

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE LETRAS

CAMILA SAVICZKI MOTTA

PERCEPÇÃO DE PLOSIVAS SURDAS DO INGLÊS SOB MÚLTIPLAS  
MANIPULAÇÕES DE *VOICE ONSET TIME* (VOT) EM TAREFA DE  
IDENTIFICAÇÃO POR BRASILEIROS E AMERICANOS

Porto Alegre

2014

CAMILA SAVICZKI MOTTA

PERCEPÇÃO DE PLOSIVAS SURDAS SOB MÚLTIPLAS MANIPULAÇÕES DE  
*VOICE ONSET TIME* (VOT) EM TAREFA DE IDENTIFICAÇÃO DO INGLÊS POR  
BRASILEIROS E AMERICANOS

Monografia apresentada como requisito parcial para a obtenção do grau de Licenciada em Letras Português/Inglês e Respectivas Literaturas da Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

Porto Alegre

2014

*Este trabalho é dedicado ao Felipe,  
pela devoção que a ele devo  
e por ser minha razão de seguir adiante.*

## **AGRADECIMENTOS**

Primeira, e mais especialmente, ao Felipe, por ser a minha maior inspiração e motivação, inclusive acadêmica. Pelo exemplo de dedicação que me representa, por me motivar e me impulsionar. Por acreditar em mim nos momentos mais importantes e menos esperados. Por estar ao meu lado quando esmoreci e, principalmente, por nunca me deixar desistir. O apoio, a força e o carinho por ele dedicados a mim foram de fundamental importância para a conclusão dessa etapa de minha vida.

À família Alves Pereira Avila pela motivação, e, em especial à Raquel, por ter sido suave e bela o bastante para me tranquilizar o coração nos momentos de grande preocupação e por me demonstrar força e determinação para me levar adiante.

Aos participantes dessa pesquisa, por dedicarem horas de seus dias a responder os testes que foram a eles pedidos e, sem os quais, não haveria trabalho.

Ao meu pai, por ser o responsável pela minha entrada no mundo das letras, e da Letras.

Ao Bira, pela orientação que me dedicou. Pela oportunidade de crescimento acadêmico que me proporcionou e também por ser responsável pela ampliação dos meus horizontes no campo do conhecimento, no qual ainda muito tenho a perseguir.

*“O jardim já é lindo*

*O vento soprou*

*Semente nasceu.”*

Felipe Alves Pereira Avila

## RESUMO

Este trabalho tem como principal objetivo verificar e analisar de que forma brasileiros aprendizes de inglês (L2) e americanos identificam o vozeamento das consoantes plosivas iniciais da língua inglesa em palavras monossilábicas; nesta língua, para a diferenciação de vozeamento, tem-se como principal pista acústica a aspiração, que pode ser medida a partir dos valores de *Voice Onset Time* (VOT). O VOT é medido em milissegundos e pode ser classificado em 3 diferentes padrões: VOT NEGATIVO, característico de /b/, /d/ e /g/ em português; VOT ZERO, encontrado, concomitantemente, em /p/, /t/ e /k/ no português e em /b/, /d/ e /g/ no inglês e, finalmente, o VOT POSITIVO, encontrado em [p<sup>h</sup>], [t<sup>h</sup>] e [k<sup>h</sup>] no inglês. Estudos anteriores (SCHWARTZHAUPT, ALVES & FONTES, 2013; ALVES & MOTTA, 2013) demonstraram que a manipulação do intervalo de VOT de valor positivo, a fim de que parecesse com um VOT de valor zero, pode resultar em diferentes índices de identificação e discriminação, dependendo da natureza do participante: brasileiros aprendizes de inglês como língua adicional tendem a não diferenciar os segmentos com VOT positivo daqueles que receberam manipulação acústica, o que sugere que a pista acústica VOT provavelmente não se mostre primordial ou prioritariamente distintiva para brasileiros. Participantes americanos, por outro lado, tendem a identificar as consoantes manipuladas de acordo com os padrões previstos para a língua inglesa. Essa identificação, contudo, não é categórica. Assim sendo, o instrumento desse estudo foi baseado na manipulação múltipla dos valores de VOT positivo num contínuo temporal, para que houvesse 5 diferentes padrões de duração de VOT manipulado, com corte de aspiração gradual e proporcional à produção original, a fim de que se verificasse a possível existência de um limiar fonético que influenciasse ou demarcasse uma porcentagem mínima de aspiração necessária para que os americanos e brasileiros identificassem uma determinada consoante como sendo surda ou sonora. O instrumento de coleta consiste em uma tarefa de identificação. Este teste de identificação foi composto por 60 estímulos-alvo (6 *types*, subdivididos em 5 estímulos diferentes entre si, oriundos de um estímulo original em comum = 30 – 2 para cada ponto de articulação – repetidos em duas rodadas = 60) mais 12 estímulos com consoantes sonoras, sendo 4 por ponto de articulação, totalizando 72 *tokens* por participante. Os resultados indicam que a manipulação múltipla das consoantes plosivas surdas se mostra capaz de confundir os

participantes, principalmente os americanos. Foram encontrados, também, indícios de que, além da pista acústica *Voice Onset Time*, outras pistas podem atuar em conjunto com o VOT no momento da identificação das consoantes, tanto por parte dos participantes americanos quanto pelos brasileiros. O trabalho teve como meta evidenciar os limiares fonéticos, referentes à fronteira entre 'surdo' e 'sonoro', para brasileiros e americanos. Espera-se que este estudo tenha apresentado evidências adicionais a respeito do caráter fundamental desta pista para tais falantes nativos, bem como confirmar o caráter não-primordial ou suficiente desta pista entre brasileiros.

*Palavras-chave:* *Voice Onset Time*; manipulação de VOT; percepção de sons em L2; aquisição fonético-fonológica.

## ABSTRACT

The objective of this study is to verify and analyze how Brazilian learners of English (L2) and Americans identify voicing patterns in stop consonants in English monosyllabic words, taking aspiration as main acoustical cue, which can be measured by *Voice Onset Time* (VOT) values. The VOT is measured in milliseconds and can be classified within three different patterns: NEGATIVE VOT, which is characterized by pre-voicing produced before voiced consonants in Brazilian Portuguese /b/, /d/, /g/ (which can also be found sometimes in the English language); ZERO VOT (which is characterized by low values of aspiration, found in voiceless plosives in Brazilian Portuguese /p/, /t/, /k/, and, at the same time, in the voicing consonants of English /b/, /d/, /g/. Finally, POSITIVE VOT, which is characteristic of voiceless stops in English [p<sup>h</sup>], [t<sup>h</sup>] and [k<sup>h</sup>]. Some previous studies (SCHWARTZHAUPT, ALVES & FONTES, 2013; ALVES & MOTTA, 2013) showed that manipulation in positive VOT consonants, in order to sound similar to zero VOT, may result in different consonant perceptions, depending on the group of participants: Brazilian learners of English as additional language tend not to discriminate the difference between the natural positive VOT and the manipulated VOT. This situation probably happens because VOT does not seem to be an important and distinctive acoustical cue for Brazilians. American participants, on the other hand, tend to identify the manipulated consonants according to the patterns that are normally presented for the English language. Based on this, the task developed for this study was based on multiple manipulation of VOT values in a *continuum*, in order to get 5 different manipulated slices of a positive VOT, which was gradually and proportionally manipulated according to its original VOT production, in order to investigate if there is a phonetic-phonological threshold which could influence and determine a minimal portion of VOT value necessary to consider a consonant as being voiced or voiceless by Americans and Brazilians. The type of test used in this study was an identification task. This task is composed by 60 stimuli (6 types, subdivided in 5 different categories, which came from the same original stimulus = 30. There are 2 stimuli for each place of articulation, repeated twice during the task = 60). In addition, there are 12 *stimuli* with voiced consonants, resulting in a task with 72 tokens per participant. The results indicate that multiple manipulations of

voiceless stop consonants seem to confuse the participants, especially the Americans. There is also evidence that, in addition to VOT, other cues may act in conjunction with this cue, when both American and Brazilian participants identify stop consonants in English. This study aimed to highlight the phonetic-phonological thresholds between 'voiceless' and 'voiced' stops. The results obtained not only provide additional evidence on the fundamental role of this cue for native speakers of English, but also confirm the non-essential nature of this cue in the voice distinctions by Brazilians.

Keywords: *Voice Onset Time*, VOT manipulation, sound perception in L2, phonetic-phonological acquisition.

## **LISTA DE QUADROS**

QUADRO 1 – INFORMAÇÕES DOS PARTICIPANTES BRASILEIROS (P.33)

QUADRO 2 – INFORMAÇÕES DOS PARTICIPANTES AMERICANOS (P. 40)

## LISTA DE IMAGENS

- FIGURA 1: O PÊNDULO COMO EXEMPLO DE SISTEMA DINÂMICO (P.23)
- FIGURA 2: VOT NEGATIVO COM DURAÇÃO DE 133MS, PRODUZIDO POR FALANTE NATIVO DO INGLÊS AMERICANO, DA PALAVRA "GILL" (P.27)
- FIGURA 3: VOT COM PLOSIVA DE PADRÃO ZERO, PRODUZIDA POR FALANTE NATIVO DO INGLÊS AMERICANO, DA PALAVRA "BIT" (P.28)
- FIGURA 4: VOT POSITIVO COM DURAÇÃO DE 80M, PRODUZIDA POR FALANTE NATIVA DO INGLÊS AMERICANO, DA PALAVRA "TICK" (P.29)
- FIGURA 5 – SELEÇÃO DE NÍVEL DO TESTE (P.34)
- FIGURA 6 – EXEMPLO 1 DE USE OF ENGLISH (P.35)
- FIGURA 7 – EXEMPLO 2 DE USE OF ENGLISH (P.35)
- FIGURA 8 – EXEMPLO 3 DE USE OF ENGLISH (P.36)
- FIGURA 9 – EXEMPLO 4 DE USE OF ENGLISH (P.37)
- FIGURA 10 – EXEMPLO 1 DE LISTENING (P.38)
- FIGURA 11 – EXEMPLO 2 DE LISTENING (P.38)
- FIGURA 12 – TELA DE FINALIZAÇÃO DO TESTE (P.39)
- FIGURA 13 – TABELA DE RESULTADO OOPT (P.40)
- FIGURA 14 – CONSENT FORM DISPONIBILIZADO AOS AMERICANOS ONLINE (P.42)
- FIGURA 15 – EXEMPLO DE VOT 100% SELECIONADO PARA MANIPULAÇÃO – PALAVRA "KILL", COM DURAÇÃO DE 94,2MS. (LOCUTOR 1) (P45)
- FIGURA 16 – VOT DE "KILL", COM ASPIRAÇÃO DE 94,2, COM 25% DA DURAÇÃO SELECIONADA, PARA SER POSTERIORMENTE RETIRADA (LOCUTOR 1) (P.45)
- FIGURA 17 – VOT DE "KILL", AGORA COM DURAÇÃO MÉDIA DE 70MS, COM 25% DA SUA ASPIRAÇÃO JÁ RETIRADA (LOCUTORA 1) (P.46)
- FIGURA 18: PRIMEIRA PÁGINA DO TESTE DE IDENTIFICAÇÃO PARA OS BRASILEIROS (P. 50)
- FIGURA 19: INFORMAÇÃO AOS PARTICIPANTES (P.51)
- FIGURA 20: LAYOUT DO TESTE DE IDENTIFICAÇÃO BRASILEIROS (P. 52)
- FIGURA 21: CONFIRMAÇÃO DE ENTENDIMENTO DO TESTE (P. 52)
- FIGURA 22: MONTAGEM DA RANDOMIZAÇÃO PELO RESEARCH RANDOMIZER (P. 54)
- FIGURA 23: EXEMPLO DO RANDOMIZADOR *RESEARCH RANDOMIZER* (P. 55)

## LISTA DE TABELAS

TABELA 1 – VALOR DE VOT POR *TOKEN* E LOCUTOR (P.46)

TABELA 2: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE BILABIAL PLOSIVA SURDA PARA AMERICANOS (P. 56)

TABELA 3: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE ALVEOLAR SURDA PARA AMERICANOS (P. 58)

TABELA 4: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE CORONAL SURDA PARA AMERICANOS (P. 60)

TABELA 5: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE BILABIAL PLOSIVA SURDA PARA BRASILEIROS (P.63)

TABELA 6: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE BILABIAL PLOSIVA SURDA PARA BRASILEIROS (P.64)

TABELA 7: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE BILABIAL PLOSIVA SURDA PARA BRASILEIROS (P. 65)

TABELA 8 - ÍNDICES DE IDENTIFICAÇÃO PARA /K/ E SUAS MANIPULAÇÕES POR NÍVEL DE PROFICIÊNCIA (70)

TABELA 9 - ÍNDICES DE IDENTIFICAÇÃO PARA /T/ E SUAS MANIPULAÇÕES POR NÍVEL DE PROFICIÊNCIA (71)

TABELA 10 - ÍNDICES DE IDENTIFICAÇÃO PARA /K/ E SUAS MANIPULAÇÕES POR NÍVEL DE PROFICIÊNCIA (71)

## **LISTA DE GRÁFICOS**

GRÁFICO 1 – ÍNDICES DE IDENTIFICAÇÃO POR PONTO DE ARTICULAÇÃO AMERICANOS (P.62)

GRÁFICO 2 – ÍNDICES DE IDENTIFICAÇÃO POR PONTO DE ARTICULAÇÃO BRASILEIROS (P. 67)

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>16</b>
1.2 Objetivos Específicos.....	18
1.3 Hipóteses.....	18
<b>2 REFERENCIAL TEÓRICO</b> .....	<b>20</b>
2.1 Uma Perspectiva Dinâmica de Aquisição de Segunda Língua .....	23
2.2 <i>Voice Onset Time</i> (VOT).....	28
<b>3 MÉTODO</b> .....	<b>34</b>
3.1 Participantes .....	34
3.2 Participantes Brasileiros .....	34
3.3 <i>Oxford Online Placement Test</i> (OOPT) .....	36
3.4 Participantes Americanos .....	42
3.5 Palavras-alvo .....	44
3.6 Gravação dos Estímulos.....	45
3.7 Da Manipulação dos Estímulos .....	46
<b>4 INSTRUMENTO</b> .....	<b>50</b>
4.1 Teste de Identificação.....	50
4.2 A Plataforma Online Utilizada: <i>Surveygizmo</i> .....	51
<b>5 ANÁLISE</b> .....	<b>59</b>
5.1 Descrição Geral dos Dados de Americanos .....	60
5.2 Descrição Geral dos Dados de Brasileiros .....	68
5.3 Descrição dos Dados de Brasileiros em relação à proficiência .....	76
5.4 Discussão dos Resultados.....	80
<b>6 CONSIDERAÇÕES FINAIS</b> .....	<b>86</b>
<b>7 REFERÊNCIAS</b> .....	<b>89</b>



## 1 INTRODUÇÃO

Estudos dirigidos à análise do componente fonético-fonológico na aquisição de segunda língua, mais especialmente da língua inglesa, têm se tornado cada vez mais especializados e rigorosos. Parte dessa especialização e rigor se devem ao acesso universal aos *softwares* de análise livres e de fácil manipulação. Parte se deve, também, ao interesse de caráter linguístico, que evoluiu e se modificou com o passar do tempo, voltando-se para a complexidade do processo de aquisição fonético-fonológica de uma nova língua. Essa complexidade reside na existência de dois sistemas fonológicos que coexistem e se afetam na configuração da língua do aprendiz. Quando se aprende uma língua estrangeira<sup>1</sup>, há uma gama de fatores que precisam ser, de certa forma, ajustados a fim de se adaptar ao novo sistema linguístico que está sendo adquirido. Segundo Flege e Munro (1995), em seu Speech Learning Model (SLM), e Best e Tyler (2007) no seu Perceptual Assimilation Model (PAM-L2), os aprendizes brasileiros de inglês como segunda língua tenderiam, por exemplo, a não fazer distinções de vozeamento entre plosivas surdas e sonoras da língua-alvo, levando-se em consideração que segmentos aspirados não ocorrem na língua materna desses aprendizes. Isso seria uma possível explicação para processos de transferências fonético-fonológicas feitas por aprendizes de língua estrangeira, no que diz respeito aos baixos índices de produção de aspiração nos segmentos plosivos iniciais do inglês, fenômeno esse já vastamente citado na literatura (VEIGA-FRANÇA, 2011; ALVES, SCHWARTZHAUPT, BARATZ, 2011; SCHWARTZHAUPT, 2013). Na tentativa de se identificarem e observarem tais fenômenos, o desenvolvimento de testes perceptuais pode ser uma forma razoavelmente eficaz de se apontar a que tipos de transferências aprendizes brasileiros podem estar sujeitos quando expostos a estímulos específicos de língua inglesa.

