

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE LETRAS

Felipe Rodrigues Kampff

**DESENVOLVIMENTO DE PADRÕES DE *VOICE ONSET TIME* (VOT) EM  
PLOSIVAS INICIAIS DO INGLÊS (L2) POR APRENDIZES DO SUL DO BRASIL  
EXPOSTOS OU NÃO A TREINAMENTO PERCEPTUAL: UMA VERIFICAÇÃO  
LONGITUDINAL EM QUATRO ETAPAS**

Porto Alegre

2018

Felipe Rodrigues Kampff

**DESENVOLVIMENTO DE PADRÕES DE *VOICE ONSET TIME* (VOT) EM  
PLOSIVAS INICIAIS DO INGLÊS (L2) POR APRENDIZES DO SUL DO BRASIL  
EXPOSTOS OU NÃO A TREINAMENTO PERCEPTUAL: UMA VERIFICAÇÃO  
LONGITUDINAL EM QUATRO ETAPAS**

Trabalho de conclusão de curso,  
apresentado como requisito parcial para a  
obtenção do título de Licenciado em Letras  
pela Universidade Federal do Rio Grande  
do Sul.

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

Porto Alegre

2018

Felipe Rodrigues Kampff

**DESENVOLVIMENTO DE PADRÕES DE *VOICE ONSET TIME* (VOT) EM  
PLOSIVAS INICIAIS DO INGLÊS (L2) POR APRENDIZES DO SUL DO BRASIL  
EXPOSTOS OU NÃO A TREINAMENTO PERCEPTUAL: UMA VERIFICAÇÃO  
LONGITUDINAL EM QUATRO ETAPAS**

Trabalho de conclusão de curso,  
apresentado como requisito parcial para a  
obtenção do título de Licenciado em Letras  
pela Universidade Federal do Rio Grande  
do Sul.

Aprovado em Porto Alegre, \_\_\_\_ de \_\_\_\_\_ de 2018.

BANCA EXAMINADORA:

---

Prof<sup>a</sup>. Dra. Letícia Pereyron (ESPM-Sul)

---

Prof. Dr. Reiner Vinicius Perozzo (UFBA)

---

Prof. Dr. Ubiratã Kickhöfel Alves (UFRGS - Orientador)

## AGRADECIMENTOS

Aos que lerem esses agradecimentos, ouçam os estrondosos aplausos de gratidão que de mim emanam.

Minha gratidão eterna por concluir uma etapa tão importante não poderia ser para outras pessoas que não as que mais me incentivaram a chegar aqui: meus pais, Ayrton e Lara, e minha madrinha, Melissa. Aos dois primeiros, por nunca deixarem me faltar nada, por me guiarem, acolherem e aceitarem do jeito que sou. E eu os acolhi e os aceitei do jeito que são. À minha dinda, pelo eterno companheirismo e pelas experiências que me proporcionou, desde os cursos de inglês e as viagens, ao cursinho preparatório para o vestibular. A eles, minha mais eterna e impagável gratidão. Agradeço, também, à minha irmã, Patrícia; sem o exemplo de força e superação que a vi demonstrando dia após dia também não teria feito diversas reflexões. Não é o sangue que te faz minha irmã, mas a admiração por quem tu é. Finalmente, agradeço aos meus avós, principalmente Osny e Iracema, sempre presentes na minha vida e criação. Sem eles, também diversas oportunidades não teriam sido possíveis.

Agradeço à Letras, enquanto curso e espaço que contribuiu para que eu seja quem sou hoje. E me orgulho de ser. Letrista, licenciando, licenciado (em breve). Saio, também, mais humano, mais consciente e, após todo o conhecimento adquirido, mais humilde. A batalha de quem educa é difícil, mas é a que produz os resultados mais lindos. Porém, o curso não poderia ter me mudado se não fosse pelas pessoas que nele estão.

Aos professores, tenho diversos agradecimentos. Agradeço à Lúcia Rebello, minha primeira orientadora e razão pela minha paixão pela língua latina; à Lia, por ser luz da cabeça aos pés de tanto que brilha quando educa falando em educação - ser luz é para poucos, e ela é; à Sabrina, por, sem saber, ter me feito continuar na licenciatura após me proporcionar um dos melhores semestres da graduação; à Sandra Maggio, Ana Welp e Simone, três professoras que, além de tudo, tornam-se amigas por conta de todo o carinho e apoio que nos fornecem dentro e fora da sala de aula; finalmente, ao Ubiratã, o Bira. Com ele, inicio e termino minha vida de pesquisador enquanto graduando. Sem ele, este trabalho não estaria aqui, nem a paixão pela fonologia, nem as risadas proporcionadas em reuniões e encontros entre orientador e orientandos.

Porém, não é com os professores que nós, alunos, convivemos diariamente. Aos meus colegas, agradeço tudo - as risadas, os choros, os abraços, os beijos, os RUs, as provas, os trabalhos e até mesmo a Ocupação. À Laura, por ser a pessoa mais amavelmente pura que conheci durante esses anos, me acompanhando desde 2013 pelas idas e vindas do curso. Não há palavras que transcrevam meu amor e admiração por ti; à Thamís, por ser, literalmente, a primeira pessoa do curso que conheci, numa tarde escaldante porto-alegrense, e ter me proporcionado momentos de felicidade imensa; à Thielle, *my dearest*, que chegou um ano depois da minha entrada e, mesmo assim, nos conhecemos uma vida toda - muito, muito obrigado!; à Lauris, com quem dividi momentos de alegria, assim como viagens e o início, juntos, da vida de pesquisa acadêmica; assim como agradeço à Jube e à Carol, minhas guaipecas. Amo vocês!

Agradeço também ao Patrick, que representa uma amizade tão forte que, mesmo entrando no mesmo ano, só fomos nos aproximar depois da metade da

graduação. Poucos chegam no auge de me aguentar que nem tu aguenta, então muito obrigado! À Lari, presente em alto e bom som desde 2013, da sala de aula à beira da praia, obrigado por ser sempre companheira. Junto com os dois, meu muito obrigado aos cafonas Viviane, Juliana, Mirvana, Filipe e Pedro pela amizade e os bons momentos juntos.

Além disso, há pessoas especiais que chegam quando já estamos prestes a sair pela soleira da porta. Por isso, meu mais sincero agradecimento à Luiza e ao seu coração de ouro. Poucas pessoas são que nem ela, que dão a mão mesmo quando estão se afogando, que dividem cama, comida e, principalmente, afeto sem cobrar nada em troca. Não muda nunca, por favor! Do mesmo jeito que quem chega depois deixa sua marca, quem chega e sai antes também. Por isso, agradeço ao Gabs pelo tipo de amizade que pensei que nunca teria. Obrigado por ser essa pessoa especial. A vida fez os nossos caminhos se cruzarem e isso me faz feliz demais.

Há, também, os amigos que nunca chegaram perto de estar na Letras, mas acompanharam todo o percurso. Esses amigos são o significado de amizade verdadeira, aquela em que os anos passam e mesmo assim nada muda. À Bigah, companheira desde o colégio, esse ser único e especial de tantas maneiras que o meu muito obrigado não é suficiente; à Dáy, minha Beyoncé, minha amiga e, verdadeiramente, minha irmã. Ter te conhecido foi uma felicidade que vale por essa vida e mais três. E espero que continuemos juntos em todas elas. E, por fim, ao Afonso, que conheci pelas andanças da vida, e de vez em quando, entre uma visita ou outra, caminhamos juntos, meu muito obrigado na constante luta por amadurecimento.

À Letícia e ao Reiner também sou grato. Além de membros de banca, foram amigos. É muito simbólico e importante que vocês, que me viram iniciar a vida de pesquisa acadêmica, façam parte do fechamento deste ciclo. Junto com eles, agradeço à Jeniffer e à Bruna, sempre presentes e dispostas a ajudar um nervoso iniciando científico antes de apresentações de trabalho.

Finalmente, agradeço a mim. Permitir-se refletir e mudar é algo que deve ser trabalhado. Se não tivesse me permitido reflexões e mudanças, não seria quem sou hoje; e sou feliz.

## RESUMO

Partindo de uma visão de Língua como Sistema Adaptativo Complexo (DE BOT *et al.*, 2009; LARSEN-FREEMAN; CAMERON, 2008; LARSEN-FREEMAN, 2015), pretendemos investigar o desenvolvimento de padrões de percepção e de produção do inglês (L2) por aprendizes brasileiros, expostos ou não a uma tarefa de treinamento perceptual, a partir da descrição dos dados em grupo e individuais dos participantes. Estudos recentes (LOWIE; VERSPOOR, 2015, 2018) sugerem uma análise da variabilidade dos estudantes, a partir das suas trajetórias individuais, uma vez que o desenvolvimento de uma segunda língua não se dará igual para todos, o que não aparece em estudos que levam em conta somente os resultados do grupo. Neste trabalho, partimos do princípio de que, no desenvolvimento da fonologia de uma língua estrangeira, os aprendizes podem focar em pistas acústicas que não desempenham um papel decisivo no estabelecimento das diferenças funcionais na sua L2 (ALVES; MOTTA, 2014; ZIMMER; ALVES, 2015; ALVES; LUCHINI, 2016). Como esta "sintonização" representa um desafio importante, tarefas de treinamento perceptual podem ter papel benéfico em orientar aprendizes para os aspectos fonéticos que são atentados pelos falantes nativos da língua alvo. Uma destas pistas é o *Voice Onset Time* (VOT), definido como o momento entre a soltura de uma plosiva e a vibração das cordas vocais referentes ao segmento seguinte. Há três padrões de VOT nas línguas naturais (LISKER; ABRAMSON, 1964): o padrão Negativo, em que as cordas vocais vibram antes da soltura; o Positivo, em que as cordas vocais vibram após a soltura, produzindo um momento de surdez conhecido como 'aspiração'; e o Zero Natural, quando a vibração das cordas vocais ocorre simultaneamente, ou quase simultaneamente, à soltura. Falantes nativos de inglês utilizam os padrões Zero e, variavelmente, Negativo, para a produção de plosivas sonoras, e o padrão Positivo para a produção de plosivas surdas. Um quarto padrão, chamado Zero Artificial, foi manipulado, neste estudo, a partir da retirada da aspiração de palavras com VOT Positivo, conservando todas as suas características acústicas do segmento, menos a aspiração. As sessões de treinamento, elaboradas a partir do *software TP* (RAUBER *et al.*, 2013), consistiram em apresentar estímulos de fala natural e estímulos manipulados aos alunos, visando a levá-los a focar na presença ou ausência do VOT Positivo. Para isso, 15 participantes foram divididos em dois grupos: Grupo Experimental, cujos alunos participaram de três sessões de treinamento das plosivas iniciais da língua alvo; e Grupo Controle, que não participou das sessões de treinamento. Todos os participantes realizaram um Pré-teste, um Pós-teste (uma semana após a última sessão de treinamento), um Pós-teste Postergado 1, realizado um mês depois da última sessão, e um Pós-teste Postergado 2, realizado três anos após o treinamento. No Teste de Identificação administrado no Pré-teste e nos três Pós-testes, aos aprendizes foram apresentados estímulos em palavras individuais, a partir dos quais eles eram solicitados a clicar no botão que indicava a consoante inicial da palavra ouvida (/p/, /t/, /k/, /b/, /d/, /g/). Neste teste, estímulos com quatro padrões de VOT (Negativo, Zero, Positivo e Artificial) foram incluídos e apresentados em ordem aleatória. No teste de produção, também aplicado nas quatro etapas, os aprendizes eram solicitados a ler listas de palavras iniciadas pelos segmentos-alvo, apresentadas, isoladamente, em *slides* de *Powerpoint*. Nosso estudo corrobora outros que atestaram efeitos positivos do treinamento na percepção dos participantes, assim como a não generalização dos mesmos para a produção, considerando os resultados do grupo como um todo (RODRIGUES, 2015; KAMPFF; ALVES, 2016; ALVES; LUCHINI, 2016). Em relação à análise individual, nosso estudo encontrou possíveis

efeitos positivos do treinamento em trajetórias individuais, em termos descritivos, para a produção.

**PALAVRAS-CHAVE:** Desenvolvimento fonético-fonológico de L2; treinamento perceptual; variabilidade; *Voice Onset Time*.

## ABSTRACT

Departing from a view of Language as a Complex Adaptive System (DE BOT *et. al.*, 2009; LARSEN-FREEMAN; CAMERON, 2008; LARSEN-FREEMAN, 2015), we aim to investigate the development of perception and production patterns of English (L2) by Brazilian learners, exposed or not to perceptual training sessions, both in the group data and in the participants' individual trajectories. Recent studies (LOWIE; VERSPOOR, 2015, 2018) suggest an analysis of the students' variable patterns, which come from their individual trajectories, since the development of a second language will not be the same to everyone, and this fact is not taken in consideration in group studies. We assume that, while developing the phonology of a foreign language, learners can focus on acoustic cues that do not play a decisive role in the functional distinctions in their L2 (ALVES; MOTTA, 2014; ZIMMER; ALVES, 2015; ALVES; LUCHINI, 2016). As this "tuning in" represents an important challenge, perceptual training tasks can play a beneficial role in guiding learners to focus on the phonetic aspects the same way native speakers do. One of those cues is Voice Onset Time (VOT), defined as the interval of time between the release of a plosive and initial vibration of the vocal folds of the following segment. There are three VOT patterns in natural languages (LISKER; ABRAMSON, 1964): Negative, when the vocal folds vibrate before the release; Positive, when the vocal folds vibrate after the release, creating an interval of voicelessness known as aspiration; and Zero VOT, when the vibration of the vocal folds happens simultaneously, or almost simultaneously, to the release. Native speakers of English produce Zero VOT and, variably, the Negative pattern, in voiced plosives, and Positive VOT in the voiceless ones. A fourth pattern, named Artificial Zero VOT, was manipulated in this study by cutting of an interval of aspiration from the voiceless stops; this caused these consonants to retain all their acoustic characteristics as representative of a voiceless segment, except for aspiration. The training sessions, elaborated on TP software (RAUBER *et al.*, 2013), consisted of presenting natural and manipulated stimuli to the students, aiming to guide them to focus on the presence or absence of the Positive VOT pattern. Fifteen students were divided in two groups: the Experimental Group, whose students took part in three training sessions; the Control Group, which did not take part in training sessions. All participants took part in a Pre-test, a Post-test, and in two Delayed Post-tests. In the Identification Task that took place in the Pre-test and in the three Post-tests, learners were presented to stimuli in individual words, and were asked to click on a button to identify the initial consonant they had heard (/p/, /t/, /k/, /b/, /d/, /g/). In this test, stimuli with the four VOT patterns (Negative, Positive, Zero and Artificial Zero) were included and presented in a random order. In the production test, also applied in all four stages of the experiment, learners were asked to read lists of words with plosives in word-initial position; these words were presented automatically, one by one, in PowerPoint slides. Our study corroborates previous experiments that showed positive effects of perceptual training on learners' perception, but no generalization of these effects in production, considering group data (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; RODRIGUES, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; KAMPFF; ALVES, 2016; ALVES; LUCHINI, 2016; ALVES; LUCHINI, 2017). As for the individual analysis, our study has found possible positive effects of training on the learners' individual trajectories for production.

**KEYWORDS:** L2 Phonetic-phonological development; perceptual training; variability; Voice Onset Time.

## LISTA DE FIGURAS

Figura 1 – As opções de escolha da tarefa de Identificação no TP.....	53
Figura 2 – Tarefa de Identificação com <i>feedback</i> positivo.....	54
Figura 3 – Tarefa de Identificação com <i>feedback</i> negativo.....	55

## LISTA DE TABELAS

Tabela 1 – Índices de acuidade de cada um dos padrões de VOT e resultados dos testes de Friedman referentes à Tarefa de Identificação.....	65
Tabela 2 – Resultados do teste post-hoc de Wilcoxon (com correção de Bonferroni) para o Teste de Identificação.....	64
Tabela 3 – Resultados do teste Mann-Whitney para o Teste de Identificação.....	66
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 Friedmann.....	68
Tabela 5 – Resultados do teste post-hoc Wilcoxon (Bonferroni) para a Tarefa de Produção.....	70
Tabela 6 – Resultados do teste post-hoc de Mann-Whitney para a Tarefa de Produção.....	72
Tabela 7 – Resultados referentes ao Questionário de Experiência e Proficiência Linguística de Scholl e Finger para o Grupo Experimental.....	129
Tabela 8 – Resultados referentes ao Questionário de Experiência e Proficiência Linguística de Scholl e Finger para o Grupo Controle.....	130
Tabela 9 – Valores de produção (valores mínimos e máximos, médias e desvio padrão) de ambos os grupos para /p/.....	132
Tabela 10 – Valores de produção (valores mínimos e máximos, médias e desvio padrão) de ambos os grupos para /t/.....	134
Tabela 11 – Valores de produção (valores mínimos e máximos, médias e desvio padrão) de ambos os grupos para /k/.....	136

## LISTA DE GRÁFICOS

Gráfico 1 – Número de acertos, por parte dos participantes do Grupo Experimental, na Tarefa de Identificação referente ao padrão VOT Zero Natural, em cada uma das etapas do estudo.....	74
Gráfico 2 – Número de acertos, por parte dos participantes do Grupo Controle, na Tarefa de Identificação referente ao padrão VOT Zero Natural, em cada uma das etapas do estudo.....	75
Gráfico 3 – Número de acertos, por parte dos participantes do Grupo Experimental, na Tarefa de Identificação referente ao padrão VOT Zero Artificial, em cada uma das etapas do estudo.....	77
Gráfico 4 – Número de acertos, por parte dos participantes do Grupo Controle, na Tarefa de Identificação referente ao padrão VOT Zero Artificial, em cada uma das etapas do estudo. ....	77
Gráfico 5 – Médias individuais referentes às produções de /p/ dos participantes do Grupo Experimental, durante as etapas do estudo.....	80
Gráfico 6 – Médias individuais referentes às produções de /p/ dos participantes do Grupo Controle, durante as etapas do estudo. ....	80
Gráfico 7 – Médias individuais referentes às produções de /t/ dos participantes do Grupo Experimental, durante as etapas do estudo.....	83
Gráfico 8 – Médias individuais referentes às produções de /t/ dos participantes do Grupo Experimental, durante as etapas do estudo.....	84
Gráfico 9 – Médias individuais referentes às produções de /k/ dos participantes do Grupo Experimental, durante as etapas do estudo.....	87
Gráfico 10 – Médias individuais referentes às produções de /k/ dos participantes do Grupo Controle, durante as etapas do estudo.....	87

Gráfico 11 – Variabilidade dos grupos, Experimental e Controle, referentes às produções de VOT da consoante /p/, em cada um dos quatro momentos de coleta de dados.....	92
Gráfico 12 – Variabilidade dos grupos, Experimental e Controle, referentes às produções de VOT da consoante /t/, em cada um dos quatro momentos de coleta de dados.....	94
Gráfico 13 – Variabilidade dos grupos, Experimental e Controle, referentes às produções de VOT da consoante /k/, em cada um dos quatro momentos de coleta de dados.....	96
Gráfico 14 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pré-teste.....	98
Gráfico 15 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pós-teste.....	99
Gráfico 16 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pós-teste Postergado 1.....	100
Gráfico 17 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pós-teste Postergado 2.....	101
Gráfico 18 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pré-teste.....	104
Gráfico 19 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pós-teste.....	105
Gráfico 20 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pós-teste Postergado 1.....	106
Gráfico 21 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pós-teste Postergado 2.....	107

Gráfico 22 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pré-teste.....	110
Gráfico 23 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pós-teste.....	111
Gráfico 24 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pós-teste Postergado 1.....	112
Gráfico 25 – Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pós-teste Postergado 2.....	113

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	16
<b>2 REFERENCIAL TEÓRICO</b> .....	20
2.1 CONCEPÇÃO DE LÍNGUA COMO UM SISTEMA ADAPTATIVO COMPLEXO ( <i>Complex Adaptive System</i> ) .....	20
2.2 VOT ( <i>Voice Onset Time</i> ), O FENÔMENO INVESTIGADO.....	25
2.3 O TREINAMENTO PERCEPTUAL.....	27
2.4 ESTUDOS ANTERIORES DE NOSSO GRUPO DE PESQUISA ENVOLVENDO VOT E TREINAMENTO PERCEPTUAL.....	28
2.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO.....	50
<b>3 METODOLOGIA</b> .....	51
3.1 PARTICIPANTES.....	51
3.2 SESSÕES DE TREINAMENTO PERCEPTUAL.....	53
3.3 INSTRUMENTOS DE COLETA DE DADOS (PRÉ E PÓS-TESTES).....	55
3.4 ANÁLISE DE DADOS.....	57
<b>4 DESCRIÇÃO E DISCUSSÃO DOS RESULTADOS</b> .....	59
4.1 DESCRIÇÃO DOS DADOS DOS GRUPOS E ANÁLISE INFERENCIAL.....	59
4.2 ANÁLISES INDIVIDUAS.....	73
4.3 SOBRE A VARIABILIDADE NAS PRODUÇÕES DOS APRENDIZES: INDÍCIOS DE DESENVOLVIMENTO LINGUÍSTICO.....	91
4.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO.....	116
<b>5 CONSIDERAÇÕES FINAIS</b> .....	117
<b>REFERÊNCIAS</b> .....	121
<b>ANEXOS</b> .....	124
ANEXO A.....	125
ANEXO B.....	128
ANEXO C.....	131
ANEXO D.....	138

## 1 INTRODUÇÃO

As tarefas de treinamento perceptual têm como objetivo fazer com que o aprendiz de segunda língua (L2) possa desenvolver suas habilidades fonéticas, com foco na língua-alvo, em termos de percepção e/ou produção. No Brasil, estudos já mostram o papel do treinamento perceptual no desenvolvimento de tais habilidades (NOBRE-OLIVEIRA, 2007; BETTONI-TECHIO, 2008; REIS; NOBRE-OLIVEIRA, 2008; BRAWERMAN-ALBINI, 2012). Estudos mais recentes, como os de Alves e Luchini (2017), buscam verificar se os benefícios possibilitados pelo treinamento à percepção apresentam caráter duradouro, e se tais benefícios podem ser generalizados para a produção. Todos esses estudos, porém, contam com, no máximo, três momentos de teste (Pré-teste, Pós-teste Imediato e Pós-teste Postergado, geralmente um ou dois meses após o Pós-Teste Imediato).

Além disso, estudos como os de De Bot *et al.* (2007) mostram que o desenvolvimento de uma L2 se dá de forma dinâmica, assim como qualquer outro processo cognitivo, tendo a influência de diversos fatores - que não devem ser considerados isoladamente, justamente por não serem percebidos de forma isolada pelos aprendizes. Ou seja, o desenvolvimento linguístico vai ser um processo e não um produto. Lowie e Verspoor (2015) ressaltam a necessidade de um olhar atento à variabilidade das trajetórias individuais dos aprendizes. Isso se dá justamente pela concepção dinâmica do desenvolvimento de uma L2; se o desenvolvimento linguístico se dá por diversos fatores em um sistema e cada sistema reage diferentemente para cada estímulo, pensar em uma universalidade estanque para o desenvolvimento da linguagem é impraticável.

Podemos entender, então, que, no âmbito fonético-fonológico, a partir dessa concepção dinâmica de língua, diversas pistas acústicas poderão influenciar a percepção do aprendiz (ZIMMER; SILVEIRA; ALVES, 2009; ZIMMER; ALVES, 2012). Além disso, uma pista acústica pode não ter a mesma relevância funcional para a produção e percepção em uma determinada primeira língua (L1) (SCHWARTZHAUPT; ALVES; FONTES, 2015; ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015). O VOT (*Voice Onset Time*), definido como o intervalo de tempo entre a explosão de uma consoante plosiva e o início de pulsação glotal (anterior ou posterior à explosão) (LISKER; ABRAMSON, 1964; ABRAMSON; WHALEN, 2017), é uma dessas pistas, e será a pista utilizada em nosso trabalho. Há três padrões de

VOT encontrados nas línguas do mundo (LISKER; ABRAMSON, 1964; COHEN, 2004; REIS; NOBRE OLIVEIRA, 2008): VOT Negativo, também chamado pré-vozeamento, quando as cordas vocais vibram durante o processo de closura da consoante; VOT Zero, padrão no qual o início da vibração das pregas vocais ocorre de forma simultânea, ou quase simultânea, à soltura do segmento plosivo; e VOT Positivo, caracterizado por um intervalo de surdez maior entre a explosão e o segmento seguinte, também conhecido como aspiração. Os falantes de inglês utilizam predominantemente dois padrões para suas plosivas iniciais: o padrão Positivo para as plosivas surdas (/p/, /t/ e /k/), e o padrão Zero para as plosivas sonoras (/b/, /d/, /g/)<sup>1</sup>. Já os falantes brasileiros utilizam o padrão Zero para suas consoantes surdas e o padrão Negativo para as plosivas sonoras. Uma série de estudos, realizados em contexto brasileiro, mostrou que o VOT não é a principal pista utilizada por brasileiros para a distinção entre as plosivas surdas e sonoras do inglês (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; SCHWARTZHAUPT, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015). Um quarto padrão, chamado de VOT Zero Artificial, então, foi criado. O nome se dá pelo fato de este padrão ser originado do corte da aspiração de uma palavra em inglês iniciada com um plosiva surda. Desse modo, tal palavra mantém todas as pistas de uma palavra iniciada por uma plosiva surda, exceto a aspiração. Tal padrão foi usado para atestar o uso do VOT por falantes nativos do inglês norte-americano (SCHWARTZHAUPT; ALVES; FONTES, 2015). O estudo, então, concluiu que, diferentemente dos brasileiros, falantes nativos do inglês tendem a guiar-se pelo VOT para a distinção da sonoridade das plosivas.

Os brasileiros, por não se utilizarem do VOT para distinção entre surdez e sonoridade, tendem não a produzir a aspiração das plosivas surdas iniciais do inglês, o que pode implicar dificuldades de identificação de tais consoantes por parte de ouvintes nativos. Ressaltamos a necessidade, dessa forma, de que sejam pensadas estratégias e práticas didáticas e laboratoriais para chamar a atenção do aprendiz para a necessidade e o uso de tal pista acústica.

A partir desses estudos, então, o presente TCC objetiva avaliar a eficiência do treinamento perceptual (com estímulos com VOT Positivo e VOT Zero Artificial) sobre a percepção e a produção da aspiração de /p/, /t/ e /k/ iniciais do inglês por estudantes brasileiros. Objetivamos, também, discutir a contribuição do treinamento perceptual

---

<sup>1</sup> O padrão VOT Negativo também pode ser produzido variavelmente com o Zero, nas plosivas iniciais (LISKER; ABRAMSON, 1964).

nos índices de acuidade de percepção para os padrões de VOT testados, bem como verificar se o treinamento perceptual é capaz de contribuir com um aumento dos índices de produção de VOT das plosivas surdas iniciais. Além disso, também esperamos contribuir com um novo segmento de investigações que propõe a analisar as produções individuais dos participantes, discutindo os efeitos que o treinamento perceptual teria sobre o desenvolvimento da percepção e da produção de cada indivíduo.

O presente trabalho teve início durante o período em que o autor desenvolveu atividades de Iniciação Científica do Projeto de Pesquisa "Efeitos de treinamento perceptual e instrução explícita na aquisição da distinção entre plosivas surdas e sonoras iniciais do inglês (LE) por aprendizes argentinos e brasileiros: o papel de pistas acústicas primordiais na L1", sob a coordenação do Prof. Dr. Ubiratã Kickhöfel Alves. Ao longo de um mês, em 2015, as três primeiras etapas de coleta de dados foram realizadas: Pré-teste, Pós-teste e Pós-teste Postergado 1 (naquela época, reportado apenas como Pós-teste Postergado). Cada uma das etapas contou com a coleta de dados de percepção e produção por parte dos participantes, que estavam divididos em dois grupos de 10 membros: Experimental e Controle. Apenas o primeiro fez parte do treinamento perceptual. Os resultados desta primeira parte foram apresentados em Kampff e Alves (2016), que será resenhado mais profundamente no próximo capítulo, sobre Referencial Teórico.

Entretanto, permaneceu a dúvida se os efeitos verificados nos dois pós-testes se manteriam duradouros, sobretudo porque o último deles foi realizado apenas um mês após o período de treinamento. Foi este questionamento, então, que motivou a realização deste TCC. Dessa forma, para que tal questionamento fosse respondido, realizamos uma quarta etapa, durante o ano de 2018, com uma quarta coleta, três anos após a realização do Pós-teste Postergado 1. Ainda que, em termos metodológicos, reconheçamos que o período entre os dois Pós-Testes Postergados seja longo, dando margem para que diversos fatores não contabilizados pudessem afetar o desenvolvimento dos participantes, julgamos tal continuação interessante para a verificar se e quais mudanças estariam presentes nos índices dos grupos. Para a quarta etapa de coleta, contamos com 7 participantes no Grupo Experimental e 8 no Grupo Controle, cujas trajetórias serão analisadas ao longo das quatro etapas temporais.

Além desta quarta etapa, o trabalho será caracterizado por outro aspecto que o diferencia da análise então realizada em Kampff e Alves (2016) e dos demais trabalhos sobre treinamento realizadas em nosso grupo de pesquisa (RODRIGUES, 2015; ALVES; LUCHINI, 2017): a verificação das trajetórias individuais e da variabilidade dos aprendizes ao longo das etapas de coleta, contribuindo com um novo segmento de estudos que propõe a analisar a individualidade dos participantes e discutir os efeitos que o treinamento perceptual teria sobre ela.

Dessa forma, o trabalho apresenta como objetivo geral analisar, ao longo de quatro pontos de coleta e em uma janela de tempo de três anos, as trajetórias individuais e em grupo de alunos do curso de Letras que participaram, ou não, das sessões de treinamento perceptual, cujos resultados foram apresentados em congresso por Kampff e Alves (2016). Para se atingir o objetivo geral, pretendemos responder a três questões norteadoras:

- I. O que as abordagens inferenciais e descritivas podem revelar a respeito do desenvolvimento de cada um dos grupos analisados?
- II. O que uma verificação de dados individuais pode revelar sobre as trajetórias dos aprendizes?
- III. Que informações sobre a trajetória de desenvolvimento linguístico a variabilidade dos aprendizes em um dado momento temporal e a variação ao longo do tempo podem revelar?

Nosso trabalho será organizado da seguinte forma: o presente capítulo contará com a introdução; a seguir, o segundo capítulo trará o referencial teórico; o terceiro, a metodologia usada para a realização do trabalho; o quarto contará com a análise e discussão dos dados coletados; e o quinto, com as considerações finais.

Esperamos, com o presente trabalho, auxiliar na solidificação em três pontos teóricos: primeiro, agregar conhecimentos e reflexões aos estudos referentes às tarefas de treinamento perceptual, e seus benefícios; segundo, também agregar conhecimento às pesquisas brasileiras que tomam o desenvolvimento linguístico como SACs; terceiro, ajudar a desenvolver, no país, a possibilidade de análise individual dos dados dos participantes envolvidos em experimentos fonético-fonológicos.

## **2 REFERENCIAL TEÓRICO**

Nesta seção nos destinamos a trazer os pressupostos teóricos que permeiam e sustentam nosso trabalho. Primeiramente, traremos as seções 2.1 e 2.1.1, que tratam sobre a visão de língua sob o prisma da teoria da complexidade, seguida da reflexão, vital para o trabalho, da importância de um olhar para a variabilidade individual dos participantes. Após tal discussão, na seção 2.2, descreveremos nosso objeto de estudo e suas características, o VOT, enquanto na seção 2.3 falaremos do treinamento perceptual. Finalmente, na seção 2.4 traremos estudos anteriores, sobre as questões referentes ao treinamento perceptual e o desenvolvimento de padrões de VOT, realizados por pesquisadores brasileiros e pelos grupos de pesquisa nos quais nos inserimos.

### **2.1 CONCEPÇÃO DE LÍNGUA COMO UM SISTEMA ADAPTATIVO COMPLEXO (*Complex Adaptive System*)**

Diversos estudos (BECKNER *et al.*, 2009; DE BOT *et al.*, 2013; LARSEN-FREEMAN; CAMERON, 2008) apontam para uma perspectiva que vê a língua como um Sistema Adaptativo Complexo (SAC). Um "sistema" é definido como tal pelo fato de apresentar diversos componentes que interagem entre si - são sistemas, por exemplo, o sistema de transporte público, o sistema gramatical, uma célula viva e, obviamente, a língua. Além disso, é empregado o termo "complexo" pelo fato de apresentar um grande número de elementos, heterogêneos e de naturezas diversas, interagindo de diferentes formas - por exemplo, o próprio ecossistema é um sistema complexo, com diferentes agentes em interação. Finalmente, o termo "adaptativo" é empregado pelo fato de as mudanças presentes em uma área do sistema afetarem todo o restante, e pelos elementos do sistema estarem em constante adaptação em função das mudanças que naturalmente ocorrem entre tais elementos.

Além disso, outra característica dos SACs é o dinamismo. Ou seja, o papel do tempo nas mudanças de determinado sistema (LARSEN-FREEMAN, 1997; BECKNER *et al.*, 2009, ZIMMER; ALVES, 2012) é fundamental. Em um sistema complexo, a mudança está ocorrendo o tempo todo - como uma planta, por exemplo,

que está em crescimento contínuo. O mesmo acontece com o sistema linguístico (os subsistemas fonético, semântico, sintático, pragmático etc. funcionam em conjunto e influenciam uns aos outros). Com o passar do tempo, esses subsistemas vão sofrendo ações de outros subsistemas que vão influenciar suas mudanças e, conseqüentemente, influenciar o sistema como um todo. Por isso são relevantes estudos como o nosso, que levam o tempo em consideração, testando e analisando os participantes em um momento posterior ao da realização do experimento de Pré-teste e dos dois Pós-testes.

Ademais, é importante ressaltar que estudos como Beckner *et al.* (2009) trazem a noção de que o sistema linguístico também é emergente. Por ser um sistema, por ser complexo, e por ser dinâmico, ao passar do tempo, a estrutura da língua vai emergir através das interações dos agentes de tal sistema. Por exemplo, um idioleto é emergente do uso de língua individual em interações sociais com outros indivíduos de uma comunidade linguística, enquanto a mesma comunidade linguística vai fazer emergir sua fala como resultado da interação de idioletos individuais (BECKNER *et al.*, 2009).

Outros três aspectos da língua que são importantes para o nosso trabalho são a 'não-linearidade', o fato de os sistemas serem 'abertos' e, também, serem de 'domínio-geral' (*domain-general*) (BECKNER *et al.*, 2009; LARSEN-FREEMAN, 2015). A não-linearidade está ligada ao famoso "efeito borboleta", de acordo com o qual uma pequena mudança em uma parte do sistema pode levar a mudanças de níveis muito maiores em outra parte do mesmo. Em termos linguísticos, quando consideramos a língua como um sistema, não há certeza sobre como uma mudança em determinada parte deste sistema vai refletir no todo - por exemplo, como estudos que serão aprofundados ao desta seção, temos o fato de não haver uma linearidade nos resultados de percepção e produção dos participantes do Grupo Experimental, exposto ao treinamento, uma vez que a relação de causa-efeito não é direta, tampouco garantida. Já o sistema ser 'aberto' remete ao fato de que o mesmo continua a receber energias externas e se adapta às mesmas, e se opõe a sistemas ditos 'fechados', ou seja, que já não aceitam mais energias externas, tendendo a um estado de equilíbrio. Aplicando este princípio ao nosso trabalho, podemos ver os estímulos que os participantes receberam, através do treinamento (maiores explicações na próxima seção, referente à Metodologia) como a energia externa que pode afetar seus sistemas linguísticos. Ao analisarmos os dados, veremos se os sistemas aparentam

estar abertos, apresentando variação e variabilidade, ou fechados, já em estado de possível estabilização. Finalmente, os sistemas são ditos de domínio-geral pela não adoção da concepção dicotômica que diferencia "conhecimento puramente linguístico" do chamado "conhecimento extralinguístico". Dessa forma, o sistema é aberto a qualquer tipo de influência social ou pessoal do aprendiz, inclusive sentimentos e emoções.

No que diz respeito ao componente fonético-fonológico, um modelo de percepção e produção dos sons que pode ser lido a partir de uma noção dinâmica é o *Speech Learning Model*, de Flege (1995). O autor sugere que as categorias de sons de falantes bilíngues ocupam simultaneamente o mesmo espaço fonético-fonológico. Em outras palavras, ambas L1 e L2 vão coexistir num mesmo espaço, influenciando uma à outra. Quando estabelecemos um paralelo entre estas considerações e a teoria da língua como um SAC, podemos levar em conta a influência da L1 sobre a L2 como um outro fator pertinente para o desenvolvimento linguístico, que vai influenciar a forma como o participante vai sofrer as possíveis desestabilizações linguísticas.

O modelo sugerido por Flege (1995) conversa diretamente com outro fator importante encontrado nos estudos de Larsen-Freeman e Cameron (2008). Ele diz respeito aos estados chamados de 'atratores'. Pelo sistema estar sempre se auto-organizando e adaptando dinamicamente, ele será 'atraído' para o estado de equilíbrio. Dessa forma, por ambas as línguas coexistirem num mesmo espaço fonológico, ao longo do desenvolvimento da L2, o aprendiz será levado a produzir padrões atratores presentes em sua L1. Portanto, o processo de desenvolvimento da L2 deve ser voltado para o estabelecimento de novos estados atratores, presentes na língua-alvo, o que se mostra como um dos objetivos do trabalho de treinamento perceptual realizado<sup>2</sup>.

Em suma, de acordo com os autores supracitados (LARSEN-FREEMAN; CAMERON, 2008; BECKNER *et al.*, 2009; DE BOT *et al.*, 2013; LARSEN-FREEMAN,

---

<sup>2</sup> Outro modelo perceptual bastante em voga na área de Aquisição Fonético-Fonológica de L2 é o Modelo de Assimilação Perceptual de L2 (*L2 Perceptual Assimilation Model*), proposto por Best & Tyler (2007) e recentemente revisitado por Perozzo (2017). Entretanto, uma vez que o presente trabalho não tem por objetivo discutir as limitações dos modelos perceptuais da área, e que os estudos prévios sobre o tema, desenvolvidos em nosso grupo de investigação (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; RODRIGUES, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; KAMPFF; ALVES, 2016; ALVES; LUCHINI, 2016; ALVES; LUCHINI, 2017), foram estabelecidos a partir do Modelo de Aprendizagem da Fala (*Speech Learning Model*), optamos, neste trabalho, por seguir o modelo de Flege, com base psicoacústica, para o entendimento da percepção dos padrões de *Voice Onset Time*, dado que o primitivo com o qual estamos operando é efetivamente acústico.

