i>	-----	-----	-----
<i>Pré-teste vs. Pós-teste Postergado 1</i>	-----	-----	n.s.
<i>Pré-teste vs. Pós-teste Postergado 2</i>	-----	-----	-----
<i>Pós-teste vs.</i>			
<i>Pós-teste Postergado 1</i>	-----	-----	-----
<i>Pós-teste vs.</i>			
<i>Pós-teste Postergado 2</i>	-----	-----	-----
<i>Pós-teste Postergado 1 vs. Pós-teste Postergado 2</i>	-----	-----	-----

**Nota:** ----- não se aplica (os resultados do teste de Friedman não foram significativos), n.s. = não significativo, \*? = marginalmente significativo ( $p < 0,02$ ), \*  $p < 0,017$ .

As etapas que apresentaram valores marginalmente significativos para /p/, no GE, são entre Pré-teste e Pós-teste Postergado 2, e entre Pós-teste Postergado 1 e Pós-teste Postergado 2. Estas diferenças possivelmente devem-se à diminuição expressiva nos valores das médias da última etapa do experimento. Com isso podemos concluir que o GE, mesmo após receber o treinamento perceptual e apresentar, num sentido geral, melhora inferencial significativa na percepção, não a generalizou para a produção. Pelo contrário, além de não ter apresentado evolução nas etapas logo após o treinamento, houve ainda uma queda evidente em seus valores três anos mais tarde, com os valores abaixo dos do Pré-teste.

A Tabela 6 traz os resultados do Teste Mann-Whitney, para que possamos verificar as diferenças entre-grupos.

**Tabela 6:** Resultados do teste Mann-Whitney para a Tarefa de Produção.

	Pré-teste U	Pós-teste U	Pós-teste Postergado 1 U	Pós-teste Postergado 2 U
/p/	<b>10,000(*)</b>	17,000	19,000	25,000
/t/	23,000	23,000	24,000	21,000
/k/	28,000	25,000	25,000	19,000

**Nota:** \*  $< 0,05$ , \*\*  $p < 0,01$ .

Voltando à Tabela 4, vemos que, para /p/, o GE traz a média de 25,39 ms no Pré-teste, enquanto o GC registra 17,56 ms durante a mesma etapa, o que pode explicar a diferença significativa, apresentada na tabela anterior, desde o começo do experimento (no Pré-teste). Também é importante destacar como ambos os grupos mantiveram médias parecidas em todos os pontos de articulação ao longo das quatro etapas. Em outras palavras, a análise inter-grupos não nos



permite que façamos afirmações contundentes sobre o papel do treinamento na produção de segmentos, sobretudo porque a consoante labial (a única que, na análise intra-grupos, apresentava diferenças significativas entre as etapas no Grupo Experimental) demonstrou diferenças significativas entre os dois grupos já no Pré-teste.

Em suma, no que diz respeito às três primeiras coletas, semelhantemente a verificado em Alves; Luchini (2017), a verificação dos dados de produção não nos permite pensar em uma generalização do treinamento perceptual à produção de VOT de /t/ e /k/ iniciais. Talvez tal resultado se deva ao fato de que, mesmo no Pré-teste, os valores de /k/ já se encontravam dentro do esperado por falantes nativos (em função do alto valor de VOT, ou “semi-aspiração”, que ocorre em tal segmento na L1 dos participantes, conforme apontado por M. Alves, 2015 e Kupske, 2016). Semelhantemente, os valores de /t/, ainda que não alcançassem a média nativa, já se encontravam muito próximos de tal valor desde o Pré-teste. No que diz respeito ao segundo Pós-teste Postergado, realizado três anos após a segunda etapa, os valores de VOT tendem a ser ainda mais baixos do que aqueles encontrados no terceiro momento de coleta de dados. Quanto à produção de /p/, assim como em Alves; Luchini (2017), em que foram encontrados resultados significativos no teste de Friedman, também no presente trabalho, com quatro momentos temporais, foram encontrados resultados significativos. Através dos testes *post-hoc*, verificamos um decréscimo significativo entre o Pré-teste e o último Pós-teste, e entre esse e o Pós-teste Postergado 1. Com tal decréscimo, não podemos falar em efeitos duradouros do treinamento na produção de tal segmento.