O VOT (*Voice Onset Time* - a ser melhor definido e exemplificado na seção do Referencial Teórico, capítulo 2) corresponde ao período de tempo existente entre a soltura de uma consoante plosiva e a vibração das cordas vocais do segmento vocálico que segue tal consoante (YAVAS, 2008). Existem três padrões de VOT, que se cruzam, quando comparados português e inglês, nos padrões definidos para

---

<sup>1</sup> Neste trabalho, os termos “Língua Adiciona (LA)”, “Língua Estrangeira (LE)” e “Segunda Língua (L2)” serão tratados como sinônimos.

plosivas surdas e sonoras: enquanto /b/, /d/ e /g/ são enquadrados no padrão de VOT negativo na língua portuguesa, tais consoantes recebem classificação de VOT zero na língua inglesa. Soma-se, ainda, a classificação de /p/, /t/ e /k/ como sendo pertencente ao padrão de VOT zero no português e, ao mesmo tempo, ao padrão positivo na língua inglesa, sendo caracterizadas pela aspiração que as segue, sendo pronunciadas como [p<sup>h</sup>], [t<sup>h</sup>] e [k<sup>h</sup>].

Segundo Alves e Motta (2013), ainda se fazem necessários estudos adicionais para que se encontrem mais evidências sobre as pistas acústicas prioritariamente relevantes para cada língua no que diz respeito à distinção de vozeamento de plosivas entre L1 e L2. O trabalho supracitado tinha por objetivo a investigação sobre o papel da pista acústica *Voice Onset Time* (VOT) e o *status* fonológico atribuído a tal pista após sua manipulação. Apenas participantes brasileiros aprendizes de língua inglesa, nivelados como de nível de proficiência intermediária e avançados, segundo o Oxford Online Placement Test (PURPURA, 2007), fizeram parte do estudo. No estudo de 2013, chegou-se à conclusão de que, muito possivelmente, brasileiros de diferentes níveis de proficiência não seguem a pista acústica VOT como primordial no momento de identificar e discriminar consoantes plosivas iniciais da língua inglesa com VOT manipulado (cortado). Aparentemente, os aprendizes brasileiros usariam a intensidade do *burst* como pista acústica primordial no momento de identificar as consoantes entre *surdas* ou *sonoras*, enquanto os americanos tenderiam a usar o VOT como pista primordial para tal distinção (SCHWARTZHAUPT, ALVES & FONTES, 2013).

Tendo os estudos supracitados como um dos motivadores do presente trabalho, além dos anteriormente desenvolvidos com propósitos semelhantes, como o de Alves, Schwartzhaupt e Baratz (2011), com manipulação do VOT de padrão positivo a fim de se soasse o mais próximo possível do padrão zero, o presente trabalho diferencia-se, principalmente, pela manipulação múltipla da pista acústica VOT, enquanto, até então, as manipulações de tal pista eram feitas de forma única, sem gradiência (ou seja, cortava-se todo o intervalo de aspiração). Pretende-se, com o presente estudo, a busca de argumentos adicionais para a verificação de que, de fato, americanos seguem a pista acústica VOT no momento da distinção de

sonoridade para plosivas, enquanto brasileiros usem o *burst* como pista primordial, como proposto por Schwartzhaupt, Alves e Fontes (2013).

Assim sendo, o objetivo principal deste estudo é verificar o *status* da pista *Voice Onset Time* e a maneira como essa pista pode influenciar a identificação de sonoridade entre plosivas surdas e sonoras do inglês, por parte de brasileiros e americanos.

Partindo-se do objetivo geral já mencionado, adicionam-se, no que segue, os objetivos específicos do estudo, seguidos de suas hipóteses.

## 1.2 Objetivos Específicos

(a) Verificar, nos dados dos sujeitos americanos, se graus distintos de manipulação de VOT implicam diferenças na identificação dos segmentos como "surdos" ou "sonoros";

(b) Verificar, nos dados dos sujeitos brasileiros, se graus distintos de manipulação de VOT implicam diferenças na identificação dos segmentos como "surdos" ou "sonoros";

(c) Investigar, nos dados dos participantes brasileiros, se diferenças nos níveis de proficiência dos aprendizes implicam índices distintos de identificação dos segmentos como "surdo" ou "sonoro".

## 1.3 Hipóteses

Tomando-se por consideração os estudos semelhantes já mencionados acima, apresentam-se, abaixo, as hipóteses referentes aos objetivos supracitados:

**H1:** Considerando-se as respostas dos sujeitos americanos em cada um dos pontos de articulação individualmente, haverá diferenças significativas em cada um dos índices de resposta ("surda", "sonora" e "erro") a partir dos diferentes graus de manipulação do intervalo de VOT;

**H2:** Considerando-se as respostas dos sujeitos brasileiros, em cada um dos pontos de articulação individualmente, não haverá diferenças significativas em cada um dos índices de resposta ("surda", "sonora" e "erro") a partir dos diferentes graus de manipulação do intervalo de VOT.

**H3:** Considerando-se cada um dos pontos de articulação individualmente, não haverá diferenças significativas nos índices de resposta "surda", "sonora" e "erro" entre os dois níveis de proficiência dos aprendizes.

Para que a estrutura do trabalho ficasse organizada de forma a torná-lo mais compreensível para o leitor, dividiu-se o presente estudo em capítulos, seções e sub-seções, brevemente apresentadas abaixo:

O primeiro capítulo, nomeado de **Introdução**, pretende apresentar ao leitor uma breve apresentação das bases teóricas que sustentam este trabalho, assim como uma pequena listagem de estudos semelhantes já desenvolvidos, seguidos do objetivo principal, dos objetivos específicos e de suas respectivas hipóteses.

No segundo capítulo, intitulado de **Referencial Teórico**, apresenta-se uma discussão de estudos mais específicos sobre as teorias fonéticas e fonológicas para a aquisição de segunda língua, assim como a definição e exemplificação do que vem a ser a pista acústica *Voice Onset Time* e a concepção de língua adotada neste estudo, que se enquadra na teoria dos sistemas dinâmico, sendo a língua um sistema complexo e de âmago cognitivo.

No capítulo seguinte, será apresentada a metodologia do presente trabalho. O capítulo sobre o **Método** versará sobre as características dos informantes consultados neste estudo, a apresentação da plataforma online utilizada nas coletas, bem como de que forma se deu a gravação dos estímulos utilizados e a explicação detalhada da manipulação múltipla que foi aplicada à pista aqui estudada.

No capítulo de **Análise**, os dados serão descritos e analisados à luz das hipóteses propostas. As análises serão apresentadas por forma de tabelas, gráficos e explicações, para que a linguagem gráfica ajude o leitor a visualizar de forma gradiente os dados analisados, já que seria esse o objetivo principal do trabalho. Finalmente, no capítulo das **Considerações Finais**, haverá uma reflexão e um balanço do trabalho como um todo, a explicitação das contribuições deste estudo para as pesquisas sobre aquisição em segunda língua e também das limitações, das falhas e dos pontos a serem melhorados em estudos vindouros.

## 2 REFERENCIAL TEÓRICO

As consoantes plosivas surdas /p/, /t/ e /k/ e as sonoras /b/, /d/ e /g/ são segmentos muito presentes em várias, senão em todas, línguas do mundo (LADEFOGED & MADDIESON, 1996; YAVAS, 2008). Apesar de compartilharem características fonético-fonológicas, elas apresentam particularidades em suas realizações a depender do sistema lingüístico em que ocorrem. Vários estudos, realizados em diferentes línguas, como português, inglês e espanhol, dentre outras (LISKER & ABRAMSON 1964; COHEN, 2004; YAVAS, 2008; BANDEIRA & ZIMMER, 2012; SCHWARTZHAUPT, 2013; ALVES & MOTTA, 2013) mostram que a investigação acústica de tais segmentos pode ser de grande relevância para o maior entendimento das relações fonético-fonológicas estabelecidas pelos falantes de línguas estrangeiras, especialmente àqueles que adquirem a língua inglesa como língua adicional. Levando-se em consideração o português e o inglês, podemos dizer que, em ambas as línguas, tais consoantes se fazem presentes nos seus respectivos inventários fonológicos, em três diferentes pontos de articulação e sendo distinguidas por vozeamento, sendo capazes de formar pares mínimos (Português = ‘bata’ – ‘pata’; ‘deu’ – ‘teu’; ‘gata’ – ‘cata’, e Inglês ‘bit’ – ‘pit’; ‘dip’ – ‘tip’; ‘git’ – ‘kit’) (YAVAS, 2008). Apesar da similaridade que possuem, no que diz respeito à categorização fonológica, a realização fonética desses fonemas difere de uma língua para a outra, conforme será visto neste capítulo.

Apesar de o estudo-base e promissor referente à produção dos segmentos plosivos ser da primeira metade dos anos 60 (LISKER e ABRAMSON, 1964), as técnicas utilizadas nas análises foram, com o passar do tempo, se tornando cada vez mais aprimoradas, o que tornou os estudos ainda mais ricos e rigorosos. Sabe-se que, em contextos reais de fala, onde, de fato, as línguas “acontecem”, não é possível se chegar - de forma oitiva - a pistas acústicas tão finas, de forma consciente e clara, como fazemos quando temos o auxílio de programas de computador que são capazes de fragmentar a massa sonora de palavras em ondas e transformá-las em unidades passíveis de análise. Entretanto, ao mesmo tempo, sabe-se que nossos ouvidos são capazes de captar, inconscientemente, aspectos tão finos que possivelmente não saibamos, ainda, classificar, ou que cujos efeitos

somente podem ser verificados em conjunto com outras pistas acústicas. Nesta incerteza de pistas cujos papéis funcionais talvez ainda não se encontrem plenamente descobertos, encontram-se estudos dirigidos à inteligibilidade e à compreensibilidade, que discutem que a forma de produzir os sons de uma língua estrangeira deve, de alguma forma, se adequar a alguma convenção, a fim de se fazer a comunicação entre locutor e interlocutor bem-sucedida. Apesar de já existirem inúmeros estudos dedicados a esse tema, ainda há, atualmente, um choque na mais básica definição do que vem a ser esse código que é responsável por fazer um falante entender o outro. Talvez esse impasse se deva, em certa parte, pela dificuldade de mensurar o que é “inteligível” e “compreensível”<sup>2</sup>, o que nos faz partir para a análise de unidades cada vez menores, capazes de se mensurar, para que se encontrem possíveis pistas que interfiram na comunicação entre falantes de inglês nativos e aprendizes. Para analisar tais pistas, a ajuda de programas de computador se faz necessária, já que eles são capazes de explicitar detalhes acústicos que nosso cérebro, como um complexo sistema fisiológico, não consegue, conscientemente, processar e analisar.

A disponibilidade de se encontrarem programas de alta qualidade e de forma gratuita na internet facilitou ainda mais o avanço na área de estudos de fonética e fonologia, sejam os pesquisadores linguistas ou fonoaudiólogos. Como exemplo de programa muito utilizado temos o *Praat* (do holandês, “fala”), desenvolvido, no ano de 1995, pelos pesquisadores holandeses Paul Boersma e David Weenink. Com atualizações recentes e recorrentes, o programa pode ser facilmente baixado pelo site do *software*<sup>3</sup>. Com tais programas, especificamente utilizados para se fazerem análises acústicas de sons da fala, é possível chegar a imagens que representam os sons de forma física e passíveis de análises mensuráveis. No que diz respeito aos segmentos plosivos, uma dessas pistas acústicas é o *Voice Onset Time* (VOT), que, como o nome em inglês sugere, seria o “tempo de início de vozeamento” após a explosão dos segmentos plosivos. Esse tempo pode diferir bastante entre as línguas e, por isso, vem se tornando cada vez mais analisado em estudos de aquisição de

---

<sup>2</sup> Definir-se-á, brevemente, os termos “inteligibilidade” como o quanto de uma conversação pode ser compreendida e “compreensibilidade” como sendo a qualidade atribuída àquilo que foi compreendido.

<sup>3</sup> Disponível em [www.praat.org](http://www.praat.org).

línguas estrangeiras, já que pode vir a ser uma pista acústica fundamental para a distinção entre segmentos surdos e sonoros, nas línguas do mundo.

Por ser o VOT uma importante pista de caráter gradiente, estudos dirigidos a esse fenômeno podem nos ajudar a entender e verificar de que forma se dá e se desenvolve a percepção e a produção dos segmentos plosivos, tanto em língua materna quanto em língua estrangeira. Essa pista, e a forma como ela é percebida e produzida, pode indicar possíveis transferências fonético-fonológicas, muito comuns a aprendizes de qualquer língua, já que os aprendizes tendem a tentar adaptar o sistema fonológico da segunda língua ao da língua materna (*cf.* BEST e TYLER, 2007).

De acordo com Zimmer, Silveira e Alves (2009), há dois possíveis tipos de transferências que parecem ser muito recorrentes em estágios de aprendizagem de uma segunda língua, sejam elas: 1) a transferência fonético fonológica, que se instancia durante a percepção e a produção de uma segunda língua; 2) a transferência grafo-fônico-fonológica, que se refere à relação entre ortografia, representação fonológica e fones produzidos em voz alta, ou percebidos e identificados quando ouvidos pelo aprendiz. Conforme Flege & Munro (2005) e Best & Tyler (2007), quando a transferência fonético-fonológica acontece, essa se deve ao fato de os aprendizes estarem tentando adequar a identificação dos sons da segunda língua a padrões já conhecidos, explicitando certa dificuldade em diferenciar os padrões pertencentes à segunda língua e no que tais padrões diferem da sua língua materna.