2015; entre outros) o desenvolvimento linguístico é complexo, por apresentar variados fatores que interagem entre si; adaptativo, pois a partir da interação esses sistemas se organizam e reorganizam; e dinâmico, por causa da mudança ao longo do tempo causada por tais interações.

### 2.1.1 O papel da variabilidade

Em Lowie e Verspoor (2015), os autores reforçam a ideia de que o sistema linguístico é, de fato, um sistema complexo, adaptativo, não-linear e dinâmico; logo, o desenvolvimento de uma língua torna-se um processo, e não um produto ou resultado final. Porém, cada sistema, por ser formado por diversos fatores, acaba por tornar-se um sistema único. O mesmo acontece com cada aprendiz de uma L2. O desenvolvimento dessa L2 vai se dar através de todas as interações passadas e presentes que vão ocorrer dentro do sistema linguístico do falante. Portanto, seriam questionáveis estudos que tentam "enquadrar" o desenvolvimento dos aprendizes em estágios universais. Por mais que uma média de grupo seja significativa, ela não abarca todas as possíveis causas da emergência do sistema linguístico de um falante. Por isso, os autores afirmam que o foco de observação deve ser outro: a variabilidade.

A variabilidade vai ser caracterizada, então, pela variação individual de cada participante. Mais ainda, ela vai ser o "atestado" de que o aprendiz está em desenvolvimento, pois a mesma vai ser uma propriedade inerente de um sistema que se auto-organiza, como o linguístico. Os autores trazem, então, que uma variabilidade maior vai acontecer em estágios iniciais de desenvolvimento, enquanto, à medida em que o sistema se organiza, essa variabilidade diminui. Essa relação entre variabilidade e desenvolvimento vai se dar justamente por ela indicar que o sistema do aprendiz está desestabilizado, ou seja, recebendo *inputs* e energias externas que o modifiquem, e tentando organizá-los dentro desse sistema. Em outras palavras, "a variabilidade não é um simples subproduto do desenvolvimento, mas uma força motriz e um impulsionador de mudança"<sup>3</sup> (LOWIE; VERSPOOR, 2015 p. 76).

Finalmente, os autores trazem, também, o reforço de que, pela dinamicidade do sistema, as formas de mensurar tal desenvolvimento devem envolver estudos longitudinais, medindo o estado do sistema linguístico do aprendiz em diferentes momentos temporais. Além disso, eles também recomendam a professores de ASL

---

<sup>3</sup> Tradução do autor;

(Aquisição de Segunda Língua) levar em consideração que a variabilidade vai estar presente na sala de aula, com os alunos variando entre ocorrências mais próximas e mais distantes da língua-alvo, e isso não pode ser considerado como um retrocesso.

Já seu recente artigo, intitulado *Individual Differences and the Ergodicity Problem*<sup>4</sup> (2018), Lowie e Verspoor aprofundam a necessidade do olhar para as diferenças individuais dos aprendizes de L2, ainda sob o espectro da língua enquanto Sistema Complexo. Segundo os autores, pelo fato de o sistema linguístico se tratar de um grande número de fatores interagindo, e também pelo fato de que tais fatores são diferentes para cada falante, o processo de desenvolvimento de uma L2 deve ser caracterizado individualmente. Por esse motivo, a língua não pode ser considerada a partir do princípio da ergodicidade, que parte de duas premissas: a primeira é de que um possível grupo estudado deve ser homogêneo; a segunda é de que os dados devem ser estáveis, com médias e variações sendo imutáveis ao longo das medições. Entretanto, sabemos que, sendo a língua um sistema complexo e dinâmico, tais premissas não poderiam ser atingidas. O experimento trazido pelos autores, no qual os participantes foram cuidadosamente selecionados a fim de alcançar o máximo de homogeneidade possível, mostrou que mesmo aprendizes que, a uma primeira vista, poderiam vir a ser considerados extremamente similares em termos de diferenças individuais, mostram trajetórias diferentes através do tempo. Além disso, através dos resultados, verificados ao longo do tempo, foi possível perceber que aqueles que demonstraram maior grau de variabilidade também demonstraram maior proficiência. Finalmente, os autores indicam que estudos de grupo e estudos individuais são complementares, sendo os estudos de grupo necessários para os fatores individuais que tenham papéis de destaque no desenvolvimento da L2 e estudos longitudinais, de caso, necessários para entender o processo individual do desenvolvimento do aprendiz.

Com base nas afirmações feitas ao longo desta seção, no presente trabalho, utilizaremos a perspectiva dinâmica de língua, somada à análise das trajetórias individuais dos participantes, para verificar a influência do treinamento perceptual, em grupo e individualmente, em relação à distinção de plosivas surdas e sonoras através do VOT, apresentado a seguir.

---

<sup>4</sup> "Diferenças individuais e o problema da ergodicidade", tradução do autor.

## 2.2 VOICE ONSET TIME (VOT), O FENÔMENO INVESTIGADO

O *Voice Onset Time* (VOT), fenômeno investigado em nosso trabalho, é definido como o intervalo de tempo entre a explosão de uma consoante plosiva e o início de pulsação glotal (anterior ou posterior à explosão) (LISKER; ABRAMSON, 1964; ABRAMSON; WHALEN, 2017). Os três principais padrões de VOT encontrados nas línguas naturais (LISKER; ABRAMSON, 1964; COHEN, 2004; REIS; NOBRE OLIVEIRA, 2008) são:

- a) VOT Negativo (que varia entre -125 ms e - 75 ms), também chamado pré-vozeamento, quando as pregas vocais vibram durante o processo de closure da consoante;
- b) VOT Zero (que pode variar de 0 ms a +35 ms), padrão no qual o início da vibração das pregas vocais ocorre de forma simultânea, ou quase simultânea, à soltura do segmento plosivo; e
- c) VOT Positivo (que pode variar de 35 ms a 100 ms), também conhecido como ‘aspiração’, que é caracterizado por um momento de surdez mais longo entre a soltura da plosiva e a vibração das pregas vocais.

Conforme já afirmado na Introdução, os falantes nativos de inglês utilizam sobretudo dois padrões para suas plosivas iniciais: o padrão VOT Positivo para as plosivas surdas (/p/, /t/ e /k/)<sup>5</sup>, e o padrão Zero para as plosivas sonoras (/b/, /d/, /g/)<sup>6</sup>. Já os falantes brasileiros utilizam o padrão Zero para suas consoantes surdas e o padrão Negativo para as plosivas sonoras. Um dos desafios desses aprendizes é realizar os segmentos plosivos surdos iniciais do inglês com aspiração<sup>7</sup>, dado, inclusive, que a não realização da mesma exerce efeitos na identificação da sonoridade de tais consoantes por parte de ouvintes nativos do inglês; tais efeitos

---

<sup>5</sup> Médias de aspiração do VOT Positivo: 58 ms para [p<sup>h</sup>], 70 ms para [t<sup>h</sup>] e 80 ms para [k<sup>h</sup>] (LISKER; ABRAMSON, 1964);

<sup>6</sup> Conforme já afirmado na Introdução, o padrão VOT Negativo também pode ser produzido variavelmente com o Zero, nas plosivas iniciais do inglês (LISKER; ABRAMSON, 1964).

<sup>7</sup> É importante ressaltar que trabalhos desenvolvidos em nosso país (M. ALVES, 2015; ZIMMER; ALVES, 2015) têm mostrado valores de VOT mais altos, no português, para a plosiva /k/, que caracterizam uma “semi-aspiração” (M. ALVES, 2015), variando entre 50 e 60 ms. Tais valores, porém, não se igualam aos encontrados no inglês (L2).

podem ser considerados negativos, uma vez que podem afetar a inteligibilidade referente à produção oral do falante.

Em uma série de trabalhos sobre percepção dos padrões de VOT realizados em nosso país (ALVES; SCHWARTZHAUPT; BARATZ, 2011; ALVES; MOTTA, 2013; ALVES; ZIMMER, 2015; SCHWARTZHAUPT, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015), foi verificado o *status* do VOT como pista acústica fundamental para as distinções funcionais de sonoridade tanto no português quanto no inglês. Esta questão de investigação se mostra importante uma vez que, de acordo com os referidos pesquisadores, havia a possibilidade de os aprendizes brasileiros não produzirem a aspiração das plosivas iniciais pelo fato de tal pista não ser tomada como fundamental por estes aprendizes nas distinções perceptuais entre segmentos surdos e sonoros. Para a realização de tal verificação, foi criado um quarto padrão de VOT, chamado VOT Zero Artificial. Com o auxílio do *software Praat* (BOERSMA; WEENINK, 2011), a aspiração do padrão Positivo foi cortada, de forma a que se assemelhasse, em termos de VOT, ao padrão Zero Natural, criando, assim, um padrão "híbrido", pois continha todas as características de uma plosiva surda do inglês, menos a aspiração inicial. Os resultados dos estudos mostraram que os falantes nativos de inglês identificavam o padrão Zero Artificial da mesma forma como o padrão Zero Natural (ou seja, como instâncias de segmentos sonoros), concluindo que o VOT, na língua inglesa, tem um papel fundamental para a distinção de plosivas surdas e sonoras (SCHWARTZHAUPT; ALVES; FONTES, 2015). Por outro lado, estudos com o mesmo padrão, aplicados em aprendizes brasileiros de língua inglesa (ALVES; SCHWARTZHAUPT; BARATZ, 2011; ALVES; ZIMMER, 2013, 2015; ALVES; MOTTA, 2014) e argentinos (ALVES; LUCHINI, 2016), chegaram à conclusão de que o VOT não é a principal pista acústica levada em consideração para a distinção entre plosivas, devido o fato de o padrão Zero Artificial ser predominantemente identificado, por tais ouvintes estrangeiros, como instâncias de consoantes surdas; tal verificação deixa claro que, diferentemente dos norte-americanos, os ouvintes em questão priorizaram outras pistas acústicas que não o VOT. Outras pistas acústicas, dessa forma, foram sugeridas como majoritárias para a distinção entre plosivas surdas e sonoras por tais aprendizes de inglês, como a intensidade da explosão do segmento plosivo e a F0 da vogal seguinte. Além disso, os trabalhos supracitados também mostram altos índices de acuidade em relação à identificação dos padrões VOT Negativo e VOT Positivo como instâncias de segmentos sonoros e surdos,

respectivamente, o que resulta em um maior foco, em nosso trabalho, na verificação dos índices de identificação dos padrões Zero Natural e Zero Artificial.

### 2.3 O TREINAMENTO PERCEPTUAL

Conforme já explicado na Introdução, o objetivo das tarefas de treinamento perceptual é fazer com que os aprendizes de uma L2 possam desenvolver suas habilidades, com foco na língua-alvo, em termos de percepção e/ou produção. Alguns estudos desenvolvidos no Brasil, apesar de ainda serem escassos, já mostram o papel do treinamento perceptual em níveis segmentais e suprasegmentais (NOBRE-OLIVEIRA, 2007; BETTONI-TECHIO, 2008; BRAWERMAN-ALBINI, 2012; REIS; NOBRE-OLIVEIRA, 2007; RODRIGUES, 2015).

Conforme trazido por Brawerman-Albini (2012), há diferentes tarefas de percepção que podem ser realizadas em um treinamento perceptual. As duas principais são 'discriminação' e 'identificação'. Em sua tese, Perozzo (2017) define as tarefas de discriminação como a "comparação de dois ou mais itens fônicos não nativos em um único estímulo auditivo com base em opções de múltipla escolha" (PEROZZO, 2017, p. 186). Em outras palavras, o participante determina, entre os estímulos a que foi exposto, se eles correspondem ao mesmo segmento ou não (no caso de serem dois os estímulos). Essas tarefas podem ter os seguintes paradigmas: AX (com dois estímulos), AXB, ABX, XAB (com três estímulos), devendo o participante responder se o elemento X é igual a outro; além disso, os três últimos paradigmas juntos num mesmo teste formam a tarefa conhecida como *Oddity Discrimination*, no qual o participante deve assinalar o elemento diferente. Já nas tarefas de identificação, o autor as define como "rotulação de um item fônico não nativo em um único estímulo auditivo com base em opções de múltipla escolha" (PEROZZO, 2017, p.185); ou seja, há uma escolha entre as opções de respostas dadas após a apresentação de um determinado estímulo.

Por sua vez, segundo o autor, uma tarefa de identificação vai necessariamente envolver um elemento fônico não nativo e levar o participante a responder de acordo com o seu conhecimento fônico também não nativo; essa é a tarefa empregada nos estudos de nosso grupo de pesquisa e no presente trabalho, como ficará claro nas seções a seguir. Além disso, conforme atesta Brawerman-Albini (2012), as tarefas de identificação mostraram-se mais eficientes que as de discriminação durante o

treinamento perceptual, por apresentarem melhoras significativas na percepção. Tal fato também motivou o uso de testes de identificação neste trabalho.

Outro ponto importante diz respeito a resultados de estudos como os de Nobre-Oliveira (2007), Rodrigues (2015), Kampff e Alves (2016), e Alves e Luchini (2017)<sup>8</sup>, que investigam a generalização dos possíveis efeitos benéficos do treinamento na percepção para a produção. Até o presente momento, tal generalização à produção mostrou-se problemática nestes trabalhos. Esperamos, com o presente trabalho, contribuir positivamente para estudos que se dediquem a tal verificação.

## **2.4 ESTUDOS ANTERIORES DE NOSSO GRUPO DE PESQUISA ENVOLVENDO VOT E TREINAMENTO PERCEPTUAL**

Nesta seção, dedicamo-nos a recapitular os principais estudos de nosso grupo de pesquisa que trabalharam com percepção de padrões de VOT ou atividades de treinamento perceptual e seus efeitos na percepção, ou percepção e produção. Tal retomada cronológica eclodirá em nosso trabalho, fruto da construção feita ao longo dos anos, mostrada a seguir. A seção será dividida em subseções, em ordem cronológica. É importante ressaltar que todos os trabalhos descritos nas subseções a seguir tomam o sistema linguístico como um SAC, que teve suas características descritas no começo deste capítulo; portanto, não retomaremos essa parte do referencial teórico na descrição de dos estudos a seguir.

### **2.4.1 Alves e Motta (2014)**

Em 2014, com os objetivos de (i) verificar se a diferença de proficiência entre participantes implicaria em diferentes respostas para identificação e discriminação dos padrões de VOT investigados; e (ii) se, em algum dos níveis de proficiência, os aprendizes fariam uso do VOT como pista principal para distinção entre plosivas surdas e sonoras do inglês, Alves e Motta realizaram um dos primeiros estudos perceptuais de nosso grupo de pesquisa sobre padrões de VOT com aprendizes brasileiros.

---

<sup>8</sup> Os três últimos serão melhor resenhados nas próximas subseções.

Baseando-se nos resultados de Alves *et al.* (2011) e Alves e Zimmer (2012), que sugeriram que os aprendizes, apesar de discriminarem plosivas surdas e sonoras do inglês, e identificarem o VOT Positivo como /p, t, k/ e os padrões de VOT Zero, ou Negativo, como /b, d, g/, talvez não utilizassem o VOT para fazer tal distinção, foram criadas tarefas de discriminação e identificação com quatro padrões de VOT. Os três primeiros padrões, encontrados nas línguas naturais, são os VOTs Positivo, Negativo e Zero; o quarto padrão foi chamado de Manipulado (posteriormente chamado de VOT Zero Artificial), por ser um estímulo com uma plosiva surda inicial do inglês, ou seja, aspirada, que teve sua aspiração cortada, tornando-se, portanto, uma plosiva com VOT Zero (característica das plosivas sonoras do inglês), porém mantendo todas as demais características acústicas de uma plosiva surda da L2.

As premissas seguidas pelos autores eram de que a identificação do VOT Zero não-manipulado como surdo indicaria que os aprendizes estariam se baseando na L1, já que plosivas surdas em português apresentam tal padrão, como ressaltado anteriormente. Já a identificação do mesmo padrão como sonoro indicaria duas possibilidades: a primeira era que os aprendizes já estariam associando a ausência de VOT Positivo como caracterização de plosivas sonoras, de acordo com a língua-alvo; e a segunda seria de que o VOT não é a pista principal usada pelos aprendizes. Neste caso, nem a L1 nem a L2 dos aprendizes estariam sendo seguidas, indicando que os mesmos possivelmente estariam seguindo outras pistas acústicas além do VOT, como F0 ou intensidade da explosão. Em relação ao padrão Manipulado, a identificação da plosiva como surda também indicaria duas possibilidades: a primeira é que os aprendizes seguiriam o padrão de VOT da L1, já que padrões de VOT de curta duração indicam plosivas surdas no português brasileiro. Neste caso, o padrão VOT Zero que não foi manipulado também seria identificado como surda, uma vez que ambos os padrões apresentam o mesmo comportamento em relação à aspiração mais curta; a segunda é de que os aprendizes baseiam sua identificação em outras pistas acústicas, já que o padrão Manipulado apresenta todas as outras pistas características de uma plosiva surda do inglês, exceto pela aspiração. Neste caso, a identificação entre o padrão Manipulado e não-manipulado seriam diferentes. Já a identificação desse padrão manipulado como sonora indicaria que os aprendizes já haveriam desenvolvido a distinção entre plosivas de acordo com os padrões da língua-alvo.

Para tal experimento, 20 alunos de graduação foram divididos em dois grupos. O primeiro, formado por 10 participantes de níveis de proficiência A1 e A2, era o de iniciantes; o segundo, formado pelos outros 10, com níveis de proficiência C1 e C2, era o de nível avançado<sup>9</sup>. Ambos os grupos participaram tanto de uma tarefa de discriminação quanto de uma tarefa de identificação.

As tarefas tiveram seus estímulos gravados por seis falantes (três homens e três mulheres) nativos de inglês norte-americano que residiam no sul do Brasil por menos de três meses. Em um estúdio profissional, os falantes gravaram três pares mínimos: *bit - pit*; *dick - tick*; e *gill - kill*. Cada par apresentava um ponto de articulação diferente e era seguido da vogal /i/, que ocasiona maiores valores de VOT (YAVAŞ; WILDERMUTH, 2006; YAVAŞ, 2008). Cada estímulo foi gravado três vezes, de modo que os estímulos que melhor representassem a língua-alvo fossem selecionados. Assim, os estímulos produzidos apresentavam três padrões de VOT: o Positivo, para plosivas surdas (/p, t, k/), e Zero e/ou Negativo, para plosivas sonoras (/b, d, g/). Além disso, após produções positivas terem suas aspirações reduzidas, ao ponto de tornarem-se um padrão Zero, através do *software* Praat (BOERSMA; WEENINK, 2011), o padrão Manipulado foi adicionado às tarefas. Logo, ambas as tarefas de identificação e discriminação contaram com os quatro padrões de VOT citados, sendo desenvolvidas e aplicadas através do *Praat*.

A tarefa de identificação consistia em que os aprendizes ouvissem um dos estímulos descritos acima de cada vez, assinalando qual a consoante inicial que haviam ouvido - /p/, /b/, /t/, /d/, /k/ e /g/. No começo da tarefa, eram oferecidas três rodadas iniciais para que os participantes se acostumassem com a tarefa. Após isso, estímulos com os quatro padrões de VOT mencionados apareciam em ordem aleatória. Na tarefa, que contava com 24 estímulos para serem identificados, cada um dos quatro padrões foi apresentado em 6 *tokens* (dois para cada ponto de articulação, com a mesma palavra produzida por um falante diferente, como [b]it, [b]it, [d]ick, [d]ick, e [g]ill, [g]ill para o padrão Negativo, por exemplo). As tarefas foram realizadas individualmente, com o uso de fones de ouvido, em uma sala silenciosa.

Para esta tarefa, três hipóteses foram levantadas: a primeira era a de que a identificação dos padrões VOT Negativo e VOT Positivo não seria diferente entre os

---

<sup>9</sup> Ambos os grupos foram testados através do *Oxford Online Placement Test*, que categoriza os participantes de A1 a C2, sendo o primeiro o mínimo e o segundo máximo, em questão de proficiência, de acordo com o Marco Comum Europeu (*Common European Framework*).

dois grupos de proficiência, pois mesmo que os aprendizes iniciantes não se guiassem pelo VOT na identificação, os mesmos se guiariam através de pistas adicionais para fazer esta identificação; a segunda era de que ambos os grupos identificariam o padrão Zero como instância de plosivas sonoras, uma vez que os participantes do grupo de iniciantes se guiariam pela intensidade da explosão e os do grupo avançado já se guiariam através dos padrões da L2; a terceira inferia que haveria uma diferença entre a identificação dos dois grupos em relação ao padrão Manipulado - os participantes iniciantes seguiriam o padrão de VOT da L1, identificando o padrão como surdo, e os avançados, por sua experiência e proficiência com a L2, identificariam o segmento como sonoro, seguindo o padrão de identificação da língua-alvo.

Já a tarefa de discriminação consistia em uma tarefa AXB. Os participantes ouviam três estímulos e recebiam múltiplas escolhas para responderem se a consoante inicial do segundo estímulo era igual à do primeiro (por exemplo, *pit - pit - bit*), à do terceiro (por exemplo, *pit - bit - bit*) ou se as três consoantes eram iguais (por exemplo, *pit - pit - pit*). Três tipos de contraste eram testados neste tipo de tarefa: VOT Negativo vs. Zero Manipulado (12 questões - quatro para cada ponto de articulação); VOT Negativo vs. VOT Positivo (12 questões); VOT Zero Manipulado vs. VOT Positivo (12 questões), totalizando 36 questões na tarefa. Contrastos como VOT Zero vs. VOT Negativo e VOT Zero vs. VOT Zero Manipulado não foram incluídos, pois Alves *et al.* (2011) já havia mostrado que, nestes casos, os aprendizes tendiam a discriminar o segundo, mas não o primeiro, desses pares. Além disso, para testar a atenção dos participantes à tarefa, as sequências com as mesmas consoantes iniciais eram apresentadas (*pit - pit - pit*). Novamente, as tarefas foram realizadas individualmente, com o uso de fones de ouvido e em uma sala silenciosa.

Para esta tarefa, também foram levantadas três hipóteses: a primeira era de que ambos os grupos discriminaríamos os padrões Negativo e Positivo, mesmo que através de pistas diferentes, apenas esperando que os do grupo avançado discriminassem as sonoridades das consoantes através dos padrões da L2; a segunda era de que a discriminação entre o padrão Negativo e o padrão Zero Manipulado seria diferente entre os dois grupos, dado que os iniciantes, guiados pela L1, considerariam o VOT Negativo como sonoro e o Zero Manipulado como surdo, e os avançados, sendo guiados pela L2, não discriminaríamos os dois padrões com VOT curto; por sua vez, a terceira previa que, entre os padrões Zero Manipulado e Positivo, haveria diferença de discriminação entre os dois grupos, com o grupo iniciante

apresentando níveis menores de discriminação, sendo guiados pela L1 e não tendo o VOT como principal pista acústica, e os avançados, guiados pela L2, usando a diferença de VOT curto e a presença de aspiração para a discriminação.

Os resultados obtidos pelo estudo foram diversos. Primeiramente, na tarefa de identificação, ambos os grupos identificaram os VOTs Negativo e Positivo como sonoro (98,33% no grupo iniciante e 100% no grupo avançado) e surdo (83,33% no grupo iniciante e 95% no grupo avançado), respectivamente, confirmando a primeira hipótese em relação à tarefa, de que os dois níveis de proficiência conseguiriam identificar consoantes aspiradas como surdas, mesmo não sendo a aspiração produzida na L1.

O padrão Zero, então, foi identificado como sonoro por ambos os grupos (68,33% no grupo iniciante e 73,32% no grupo avançado), confirmando a segunda hipótese desta tarefa, de que não haveria diferença de identificação entre os dois grupos. Porém, há duas possibilidades de explicação teórica para o que levou esse padrão a ser identificado como sonoro: (i) os aprendizes ainda não usam o VOT para fazer a identificação dessas consoantes como sonoras, mas sim outras pistas acústicas; (ii) os aprendizes já utilizam o VOT e aprenderam os padrões da língua-alvo.

Já no padrão Zero Manipulado, não houve diferenças inferenciais na identificação entre surdas e sonoras em ambos os grupos. O grupo avançado, inclusive, identificou o padrão como surdo mais vezes do que o grupo iniciante o fez, sendo 70% de identificação para surdo no grupo avançado e 58,32% no iniciante. A terceira hipótese desta tarefa, então, não foi confirmada. Novamente, a baixa identificação do padrão como sonoro levanta duas possibilidades de explicação teórica: a primeira era de que aprendizes avançados não utilizam o VOT como pista principal para a distinção de sonoridade entre plosivas, e ao invés disso utilizam outras pistas, como força de explosão; e a segunda era de que, mesmo guiados pela pista do VOT, os aprendizes ainda não haviam aprendido esses padrões para a L2, sendo guiados pelo padrão de VOT da L1, que identifica o padrão Zero como surdo.

Em suma, em relação à tarefa de identificação, os resultados indicam que, de fato, o VOT não vai ser a principal pista acústica usada pelos aprendizes brasileiros, independentemente do nível de proficiência, para a distinção entre plosivas surdas e sonoras do inglês. Sendo assim, os autores sugerem que força de explosão poderia ser a pista principal que guiasse os aprendizes, já que o padrão Manipulado havia sido

identificado como surdo; o padrão Zero, com uma explosão mais fraca, semelhante à do português brasileiro, havia sido identificado como sonoro; o padrão Negativo, também caracterizado com uma explosão fraca, mas com pré-vozeamento, havia sido identificado como sonoro; e o padrão Positivo identificado como surdo, com uma explosão mais forte.

Em relação aos resultados da tarefa de discriminação, a primeira hipótese foi confirmada parcialmente. Como previsto, a discriminação entre os padrões Negativo e Positivo teve altos níveis de acurácia (sendo 84,16% no grupo iniciante e 87,5% no grupo avançado). Porém, ao contrastar com os resultados obtidos na tarefa de identificação com os índices de acurácia da tarefa de discriminação, os autores corroboram a sugestão de que os participantes não usam o VOT para a distinção de sonoridade das plosivas, independente do nível de proficiência.

Já em relação à discriminação entre os padrões Negativo e Manipulado, os dados descritivos mostraram uma tendência dos dois grupos a discriminarem os padrões (sendo 62,5% de discriminação para o grupo iniciante e 66,66% para o grupo avançado), apesar de os resultados inferenciais não apresentarem significância. A acurácia aproximada entre ambos os grupos rejeita a segunda hipótese, referente a esta tarefa, que afirmava que os aprendizes avançados, baseando-se no padrão VOT da língua-alvo, teriam maior dificuldade para discriminar os dois padrões. Os índices sugeriram que os aprendizes, independente do nível de proficiência, ou baseavam-se nos padrões da L1, ou, ainda, que outras pistas acústicas estavam sendo tomadas como prioritárias.

A última hipótese, que previa uma diferença significativa entre os grupos na discriminação dos padrões Manipulado e Positivo, também não foi confirmada, uma vez que o índice de não-discriminação do grupo iniciante foi de 48,33%, e o do grupo avançado, de 58,33%. Os participantes de nível iniciante, segundo os autores, não iriam discriminar entre os dois padrões citados, pois estariam seguindo os padrões da L1 e fazendo a identificação através da força de explosão (que é semelhante nos dois padrões); já os participantes avançados discriminaram os padrões pois, por serem mais proficientes, já seguiriam os padrões de discriminação da L2. Como os resultados trazidos mostram, ambos os grupos de proficiência mostraram a mesma tendência em suas respostas.

Em suma, somando as conclusões dos resultados de ambas as tarefas, reforçou-se a afirmação de que aprendizes brasileiros, independentemente do nível

de proficiência, não seguem o VOT como pista principal para distinção entre plosivas surdas e sonoras. Além disso, os autores abriram maiores caminhos para estudos referentes ao desenvolvimento dos padrões de VOT da L2, tais como estudos que continuassem a investigação acerca de qual seria a principal pista acústica usada por brasileiros para a distinção entre a sonoridade das plosivas, assim como a conexão entre a percepção de tais pistas e sua influência na produção.

#### 2.4.2 Alves e Zimmer (2015)

Já em 2015, Alves e Zimmer dão continuidade aos estudos referentes à percepção e produção dos padrões de VOT do inglês por brasileiros, com alguns pontos semelhantes ao de Alves e Motta (2014). Dentre seus objetivos, estavam: (i) avaliar se aprendizes, em diferentes níveis de proficiência, se mostrariam capazes de distinguir os diferentes padrões de VOT produzidos por falantes nativos de inglês; (ii) investigar se os aprendizes destes dois grupos produzem valores de VOT semelhantes aos encontrados na língua-alvo, no caso, o inglês; (iii) discutir o papel do VOT como pista acústica para aprendizes brasileiros distinguirem plosivas surdas e sonoras do inglês.

Para o estudo, os autores contaram com 34 participantes naturais da cidade de Porto Alegre - RS. 24 deles foram selecionados para o grupo de nível básico, ou iniciante, e 10 para o grupo de nível avançado<sup>10</sup>. Até a data da realização das tarefas, nenhum dos participantes havia recebido instrução formal sobre pronúncia. Além disso, todos participaram das tarefas de percepção, divididas entre identificação e discriminação, e de produção.

Para as tarefas de percepção, foram usadas as mesmas tarefas do estudo de Alves e Motta (2014), reportado na subseção anterior. Ou seja, foi realizada uma tarefa de identificação, que consistia em fazer com que o participante ouvisse aleatoriamente um estímulo com dos padrões VOT (Negativo, Positivo, Zero Natural e Zero Artificial, antes conhecido como Manipulado) e assinalasse, entre múltiplas escolhas, qual a consoante inicial da palavra. Também foi aplicada a mesma tarefa de discriminação

---

<sup>10</sup> Assim como em Alves e Motta (2014), o teste de proficiência utilizado foi o *Oxford Online Placement Test*, que dividiu os participantes A1 e A2, e C1 e C2 (de acordo com o Marco Comum Europeu, ou *Common European Framework*) entre o grupo de nível básico e de nível avançado, respectivamente.

do estudo anterior, no formato AXB, contrastando o padrão VOT Negativo vs. VOT Zero Artificial, VOT Negativo vs. VOT Positivo, e VOT Zero Artificial vs. VOT Positivo.

O teste de produção, por sua vez, era dividido em duas partes. A primeira era a produção de palavras em português brasileiro. Os aprendizes eram solicitados a ler um arquivo no formato *PowerPoint* (ppt), com palavras dissilábicas, cada uma em um *slide* diferente e de continuidade automática, que iniciassem com as consoantes /p/, /b/, /k/ e /g/. Palavras iniciadas com /t/ e /d/ não foram incluídas na tarefa porque os aprendizes apresentam palatalização das mesmas antes da vogal /i/, que era a vogal presente após todas as consoantes das palavras produzidas no experimento. O teste contou com quatro repetições de cada palavra-alvo, além de oito palavras distratoras iniciadas por segmentos não plosivos. A segunda tarefa, em inglês, seguindo a mesma linha da tarefa de português, com as palavras passando automaticamente no arquivo *PowerPoint*, contava com palavras-alvo iniciadas pelas plosivas /p/, /t/, /k/, /b/, /d/ e /g/, também seguidos por uma vogal frontal alta. Cada palavra foi produzida seis vezes, além de também conter palavras distratoras.

O estudo trouxe oito hipóteses, todas partindo da premissa de que alunos de nível mais básico não utilizariam o VOT como principal pista acústica, enquanto os mais avançados, sim. As três primeiras hipóteses, referentes à tarefa de identificação, eram: 1 - não haveria diferença significativa entre os dois grupos na identificação dos padrões VOT Negativo e VOT Positivo, uma vez que os aprendizes básicos identificariam corretamente cada padrão através de outras pistas acústicas que não o VOT, como intensidade da explosão, e os aprendizes avançados fariam a mesma identificação através do uso do VOT como pista principal; 2 - não haveria diferença significativa entre os dois grupos na identificação do VOT Zero Natural como uma consoante surda, uma vez que a utilização da explosão seria a principal pista dos aprendizes de nível básico e os de nível avançado estariam, novamente, utilizando-se do VOT para tal identificação; e 3 - haveria uma diferença significativa entre os dois grupos na identificação do padrão VOT Zero Artificial, uma vez que os aprendizes avançados, guiando-se pelos padrões da L2, utilizariam a ausência de VOT para identificar o padrão como uma plosiva sonora, enquanto os aprendizes avançados, seguindo outras pistas acústicas, como força de explosão, continuariam a identificar tal padrão como surda.

As outras três hipóteses, referentes à tarefa de discriminação, eram: 4 - não haveria diferença significativa entre ambos os grupos de proficiência na discriminação

dos padrões VOT Negativo e VOT Positivo, uma vez que ambos os grupos seriam capazes de discriminar entre os padrões, mesmo que utilizando-se de diferentes pistas acústicas; 5 - haveria uma diferença significativa entre os dois grupos na discriminação entre os padrões VOT Negativo e VOT Zero Artificial, uma vez que os aprendizes de nível básico seriam guiados por pistas que não o VOT e as diferenciariam nos dois padrões, enquanto o nível avançado, pela curta duração dos padrões, não conseguiria diferenciá-los; e 6 - haveria uma diferença significativa entre os dois grupos no que diz respeito à discriminação dos padrões VOT Zero Artificial e VOT Positivo, uma vez que, por não tomarem o VOT como pista principal, o grupo mais básico não os diferenciaria, enquanto o grupo avançado, guiado pelo VOT, diferenciaria entre a ausência e a presença de aspiração.

Finalmente, as últimas duas hipóteses são referentes à produção: 7 - haveria uma diferença significativa, em cada um dos grupos, considerados separadamente, entre as durações de VOT de /p/ e /k/ encontradas em português e inglês, uma vez que, ainda que os aprendizes não tenham atingido valores de VOT semelhantes aos encontrados na língua inglesa, os mesmos já fazem uso parcial da pista para assinalar as plosivas surdas do inglês; e 8 - em relação à produção do VOT Positivo, não haveria diferenças significativas nas durações de tal padrão, uma vez que resultados prévios da literatura (ALVES; SCHWARTZHAUPT; BARATZ, 2011) mostravam que aprendizes, independentemente do nível de proficiência, não produzem valores de acordo com o padrão nativo da língua-alvo.

Nos resultados do estudo, então, a Hipótese 1 foi confirmada, pois ambos os grupos fizeram a identificação do VOT Negativo e Positivo sem diferenças entre si, independentemente da pista majoritária seguida, com 99,31% do Grupo Básico e 96,67% do Grupo Avançado atribuindo *status* sonoro ao Negativo e 91,67% e 98,33%, nos respectivos grupos, atribuindo *status* surdo para o VOT Positivo. A Hipótese 2 também foi confirmada, com ambos os grupos atribuindo *status* sonoro para o VOT Zero Natural (69,1% para o Grupo Básico e 71,67% para o Grupo Avançado) - cabe ressaltar aqui a hipótese de os alunos usarem diferentes pistas para chegarem à mesma identificação; enquanto o nível mais básico seguiria a força de explosão, o nível avançado já seguiria os padrões da L2. Para corroborar tal afirmação, os resultados da terceira hipótese seriam importantíssimos. No entanto, os resultados acerca desta hipótese contradizem tais conjecturas. De acordo com a Hipótese 3, o padrão Zero Artificial seria identificado como sonoro pelos falantes avançados,

seguindo a ausência de aspiração dos estímulos, e surdo pelos falantes de nível básico, uma vez que o VOT ainda não se apresentaria como principal pista acústica. Porém, 76,67% do Grupo Avançado e 57,29% do Grupo Básico identificaram tal padrão como surdo. Ou seja, o grupo de nível mais avançado apresentou, inclusive, maior tendência a identificar o padrão como surdo, não confirmando, assim, a terceira hipótese.

Em relação aos resultados de discriminação, a Hipótese 4 foi confirmada, com altos índices de acuidade nos dois grupos na discriminação entre o VOT Negativo e Positivo, sendo 76,74% no Grupo Básico e 93,33% no Grupo Avançado. A Hipótese 5, por sua vez, não foi confirmada, uma vez que ambos os grupos foram capazes de discriminar os padrões VOT Negativo e VOT Zero Artificial (46,49% no Grupo Básico e 64,17% no Grupo Avançado), enquanto era esperado que apenas o nível básico o fizesse. A Hipótese 6 também foi contrariada, uma vez que a mesma previa uma diferença significativa nos índices de acuidade dos dois grupos na discriminação entre VOT Zero Artificial e Positivo, porém ambos os grupos mostraram preferência por identificar os padrões como similares (45,49% no Grupo Básico e 50,83% no Grupo Avançado)

Em suma, ao serem considerados todos os resultados das tarefas de percepção, o estudo corrobora os resultados de Alves e Motta (2014), mostrando o não uso do VOT como principal pista acústica utilizada pelos aprendizes brasileiros, independentemente do nível de proficiência.

Entre as hipóteses dos resultados de produção, por sua vez, a número 7 mostrou-se parcialmente corroborada. A mesma previa que, em cada um dos níveis de proficiência, haveria uma diferença estatisticamente significativa entre os valores de VOT de /p/ e /k/ da L1 e da L2 de ambos os grupos. Para /k/, ambos os grupos mostraram tal diferença (55,71 ms no PB do Grupo Básico e 51,7 ms no PB do Grupo Avançado, contra 68,87 ms na L2 do Grupo Básico e 79,80 ms na L2 do Grupo Avançado). Já para /p/, apenas o nível básico mostrou diferença significativa, enquanto o nível avançado não, provavelmente por causa dos valores já altos nesta consoante na L1 (24,13 ms no PB do Grupo Básico e 33 ms no PB do Grupo Avançado, e 45,04 ms na L2 do Grupo Básico e 34,40 ms na L2 do Grupo Avançado). Segundo a Hipótese 8, por sua vez, não haveria diferenças significativas entre as durações de VOT produzidas por aprendizes básicos e avançados de inglês. Tal hipótese foi corroborada, uma vez que os dados de /p/ (45,04 ms no Grupo Básico e

34,40 ms no Grupo Avançado), /t/ (59,04 ms no Grupo Básico e 58,40 ms no Grupo Avançado) ou /k/ (68,87 ms no Grupo Básico e 79,80 ms no Grupo Avançado) não mostraram diferenças estatisticamente significativas.

Os resultados referentes à tarefa de produção, então, evidenciaram que os aprendizes já produzem maiores valores de VOT na língua-alvo do que na L1. Porém, tais valores não se assemelham aos valores nativos, o que leva os autores a sugerirem que, embora os aprendizes já reconheçam a necessidade da produção de um VOT mais longo para a redução do sotaque, a pista adquira um caráter "alofônico", não sendo tomado como principal detalhe fonético para a distinção entre segmentos surdos e sonoros.