## Conclusão

No que diz respeito à percepção, verificamos, na identificação do padrão Zero Natural como sonoro, um aumento nos índices de acuidade no Grupo Experimental, resultando em altos índices que se mantiveram, inclusive, na última etapa de coleta de dados, realizada em 2018, superando o valor de 95%. No que diz respeito à identificação do padrão Zero Artificial como sonoro, o teste de Friedman também apenas mostrou diferenças significativas no Grupo Experimental, apesar de o índice de acuidade nunca haver superado 59,52% em nenhuma das etapas (verificáveis no Pós-teste Imediato e no primeiro Pós-teste Postergado), em oposição ao índice máximo de 28,13%, encontrado no Grupo Controle, na coleta de 2018. Em outras palavras, parece que o treinamento contribuiu, no Grupo Experimental, para um aumento no índice de respostas que evidenciassem uma tomada do VOT como pista prioritária na distinção entre segmentos surdos e sonoros.

Uma vez que os índices de acerto referentes à identificação do padrão Zero Artificial não superaram o valor de 60%, parece não ter havido uma estabilização no que diz respeito à tomada do VOT como pista acústica prioritária nas distinções de sonoridade. Tal fato talvez explique os resultados de produção: não sendo o VOT a pista prioritariamente empregada na percepção, tampouco os alunos

sentem a necessidade de produzir tal pista<sup>19</sup>. Explicam-se, assim, os índices de produção encontrados nos dois grupos. Sobretudo na oclusiva labial /p/, que era a que se encontrava mais distante dos valores nativos (cuja média de VOT se encontra por volta dos 60ms, cf. CHO; LADEFOGED, 1999; SCHWARTZHAUPT, 2012; KUPSKE, 2016), não encontramos um aumento nos índices de VOT, em nenhum dos dois grupos, entre o Pré-teste e o último Pós-teste. De fato, no Grupo Experimental, ainda que tenhamos encontrado um aumento nas médias entre o Pré-teste e o Pós-teste Postergado 1 (o que, em princípio, implicaria um fator motivador para a realização de um segundo pós-teste postergado três anos depois, já que, com a contínua exposição dos aprendizes ao *input*, seriam esperados índices ainda maiores em 2018), verificamos uma queda considerável nos índices de produção na última coleta, índices esses ainda mais baixos do que aqueles encontrados no Pré-teste. Em outras palavras, a suposta melhora verificada um mês após o treinamento não parece ter sido mantida três anos após tal intervenção, quando os alunos se encontravam no final do curso.<sup>20</sup>

Em linhas gerais, podemos dizer que o treinamento perceptual contribuiu modestamente com a percepção, mas não com a generalização à produção, referente aos padrões de VOT das plosivas iniciais do Inglês. Ainda que efeitos na produção tenham sido verificados para /p/ nas produções do Grupo Experimental no Pós-teste Postergado 1, tais índices, ao invés de aumentar, acabaram por diminuir no segundo Pós-teste Postergado. A uma primeira vista, considerando-se, conforme apontam Sakai; Moorman (2018), que uma contribuição importante do treinamento perceptual é promover efeitos significativos na produção dos aspectos-alvo, poderíamos pensar que tal metodologia de intervenção pedagógica acabou por não exercer os efeitos desejados. É provável que tal fato tenha se dado, justamente, pelo fato de os efeitos referentes à percepção terem se mostrado modestos. Uma vez que os alunos não foram treinados na produção, e que tal produção seria decorrência do trabalho perceptual (a partir do qual os aprendizes passariam a tomar o VOT como pista prioritária e, a partir de tal atribuição de *status* fundamental a tal pista, poderiam, com a oposição contínua ao *input* e o seu decorrente processamento, passar a realizar tal pista na produção.<sup>21</sup>