No que diz respeito ao segundo tipo de transferência, Zimmer (2004) afirma, entretanto, que quando a hipótese da transferência grafo-fônico-fonológica se faz presente, há um sinal de que não apenas o conhecimento fonológico de ambas as línguas (L1 e L2) está por trás da percepção e produção de sons de uma segunda língua, mas, também, os sistemas alfabéticos de cada uma delas está em jogo. A partir de testes de percepção e produção, é possível detectar se tais transferências estão sendo realizadas pelos aprendizes e especular o porquê disso acontecer. É por isso que o objetivo desse trabalho está focado na percepção do VOT por brasileiros aprendizes de inglês de diferentes níveis de proficiência, e também por americanos, falantes nativos, a fim de que se comparem de que forma os *stati*

atribuídos à questão de sonoridade dos segmentos /p/, /b/, /t/, /d/, /k/ e /g/ podem diferir entre falantes que compartilham diferentes sistemas de L1.

## 2.1 Uma Perspectiva Dinâmica de Aquisição de Segunda Língua

De acordo com alguns estudiosos da área de aquisição de segunda língua (ELLIS, 2005; ZIMMER e ALVES, 2012), a aquisição de uma língua, sendo ela estrangeira ou materna, faz parte de um processo cognitivo, e não “apenas” puramente linguístico (visto que não se poderia pensar a capacidade linguística como isolada de outras habilidades). Ellis (2005) afirma, inclusive, que a aquisição de uma língua é um processo, por excelência, cognitivo. A experiência adquirida e vivida, o *input* recebido, a maneira como a linguagem se desenvolve no cérebro não pode ser considerada como estando à parte do desenvolvimento biológico e comportamental de uma criança (como o processo de aprender a caminhar, por exemplo), e também das novas experiências que continuamos a ter mesmo depois de adolescentes e adultos (como dirigir, andar de bicicleta, aprender a desenhar, etc.). Na área linguística, encontra-se a corrente *emergentista*, que considera que o processo de aquisição de uma L2, é, como qualquer outra capacidade cognitiva, dinâmica. O que interfere, entretanto, nesse processo, são múltiplas variáveis que não são estáticas (como o próprio modelo dinâmico prevê), devendo, portanto, ser ajustadas de acordo com o *corpus* que se deseja analisar. Assim sendo, é sempre importante estar ciente de que os resultados a que temos acesso, a interpretação que fazemos deles e as consequências de nossos estudos em relação ao vasto campo da linguagem deve se limitar àquele grupo estudado, não podendo, portanto, ser exatamente aplicado, com idênticas pressuposições, em diferentes grupos.

## 2.2 Porque a Língua pode ser considerada um Sistema Dinâmico

Segundo Monteiro (2011), “Um *sistema* pode ser definido como um conjunto de objetos agrupados por alguma interação ou interdependência, de modo que existam relações de causa e efeito nos fenômenos que ocorrem com os elementos

desse conjunto”. A língua, basicamente, pode ser definida como um sistema dinâmico por ser um emaranhado de fenômenos sistemáticos que ocorrem, de certa forma, dentro de alguma regularidade interna. Não se pode, entretanto, considerar esse sistema como fechado, já que está intermitentemente sendo afetado e desequilibrado por fatores a ele externos. Não é possível, por exemplo, definir de forma definitiva quais seriam esses fatores, já que os mesmos podem mudar de acordo com o **tempo**, que é a variável responsável por suportar essas mudanças. Pode-se, por exemplo, tentar inferir quais seriam as variáveis que podem vir a interferir num específico grupo de sujeitos (americanos **ou** brasileiros, como neste estudo, por exemplo) ou, até mesmo, pode ser considerado apenas um sujeito como um sistema complexo único.

Aceitando-se a natureza dinâmica da língua, quais seriam as vantagens de analisar um determinado estágio dessa dinamicidade, como se tirássemos uma foto de um específico momento da língua? Inúmeras podem ser as respostas do porquê se analisar sistemas desse tipo, já que, por natureza, sistemas dinâmicos não necessariamente são físicos (“palpáveis”): sempre trabalhar-se-á com base em acontecimentos sistemáticos e possíveis probabilidades. Da mesma forma como é importante (menos perigoso e mais barato) analisar a possibilidade de uma nova droga desenvolvida em laboratórios causar determinada porcentagem de efeitos colaterais antes de tal droga estar no mercado e, então, ter de verificar quais foram as consequências em pacientes, podemos, na língua, fazer um caminho semelhante, tentando observar e analisar o comportamento de aprendizes de inglês quando comparados a nativos da língua inglesa, por exemplo.

Não se pode escolher analisar determinado estágio de uma língua sem observar o comportamento dos que a falam, pois a língua não é um fenômeno unicamente físico, muito embora tentemos mensurá-la de diversas formas. Embora tal tentativa esteja sendo bem sucedida em diversos campos (como a medição de aspiração em milissegundos, por exemplo), ainda assim, não se tem controle sobre o que é analisado. O propósito, então, é uma tentativa de descrição e compreensão do(s) fenômeno(s) previamente proposto(s), perguntando-se se determinado resultado ou comportamento pode ou não ser esperado de um determinado grupo analisado. Ainda segundo Monteiro (2011), o estudo de um sistema dinâmico

deveria ser dividido em duas etapas: 1) proposta de um possível modelo “adequado” e 2) análise de tal modelo. Normalmente, não há apenas um tipo de modelo que pode ser considerado como sendo correto, pois, dependendo da motivação, mais de um pode se adequar a determinada análise. A língua, finalmente, pode ser considerada um sistema dinâmico *per se* por ser um fenômeno onde apenas uma variável independente se faz presente, sendo esta variável o **tempo**. Assim sendo, os dados utilizados neste trabalho, se coletados novamente, com os mesmos sujeitos, pouco provavelmente apresentariam resultados iguais, embora pudessem apresentar escores semelhantes, levando-se em consideração que tais semelhanças poderiam se dever pelo fato de existirem sistemas atratores, que dificilmente mudariam em pouco espaço de tempo. Por isso, é importante salientar que tais repostas podem ser consideradas como “retratos” de uma situação X onde o participante respondeu/estava respondendo às questões que a eles foram propostas.

Estudiosos da aquisição de segunda língua, quando consideram-na um sistema dinâmico, muito fazem alusão do sistema linguístico em analogia com um pêndulo, por esse ser um objeto (mesmo que retrate algo abstrato) que ilustra de forma didática e mais prática a maneira como o tempo interfere de forma inevitável e infinita as variáveis que dele dependem. Dessa forma, no que segue, apresenta-se um pêndulo e uma tentativa de aproximação de seus movimentos com a sistematicidade inesperada, porém, organizada, da língua:

FIGURA 1: O Pêndulo como Exemplo de Sistema Dinâmico

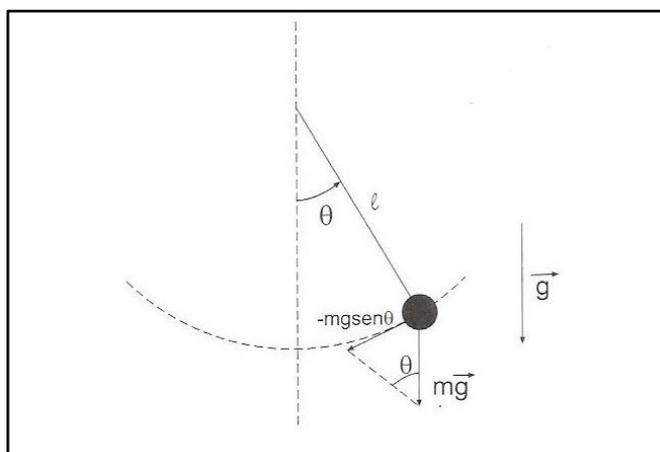


Figura adaptada de Monteiro (2011)

Um pêndulo simples, como o apresentado acima, pode ser uma exemplificação do desenvolvimento da língua e sua consequente e intermitente aquisição.

Sendo o tempo uma variável de evolução independente (assim considerado pela impossibilidade de se interferir na sua velocidade, sendo apenas possível mensurá-lo), chama-se ele de variável independente. Levando isso em consideração no campo de aquisição da segunda língua, pode-se assumir que o tempo interfere de forma inevitável no processo de aquisição e desenvolvimento do sistema linguístico, sem que possamos interferir nele a fins de se projetar o quanto queremos adquirir em um processo de aprendizagem ou quanto tempo vamos levar para adquirir um aspecto específico da língua segunda que se estuda.

O ângulo  $\theta$ , apresentado na figura 1, é aquele que sofre com a variação do tempo. Essa variável é, portanto, a variável dependente, já que depende do tempo para variar. Esse ângulo pode ser considerado como uma tentativa de se ilustrar o desenvolvimento da língua e da aquisição de seus vários componentes (de léxico, de sentidos, de sons, de línguas estrangeiras, por exemplo). O modelo de sistemas dinâmicos se presta à tentativa de determinar qual dependência é essa que muda apenas com o tempo e sem chances de se manter estática.

Segundo a figura mostrada anteriormente,  $l$  e  $g$  são quantidades que podem influenciar o comportamento do sistema. Como pode-se perceber, sendo a figura um pêndulo, pressupõe-se, nele, movimento. Logo, o valor (seja ele matemático ou não, como uma determinada fase de aquisição de uma língua, por exemplo) de tais quantidades variam de acordo com a posição do pêndulo. Ainda se valendo de uma teoria para a linguagem, usando a aquisição em segunda língua como exemplo, determinado desenvolvimento em um momento específico do aprendizado pode retroceder, ou evoluir, e ainda, retroceder novamente. Não se pode adivinhar onde o pêndulo estará depois de 5 anos de estudo em língua inglesa, pois ele pode estar em qualquer lugar, dependendo de quem aprende, da forma que aprende, entre inúmeros outros fatores.

O processo de aquisição de uma segunda língua pode ser considerado um processo dinâmico não-linear (Alves & Zimmer, 2012). Em um sistema dinâmico,

sabe-se que há uma força interna que tende a procurar o equilíbrio, dada a desorganização que lhe é característica. Esses pontos de equilíbrio, porém, são raros e instáveis e as forças atratoras que atuam na aquisição de uma segunda língua podem mudar de maneira não-linear (estudantes avançados de língua inglesa podem, por esquecimento ou sem qualquer explicação aparente, esquecer de conjugar com “-s” a terceira pessoa do singular no inglês, por exemplo). Há, inclusive, o processo de *backsliding* (que, em português, corresponderia a um termo como “cair para trás”), onde o aprendiz de uma língua chega a níveis cada vez mais avançados, já praticando a conversação, e acaba por apresentar problemas em se lembrar como se expressar diante de uma situação em que poderia lhe parecer simples quando ainda iniciava os estudos. Essa “queda”, geralmente, precede uma nova fase para o aprendiz, que logo recobra tais memórias e continua a avançar em conhecimentos da língua-alvo. Esse pode ser um exemplo de tentativa de busca pelo equilíbrio do sistema (e do movimento pendular que a língua percorreria, se enquadrada numa teoria dinâmica), que percorre um caminho aparentemente controlado e que pode, de maneira repentina, apresentar uma oscilação, que tende a ser encoberta por outro momento de organização e que pode prever outra oscilação, e assim por diante. O sistema tende a se estabilizar por algum tempo e assim também pode ser considerado o processo de aquisição de uma nova língua: mesmo buscando esses pontos de equilíbrio, a linguagem tende a movimentos que não podem ser previstos e podem perseguir caminhos não usuais, ou não-lineares (cf. PORT & VAN GELDER, 1995). Esse processo cognitivo não-linear pressupõe que diferentes fatores podem influenciar o processo de aprendizagem, sendo esses fatores linguísticos ou não-linguísticos, que, portanto, não podem ser tratados de forma isolada (DERWING e MUNRO, 2007).

O aprendiz, entretanto, necessita de regularidades para adquirir aspectos fundamentais da língua-alvo. É por isso que, de acordo com essa perspectiva, o *input* é considerado rico, já que é a partir dele que os aprendizes criam regularidades e diretrizes por onde devem percorrer para seguir em direção à fluência. É a partir do *input*, e das diferentes formas que ele pode ser apresentado aos aprendizes, que eles, expostos a esses excertos da língua estudada, conseguem criar e perceber regularidades e restrições, equivalentes a regras e

exceções, desenvolvendo melhor percepção em relação ao sistema da segunda língua (ZIMMER, SILVEIRA e ALVES, 2009).

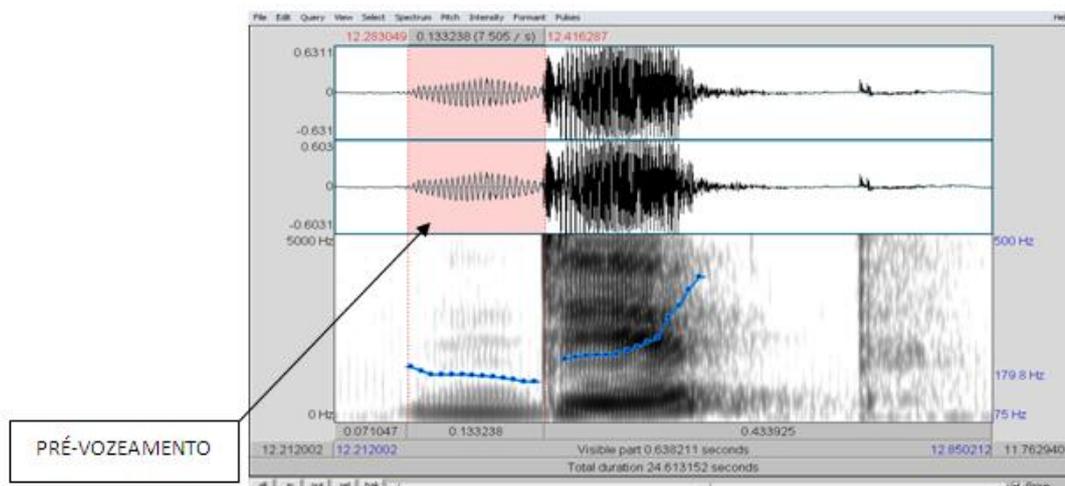
Tratando de estudos em percepção de sons em segunda língua, é preciso lembrar que, segundo Best (1995) afirma, características fonêmicas, de *status* fonológico ou funcional, são de domínio específico de cada língua, não podendo ser tratados, portanto, como universais. Aqui reside a barreira entre o fonético e o fonológico deste trabalho: trabalha-se não apenas com as características físicas de um som em um determinado estímulo, mas sim, como este estímulo é percebido e, ainda posteriormente, identificado pelos informantes, em termos de distinções funcionais. O limiar entre o fonético e o fonológico se faz evidente, sendo uma espécie de “batalha entre amigos” que ocorre no cérebro de cada falante a todo momento, de forma não-consciente. O VOT pode ser uma dessas pistas acústicas de domínio da língua. O exemplo mais claro disso se dá na categorização do VOT zero em ambas as línguas: as plosivas sonoras [b], [d] e [g] são enquadradas como pertencentes ao VOT zero, ou seja, sem apresentar pré-vozeamento e com aspiração com duração menor que 35 milissegundos. Para o português, entretanto, o VOT zero engloba as plosivas surdas [p], [t] e [k] – que seriam pertencentes ao padrão positivo no inglês – por não apresentarem altos índices de aspiração, ao contrário do que é previsto para a língua inglesa. Assim sendo, é possível que produções de fala de aprendizes brasileiros de plosivas surdas com pouca ou nenhuma aspiração pudessem ser percebidas como sendo suas respectivas representações sonoras por americanos. A seguir, pode-se ver melhor explicação sobre o que vem a ser o VOT (*Voice Onset Time*) e de que forma ele é classificado em inglês e em português.