Em suma, os resultados encontrados neste estudo corroboram os de Alves e Motta (2014), com o VOT não desempenhando o papel de principal pista para a distinção das plosivas entre aprendizes brasileiros. Novamente, os resultados e as reflexões do artigo abrem caminhos para outros possíveis estudos, como a influência da L2 sobre a L1, posteriormente trabalhada por nosso grupo de pesquisa (KUPSKE, 2016; PEREYRON, 2017; SCHERESCHEWSKY; ALVES; KUPSKE; 2017; SCHERESCHEWSKY, 2018) - não resenhados no presente trabalho por delimitações do mesmo -. Os autores também sugerem a continuidade da investigação de qual pista, ou conjunto de pistas, seriam primordiais para os aprendizes fazerem a distinção de sonoridade já citada, assim como estudos que prezem pela tentativa de fazer os aprendizes atentarem ao VOT como pista principal.

#### 2.4.4 Schwartzhaupt, Alves e Fontes (2015)

Após os estudos anteriores (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015) apresentarem os resultados referentes ao uso dos padrões de VOT por aprendizes brasileiros para as distinções de sonoridade das plosivas do inglês, Schwartzhaupt, Alves e Fontes, em 2015, continuaram os estudos de percepção não apenas com aprendizes brasileiros, mas, desta vez, também com falantes nativos do inglês. O objetivo de tal estudo era investigar como os diferentes padrões do VOT seriam categorizados pelos falantes nativos da língua-alvo e pelos aprendizes brasileiros.

Dois grupos participaram do estudo. O primeiro contava com 20 falantes nativos de inglês norte-americano, todos nascidos no estado da Pensilvânia, sendo que todos haviam aprendido somente a língua inglesa até os seis anos de idade. O segundo era

constituído por 17 aprendizes brasileiros de inglês, nascidos e residentes na cidade de Porto Alegre - RS, que também haviam aprendido somente o português brasileiro até os seis anos. Além disso, o segundo grupo era formado por aprendizes com nível de proficiência avançado em língua inglesa<sup>11</sup>.

Como o estudo visava apenas testar a percepção dos participantes, foram feitas apenas tarefas de identificação e discriminação. Ambas as tarefas foram as mesmas já empregadas em nosso grupo de pesquisa até o momento do estudo (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015). Em outras palavras, a tarefa de identificação trabalhava com estímulos com os quatro padrões de VOT (Negativo, Positivo, Zero Natural e Zero Artificial) e a de discriminação contrastava VOT Negativo vs. Zero Artificial, VOT Negativo vs. Positivo, e VOT Zero Artificial vs. Positivo.

Duas hipóteses foram levantadas: a primeira era de que na tarefa de discriminação haveria diferenças significativas entre os dois grupos nos índices de acurácia contrastando VOT Negativo e Zero Artificial, e Zero Artificial e VOT Positivo. Os falantes nativos não discriminariam os padrões do primeiro contraste, mas discriminariam no segundo. O oposto ocorreria com os aprendizes brasileiros. A segunda hipótese, referente à tarefa de identificação, previa diferenças significativas entre os dois grupos somente em relação ao padrão Zero Artificial. Em relação à identificação dos quatro padrões, o grupo de americanos identificaria apenas o VOT Positivo como representante de uma plosiva surda, enquanto os aprendizes brasileiros identificariam tanto o padrão Positivo quanto o Artificial como surdo.

Os resultados da tarefa de discriminação, então, mostraram que os grupos não mostraram diferenças entre si na discriminação entre os padrões VOT Negativo e Positivo, pois ambos os grupos mostraram altos níveis de acurácia (92,08% no grupo nativo e 91,66% no grupo de aprendizes). Ambos os grupos mostraram diferenças significativas na diferenciação entre VOT Negativo e VOT Zero Artificial, com os aprendizes brasileiros trazendo maiores índices de discriminação que os falantes nativos (65,68% no grupo de aprendizes e 10,41% no grupo de falantes nativos). Finalmente, entre o contraste VOT Zero Artificial e Positivo, diferenças significativas entre os dois grupos também foram encontradas, com o grupo de falantes nativos apresentando índices maiores de discriminação do que os aprendizes brasileiros

---

<sup>11</sup> À semelhança dos estudos descritos nas subseções anteriores, o teste de proficiência usado foi o *Oxford Online Placement Test*.

(77,50% no grupo nativo e 39,21% no grupo de aprendizes). A hipótese 1 foi, portanto, confirmada pelos autores.

Já os resultados para a tarefa de identificação mostraram que não houve diferenças significativas entre os grupos na identificação do VOT Negativo como sonoro (100% para aprendizes brasileiros e 99,16% para falantes nativos) e do VOT Positivo como surdo (99,01% para aprendizes e 99,16% para falantes nativos). O VOT Zero Natural, por sua vez, apresentou diferenças significativas entre os grupos. Apesar de ambos os grupos associarem o padrão como uma plosiva sonora, ela é maior entre os falantes nativos (95,83%) do que entre os aprendizes brasileiros (75,49%), com 22,54% deles identificando tal padrão como surdo, o que gerou uma diferença significativa. Já em relação ao VOT Zero Artificial, também houve diferenças significativas entre os dois grupos (com 76,66% dos falantes nativos identificando o padrão sonoro e 76,47% dos aprendizes como surdo). A hipótese 2, então, foi parcialmente corroborada. Além disso, foi confirmada a hipótese que sugeria que os falantes nativos identificariam apenas o padrão Positivo como surdo, enquanto os aprendizes identificariam como instâncias de consoantes surdas os padrões Positivo e Zero Artificial.

Em suma, tal estudo trouxe resultados importantíssimos para a área. O principal deles, obviamente, diz respeito à descoberta de que os falantes nativos de inglês norte-americano se guiam através do VOT para a distinção entre a sonoridade das plosivas. Além disso, os autores também reforçam os resultados já encontrados (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015), que sugeriam que a mesma pista não é a principal para a distinção de sonoridade por parte dos aprendizes brasileiros, o que reforça a necessidade de tarefas que visem a fazê-los atentar ao VOT como pista principal, dado o caráter fundamental desta pista para a inteligibilidade da fala em inglês (SCHWARTZHAUPT, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015) como o trabalho descrito a seguir.

#### 2.4.3 Rodrigues (2015)

Com os objetivos de a) verificar se haveria diferença nas médias de produção de VOT entre os grupos de participantes do experimento, que serão descritos em maior detalhe a seguir, em cada uma das etapas de coleta de dados; b) analisar, na observação de cada grupo tomado individualmente, se haveria diferenças nas médias

de VOT anterior e posteriormente ao treinamento perceptual, bem como se essa diferença seria mantida um mês após a realização do treinamento; e c) discutir os possíveis fatores, à luz de uma concepção dinâmica de língua, que poderiam estar sendo relevantes para a verificação dos efeitos ou não do treinamento, Rodrigues (2015), em seu Trabalho de Conclusão de Curso sobre o papel do treinamento perceptual, reforça a continuidade dos estudos de percepção e produção de padrões de VOT do inglês por brasileiros.

Ao todo, 67 participantes do curso de Letras da Universidade Federal do Rio Grande do Sul começaram como parte do estudo – os participantes do presente estudo, inclusive, encontram-se entre esses participantes primários. Destes 67, após aplicados os controles metodológicos necessários para a delimitação do estudo (como (i) a realização do *Oxford Online Placement Test* e o atestado de, no mínimo, proficiência nível B1 de acordo com o *Common European Framework*, (ii) a necessidade de participação nas três fases do experimento (Pré-teste, Pós-teste e Pós-teste Postergado) e (iii) a participação de, no mínimo, duas das três sessões de treinamento), restaram apenas 19 participantes, que foram divididos entre Grupo Controle (seis participantes), Grupo Experimental 1 (formado por oito participantes em duas sessões de treinamento) e Grupo Experimental 2 (formado por cinco alunos que participaram de três sessões de treinamento).

O estudo foi dividido em três etapas de coleta de dados: Pré-teste, Pós-teste e Pós-teste Postergado. Cada uma dessas sessões contou com uma tarefa de identificação e uma tarefa de produção. As sessões de treinamento foram feitas durante três sessões ao longo de uma semana, através do *software* TP (RAUBER *et al.*, 2013), entre as etapas de Pré-teste e Pós-teste. Os instrumentos aplicados no pré e nos pós-testes, inclusive, foram os mesmos utilizados em Alves e Motta (2014), no caso da tarefa de identificação, e em Alves e Zimmer (2015), no caso da tarefa de produção. A tarefa de treinamento, por sua vez, é a mesma utilizada no presente estudo<sup>12</sup>. Vale ressaltar que todas as etapas e instrumentos utilizados são os mesmos que compõem a primeira fase de nosso trabalho. Por questões de delimitação do estudo, o autor analisou somente os dados referentes à produção das consoantes /p/,

---

<sup>12</sup> Aprofundaremos as explicações sobre a tarefa de treinamento no capítulo 3, referente à Metodologia, do atual trabalho.

/t/ e /k/ de palavras da L2 e de /p/ e /k/ para o português, pela semi-aspiração já atestada em estudos anteriores, supracitados no presente estudo.

No Pré-teste, os valores de produção de /p/ e /k/ para o grupo Controle revelaram resultados descritivamente mais baixos que os dos outros dois grupos. O autor salienta, ainda, que participantes de tal grupo apresentaram níveis de proficiência menores que os dos grupos 2 e 3. Também foram apontados, para /p/, resultados de diferença significativa entre os grupos Controle e Experimental 1, e Controle e Experimental 2, mostrando desvantagem entre o Controle e os outros dois grupos desde o Pré-teste. Já em relação à produção da consoante /t/, em termos descritivos, a média mais baixa foi encontrada no Experimental 1, enquanto a do Experimental (82,08 ms) se mostrou, inclusive, mais alta que as encontradas na língua-alvo, que tende a ser de 75ms, aproximadamente (CHO; LADEFOGED, 1999). A consoante /k/, por sua vez, não se mostrou estatisticamente significativa, pois os valores dos três grupos estavam bastante próximos, sendo todas as médias bastante semelhantes, em termos descritivos (74,91 ms para o Grupo Controle, 81,6 ms para o Experimental 1 e 90,71 ms para o Experimental 2), ao apontado pela literatura (entre 80 e 85 ms, cf. CHO; LADEFOGED, 1999). Em outras palavras, a consoante /k/ foi a primeira que exhibe os valores mais próximos do padrão nativo. O autor, então, sugere que talvez a realização de um treinamento perceptual se faça mais necessária com segmentos labiais e coronais do que com dorsais. Uma possível explicação para tal fenômeno seria o fato de que o /k/, no português brasileiro, já aparece com um certo grau de aspiração (M. ALVES, 2015; KUPSKE, 2016), porém menos acentuado do que o encontrado em inglês. Tal premissa corrobora os resultados encontrados em nosso estudo, que serão discutidos em maior profundidade no capítulo 4.

Já no Pós-teste, para /p/, os valores apresentados para o grupo Experimental 1 continuaram sendo mais elevados, em termos descritivos, do que dos demais grupos da pesquisa. Em termos estatísticos, foi apontada uma diferença significativa entre as médias de /p/ dos grupos Controle e Experimental 1. Esses resultados ressaltaram novamente a suposta "desvantagem" do Controle em relação aos outros dois grupos Experimentais para este ponto de articulação. O autor não atribuiu tal diferença ao treinamento, uma vez que essa desvantagem já havia sido vista na primeira etapa. Em relação à consoante /t/, o Controle apresentou, novamente, uma média de VOT maior que a do Experimental 1. Quanto à consoante /t/, assim como já havia sido apontado no pré-teste, também nesta etapa da pesquisa, o Grupo Controle apresenta

uma média de VOT maior do que a obtida para o grupo Experimental 1. Além disso, também houve diferença estatisticamente significativa entre os valores dos grupos Experimentais, o que também havia ocorrido no Pré-Teste. Para a consoante /k/, não foi evidenciada nenhuma diferença significativa, apesar de, em termos descritivos, os valores apresentados pelo Experimental 2 serem maiores do que os dos outros grupos (79,53 ms para o Controle, 87,16 ms para o Experimental 1 e 95,01 ms para o Experimental 2).

Finalmente, o Pós-teste Postergado mostrou que nas consoantes /t/ e /k/ os valores do Grupo Controle estavam acima dos valores apresentados pelo Grupo Experimental 1, em termos de estatística descritiva. Tal vantagem do Controle para /t/ já havia sido notada desde a primeira etapa. A consoante /p/, por sua vez, manteve, em termos descritivos, o valor da média abaixo do das médias dos outros grupos. Porém, uma vez que a diferença se mostra notável desde o Pré-teste, o autor não a atribuiu ao treinamento perceptual<sup>13</sup>. Ainda para /p/ foi apontada diferença significativa, na terceira etapa, entre as médias dos grupos Controle e Experimental 1, e grupo Controle e Experimental 2. Tais diferenças também foram verificadas na primeira e segunda etapas, o que, segundo o autor, impede a atribuição dessas diferenças aos efeitos do treinamento perceptual. Para a consoante /t/, também foi evidenciada diferença significativa entre o grupo Controle e o grupo Experimental 1, entre os grupos Experimentais 1 e 2, e o grupo Controle e o grupo Experimental 2. O autor especula que a diferença entre os grupos Controle e Experimental 1 deu-se em função da superioridade das médias do Grupo Controle, superioridade essa que já fora apontada nas etapas anteriores. Também foi apresentada uma diferença significativa entre os valores do grupo Controle e do Grupo Experimental 2, devido aos altos valores de VOT apresentados pelo último grupo. Para a consoante /k/, novamente não foi evidenciada nenhuma diferença que fosse significativa, devido aos valores das médias dos três grupos estarem próximos.

Em suma, o autor conclui que, após o treinamento perceptual, os resultados apontaram um aumento descritivo das médias de VOT dos dois grupos experimentais, em todos os três pontos. O grupo controle, por sua vez, não se manteve estável em todos os pontos, pois foi observado o acréscimo de VOT somente para /p/ e para /k/. O autor creditou este aumento, principalmente em relação a /k/, à semi-aspiração

---

<sup>13</sup> Para maiores informações em relação ao valor das médias citadas, cf. Rodrigues (2015, p. 48).

deste ponto no português brasileiro, conforme previsto em M. Alves (2015) e Kupske (2016).

O trabalho de Rodrigues, então, é um dos que mais se assemelha ao presente estudo por diversos fatores, mas ao mesmo difere-se por detalhes dentro das semelhanças. Primeiro, a mesma metodologia é empregada entre o trabalho do autor e a primeira fase do nosso. Porém, aqui, a diferença se dá pela adição de uma quarta etapa de coleta ao nosso trabalho, ainda fora dos parâmetros sugeridos para estudos longitudinais que visem estudar SACs. Segundo, compartilhamos o escasso número de mesmos participantes, ainda que no nosso estudo os mesmos estejam ainda mais reduzidos. Tal fator, porém, não se torna um problema, uma vez, que, à luz de Lowie e Verspoor (2015), que traz a relação da língua como um SAC e a variabilidade, partimos para a análise individual de cada sujeito. Terceiro, visamos contribuir com os estudos acerca dos efeitos que o treinamento perceptual terá na percepção dos participantes, o que não foi feito, por delimitações do trabalho, pelo autor. E, quarto, ambos apresentam grupos que não são selecionados de acordo com a proficiência, fato a que o autor se refere como uma limitação de estudo. Tal limitação, para nós, não se mostra necessariamente desse modo. Conforme mostrado no estudo de Lowie e Verspoor (2018), mesmo aprendizes de um grupo homogêneo mostram trajetórias heterogêneas. Enquanto o autor traz seu estudo numa perspectiva de grupo sob a lógica causa-efeito, nosso estudo, além disso, tenta ver os aprendizes individualmente, em suas médias e variações, além de investigá-los após o longo período já citado.

#### 2.4.5 Kampff e Alves, 2016<sup>14</sup>

Foi deste trabalho, apresentado em 2016 no II Encontro do Núcleo de Pesquisa em Fonética e Fonologia Aplicada à Língua Estrangeira (NUPFFALE), em Curitiba-PR, que se originou o presente TCC. O objetivo do estudo era avaliar a eficiência do treinamento perceptual (com estímulos com VOT Positivo e VOT Zero Artificial) sobre a percepção e a produção da aspiração de /p/, /t/ e /k/, em posição inicial, no inglês, por aprendizes brasileiros. Também era visado discutir a contribuição do treinamento

---

<sup>14</sup> Considerando que o estudo de Kampff e Alves (2016) constituiu a primeira fase do presente trabalho (com as primeiras três coletas de dados), detalhes e explicações mais aprofundadas em questão de Metodologia, como instrumentos, etapas e participantes, estarão presentes no Capítulo 3 deste estudo.

nos índices de acuidade de percepção dos quatro padrões de VOT testados (VOT Zero Natural, VOT Zero Artificial, VOT Negativo e VOT Positivo), além de se verificar se o treinamento era capaz de contribuir com um aumento dos índices de produção de VOT das plosivas surdas iniciais, investigando-se se os possíveis efeitos encontrados no Pós-teste teriam efeitos duradouros.

Para isso, 20 alunos do primeiro semestre de graduação em Letras - Português e Inglês, da Universidade Federal do Rio Grande do Sul (UFRGS), foram divididos em dois grupos, o Grupo Controle e o Grupo Experimental, resultando em 10 participantes para cada grupo. A divisão ocorreu de acordo com a turma em que os aprendizes estavam, e não por nível de proficiência dos participantes. Enquanto o grupo Experimental participou de três sessões de treinamento perceptual, o grupo Controle não recebeu qualquer tipo de instrução ou de sessão de treinamento acerca dos padrões de VOT. Ambos os grupos participaram das três etapas de coleta de dados (Pré-teste, Pós-teste e Pós-teste Postergado), à semelhança de Rodrigues (2015). Em cada etapa, foram coletados dados referentes à percepção, com a mesma tarefa de identificação usada em Alves e Motta (2014), e Alves e Zimmer (2015), e a tarefa de produção sendo a mesma usada em Alves e Zimmer (2015) e Rodrigues (2015). Ambas as tarefas foram aplicadas nos mesmos moldes que os estudos supracitados.

Os resultados deste estudo, então, corroboraram Alves e Zimmer (2015). Em relação ao VOT Negativo, por exemplo, ambos os grupos mostraram índices altos de acuidade nas três etapas do estudo (sendo 95,83% no Pré-teste e 100% no Pós-teste e Pós-teste Postergado para o Grupo Experimental e 100%, 100% e 99,16% para o Grupo Controle nas respectivas etapas), assim como o VOT Positivo (97,5% no Pré-teste, 94,16% no Pós-teste e 95,83% no Pós-teste Postergado para o Grupo Experimental e 99,16%, 100% e 99,16% para o Grupo Controle nas respectivas etapas). Já para o VOT Zero Natural, os índices de acuidade foram 69,16% no Pré-teste, 98,33% no Pós-teste e 97,5% no Pós-teste Postergado. Tais índices levaram à significância estatística entre Pré-teste e Pós-teste e Pré-teste e Pós-teste Postergado. O Grupo Controle, para o mesmo padrão, não apresentou significância em nenhuma das etapas. O mesmo vai acontecer com o padrão VOT Zero Artificial, enquanto o Grupo Experimental aumenta seus índices de acuidade de 31,66%, no Pré-teste, para 62,5% no Pós-teste e para 68,33% no Pós-teste Postergado, resultando em uma diferença marginalmente significativa entre Pré-teste e Pós-teste e diferença significativa entre Pré-teste e Pós-teste Postergado, o Grupo Controle não

apresenta nenhuma significância em nenhuma das três etapas. Além disso, em nenhuma das etapas há diferenças inferenciais significativas entre os índices de acuidade. Tais diferenças aparecerão apenas a partir do Pós-teste e perdurarão até o Pós-teste Postergado, nos padrões VOT Zero Natural e Zero Artificial. Esses resultados, então, levaram os autores a atestar efeitos positivos causados pelo treinamento perceptual na identificação de plosivas surdas e sonoras do inglês.

Já em relação aos dados de produção, os resultados corroboram os resultados de Rodrigues (2015) na questão da nebulosidade sobre os efeitos do treinamento sobre a produção dos aprendizes. Os dados não se apresentaram inferencialmente significantes no Grupo Experimental em nenhuma consoante. Surpreendentemente, no Grupo Controle, houve uma diferença significativa para a plosiva /k/, que se deu entre Pré-teste e Pós-teste Postergado<sup>15</sup>.

Em suma, tal estudo corrobora premissas já trazidas nos trabalhos anteriormente citados, como, por exemplo, a identificação dos padrões Negativo e Positivo não se mostrar um empecilho para aprendizes brasileiros. Além disso, o estudo lança mais luz sobre o papel do treinamento perceptual para a distinção entre plosivas surdas e sonoras, que, ao que tudo indica, é positiva. Também, seguiu-se a reflexão iniciada em Rodrigues (2015) sobre os efeitos aos quais o treinamento pode ter na produção dos aprendizes, de modo que sugeríssemos que tais efeitos não podem ser considerados imediatos, ou de curto prazo. Esta é uma das principais questões que motivam, justamente, a realização de uma nova verificação de dados anos após a realização do treinamento perceptual.

#### 2.4.6 Alves e Luchini (2016)

Em 2016, então, Alves e Luchini, publicam os resultados de um estudo realizado em 2014, no qual aplicaram as tarefas já utilizadas nos estudos anteriormente descritos nesta seção (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; RODRIGUES, 2015), porém com aprendizes argentinos estudantes de inglês. O objetivo do estudo era investigar se

---

<sup>15</sup> Considerando que o presente trabalho analisa parte dos mesmos dados apresentados, por utilizarmos 15 dos participantes utilizados no estudo descrito, e resultados semelhantes serão encontrados em nossa seção de análise de dados, deixamos maiores detalhes e reflexões para o capítulo 4, dedicado a tal análise.

aprendizes, com a variedade rio-platense de espanhol como L1, utilizam-se do VOT como pista principal para fazer a distinção entre plosivas iniciais surdas e sonoras do inglês, assim como os americanos (SCHWARTZHAUPT; ALVES; FONTES, 2015), ou se utilizam outras pistas, tais como os brasileiros (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; RODRIGUES, 2015). Cabe ressaltar que a o VOT é utilizado para plosivas surdas e sonoras no espanhol assim como é no português.

Para isso, 15 estudantes argentinos do ensino superior da cidade de Mar del Plata (Buenos Aires, Argentina) foram selecionados. Todos tiveram sua proficiência atestada através do *Oxford Online Placement Test*, atingindo níveis C1 e C2; portanto, todos eram do nível avançado. Além disso, todos participaram das tarefas de identificação e discriminação (contendo os três padrões de VOT naturais, Negativo, Positivo e Zero Natural, e o padrão manipulado, Zero Artificial), já descritas anteriormente ao longo da seção.

Sete hipóteses foram levantadas pelos autores: a primeira previa que os aprendizes identificariam o padrão VOT Negativo como sonoro; a segunda previa a identificação do padrão Positivo como surdo; a terceira seria que os índices de identificação do padrão Zero Natural seriam maiores para "sonoro" do que para "surdo; já a quarta, que os índices de identificação do VOT Zero Artificial seriam maiores para "surdo" do que "sonoro". A quinta hipótese, já referente à discriminação, dizia que os índices de acuidade de discriminação entre o VOT Negativo e VOT Positivo seriam altos; a sexta previa os mesmos resultados, porém na discriminação entre VOT Negativo e o Zero Artificial; e a sétima hipótese previa um maior índice de igualdade entre os padrões Zero Natural e Positivo.

Em relação aos resultados de identificação, as hipóteses 1, 2, 3 e 4 foram confirmadas, sendo 97,78% o índice de identificação do padrão Negativo como sonoro e 96,11% o do padrão Positivo como surdo. Os resultados do padrão Zero Natural mostraram 77,78% dos aprendizes identificando o padrão como sonoro e 67,78% deles identificaram o padrão Zero Artificial como surdo. Todos os resultados foram estatisticamente significativos.

Com os resultados de discriminação, então, as hipóteses 5 e 6 foram corroboradas. Os padrões VOT Negativo e Positivo apresentaram índices de acuidade de 95%, enquanto a discriminação do VOT Negativo e Zero Artificial foi de 63,33%, ambos apresentando resultados estatisticamente significativos. A sétima hipótese,

referente à discriminação dos padrões Zero Artificial e Positivo, por sua vez, não foi confirmada, uma vez que a mesma previa uma significância alta nos índices de igualdade entre os padrões. Porém, este índice foi de 45,56%, enquanto o de acuidade foi 42,78%.

Novamente, o estudo corrobora outros estudos previamente realizados pelo grupo de pesquisa, que mostravam que o VOT não era uma pista primordial para aprendizes de inglês. Este trabalho, por sua vez, aumenta ainda mais a quantidade de línguas conhecidas em que tal pista não é tida como principal. Os autores lançam, também, uma série de sugestões de estudos a serem realizados, como a procura pela pista, ou pistas, principal usada para a distinção de plosivas na L1 dos aprendizes, assim como a necessidade de verificar a produção dos indivíduos em relação ao VOT, por exemplo.

#### 2.4.7 Alves e Luchini (2017)

Em 2017, então, os autores publicam novos resultados referentes aos estudos que haviam desenvolvido com aprendizes argentinos. Tal trabalho objetivou testar os efeitos do treinamento perceptual na percepção e produção dos participantes, assim como em Rodrigues (2015), e Kampff e Alves (2016). Um dos diferenciais deste estudo, porém, é a adição de um terceiro grupo, que além de receber o treinamento, também recebeu instrução explícita sobre os itens da língua-alvo estudados.

Para isso, 24 alunos foram divididos em três grupos diferentes. Cada grupo, no final, contou com oito participantes. O primeiro, chamado de Grupo 1, recebeu apenas o treinamento; o segundo, chamado de Grupo 2, além do treinamento, também recebeu, no início de cada sessão de treinamento, instrução explícita sobre o que é aspiração e como as plosivas iniciais surdas (/p, t, k/) do inglês são aspiradas. O terceiro grupo, Grupo 3, foi o grupo Controle, não tendo recebido nenhum tipo de instrução ou treinamento. Todos apresentavam índice de proficiência C1 e C2, de acordo com o *Oxford Online Placement Test*.

Todos os instrumentos usados para o treinamento, assim como as tarefas de identificação e produção, além do número de etapas de coleta (Pré-teste, Pós-teste e Pós-teste Postergado), são os mesmos usados em trabalhos descritos nas subseções anteriores (ALVES; MOTTA, 2014; ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; RODRIGUES, 2015; KAMPFF; ALVES, 2016; ALVES;

LUCHINI, 2016). Além disso, três hipóteses foram levantadas pelos autores: a primeira era de que o treinamento resultaria em maiores índices de identificação do VOT Zero Natural e Zero Artificial em ambos os grupos 1 e 2; a segunda era de que os possíveis efeitos positivos do treinamento, na percepção, seriam generalizados para a produção, especialmente pelo segundo grupo, que recebeu instrução; e a terceira era de que os efeitos positivos do treinamento, tanto na produção quanto percepção, estariam presentes um mês depois das sessões de treinamento.

A primeira hipótese foi, então, corroborada, uma vez que ambos os grupos experimentais mostraram diferenças significativas nas identificações dos padrões VOT Zero Natural, Zero Artificial e Positivo ao longo das etapas de coleta. A segunda hipótese, por sua vez, não foi totalmente confirmada. De fato, apenas diferenças marginalmente significativas foram encontradas no Grupo 1. O Grupo 2, porém, apresentou um aumento significativo na produção de /p/ e /t/, então a possibilidade de que a instrução produziu efeitos na produção nos aprendizes não pode ser desconsiderada. A terceira hipótese também foi parcialmente corroborada, com diferenças significativas, na percepção, entre o Pré-teste e Pós-teste Postergado do Grupo 2. Para o VOT Zero Natural, no mesmo grupo, diferenças significativas aparecem entre o Pré-teste e em cada um dos Pós-testes, o que corrobora a hipótese. Porém, em relação à percepção do padrão Zero Artificial, a diferença aparece apenas entre o Pré-teste e o Pós-teste Postergado. Ambos os resultados, segundo os autores, levaram a acreditar que a aplicação do treinamento perceptual e instrução explícita resulta em mudanças imediatas na percepção dos aprendizes. Tais mudanças, porém, não parecem manter-se por muito tempo, e os autores sugerem que talvez os aprendizes necessitem de um intervalo maior para "sintonizar" a pista acústica do VOT. Em relação à produção, os autores encontraram indícios de que as diferenças significativas na produção de /p/ e /t/ do segundo grupo permaneceram em longo prazo<sup>16</sup>.

Em suma, os autores trouxeram novas questões que, de fato, são relevantes tanto para o estudo da área quanto para o presente trabalho, como a adição de etapas de coleta e a importância de estudos futuros que envolvam o uso tanto do treinamento quanto da instrução explícita, inclusive separados. As limitações do estudo, inclusive, se mostram relevantes. Tais limitações, porém, por serem muito semelhantes às do

---

<sup>16</sup> Para conferir resultados em termos descritivos, cf. Alves; Luchini (2017).

presente estudo, estão melhor discutidas e elencadas no capítulo 5, das considerações finais.

## **2.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO**

Este capítulo teve por objetivo retomar os pressupostos teóricos que permeiam o presente estudo. A partir dele é possível perceber que nosso trabalho propõe tratar de dois fatores já presentes em estudos brasileiros: o primeiro é dar continuidade aos estudos que concebem a língua sob a perspectiva da teoria dos Sistemas Adaptativos Complexos, refletindo sobre o que tal visão implica em nosso desenvolvimento linguístico; a segunda é a continuação dos estudos de nosso grupo de pesquisa, considerando a relação próxima com os trabalhos anteriores, como o uso dos mesmos instrumentos para tarefa de identificação e de produção, participantes, e etapas de coleta.

Porém, é de extrema importância destacar, também, os motivos pelos quais nosso estudo se difere do que já foi realizado em nosso grupo de estudo e, inclusive, no país, no que diz respeito a estudos sobre efeitos do treinamento perceptual. O primeiro motivo é a adição de uma quarta etapa de coleta. Até então, nenhum de nossos trabalhos havia apresentado esta etapa. Discutiremos o mérito desta adição, e o demérito em relação ao controverso intervalo de tempo entre ela e a terceira etapa, mais adiante. O segundo, e, a nosso ver, principal motivo é o começo, no país e em nosso grupo de pesquisa, das análises referentes ao papel do treinamento perceptual que levam em consideração as trajetórias individuais dos participantes, corroborando a visão dinâmica de língua, de acordo com a qual cada sujeito e aprendiz possui seu próprio sistema e reage de maneira diferente a mudanças externas exercidas sobre o mesmo.

Partiremos, agora, para o capítulo referente à Metodologia utilizada no presente trabalho, seguida da análise de dados e das considerações finais do estudo.

### **3. METODOLOGIA**

Esta seção dedica-se à metodologia utilizada para a realização do trabalho. Primeiramente, apresentaremos informações sobre os participantes; em seguida, relataremos como foi realizado o treinamento perceptual, seguido dos instrumentos para a coleta de dados no Pré-teste e nos três Pós-Testes; finalmente, descreveremos os cuidados metodológicos adotados na análise dos resultados.

Retomando a Introdução, é importante lembrar que o trabalho foi realizado na seguinte ordem: em abril de 2015, na primeira fase do projeto, foram realizados o Pré-teste (uma semana após tal coleta, foi aplicada a etapa de treinamento perceptual); o primeiro Pós-teste, realizado imediatamente após o treinamento; e, finalmente, no que diz respeito à primeira fase, o Pós-teste Postergado 1 foi levado a cabo uma semana após o Pós-teste. Em junho de 2018, na segunda fase, foi realizado o Pós-teste Postergado 2.

#### **3.1 PARTICIPANTES**

Na primeira fase do projeto, cerca de setenta alunos de três turmas da graduação em Letras - Português e Inglês, da Universidade Federal do Rio Grande do Sul (UFRGS), foram divididos em dois grupos: Experimental, daqui em diante referido como GE, e Controle, daqui em diante referido como GC. O primeiro contava com um total de 40 alunos e o segundo, com 27. Enquanto o grupo Experimental participou das sessões de treinamento, o grupo Controle não recebeu qualquer tipo de instrução ou treinamento na ocasião do experimento<sup>17</sup>. Para que a participação dos aprendizes fosse válida, os mesmos precisavam estar presentes em todas as três etapas. Com isso, ao final da terceira etapa, apenas 10 participantes permaneceram em cada um dos grupos, totalizando 20 alunos. Em 2018, na realização da segunda fase, cinco deles haviam desistido do curso ou trocado de ênfase em língua estrangeira, o que os desclassificou para continuarem participando do experimento.

---

<sup>17</sup> Reconhecemos, indiscutivelmente, a possibilidade de que, ao longo dos três anos de graduação em Letras, os alunos do GC, apesar de não terem participado das sessões de treinamento, possam ter sido instruídos sobre o fenômeno de aspiração em suas aulas, assim como terem tomado conhecimento dos objetivos da pesquisa. Tal fato, em nossa opinião, não deve ser visto negativamente, o que caracteriza um argumento adicional para que realizemos um acompanhamento de caráter mais individual para cada participante.

Ao final do trabalho, então, o Grupo Experimental contou com sete participantes, enquanto o Controle contou com oito.

Todos os participantes receberam e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE)<sup>18</sup> (Anexo D). O termo explicava aos participantes como seria a condução do trabalho, além de trazer os efeitos benéficos, como uma possível melhora na língua inglesa, e os possíveis riscos, como cansaço e nervosismo.

### 3.2 SESSÕES DE TREINAMENTO PERCEPTUAL

As sessões de treinamento consistiram em um teste de identificação com *feedback* imediato, através do *software* TP (RAUBER *et al.*, 2013), sendo realizadas 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 atribuíssem *status* de surdo, e VOT Zero Artificial, esperando-se que atribuíssem *status* sonoro, indicaria que os aprendizes estariam tomando o VOT como pista acústica fundamental para a distinção dos pares mínimos, uma vez que estariam priorizando o VOT sobre o padrão híbrido. Ou seja, a função do treinamento era ensinar o aprendiz a seguir o VOT como pista acústica principal na distinção entre oclusivas surdas e sonoras.

Os estímulos foram gravados em um estúdio profissional particular da cidade de Porto Alegre, RS, no ano de 2011, com o início do projeto que deu insumos ao estudo de treinamento perceptual em que se encontra este trabalho, por seis falantes nativos de inglês americano, residentes no Brasil há menos de três meses antes da coleta dos dados (três homens e três mulheres).

Seguindo Alves e Luchini (2017), os itens lexicais usados como estímulos no treinamento foram '*pee*', '*tip*' e '*kit*', seguindo os estudos de Yavas e Wildermuth (2006) e Schwartzhaupt (2012), que afirmam um valor de aspiração maior nas plosivas surdas quando seguidas de uma vogal alta, o que facilitaria a percepção<sup>19</sup>. Cada falante gravou diversas vezes os itens lexicais supracitados; ao final, foram

---

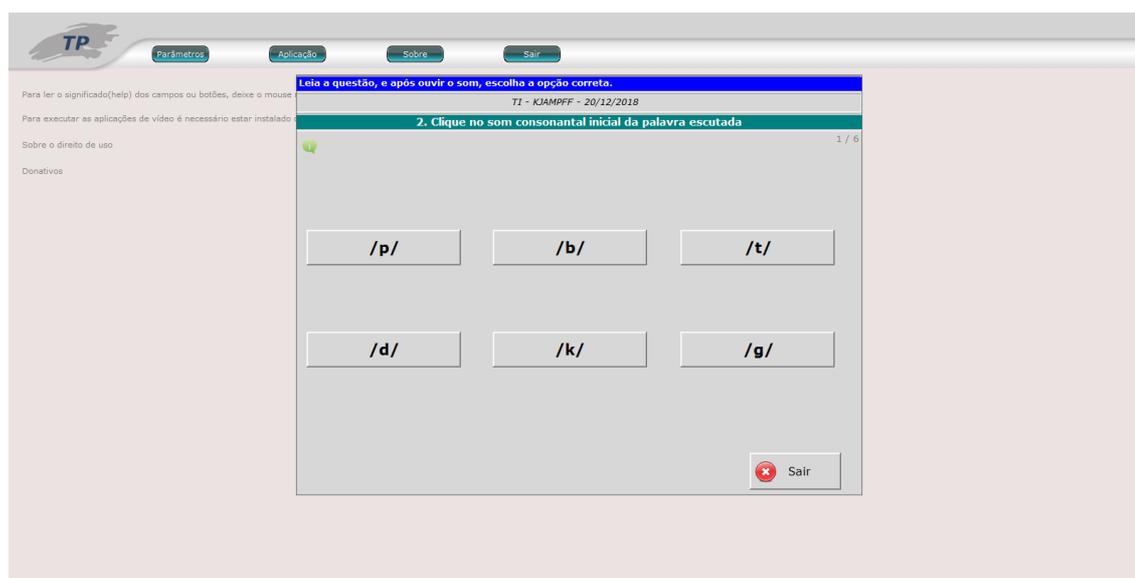
<sup>18</sup> O projeto de pesquisa no qual o presente trabalho se insere, "Efeitos de treinamento perceptual e instrução explícita na aquisição da distinção entre plosivas surdas e sonoras iniciais do inglês (LE) por aprendizes argentinos e brasileiros: o papel de pistas acústicas primordiais na L1", sob a coordenação do Prof. Dr. Ubiratã Kickhöfel Alves (UFRGS), teve parecer favorável por parte do Comitê de Ética em Pesquisa da Instituição (número do parecer: 648.674, emitido em 24/04/2014).

<sup>19</sup> Para fins de treinamento, somente foram utilizados os padrões VOT Positivo e VOT Zero Artificial.

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. Metade desses itens tiveram sua aspiração cortada para que virassem VOT Zero Artificial. 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 eram dadas a cada 90 *tokens*.

As sessões de treinamento, que consistiam em uma tarefa de Identificação, foram realizadas no laboratório de informática do prédio de aula da Letras da UFRGS, com todos os participantes sendo testados ao mesmo tempo, um em cada computador, e usando fones de ouvido. Os aprendizes deveriam sinalizar a consoante inicial da palavra que eles ouviram, como é mostrado na Figura 1, a seguir.