Ao apresentarmos estes dados, gostaríamos de discutir possíveis explicações para tais índices, bem como apontar as implicações de tais resultados para o universo de sala de aula. Primeiramente, podemos dizer que, pelo menos no Pós-teste Postergado 1, tanto os índices de acuidade de identificação do Padrão Zero Artificial quanto os valores de VOT de /p/ atingem o seu pico. Caberia, portanto, uma maior investigação sobre como foi a exposição desses aprendizes ao longo do curso após tal momento de coleta, ou seja, do quarto ao sexto semestre de Letras.

Frente a tal questão, mencionamos, aqui, uma das limitações do estudo: não termos acompanhado o desenvolvimento dos aprendizes durante todo o curso de graduação. De fato, o longo intervalo de tempo entre o primeiro e o último ano de curso dos estudantes, ainda que seja caracterizado por dois marcos importantes durante a vida acadêmica dos participantes, acaba por perder um acompanhamento de como foi a trajetória dos aprendizes ao longo dos anos

na Universidade. Tal acompanhamento longitudinal teria permitido estabelecer conjunturas, inclusive, sobre a qualidade da exposição desses aprendizes ao *input* com os padrões-alvo,<sup>22</sup> além de possibilitar a verificação da realização de algum tipo de atividade de tomada de consciência dos aprendizes para tais aspectos, o que parece não ter sido o caso. Em recentes trabalhos, De Bot (2015) e Lowie; Verspoor (2015, 2018) chamam a atenção para a necessidade de um acompanhamento longitudinal dos aprendizes em intervalos contínuos e em janelas de tempo menores, para que se possam verificar as oscilações nos sistemas em desenvolvimento dos aprendizes. Ao reconhecermos tal limitação de nosso trabalho, chamamos a necessidade de estudos futuros que, ainda que verifiquem as produções dos aprendizes no início e no final do curso de Letras, acompanhem os aprendizes em seu desenvolvimento em escalas menores, tais como as semestrais (conforme realizado em Lima Jr., 2017), visando apontar possíveis fatores que mostrem que os resultados dos aprendizes em dois pontos distintos não são abruptos, mas refletem mudanças graduais.

Ao termos levantado, também, a possibilidade de que, ao longo de seu curso de graduação, os aprendizes possam não ter tido uma exposição massiva aos itens-alvo (ou, pelo menos, a atividades que ressaltassem e os levassem a prestar atenção a tais itens alvo no *input*), chamamos a atenção para uma das principais implicações do presente estudo para a sala de aula: a necessidade não somente de exposição massiva às formas-alvo, mas também de oportunidades que levem o aprendiz a atentar à necessidade efetiva de tais aspectos fonético-fonológicos para a inteligibilidade. Tais oportunidades não devem resumir, unicamente, a um período instrucional: tarefas de tomada de consciência, implícitas ou explícitas (cf. ELLIS, 2005; ANDRINGA; REBUSCHAT, 2015; LEOW, 2015), são fundamentais para, no caso em questão, ressaltar a necessidade de o aprendiz atentar a tais pistas acústicas, produzindo um VOT longo de modo a evitar problemas de comunicação. Em sua meta-análise sobre os efeitos do treinamento perceptual, Sakai; Moorman (2018) consideram como válida a atividade de chamar explicitamente a atenção do aprendiz para o aspecto fonético-fonológico que está sendo treinado. Tal consideração vai ao encontro dos resultados verificados em Alves; Luchini (2017). No estudo em questão, além dos grupos experimental e controle aqui revisitados, contávamos com um grupo adicional, cujos participantes, além de expostos ao treinamento perceptual, eram instruídos explicitamente sobre o aspecto-alvo para o qual deveriam prestar atenção. Tal grupo foi o único que apresentou melhoras significativas, também, na produção. Em outras palavras, em termos de implicações para o ensino, os resultados dos trabalhos aqui discutidos chamam a atenção para a necessidade constante de exposição e de utilização de uma diversidade de estratégias, tanto de maior quanto de menor teor explícito, que levem o aprendiz a atentar às formas-alvo. No caso do presente estudo, cujos participantes são alunos do curso de graduação em Letras, tais práticas cobram fundamental importância, por contribuir tanto com a formação teórica quanto com o desenvolvimento linguístico desses futuros profissionais da área.