## 2.2 Voice Onset Time (VOT)

O *Voice Onset Time* é uma pista acústica relevante a ser analisada no processo de aquisição fonético-fonológica de uma segunda língua, já que é uma pista que apresenta diferentes realizações entre sistemas linguísticos distintos. É também uma boa forma de identificar estágios de aprendizagem (levando-se em conta a sua aquisição ou não por parte dos aprendizes), ao evidenciar de que forma os aspectos fonéticos e fonológicos vão sendo adquiridos (ou adequados pelo falante, a partir da sua L1). Conforme Lisker & Abramson (1964), Ashby (2011), Gussenhoven & Jacobs (2011), Cohen (2004), Yavas (2008) e Reis & Nobre-Oliveira (2008), o VOT pode ser definido como o tempo entre a soltura de uma consoante plosiva e o início da vibração das cordas vocais do segmento vocálico que segue tal consoante. Existem três padrões de VOT encontrados nas línguas do mundo (LISKER & ABRAMSON, 1964; ASHBY, 2011; GUSSENHOVEN & JACOBS, 2011). Os padrões de VOT são classificados de acordo com a duração do tempo entre a soltura da consoante e a vibração das cordas vocais do segmento vocálico, por isso, há uma gradiente específica para classificar cada padrão, como exemplificado nas figuras a seguir.

- **VOT NEGATIVO: duração de -125ms a -75ms (/b/, /d/ e /g/ em português):** é caracterizado pelo pré-vozeamento que precede a produção de uma consoante sonora. Nesse caso, as cordas vocais cessam brevemente a vibração antes de o estreitamento da obstruente estar completamente articulado. Pode ser também chamado de “pré-aspiração” (cf. GUSSENHOVEN & JACOBS, 2011). É característico da realização das consoantes sonoras /b/, /d/ e /g/ na língua portuguesa;

Figura 2: VOT negativo com duração de 133ms, produzido por falante nativo do inglês americano, da palavra “gill”

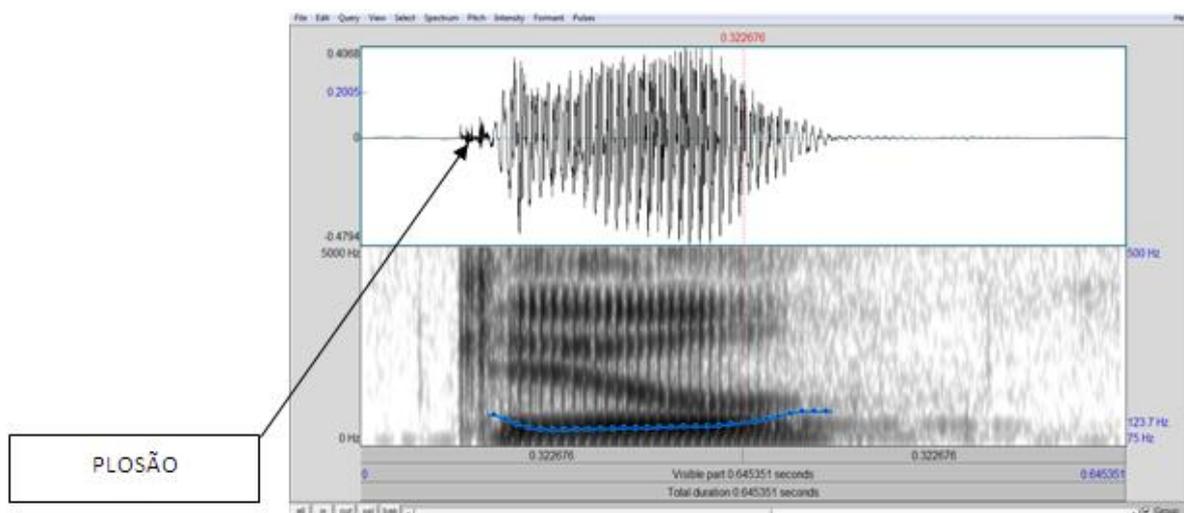


Fonte: banco de dados da autora.

- **VOT ZERO: duração de 0ms a +35ms (/p/, /t/ e /k/ no português e /b/, /d/ e /g/ no inglês):** também chamado de “não-aspirado”, ocorre quando as cordas vocais começam a vibrar quase simultaneamente com o fim da constrição oral (closura total para consoantes) – a vibração das cordas vocais se inicia após a produção da consoante surda, devido à consoante que a segue, porém, sem soltura de ar subsequente. É encontrado na produção de /p/, /t/ e /k/ em português e, também, na produção de /b/, /d/ e /g/ do inglês<sup>4</sup>;

<sup>4</sup> É importante enfatizar aqui que a produção das plosivas sonoras na língua inglesa pode encontrar padrão variável – inclusive como proposto na literatura da área (ALVES, SCHWARTZHAUPT & BARATZ, 2011; ALVES & MOTTA, 2013): /b/, /d/ e /g/ podem, variavelmente, dependendo do locutor, se encaixar no padrão Zero (caso não apresentem pré-vozeamento) ou no padrão Negativo (caso apresente pré-aspiração, que pode, também, ser chamada de pré-vozeamento).

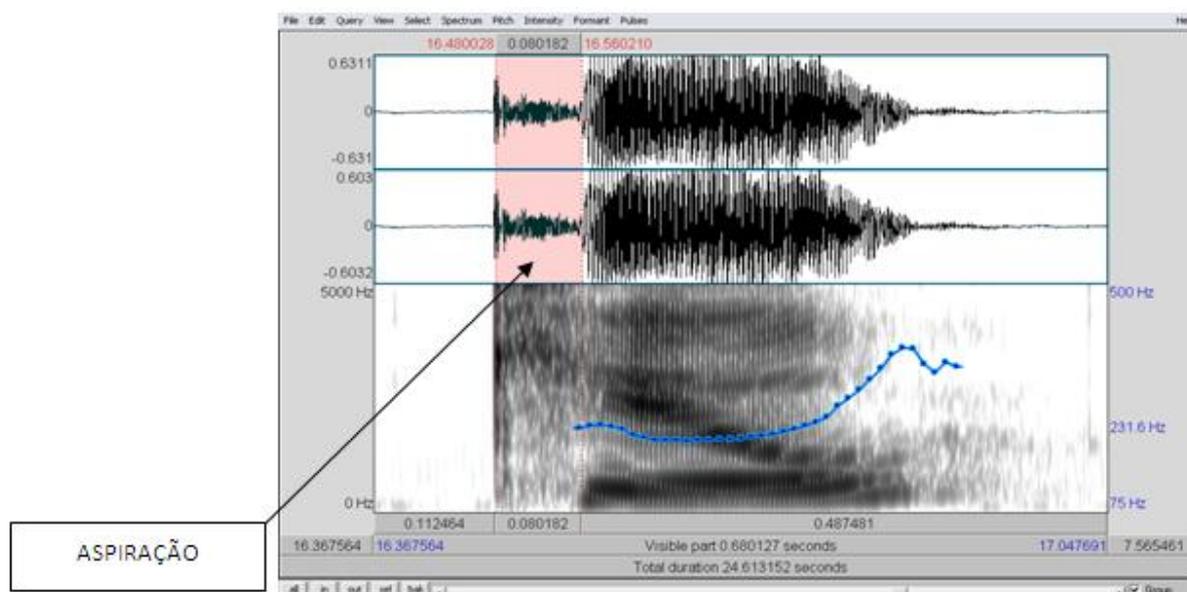
Figura 3: VOT com plosiva de padrão zero, produzida por falante nativo do inglês americano, da palavra “bit”



Fonte: banco de dados da autora.

- **VOT POSITIVO: duração de +35ms a +100ms ([p<sup>h</sup>], [t<sup>h</sup>] e [k<sup>h</sup>] no inglês):** também classificado como “longo”, é caracterizado por uma saída breve de ar após a produção de plosivas surdas. Ocorre na produção de [p<sup>h</sup>], [t<sup>h</sup>] e [k<sup>h</sup>] na língua inglesa. Em termos de percepção, essa saída de ar pode ser percebida como **aspiração**.

Figura 4: VOT positivo com duração de 80m, produzido por falante nativo do inglês americano, da palavra “tick”



Fonte: banco de dados da autora.

Estudos dirigidos desde o ano de 2009 pelo GEAFFLE<sup>5</sup> (ALVES, SCHWARTZHAUPT e BARATZ, 2011; SCHWARTZHAUPT, 2013; ALVES e MOTTA, 2013; SCHWARTZHAUPT, ALVES e FONTES, 2013) mostraram que o VOT se mostra como uma pista importante na distinção entre plosivas surdas e sonoras para os americanos. Estudos feitos a partir da manipulação do VOT positivo, que consistia na extração de toda a aspiração de consoantes surdas para que soassem, artificialmente, como sendo pertencente ao padrão zero, mostraram que participantes americanos tendem a identificar como sonoras as plosivas que tiveram seu VOT positivo cortado para o padrão zero (SCHWARTZHAUPT, ALVES e FONTES, 2013). Ainda que os resultados desses estudos tenham apontado a importância do VOT como pista acústica para as diferenças funcionais entre consoantes surdas e sonoras entre falantes do inglês, faz-se necessário, ainda, verificar se há valores limiares de VOT para a distinção surdo/sonoro entre esse grupo de participantes, tarefa essa a ser perseguida no presente estudo.

<sup>5</sup>Grupo de Estudos em Aquisição Fonético-Fonológica de Língua Estrangeira – (<http://www.geaffle.com>), com sede na Universidade Federal do Rio Grande do Sul.

Por outro lado, os estudos supracitados demonstraram que o mesmo *status* para o VOT não foi atribuído entre aprendizes brasileiros, que, independente do grau de proficiência, continuaram a identificar como surdas as plosivas com o VOT manipulado. Isso nos leva a acreditar que, ao contrário do que acontece com os americanos, o VOT pode não ser a pista acústica primordialmente seguida pelos brasileiros, de modo que possa existir alguma(s) outra(s) pista(s), que não o VOT, que seja tomada como mais importante pelos falantes nativos de português brasileiro. Como sugerido por Alves e Motta (2013), há a possibilidade de os brasileiros atribuírem maior importância ao *burst* (ou seja, à força da explosão em si) do que à aspiração, que caracteriza o VOT positivo. Estudos lidando com a manipulação da pista de burst ainda precisam ser feitos para que haja maiores evidências que pudessem comprovar tal hipótese. Por ora, destaca-se o fato de que uma pista acústica pode vir a exercer um papel distintivo fundamental em uma língua, mas não em outra.

Essa variabilidade está prevista também na literatura da área (LISKER & ABRAMSON, 1964), que admite um valor gradiente para cada um dos padrões. Foi possível perceber, a partir dos estudos desenvolvidos por nosso grupo de pesquisa, que muitos americanos não se “deixavam enganar” por alguns estímulos manipulados (ou seja, mesmo com o VOT manipulado, os nativos continuaram atribuindo características de consoantes surdas às consoantes surdas manipuladas, que deveriam soar como sonoras – SCHWARTZHAUPT, ALVES & FONTES, 2013). Como não houve um resultado mais específico em relação a uma determinada consoante ou estímulo, sentiu-se a necessidade, conforme já explicitado na Introdução, de se manipular essa aspiração em múltiplas “fatias”, para que se analisasse a possibilidade de uma “aspiração mínima” necessária para que os nativos de inglês considerassem uma ou outra consoante como sendo surda. De acordo com a literatura, o VOT positivo apresenta aspiração que pode variar de 35 a 100 milissegundos. Assim sendo, é possível que o mesmo americano identifique a produção de uma palavra com 100 milissegundos como sendo iniciada por uma consoante surda e, logo após a manipulação – e ainda pertencente ao padrão positivo – atribuisse valor de um VOT zero a um VOT positivo manipulado, que tivesse, por exemplo, 40 milissegundos. Dessa maneira, mesmo que de acordo com a literatura, é possível que um VOT pertencente ao padrão POSITIVO, devido à

duração da aspiração, não seja longo o suficiente para que um nativo de língua inglesa identifique uma consoante como sendo surda. Em outras palavras, com o presente estudo, tem-se a intenção de verificar os limiares, dentro do *continuum* fonético, responsáveis pela diferenciação na categorização fonológica entre segmentos surdos e sonoros, de modo a verificar a possibilidade de diferentes limiares tanto com falantes nativos quanto com aprendizes de inglês, em dois diferentes níveis de proficiência.

Infere-se, entretanto, que brasileiros provavelmente oscilem menos quando deparados à manipulação, ou seja, normalmente atribuem as mesmas características como um grande grupo em si a um mesmo estímulo, manipulado ou não (maioria identificou “gill” manipulado como sendo “g” ou “k”, enquanto, entre os americanos, os resultados podiam apresentar maior índice de erro e respostas mais distribuídas entre todas as opções possíveis: p, b, t, d, k, g), como nos estudos de E de Alves, Schwartzhaupt e Motta (2012) e Alves e Motta (2013). Como até mesmo o nível de proficiência não foi um fator determinante encontrado nos primeiros estudos já mencionados (ainda que isso fosse o esperado, porém, não verificado), espera-se que, à medida em que o aprendiz avança em direção das pistas da segunda língua, haja tendências de que ele caminhe também em direção à percepção de pistas dessa segunda língua, mas não há, entretanto, estudos que corroborem ou que desqualifiquem tal possibilidade. Por esses dois cenários estudados, foram propostos os objetivos e as hipóteses para o presente estudo, já apresentadas no primeiro capítulo.

## 3 MÉTODO

### 3.1 Participantes

Para este estudo, foram considerados dois grupos distintos de participantes: um de brasileiros aprendizes de inglês como língua adicional, subdivididos entre *intermediários* e *avançados* (totalizando um número de 16 participantes brasileiros, sendo 7 deles de nível de proficiência avançado e 9 de nível intermediário) e um de falantes nativos do inglês norte-americano (totalizando 12 informantes estrangeiros).

### 3.2 Participantes Brasileiros

Os brasileiros participantes deste estudo eram residentes de Porto Alegre ou região metropolitana. Informantes oriundos de outras cidades do Rio Grande do Sul não foram considerados para que se evitassem influências dialetais, mesmo que pequenas e/ou desconhecidas. Todos haviam estudado inglês por, no mínimo, 5 anos, contando-se cursos livres de língua inglesa e graduação. A maior parte deles foi composta de estudantes de Letras, com ênfase em língua inglesa, na Universidade Federal do Rio Grande do Sul. O total de participantes brasileiros foi de 16 e a eles foi pedido que se assinasse um Termo de Consentimento Livre e Esclarecido, que continha informações sobre a pesquisa desenvolvida, o que assegurou que todos estavam cientes e de acordo com as tarefas das quais iriam participar.