**FIGURA 1.** As opções de escolha da tarefa de Identificação no TP.

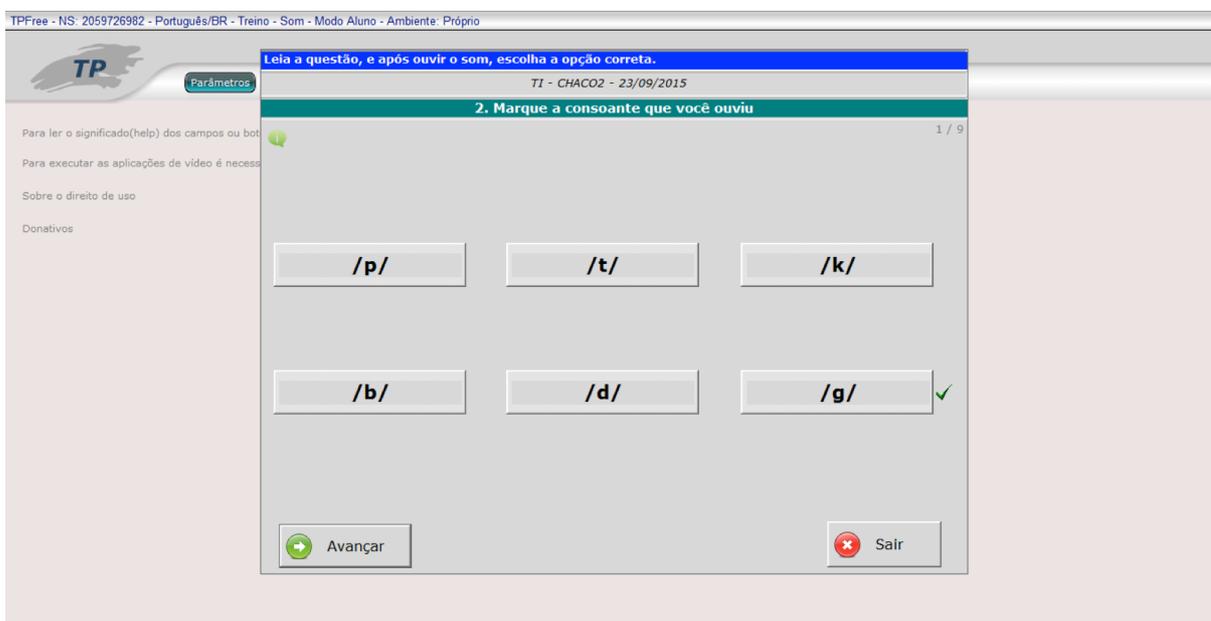


**Fonte:** Arquivo do autor.

Como dito anteriormente, o *feedback* era dado imediatamente após cada uma das respostas escolhidas pelos aprendizes; 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 na 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. Ao fazer isso, era esperado 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.

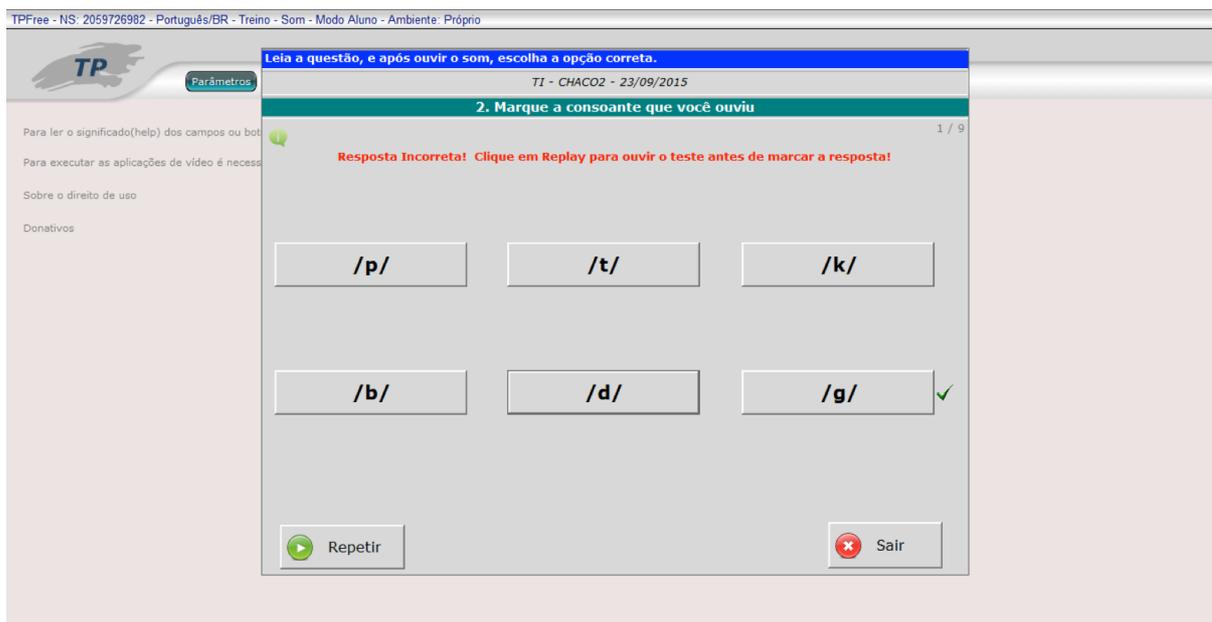
**FIGURA 2.** Tarefa de Identificação com *feedback* positivo.



**Fonte:** Arquivo do autor.

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

**FIGURA 3.** Tarefa de Identificação com *feedback* negativo.



Fonte: Arquivo do autor.

Após a finalização do teste, uma tabela em Excel era criada automaticamente pelo programa. Nela, estavam detalhados os números de erros e acertos de acordo com os critérios mencionados anteriormente nesta seção.

### 3.3 INSTRUMENTOS DE COLETA DE DADOS (PRÉ E PÓS-TESTES)

As tarefas do experimento se dividiram nas seguintes etapas: Pré-teste, com os testes de produção em português e inglês, e percepção em inglês; treinamento perceptual, com três sessões divididas ao longo de uma semana (cujos procedimentos já foram descritos na seção anterior); Pós-teste Imediato, realizado dois dias após o treinamento, com percepção e produção em inglês; Pós-teste Postergado 1, um mês após o pré-teste, com percepção e produção em inglês; e Pós-teste Postergado 2, cerca de três anos após o Pós-teste Postergado 1, com percepção e produção em inglês. Nesta seção, serão descritas as tarefas de percepção e produção utilizadas no pré-teste e nos três pós-testes.

### 3.3.1 Tarefa de Identificação

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

Na Tarefa de Identificação aplicada no pré-teste e pós-testes, os aprendizes ouviam estímulos individuais e escolhiam a consoante plosiva inicial da palavra ouvida, sem qualquer tipo de *feedback* após a escolha da resposta. Neste teste, os seguintes estímulos, com quatro padrões de VOT (Negativo, Zero, Positivo e Artificial), foram incluídos e apresentados em ordem aleatória: '*pit*', '*bit*', '*tick*', '*dick*', '*kill*' e '*gill*'<sup>20</sup>. Além disso, o teste consistia em 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>21</sup>).

Após a finalização do teste, à semelhança do que o programa realizava com os dados do treinamento, uma tabela em Excel, contendo os acertos e desvios do participante, também era criada automaticamente.

### 3.3.2 Tarefa de Produção

A tarefa de produção também segue o padrão de estudos anteriores, como o de Alves e Zimmer (2015). Em uma sala silenciosa, individualmente, os aprendizes liam um arquivo de .ppt com palavras em ordem aleatória, dentre elas as palavras alvo '*peer*', '*pit*', '*pee*', '*team*', '*tick*', '*tip*', '*kit*', '*keel*' e '*kill*', com cada palavra alvo 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.

---

<sup>20</sup> 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, ela indicasse a capacidade de generalização para outros itens lexicais.

<sup>21</sup> Mesmos falantes dos estímulos presentes no treinamento perceptual.

### 3.3.3 Questionário de Experiência e Proficiência Linguística

Para que conseguíssemos responder às segunda e terceira perguntas norteadoras, que propõem refletir sobre o que individualidade de cada participante pode revelar acerca de suas trajetórias de desenvolvimento linguístico, era importante que levássemos em consideração, além dos resultados estatísticos e descritivos das tarefas de identificação e produção, a familiaridade dos participantes com a língua inglesa, a partir de uma visão mais qualitativa e holística de cada participante. Aplicamos, então, o Questionário de Experiência e Proficiência Linguística (Anexo A), criado por Scholl e Finger (2013). A partir de tal instrumento, os próprios estudantes avaliavam e refletiam sobre seu *background* linguístico através de perguntas sobre o número de línguas com que tinham contato; a idade em que começaram a aprender tais línguas; os ambientes em que as usavam; e a quantidade de horas por dia em que tais línguas eram utilizadas, por exemplo. Este questionário se mostrou importante porque, a partir dele, obtivemos informações mais detalhadas sobre cada aprendiz, o que nos será útil, sobretudo, quando forem discutidas as trajetórias individuais de cada participante, para que possamos ter um entendimento mais global, de caráter mais qualitativo, de cada indivíduo, conforme advogado pelos teóricos da visão dinâmico-complexa de língua (LARSEN-FREEMAN, 2017; LOWIE, 2017; LOWIE; VERSPOOR, 2015, 2018, dentre outros).

### 3.4 ANÁLISE DE DADOS

Após a coleta, os dados de produção foram analisados acusticamente no *software Praat* (BOERSMA; WEENIK, 2015)<sup>22</sup>. As medidas da aspiração das três plosivas surdas - /p/, /t/ e /k/ - foram primeiramente agrupadas em tabelas Excel para a rodagem final das estatísticas descritivas e a posterior verificação inferencial. Os dados referentes à percepção, a partir das tabelas geradas automaticamente pelo *software TP*, foram analisados e agrupados em uma nova tabela, contabilizando os acertos, ainda de acordo com os critérios trazidos na seção 3.2.

As estatísticas descritivas e inferenciais foram rodadas no programa SPSS versão 21 para Windows (SPSS Inc.; Chicago, IL, USA). Foi realizada a análise

---

<sup>22</sup> Versão 6.0 nas três primeiras etapas, e versão 6.0.41 na quarta.

descritiva (médias, valor máximo, valor mínimo, desvios padrão e mediana) para cada consoante, além das verificações inferenciais através dos testes de Friedman, Wilcoxon e Mann-Whitney.

## 4 DESCRIÇÃO E DISCUSSÃO DOS RESULTADOS

Para que comecemos a responder às questões norteadoras, este capítulo será dividido em três partes. A primeira será dedicada à análise das estatísticas inferenciais, de modo a responder à primeira questão - o que tal abordagem de análise de dados pode revelar a respeito do desenvolvimento dos dois grupos; em seguida, analisaremos e discutiremos os resultados individuais, para responder à segunda pergunta - o que as variações das trajetórias individuais podem revelar. Na terceira, aprofundaremos as análises individuais para abordar a terceira questão - quais informações sobre a trajetória de desenvolvimento linguístico a variabilidade dos aprendizes em um dado momento temporal e a variação ao longo do tempo podem revelar.

### 4.1 DESCRIÇÃO DOS DADOS DOS GRUPOS E ANÁLISE INFERENCIAL

A presente seção, que trata dos resultados descritivos dos grupos e das verificações inferenciais, também será subdividida - analisaremos, primeiro, os resultados referentes à percepção e, após isso, os de produção. Tal seção abordará a discussão acerca da primeira Questão Norteadora ('O que as abordagens inferenciais e descritivas podem dizer a respeito do desenvolvimento de cada um dos grupos analisados?').

Conforme já explicitado na seção 3, sobre a Metodologia, os testes referentes à estatística inferencial foram realizados a partir do programa SPSS, levando em conta o comportamento dos grupos ao longo das etapas do treinamento.

#### 4.1.1 Descrição dos resultados das tarefas de identificação

Começamos nossa análise a partir da descrição dos índices de acurácia<sup>23</sup> nas Tarefas de Identificação (Pré-teste, Pós-Teste, Pós-Teste Postergado 1 e Pós-teste

---

<sup>23</sup>Para os padrões VOT Negativo, VOT Zero Natural e Zero Artificial, foi considerada correta a atribuição do *status* sonoro; no VOT Positivo, foi considerada correta a atribuição do *status* surdo. 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.

Postergado 2) e dos resultados dos testes de Friedman<sup>24</sup> (verificações intra-grupo, ao longo de diferentes momentos) para cada um dos dois grupos:

---

<sup>24</sup> 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 das variáveis dependentes.

**Tabela 1.** Índices de acuidade de cada um dos padrões de VOT e resultados dos testes de Friedman referentes à Tarefa de Identificação.

<b>Grupo 1 - Experimental</b>					
	Pré-teste	Pós-teste	Pós-teste Postergado 1	Pós-teste Postergado 2	X2 (df)
<i>VOT Negativo</i>	<b>97,62%</b> 82/84 (0,78)	<b>100%</b> 84/84 (0)	<b>100%</b> 84/84 (0)	<b>100%</b> 84/84 (0)	9,00 (3) (*)
<i>VOT Positivo</i>	<b>96,43%</b> 81/84 (1,13)	<b>94,05%</b> 79/84 (1,15)	<b>96,43%</b> 81/84 (0,75)	<b>100%</b> 84/84 (0)	7,65 (3) *?
<i>VOT Zero Natural</i>	<b>70,24%</b> 59/84 (1,25)	<b>100%</b> 84/84 (0)	<b>96,43%</b> 81/84 (0,48)	<b>97,62%</b> 82/84 (0,75)	18,42 (3) (***)
<i>VOT Zero Artificial</i>	<b>35,71%</b> 30/84 (3,28)	<b>59,52%</b> 50/84 (2,23)	<b>59,52%</b> 50/84 (1,73)	<b>51,19%</b> 43/84 (2,11)	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> 96/96 (0)	<b>100%</b> 96/96 (0)	<b>100%</b> 96/96 (0)	<b>100%</b> 96/96 (0)	,000 (3)
<i>VOT Positivo</i>	<b>100%</b> 96/96 (0)	<b>100%</b> 96/96 (0)	<b>98,96%</b> 95/96 (0,35)	<b>97,92%</b> 94/96 (0,70)	2,00 (3)
<i>VOT Zero Natural</i>	<b>79,17%</b> 76/96 (2,0)	<b>81,25%</b> 78/96 (1,9)	<b>77,08%</b> 74/96 (2,1)	<b>83,33%</b> 80/96 (2,0)	1,88 (3)
<i>VOT Zero Artificial</i>	<b>21,88%</b> 21/96 (1,18)	<b>20,83%</b> 20/96 (2,8)	<b>25%</b> 24/96 (2,0)	<b>28,13%</b> 27/96 (1,68)	2,00 (3)

**Nota:** \*? p < ,10 (marginalmente significativo); \*p < ,05; \*\* p < ,01; \*\*\*p < ,001; DP = desvio padrão.

**Fonte:** Elaborada pelo autor.

Analisando a Tabela 1, vemos que três dos padrões apresentam diferenças inferenciais significativas entre as quatro etapas no GE, 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, podemos concluir, então, que à primeira vista o treinamento perceptual parece apresentar, em termos de estatística inferencial, resultados que podemos considerar como benéficos.

Ao olharmos os resultados descritivos, fica mais clara a contribuição do treinamento, pois, além da significância, os aprendizes aumentaram o índice de acurácia de 97,62% no Pré-teste para 100% no Pós-teste, e assim o mantiveram no Pós-teste Postergado 1 e no Pós-teste Postergado 2, no que diz respeito ao padrão VOT Negativo do GE. Além disso, cabe mencionar que tal diferença significativa não era esperada, uma vez que os índices do Pré-teste já eram altos, porém o treinamento contribuiu para o aumento do mesmo. São destacáveis, também, os resultados referentes aos índices de acuidade do GE que aumentaram entre as etapas, inclusive entre as duas últimas, com o maior espaço de tempo, como os casos dos VOTs Positivo (indo de 96,43% para 100%) e Zero Natural (indo de 96,43 para 97,62%); e resultados descritivos que mostram um aumento entre Pré-teste e Pós-teste Postergado 1, indo de 96,43% para 100%, no Positivo, e de 70,24% para 100%, no caso do Zero Natural.

Além disso, de acordo com os resultados encontrados em estudos anteriores (ALVES *et al.* 2011; ALVES; ZIMMER, 2012, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015), aprendizes brasileiros tendem a identificar o padrão Zero Artificial como surdo; na Tabela 1, através da estatística descritiva, vemos que em 35,71% dos dados dos participantes do GE, no Pré-teste, já se identificava a plosiva como sonora. Em outras palavras, 64,29% dos dados referentes a este padrão eram identificados como surdos, corroborando os estudos supracitados. Ao aumentar o índice de acuidade para 59,52% no Pós-teste e mantê-lo no Pós-teste Postergado 1, vemos que, nesta etapa do experimento, o treinamento contribuiu positivamente para a percepção dos aprendizes.

É importante, também, destacar que o GC, não exposto ao treinamento, não apresentou melhora significativa em nenhum dos padrões. Apesar disso, há oscilações, em termos de estatística descritiva, nos índices de acuidade deste grupo, como podemos ver nos padrões Zero Natural e Artificial, indo de 79,17% no Pré-teste para 81,25% no Pós-teste, alcançando os valores de 77,08% no Pós-teste Postergado

1 e 83,33% no Pós-teste Postergado 2, no caso do primeiro padrão, e de 21,88% para 20,83%, e desse valor a 25% e 28,15%, nas mesmas etapas, para o segundo. É possível que isso se dê pelo contato do grupo com os próprios experimentos perceptuais, que podem acabar por afetar o sistema dos participantes. Além disso, também, é possível que este aumento seja fruto de uma simples variabilidade característica dos SACs, conforme já expresso no Referencial Teórico deste trabalho.

Na próxima tabela, trazemos a comparação dos resultados dos testes Post-hoc de Wilcoxon (com correção de Bonferroni) da Tarefa de Identificação apenas do GE (por ser o único a apresentar diferenças significativas entre as etapas).

**Tabela 2.** Resultados do teste post-hoc Wilcoxon (com correção de Bonferroni) para o Teste de Identificação.

Grupo 1 - Experimental				
	VOT Negativo	VOT Positivo	VOT Zero Natural	VOT Zero Artificial
Pré-teste vs. Pós-teste	n.s.	n.s.	p = 1,7 *?	n.s.
Pré-teste vs. Pós-teste Postergado 1	n.s.	n.s.	p = 1,7 *?	n.s.
Pré-teste vs. Pós-teste Postergado 2	n.s.	n.s.	p = 1,7 *?	p = 1,8 *?
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 < .017, \*\*p < ,001.

**Fonte:** Elaborada pelo autor.

Como podemos ver na tabela, em relação ao padrão de 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 = 1,7). O resultado diferencia-se do apresentado em Kampff e Alves (2016), experimento esse com os mesmos participantes, porém com um número de participantes maior em cada grupo e sem a última etapa temporal de coleta de dados. O trabalho em questão trazia, nos resultados inferenciais, diferenças significativas entre Pré-teste e Pós-teste e Pré-teste e Pós-teste Postergado 1, talvez pelo maior número de participantes do grupo experimental (10 participantes). 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, o que é de suma importância, no Pós-teste Postergado 2, realizado cerca de três anos após Pós-teste Postergado 1. Voltando à Tabela 1, os valores referentes à estatística descritiva deixam essa diferença mais expressiva: dos 70,24% de acurácia na primeira etapa, há um aumento para 100% na segunda, com uma pequena diminuição para 96,43% na terceira e, na quarta, subindo para 97,62%.

Cabe destacar, também, que, 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 entre as etapas. 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, dessa forma, uma diferença estatisticamente significativa entre etapas de coleta em específico.

Em relação ao VOT Zero Artificial, ao contrário do que poderíamos esperar a partir da verificação das percentagens, que mostram um acréscimo notável entre Pré-teste e Pós-teste, subindo de 35,71% para 59,52%, e de trabalhos anteriores, como o de Kampff & Alves(2016), que se mostraram inferencialmente significativos ( $p < ,01$ ), 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. Tais dados reforçam a importância da reflexão acerca dos valores descritivos, já que eles mostram claramente um aumento considerável nos índices de acuidade.

A partir de agora, então, passamos para as 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 o Teste 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 *? $p = 7,2$	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, \* < ,05, \*\* p < ,01, \*\*\*p < ,001.

**Fonte:** Elaborada pelo autor.

Pela Tabela 3, podemos ver que as maiores diferenças entre os grupos, nos índices de acuidade referentes à identificação, se deram no Pós-teste, Pós-teste Postergado 1 e Pós-teste Postergado 2 para os padrões VOT Zero Natural e Zero Artificial. Isso indica que ambos os grupos se apresentaram estatisticamente equivalentes na primeira etapa do experimento, mas, a partir das seguintes etapas, já com o treinamento perceptual, a diferença entre eles aumentou, como já atestado anteriormente (KAMPFF; ALVES, 2016). É válido mencionar, também, que diferenças entre os dois grupos não foram encontradas nos padrões Negativo e Positivo em nenhuma das etapas, porque, justamente, 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 no Referencial Teórico deste trabalho.

Se observarmos novamente a Tabela 1, tais resultados se mostram claros a partir das porcentagens trazidas. Observemos os números referentes ao VOT Zero Natural. O GE começa o Pré-teste com 70,24% de acuidade. Já o GC começa com um índice ainda maior: 79,17%. Porém, nas seguintes etapas, enquanto o GE salta dos 70,24% para 100%, o GC mantém-se em torno dos 80%, com algumas pequenas variações.

Ao analisarmos os valores do padrão Zero Artificial, o mesmo comportamento referente ao Zero Natural ocorre. Enquanto o índice de acuidade do GE salta de 35,71% no Pré-teste para 59,52% no Pós-teste, mantendo-se no Pós-teste Postergado 1 e fechando com 51,19% no Pós-teste Postergado 2, o GC apresenta 21,88% no Pré-teste, diminuição para 20,83% no Pós-teste, aumento para 25% no Pós-teste Postergado 1, e finalmente exibindo um aumento para 28,13% no Pós-teste Postergado 2. Apesar desse aumento de acuidade entre as duas últimas etapas, os valores do GC permanecem consideravelmente abaixo dos do GE.

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, mostraram-se, no mínimo, 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. 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). Tal resultado motivará uma posterior análise individual das trajetórias dos aprendizes, de modo que respondamos às outras duas questões adicionais do estudo. Antes disso, entretanto, passemos às análises inferenciais referentes à tarefa de produção, de modo a concluirmos a resposta à primeira Questão Norteadora do trabalho.

#### 4.1.2 Descrição dos resultados das tarefas de produção

Para começar nossa análise, olhemos para a próxima tabela, com as médias de VOT e os resultados do teste de Friedman (verificação intra-grupo) 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 Friedmann.

<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<,05.

**Fonte:** Elaborada pelo autor.

De acordo com estudos anteriores (RODRIGUES, 2015; ALVES; LUCHINI, 2017), alguns, inclusive, com os mesmos participantes do atual trabalho (KAMPPFF; ALVES, 2016), a generalização dos efeitos do treinamento perceptual para a produção se mostra problemática em termos inferenciais, com resultados apenas significativos para /p/ no GE e marginalmente para /k/ no GC. Nosso trabalho não foge disso.

No GE, apenas /p/ apresenta uma diferença significativa ( $p=.029$ ) nas médias de produção. Isso fica evidente quando destacamos a diferença entre a média de 25,39 ms (mediana de 29,17 ms)<sup>25</sup> do Pré-teste com a do Pós-teste Postergado 1, de 38,48 ms (mediana de 24,17 ms). Também é importante frisar que a média da quarta etapa é 18,09 ms (mediana de 18,50 ms). Analisaremos este fato mais a fundo após apresentarmos os resultados dos testes *post-hoc* de Wilcoxon, a seguir.

Antes, porém, 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. No caso do /t/, nas primeiras três etapas os participantes passam de 61,09 ms (mediana de 54,65 ms) para 64,21 ms (mediana de 60,93) e para 68,99 ms (mediana de 60,76 ms), respectivamente. Já no que diz respeito à consoante /k/, o aumento nas mesmas etapas é de 79,7 ms (mediana de 66,46 ms) para 81,31 ms (mediana de 74,92 ms), e de tal valor para 84,24 ms (mediana de 84,94 ms), respectivamente. Já no que diz respeito à quarta etapa, ambas as consoantes sofrem uma diminuição na média, com a média de /t/ caindo para 51,23 ms (mediana de 51,76) e a de /k/, 77,85 ms (mediana de 76,02 ms). Ambas mostram uma diminuição não apenas em relação às duas últimas etapas, mas também evidenciam que os participantes, dois anos depois, apresentam médias menores que as do Pré-teste, em termos descritivos.

De forma surpreendente, o GC apresentou uma diferença marginalmente significativa ( $p=.098$ ) para os valores de /k/. Tais valores podem ser vistos nos valores descritivos das médias, quando o grupo começa com 77,10 ms (mediana de 77,49 ms) no Pré-teste, passa para 84,10 ms (mediana de 85,73 ms) no Pós-teste e chega em 89,76 ms (mediana de 91,72 ms) no Pós-teste Postergado 1. Mais surpreendente, ainda, é verificarmos que na quarta etapa, cerca de três anos após as três primeiras, o grupo apresentou uma média de 88,74 ms (mediana de 89,50), o que representa uma levíssima diminuição em relação à terceira etapa, contrariando o comportamento do GE.

Ao levarmos em consideração o diferencial do trabalho, que é o espaço de tempo entre Pós-teste Postergado 1 e Pós-teste Postergado 2, resta-nos analisar,

---

<sup>25</sup> Reconhecemos que testes não paramétricos se baseiam no valor da Mediana, e não da média, para seus cálculos inferenciais. Entretanto, ao considerarmos que o leitor pode se mostrar mais acostumado com os valores das médias, decidimos, na presente seção, reportar ambos os valores.

com maior atenção, as diferenças entre as etapas. Esses resultados são trazidos na próxima tabela, com o 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>	----	----	p = 0.36*?
<i>Pré-teste vs. Pós-teste Postergado 2</i>	----	----	----
<i>Pós-teste vs. Pós-teste Postergado 1</i>	----	----	----
<i>Pós-teste vs. 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<,017, \*\* p<,01. **Fonte:** Elaborada pelo autor.

Como podemos ver na Tabela 5, 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, acreditamos, devem-se pela diminuição expressiva nos valores das médias da última etapa do experimento. Como podemos ver na Tabela 4, de 38,48 ms (mediana de 24,17 ms), a média cai para 18,09 ms (mediana de 18,50 ms) entre as duas últimas etapas. Além disso, há diferença também nas médias entre Pré-teste e Pós-teste Postergado 2, caindo de 25,39 ms (mediana de 29,17 ms) para os 18,09 ms já citados. 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 após o treinamento, com a média se mantendo abaixo da média da primeira etapa, anterior ao treinamento.

Em relação ao GC, os valores de /k/ apresentam, surpreendentemente, neste grupo, um resultado marginalmente significativo<sup>26</sup> entre o Pré-teste e o Pós-teste Postergado 1. A diferença de 77,10 ms (mediana de 77,49 ms) no Pré-teste para 89,76 ms (mediana de 91,72 ms) no Pós-teste Postergado 1 indica que, de alguma forma, os participantes aumentaram a aspiração da consoante. Também, a pouca variação de 89,76 ms (mediana de 91,72 ms) no Pós-teste Postergado 1 e 88,74 ms (mediana de 89,50 ms) no Pós-teste Postergado 2 indicam que, além do aumento, os estudantes, seja por familiaridade ao experimento, seja por fatores externos aos do trabalho, mantiveram sua média em torno dos 90 ms (mediana de 89,50 ms). Além disso, é importante destacar que a diferença entre pré-teste e pós-teste pode ter se dado em função de, no Pré-teste, as médias e desvios-padrão deste grupo serem um pouco inferiores às do outro, sendo a média de 79,7 ms (mediana de 66,46 ms) para o GE e 77,1 ms (mediana de 77,49 ms) para o GC, e os desvios-padrão de 66,46 ms e 77,49 ms, respectivamente.

---

26 Na realização dos testes post-hoc com correção de Bonferroni, neste trabalho, consideramos como marginalmente significativo qualquer diferença menor do que ,005, e como significativos valores de significância menores do que ,0083 (em função dos 6 testes possíveis de realização entre os quatro momentos temporais). Cabe dizer, entretanto, que o valor de p aqui apresentado para a plosiva velar ( $p=,36$ ) se mostra bastante distinto daquele valor marginalmente significativo referente à plosiva labial ( $p=,18$ ), fato esse que merece, também, consideração por parte do leitor.

A Tabela 6, a seguir, 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:** \* < ,05, \*\* p < ,01, \*\*\*p < ,001.

**Fonte:** Elaborada pelo autor.

Voltando à Tabela 4, vemos que, para /p/, o GE traz a média de 25,39 ms (mediana de 29,17 ms) no Pré-teste, enquanto o GC registra 17,56 ms (mediana de 15,42 ms) durante a mesma etapa, o que pode explicar a diferença significativa, apresentada na tabela anterior, desde o começo do experimento (no Pré-teste), o que não era o objetivo do trabalho. 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 já no pré-teste.

Em suma, os resultados até aqui apresentados nos permitem responder, preliminarmente, à primeira questão norteadora, que indagava o que as abordagens inferenciais e descritivas podem dizer a respeito do desenvolvimento dos dois grupos. Como dito no começo desta seção, resultados de estudos anteriores (KAMPPFF; ALVES, 2016; ALVES, LUCHINI, 2017) já nos proporcionaram o possível resultado de que possíveis melhoras obtidas na percepção não seriam imediatamente generalizadas na produção. Ao adicionarmos a quarta etapa, podemos visualizar, no

que tange à percepção, que o aumento inferencial significativo obtido na primeira fase do experimento, em 2015, manteve-se até a segunda fase, em 2018, o que não acontece com a produção. Porém, quando nosso olhar recai na análise descritiva, vemos que há um aumento, mesmo que tímido, na produção.

Vamos, então, partir da análise de grupos para a análise proposta por Lowie & Verspoor (2015), que leva em consideração a trajetória individual de cada participante durante o experimento.

## **4.2 ANÁLISES INDIVIDUAIS**

À luz de Lowie e Verspoor (2015), baseamos nossa segunda questão norteadora (o que uma verificação de dados individuais pode revelar sobre as trajetórias dos aprendizes?), pois, segundo os autores, "generalizações não são garantias de desenvolvimento, já que as redes dinâmicas de variáveis que afetam a trajetória desenvolvimental não podem ser equacionadas com a variação interindividual de um momento no tempo"<sup>27</sup> (*op. cit.*, p. 69). Isso nos leva a crer que as trajetórias individuais sejam dignas de uma análise que nos permita destacar os possíveis efeitos que o treinamento perceptual tem em cada participante.

Novamente a seção será dividida em duas subseções, uma referente aos dados de identificação, e outra referente à produção.

### **4.2.1 Análise individual dos resultados de identificação**

Nesta seção, traremos, sob a forma de gráficos, os resultados referentes à identificação de dois padrões VOT: o Zero Natural e o Zero Artificial. Os padrões Negativo e Positivo, como mostrado em estudos anteriores (ALVES; MOTTA, 2014; MOTTA, 2014; 2017; ALVES, ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015; RODRIGUES, 2015; KAMPFF; ALVES, 2016), se apresentam menos problemáticos de serem identificados pelos aprendizes. Nosso trabalho, inclusive, como mostrado nos resultados trazidos na Tabela 1, na seção anterior, corrobora tais estudos. Por esse motivo, apenas os dois padrões de VOT Zero serão discutidos. Primeiramente, traremos e analisaremos o padrão VOT Zero Natural do GE, seguido

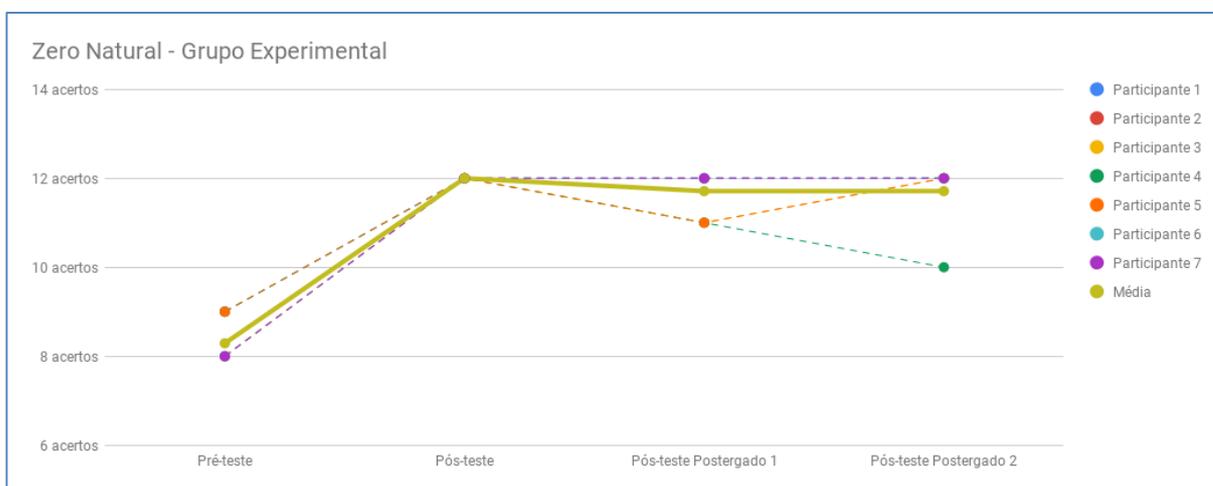
---

<sup>27</sup> Tradução do autor.

do GC. Após isso, os gráficos referentes ao padrão Zero Artificial, também na ordem de GE e depois GC, serão apresentados.

Antes de que analisemos os gráficos, é pertinente que retomemos alguns pontos já mencionados na análise de grupos, na seção 4.1.1. Primeiro, cabe mencionar o fato de que os resultados foram estatisticamente significativos nos padrões VOT Zero Natural e Zero Artificial, de acordo com os testes de Friedman para o Grupo Experimental. Além disso, no primeiro padrão de VOT, com os testes de Wilcoxon, a diferença foi marginalmente significativa entre as etapas Pré-teste e Pós-teste, Pré-teste e Pós-teste Postergado 1, e Pré-teste e Pós-teste Postergado 2. No segundo, foram encontradas diferenças marginalmente significativas entre o Pré-teste e o Pós-teste Postergado 2. O fato de as diferenças dos testes *post-hoc* terem sido todas marginais (muito provavelmente, pelo baixo número de participantes) motivam, justamente, uma análise mais individualizada, conforme já argumentamos anteriormente.

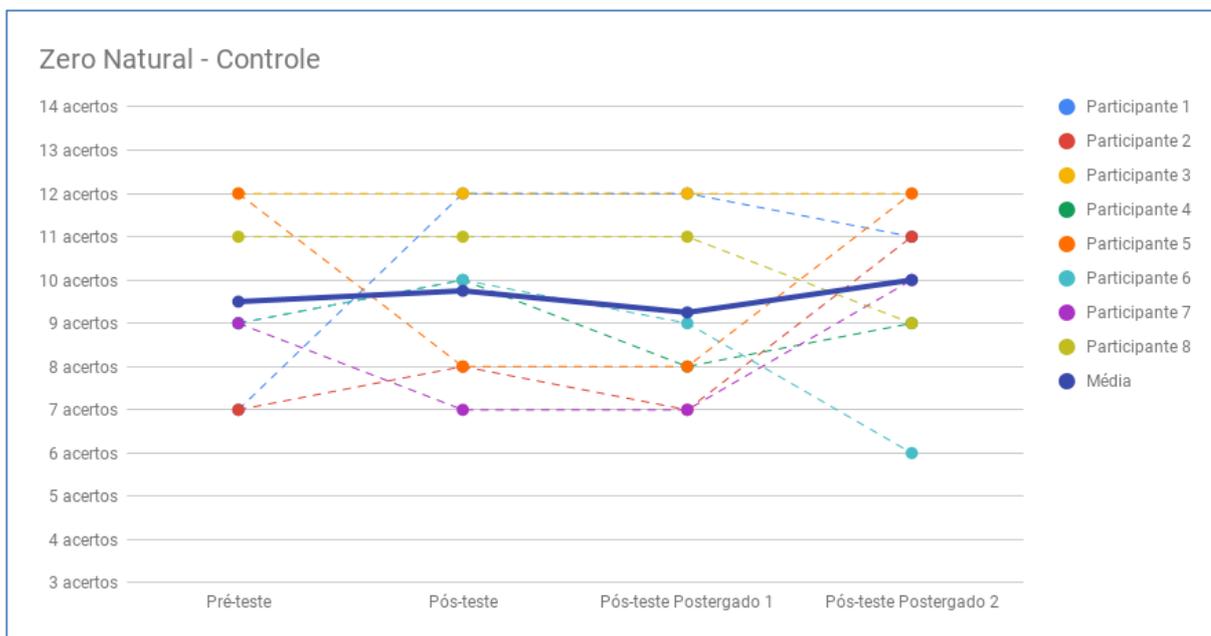
**Gráfico 1.** Número de acertos, por parte dos participantes do Grupo Experimental, na Tarefa de Identificação referente ao padrão VOT Zero Natural<sup>28</sup>, em cada uma das etapas do estudo.



**Fonte:** Elaborado pelo autor.

<sup>28</sup> Foram consideradas "acertos" as identificações de tal padrão como instâncias de consoantes sonoras.

**Gráfico 2.** Número de acertos, por parte dos participantes do Grupo Controle, na Tarefa de Identificação referente ao padrão VOT Zero Natural, em cada uma das etapas do estudo.



**Fonte:** Elaborado pelo autor.

Como destacado no começo desta seção, Lowie e Verspoor (2015) trazem a importância de olharmos para as trajetórias individuais dos participantes, separando essas trajetórias da trajetória de grupo. Uma primeira observação nos gráficos 1 e 2 já nos permite concordar com tal afirmação, pois neles fica claro como diversas trajetórias individuais (em pontilhado) seguem padrões diferentes da traçada pelas médias do grupo (linha contínua).