É importante reconhecer que o presente estudo apresentou uma série de limitações adicionais, sobretudo no que concerne ao número de sessões de treinamento e de participantes. No que diz respeito ao número de sessões, uma vez que o trabalho replica a metodologia de Alves; Luchini (2017), dispomos do mesmo número limitado de encontros para a aplicação do treinamento perceptual. Reconhecemos que este número baixo de sessões (que, inclusive, pode ter sido um dos fatores pelos quais os índices de acuidade de percepção do padrão Zero Artificial não ultrapassaram os 60%) é decorrente das limitações de tempo que têm de ser enfrentadas por todo e qualquer estudo em sala de aula, uma vez que a realização deste tipo de investigação implica que professores regentes disponibilizem suas turmas e horários de aula para a realização dos experimentos.

Já no que diz respeito ao número de participantes, ressaltamos, conforme já feito na Metodologia, que o longo intervalo de tempo entre as etapas foi responsável pela alta queda nos índices de participação. O baixo número de participantes é, possivelmente, o grande responsável por termos encontrado apenas diferenças marginalmente significativas nos testes *post hoc*. Nesse sentido, estudos recentes à luz do Paradigma da Complexidade (LOWIE; VERSPOOR, 2015; 2018; LIMA JR., 2016) têm advertido a necessidade de análises das trajetórias individuais dos participantes, agregadas às análises inferenciais e de grupo, uma vez que ambos os métodos de verificação de dados, de forma conjugada, podem revelar aspectos adicionais acerca das trajetórias dos aprendizes e do desenvolvimento individual. Ao considerarmos que cada aprendiz tem suas características individuais e seu desenvolvimento é fruto de tais características (DE BOT, 2015; LOWIE; VERSPOOR, 2015; 2018), em trabalho futuro pretendemos realizar, também, a análise da trajetória de cada um dos participantes do estudo, de modo a verificarmos se, considerados os aprendizes individualmente, o treinamento perceptual proposto pode ter seus efeitos mais claramente evidenciados em algum dos participantes.

Antes de terminarmos, é importante mencionarmos que o estudo aqui realizado teve como participantes alunos do curso de Graduação em Letras, que podem ser considerados aprendizes privilegiados no que concerne às oportunidades de exposição e uso da língua. Tal motivo talvez explique os altos índices de produção de VOT de /t/ e /k/ por parte dos dois grupos já no Pré-teste, de modo que o treinamento pudesse vir a exercer efeitos estatisticamente claros apenas na oclusiva labial /p/. É possível que, com aprendizes que não tomem a Língua Inglesa como objeto de formação profissional, o treinamento possa vir a se mostrar necessário, também, com as plosivas coronais e dorsais. Por sua vez, reconhecemos que, no caso de aprendizes fora do ambiente das Letras, será necessário um número ainda maior de sessões de treinamentos e oportunidades de exposição e de tomada de consciência acerca das formas-alvo.