Como havia aprendizes de níveis de proficiência diferenciados, foi aplicado a todos os participantes um teste de nivelamento em língua inglesa, o *Oxford Online Placement Test* (OOPT)<sup>6</sup>. Esse teste é uma ferramenta online para se avaliar o nível de proficiência em língua inglesa de aprendizes em todas as partes do mundo. O teste em questão foi escolhido para este estudo pela dificuldade de se usar um teste de nivelamento em papel, e, já que a ferramenta utilizada para a aplicação do instrumento era também online, decidiu-se usar o OOPT. Outro motivo para essa escolha foi a superioridade em relação ao *Oxford Placement Test* tradicional (ALLAN, 2004), já utilizado em muitos estudos da área (ALVES, SCHWARTZHAUPT, BARATZ, 2011): apesar de ter custos mais baixos, o teste

---

<sup>6</sup> Para maiores informações sobre o Oxford Online Placement Test, consultar Purpura (2007) e Pollitt (2007).

tradicional é muito focado na parte de listening apenas, havendo inúmeras questões que se resumem à diferenciação de pares mínimos, não havendo espaço para compreensão de textos, por exemplo. A seguir, é apresentado um breve quadro com informações sobre os informantes brasileiros:

QUADRO 1 – INFORMAÇÕES DOS PARTICIPANTES BRASILEIROS

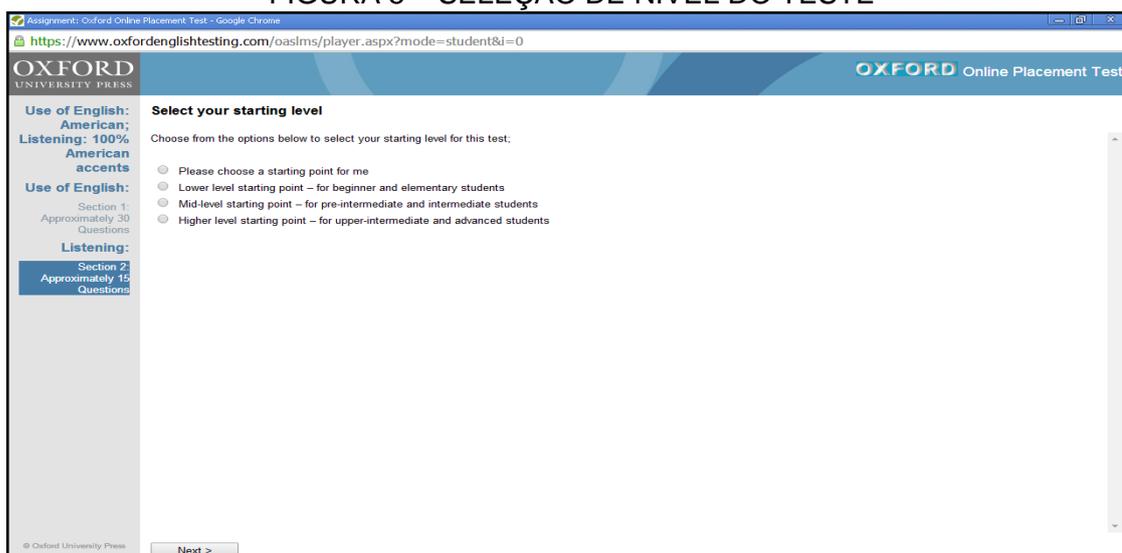
Participante	Idade	Cidade de Nascimento	Línguas Adquiridas antes dos 6 anos de Idade	Idade em que Começou a Estudar Outra Língua	Tempo de Estudo de Outra Língua	Resultado Nivelamento
1	43	Canoas	Português	20	23	<b>C1</b>
2	27	Porto Alegre	Português	20	7	<b>C1</b>
3	22	Porto Alegre	Português	16	6	<b>C1</b>
4	30	Canoas	Português	12	18	<b>C1</b>
5	30	Guaíba	Português e Espanhol	17	13	B1
6	20	Porto Alegre	Português e Espanhol	12	8	B2
7	24	Eldorado do Sul	Português	17	7	B1
8	17	Porto Alegre	Português	10	7	<b>C1</b>
9	23	Viamão	Português	15	8	<b>C2</b>
10	24	Viamão	Português	16	8	B2
11	19	Porto Alegre	Português	10	9	B2
12	22	Porto Alegre	Português	15	7	B1
13	24	Porto Alegre	Português	15	9	<b>C1</b>
14	22	Porto Alegre	Português	12	10	B2
15	36	Porto Alegre	Português	30	6	B2
16	23	Porto Alegre	Português	18	5	B2

Conforme apresentado no quadro acima, o total de participantes brasileiros foi de 16, sendo 9 deles enquadrados no nível de proficiência *intermediário* e 7 deles no nível *avançado*. O parâmetro utilizado para tal categorização será melhor explorado a seguir.

### 3.3 Oxford Online Placement Test (OOPT)

O *Oxford Online Placement Test* (OOPT) é acessado pelos participantes através de uma senha, a partir da qual os alunos têm acesso à plataforma de testes. Cada teste leva em torno de 1 hora, tendo tempo máximo de 1 hora e meia para ser finalizado. O tempo máximo de resposta ao teste pode ser ajustado pelo pesquisador, variando entre 60 e 90 minutos. Antes de iniciar o teste, o informante é solicitado a informar em que nível ele se encontraria, a fim de que o teste seja montado eletronicamente. Embora houvesse a possibilidade de escolha de o próprio site definir o nível dos participantes (levando em consideração as primeiras questões), os informantes deste estudo foram orientados a marcar a terceira opção disponível para níveis (*mid-level starting point – for pre-intermediate and intermediate students* – FIGURA 5), já que se imaginou que seria o mais equilibrado dentre todos. Outro ponto levado em consideração para a escolha dessa opção e não da primeira (*“please, choose a starting point for me”*) onde a calibragem inicial seria feita pelo site da *Oxford*, foi o fato de esta primeira opção acarretar em muito tempo de questões-teste, o que poderia comprometer o tempo máximo de 90 minutos estipulado.

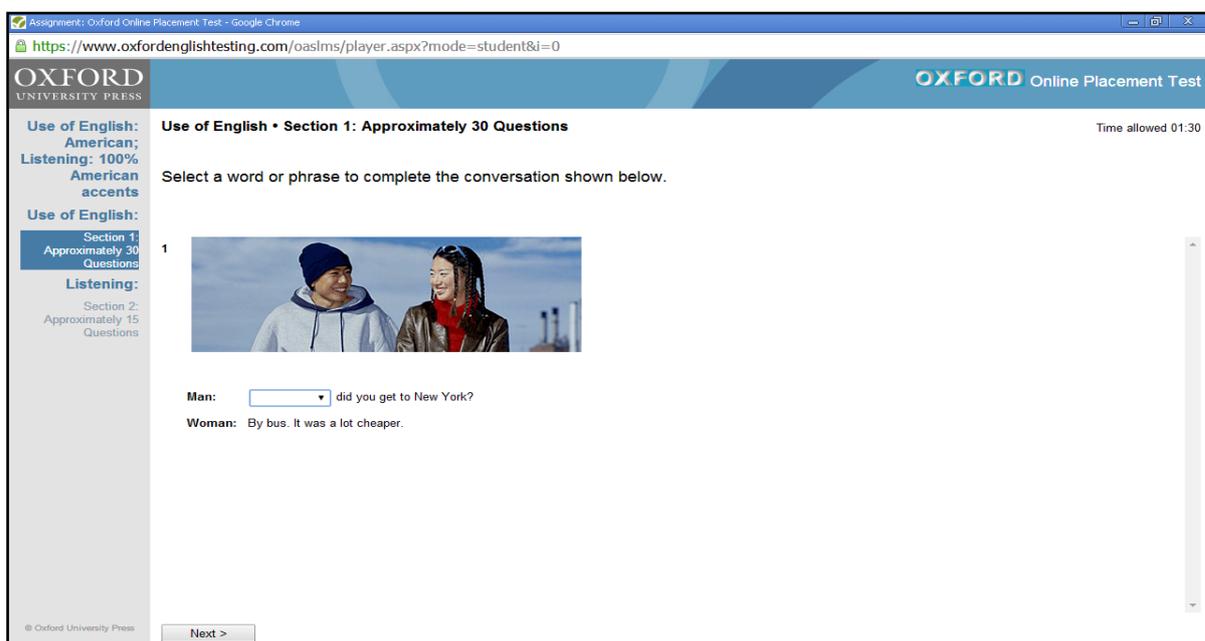
FIGURA 5 – SELEÇÃO DE NÍVEL DO TESTE



Fonte: adaptado do site Oxford EnglishTesting.

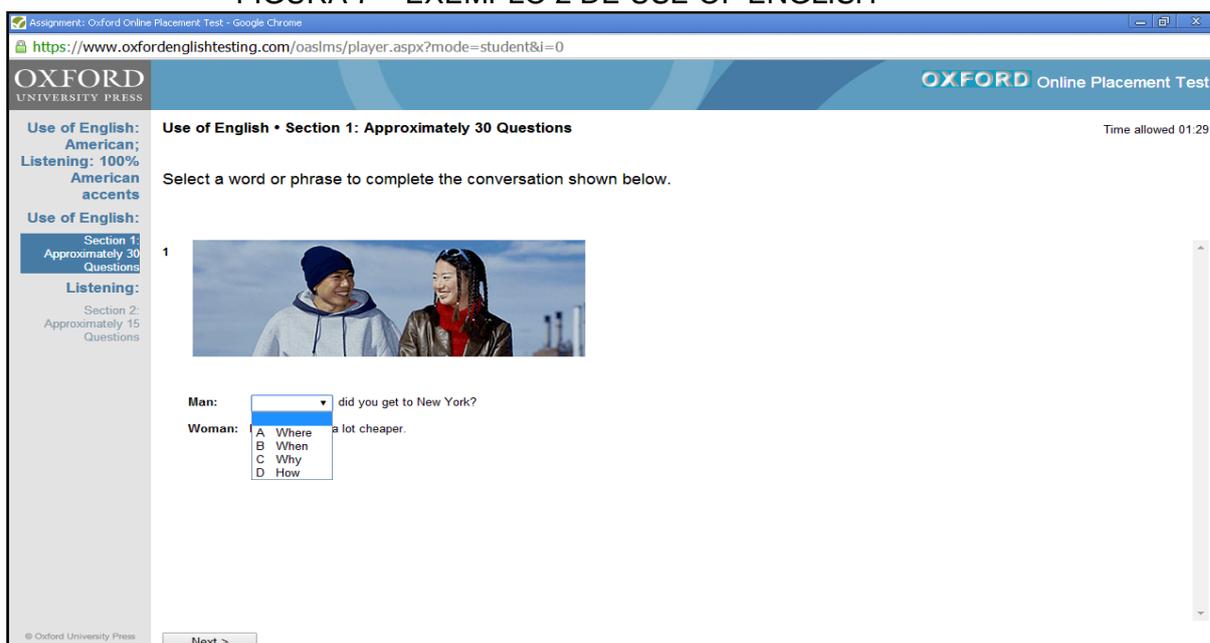
O teste é dividido em partes de *use of English* e *listening*, e ainda subdividido com exercícios diferentes dentro de cada uma dessas partes. A parte de *use of English* é a primeira e tem, em média, 30 questões. O teste, de certa forma, “calibra” o número de questões por participante, logo, o número de questões nunca é igual para cada participante. Essa primeira seção é subdividida em três partes: a primeira tem por objetivo verificar a acurácia do aprendiz em relação a palavras funcionais e pequenas, como preposições e advérbios, por exemplo (FIGURA 6).

FIGURA 6 – EXEMPLO 1 DE USE OF ENGLISH



Fonte: adaptado do site Oxford EnglishTesting.

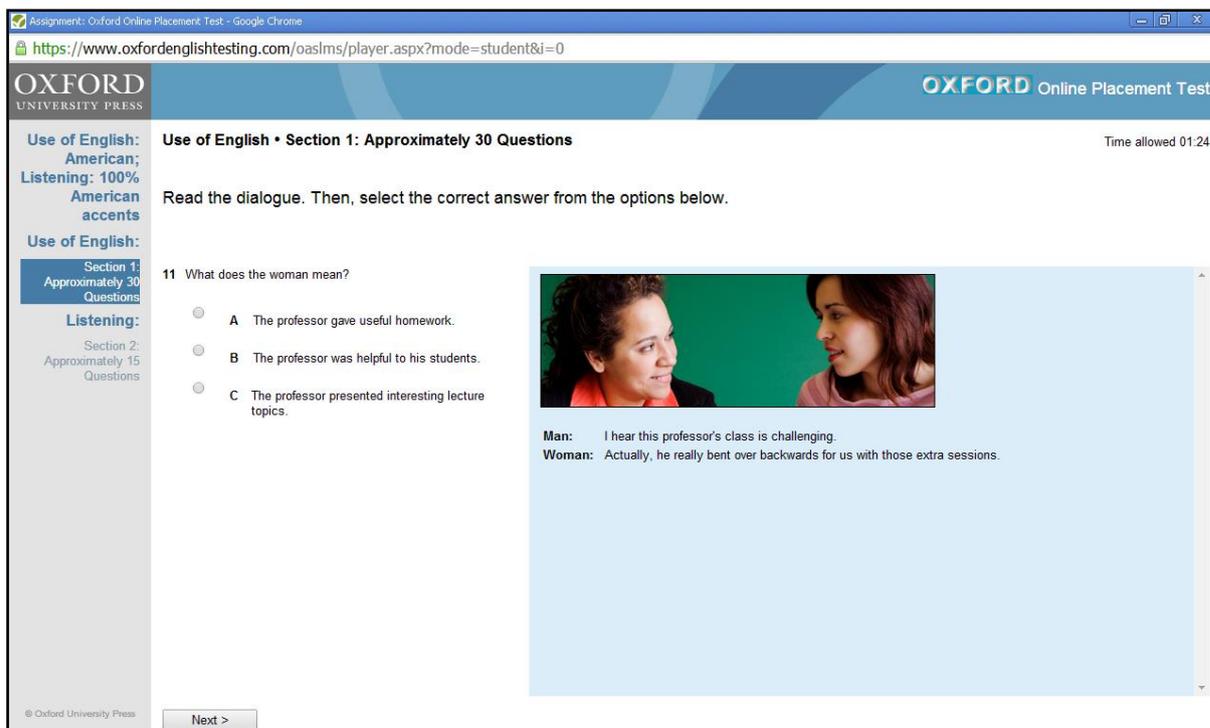
FIGURA 7 – EXEMPLO 2 DE USE OF ENGLISH



Fonte: adaptado do site Oxford EnglishTesting.

A segunda parte do *use of English* consiste em apresentar um breve diálogo ao participante e fazê-lo escolher entre opções de interpretação sobre tal diálogo (FIGURA 8).