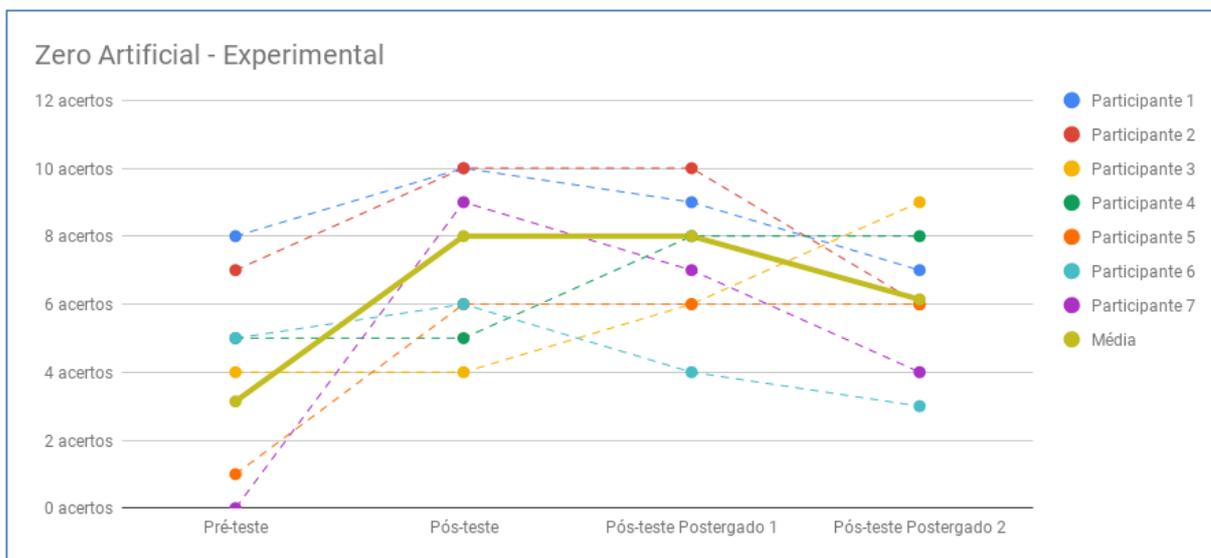
Em relação ao VOT Zero Natural, vemos que todos os participantes do GE aumentaram o número de acertos entre o Pré-teste e o Pós-teste, variando entre 8 e 9 para 10 e 12, sendo 12 o máximo de acertos possíveis. Como vimos através dos testes de Wilcoxon (Tabela 2), diferenças marginalmente significativas (que, por serem marginais, não levam a uma conclusão sólida, e motivam a verificação individual) entre cada uma das etapas e o Pré-teste foram verificadas, com a maioria dos participantes (1, 2, 4, 6 e 7) mantendo o máximo de 12, e o restante (3 e 5) variando entre 10 e 11 acertos. Isso ressalta bem o fato de que nenhum dos participantes que recebeu o treinamento perceptual terminou o experimento com menos acertos do que no começo (etapa de Pré-teste).

Em contrapartida, o GC mostrou um comportamento bem diferente do GE. Apesar de apenas os participantes 8 e 9 trazerem índices de acertos relativamente

baixos (7), o restante dos participantes trouxe, no Pré-teste, acertos tão altos, às vezes até mais, quanto os do GE. Porém, esses valores não foram mantidos durante as etapas seguintes, com exceção dos participantes 10 e 12, que terminaram o experimento com o mesmo número de acertos com que começaram. O participante 10, por sua vez, foge completamente à média trazida no Gráfico 2, pois mantém seus acertos no índice máximo durante as quatro etapas. Vale lembrar que, de acordo com os resultados do Questionário de Experiência e Proficiência Linguística de Scholl e Finger (2013) (ver Anexo B), o participante soma 14 horas por dia de contato com a língua inglesa (com a média do grupo sendo 7,7h e o desvio-padrão, 3,4h), além de ter começado a aprender a língua aos 6 anos (com a média do grupo sendo 9,75 anos e o desvio-padrão, 2,25 anos), frequentando por 16 anos curso ou escola em que a mesma é falada (sendo a média do grupo de 8,57 anos e o desvio-padrão, 5 anos). Este pode ser um fator importante para a sensibilidade do participante com o inglês. Além disso, tanto o participante 13 quanto o 15 apresentaram número de acertos no Pós-teste Postergado 2 (participante 13, 6 acertos, e participante 15, 9 acertos) menor do que no Pré-teste (participante 13, 9 acertos, e participante 15, 11 acertos), o que não houve no GE.

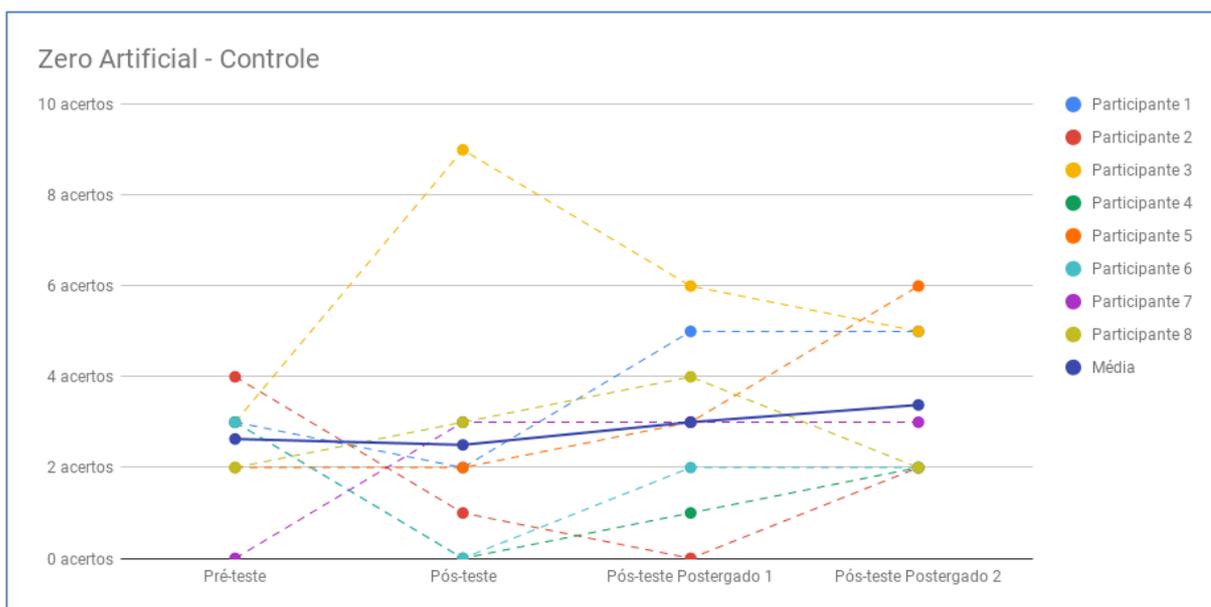
A seguir, olhemos como ambos os grupos se comportaram em relação ao padrão VOT Zero Artificial. Os valores estão presentes nos gráficos a seguir.

**Gráfico 3.** Número de acertos, por parte dos participantes do Grupo Experimental, na Tarefa de Identificação referente ao padrão VOT Zero Artificial, em cada uma das etapas do estudo.



Fonte: Elaborado pelo autor.

**Gráfico 4.** Número de acertos, por parte dos participantes do Grupo Controle, na Tarefa de Identificação referente ao padrão VOT Zero Artificial, em cada uma das etapas do estudo.



Fonte: Elaborado pelo autor.

Novamente, cabe-nos destacar, aqui, como as trajetórias individuais, representadas pelas linhas pontilhadas, se destacam da trajetória traçada, na linha

contínua, pelas médias do grupo. Como vemos a partir dos gráficos, os participantes do GE seguiram o mesmo padrão apresentado nos resultados do VOT Zero Natural; com exceção dos participantes 3 e 4, todos os outros aumentaram o número de acertos para o padrão Zero Artificial entre o Pré-teste e o Pós-teste. Inclusive, destacamos o participante 7, que não obteve nenhum acerto na primeira etapa, após a primeira sessão de treinamento perceptual, chegando ao Pós-teste com 9 acertos, de modo a mostrar a possível desestabilização que o seu sistema em desenvolvimento sofreu. Já entre o Pós-teste e o Pós-teste Postergado 1, os participantes mostram comportamentos bem diversificados, com 2 deles (participantes 2 e 5) mantendo o número de acertos, 2 aumentando (participantes 3 e 4) e 3 diminuindo (participantes 1, 6 e 7); ainda que esses três últimos citados tenham diminuído seus índices de acerto, a média do grupo se mantém entre estas duas etapas (8 acertos), o que ressalta a afirmação de que a mesma não expressa o comportamento de todos os participantes. Entre as etapas do Pós-teste Postergado 1 e 2, todos os participantes mostram menos acertos entre elas, com exceção do participante 3, que aumentou, e dos participantes 4 e 5, que mantiveram os mesmos índices de acertos. Outro fato relevante, também, é como, com exceção dos participantes 1, 2 e 6, todos apresentam o valor de acertos no Pós-teste Postergado 2 maiores do que os iniciais (do Pré-teste).

Já no GC, sem receber o treinamento, tal grupo não mostrou nenhum tipo de padrão ou mudança que pudesse ser considerado significativo, para as estatísticas inferenciais, tampouco um padrão característico nos números individuais. De acordo com o gráfico, apenas o participante 10 chegou a 9 acertos (Pós-teste), enquanto todos os outros mantiveram-se em uma média de 2, 3 e 4, com alguns raros casos chegando a 6 (novamente, o participante 6, no Pós-teste Postergado 1, e o participante 12 no Pós-teste Postergado 2).

Também vemos que quatro dos participantes, os de números 8, 10, 12 e 14, ou seja, metade do total, terminam o experimento com valores acima dos valores iniciais. Porém, a diferença entre esses valores é pouca (variando um ou dois acertos), o que pode ter contribuído com a falta de significância nas estatísticas.

Em suma, pudemos ver que todas as trajetórias individuais (linhas pontilhadas) fugiram da trajetória das médias mostradas nos gráficos (linhas contínuas). Inclusive, em linhas gerais, podemos ver que as trajetórias apontam para maiores índices de crescimento individual entre os participantes do GE do que os do GC.

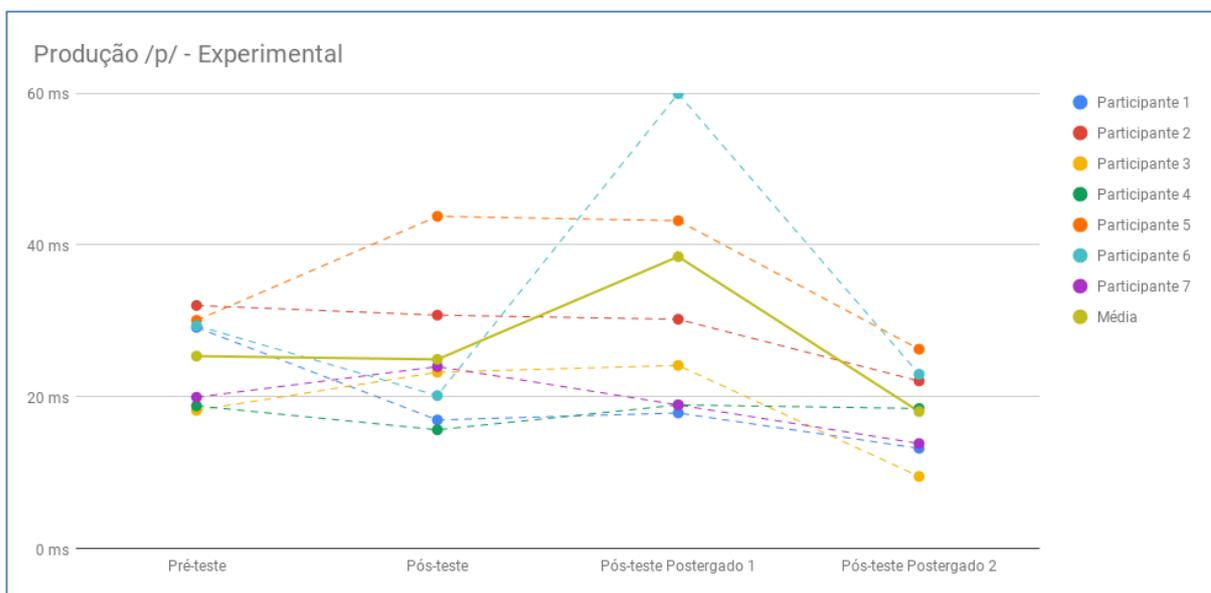
#### 4.2.2 Análise individual dos resultados de produção

Seguindo a organização anterior, traremos gráficos referentes aos valores de VOT de cada uma das plosivas surdas do inglês - /p/, /t/ e /k/ -, com uma visão geral dos valores das médias de cada participante. Após isso, destacaremos mais informações, tal como o desvio padrão de cada participante, de acordo com o que for mais pertinente para a nossa análise. A tabela mais detalhada, com todos os valores (valores mínimos, máximos, médias e desvios-padrão), pode ser encontrada no Anexo C.

##### 4.2.2.1 Análise individual dos resultados de produção para /p/

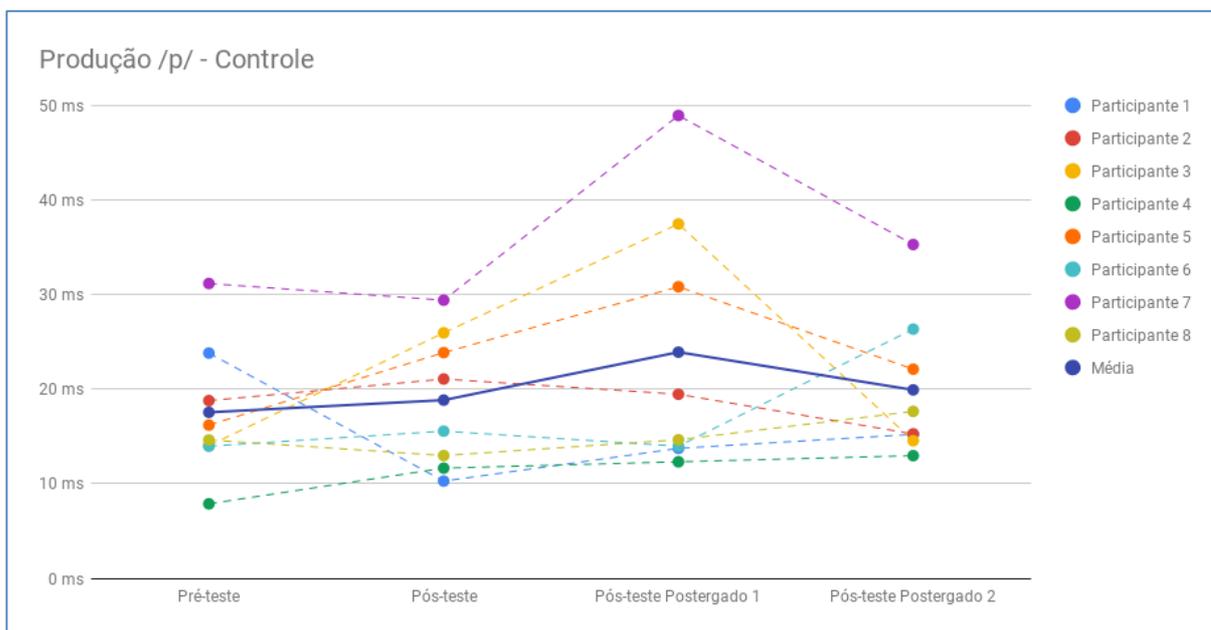
A seguir, encontraremos os seguintes gráficos: primeiramente, apresentamos as médias de produção para /p/ do GE, seguido do GC. Após isso, descreveremos as médias de /t/, sendo primeiramente apresentados os dados do GE e, posteriormente, os do GC. Finalmente, apresentaremos os valores para /k/, ainda na mesma ordem dos grupos.

**Gráfico 5.** Médias individuais referentes às produções de /p/ dos participantes do Grupo Experimental durante as etapas do estudo.



Fonte: Elaborado pelo autor.

**Gráfico 6.** Médias individuais referentes às produções de /p/ dos participantes do Grupo Controle durante as etapas do estudo.



Fonte: Elaborado pelo autor.

Retomando o que trouxemos na seção anterior, é válido destacar que os testes de Friedman (Tabela 4) mostraram diferença significativa para a consoante /p/ no GE. A possibilidade de tal significância parece residir no fato de que os valores da última etapa, Pós-teste Postergado 2, estão todos abaixo dos valores trazidos no Pré-teste,

antes de os estudantes participarem do experimento (25,39 ms no Pré-teste e 18,09 no Pós-teste Postergado 2). Inclusive, lembrando os resultados do teste de Wilcoxon (Tabela 5), foi apontada uma diferença marginalmente significativa justamente entre as duas últimas etapas do experimento; assim, os alunos demonstraram uma diminuição considerável nos valores das médias entre estas etapas. Especulamos que uma das explicações para a mudança de valor seja uma possível estabilização da língua entre o Pós-teste Postergado 1 (em 2015) e o 2 (em 2018), como trazido na seção 4, durante a análise de grupos.

São destacáveis, também, os resultados de produção do participante 6 no Pós-teste Postergado 1 (média de 59,93 ms), com um desvio-padrão de 56,63 ms, o que sugeriria uma maior variabilidade. Tal variabilidade pode nos indicar duas coisas. A primeira diz respeito a uma possível desestabilização pela qual o sistema linguístico do participante passa, considerando o "salto" da média de 20,19 ms (com desvio padrão 2,8, o que nos indicaria menor variabilidade) do Pós-teste para os 59,93 ms do Pós-teste Postergado 1. A segunda diz respeito ao fato de que esta diferença entre as médias faz com que a própria média do grupo apresente um aumento considerável, de 24,96 ms no Pós-teste para 38,48 ms no Pós-teste Postergado 1 (com os desvios-padrão aumentando de 9,69 ms para 15,77 ms), já apresentado na Tabela 4.

Já em relação ao GC, é destacável a trajetória do participante 13, que segue o caminho contrário ao mostrado para a média geral entre as duas etapas finais do experimento. Enquanto a média do grupo apresenta uma diminuição do valor de 23,92 ms no Pós-teste Postergado para 19,93 ms no Pós-teste Postergado 2, o referido participante aumenta de 13,98 ms para 26,35 ms. Ao analisarmos os desvios-padrão deste participante em cada etapa mencionada (2,26 e 23,77, respectivamente, conforme o Anexo C), também podemos verificar a variabilidade do sistema do participante. Tais resultados corroboram o que foi especulado ao longo do trabalho, referente à suposição de que o GC, por não ter participado do treinamento perceptual, em 2015, sofreu a desestabilização que leva à mudança de padrões de VOT mais tardiamente em sua trajetória desenvolvimental. Ainda sobre o participante 13, de acordo com os resultados do Questionário de Experiência e Proficiência Linguística, de Scholl e Finger (2013) (Anexo B), a média geral de horas por dia utilizando o inglês, no GC, é de 7,7; tal participante, por sua vez, utiliza o inglês por oito horas do seu dia, tendo estado em contato com o idioma por mais de 12 anos em escolas de língua

inglesa, sendo a média do grupo de 8,57 anos. Estes fatores mostram maior contato com a língua, o que podem afetar a experiência linguística do participante.

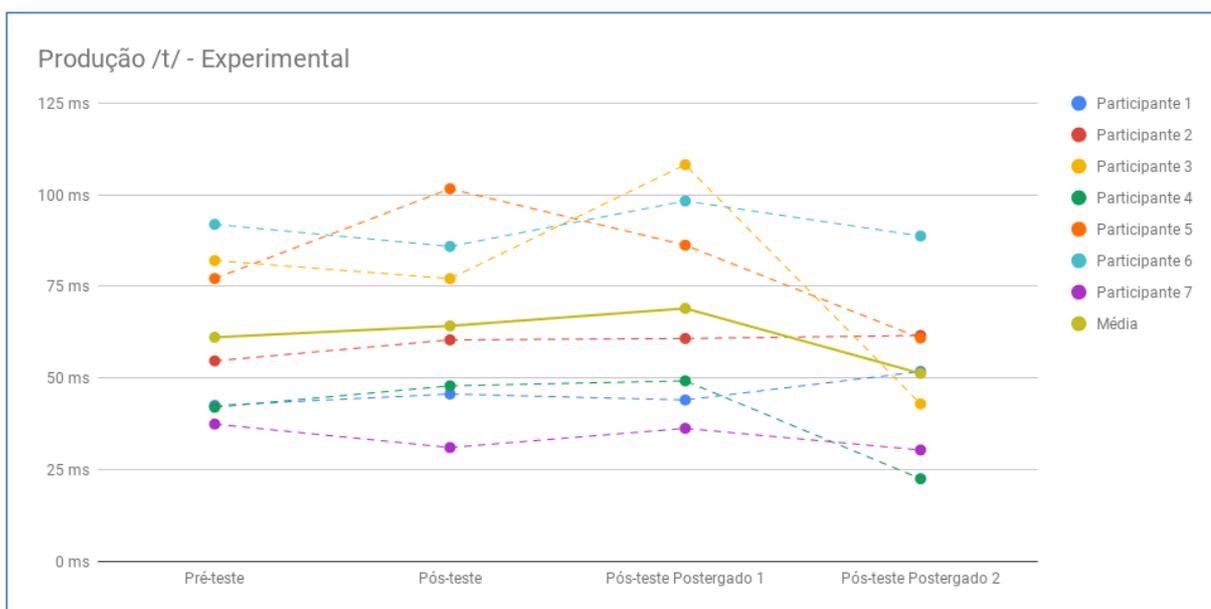
Outro resultado interessante de destacar é a do participante 14. Absolutamente em todas as etapas do experimento, seus valores de produção individuais (31,17 ms; 29,41 ms; 48,93 ms; 35,29 ms) estão acima dos mostrados pela média geral (25,39 ms; 24,96 ms; 38,48 ms; e 18,09 ms; cf. Tabela 4). No Pós-teste Postergado 1, o valor máximo do participante foi 47,75 ms. Aqui, a observação dos resultados do questionário de Scholl e Finger (2013) é preciosa, pois o participante traz alguns dos valores mais altos em relação ao tempo em escola ou trabalho em que se usa a L2 (13 anos, sendo a média do grupo 8,57 anos), percentagem de uso da língua (30% do dia, sendo a média do grupo 16,8%), e 9h/dia (sendo a média do grupo 7,7h/dia). Novamente, a maior experiência com a língua-alvo aparece refletida em participantes com trajetórias destacáveis. Finalmente, quando olhamos para o Gráfico 6, vemos que, nas duas últimas etapas, apenas três dos oito participantes (participantes 10, 12 e 14, no Pós-teste Postergado 1, e participantes 12, 13 e 15 no Pós-teste Postergado 2) do grupo estiveram acima da média geral, sendo, possivelmente, esses os que ajudaram a aumentá-la.

Em suma, é importante ressaltar que, por questões aerodinâmicas, o VOT de /p/ é menor que o de /t/ e /k/, resultando em menor possibilidade de variação na produção. Entretanto, ao fazermos uma análise em linhas gerais, vemos que o GE chega na última etapa do experimento com seus valores aparentemente mais estabilizados do que os do GC. Novamente, tal fato pode se dar justamente pelo treinamento, que se mostrou útil para esta consoante, pelo qual o primeiro grupo passou, causando a desestabilização do sistema durante as três primeiras etapas, no qual houve o crescimento dos valores de média e desvios-padrão, e a diminuição dos mesmos na quarta etapa, quando o sistema dos participantes estaria em rumo a uma estabilização. Já o contrário acontece com o GC. Por não ter participado do treinamento, em 2015, seu desenvolvimento vai se dar mais tardiamente, o que explicaria os valores maiores na última etapa do projeto, em relação ao GE.

#### 4.2.2.2 Análise individual dos resultados de produção para /t/

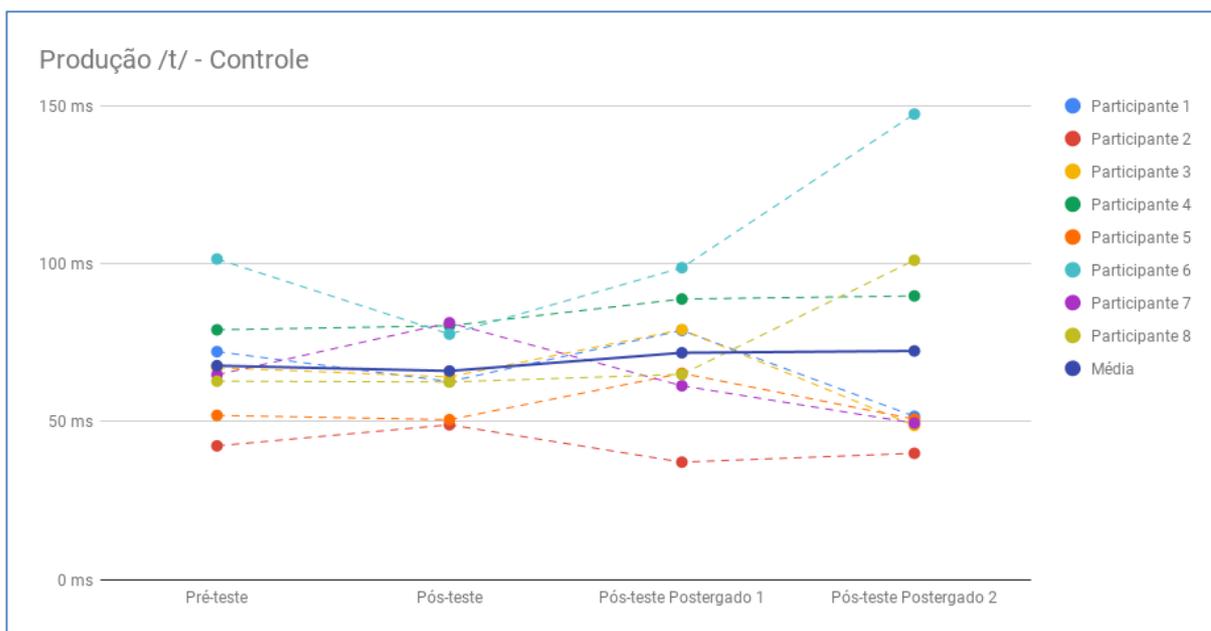
Seguindo a mesma organização para /p/, traremos os gráficos referentes às trajetórias dos participantes em relação à consoante /t/, destacando dados individuais mais detalhados, presentes no Anexo C, de acordo com o que for pertinente à análise.

**Gráfico 7.** Médias individuais referentes às produções de /t/ dos participantes do Grupo Experimental, durante as etapas do estudo.



Fonte: Elaborado pelo autor.

**Gráfico 8.** Médias individuais referentes às produções de /t/ dos participantes do Grupo Controle, durante as etapas do estudo.



**Fonte:** Elaborado pelo autor.

Em relação ao /t/, cabe-nos lembrar que as estatísticas inferenciais não mostraram nenhuma significância, tanto no GE quanto no GC. Porém, como destacado, em termos descritivos, nas primeiras três etapas do experimento, os participantes do GE passam de 61,09 ms, para 64,21 ms, e, posteriormente, para 68,99 ms. Na quarta etapa, há uma diminuição na média (51,23 ms). Considerando esses resultados, vamos analisar as trajetórias individuais de participantes que se distanciam da trajetória do grupo.

Uma das trajetórias destacáveis do GE é a do participante 3. Ao analisarmos os desvios-padrão do participante entre Pré-teste (18,82 ms) e Pós-teste (28,84 ms), vemos uma clara diferença na variabilidade, mesmo com a média caindo de 82,06 ms para 77,12 ms entre elas. Já entre as etapas do Pós-teste (média individual de 77,12 ms) e do Pós-teste Postergado 1 (média individual de 108,23 ms), há um salto considerável; ao analisarmos os desvios-padrão dessas etapas (28,84 ms e 14,56 ms, respectivamente), podemos especular que uma possível estabilização esteja em ocorrência, considerando o menor desvio-padrão.

O participante 2 também apresenta uma trajetória interessante de ser analisada, sendo esse o único a terminar o experimento com um crescimento, ainda

que tímido, nas médias individuais em cada uma das etapas (54,64 ms; 60,39 ms; 60,76 ms; 61,59 ms). Mais destacável ainda é o salto dos desvios-padrão do participante entre Pré-teste e Pós-teste (de 16,5 ms para 22,2 ms), que se manteve no Pós-teste Postergado 1 (22,04 ms) e cai no Pós-teste Postergado 2 (13,13 ms). Esses valores mostram uma possível desestabilização do sistema da primeira para a segunda etapa, que, aparentemente, se manteve na terceira e, na quarta, diminuiu, corroborando a especulação da seção anterior de que na última etapa do experimento alguns participantes já haviam estabilizado seus sistemas depois da desestabilização inicial, possivelmente causada pelo treinamento perceptual.

O participante 5 também apresenta uma possível desestabilização ao aumentar a média entre Pré-teste (77,14 ms) e Pós-teste (101,66 ms); essa mudança fica mais evidente no grande salto dos índices individuais de desvio-padrão entre as mesmas etapas (16,96 ms e 50,14 ms, respectivamente), o que parece ser evidência dos efeitos do treinamento. Porém, a possível desestabilização não aparenta manter-se nas etapas seguintes, sendo os desvios-padrão 4,18 ms e 1,66 ms no Pós-teste Postergado 1 e Pós-teste Postergado 2.

O GC, por outro lado, apresenta um comportamento levemente diferente. Três dos participantes (11, 13 e 15) apresentam valores individuais de média no Pós-teste Postergado 2 maiores do que no Pré-teste. Dentre eles, os participantes 11 e 15 são destacáveis.

O participante 11 segue o mesmo padrão atestado no GE. Ao olharmos seus desvios-padrão, vemos um aumento de 17,47 ms do Pré-teste para 18,98 ms no Pós-teste, subindo para 30,86 ms e caindo para 25,34 ms. Com esses números, podemos especular que o fato de o valor do desvio-padrão ser maior no Pós-teste Postergado 2 do que no Pré-teste representa uma continuidade na possível desestabilização do sistema linguístico do estudante. Em outras palavras, o aprendiz continua em pleno estado de desenvolvimento. Isso acontece mais tardiamente, em comparação aos participantes do GE, justamente porque tal participante não sofreu nenhuma intervenção, de modo que seu desenvolvimento esteja sendo advindo do contato do aprendiz com a língua em uso.

Inclusive, o comportamento semelhante do participante 11 aparece no participante 15. Seus desvios-padrão são de 15,28 ms (Pré-teste), 28,49 ms (Pós-

teste), 16,77 ms (Pós-teste Postergado 1) e 18,68 ms (Pós-teste Postergado 2). Novamente, o aumento entre Pré-teste e Pós-teste, que é surpreendente pelo fato de o participante ser o grupo de controle, ou seja, sem passar pelo treinamento, em 2015, possivelmente representa uma desestabilização descritivamente significativa. O fato de o desvio-padrão do Pós-teste Postergado 2 ser maior do que o do Pré-teste pode representar uma continuidade na desestabilização em relação ao início do experimento. Tal fato não se mostra surpreendente, uma vez que os participantes do GC, por não terem seu desenvolvimento "acelerado" pelo treinamento, podem passar por um período ainda mais longo e tardio de desestabilização.

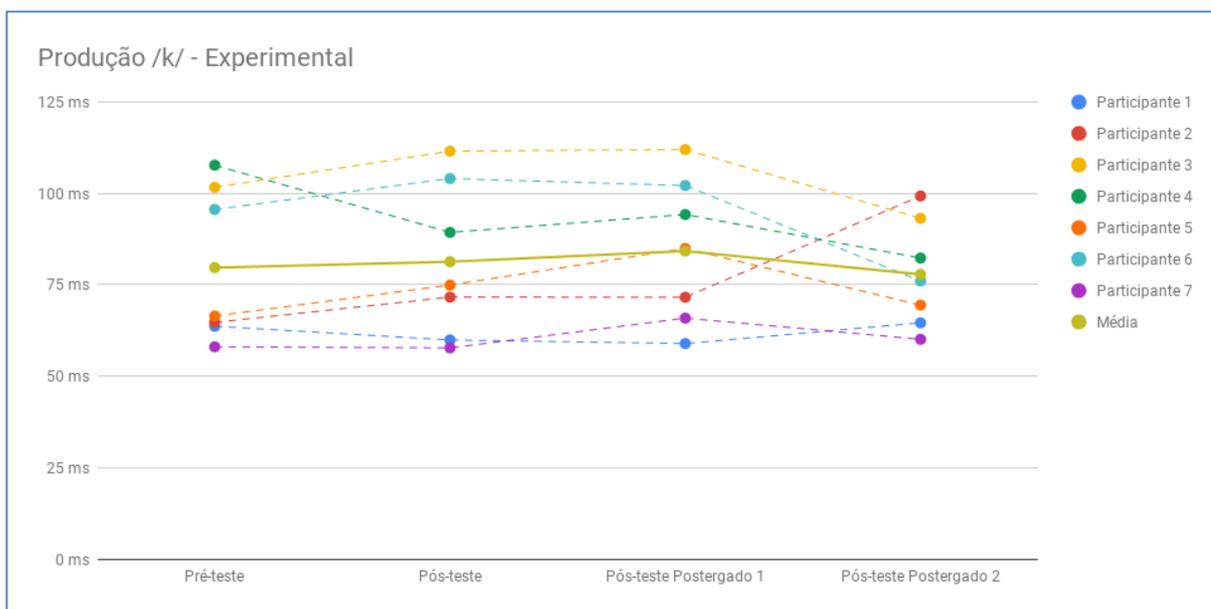
Em suma, no caso do /t/, a instabilidade entre participantes é ainda maior do que em /p/ (talvez por tal consoante inerentemente apresentar valores mais altos de VOT e, por isso, possibilitar diferenças ainda maiores entre os participantes). Estas diferenças individuais reforçam, mais uma vez, a análise individual, uma vez que a média de grupo parece não expressar o que está acontecendo com os participantes.

Além disso, ainda que as estatísticas inferenciais não tenham apresentado resultados significativos para /t/, parece que o treinamento contribuiu para instaurar uma instabilidade mais precoce entre os participantes do GE. Esta instabilidade nos dados, que também denota o início e a intensificação do desenvolvimento linguístico, também pode ser encontrada entre participantes do GC; entretanto, no caso destes participantes, tal instabilidade tende a aparecer em momentos temporais mais tardios, justamente pelo fato de tais participantes não terem tido seus processos de desenvolvimento de VOT acelerados pelo vozeamento.

#### 4.2.2.3 - Análise individual dos resultados de produção para /k/

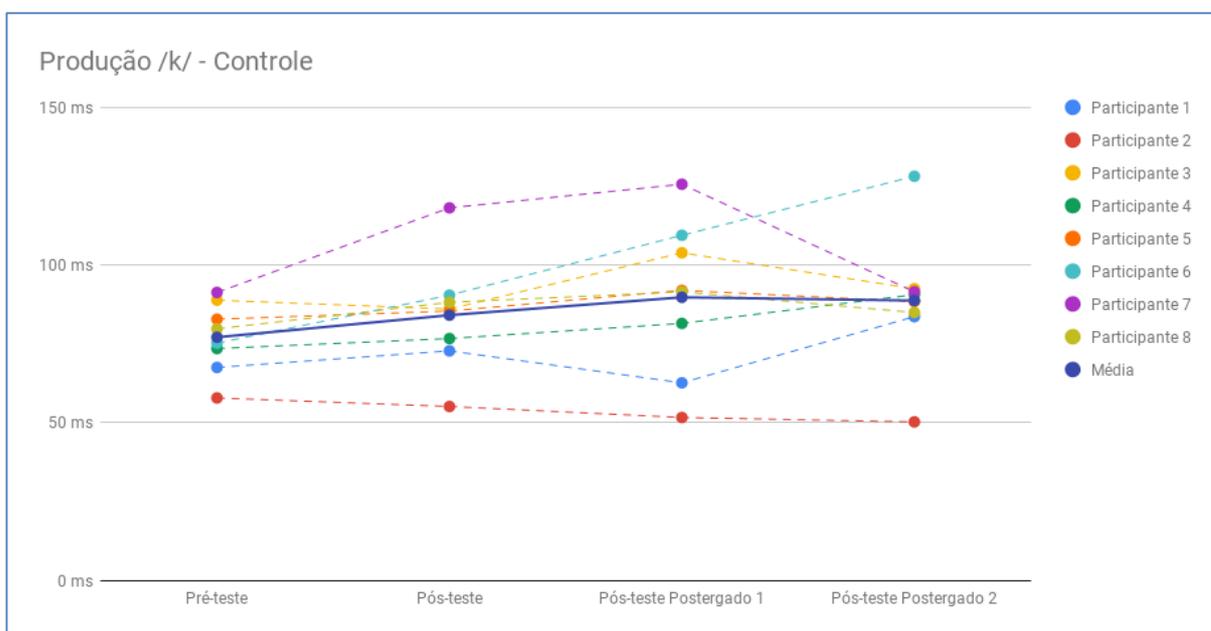
A seguir, traremos as trajetórias individuais da consoante /k/, seguindo a organização de apresentar os gráficos do GE e GC, para depois partirmos para a análise.

**Gráfico 9.** Médias individuais referentes às produções de /k/ dos participantes do Grupo Experimental, durante as etapas do estudo.



Fonte: Elaborado pelo autor

**Gráfico 10.** Médias individuais referentes às produções de /k/ dos participantes do Grupo Controle, durante as etapas do estudo.



Fonte: Elaborado pelo autor.

Como trazido na seção anterior, o comportamento do GE, em relação à consoante /k/, na média das três primeiras etapas, é de 79,7 ms (Pré-teste), 81,31 ms

(Pós-teste) e 84,24 ms (Pós-teste Postergado 1). Já na quarta etapa, há uma diminuição na média, sendo esta 77,85 ms. Apesar destas médias, quando olhamos para os desvios-padrão (Tabela 4), principais atestadores de uma possível variabilidade, maiores informações podem ser inferidas. Além disso, quando olhamos para os dados individuais, vemos que é nessa consoante que os dados individuais se mostram mais diferenciados da média do grupo, talvez, justamente, pelo fato de tal ponto de articulação implicar valores mais altos de VOT e, por conseguinte, maior possibilidade de variação.

Participantes como os de número 1 e 6 aumentam seus desvios-padrão de 7,49 ms e 9,25 ms para 16,33 ms e 21,34 ms, respectivamente, entre as duas primeiras etapas (ou seja, após o treinamento perceptual, que ocorreu entre as mesmas, como trazido na seção de Método). Curiosamente, quando vemos a média do participante 1, nas mesmas etapas, há uma diminuição de 63,67 ms para 59,91 ms. Ou seja, mesmo com a diminuição do valor médio, há um aumento no desvio padrão, o que indica uma possível variabilidade, o que também denota desestabilização do sistema linguístico, conforme será discutido com maior profundidade ao respondermos à terceira Questão Norteadora. Cabe, ainda, mencionar que ambos os participantes vão mostrar menores valores de desvio padrão nas etapas seguintes, rumo à estabilização do sistema linguístico.

Já o participante 3 apresenta o mesmo comportamento trazido pelo grupo, com as médias de 101,68 ms (Pré-teste), 111,57 ms (Pós-teste), 112 ms (Pós-teste Postergado 2) e 93,18 ms (Pós-teste Postergado 2). Porém, ao analisarmos os desvios-padrão, vemos os seguintes índices para as quatro etapas, respectivamente: 14,49 ms; 22,48 ms; 16,84 ms; e 50,91 ms. Como podemos ver, com o valor alto do desvio-padrão do Pós-teste Postergado, é possível que a principal desestabilização do sistema do participante não tenha acontecido durante a primeira parte dos testes, mas anos depois. O mesmo acontece com o participante 4, que apresenta o desvio-padrão de 37 ms na última etapa, contrastando com os das três primeiras (17,84 ms; 16,65 ms; e 24,69 ms). Isso se deve, talvez, ao fato de que, desde o Pré-teste, tais participantes já produziam padrões de VOT para /k/ semelhantes, ou até superiores, aos índices produzidos por falantes nativos. Dessa forma, o treinamento não necessitaria causar uma desestabilização que levasse à produção das formas-alvo. Por outro lado, a desestabilização que agora ocorre pode, inclusive, ser uma maneira

de "aproximar" os padrões de VOT dos aprendizes aos valores nativos, uma vez que, como no caso do participante 3, as produções realizadas tendiam inclusive a superar aqueles índices previstos pela literatura para falantes nativos do inglês. Maiores detalhes sobre a variabilidade dos índices de VOT dos participantes serão fornecidos em breve, ao abordarmos a Questão Norteadora 3.