Com base no aqui discutido, consideramos que o trabalho que aqui se encerra abre espaço para uma série de estudos novos, os quais, a partir dos resultados aqui apontados, têm aplicação direta no ensino. Sugerimos a proposição de tarefas de treinamento perceptual conjugadas à explicitação dos aspectos a serem

atentados, com um número maior de sessões e com um intervalo contínuo entre as sessões de treinamento. Além disso, chamamos a atenção para a necessidade de um acompanhamento da trajetória dos aprendizes em intervalos de tempos menores e regulares, nos quais o professor possa se certificar de estar provendo insumos necessários para que a pista de VOT possa ser tomada como prioritária pelos aprendizes. Acreditamos que, a partir de tais estudos futuros e das reflexões aqui lançadas, poderemos contribuir para a produção de padrões fonético-fonológicos que resultem em uma fala mais inteligível em Língua Inglesa.

### Notas

1. Em Inglês, a plosiva sonora pode ser produzida variavelmente, tanto com o padrão de VOT Zero quanto com VOT Negativo (LISKER; ABRAMSON, 1964; CHO; LADEFOGDED, 1999). Estudos desenvolvidos por nosso grupo de pesquisa (AUTOR 5) evidenciam que ambos os padrões são identificados como instâncias de plosivas sonoras por falantes norte-americanos e aprendizes brasileiros.
2. Além destes dois padrões, os experimentos de identificação realizados também testaram a percepção dos padrões VOT Negativo e VOT Positivo, os quais, independentemente de os ouvintes serem norte-americanos, brasileiros ou hispânicos (esses dois últimos grupos com grau de proficiência alto na L2), foram identificados em quase 100% dos casos como instâncias de segmentos sonoros e surdos, respectivamente.
3. Os participantes em questão apresentavam minimamente o nível C1 de proficiência, de acordo com o Marco Comum Europeu de Referência para o Ensino de Línguas. Para maiores informações sobre a caracterização de tal nível, veja-se <https://www.coe.int/en/web/common-european-framework-reference-languages>.
4. Os estudos de inteligibilidade supracitados foram realizados com ouvintes cuja L1 é de variedade norte-americana de inglês. Entretanto, com base nas considerações de Jenkins (2000) e Walker (2010), podemos depreender que a aspiração é um aspecto-chave para a inteligibilidade independentemente do dialeto de língua inglesa em questão.
5. Para uma revisão do estado da arte e uma meta-análise sobre estudos referentes à instrução de pronúncia, vejam-se, respectivamente, Thomson; Derwing (2015) e Lee; Jang; Plonsky (2015).
6. Para uma meta-análise sobre estudos de treinamento perceptual, veja-se Sakai; Moornan (2018).
7. Para maiores detalhes sobre os estímulos, produzidos por falantes norte-americanos, veja-se a seção referente à Metodologia.
8. É importante mencionar que o estudo de AUTOR 4 contava, ainda, com um terceiro grupo, no qual as sessões de treinamento perceptual eram acrescidas de instrução explícita sobre as formas-alvo. Uma vez que o referido grupo não compreende o foco deste trabalho, dado que o grupo em questão não será replicado na investigação empírica aqui proposta, sugere-se a leitura de AUTOR 4 para uma exposição mais detalhada dos resultados referentes a tal grupo.
9. Os acadêmicos do referido curso podem optar pela carreira de Licenciatura ou Bacharelado em Letras, bem como pela ênfase única (somente em Inglês) ou dupla (em Português-Inglês). Nas aulas de Inglês, os acadêmicos do Bacharelado e da Licenciatura participam dos encontros em uma mesma turma.