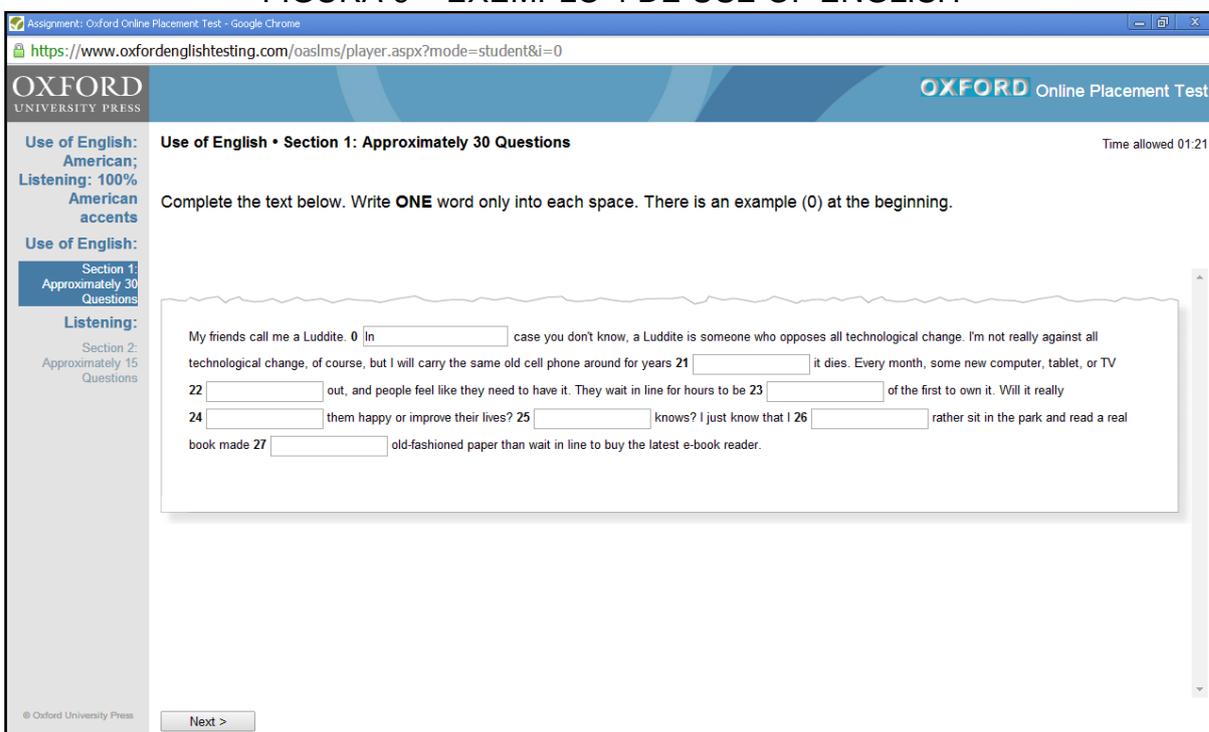
FIGURA 8 – EXEMPLO 3 DE USE OF ENGLISH



Fonte: adaptado do site Oxford EnglishTesting.

Diferentemente da primeira parte, essa segunda está mais ligada à questão semântica e pragmática da língua. Finalmente, a terceira parte do *Use of English* apresenta um grupo de pequenos textos com blocos de sentido fechados em si, onde se pede que o informante complete os *blanks* de acordo com as palavras que achar melhor (FIGURA 9). Ao contrário da primeira parte, aqui não existem respostas pré-definidas a serem escolhidas pelos aprendizes.

FIGURA 9 – EXEMPLO 4 DE USE OF ENGLISH



Fonte: adaptado do site Oxford EnglishTesting.

A segunda seção do OOPT, intitulada como *listening*, relaciona-se com a compreensão mais ampla do áudio que o teste disponibiliza; diálogos de 15 segundos são tocados pelos participantes, que, após lerem primeiramente as questões e posteriormente ouvir o diálogo, devem escolher qual das opções melhor se enquadra no assunto abordado no diálogo (FIGURA 10).

FIGURA 10 – EXEMPLO 1 DE LISTENING

The screenshot shows a web browser window with the URL <https://www.oxfordenglishtesting.com/oaslms/player.aspx?mode=student&i=0>. The page header includes the Oxford University Press logo and 'OXFORD Online Placement Test'. On the left, a sidebar indicates 'Use of English: American; Listening: 100% American accents' and 'Section 2: Approximately 15 Questions'. The main content area is titled 'Listening • Section 2: Approximately 15 Questions' with a 'Time allowed 01:20'. Below the title, instructions state: 'Read the sentences below. Then, listen to the short conversation. Select the correct answer from the options below. You will have time to play the recording twice.' A play button is visible. Question 28 asks: 'Two friends are talking about cell phones. The man tells the woman that...'. The options are: A he can't find his cell phone, B he uses his new cell phone a lot, C he thinks he should buy a cell phone, and D he's not comfortable with his new cell phone. A 'Next >' button is at the bottom.

Fonte: adaptado do site Oxford EnglishTesting.

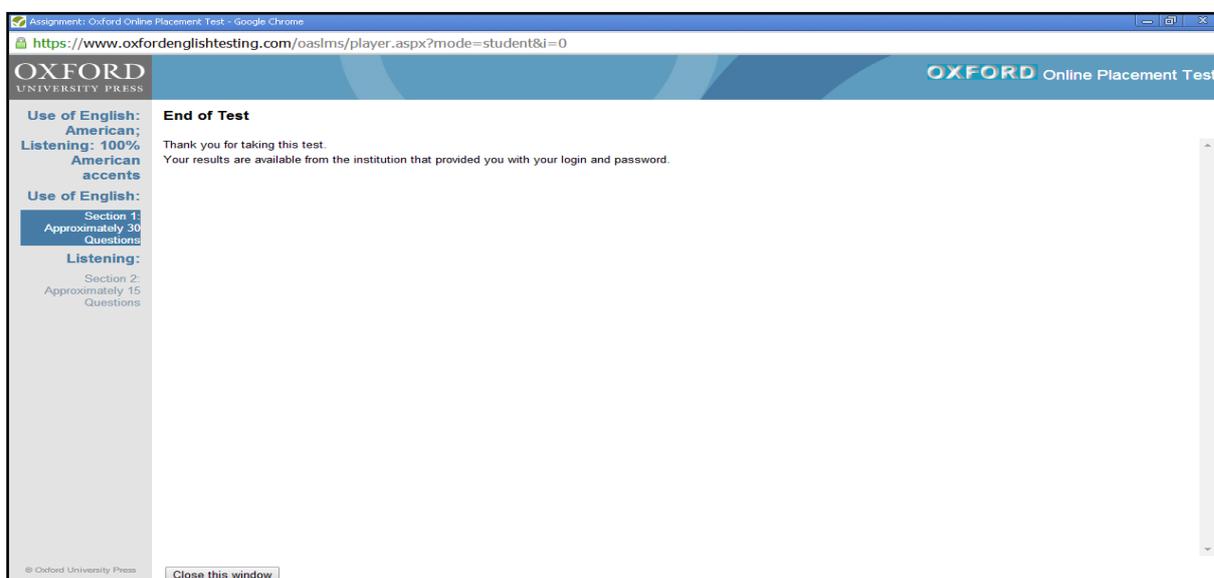
FIGURA 11 – EXEMPLO 2 DE LISTENING

The screenshot shows a web browser window with the URL <https://www.oxfordenglishtesting.com/oaslms/player.aspx?mode=student&i=0>. The page header includes the Oxford University Press logo and 'OXFORD Online Placement Test'. On the left, a sidebar indicates 'Use of English: American; Listening: 100% American accents' and 'Section 2: Approximately 15 Questions'. The main content area is titled 'Listening • Section 2: Approximately 15 Questions' with a 'Time allowed 01:19'. Below the title, instructions state: 'Read the sentences below. Then, listen to the recording. Select the correct answer from the options for each question below. You will have time to play the recording twice.' A play button is visible. Question 34 asks: 'Two friends are talking about how they get their news. Why doesn't the man watch TV news?'. The options are: A Most of the information is local, B The stories are poorly researched, C There isn't enough international news, and D The news stories aren't recent enough. Question 35 asks: 'What type of news is the man interested in?'. The options are: A local news, B financial news, C international news, and D entertainment news. A 'Next >' button is at the bottom.

Fonte: adaptado do site Oxford EnglishTesting.

O OOPT possibilita ao administrador dos testes que seja feita a escolha da variedade de inglês a ser testada (100% inglês americano; 50% americano e 50% britânico ou 100% britânico). Neste estudo, a variedade americana foi a escolhida, haja vista que os estímulos utilizados nos instrumentos foram gravados por falantes nativos dessa variedade. Ao final do teste, o participante é nivelado de acordo com o Quadro ou Marco Comum Europeu (*Common European Framework* - FIGURA 13)<sup>7</sup>. Existem três notas: uma para “Use of English”, uma segunda para “Listening” e, finalmente, uma nota considerada final, que, supostamente, enquadraria o participante em um dos seis níveis possíveis do Marco Comum Europeu (A1, A2, B1, B2, C1, C2). Para fins práticos, e de rigor metodológico, os participantes brasileiros deste estudo foram divididos em dois grupos: Avançado (que atingiram, no mínimo, C1 em cada uma das bandas do teste); e Iniciante (que atingiram A1 e A2 no total dos testes).

FIGURA 12 – TELA DE FINALIZAÇÃO DO TESTE



Fonte: adaptado de Oxford EnglishTesting.

<sup>7</sup>O Marco ou Quadro Comum Europeu (de sigla QECRL para “Quadro Europeu Comum de Referência para as Línguas) foi uma padronização proposta pelo Conselho da Europa a fins de internacionalização de níveis de aprendizagem para línguas adicionais. Tal medida se fez necessária devido à cultura plurilíngue fomentada e incentivada nos países europeus para o meio acadêmico. Com a finalidade de se seguir um padrão internacional, diversos testes de nivelamento optam por utilizar o QECRL como parâmetro, já que cada um dos seis níveis propostos enquadram a habilidade global e de uso real da língua, em contraponto a testes que tenham bases de análise puramente quantitativas. O TOEFL (*Test of English as a Foreign Language*), muito conhecido por ser um dos mais utilizados como pré-requisito para ingresso em diversas universidades do mundo, utiliza-se do QECRL como padrão de categorização dos aprendizes.

FIGURA 13 – TABELA DE RESULTADO OOPT

Oxford English Testing LMS							
Results [REDACTED]							
Component	Score	Time taken	CEF	Note	Date taken	Status	
Oxford Online Placement Test	79	00:47	B2	Hover over to view additional information...	2014-04-07	Normal	
Use of English:	66	00:20	B2	Hover over to view additional information...			
Section 1: Approximately 30 Questions							
Listening:	92	00:27	C1	Hover over to view additional information...			
Section 2: Approximately 15 Questions							

Fonte: adaptado do site Oxford EnglishTesting.

### 3.4 Participantes Americanos

Os participantes americanos totalizaram 12, oriundos de diferentes estados do Estados Unidos (ver Quadro 2, sobre informações dos participantes). Não foi possível controlar os informantes a partir do número de anos vividos apenas nos Estados Unidos ou o não contato com outras línguas, pois todos os participantes, com exceção de 1, tiveram contato com outras línguas e fizeram viagens para países estrangeiros, em proporções diferenciadas. O tempo limite considerado como corte para manter um informante nessa pesquisa foi de até 2 meses de moradia ou viagem em um único lugar. Definiu-se essa quantidade de tempo porque era o que mais se aproximava da média total de viagem, se contando todos os participantes. A seguir, apresenta-se um quadro ilustrativo com as informações dos participantes americanos.

QUADRO 2 – INFORMAÇÕES DOS PARTICIPANTES AMERICANOS

Participante	Idade	Cidade e Estado de Nascimento	Línguas Adquiridas antes dos 6 anos de Idade	Idade em que Começou a Estudar Outra Língua	Tempo de Estudo de Outra Língua
1	26	Fort Myers, Flórida	Inglês e um pouco de árabe	11	15
2	19	Fairbury, Illinois	Inglês	14	5
3	23	Oakdale, Califórnia	Inglês e um pouco de espanhol	13	6
4	22	Mill Valley, Califórnia	Inglês e um pouco de espanhol	Não informou	13
5	26	Miramar, Flórida	Inglês	Espanhol 15; Francês 20	Espanhol, 10; Francês, 1
6	30	Salt Lake City, Utah	Inglês	12	18
7	50	Cidade de Nova Iorque, Nova Iorque	Inglês	25	17
8	28	Los Angeles, Califórnia	Inglês	14	14
9	57	Urbana, Illinois	Inglês	Não informou	4
10	23	Merchantville, Nova Jersey	Inglês	14, Espanhol	9, Espanhol
11	38	Imperial, Nebraska	Inglês	16	2
12	40	Pittsburgh, Pensilvânia	Inglês	12	3

Os participantes americanos responderam ao Termo de Consentimento Informado (*consent form*) que estava contido junto aos testes, de forma online (ver figura 14). Mais informações sobre a plataforma online de testes serão fornecidas na seção específica *A Plataforma Online Utilizada: Surveygizmo*.

FIGURA 14 – CONSENT FORM DISPONIBILIZADO AOS AMERICANOS ONLINE

1. PARTICIPANT CONSENT FORM - You will fill in a form in order to provide some information about you. Your participation is very important and your data are confidential.

This study aims to contributing with studies related to the acquisition of English as a foreign language and researching how native speakers and learners behave according to some features of the English language.

Do you accept participating in this study? \*

Yes, I accept.

No, I do not accept.

Fonte: adaptado do site Oxford EnglishTesting.

### 3.5 Palavras-alvo

Como palavras-alvo, foram utilizadas palavras monossilábicas da língua inglesa, com estrutura CVC (com exceção de *pee*), onde a vogal sempre seria alta (/i/ ou /ɪ/), contexto esse que contribui para um aumento, e conseqüente percepção, do valor de aspiração do VOT das consoantes (cf. YAVAS, 2008). As consoantes iniciais alternavam entre as plosivas surdas /p/, /t/ e /k/, já que o objetivo principal deste estudo está focado na percepção de segmentos que possibilitem a manipulação do VOT positivo que ocorre em posição inicial de palavra.

O total de palavras-alvo (*types*) escolhido para o estudo foi de 6 (seis): *pee*, *pit*, *tick*, *tip*, *kit*, *kill*. Esses estímulos foram os mesmos utilizados nos estudos de Alves, Schwartzhaupt, Baratz (2011) e Alves e Motta (2013), já mencionados anteriormente. Nos estudos anteriores, a escolha dessas palavras levou em conta diversos fatores, tais como: ocorrência semelhante das palavras frequentes e infrequentes, o número de sílabas e, inclusive, as possíveis correspondências grafo-fônico-fonológicas a serem invocadas em cada palavra. Por esse motivo, encontrase, fora da estrutura CVC, a palavra *pee*, em detrimento de *pete* (que poderia vir a propiciar uma epêntese final), por exemplo. A delimitação dessas 6 palavras foi feita para que todos os pontos de articulação fossem contemplados: dessa forma, há duas palavras por ponto de articulação. Cada uma delas foi manipulada por quatro vezes, o que resultou em 5 diferentes amostras da cada estímulo: o original (considerado como VOT positivo 100%) e os demais, com aspiração cortada a cada 25% do valor total dos milissegundos, chegando-se, assim, até o VOT zero de caráter artificial. Cabe salientar que este último corte teve por objetivo o **padrão** VOT

zero (ou seja, dentro de tal padrão, poderiam ser encontrados VOTs com valores de duração de 12 milissegundos e também de 25 milissegundos. Mesmo que seja praticamente um o dobro do valor do outro, ainda assim, ambos se enquadravam no padrão de VOT zero)<sup>8</sup>, já que uma tentativa de corte total de aspiração poderia levar à alteração de pistas acústicas não desejadas (como transição formântica das vogais e closure, por exemplo), e que poderiam influenciar na percepção dos participantes.