Em relação ao GC, o mesmo apresentou, de forma surpreendente, uma diferença estatística inferencial marginalmente significativa ( $p=.098$ ) para os valores de /k/ (Tabela 4), sendo a diferença, também marginal, entre as etapas Pré-teste e Pós-teste Postergado 1. Tais valores se refletem nos valores descritivos das médias, quando o grupo começa com 77,10 ms no Pré-teste e chega a 89,76 ms no Pós-teste Postergado 1.

A trajetória do participante 13 se torna uma das mais destacáveis, pois, concordando com as médias do grupo, seus índices médios individuais são de 75,22 ms (Pré-teste), 90,44 ms (Pós-teste), 109,42 ms (Pós-teste Postergado 1) e 128,14 ms (Pós-teste Postergado 2). Porém, à medida em que suas médias foram aumentando, seus desvio-padrão foram diminuindo entre etapas (com exceção do desvio da última, havendo um salto nesta em relação à terceira, que apresentava uma queda bastante acentuada), sendo eles, respectivamente, 48,65 ms; 40,13 ms; 13,27 ms; e 33,81 ms. O que podemos verificar aqui é que este participante já iniciou o experimento com o sistema desestabilizado, mantendo a desestabilização na segunda etapa e, na terceira, mostrando um possível início de estabilização. Ao analisarmos o novo salto entre a terceira e a quarta etapa, podemos recorrer ao Questionário de Experiência e Proficiência Linguística, de Scholl e Finger (2013). O participante apresenta o segundo maior tempo de trabalho ou escola em que se fala a L2 (12,6 anos), oito horas por dia de contato com a L2 (enquanto a média do grupo é 7,7) e 20% de tempo de uso da L2 por dia (sendo a média do grupo 16,87%, cf. Anexo B). Novamente, a relação de envolvimento com a L2 parece ser um elemento importante ao analisarmos participantes que já iniciam o experimento com os sistemas desestabilizados. Comportamento semelhante vai se apresentar no participante 11. Enquanto suas médias aumentam ao longo das etapas do experimento (73,51 ms; 76,69 ms; 81,49 ms; e 90,6 ms), seus desvios-padrão começam em 20,75 ms no Pré-teste, diminuindo para 16,58 ms no Pós-teste, mantendo-se em 16,85 ms no Pós-teste Postergado 1 e aumentando para 23,03 ms no Pós-teste Postergado 2.

Em suma, temos que considerar que, de todas as consoantes, essa é a que permite os maiores graus de diferenças entre participantes, e que justifica com mais clareza a necessidade de análises individuais. Ao contrário das outras consoantes, em função da semi-aspiração que a velar já sofre no próprio PB (cf. M. ALVES, 2015), os índices de VOT tendem a ser altos desde o Pré-Teste, o que justifica a não-verificação de diferenças significativas no GE. Ao mesmo tempo, as diferenças individuais mostram que alguns participantes se mostram com o sistema estabilizado desde o início do experimento, e outros, ainda que apresentem já valores nativos em suas produções, apresentam grande variabilidade (a ser discutida com mais profundidade ao tratarmos da Questão Norteadora 3). Há casos, inclusive, de participantes (3, 4 e 6) que iniciam com índices muito superiores aos nativos e que cujo processo de desenvolvimento, evidenciado, também, pela variabilidade, implica encontrar médias menores de VOT, que se aproximem às naturalmente produzidas por falantes nativos da língua.

Finalmente, fica claro que as respostas para a Questão Norteadora 2, que traz a análise das trajetórias individuais dos participantes e o que elas podem revelar, são amplas. Por exemplo, parece haver uma relação entre a variação na duração de VOT e ponto de articulação: quanto mais posterior o ponto, maior o valor de VOT, e, por conseguinte, maiores as possibilidades de diferentes padrões daquele encontrado para todo o grupo. É a partir da comparação entre os padrões individuais e de grupo que se justifica nossa análise: é notável que as trajetórias individuais dos aprendizes podem ser bastante diferentes daquela tomada como a de grupo.

Independentemente do ponto de articulação, verificamos que os diferentes padrões encontrados entre os aprendizes, bem como a variabilidade encontrada nos dados, podem ter sido um dos fatores que resultou na não verificação de diferenças estatisticamente significativas nas médias dos grupos como um todo. Apesar de a estatística inferencial não ter mostrado diferenças, verificamos que: (i) nos casos de alguns participantes, as diferenças entre Pré-Teste e Pós-Teste eram nítidas; (ii) mudanças verificadas entre os participantes do GE nas duas primeiras etapas de coleta puderam ser verificadas, mais tardiamente, entre os aprendizes do GC, sobretudo no que diz respeito à plosiva /t/. Tal fato serve como uma espécie de evidência de que, ainda que o desenvolvimento dos padrões de VOT compreenda uma etapa inegável da trajetória em direção à L2, o treinamento parece exercer o

papel de "acelerar" tal desenvolvimento daqueles que dele participam. Esta aceleração pode ser verificada não somente nas médias individuais, mas, também, nos próprios desvios-padrão individuais dos aprendizes. Ao considerarmos que a variabilidade exerce papel de evidência fundamental para o desenvolvimento linguístico, a Questão 3, ao se concentrar na análise dos valores de mediana, bem como na discussão dos valores máximos e mínimos de produção dos aprendizes, tentará, através da variabilidade, evidenciar adicionais para o papel do treinamento perceptual.

Em suma, quando relembremos a visão do desenvolvimento linguístico de uma L2 a partir da noção de Sistemas Complexos (BECKNER *et al.*, 2009; LARSEN-FREEMAN; CAMERON, 2008), principalmente buscando o olhar para o dado individual, como proposto por Lowie e Verspoor (2015; 2018), os resultados apresentados ficam mais claros. Nem sempre, enquanto grupo, os resultados inferenciais mostrar-se-ão significativos, mas uma análise mais individual, como a trazida neste trabalho, mostrará as possíveis mudanças que afetam os elementos do sistema do falante, assim como a reação às mesmas.

#### **4.3 SOBRE A VARIABILIDADE NAS PRODUÇÕES DOS APRENDIZES: INDÍCIOS DE DESENVOLVIMENTO LINGUÍSTICO**

Nesta seção, a fim de responder à terceira Questão Norteadora, apresentaremos a variabilidade referente aos índices de produção do grupo e, em seguida, a dos participantes, tomada individualmente. A seção será dividida na análise dos gráficos *boxplots* de /p/, /t/ e /k/ referentes aos grupos e, em seguida, a análise dos gráficos individuais para as mesmas consoantes. Aprofundaremos estas análises trazendo informações, como valores mínimos e máximos de produção de participantes específicos, que podem ser encontrados no Anexo C, visualizáveis através dos próprios gráficos a serem apresentados.

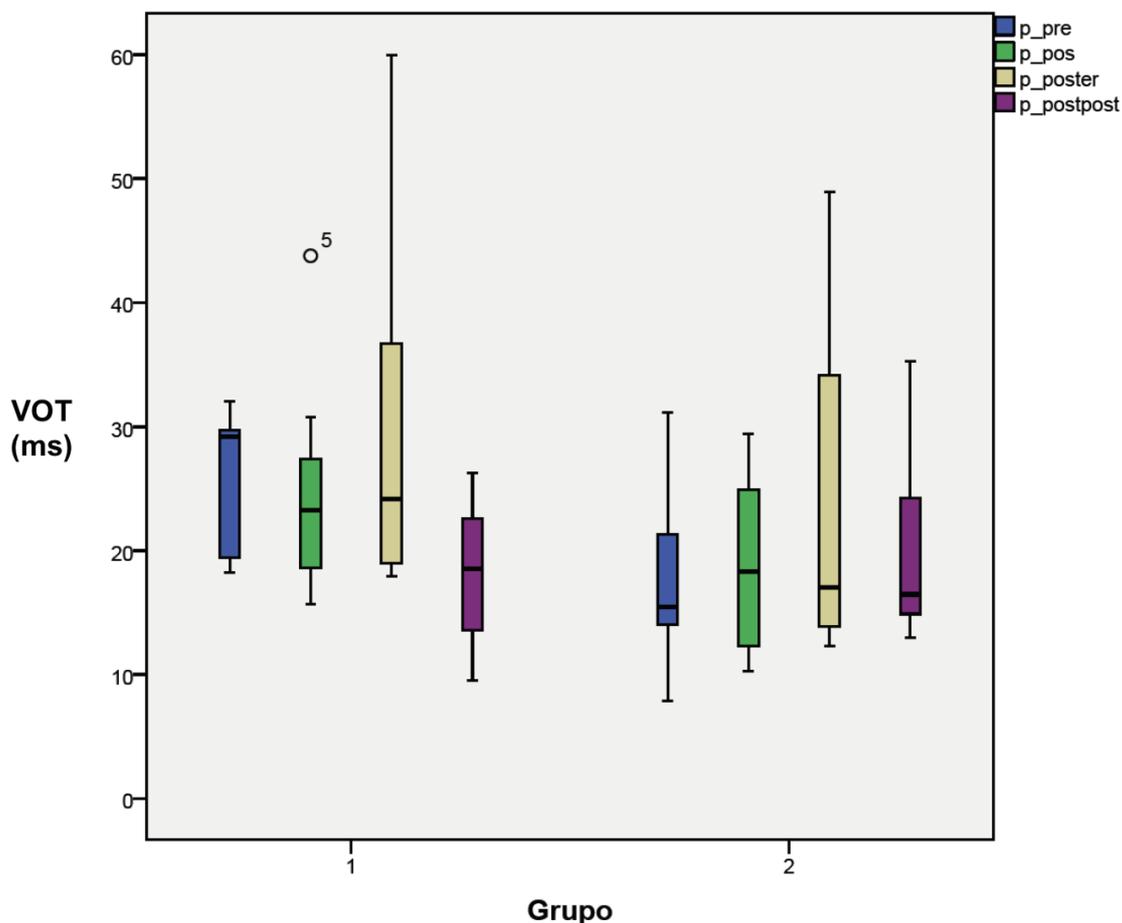
Como explicado na Introdução, focaremos somente na variabilidade da produção. A mesma, como já atestado ao longo do trabalho, tende a não mostrar resultados significativos em termos inferenciais, com apenas a vogal /p/, no GE, apresentando uma diferença significativa (Tabela 4). Por isso, esperamos que, através

da análise descritiva individual, encontremos algum padrão de comportamento que evidencie uma efetiva contribuição do treinamento entre os participantes.

#### 4.3.1 - Análise dos índices de variabilidade dos dois grupos

Nesta seção, cada uma das consoantes será analisada individualmente. A seguir, apresentaremos o *boxplot* referente à produção de /p/ por ambos os grupos - Experimental e Controle, em cada um dos quatro momentos de coleta de dados.

**Gráfico 11** - Variabilidade dos grupos, Experimental e Controle, referentes às produções de VOT da consoante /p/, em cada um dos quatro momentos de coleta de dados.



**Nota:** Grupo 1 = Grupo Experimental; Grupo 2 = Grupo Controle.

**Fonte:** Elaborado pelo autor.

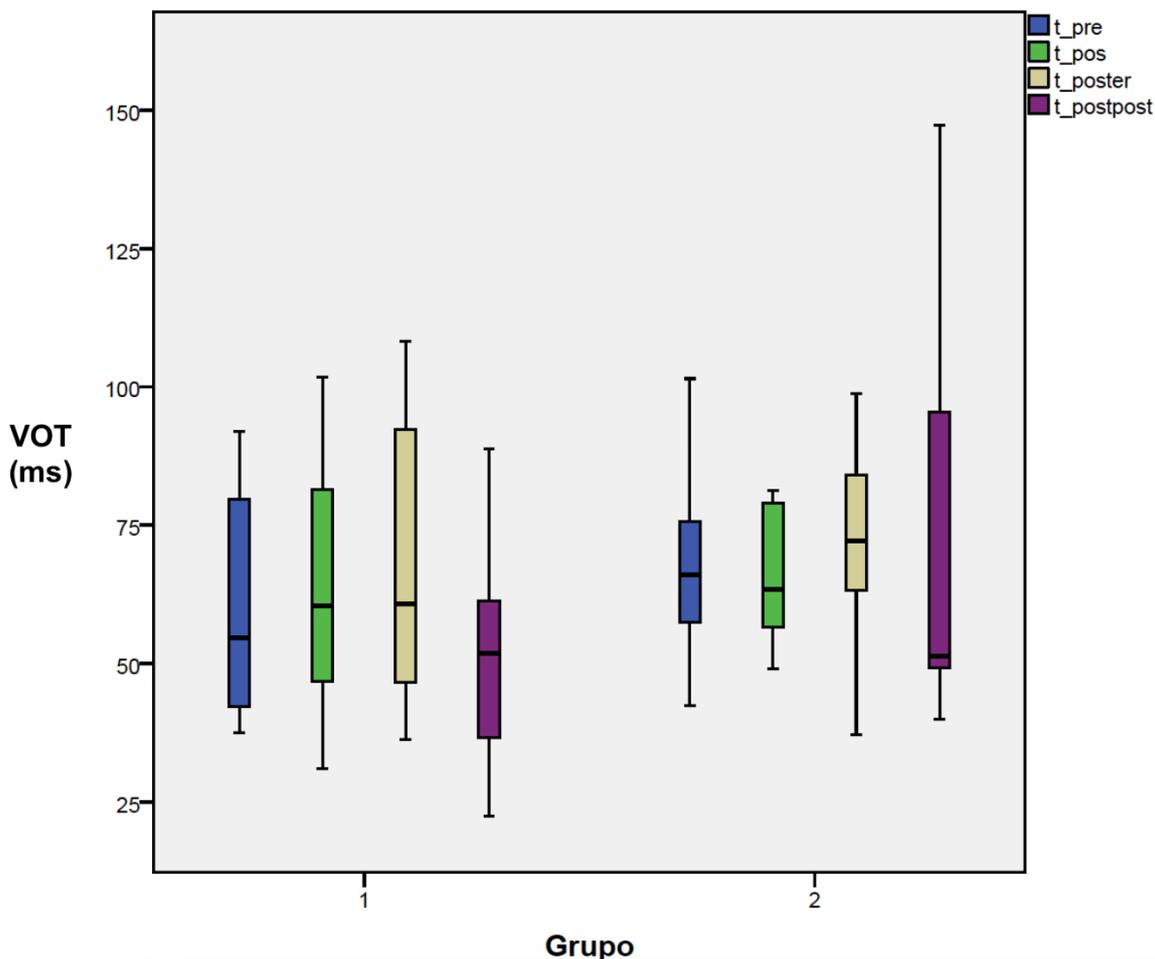
Como visto na seção 4.1, apenas a consoante /p/ mostrou diferença significativa em termos inferenciais (cf. Tabela 4) no GE; essa significância foi atribuída à diminuição da média entre as etapas de Pós-teste Postergado 1 e Pós-

teste Postergado 2. Tal significância pôde ser atestada na análise das estatísticas descritivas trazidas na seção 4.2.2. Considerando a variação entre o valor mínimo e máximo trazidos em cada etapa e apresentadas no gráfico acima, também podemos fazer inferências sobre a variabilidade dos grupos. Ao observarmos os valores de /p/, vemos que no Pós-teste e no Pós-teste Postergado 1, etapas logo após o treinamento, o GE apresenta a média mínima de 15,7 ms e máxima de 43,78 no Pós-teste, e mínima de 17,91 ms e máxima de 59,93 ms no Pós-teste Postergado 1, enquanto o GC apresenta valores mínimos de 10,29 ms e 12,33 ms e máximos de 29,41 ms e 48,93 ms para as mesmas etapas, respectivamente, que representam uma maior variabilidade do Grupo Experimental nestas etapas.

Outro dado interessante é o fato de os valores mínimos e máximos variarem mais no GC do que no GE nas etapas de Pré-teste (mínimo de 7,89 ms e máximo de 31,17 ms para o GC, e mínimo de 18,84 ms e máximo de 32,05 ms para o GE) e Pós-teste Postergado 2 (mínimo de 12,98 ms e máximo de 35,29 ms para o GC, e mínimo de 9,56 ms e máximo de 26,3 para o GE). Duas possíveis conclusões podem ser tiradas destes resultados. A primeira é a de que o GC, neste estudo, já começa o experimento com uma variabilidade maior do que o GE; portanto, o mesmo seria mais suscetível a uma desestabilização do sistema - no nosso caso, possibilitada pelo treinamento perceptual. A segunda, em relação ao Pós-teste Postergado 2, permite especular que o GE vai apresentar uma menor variabilidade por já ter tido a desestabilização do sistema durante o período do treinamento, enquanto o GC apresenta a maior variabilidade mais tarde (neste caso, na última etapa temporal de coleta).

A seguir, apresentaremos e analisaremos o gráfico de /t/ para as quatro etapas dos dois grupos.

**Gráfico 12** - Variabilidade dos grupos, Experimental e Controle, referentes às produções de VOT da consoante /t/, em cada um dos quatro momentos de coleta de dados.



**Nota:** Grupo 1 = Grupo Experimental; Grupo 2 = Grupo Controle.

**Fonte:** Elaborado pelo autor.

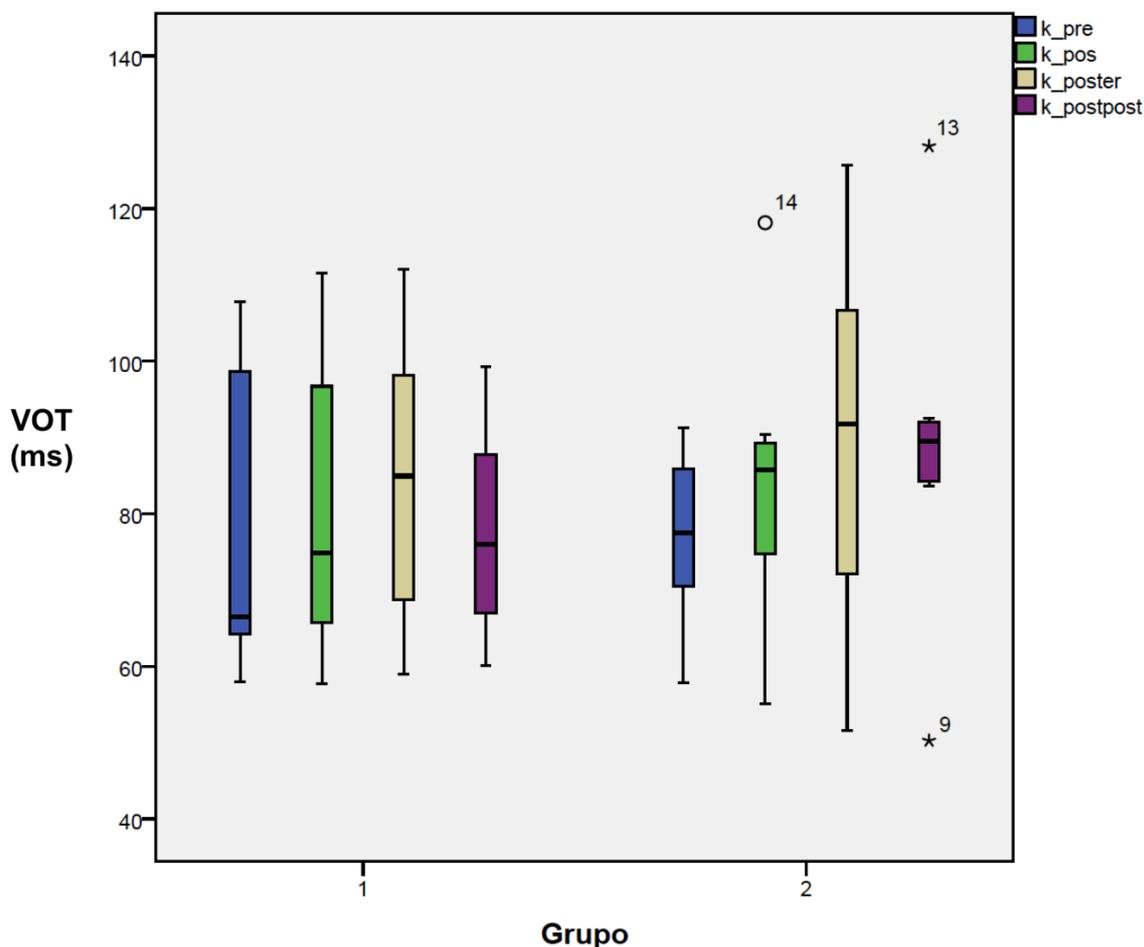
Em relação aos valores de /t/, algo a mais pode ser analisado. Retomemos, com base na Tabela 4, as médias dos dois grupos nas quatro etapas: GE - 61,09 ms; 64,21 ms; 68,99 ms; 51,23 ms; GC - 67,71 ms; 66,04 ms; 71,79 ms; 72,37 ms. Desde o Pré-teste, todas as médias do GC já apresentam valores maiores do que os do GE. Em relação às medianas, com a exceção do Pós-teste Postergado 2, o mesmo padrão se repete, pois temos, nas quatro etapas, os seguintes números: GE - 54,64 ms; 60,93 ms; 60,76 ms; e 51,76 ms; GC - 65,96 ms; 63,42 ms; 72,08 ms; e 51,23 ms.

Quando olhamos para os valores mínimos e máximos, a partir do Gráfico 12, temos: GE - mínimo de 37,39 ms e máximo de 91,94 ms no Pré-teste; 30,98 ms e 101,66 ms no Pós-teste; 36,22 ms e 108,23 ms no Pós-teste Postergado 1; 22,47 ms e 88,75 ms no Pós-teste Postergado 2; GC - mínimo de 51,98 ms e máximo de 101,48

ms no Pré-teste; 49,01 ms e 81,28 ms no Pós-teste; 37,18 ms e 98,72 ms no Pós-teste Postergado 1; 39,98 ms e 101,08 ms no Pós-teste Postergado 2. O que podemos inferir a partir destes dados é que, com valores menores, portanto mais distantes dos valores da língua-alvo e, ao mesmo tempo, valores maiores de desvio padrão, o GE apresentava, desde o início (apesar de não demonstrado inferencialmente), maior potencial de desenvolvimento, enquanto o GC, por já ter valores mais próximos ao nativo, não. Além disso, podemos reafirmar a especulação de que a mudança na última etapa, com os valores mínimos e máximos do GC maiores do que os do GE, seja por causa do treinamento, devido ao qual, após o período de desestabilização das primeiras três etapas, o grupo chegou na última etapa com seu sistema mais estabilizado. Em contrapartida, o oposto pode ter acontecido com o GC; pelo fato de os participantes já terem valores próximos dos nativos, além de não terem tido o treinamento, a variação não ocorreu em tamanha intensidade, tendo ocorrido mais tardiamente, como mostra o desvio padrão do grupo no Pós-teste Postergado 2.

A seguir, apresentaremos e analisaremos o gráfico de /k/ para as quatro etapas dos dois grupos.

**Gráfico 13** - Variabilidade dos grupos, Experimental e Controle, referentes às produções de VOT da consoante /k/, em cada um dos quatro momentos de coleta de dados.



**Nota:** Grupo 1 = Grupo Experimental; Grupo 2 = Grupo Controle.

**Fonte:** Elaborado pelo autor.

Retomando a seção 4.1, a produção da consoante /k/ não se mostrou significativa para o GE em nenhuma etapa, de acordo com o teste de Friedman; porém, de forma surpreendente, mostrou-se marginalmente significativa no GC. De acordo com os resultados do teste de Wilcoxon (Tabela 5), essa diferença marginal se dá entre o Pré-teste e o Pós-teste Postergado 1. Em termos descritivos, uma série de fatores nos permitem entender os resultados inferenciais. Primeiramente, podemos retomar os valores da média e mediana do grupo em cada etapa: média 77,10 ms e mediana de 77,49 ms no Pré-teste, média de 84,10 ms e mediana de 85,73 ms no Pós-teste, média de 89,76 ms e mediana de 91,72 ms no Pós-teste Postergado 1, e média de 88,74 ms e mediana de 89,50 ms no Pós-teste Postergado 2 (Tabela 4). Agora, olhemos os valores mínimos e máximos do GC nas quatro etapas do trabalho:

Pré-teste, mínimo de 57,86 ms e máximo de 91,31 ms; Pós-teste, mínimo de 55,15 ms e máximo de 118,12 ms; Pós-teste Postergado 1, mínimo de 51,66 ms e máximo de 125,61 ms; Pós-teste Postergado 2, mínimo de 50,23 ms e máximo de 128,15 ms. Ao estabelecermos um paralelo entre esses valores e os índices do Gráfico 13, fica clara a variabilidade principalmente nas três últimas etapas.

O GE, como assinalado, não mostrou significância inferencial para esta consoante no teste de Friedman. Porém, como podemos ver no gráfico 13, já era presente uma variabilidade grande dentro do grupo desde o começo do experimento. Quando olhamos para os valores mínimos e máximos do grupo nas quatro etapas, tal variabilidade fica mais evidente: Pré-teste, mínimo de 44,13 ms e máximo de 127,2 ms; Pós-teste, mínimo de 31,3 ms e máximo de 154,73 ms; Pós-teste Postergado 1, mínimo de 31,65 ms e máximo de 129,7 ms; e Pós-teste Postergado 2, mínimo de 16,52 ms e máximo de 124,14 ms.

Em suma, os gráficos parecem sugerir que as três consoantes aumentam a variação no GE entre as etapas Pré-teste e Pós-teste. Tal aumento sugere que o treinamento tenha, mesmo que apenas em termos descritivos, afetado positivamente a produção dos participantes. Ao mesmo tempo, quando lançamos o olhar para a quarta etapa, a variação do GC é maior do que a do GE em todas as consoantes. Tais resultados corroboram a especulação de que o Controle tenha tido sua desestabilização do sistema posteriormente ao Experimental, que possivelmente teve sua desestabilização ocasionada pelo treinamento. Passemos, então, para a análise individual dos participantes.

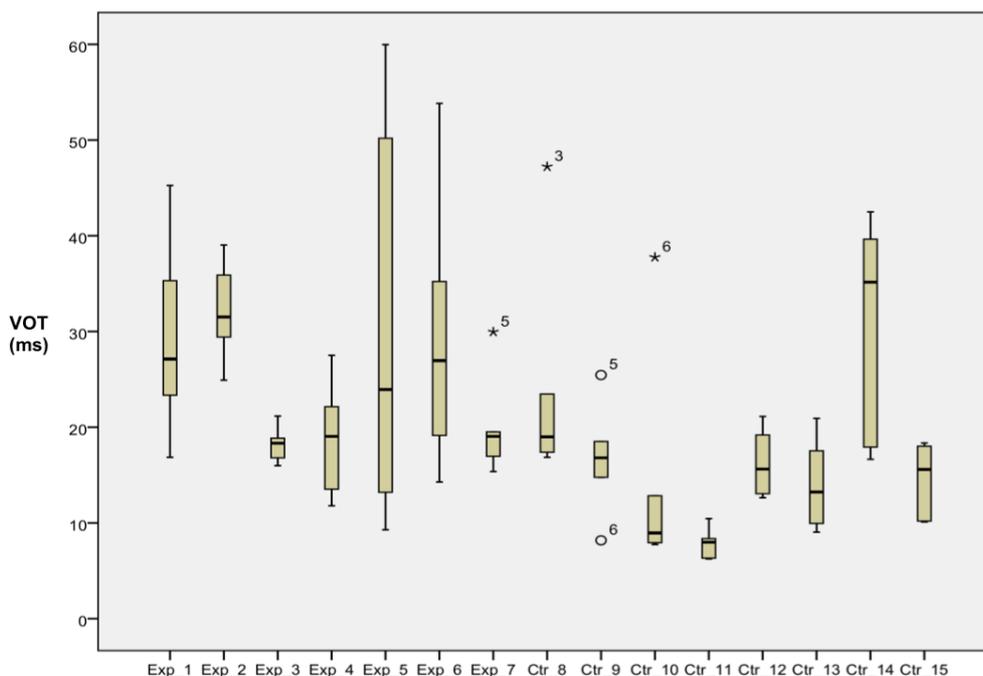
#### 4.3.1 Análise dos índices de variabilidade individuais dos participantes

A seguir, serão apresentados os gráficos com a variabilidade individual dos participantes dos dois grupos do experimento. Esta seção será dividida da seguinte forma: primeiramente, traremos gráficos individuais para cada uma das quatro etapas de coleta referentes à consoante /p/, com cada um deles seguido de sua respectiva análise; em seguida, os gráficos das quatro etapas de /t/ serão apresentados, junto com suas análises correspondentes; finalmente, os gráficos e análise referentes à consoante /k/ serão apresentados e discutidos.

#### 4.3.1.1 Análise dos índices de variabilidade individuais dos participantes para a consoante /p/

A seguir, trazemos o gráfico referente às produções individuais da consoante /p/ para os participantes dos dois grupos no Pré-teste. Em seguida, traremos o gráfico referente ao Pós-teste, seguido do gráfico referente ao Pós-teste Postergado 1 e, finalmente, o do Pós-teste Postergado 2. Após uma breve análise de cada um dos gráficos separadamente, faremos a análise geral, entre as quatro etapas, das trajetórias.

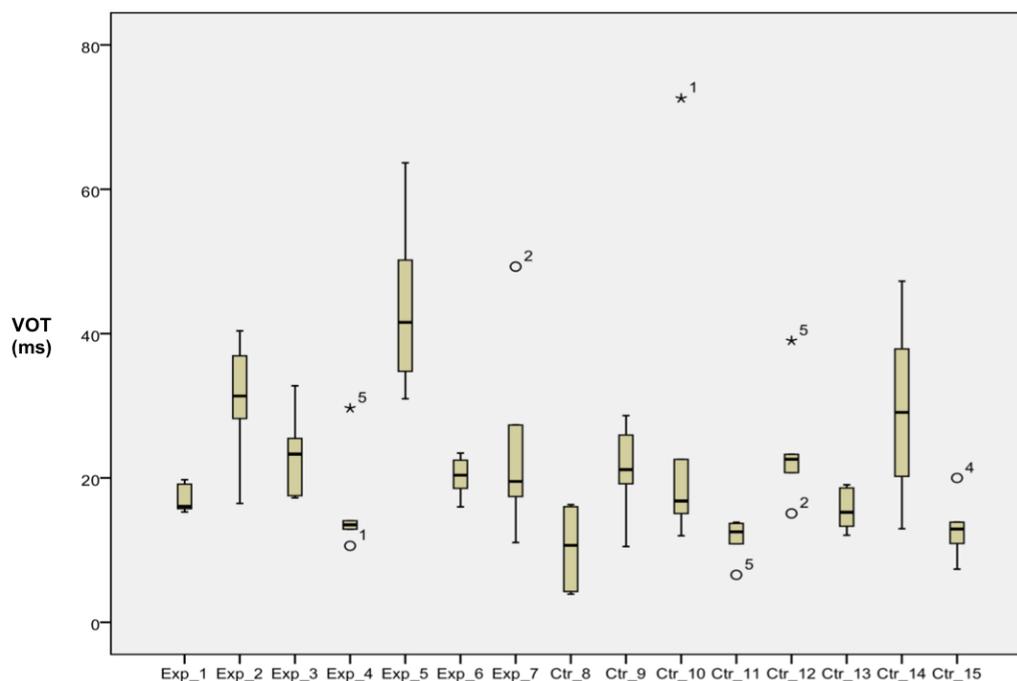
**Gráfico 14** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pré-teste.



Fonte: Elaborado pelo autor.

O gráfico mostra-nos como os participantes do GE, nesta etapa de coleta, já se mostram mais dispostos à variabilidade dos que os do GC. Participantes como o 1, 5 e 6, principalmente, demonstram tal variabilidade. Os participantes 8, 10 e 14 destacam-se no GC, especialmente o 14, que será analisado em todas as consoantes, devido à alta variabilidade que apresentou ao longo de todo o estudo. A seguir, veremos a variabilidade dos grupos para /p/ na etapa referente ao Pós-teste.

**Gráfico 15** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pós-teste.



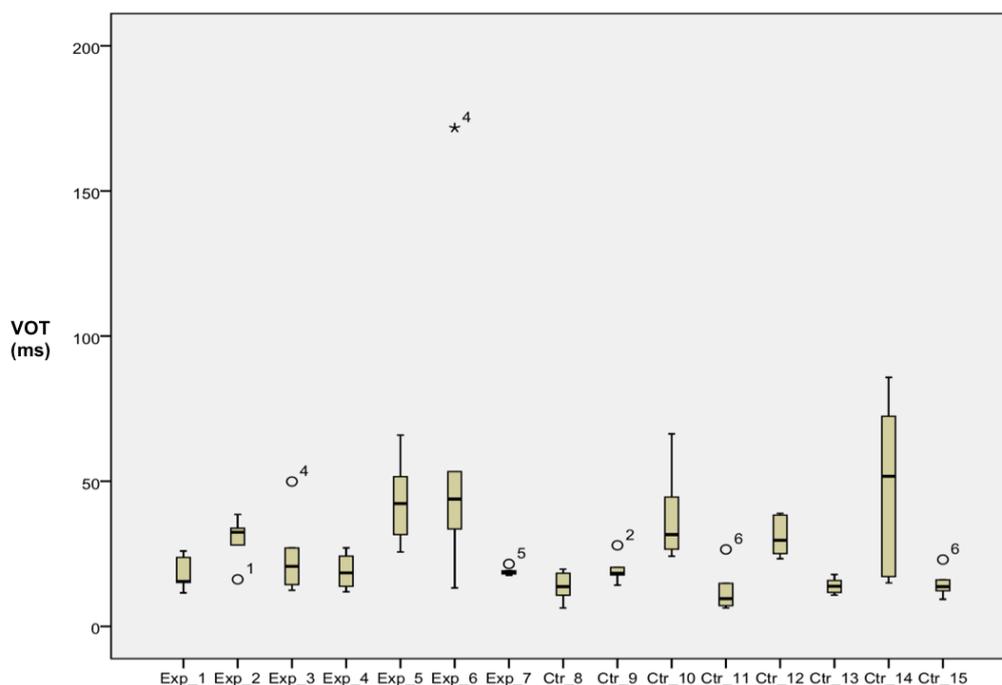
**Fonte:** Elaborado pelo autor.

O GE, na segunda etapa de coleta, apresenta uma menor variabilidade dos participantes, fato que se refletiu em termos estatísticos na Tabela 5, que não mostra um aumento significativo da etapa anterior para esta etapa. Porém, a partir da análise individual, podemos ver que o participante 7 apresenta uma variabilidade considerável, assim como o participante 5, contrariando a perspectiva de grupo.

O GC, por sua vez, ainda traz a variabilidade do participante 14, assim como o participante 10 mostra ter uma ampla variação nesta etapa.

A seguir, veremos a variabilidade individual dos participantes dos dois grupos para /p/, na etapa referente ao Pós-teste Postergado 1.

**Gráfico 16** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pós-teste Postergado 1.



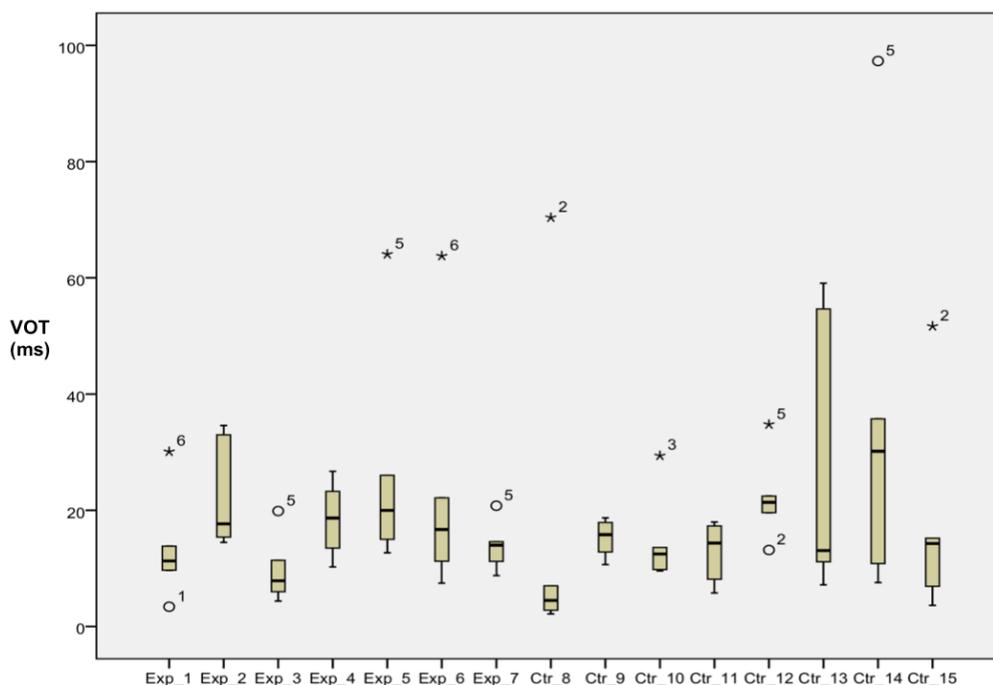
**Fonte:** Elaborado pelo autor.

Nesta etapa, para o GE, é destacável o gráfico do participante 6, que apresenta uma variabilidade muitíssimo ampla. Os participantes 3 e 5 também apresentam variabilidade, apesar de mais tímida do que a do participante 6. O restante do grupo, porém, aparenta representar uma possível estabilização em seus valores.

Já no GC, o participante 14 continua com uma variabilidade considerável, acompanhado do participante 10. O restante do grupo, porém, apresenta valores de variabilidade menores.

A seguir, veremos a variabilidade dos grupos para /p/ na etapa Pós-teste Postergado 2.

**Gráfico 17** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /p/ no Pós-teste Postergado 2.



**Fonte:** Elaborado pelo autor.

Na quarta e última etapa, corroborando o que já foi trazido ao longo do trabalho, o que se destaca, em questão de variabilidade, é que a mesma está presente em ambos os grupos. No GE, tal variabilidade resultou em uma diferença estatisticamente significativa (Tabela 4). Porém, um fato importante é apresentado neste gráfico: a maior variabilidade apresenta-se nos participantes 5 e 6. Os outros, apesar de apresentarem *outliers*, que inicialmente poderiam ser vistos como sinal de alta variabilidade, na verdade mostraram menor variabilidade, corroborando a especulação de os sistemas da maioria dos participantes que passaram pelo treinamento já apresentarem indícios de estabilização anos após o mesmo.