10. O projeto de pesquisa no qual o presente trabalho se insere foi aprovado pelo Comitê de Ética em Pesquisa da Instituição (parecer 648.674, de 24/04/2014).
11. Reconhecemos que o baixo número de participantes representa uma das limitações do estudo, conforme será apontado ao final do trabalho. Ao mesmo tempo, também consideramos a perda amostral como inevitável, dado o intervalo de tempo de três anos entre a primeira e a última coleta de dados do estudo.
12. Retomamos a informação de que, para fins de treinamento, somente foram utilizados os padrões VOT Positivo e Zero Artificial.
13. Os estímulos usados na tarefa de identificação são diferentes dos usados nas sessões de treinamento para que, caso houvesse uma melhora nos índices de acurácia da tarefa de identificação, essa indicasse a capacidade de generalização para outros itens lexicais.
14. Mesmos falantes dos estímulos presentes no treinamento perceptual.
15. Seguindo-se AUTOR 4, dos três *types* de cada ponto de articulação, um deles havia sido usado durante o treinamento ('*pee*', '*tip*', '*kit*'), um havia sido aplicado nos pré e nos pós-testes ('*pit*', '*tip*', '*kill*') e outro representava um novo item lexical ainda não testado ('*peer*', '*team*', '*keel*'). A partir deste *design*, esperávamos investigar se os itens novos implicariam menores valores de VOT dos que os já conhecidos. Por razões de delimitação, deixaremos esta verificação para um trabalho futuro.
16. Versão 6.0 nas três primeiras etapas, e versão 6.0.41 na quarta.
17. Para os padrões VOT Negativo, VOT Zero Natural e VOT Zero Artificial, foi considerada correta a atribuição do *status* sonoro; no VOT Positivo, foi considerada correta a atribuição do *status* surdo. Conforme já esclarecido, somente foram consideradas corretas as escolhas cujos pontos de articulação correspondessem ao ponto da consoante do estímulo. Por exemplo, se o aprendiz ouviu um padrão VOT Negativo de /b/ e marcou /d/, tal resposta não foi considerada correta.
18. Neste estudo foram realizados testes não-paramétricos, uma vez que os Testes de Normalidade de Kolmogorov-Smirnov e Shapiro-Wilk indicaram uma distribuição não normal dos dados.
19. Tal constatação é acentuada, ainda, pelo fato de que os aprendizes nunca foram treinados a produzir tal pista (dado que, em nosso estudo, a produção seria consequência do aumento dos índices de acuidade na percepção, ocasionado pelo treinamento). Não tendo havido um aumento categórico na percepção, os resultados não-significativos na produção são, também, esperados.
20. Conforme será discutido ao final do trabalho, se, por um lado, a coleta três anos após a realização do primeiro Pós-teste Postergado pôde ser capaz de prover informações claras sobre os momentos desenvolvimentais dos participantes na entrada e no final do curso, por outro, o longo intervalo de tempo entre o primeiro Pós-teste Postergado e o último caracterizam uma limitação do estudo. Tal longo espaço de tempo deu-se em função de dificuldades logísticas referentes a um novo recrutamento dos mesmos participantes. Acreditamos, dessa forma, que um acompanhamento longitudinal mais frequente dos participantes, ao longo de todo o curso de graduação, possa revelar importantes aspectos referentes à trajetória do aprendiz.
21. Consideramos que a produção da aspiração não constitui grande dificuldade articulatória para os aprendizes brasileiros e hispânicos de inglês. Como evidência de tal fato, mencionamos o estudo de Alves; Luchini (2017). No referido estudo, aprendizes que foram instruídos explicitamente acerca de tal pista passaram a produzi-la significativamente, mesmo com índices baixos de acuidade na tarefa de percepção. Dessa forma, atribuímos os índices baixos de produção do presente estudo ao fato de que a exposição perceptual (sem explicitação) propiciada

pelo treinamento constituiu o único meio a partir do qual o aprendiz passaria a atender a tal pista na percepção e, por conseguinte, na produção. Uma vez que os resultados da percepção mostraram alta variabilidade nos dados, é possível que o VOT não tenha sido plenamente tomado como pista perceptual prioritária pelos aprendizes, o que, por conseguinte, acarreta os índices baixos da produção.

22. Conforme aponta França (2015), professores brasileiros de Inglês, com alto grau de proficiência na L2, tendem a não aspirar as plosivas iniciais. Não descartamos a possibilidade de que, ao longo do curso, os aprendizes possam ter sido expostos, em sala de aula, a um insumo sem aspiração na língua estrangeira.