### 3.6 Gravação dos Estímulos

Os estímulos foram todos gravados por falantes nativos da língua inglesa, da variedade estadunidense. Originalmente, para fins de utilização em estudos anteriores (ALVES, SCHWARTZHAUPT e MOTTA – 2012; ALVES, SCHWARTZHAUPT e BARATZ, 2011), seis informantes nativos (3 homens e 3 mulheres) foram os responsáveis pela gravação dos estímulos, que seriam, posteriormente, utilizados em estudos sobre VOT para o grupo de pesquisa de fonética e fonologia da Universidade Federal do Rio Grande do Sul (SCHWARTZHAUPT, 2012; ALVES & MOTTA 2013).

Para este estudo em específico, entretanto, o número de falantes nativos foi diminuído para 2 (ambos mulheres), já que todos os áudios foram pré-analisados acusticamente no software *Praat* e esses dois informantes nativos eram os que mais se aproximavam aos dados de produção encontrados na literatura da área. (veja-se Referencial Teórico, seção 2). Chegou-se a cogitar a possibilidade de utilizar estímulos de 4 informantes, porém um deles apresentava produções exageradas para /k/ (com valores de VOT mais altos do que 130ms, indo além do previsto pela literatura). O segundo locutor excluído apresentava valores que poderiam ser considerados normais, em termos de duração, porém era possível analisar, acústica e oitivamente, hiperarticulação para a produção de /k/ e /p/, que resultava em valores exagerados de duração de VOT. Finalmente, para a montagem do instrumento piloto, foram utilizados os estímulos desses dois falantes. A primeira locutora era originária de Indianola, estado de Iowa, ao sul dos Estados Unidos e tinha 29 anos. A segunda informante era oriunda de Western Washington,

---

<sup>8</sup> É de suma importância ressaltar, novamente, que PADRÃO zero e DURAÇÃO zero pertencem a conceitos distintos. Para maiores detalhes, veja-se seção 2.2 do Referencial Teórico.

pertencente ao estado de Washington, a Oeste das Montanhas Cascade, próximo à capital do estado, Olympia, e tinha entre 27 e 29 anos. Ambas haviam vindo dos Estados Unidos para estudar no Brasil e estavam residindo na cidade de Porto Alegre por aproximadamente 3 meses.

### 3.7 Da Manipulação dos Estímulos

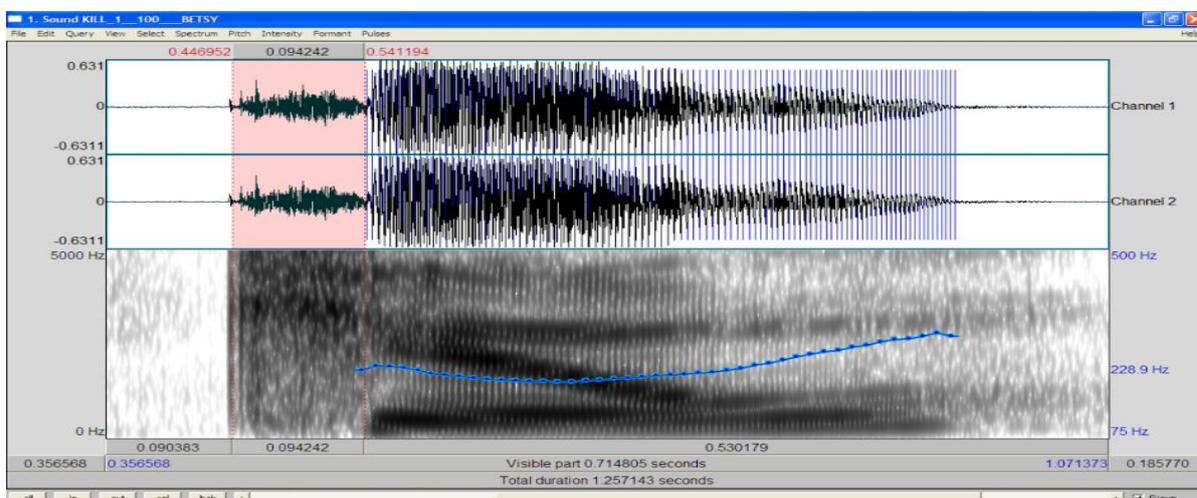
Para que se evitassem possíveis equívocos, ao se manipularem os áudios manualmente, várias tentativas (não bem-sucedidas) de manipulação por *scripts* do software *Praat* (BOERSMA e WEENINK, 2001) foram postas em prática, para que se minimizassem possíveis manipulações de pistas outras, como já mencionado. O mais próximo que se chegou de um *script* capaz de fazer múltipla manipulação de corte do VOT implicava marcar o intervalo de VOT a ser manipulado como um intervalo em um *tier*. O necessário a ser feito seria uma marcação de todo o intervalo de VOT e, então, que se rodasse um *script* que fosse capaz de cortar a porcentagem desejada em cada uma das manipulações.

Após as tentativas de manipulação múltipla e automática, partiu-se, então, para a manipulação manual, efetivamente realizada neste trabalho, que foi feito da seguinte forma:

- o arquivo genuíno foi aberto no software *Praat* e salvo como o original, para que houvesse o áudio com o VOT 100%, ou seja, padrão positivo, sem manipulação;
- as manipulações subsequentes sempre se deram a partir do ponto de corte do meio da duração total do VOT em direção às bordas, até que se chegasse ao padrão de VOT zero. Dessa forma, as manipulações visavam ao corte apenas da pista VOT, para que evitasse o corte de outras pistas, tais como transição formântica das vogais, por exemplo;
- as manipulações levaram em conta a porcentagem sobre o valor bruto inicial; assim sendo, o corte de 75% de *pit* do locutor 1, por exemplo, era relativo ao valor inicial de 100%, e não ao de 50%.

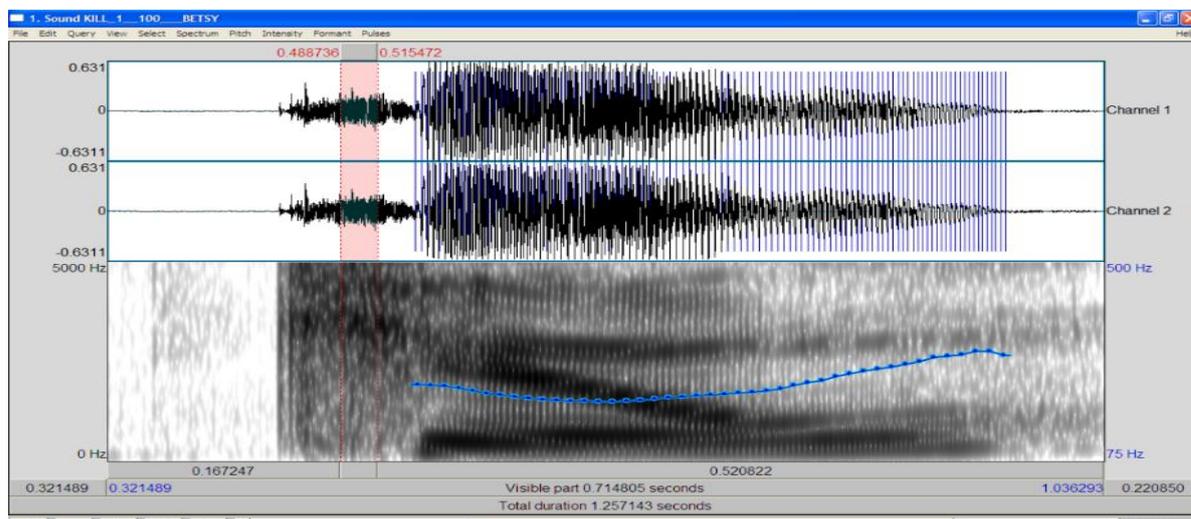
A seguir, são apresentadas algumas imagens que ilustram como as manipulações foram feitas:

FIGURA 15 – Exemplo de VOT 100% selecionado para manipulação – Palavra “Kill”, com duração de 94,2 ms. (Locutor 1)



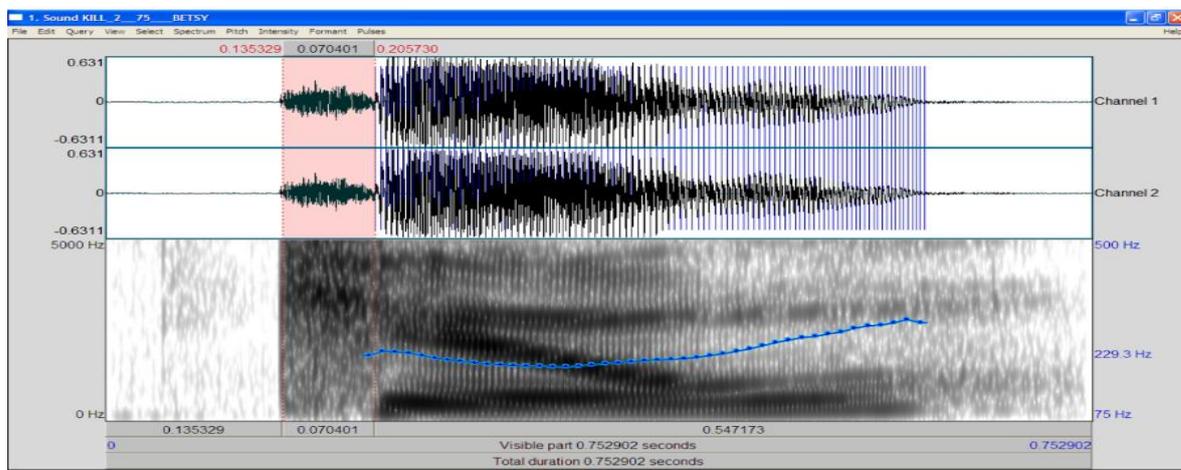
Fonte: banco de dados da autora.

FIGURA 16 – VOT de “Kill”, com aspiração de 94,2, com 25% da duração selecionada, para ser posteriormente retirada (Locutor 1)



Fonte: banco de dados da autora.

FIGURA 17 – VOT de “Kill”, agora com duração média de 70ms, com 25% da sua aspiração já retirada (Locutora 1)



Fonte: banco de dados da autora.

Abaixo, apresenta-se uma tabela ilustrativa com todos os valores finais atingidos, que constituíram o total de estímulos utilizados nos testes, após a manipulação de todos os estímulos originais (100%):

TABELA 1 – Valor de VOT por *token* em milissegundos e locutor

<b>TOKENS</b>	<b>LOCUTORA 1 (em ms)</b>	<b>LOCUTORA 2 (em ms)</b>
PIT (100%)	73	124
PIT (75%)	54	92,7
PIT (50%)	36,78	60
PIT (25%)	18,44	34
PIT (0 Art)	7,23	7,8
PEE (100%)	107	124
PEE (75%)	78	91,18
PEE (50%)	52,38	63
PEE (25%)	25,13	30
PEE (0 Art)	9,19	10
TICK (100%)	85,77	100
TICK (75%)	60,18	75,18
TICK (50%)	42,3	49,33
TICK (25%)	21,25	26,3
TICK (0 Art)	12,12	12
TIP (100%)	94	78,33

TIP (75%)	70,75	59,2
TIP (50%)	47,88	38
TIP (25%)	23,6	19,4
TIP (0%)	8	6,99
KILL (100%)	94,66	91,12
KILL (75%)	45	68,23
KILL (50%)	22,2	45,6
KILL (25%)	12	21,2
KILL (0 Art)	8	10,11
KIT (100%)	77	68,13
KIT (75%)	58	49,2
KIT (50%)	40	33
KIT (25%)	19,8	18,98
KIT (0 Art)	6	11,12

É importante salientar que o estímulo “original”, contendo 100% da aspiração, foi exatamente o mesmo utilizado para as manipulações subsequentes. Parece ser evidente que cortes de valores de VOT muito próximos (como de pouco mais de 20ms de aspiração) ficassem, já nas últimas extrações, demasiadamente parecidos, mas tal escolha foi feita devida à análise prévia realizada nos áudios. Se fossem escolhidos estímulos de locutores diferentes para a mesma palavra e aspiração cortadas (locutora 1 para “pit” 75% e locutora 2 para “pit” 50%), tal fato poderia gerar incerteza por parte dos participantes do estudo, já que diferenças suprasegmentais foram analisadas nas curvas de *pitch* quando os estímulos foram comparados. O mesmo problema ocorria quando estímulos de rodadas diferentes, oriundos de uma mesma locutora, foram analisados. Mesmo que aspectos suprasegmentais não fossem o foco deste estudo, a intenção principal foi deixar os estímulos o mais equilibrados entre si. Para que se evitasse a possibilidade de que áudios manipulados em sequência serem tocados na hora do teste, os estímulos passaram por randomização, a ser explicada no que segue.

## 4. INSTRUMENTO DE COLETA DE DADOS

### 4.1 Teste de Identificação

Conforme a descrição do processo de elaboração de estímulos apresentada na seção anterior, o teste de identificação consistia de 60 *tokens* (veja-se Tabela 1), onde a tarefa principal dos participantes era identificar qual a consoante inicial da palavra que ouviam. Se o participante ouvisse o estímulo “pit” com 100% de aspiração, por exemplo, esperava-se que os informantes, de ambas as línguas, identificassem um /p/ como sendo a consoante inicial da palavra ouvida (os americanos, por seguirem a pista do VOT e os brasileiros, a intensidade do *burst*, como mencionado no capítulo de introdução, seguindo o estudo de Alves e Motta, 2013). É importante salientar que foram usados os grafemas<sup>9</sup> no lugar da simbologia dos fones, já que os participantes, como pré-requisito básico, não poderiam possuir conhecimentos sobre fonética, e não estariam aptos a reconhecer, portanto, símbolos do IPA<sup>10</sup>. Os participantes estavam aptos a optar pela opção *replay* no áudio quantas vezes pudessem carregar a página, embora essa característica não fosse a desejada para este estudo (este fator será melhor explicado a seguir) e não fosse possível, também, contabilizar o número de repetições efetuadas. Esse, inclusive, pode ser considerado como um ponto desfavorável na plataforma usada para hospedar o teste (a ser apresentada melhor a seguir): o site disponibiliza a possibilidade de deixar um *player* visível para os participantes ou não. Caso o player esteja visível, o participante fica mais consciente de que pode dar *play* inúmeras vezes. Caso o administrador decida não colocar o *player*, o participante, aparentemente, não teria como tocar o mesmo estímulo quantas vezes quisesse, porém, ao se passar para a próxima questão, caso o participante não tenha escolhido nenhuma das opções disponíveis (/p/, /t/, /k/, /b/, /d/, /g/), o site recarrega a página, o que acaba por tocar o estímulo novamente, fazendo com que o participante clique em *next* sem ter marcado nenhuma das opções possíveis.

---

<sup>9</sup> Grafema, entenda-se “letra”.

<sup>10</sup> Sigla em inglês para o referente português AFI (Alfabeto Fonético Internacional), o IPA (*International Phonetic Alphabet*) consiste em um sistema de notação fonética, criado pela Associação Fonética Internacional, a fim de fazer possível uma representação internacional padronizada de qualquer som da fala.