O GC, apesar de não apresentar esta diferença significativa na estatística inferencial, a apresenta em termos descritivos, que são refletidos diretamente neste último gráfico. Novamente, a trajetória do participante 14 é destacável. Participantes como 8, 13 e 15 também se destacam nesta etapa, corroborando as especulações de que o GC, por não ter passado pela desestabilização causada pelo treinamento, apresenta maior variabilidade do que o GE. A seguir, então, faremos a análise geral da trajetória dos indivíduos para a produção de /p/ nos quatro pontos de coletas.

Considerando a trajetória dos indivíduos nos quatro pontos, na seção 4.2.2.1, destacamos a trajetória do participante 6. No Pós-teste Postergado 1, o participante trouxe a média de 59,93 ms, com desvio-padrão de 56,63 (Anexo C). Ao analisarmos o valor mínimo de 13,24 ms e chegando ao valor máximo de 171,72 ms nesta etapa, e ao fazemos a relação com o Gráfico 16, fica clara a variabilidade do participante. O Gráfico 17 também deixa isso claro; ambos, inclusive, apresentam valores extremos tanto no Pós-teste Postergado 1 quanto Pós-teste Postergado 2.

Outro participante do GE destacável é o número 5, principalmente no Pré-teste, Pós-teste e Pós-teste Postergado 2. No primeiro, seus valores variam do mínimo de 9,3 ms ao máximo de 59,19 ms, que se destaca no Gráfico 14; no segundo, de 30,98 ms a 63,66 ms; e no quarto, de 12,7 ms a 64,06 ms. Esta última etapa traz, à semelhança do participante 6, um valor extremo.

Já em relação ao GC, foi destacada a trajetória do participante 13, que segue o caminho contrário ao mostrado para a média geral entre as duas etapas finais do experimento. Enquanto a média do grupo apresenta uma diminuição de 23,92 ms no Pós-teste Postergado para 19,93 ms no Pós-teste Postergado 2, o participante aumenta de 13,98 ms para 26,35 ms. Ao analisarmos os valores mínimos e máximos de cada etapa mencionada (mínimos de 10,79 ms e 7,5 ms, e máximos de 17,9 ms e 59,06 ms, respectivamente), também podemos verificar a variabilidade do sistema do participante. Tais resultados corroboram o que foi especulado nas seções 4 e 4.2.2, referente à possibilidade de que o GC, por não ter participado do treinamento perceptual, sofreu a desestabilização mais tardiamente.

O participante 14 traz, também, uma trajetória muito importante. No Pré-teste, seus valores de produção individual variam do mínimo de 16,66 ms ao máximo de 42,49 ms. No Pós-teste, de 12,96 ms a 47,25 ms. No Pós-teste Postergado 1, de 14,95 ms a 85,75 ms. No Pós-teste Postergado 2, 7,57 ms a 97,31 ms. Ao relacionarmos esses valores com os gráficos 14 a 17, fica evidente a variabilidade do sistema do participante. Podemos ver que o mesmo, apesar de não ter participado do treinamento perceptual, já começou o experimento com o sistema desestabilizado, o que se manteve durante o restante do trabalho.

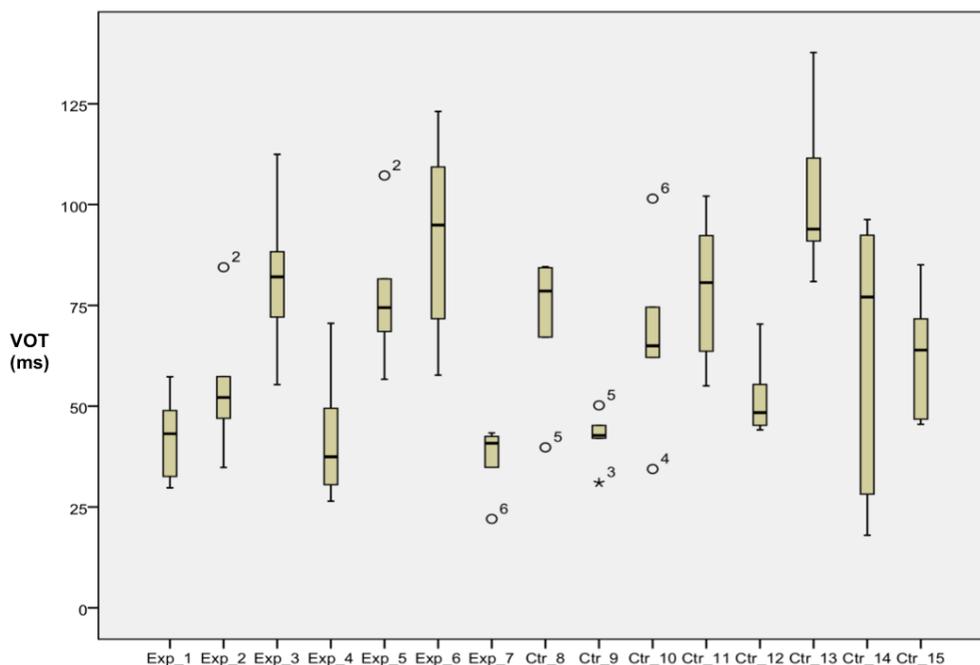
Ao fazermos uma análise em linhas gerais, reforçamos os achados da seção 4.2.2.1, de que a maioria dos participantes do GE chega na última etapa do experimento com seus valores mais estabilizados que os do GC. Novamente, tal fato pode se dar justamente pelo treinamento pelo qual o primeiro grupo passou. Tal

treinamento parece ter se mostrado útil para esta consoante tanto em termos inferenciais quanto descritivos, causando a desestabilização do sistema durante as três primeiras etapas, no qual houve o crescimento dos valores de média e desvios-padrão, com maior variação entre valores mínimos e máximos, e a diminuição dos mesmos na quarta etapa, quando o sistema dos participantes estaria estabilizado. Já o contrário acontece com o GC: por não participar de nenhum tipo de treinamento, seu desenvolvimento vai se dar mais tardiamente, o que explicaria os valores maiores na última etapa do projeto, em relação ao GE.

#### 4.3.1.2 Análise dos índices de variabilidade individuais dos participantes para a consoante /t/

A seguir, trazemos o gráfico referente às produções individuais da consoante /t/ para os participantes dos dois grupos no Pré-teste. Em seguida, traremos o gráfico referente ao Pós-teste, seguido do gráfico referente ao Pós-teste Postergado 1 e, finalmente, o do Pós-teste Postergado 2. Após uma breve análise de cada um dos gráficos separadamente, faremos a análise geral, entre as quatro etapas, das trajetórias.

**Gráfico 18** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pré-teste.

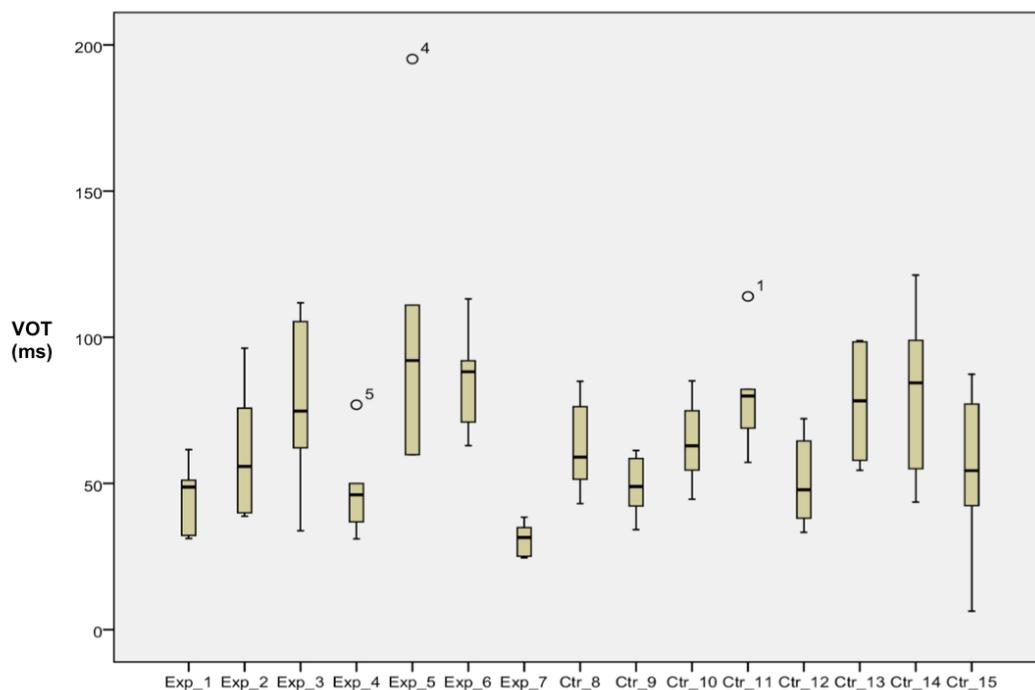


**Fonte:** Elaborado pelo autor.

No GE, vemos que ambos os participantes 3 e 5 já começam com uma variabilidade considerável. Apesar de o restante do grupo não apresentar tamanha variabilidade, ela também está presente, exceto pelos participantes 1 e 7. Além disso, por questões de produção próprias da consoante, não é surpresa que a mesma já vá apresentar maior variação desde o começo. Falaremos mais disso nas análises a seguir. O GC, por sua vez, também traz alunos que individualmente mostram variabilidade, como os participantes 10, 11, 13 e 14.

A seguir, traremos o gráfico referente às produções individuais de /t/ pelos participantes dos dois grupos no Pós-teste.

**Gráfico 19** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pós-teste.



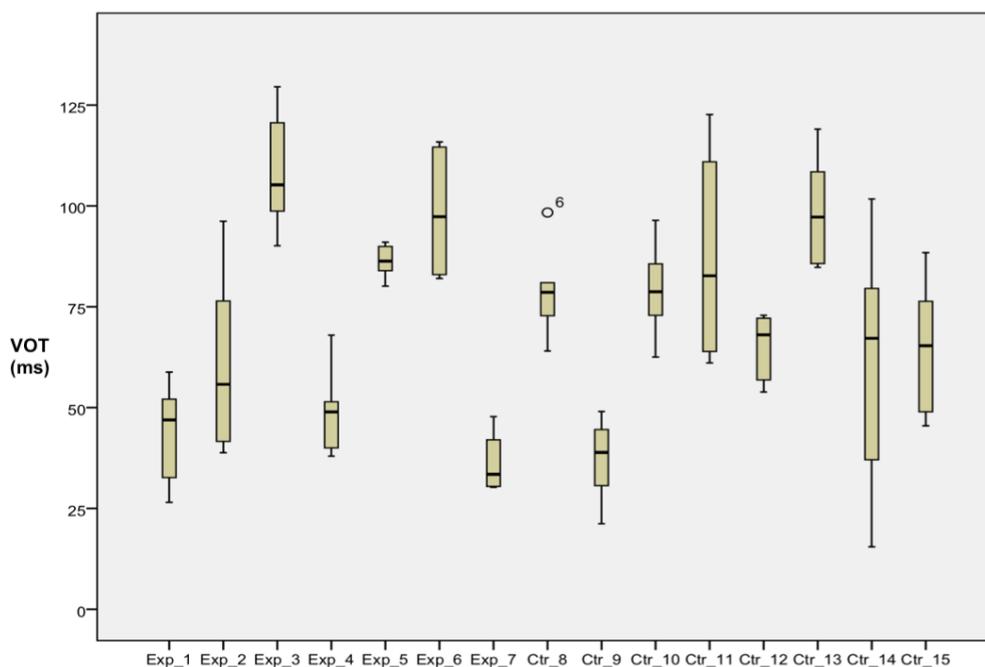
**Fonte:** Elaborado pelo autor.

O gráfico acima, por sua vez, reflete o que acontece, em relação ao GE, com os participantes do mesmo após passarem pelo treinamento. É possível identificarmos maior variabilidade nos participantes 3 e 5. Outros participantes, como os de número 2, 4 e 6, também apresentam variabilidade, apesar de mais tímida do que a dos primeiramente citados.

O GC, por sua vez, apresenta, também, dois participantes com clara variabilidade, os de número 14 e 15. Outros participantes, como os de número 11 e 13, apresentam uma variabilidade mais tímida, enquanto o restante dos alunos (8, 9, 10, 12), que representa a metade do grupo, apresenta uma variabilidade menor.

A seguir, traremos o gráfico referente às produções individuais de /t/ pelos participantes dos dois grupos no Pós-teste Postergado 1.

**Gráfico 20** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pós-teste Postergado 1.



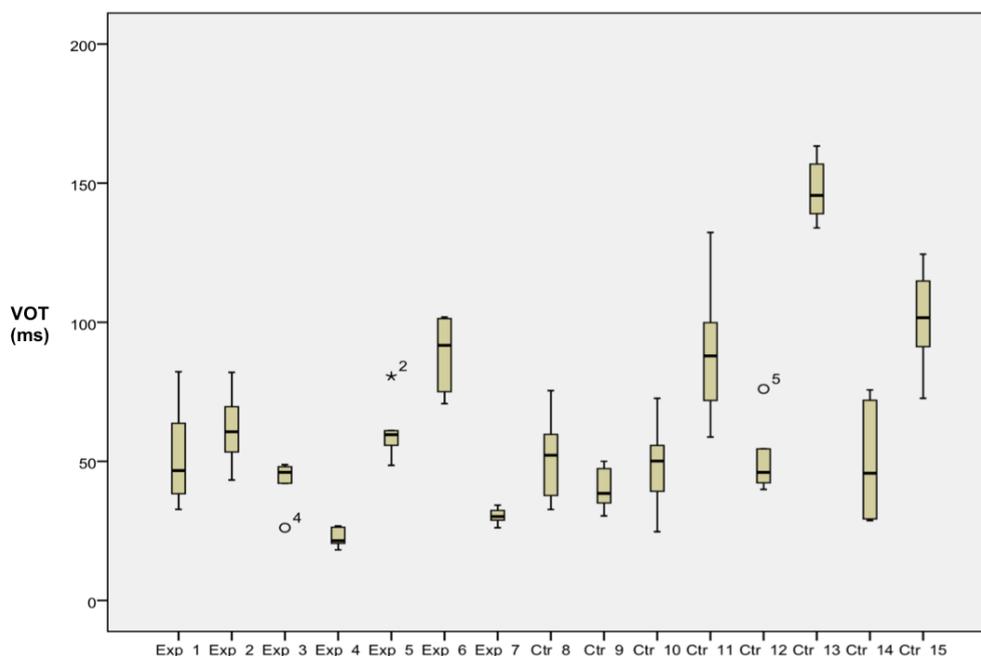
**Fonte:** Elaborado pelo autor.

Neste gráfico, destacamos, para ambos os grupos, o aumento das medianas, com os participantes trazendo valores maiores do que as médias encontradas nas produções de nativos da língua-alvo. Além disso, no GE, a maior variabilidade fica a cargo do participante 2. O restante dos participantes do GE, com variabilidades mais tímidas, pode corroborar o que já foi afirmado anteriormente, de modo a sugerirem o começo de um processo de estabilização.

O GC, por sua vez, apresenta, além do aumento das medianas, variabilidade clara em dois participantes, o de número 11 e o de número 14. Participantes como 10 e 15 também apresentam variação, mas em menor grau, enquanto o restante do grupo não aparenta variar notavelmente.

A seguir, traremos o gráfico referente às produções individuais de /t/ pelos participantes dos dois grupos no Pós-teste Postergado 2.

**Gráfico 21** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /t/ no Pós-teste Postergado 2.



Fonte: Elaborado pelo autor.

Neste último gráfico para /t/, fica evidente uma das especulações que foi feita ao longo de todo o trabalho: a estabilização que o GE aparenta apresentar na última e quarta etapa. O GC, surpreendentemente, apresenta participantes com maiores variabilidades do que os do GE, tais como os participantes 8, 10, 11, 14 e 15. A seguir, então, faremos a análise geral da trajetória dos indivíduos para a produção de /t/ nos quatro pontos de coleta.

Na seção 4.2.2.2, uma das trajetórias destacadas do GE foi a do participante 3. Os desvios-padrão entre Pré-teste (18,82) e Pós-teste (28,84) mostraram variabilidade entre as etapas. Evidências maiores podem ser encontradas quando observamos os valores mínimos e máximos desse participante nessas etapas - mínimos de 55,38 ms e 33,86 ms, e máximos de 112,42 ms e 111,82 ms, respectivamente. Trata-se de uma maior variação entre etapas, evidenciando uma maior desestabilização no sistema linguístico do participante. Entre as etapas Pós-teste (média de 77,12 ms) e Pós-teste Postergado 1 (média de 108,23 ms), vimos um aumento considerável no que diz respeito às médias; ao analisarmos os valores mínimos e máximos destas etapas (mínimos de 33,86 ms e 90,13 ms, máximos de 111,82 ms e 129,49 ms, respectivamente), refletidos diretamente nos quatro últimos

gráficos, podemos especular que um possível início de estabilização esteja em ocorrência.

O participante 2 também apresentou uma trajetória interessante. Podemos analisar os valores mínimos e máximos ao longo das quatro etapas: mínimo de 34,81 ms e máximo de 84,45 ms no Pré-teste; 38,77 ms e 96,26 ms no Pós-teste; 38,8 ms e 96,2 ms no Pós-teste Postergado 1; e 43,34 ms e 82 ms no Pós-teste Postergado 2. Tais valores reforçam a afirmação de uma possível desestabilização do sistema da primeira para a segunda etapa, que se manteve na terceira e diminuiu na quarta, corroborando a especulação da seção anterior de que na última etapa do experimento alguns participantes já haviam estabilizado seus sistemas depois da desestabilização inicial.

O participante 5 também apresenta uma possível desestabilização ao aumentar a média entre Pré-teste (77,14 ms) e Pós-teste (101,66 ms); ao observarmos os gráficos 18 e 19, referentes às etapas supracitadas, vemos que o participante apresenta, inclusive, valores máximos extremos, sendo 107,22 ms na primeira etapa e 195,21 ms na segunda. Porém, como ressaltado na análise da seção 4.2.2, a possível desestabilização não aparenta manter-se nas etapas seguintes.

Já para o GC, destacamos os participantes 11, pois o participante segue o mesmo padrão apresentado no GE. Em questão de valores mínimos e máximos, seus índices são de 55,07 ms e 102,11 ms no Pré-teste; 57,16 ms e 113,96 no Pós-teste; 61,04 ms e 122,67 ms no Pós-teste Postergado 1; e 58,78 ms e 132,2 ms no Pós-teste Postergado 2. Com esses valores, especulamos que a variação maior no Pós-teste Postergado 2 do que no Pré-teste representa uma continuidade na possível desestabilização do sistema fonético-fonológico do estudante. Ou seja, o sistema do participante continua bastante instável e sensível a alterações, que possivelmente acontecem mais tardiamente no participante do GC justamente por esse não ter passado pelo treinamento perceptual.

Novamente, cabe destacar a trajetória do participante 14. Nos gráficos 18 e 20, referentes ao Pré-teste e ao Pós-teste Postergado 1, vemos que há uma variabilidade considerável. Trazendo seus valores mínimos e máximos, vemos um valor mínimo de 18,03 ms e um valor máximo de 96,26 ms no Pré-teste, e de 15,46 ms e 101,75 no Pós-teste Postergado 1, de modo que essa variabilidade fica ainda mais evidente.

Como destacado na seção 4.2.2, a maior experiência com a língua-alvo aparece refletida em participantes com trajetórias destacáveis; esse é o caso do participante em questão, que, de acordo com as respostas referentes ao Questionário de Experiência e Proficiência Linguística, de Scholl e Finger (2013), traz alguns dos valores mais altos em relação ao tempo em escola ou trabalho em que se usa a L2 (13 anos), percentagem de uso da língua (30% do dia), e nove horas diárias de uso da língua.

Em suma, nossa análise reforça a instabilidade de /t/<sup>29</sup> ser maior do que a de /p/. Tal instabilidade talvez se dê pelo fato de a própria consoante /t/ ser mais suscetível à variação, uma vez que dispõe de maior capacidade aerodinâmica de realização. As diferenças individuais reforçam, mais uma vez, a análise individual, uma vez que a média de grupo parece não expressar o que está acontecendo com os participantes. Além disso, mesmo as estatísticas inferenciais não tendo apresentado resultados significativos para /t/, o treinamento parece contribuir para instaurar uma instabilidade mais precoce entre os participantes do GE. Esta instabilidade nos dados, que além disso denota o início e a intensificação do desenvolvimento linguístico, também pode ser encontrada entre participantes do GC; porém, no caso de tais participantes, a instabilidade tende a aparecer em momentos temporais mais tardios, justamente pelo fato de não terem tido seus processos de desenvolvimento de VOT acelerados pelo treinamento.

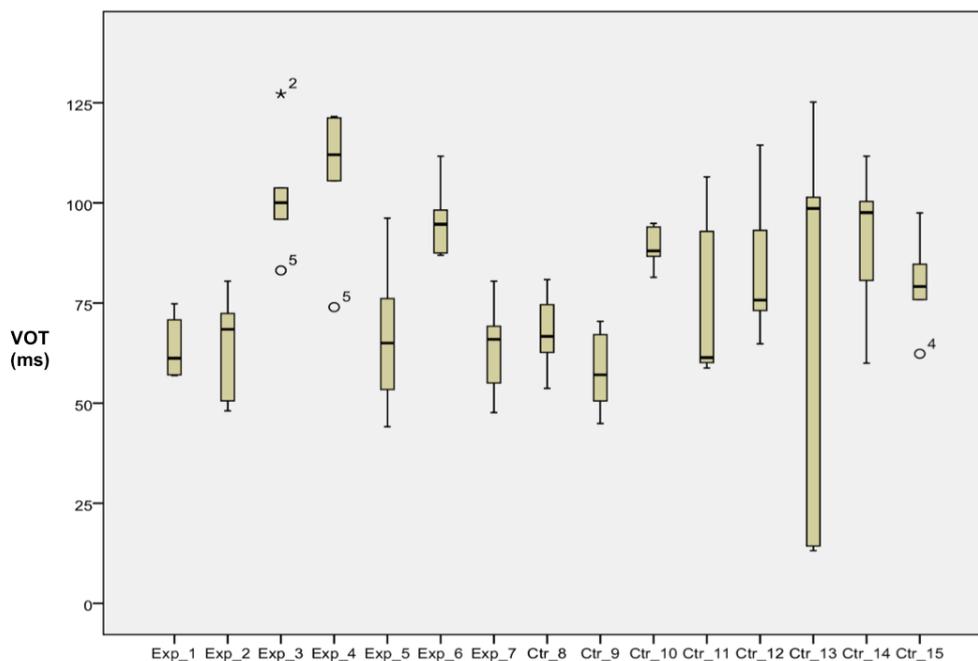
#### 4.3.1.3 Análise dos índices de variabilidade individuais dos participantes para a consoante /k/

A seguir, trazemos o gráfico referente às produções individuais da consoante /k/ para os participantes dos dois grupos no Pré-teste. Em seguida, traremos o gráfico referente ao Pós-teste, seguido do gráfico referente ao Pós-teste Postergado 1 e, finalmente, o do Pós-teste Postergado 2. Após uma breve análise de cada um dos gráficos separadamente, faremos a análise geral, entre as quatro etapas, das trajetórias.

---

<sup>29</sup> Exemplos adicionais da instabilidade de /t/ em outros fenômenos fonético-fonológicos podem ser encontrados em Perozzo (2013).

**Gráfico 22** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pré-teste.



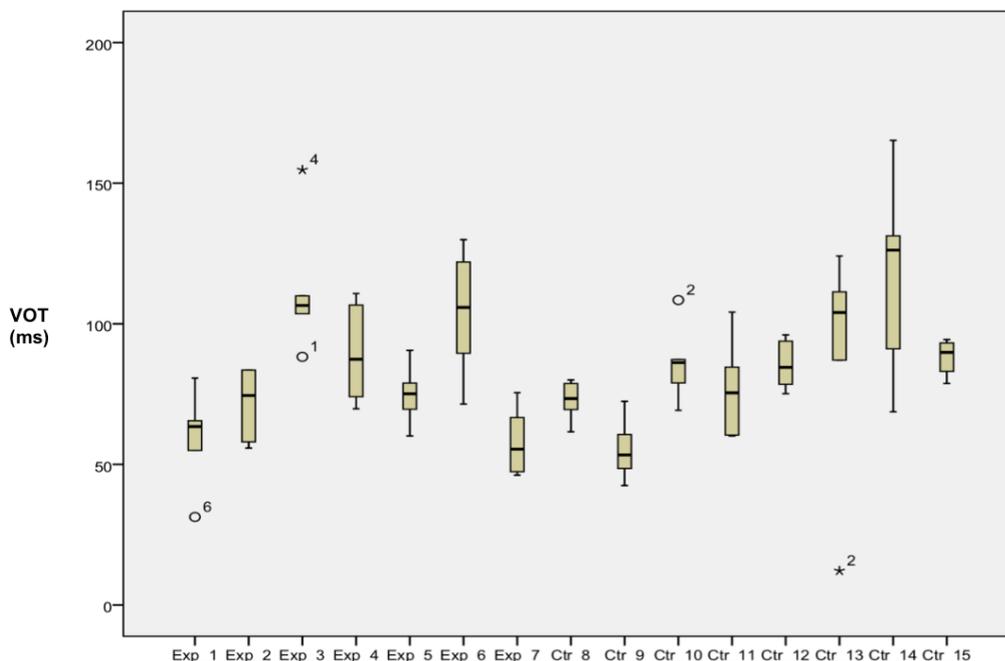
Fonte: Elaborado pelo autor.

Como já destacado ao longo do trabalho, a consoante /k/ teve uma significância marginal em termos estatísticos (Tabela 4). Ao observarmos o gráfico acima, vemos que no Pré-teste, antes do treinamento, a variabilidade dos participantes do GE é tímida, o que demonstra um potencial para que o treinamento promova mudanças em seus sistemas.

O GC, por sua vez, apresenta-se bem dividido em relação à variabilidade. Enquanto participantes como os de número 11, 12, 13 e 14 trazem variabilidades consideráveis, o restante apresenta graus menores de variação.

A seguir, traremos o gráfico referente à produção de /k/ pelos dois grupos no Pós-teste.

**Gráfico 23** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pós-teste.



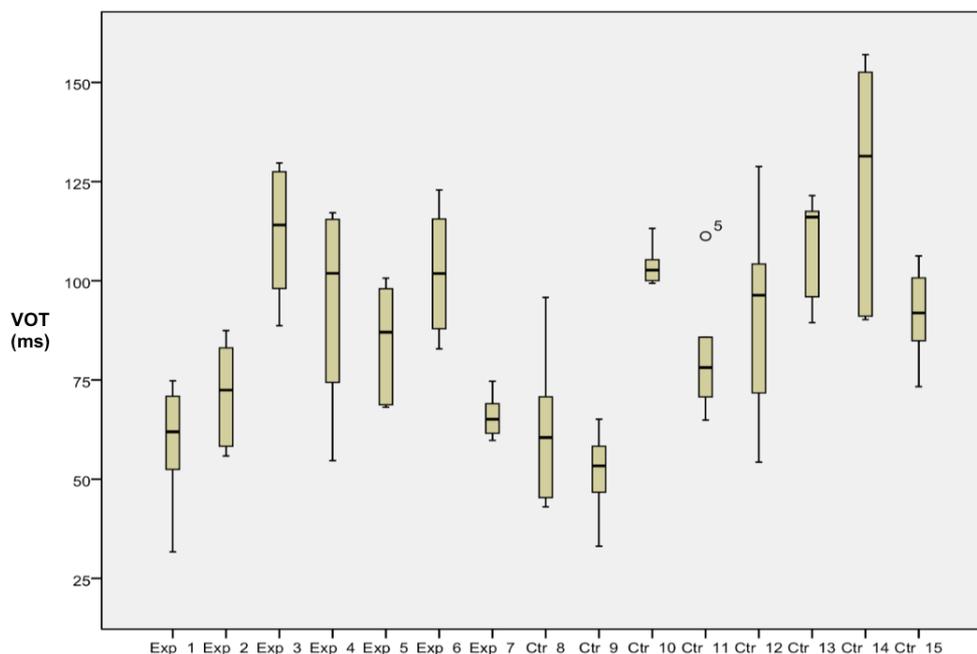
Fonte: Elaborado pelo autor.

Neste gráfico, vemos que poucos participantes apresentam variabilidades destacáveis. Tal resultado não surpreende, considerando que dados anteriores indicavam que a diferença estatisticamente significativa se deu apenas entre Pré-teste e Pós-teste Postergado 1. Porém, voltando nosso olhar para as trajetórias individuais, vemos que as maiores variabilidades ocorreram nos participantes 3, 4 e 6, que podem ter sido causadas pelo treinamento perceptual.

O GC, por sua vez, apresenta apenas dois participantes com variabilidade destacável, os participantes 13 e 14. É interessante ressaltar que ambos são destacados ao longo do trabalho. Retomaremos informações referentes aos mesmos a partir do Questionário de Experiência e Proficiência Linguística, de Scholl e Finger (2013).

A seguir, traremos o gráfico referente à produção de /k/ pelos dois grupos no Pós-teste Postergado 1.

**Gráfico 24** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pós-teste Postergado 1.

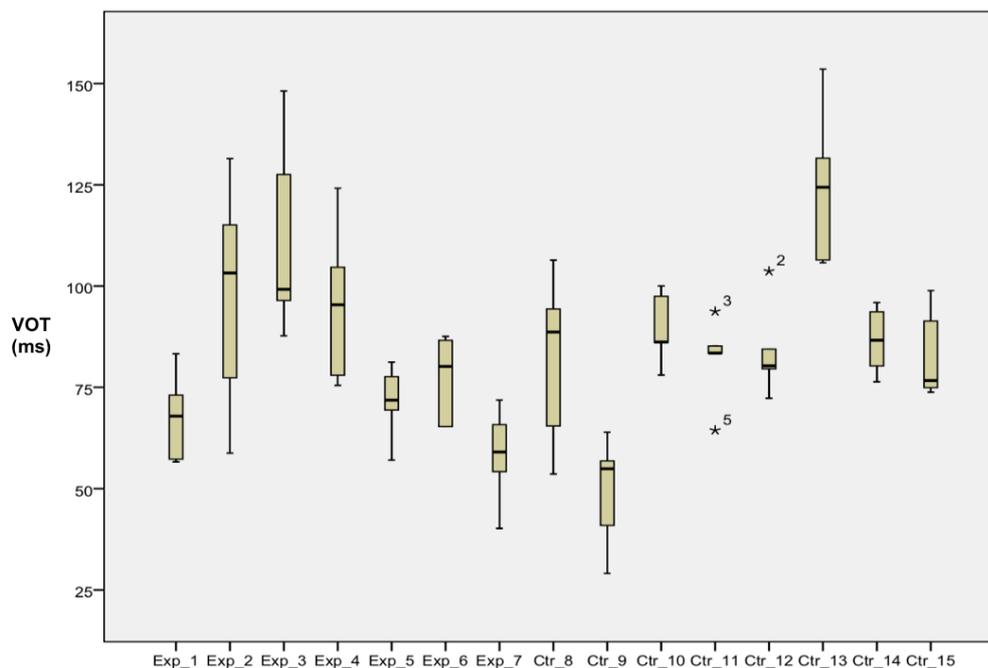


Fonte: Elaborado pelo autor.

Neste gráfico, com exceção do participante 7, todos os participantes aparentam trazer maior variabilidade em suas produções. Tal variabilidade pode significar que os participantes necessitassem de um período de tempo maior que o decorrido entre Pré e Pós-teste para "adequar" suas produções de acordo com o padrão nativo da língua-alvo.

O GC, por sua vez, apresenta os participantes 8, 12 e 14 com variabilidades bem marcantes, enquanto o restante permanece com uma variabilidade mais tímida. A seguir, traremos o gráfico referente à produção de /k/ pelos dois grupos no Pós-teste Postergado 2.

**Gráfico 25** - Variabilidade individual dos participantes de ambos os grupos, Experimental e Controle, para a consoante /k/ no Pós-teste Postergado 2.



Fonte: Elaborado pelo autor.

Finalmente, o gráfico acima demonstra que, na última etapa do experimento, no GE, os participantes 2, 3 e 4 apresentam as maiores variabilidades. Os três citados são os que se destacam quando consideramos que, por representarem a maioria, os quatro participantes restantes criariam uma possível representação de que o grupo, em si, apareceria estabilizado nesta última etapa, anos depois no Pós-Teste Postergado 2.

Já no GC, apenas os participantes 8 e 13 apresentam uma variabilidade destacável. Os outros, por sua vez, apresentam variabilidades em grau menor. A seguir, então, faremos a análise geral da trajetória dos indivíduos para a produção de /k/ nos quatro pontos de coleta.

Na seção 4.2.2, foram destacados, no GE, os participantes 1, 3 e 6. Os participantes 1 e 6 aumentaram a diferença entre seus valores mínimos e máximos entre as duas primeiras etapas, ou seja, após o treinamento perceptual. Em termos de valores descritivos, o participante 1 apresenta o valor mínimo de 58,88 ms e máximo de 74,83 ms no Pré-teste, e mínimo de 31,3 ms e máximo de 80,71 ms no Pós-teste, enquanto o participante 6 traz o mínimo de 86,94 ms e máximo de 111,68

ms no Pré-teste, e mínimo de 71,48 ms e máximo de 129,88 ms no Pós-teste. Como dito anteriormente, quando vemos a média do participante 1, nas mesmas etapas, há uma diminuição de 63,67 ms para 59,91 ms. Ou seja, mesmo com a diminuição da média, a variabilidade de valores mínimos e máximos aumenta. No que diz respeito às medianas expressas nos gráficos, o valor da mediana do participante 1 também diminui entre as mesmas etapas, assim como as do participante 6. Como foi ressaltado na seção supracitada, ambos os participantes vão mostrar menor variação entre valores mínimos e máximos, sugerindo a estabilização do sistema linguístico.

Já em relação ao participante 3, ao analisarmos os valores mínimos e máximos, temos os seguintes índices para as quatro etapas, respectivamente: mínimos de 88,14 ms; 88,26 ms; 88,71 ms; e 87,72 ms; máximos de 127,2 ms; 154,73 ms; 129,7 ms; e 148,14 ms. Pela variabilidade entre valores mínimos e máximos, especulamos que a principal desestabilização do sistema do participante não tenha ocorrido durante a primeira parte dos testes, mas anos depois (verificável a partir do segundo pós-teste postergado). O mesmo parece ter acontecido com o participante 4. Inclusive, ao observarmos os gráficos 24 e 25, referentes às etapas Pós-teste Postergado 1 e Pós-teste Postergado 2, a desestabilização do sistema fica evidente, principalmente em termos de valores mínimos e máximos, sendo estes 54,68 ms e 117,11 ms no Postergado 1 e 16,25 ms e 124,14 ms no Postergado 2. Isso é diferente do que ocorreu para /p/ e /t/; tal fato talvez se deve porque, no caso de /k/, a estabilização do padrão implicasse uma diminuição, e não um aumento, como a duas primeiras, dos valores médios de VOT.

Já para o GC, a trajetória do participante 13 tornou-se uma das mais destacáveis, pois, conforme já apresentado na seção 4.2.2.3, suas médias foram 75,22 ms (Pré-teste), 90,44 ms (Pós-teste), 109,42 ms (Pós-teste Postergado 1) e 128,14 ms (Pós-teste Postergado 2). Em questão de índices mínimos e máximos, os valores são os seguintes: mínimo de 13,21 ms e máximo de 125,19 ms no Pré-teste; mínimo de 12,15 ms e máximo de 124,08 ms no Pós-teste; mínimo de 89,43 ms e máximo de 121,51 ms no Pós-teste Postergado 1; e mínimo de 105,73 ms e máximo de 153,55 ms no Pós-teste Postergado 2. Esta diferença é claramente presente no Gráfico 22, de acordo com o qual, de todos os participantes, o 13 é o que apresenta a maior variabilidade. Podemos reforçar a ideia trazida na seção 4.2.2, de que o participante já iniciou o experimento com o sistema desestabilizado, mantendo a

desestabilização na segunda etapa e, na terceira, mostrando um possível início de estabilização.

Novamente, a trajetória do participante 14 destaca-se das demais. Nas primeiras três etapas, seus valores mínimos e máximos são de 59,89 ms; 68,67 ms; e 90,26 ms para os mínimos, e de 111,7 ms; 165,25 ms; e 156,98 ms para os máximos. Este crescimento da diferença entre valores do Pré-teste para o Pós-teste possivelmente significa uma maior desestabilização, e pode ser verificado também nos gráficos 22 e 23, assim como o 24 pode indicar que o participante manteve-a no Pós-teste Postergado 1. Já os números referentes à última etapa indicariam um princípio de estabilização do sistema, como podemos ver no gráfico 25.

Em suma, nossa análise reforça, novamente, que essa é a consoante que permite os maiores graus de variabilidade entre os participantes, e que justifica com mais clareza a necessidade de análises individuais. Essas diferenças mostram que alguns participantes apresentam o sistema estabilizado desde o início do experimento. Outros participantes (3, 4, 6, 11, 13 e 14), por sua vez, ainda que apresentassem já valores nativos em suas produções, apresentaram grande variabilidade, que, possivelmente, é caracterizada pela busca por uma diminuição dos valores das médias (para se aproximar do atrator nativo), como dito anteriormente. Quando reforçamos nossa análise à luz dos SACs e da análise individual, é importante ressaltar que já havia a perspectiva de resultados diversos e, aparentemente, aleatórios. A não linearidade, uma das propriedades principais da língua como um sistema complexo e dinâmico, dá conta da explicação para que tais resultados se apresentem tão variados.

Em suma, a análise das trajetórias individuais parece corroborar a análise dos grupos, mostrando que os membros do GE, em sua maioria, mesmo que apenas em termos descritivos, apresentam um aumento na variabilidade na etapa do Pós-teste, logo após as sessões de treinamento perceptual. As análises individuais do GC, porém, mostraram essa suscetibilidade à variabilidade já na quarta etapa, anos após passarem pelas primeiras etapas de coleta.

#### 4.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Este capítulo propôs descrever e analisar todos os dados coletados ao longo das quatro etapas de nosso experimento. Ele é o resultado de longas horas não apenas de organização de calendário para que conseguíssemos participantes, mas também a organização das datas para coletá-los e treiná-los, além das outras tantas horas e pessoas envolvidas na transcrição dos dados coletados, organização dos mesmos nas tabelas, rodagem estatística e, finalmente, uma reflexão que fizesse sentido ao trabalho proposto e, sobretudo, que conversasse com o leitor.