### Referências

- ABRAMSON, A. S. WHALEN, D. H. Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, 63, p. 75-86, 2017.
- AHN, S. The role of tongue position in laryngeal contrasts: an ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics*, 71, p. 451-467, 2018.
- AL-TAMIM, J.; KHATTAB, G. Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, 71, p. 306-325, 2018.
- ALVES, M. A. *Estudo dos parâmetros acústicos relacionados à produção das plosivas do português brasileiro na fala adulta: análise acústico-quantitativa*. Tese (Doutorado em Letras). Florianópolis: Universidade Federal de Santa Catarina, 2015.
- ALVES, U. K.; MOTTA, C. S. Focusing on the right cue: Perception of voiceless and voiced stops in English by Brazilian learners. *Phrasis – Studies in Language and Literature*, 5, 31-50, 2014.
- ALVES, U. K.; ZIMMER, M. C. Perception and production of English VOT patterns by Brazilian learners: the role of multiple acoustic cues in a DST perspective. *Alfa: Revista de Linguística* (UNESP. Online), 59, p. 155-175, 2015.
- ALVES, U. K.; LUCHINI, P. L. Percepción de la distinción entre oclusivas sordas y sonoras iniciales del inglés (LE) por estudiantes argentinos: datos de identificación y discriminación. *Revista Lingüística (ALFAL)*, 32, p. 25-29, 2016.
- ALVES, U. K.; LUCHINI, P. L. Effects of perceptual training on the identification and production of word-initial voiceless stops by Argentinean learners of English. *Ilha do Desterro*, 70, n. 3, p. 15-32, 2017.
- ANDRINGA, S.; REBUSCHAT, P. New directions in the study of implicit and explicit learning – an introduction. *Studies in Second Language Acquisition*, 37, p. 185-196, 2015.
- BOERSMA, P.; WEENINK, D. *Praat: Doing Phonetics by Computer*. Version 5.3.48. 2013. Disponível em: <www.praat.org>. Acesso em: 20 jan. 2015.
- CHO, T.; LADEFOGED, P. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, n. 27, p. 207-229, 1999.
- DE BOT, Kees. Rates of Change: Timescales in Second Language Development. In: DÖRNYEI, Z.; MacINTYRE, P. D.; HENRY, A. (eds). *Motivational Dynamics in Language Learning*. Bristol: Multilingual Matters, 2015, p. 29-37.
- ELLIS, N. At the interface: dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, p. 305-352, 2005.
- FLEGE, J. Second language speech learning: theory, findings, and problems. In: STRANGE, W. (ed). *Speech perception and language experience: issues in cross-language research*. Maryland: York Press, 1995, p. 233-278.

- FRANÇA, K. V. *A aquisição da Aspiração*. Curitiba: Appris, 2015.
- HOLT, L.; LOTTO, A. Cue weighting in auditory categorization: implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, v. 119, n. 5, p. 3059-3071, 2006.
- JENKINS, J. *The phonology of English as a Lingua Franca*. Oxford: Oxford University Press.
- KLEBER, F. VOT or quantity: what matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. *Journal of Phonetics*, v. 71, p. 468-486, 2018.
- KUPSKE, Felipe Flores. *Imigração, Atrito e Complexidade: a produção das oclusivas surdas iniciais do Inglês e do Português por Sul-Brasileiros residentes em Londres*. Tese (Doutorado em Letras). Porto Alegre: Universidade Federal do Rio Grande do Sul, 2016.
- LADD, D. R.; SCHMID, S. Obstruent voicing effects on F0, but without voicing: phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics*, 71, p. 229-248, 2018.
- LEE, J.; JANG, J.; PLONSKY, L. The effectiveness of second language pronunciation instruction: a meta-analysis. *Applied Linguistics*, 36, n. 3, p. 354-366, 2015.
- LEOW, R. P. *Explicit learning in the L2 classroom: a student-centered approach*. New York: Routledge, 2015.
- LEOW, R. P.; ADRADA-RAFAEL, S. La atención y la concienciación en el campo de la adquisición de segundas lenguas. In: ORTIZ-PREUSS, E.; FINGER, I. (org.). *A dinâmica do processamento bilíngue*. Campinas: Pontes Editores, 2018, p. 191-231.
- LIMA JR., R. M. A necessidade de dados individuais e longitudinais para análise do desenvolvimento fonológico de L2 como sistema complexo. *Revista Virtual de Estudos da Linguagem*, 14, n. 27, p. 203-225, 2016.
- LIMA JR., R. M. The influence of metalinguistic knowledge of segmental phonology on the production of English vowels by Brazilian undergraduate students. *Ilha do Desterro*, v. 70, n. 3, p. 117-130, 2017.
- LISKER, L.; ABRAMSON, A. S. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, n. 3, p. 384-422, 1964.
- LOWIE, W. M.; VERSPOOR, M. H. Variability and Variation in Second Language Acquisition Orders: A Dynamic Reevaluation. *Language Learning*, 65, n.1, p. 63-88, 2017.
- LOWIE, W. M.; VERSPOOR, M. H. Individual Differences and the Ergodicity Problem. *Language Learning*, p. 1-23, 2018.
- MOTTA-ÁVILA, C. *A formalização fonético-fonológica da percepção de plosivas surdas sob múltiplas manipulações de Voice Onset Time (VOT) por brasileiros e americanos à luz do modelo "Biphon"*. Dissertação (Mestrado em Letras). Pelotas: Universidade Católica de Pelotas, 2017.
- RAUBER, A. S.; RATO, A.; SANTOS, G. R.; KLUGE, D. C.; FIGUEIREDO, M. *TP: Testes de Percepção e Treinamento Perceptual com Feedback Imediato – Versão 3.1*. 2013. Disponível em [http://www.worken.com.br/tp\\_regfree.php](http://www.worken.com.br/tp_regfree.php). Acesso em 04 de janeiro de 2019.
- SAKAI, M.; MOORMAN, C. Can perceptual training improve the production of second language phonemes? A meta-analytic review of 25 years of perceptual training research. *Applied Psycholinguistics*, 39 (1), 187-224, 2018.



- SCHWARTZHAUPT, B. M. *Factors influencing Voice Onset Time: analyzing Brazilian Portuguese, English and interlanguage data*. Monografia (Trabalho de Conclusão de Curso). Porto Alegre: Universidade Federal do Rio Grande do Sul, 2012.
- SCHWARTZHAUPT, B. M. *Testing intelligibility in English: the effects of Positive VOT and contextual information in a sentence transcription task*, 2015. Dissertação (Mestrado em Letras). Porto Alegre: Universidade Federal do Rio Grande do Sul, 2015.
- SCHWARTZHAUPT, B. M.; ALVES, U. K.; FONTES, A. B. A. L. The role of L1 knowledge on L2 speech perception: investigating how native speakers and Brazilian learners categorize different VOT patterns in English. *Revista de Estudos da Linguagem*. 23(2), p. 311-334, 2015.
- THOMSON, R. I.; DERWING, T. M. The effectiveness of L2 pronunciation instruction: a narrative review. *Applied Linguistics*, 36, n. 3, p. 326-344, 2015.
- WALKER, R. *Teaching the pronunciation of English as a Lingua Franca*. Oxford: Oxford University Press, 2010.
- YAVAS, M.; WILDERMUTH, R. The effects of place of articulation and vowel height in the acquisition of English aspiration stops by Spanish speakers. *IRAL*, 44, p. 251-263, 2006.

Recebido em: 09/01/2019

Aceito em: 07/03/2019