Das análises em grupo, em termos inferenciais, nossos resultados, em relação à percepção, corroboram os estudos publicados anteriormente (RODRIGUES, 2015; KAMPFF; ALVES, 2016; ALVES; LUCHINI, 2017), com o treinamento perceptual mostrando-se útil para aumentar o índice de acuidade nos padrões VOT Zero Natural e VOT Artificial. Os resultados de produção também corroboram os resultados supracitados, com a generalização dos efeitos positivos da percepção para a produção mostrando-se dificultosa na maioria das consoantes. As análises descritivas, também presentes nos estudos supracitados, mostram-se valiosas no sentido de que expressam mudanças, mesmo que tímidas, positivas nos índices e médias referentes às produções do GE, que passou pelo treinamento.

Porém, uma nova abordagem de análise mostrou-se ainda mais interessante, que é o olhar individual sobre os participantes. Este olhar permite-nos identificar, em cada um dos participantes, em termos descritivos, aproximações e distanciamentos das mudanças presentes nas estatísticas de grupo. Mais ainda, tal análise nos permitiu que, alinhando-nos à perspectiva dinâmica, pudéssemos verificar como cada sujeito apresenta uma jornada de desenvolvimento linguístico única.

Finalmente, passemos às considerações finais do trabalho.

## 5 CONSIDERAÇÕES FINAIS

O objetivo geral de nosso trabalho foi analisar, ao longo dos quatro pontos de coleta, e em uma janela temporal de três anos, as trajetórias individuais e em grupo de alunos do curso de Letras que participaram, ou não, de sessões de treinamento perceptual referentes aos padrões de VOT das oclusivas iniciais do inglês. Para tanto, três Questões Norteadoras nos guiaram: (i) o que as abordagens inferenciais e descritivas podem dizer a respeito do desenvolvimento de cada um dos grupos analisados?; (ii) o que uma verificação de dados individuais pode revelar sobre as trajetórias dos aprendizes?; (iii) que informações sobre a trajetória de desenvolvimento linguístico a variabilidade dos aprendizes em um dado momento temporal e a variação ao longo do tempo podem revelar?

Para a primeira pergunta, em termos inferenciais, no que concerne à percepção, nosso estudo corrobora outros já realizados e mencionados anteriormente ao longo do trabalho (RODRIGUES, 2015; KAMPFF; ALVES, 2016; ALVES; LUCHINI, 2017), pois os resultados mostraram-se significativos para o GE, submetido ao treinamento, nos quatro padrões de VOT, inclusive o Artificial, usado para fazer com que os participantes passassem a usar o VOT como principal pista de distinção entre surdas e sonoras. Já para a produção, apenas /p/ mostrou-se significativa para o GE. Como tal diferença de significância deu-se, em termos descritivos, por um aumento entre o Pré-teste e Pós-teste Postergado 1, e uma diminuição entre Pós-teste Postergado 1 e Pós-teste Postergado 2, especulamos que isso se deva ao fato de os participantes terem passado por um período de desestabilização por causa do treinamento e, anos depois, estarem com seus sistemas um tanto mais estabilizados, o que conversa diretamente com a visão dinâmica de língua que rege o presente trabalho (DE BOT *et. al.*, 2009; BECKNER *et al.*, 2009; LARSEN-FREEMAN, 2015).

Em relação à segunda Questão, dados ainda mais interessantes puderam ser analisados. Vimos que, a partir de uma análise mais individual, a visão dinâmica de língua se mostra claramente mais palpável, pois esta análise mostra o quanto as trajetórias individuais vão se distanciar da trajetória marcada pelo grupo como um todo. Retomando o teorema da ergodicidade (cf. LOWIE; VERSPOOR, 2018), fica claro o porquê da necessidade de avaliarmos nossos dados individualmente, uma vez que nossos participantes se apresentam heterogeneamente em diversos fatores que poderiam afetar os resultados, como nível de proficiência e experiência com a língua.

Tais resultados ainda corroboram as considerações de tais autores quando pensamos nas aplicações didáticas do que foi verificado. Ao considerarmos o desenvolvimento linguístico único, retomamos a importância do olhar individual do professor a cada um de seus alunos, sempre levando em consideração que cada um apresentará graus de desenvolvimento diferentes dos outros, (inclusive, como colocam os autores, parecendo regredir entre uma aula ou outra). O que podemos levar em consideração, por enquanto, é o dever do professor de persistir no ensino das características da língua-alvo que sejam relevantes para a comunicação do aprendiz, tais como o VOT, que corresponde a um aspecto fundamental para a inteligibilidade da L2 (SCHWARTZHAUPT, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015) uso de outras técnicas, além do treinamento, como instrução, por exemplo, será abordado posteriormente neste capítulo.

Já a terceira Questão, que analisa diretamente a variabilidade dos participantes, corrobora os achados de Lowie e Verspoor dos últimos anos (2015; 2018). Apesar de a variabilidade individual, quando considerada em conjunto, indicar que os participantes do GE chegaram à última etapa com seus sistemas menos instáveis do que os do GC, pudemos verificar que, como afirmam os autores, cada participante vai trazer uma trajetória individual, com variabilidades que nem sempre terão uma explicação específica. Portanto, uma das conclusões a que chegamos é a de que o treinamento pareceu "acelerar" o processo de desenvolvimento dos participantes do GE, enquanto o GC mostrou o mesmo desenvolvimento mais tardiamente. Tal conclusão se mostra importante principalmente para a análise dos dados de produção, que não apresentaram resultados significativos na verificação inferencial. Mesmo assim, o olhar individual nos permitiu verificar que, apesar da não significância enquanto grupo, participantes pareceram, sim, aproveitar-se significativamente do treinamento para atingirem padrões de produção próximos, ou até iguais, aos da língua-alvo.

Apesar dos resultados, que consideramos promissores, algumas limitações afetaram o trabalho. O número de sessões de treinamento foi uma delas, uma vez que contamos com períodos de aula cedidos por outros professores, para a realização dos estudos. Além disso, o baixo número de participantes caracterizou outra limitação, uma vez que não nos foi possível realizar testes estatísticos paramétricos, que se mostram mais poderosos inferencialmente. Porém, a partir de tal limitação, o número

insuficiente de participantes nos levou à análise das trajetórias individuais, uma perspectiva relativamente nova nos estudos da linguagem, sobretudo em nosso país.

Além disso, a não testagem da proficiência dos estudantes poderia vir a ser considerada, a uma primeira vista, como uma limitação. Partindo, novamente, do pressuposto de que os participantes de um grupo devem ser homogêneos para que as estatísticas funcionem, a verificação da proficiência poderia, a princípio, ser feita. Entretanto, ao se considerar novamente o trabalho de Lowie e Verspoor (2018), dois caminhos metodológicos poderiam ser seguidos: o primeiro implicaria fazer a separação dos grupos através da proficiência e experiência com a língua-alvo a ser estudada, para reforçar os resultados trazidos pelos autores, no trabalho supracitado, de que mesmo grupos supostamente homogêneos apresentam trajetórias individuais heterogêneas; o segundo diria justamente respeito à não testagem da proficiência, pois, se, em termos inferenciais, controlá-la mais detalhadamente se faça necessário, ao considerarmos a concepção que temos em nosso estudo, a não testagem não implicaria a perda da heterogeneidade, pois tal propriedade é a que caracteriza os indivíduos como sistemas complexos.

Finalmente, outra limitação, e talvez a mais importante, foi o intervalo de tempo entre primeira e a segunda fase de coletas, pois estudos longitudinais com visão dinâmica de língua sugerem uma certa periodicidade nos intervalos de tempo entre coletas. A quarta etapa, então, surgiu da curiosidade científica do autor, intrigado em saber quais seriam os resultados encontrados nos participantes após a terceira etapa de coleta, ocorrida em 2015. Ainda assim, muitos aspectos da descrição da trajetória dinâmica entre as duas últimas etapas podem ter sido prejudicados, uma vez que os participantes possam ter sofrido modificações no intervalo de tempo, como mais contato e experiência com a língua-alvo, além das próprias mudanças sentimentais e emocionais, levadas em conta na visão dinâmica, por considerar a língua como de domínio geral (BECKNER *et al.*, 2009; LARSEN-FREEMAN, 2015).

O trabalho lança luz, ademais, sobre uma série de investigações futuras. A partir de tal estudo, é possível pensar, por exemplo, em coletas a serem realizadas durante a própria fase de treinamento perceptual, uma vez que a mesma pode representar o período de maior instabilidade dos participantes. Além disso, um maior número de coletas é necessário e encorajado, tanto antes quanto depois do treinamento, inclusive em menores intervalos de tempo. Também, cabe destacar a necessidade de investigações sobre os efeitos da associação entre treinamento e

instrução, assim como a necessidade de investigação de grupos submetidos a instrução e treinamento separadamente, reforçando os estudos futuros sugeridos no trabalho de Alves e Luchini (2017). A partir de tais aspectos, deve-se, sempre, focalizar as trajetórias individuais de cada participante, assim como o que os resultados podem representar enquanto grupo.

Esperamos, com o presente trabalho, ter contribuído para o enriquecimento de estudos que analisem o papel do treinamento no desenvolvimento dos sistemas linguísticos de L2. Além disso, esperamos que o trabalho contribua para novas metodologias e análises à luz de uma visão dinâmica, e que contribua para evidenciar, através do componente empírico, considerações importantes estabelecidas pela teoria, como a heterogeneidade das trajetórias individuais e o poder explanatório da variabilidade dos dados de desenvolvimento.

## REFERÊNCIAS

ALVES, M. *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*, v. 5, 31-50, 2014.

ABRAMSON, A. S. WHALEN, D. H. Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, v. 63, p. 75-86, 2017.

\_\_\_\_\_; ZIMMER, M. C. Percepção e produção dos padrões de VOT do inglês por aprendizes brasileiros: O papel de múltiplas pistas acústicas sob uma perspectiva dinâmica. *Alfa: Revista de Linguística* (UNESP. Online), v. 59, p. 155-175, 2015.

\_\_\_\_\_; 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 Linguística* (ALFAL), v. 32, p. 25-29, 2016.

\_\_\_\_\_; 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*, v. 70 (3), p. 15-32, 2017.

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M.; CROFT, W.; ELLIS, N.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language is a Complex Adaptive System: Position Paper. *Language Learning*, 59 (1), p. 1-26, 2009.

BETTONI-TECHIO, M. *Perceptual Training and word-initial /s/- clusters in Brazilian Portuguese/English Interphology*. 2008. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal de Santa Catarina, Florianópolis, 2008.

BOERSMA, P.; WEENINK, D. *Praat: Doing Phonetics by Computer*. Version 5.3.48. 2013. Disponível em: <www.praat.org>. Acesso em: 20 jan. 2015.

BRAWERMAN-ALBINI, A. *Os efeitos de um treinamento de percepção na aquisição do padrão acentual pré-paroxítono da língua inglesa por aprendizes brasileiros*. 2012. 333 f. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal do Paraná, Curitiba, 2012.

CHO, T.; LADEFOGED, P. Variation and Universals in VOT: evidence from 18 languages. *Journal of Phonetics*, V. 27, p. 207-229, 1999.

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

DE BOT, K.; LOWIE, W.; THORNE, S. L.; VERSPOOR, M. Dynamic System Theory as a comprehensive theory of second language development. In P. García Mayo, J. Gutierrez Mangado, & M. Martínez Adrián (Eds.), *Contemporary perspectives on second language acquisition*. John Benjamins Publishers, 2013, p. 167-189.

FLEGE, J. E.; Second Language Speech Learning: Theory, Findings, and Problems. In: STRANGE, W. (ed.) *Speech perception and linguistic experience: issues in cross language research*. Timonium, MD: York press, 1995, p. 233-277.

KAMPFF, F. R; ALVES, U. K. Efeitos do treinamento perceptual na identificação e produção de plosivas surdas iniciais do inglês por aprendizes porto-alegrenses. *II Encontro do Núcleo de Pesquisa em Fonética e Fonologia Aplicada à Língua Estrangeira (NUPFFALE) (Encontro)*. Curitiba, 2016.

KUPSKE, F. *Imigração, Atrito e Complexidade: a produção das oclusivas surdas iniciais do inglês e do português por brasileiros residentes em Londres*. Tese (Doutorado em Letras). Universidade Federal Do Rio Grande do Sul, Porto Alegre, RS, 2016.

LARSEN-FREEMAN, D. Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics*, 18 (2), p. 141-165, 1997.

\_\_\_\_\_. CAMERON, L. *Complex Systems and Applied Linguistics*. Oxford University Press, 2008.

LISKER, L.; ABRAMSON, A. S. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), p. 384-422, 1964.

LOWIE, W. M.; VERSPOOR, M. H. Variability and Variation in Second Language Acquisition Orders: A Dynamic Reevaluation. *Language Learning*, v. 65 (1), p. 63-88, 2017.

\_\_\_\_\_; \_\_\_\_\_. Individual Differences and the Ergodicity Problem. *Language Learning*, p. 1-23, 2018.

NOBRE-OLIVEIRA, D. *The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers*. 2007. 198 f. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal de Santa Catarina, Florianópolis, 2007.

PEREYRON, L. *A produção vocálica por falantes de Espanhol (L1), Inglês (L2) e Português (L3): uma perspectiva dinâmica na (multi) direcionalidade da transferência linguística*. Tese (Doutorado em Letras). Universidade Federal Do Rio Grande do Sul, Porto Alegre, RS, 2017.

PEROZZO, R. V. *Percepção de oclusivas não vozeadas sem soltura audível em codas finais do inglês (L2) por brasileiros: o papel do contexto fonético-fonológico, da*

*instrução explícita e do nível de proficiência*. Dissertação (Mestrado em Letras). Universidade Federal Do Rio Grande do Sul, Porto Alegre, RS, 2013.

PEROZZO, R. V. *Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa: para além do PAM-L2*. Tese (Doutorado em Letras). Universidade Federal Do Rio Grande do Sul, Porto Alegre, RS, 2017.

REIS, M.; NOBRE-OLIVEIRA, D. 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, p. 372-381, 2008.

RODRIGUES, G. S. *Efeitos do treinamento perceptual na produção das plosivas surdas iniciais do inglês por falantes brasileiros*. Monografia (Trabalho de Conclusão de Curso) - Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2015.

SCHERESCHEWSKY, L. C.; ALVES, U. K.; KUPSKE, F. First language attrition: the effects of English (L2) on Brazilian Portuguese VOT patterns in an L1 - dominant environment. *Letrônica*. Porto Alegre, RS. V. 10, n. 2, p. 700-716

SCHWARTZHAUPT, B. M. *Factors influencing Voice Onset Time: analyzing Brazilian Portuguese, English and interlanguage data*. Monografia (Trabalho de Conclusão de Curso) - Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2012.

\_\_\_\_\_. *Testing intelligibility in English: the effects of Positive VOT and contextual information in a sentence transcription task*. 2015. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.

\_\_\_\_\_; 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*. Belo Horizonte, MG. V. 23, n. 2, p. 311-334, 2015.

ZIMMER, M. C.; ALVES, U. K. On the status of terminal devoicing as an interlanguage process among Brazilian learners of English. *Ilha do Desterro*, v. 55, p. 41-62, 2008.

\_\_\_\_\_; \_\_\_\_\_. Uma visão dinâmica da produção da fala em L2: o caso da Dessonorização Terminal. *Revista da ABRALIN*, v. 11, p. 221-272, 2012.

\_\_\_\_\_; SILVEIRA, R.; ALVES, U. K. *Pronunciation Instruction for Brazilians: Bringing Theory and Practice Together*. 1. ed. Newcastle upon Tyne: Cambridge Scholars Publishing, 2009.

## ANEXOS

ANEXO A

Questionário de Experiência e Proficiência Linguística

### Questionário de Experiência e Proficiência Linguística

Nº de matrícula UFRGS: \_\_\_\_\_ Sexo: ( ) F ( ) M Curso: \_\_\_\_\_  
 Data de nascimento: \_\_\_\_/\_\_\_\_/\_\_\_\_ Local de nascimento: \_\_\_\_\_  
 CPF: \_\_\_\_\_ E-mail: \_\_\_\_\_

**1. Liste todas as línguas que você conhece na ordem em que foram adquiridas (1 sendo sua língua nativa):**

<b>Língua 1</b>		<b>Língua 3</b>	
<b>Língua 2</b>		<b>Língua 4</b>	

**2. Indique onde você aprendeu as suas línguas (marque tantas opções quantas forem necessárias):**

Língua 1	Língua 2	Língua 3	Língua 4
<input type="checkbox"/> Casa	<input type="checkbox"/> Casa	<input type="checkbox"/> Casa	<input type="checkbox"/> Casa
<input type="checkbox"/> Escola	<input type="checkbox"/> Escola	<input type="checkbox"/> Escola	<input type="checkbox"/> Escola
<input type="checkbox"/> Curso de línguas			
<input type="checkbox"/> Sozinho	<input type="checkbox"/> Sozinho	<input type="checkbox"/> Sozinho	<input type="checkbox"/> Sozinho
<input type="checkbox"/> Outro _____			

**3. Informe (se for o caso) a idade em que você:**

	Língua 1	Língua 2	Língua 3	Língua 4
Começou a aprender	_____ anos	_____ anos	_____ anos	_____ anos
Começou a utilizar ativamente	_____ anos	_____ anos	_____ anos	_____ anos
Tornou-se fluente	_____ anos	_____ anos	_____ anos	_____ anos

**4. Indique, em uma escala de 0 a 6 (0 = nada, 3 = razoavelmente, 6 = muito), o quanto cada um destes fatores contribuiu para a aprendizagem das suas línguas:**

	Língua 1	Língua 2	Língua 3	Língua 4
Interação com a família				
Interação com os amigos				
Leitura geral				
Leitura de textos acadêmicos				
Assistir televisão e filmes				
Ouvir rádio e/ou música				
Uso da internet				
Curso de línguas				
Outro				

**5. Informe o número de anos e meses que você passou em cada um destes ambientes:**

	Língua 1	Língua 2	Língua 3	Língua 4
País em que a língua é falada	_____ anos _____ meses			
Família em que a língua é falada	_____ anos _____ meses			
Escola / trabalho em que a língua é falada	_____ anos _____ meses			

**6. Marque com um X em que língua você realiza estas atividades e circule o número correspondente à frequência com que elas acontecem:**

1 = algumas vezes por ano      2 = uma vez por mês      3 = uma vez a cada duas semanas  
 4 = uma vez por semana      5 = mais de uma vez por semana      6 = diariamente

	Língua 1	Frequência	Língua 2	Frequência	Língua 3	Frequência	Língua 4	Frequência
Fala com seu pai		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6
Fala com sua mãe		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6
Fala com familiares		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6
Fala com amigos		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6
Fala no trabalho/faculdade		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6
Lê/escreve no trabalho/faculdade		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6		1 2 3 4 5 6

7. Estime a porcentagem de tempo em que você usa cada língua diariamente (o total deve ser 100%):

	% do tempo
Língua 1	
Língua 2	
Língua 3	
Língua 4	

8. Estime em número de horas o quanto você usa cada língua para as seguintes atividades diariamente:

	Língua 1	Língua 2	Língua 3	Língua 4
Assistir TV/Filmes				
Ouvir música				
Jogar videogames				
Ler (livros, revistas...)				
Ler (textos acadêmicos)				
Escrever				
Falar				

9. Circule em uma escala de 1 a 6 seu nível de proficiência nas línguas que conhece:

1 = muito baixo    2 = baixo    3 = razoável    4 = bom    5 = muito bom    6 = proficiente

**Língua 1**

Leitura	1	2	3	4	5	6
Escrita	1	2	3	4	5	6
Compreensão auditiva	1	2	3	4	5	6
Fala	1	2	3	4	5	6

**Língua 2**

Leitura	1	2	3	4	5	6
Escrita	1	2	3	4	5	6
Compreensão auditiva	1	2	3	4	5	6
Fala	1	2	3	4	5	6

**Língua 3**

Leitura	1	2	3	4	5	6
Escrita	1	2	3	4	5	6
Compreensão auditiva	1	2	3	4	5	6
Fala	1	2	3	4	5	6

**Língua 4**

Leitura	1	2	3	4	5	6
Escrita	1	2	3	4	5	6
Compreensão auditiva	1	2	3	4	5	6
Fala	1	2	3	4	5	6

10. Marque com um X em que língua você se sente mais confiante ao:

	Língua 1	Língua 2	Língua 3	Língua 4
Ler				
Escrever				
Compreender				
Falar				

11. Caso você já tenha realizado algum teste de proficiência, indique:

Língua	Teste	Ano	Pontuação

## ANEXO B

Tabelas com os resultados do Questionário de Experiência e Proficiência Linguística,  
de Scholl e Finger

<b>Grupo 1</b>									
<i>Participante</i>	1	2	3	4	5	6	7	Média	DP
<i>Idade em que começou a aprender a L2</i>	29	6	6	10	9	7	13	11,42	8,14
<i>Idade em que começou a utilizar a L2</i>	30	13	12	18	9	17	20	17	6,87
<i>Idade em que tornou-se fluente na L2</i>	31	18	19	19	16	17	25	20,71	5,37
<i>Meses em escola/trabalho da L2 (anos)</i>	0	7	3	4	12,5	3,6	3	4,72	3,99
<i>Percentagem de tempo de uso da L2</i>	40	10	30	30	40	42	30	31,71	10,98
<i>Horas de uso/dia da L2</i>	2,5	9,5	2	9	8,5	10,5	10	7,42	3,59
<i>Autoavaliação auditiva da L2</i>	3	5	4	4	5	4	6	4,42	0,97
<i>Autoavaliação da fala da L2</i>	3	5	4	5	3	4	4	4	0,81

**Nota:** DP = Desvio Padrão.

**Fonte:** Elaborada pelo autor.

<b>Grupo 2</b>										
<i>Participante</i>	8	9	10	11	12	13	14	15	Média	DP
<i>Idade em que começou a aprender a L2</i>	10	12	6	12	11	11	9	7	9,75	2,25
<i>Idade em que começou a utilizar a L2</i>	15	18	9	14	15	15	12	23	15,12	4,12
<i>Idade em que tornou-se fluente na L2</i>	22	19	17	18	16	20	17	0*	16,12	6,79
<i>Anos em escola/trabalho da L2</i>	8	4	16	9	3,5	12,6	13	2,5	8,57	5,00
<i>Porcentagem de tempo de uso da L2</i>	5	30	15	15	15	20	30	5	16,8	9,61
<i>Horas de uso/dia da L2</i>	9	6,5	14	8	4,5	8	9	2,6	7,7	3,40
<i>Autoavaliação auditiva da L2</i>	4	4	5	6	5	5	5	3	4,62	0,91
<i>Autoavaliação da fala da L2</i>	4	4	4	6	5	4	5	3	4,37	0,91

**Nota:** DP = Desvio Padrão.

\* A participante não se considera fluente na L2;

**Fonte:** Elaborada pelo autor.

## ANEXO C

Tabelas com valores de produção de ambos os grupos (valores mínimos e máximos, médias e desvio padrão)

P	G	Etapa	/p/			
			Mínimo	Máximo	Média	DP
1	1	Pré-teste	16,88	45,25	29,17	10,03
		Pós-teste	15,28	19,75	16,99	1,92
		Pós-teste Postergado 1	11,56	25,96	17,91	5,62
		Pós-teste Postergado 2	3,4	30,09	13,26	8,96
2	1	Pré-teste	24,94	39,02	32,05	5
		Pós-teste	16,48	40,36	30,78	8,41
		Pós-teste Postergado 1	16,19	38,53	30,24	7,68
		Pós-teste Postergado 2	14,45	34,58	22,12	9,16
3	1	Pré-teste	16	21,17	18,25	1,79
		Pós-teste	17,24	32,75	23,27	5,87
		Pós-teste Postergado 1	12,47	49,41	24,16	13,69
		Pós-teste Postergado 2	4,38	19,87	9,56	5,58
4	1	Pré-teste	11,79	27,5	18,84	5,88
		Pós-teste	10,59	29,66	15,7	6,94
		Pós-teste Postergado 1	11,9	27,06	18,97	5,83
		Pós-teste Postergado 2	10,24	26,72	18,5	6,46
5	1	Pré-teste	9,3	59,19	30,09	20,44
		Pós-teste	30,98	63,66	43,78	11,79
		Pós-teste Postergado 1	25,63	65,85	43,21	14,99
		Pós-teste Postergado 2	12,7	64,06	26,3	19,31
6	1	Pré-teste	14,29	53,83	29,4	13,97
		Pós-teste	15,99	23,45	20,19	2,8
		Pós-teste Postergado 1	13,24	171,72	59,93	56,63
		Pós-teste Postergado 2	7,45	63,76	23	20,61
7	1	Pré-teste	15,37	29,96	19,98	5,13
		Pós-teste	11,05	49,27	24,01	13,46
		Pós-teste Postergado 1	17,64	21,5	18,97	1,35
		Pós-teste Postergado 2	8,74	20,78	13,89	4,04
8	2	Pré-teste	16,86	47,22	23,82	11,73
		Pós-teste	3,89	16,3	10,29	5,7
		Pós-teste Postergado 1	6,36	19,73	13,75	4,9
		Pós-teste Postergado 2	2,2	70,37	15,23	27,08

Continua

P	G	Etapa	/p/			
			Mínimo	Máximo	Média	DP
9	2	Pré-teste	8,19	25,46	18,8	5,58
		Pós-teste	10,48	28,64	21,08	6,34
		Pós-teste Postergado 1	14,22	27,92	19,47	4,59
		Pós-teste Postergado 2	10,65	18,69	15,28	3,1
10	2	Pré-teste	7,75	37,76	14,03	11,77
		Pós-teste	11,97	72,59	25,96	23,1
		Pós-teste Postergado 1	24,15	66,32	37,47	15,8
		Pós-teste Postergado 2	9,55	29,39	14,56	7,46
11	2	Pré-teste	6,25	10,47	7,89	1,55
		Pós-teste	6,57	13,84	11,66	2,73
		Pós-teste Postergado 1	6,37	26,52	12,33	29,28
		Pós-teste Postergado 2	5,75	17,98	12,98	4,96
12	2	Pré-teste	12,63	21,12	16,21	3,5
		Pós-teste	15,08	39,01	23,87	8
		Pós-teste Postergado 1	23,29	38,94	30,83	6,7
		Pós-teste Postergado 2	13,17	34,77	22,12	7,07
13	2	Pré-teste	9,03	20,91	18,98	4,49
		Pós-teste	12,04	19,04	15,57	2,89
		Pós-teste Postergado 1	10,79	17,9	13,98	2,62
		Pós-teste Postergado 2	7,15	59,06	26,35	23,77
14	2	Pré-teste	16,66	42,49	31,17	11,19
		Pós-teste	12,96	47,25	29,41	13
		Pós-teste Postergado 1	14,95	85,75	48,93	29,28
		Pós-teste Postergado 2	7,57	97,31	35,29	32,52
15	2	Pré-teste	10,1	18,36	14,64	3,91
		Pós-teste	7,37	19,99	12,99	4,12
		Pós-teste Postergado 1	9,36	23,02	14,66	4,66
		Pós-teste Postergado 2	3,63	51,67	17,66	17,3

**Nota:** P = Participante; G = Grupo; DP = Desvio Padrão.

**Fonte:** Elaborada pelo autor.

P	G	Etapa	/t/			
			Mínimo	Máximo	Média	DP
1	1	Pré-teste	29,78	57,31	42,48	10,21
		Pós-teste	31,16	61,56	45,6	11,82
		Pós-teste Postergado 1	26,51	58,77	43,97	12,13
		Pós-teste Postergado 2	32,79	82,23	51,76	18,68
2	1	Pré-teste	34,81	84,45	54,64	16,5
		Pós-teste	38,77	96,26	60,39	22,2
		Pós-teste Postergado 1	38,8	96,2	60,76	22,04
		Pós-teste Postergado 2	43,34	82	61,59	13,13
3	1	Pré-teste	55,38	112,42	82,06	18,82
		Pós-teste	33,86	111,82	77,12	28,84
		Pós-teste Postergado 1	90,13	129,49	108,23	14,5
		Pós-teste Postergado 2	26,14	48,82	42,9	8,59
4	1	Pré-teste	26,45	70,55	41,98	16,03
		Pós-teste	31,06	76,92	47,83	15,88
		Pós-teste Postergado 1	37,95	67,98	49,2	10,7
		Pós-teste Postergado 2	18,24	26,76	22,47	3,39
5	1	Pré-teste	56,67	107,22	77,14	16,96
		Pós-teste	59,77	195,21	101,66	50,14
		Pós-teste Postergado 1	80,09	90,96	86,25	4,18
		Pós-teste Postergado 2	48,58	80,57	60,87	1,66
6	1	Pré-teste	57,67	123,13	91,94	25,47
		Pós-teste	62,97	113,12	85,9	17,52
		Pós-teste Postergado 1	82,03	115,84	98,33	14,8
		Pós-teste Postergado 2	70,74	101,9	88,75	13,24
7	1	Pré-teste	22,04	43,37	37,39	8,12
		Pós-teste	24,51	38,44	30,98	5,44
		Pós-teste Postergado 1	30,26	47,75	36,22	7,08
		Pós-teste Postergado 2	26,15	34,28	30,32	2,82
8	2	Pré-teste	39,78	84,59	72,15	17,08
		Pós-teste	43,12	84,94	62,69	15,62
		Pós-teste Postergado 1	64,03	98,35	78,87	11,32
		Pós-teste Postergado 2	32,69	75,45	51,68	15,5

Continua

P	G	Etapa	/t/			
			Mínimo	Máximo	Média	DP
9	2	Pré-teste	31,02	50,21	42,33	6,3
		Pós-teste	34,21	61,23	49,01	10,04
		Pós-teste Postergado 1	21,17	49,04	37,18	10,09
		Pós-teste Postergado 2	30,4	49,99	39,98	7,45
10	2	Pré-teste	34,41	101,48	67,07	21,64
		Pós-teste	44,61	85,11	64,15	14,73
		Pós-teste Postergado 1	62,54	96,39	79,14	12,02
		Pós-teste Postergado 2	24,74	72,65	48,78	14,22
11	2	Pré-teste	55,07	102,11	79,06	17,47
		Pós-teste	57,16	113,96	80,34	18,98
		Pós-teste Postergado 1	61,04	122,67	88,81	30,86
		Pós-teste Postergado 2	58,78	132,2	89,79	25,34
12	2	Pré-teste	44,14	70,35	51,98	9,98
		Pós-teste	33,26	72,14	50,62	15,03
		Pós-teste Postergado 1	53,85	72,89	65,3	8,15
		Pós-teste Postergado 2	39,94	76,02	50,79	13,36
13	2	Pré-teste	80,91	137,7	101,48	20,37
		Pós-teste	54,5	98,84	77,69	19,14
		Pós-teste Postergado 1	84,75	119,01	98,72	13,53
		Pós-teste Postergado 2	133,87	163,55	147,36	10,98
14	2	Pré-teste	18,03	96,26	64,83	33,62
		Pós-teste	43,6	121,28	81,28	28,78
		Pós-teste Postergado 1	15,46	101,75	61,35	30,86
		Pós-teste Postergado 2	28,74	75,72	49,53	20,29
15	2	Pré-teste	45,52	85,03	62,8	15,03
		Pós-teste	6,29	87,36	62,55	28,49
		Pós-teste Postergado 1	45,49	88,41	64,97	16,77
		Pós-teste Postergado 2	72,68	124,49	101,08	18,68

**Nota:** P = Participante; G = Grupo; DP = Desvio Padrão.

**Fonte:** Elaborada pelo autor.

P	G	Etapa	/k/			
			Mínimo	Máximo	Média	DP
1	1	Pré-teste	56,88	74,83	63,67	7,49
		Pós-teste	31,3	80,71	59,91	16,33
		Pós-teste Postergado 1	31,65	74,77	58,93	15,99
		Pós-teste Postergado 2	49,38	83,28	64,59	12,49
2	1	Pré-teste	48,07	80,46	64,72	12,79
		Pós-teste	55,77	83,55	71,63	12,24
		Pós-teste Postergado 1	55,82	87,44	71,59	13,01
		Pós-teste Postergado 2	58,79	131,49	99,29	26,58
3	1	Pré-teste	83,14	127,2	101,68	14,49
		Pós-teste	88,26	154,73	111,57	22,48
		Pós-teste Postergado 1	88,71	129,7	112	16,84
		Pós-teste Postergado 2	87,74	148,14	93,18	50,91
4	1	Pré-teste	73,98	121,58	107,73	17,84
		Pós-teste	69,82	110,74	89,33	16,65
		Pós-teste Postergado 1	54,68	117,11	94,23	24,69
		Pós-teste Postergado 2	16,25	124,14	82,32	37
5	1	Pré-teste	44,13	96,21	66,46	18,18
		Pós-teste	60,16	90,58	74,92	10,18
		Pós-teste Postergado 1	68,14	100,65	84,94	14,68
		Pós-teste Postergado 2	57,02	81,25	69,45	9,63
6	1	Pré-teste	86,94	111,68	95,61	9,25
		Pós-teste	71,48	129,88	104,07	21,34
		Pós-teste Postergado 1	82,84	122,93	102,15	16,79
		Pós-teste Postergado 2	65,31	87,58	76,03	10,16
7	1	Pré-teste	47,63	80,47	58,03	11,43
		Pós-teste	46,14	75,46	57,74	11,4
		Pós-teste Postergado 1	59,77	74,69	65,88	5,43
		Pós-teste Postergado 2	40,19	71,9	60,09	11,74
8	2	Pré-teste	53,7	80,87	67,53	9,62
		Pós-teste	61,61	80,04	72,78	6,79
		Pós-teste Postergado 1	43,06	95,8	62,65	19,56
		Pós-teste Postergado 2	53,59	106,42	83,63	19,92

Continua

P	G	Etapa	/k/			
			Mínimo	Máximo	Média	DP
9	2	Pré-teste	44,95	70,42	57,86	9,79
		Pós-teste	42,51	74,24	55,15	10,38
		Pós-teste Postergado 1	33,14	65,1	51,66	10,97
		Pós-teste Postergado 2	29,09	63,93	50,27	12,79
10	2	Pré-teste	81,44	94,91	88,84	5,01
		Pós-teste	69,24	108,43	86,05	12,92
		Pós-teste Postergado 1	99,38	113,18	103,87	5,07
		Pós-teste Postergado 2	78,05	107,01	92,47	10,78
11	2	Pré-teste	58,81	106,52	73,51	20,75
		Pós-teste	60,13	104,16	76,69	16,58
		Pós-teste Postergado 1	70,31	111,25	81,49	16,85
		Pós-teste Postergado 2	64,41	133,33	90,6	23,03
12	2	Pré-teste	64,81	114,4	82,82	18,01
		Pós-teste	75,19	96,03	85,41	8,66
		Pós-teste Postergado 1	54,26	128,82	91,96	26,34
		Pós-teste Postergado 2	72,34	110,11	88,41	14,97
13	2	Pré-teste	13,21	125,19	75,22	48,65
		Pós-teste	12,15	124,08	90,44	40,13
		Pós-teste Postergado 1	89,43	121,51	109,42	13,27
		Pós-teste Postergado 2	105,73	153,55	128,14	33,81
14	2	Pré-teste	59,89	111,7	91,31	18,3
		Pós-teste	68,67	165,25	118,12	33,8
		Pós-teste Postergado 1	90,26	156,98	125,61	29,05
		Pós-teste Postergado 2	76,36	116,19	91,51	14,22
15	2	Pré-teste	62,32	97,53	79,77	11,53
		Pós-teste	78,74	94,41	88,17	6,3
		Pós-teste Postergado 1	73,32	106,25	91,49	12,25
		Pós-teste Postergado 2	73,82	98,91	84,94	11,05

**Nota:** P = Participante; G = Grupo; DP = Desvio Padrão.

**Fonte:** Elaborada pelo autor.

## ANEXO D

### Termo de Compromisso Livre e Esclarecido



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE LETRAS



Grupo de Pesquisa  
'Cognição e Aquisição/Aprendizagem de L2: Uma perspectiva psicolinguística'  
Projeto de Pesquisa: **Efeitos de Treinamento Perceptual e Instrução Explícita na Aquisição da Distinção entre Plosivas Surdas e Sonoras Iniciais do Inglês (LE) por Aprendizes Argentinos e Brasileiros: O Papel de Pistas Acústicas Primordiais na L1**  
Prof. Dr. Ubiratã Kickhöfel Alves

### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado participante,

A presente investigação tem como objetivo a produção de conhecimento sobre as dificuldades de pronúncia em língua inglesa por parte de estudantes brasileiros. Esperamos contar com a sua valiosa participação.

Ao participar deste estudo, você realizará uma tarefa de percepção em língua estrangeira. Na frente de um computador, você ouvirá palavras da língua inglesa e será convidado a escolher a consoante inicial destas palavras.

Você também realizará uma leitura em voz alta de palavras em língua portuguesa e em língua inglesa. O material a ser lido será apresentado através de slides, em um computador laptop. Você lerá as frases apresentadas nos slides em voz alta, e estará utilizando um fone de ouvido acoplado a um microfone, para a captação de sua voz. Sua voz será gravada e armazenada, para posterior análise por parte do proponente desta pesquisa. A gravação será realizada em um ambiente silencioso, de forma individual, na presença somente do investigador responsável pelo estudo.

De modo a nos anteciparmos ao risco da identificação de sua identidade, queremos deixar claro que os participantes serão identificados nas transcrições dos dados de pronúncia e nos relatos de pesquisa apenas por um número de identificação.

Não há benefício direto para você ao participar do estudo, mas as descobertas poderão servir como fonte de consulta para estudiosos do processo de aquisição da pronúncia em língua estrangeira, bem como de metodologia de ensino e aprendizagem de línguas, além de formadores de professores de línguas. Salientamos que a leitura das palavras pode causar cansaço, ou ansiedade de sua parte.

Os resultados da pesquisa serão divulgados à comunidade acadêmica e à comunidade de educadores por meio de publicações, apresentações em eventos acadêmicos, oficinas de formação de professores, entre outras formas de divulgação.

Sua participação é essencial para a realização do trabalho de pesquisa, mas você tem a liberdade para se recusar a participar ou retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo.

Em caso de dúvida relacionada a seus direitos ou sobre sua participação nesta pesquisa, por favor, entre em contato com o coordenador do estudo.

Professor Ubiratã Kickhöfel Alves  
Sala 220 – Prédio Administrativo do Instituto de Letras  
E-mail: ukalves@gmail.com

Porto Alegre, maio de 2016.

Atenciosamente,

Prof. Dr. Ubiratã Kickhöfel Alves  
(PPG-LETRAS/UFRGS)

Participante: \_\_\_\_\_