## 4.2 A Plataforma Online Utilizada: *Surveygizmo*

O Surveygizmo (disponível em <https://app.surveygizmo.com/>) é uma plataforma online que possibilita a criação de diversos tipos de enquetes, desde formulários sobre opinião até quizzes, que podem ou não dar feedback para os participantes sobre as respostas corretas (caso elas existam). Utilizado por algumas empresas internacionais para promover pesquisas de satisfação entre os usuários, o Surveygizmo acaba por ser uma ferramenta de extrema utilidade também para pesquisas acadêmicas, já que possibilita aos pesquisadores o compartilhamento online de seus testes, caso não possam coletar participantes de forma pessoal. Existem vários motivos que podem impedir uma coleta ao vivo com os informantes, que vão desde deixá-los mais à vontade no momento de responder aos testes (não comprometendo, assim, os dados, caso eles se sintam intimidados pela presença do pesquisador) até a impossibilidade de fazer um número significativo de coletas em um determinado intervalo de tempo. Para esta pesquisa, a escolha dessa plataforma se deu pela praticidade de coletar dados de informantes estrangeiros, haja vista a impossibilidade de se viajar para aplicar os testes de forma pessoal. Foi pensado, também, o uso do software *Praat*, mas este programa necessita ser instalado em cada computador em que for ser usado para responder aos testes, e isso tornaria a logística de coleta mais complicada, pois, se fosse necessário aplicar os testes em uma turma de 10 americanos ao mesmo tempo, por exemplo, seria necessária a instalação do software tantas vezes fossem os números de máquinas disponíveis. Com a plataforma online Surveygizmo, um link é gerado pelo site, que hospeda páginas como se fossem sites, o que torna o acesso possível de qualquer lugar do mundo, sem ser necessária qualquer instalação. É possível, inclusive, fazer este acesso via dispositivos móveis, como tablets ou smartphones. Para se manter um paralelismo de condições entre os participantes de ambas naturezas, foi decidido que usar a plataforma tanto com americanos quanto com brasileiros seria a forma mais equilibrada de aplicar os testes.

A plataforma possibilita inúmeras configurações de testes e tipos de pergunta, desde abertas até fechadas. Isso foi extremamente importante para que fosse possível se fazer identificação posterior dos participantes brasileiros, pois seria necessário, ao final dos testes, identificá-los a fins de separá-los por níveis de

proficiência, e isso só foi possível graças a possibilidade de se inserir uma pergunta aberta aos participantes, pedindo-se que informassem seus nomes. Sendo de outra forma, apenas uma tabela com o número de respostas por pergunta seria disponibilizado. Abaixo, apresenta-se uma pequena ilustração de como o site era apresentado para os participantes brasileiros:

Figura 18: Primeira página do teste de identificação para os brasileiros<sup>11</sup>

Teste de Identificação - Português  
Page 1 Questions

Muito obrigada por participar dessa pesquisa! Primeiramente, é preciso que você se identifique. Seus dados são confidenciais e serão consultados apenas pela pesquisadora e seu orientador. Digite, no espaço abaixo, seu nome e email. \*

Next  
0%

This student research survey is powered by  
 surveygizmo

[Planning or conducting an academic research project?](#)

Fonte: adaptado do site Surveygizmo

Após a página de identificação, o participante era advertido sobre como o teste deveria ser feito. A utilização de fones de ouvido era solicitada como fator importante, já que nenhum aplicador estaria presente no momento do teste, o que poderia acarretar uma falta de rigor por parte dos participantes:

<sup>11</sup> É importante salientar que as perguntas que continham um asterisco vermelho ao lado direito superior (\*) eram de caráter obrigatório. Assim sendo, o participante só poderia clicar em “next” e ter acesso à página seguinte caso tivesse respondido adequadamente ao tipo de questão proposta. Caso a questão ficasse em branco ou não fosse adequadamente respondida (marcar duas respostas, por exemplo), a página era novamente carregada.

Figura 19: Informação aos participantes

The image shows a screenshot of a survey page. At the top, there is a blue header with the text 'Teste de Identificação - Português' and 'New Page' below it. The main content area is white and contains the following text: '1. Este teste não pode ser repetido. É necessário, portanto, saber se você entendeu a maneira correta de respondê-lo antes de começar. É INDISPENSÁVEL O USO DE FONES DE OUVIDO. Sugere-se, gentilmente, que seja usado o Google Chrome como navegador, já que é o menos propenso a causar problemas no meio do teste, fazendo com que se percam todos os dados. Suas tarefas são: 1) OUVIR a palavra que a página irá carregar automaticamente e 2) IDENTIFICAR qual é a consoante inicial da palavra que você ouviu Você terá duas questões-teste antes de começar.' Below this text are two radio button options: 'Ok, estou usando fones de ouvido e entendi como responder ao teste.' and 'Não entendi e/ou não estou preparado para responder ao teste (se este for o seu caso, por favor, entre em contato com a pesquisadora)'. At the bottom of the page, there are two buttons labeled 'Back' and 'Next', and a progress bar showing '1%'.

Fonte: adaptado do site Surveygizmo

Após confirmar que estava ciente do uso de fones de ouvido e da maneira de responder aos testes, o participante, finalmente, chegava às duas questões de exemplo antes de responder, de fato, às questões que seriam contadas como *tokens* neste estudo. Nas duas questões-teste, foram utilizados estímulos com o fonema /g/, com alto pré-vozeamento (Gill: -30ms e Good: -32ms) para que se minimizassem dúvidas em relação à identificação da consoante por parte dos participantes antes de que o teste fosse começar, aumentando, assim, a chance de confiança no entendimento das questões.

Figura 20: Layout do Teste de Identificação Brasileiros

The screenshot shows a survey interface with a blue header containing the text "Teste de Identificação - Português" and "Page 1 Questions". The main content area displays a question: "2. Qual é a CONSOANTE INICIAL da palavra que você ouviu? \*". Below the question are six radio button options: "p", "b", "k", "g", "t", and "d". At the bottom of the question area, there are "Back" and "Next" buttons and a progress bar showing 3% completion.

Fonte: adaptado do site Surveygizmo.

Finalmente, antes de os testes começarem, após as duas questões-teste serem aplicadas, o participante deveria confirmar que entendeu como as perguntas deveriam ser respondidas e que estava familiarizado com a plataforma, podendo, portanto, continuar com o resto da tarefa. As telas seguintes eram de interface idêntica às das questões de teste antes aplicadas.

Figura 21: Confirmação de Entendimento do Teste

The screenshot shows a survey interface with a blue header containing the text "Teste de Identificação - Português" and "New Page". The main content area displays a question: "6. Ok, você entendeu? O teste pode continuar?". Below the question are two radio button options: "Sim" and "Não". At the bottom of the question area, there are "Back" and "Next" buttons and a progress bar showing 5% completion.

Fonte: adaptado do site Surveygizmo.

Apesar de se deixar claro o uso do fone de ouvido ou a certeza em relação à realização do teste, não é totalmente confiável que todos os participantes tenham, de fato, usado o acessório pedido, ou que tenham entendido completamente a

maneira de se portar diante da plataforma e das perguntas a serem respondidas. É importante salientar que três participantes brasileiros e dois americanos entraram em contato para que dúvidas fossem sanadas em relação ao que estava sendo esperado como resposta.

Existem, como já mencionados acima, alguns pontos positivos em usar esse tipo de plataforma em pesquisas acadêmicas. Não se podem esquecer, contudo, os pontos desfavoráveis em relação ao uso de tal plataforma. Além da incerteza do uso de acessórios necessários (neste estudo, o fone de ouvido) ou do entendimento do teste, existem outros pontos a serem considerados como negativos no uso dessa plataforma em específico. O tempo na montagem de cada uma das questões foi algo que se esperava ser menos demorado. De fato, para a montagem, é necessário fazer upload de todos os áudios que se pretendem usar como estímulo para uma espécie de “nuvem de arquivos” do próprio site. Após isso, ao se montar cada uma das questões, é preciso copiar cada uma das URLs que o site cria para cada um dos áudios. Há a possibilidade de copiar as questões, mas, se uma questão é copiada já com um áudio específico definido, este mesmo áudio será copiado para todas as outras questões, gerando a necessidade de se editar uma página por vez para a mudança de arquivo de áudio.

Ainda em relação aos arquivos de áudio, levando-se em consideração que o teste é online, é preciso estar ciente de que uma conexão de *internet* de qualidade razoável é necessária. Se a conexão cair, por exemplo, durante a participação de algum informante, ele não conseguirá recomeçar as perguntas de onde parou, sendo obrigado, portanto, a respondê-las todas novamente, o que não seria interessante para os objetivos deste trabalho. Outro fator em relação aos arquivos de áudio que pode ser crucial reside no fato de que, para este tipo de teste, o ideal é usar arquivos no formato .wav, que são mais pesados do que os de formato .mp3. Os arquivos em formato .mp3 são contraídos, e devem ser evitados, já que, apesar de serem mais leves e preferíveis para se evitar queda de conexão de rede, podem influenciar em alguma possível pista acústica relevante, ainda que desconhecida. Além disso, outro fator negativo era a impossibilidade de randomização dos estímulos para a ordem de apresentação das perguntas, característica mudada da

última versão do site para a atual. Essa não randomização fez com que as questões do teste aparecessem na mesma ordem para os americanos e para os brasileiros, não podendo ser evitado, dessa forma, efeitos da tarefa sobre os participantes. Por questão de organização, geralmente os estímulos são montados em ordem alfabética, para se facilitar a correção dos mesmos ao final. Para se evitar qualquer problema de confusão por parte dos informantes, foi necessário o uso de um randomizador online, o *Reasarch Randomizer* (disponível em: <http://www.randomizer.org/>). Isso evitaria que a questão 1 tivesse o estímulo *pit* com 100% de aspiração, seguido da questão 2, com o mesmo estímulo contendo 75% de aspiração em relação ao primeiro, a questão 3 com 50% da aspiração do mesmo estímulo *pit* e assim por diante. Abaixo, descrevem-se os passos para randomizar números, utilizando-se o referido site:

- 1) Define-se o número de sequências desejadas para o intervalo de números disponibilizados;
- 2) Aponta-se a quantidade de números que devem ser randomizados;
- 3) Seleciona-se qual deve ser o intervalo desses números;
- 4) Define-se se o programa deve repetir os números mais de uma vez por sequência.

Figura 22: Montagem da Randomização pelo ResearchRandomizer

The screenshot shows the Research Randomizer website interface. At the top, there is a navigation bar with links for 'Randomize', 'Tutorial', 'Links', and 'About Us'. The main content area is titled 'RESEARCH RANDOMIZER' and contains a form for generating random numbers. The form includes the following fields and options:

- To generate random numbers, enter your choices below (using integer values only):**
- How many sets of numbers do you want to generate?** Input: 1
- How many numbers per set?** Input: 72
- Number range (e.g., 1-50):** From: 1, To: 72
- Do you wish each number in a set to remain unique?** Yes
- Do you wish to sort the numbers that are generated?** No
- How do you wish to view your random numbers?** Place Markers Off

Each input field has a 'Help' link next to it. A 'Randomize Now!' button is located at the bottom of the form. On the right side, there is a 'Site Overview' sidebar with links to 'Randomize Now', 'Quick Tutorial', 'Related Links', and 'About Research Randomizer'. Below the sidebar is a 'Randomizer Box' with a small thumbnail of the tool and text: 'Add this tool to your website and generate your own number sets.'

Copyright ©1997-2008 by Geoffrey C. Urbaniak and Scott Plous | Site Statistics

Fonte: adaptado do site ResearchRandomizer.

Após a execução dos passos referidos, a sequência pode, finalmente, ser consultada:

FIGURA 23: Exemplo do Randomizador *Research Randomizer*

The screenshot shows the Research Randomizer website displaying the results of a randomization process. The page is titled 'RESEARCH RANDOMIZER' and includes a navigation bar with links for 'Print', 'Download in Excel', and 'Close'. The main content area is titled 'Research Randomizer Results' and contains the following information:

- 1 Set of 72 Unique Numbers Per Set**
- Range: From 1 to 72 -- Unsorted**
- Job Status: Finished**

Below the results, there is a section titled 'Set #1:' followed by a list of 72 unique numbers: 66, 10, 37, 1, 72, 30, 51, 36, 61, 46, 20, 58, 11, 50, 25, 55, 57, 59, 15, 32, 23, 9, 12, 33, 40, 41, 44, 53, 27, 60, 71, 48, 63, 69, 34, 7, 17, 47, 13, 24, 42, 67, 56, 65, 6, 28, 18, 52, 29, 62, 68, 2, 22, 16, 4, 5, 21, 54, 14, 39, 26, 3, 19, 38, 45, 43, 8, 35, 70, 64, 49, 31.

Fonte: adaptado do site Research Randomizer.

A versão em inglês do teste continha algumas perguntas extras, relativas a informações pessoais dos participantes, para futuras criações de categorias entre eles e um *consent form*, que se fez necessário sob a modalidade online, já que as assinaturas e informações não poderiam ser recolhidas em papéis como feito com os informantes brasileiros

## 5 ANÁLISE

Após a coleta feita com os 12 americanos e os 16 brasileiros, poderão ser vistas, a seguir, os resultados das análises que se voltam à verificação das hipóteses propostas no capítulo primeiro deste trabalho, sobre os dados coletados e apresentados.

É importante que se faça, primeiramente, a devida explanação de como as respostas foram analisadas.

Houve, na análise deste trabalho, três possibilidades de enquadramento das respostas<sup>12</sup>, sendo elas consideradas como: *surdo*<sup>13</sup> (quando o participante marcasse uma consoante surda ao identificar o estímulo), *sonoro* (quando o participante escolhesse uma consoante sonora ao identificar o estímulo) e *erro*<sup>14</sup> (que foi assim considerado quando o informante escolhesse marcar como resposta ao estímulo ouvido uma consoante que não fosse do mesmo ponto de articulação do estímulo original).

Os dados serão apresentados em tabelas e gráficos com suas respectivas interpretações, levando-se em consideração: o grupo de informantes, ponto de articulação identificado pelos participantes, grau de manipulação realizada no VOT e verificação das hipóteses apresentadas no capítulo 1.

---

<sup>12</sup> Baseado no estudo de Alves, Schwartzhaupt e Baratz (2011).

<sup>13</sup> Apenas seriam consideradas *surdas* ou *sonoras* (e corretas) as respostas que tiveram apenas o padrão VOZAMENTO como diferença entre a consoante manipulada e ouvida pelo participante. Assim sendo, se houvesse um estímulo “pit”, com 75% de sua aspiração cortada, as únicas respostas “corretas” a serem consideradas seriam a marcação de ‘p’ – que foi considerado como surdo – ou ‘b’ – considerado como sonoro.

<sup>14</sup> Sempre que houve marcação de resposta para um estímulo com ponto de articulação diferente daquele do estímulo original, a resposta do participante foi considerada errada. Logo, se o informante fosse exposto ao estímulo “kill”, com 50% de sua aspiração retirada, por exemplo, as únicas respostas corretas seriam ‘k’ ou ‘g’, sendo ‘t’, ‘d’, e ‘p’, ‘b’ consideradas erradas.

## 5.1 Descrição Geral dos Dados de Americanos

A seguir, serão verificados, de forma geral, como o grupo de americanos se comportou perante os estímulos ouvidos, levando-se em consideração cada um dos pontos de articulação das consoantes, individualmente, e seus graus de manipulação.

Considerando-se, primeiramente, a consoante bilabial surda e suas subseqüentes manipulações, apresenta-se, abaixo, a tabela referente aos índices de identificação de tal consoante, seguida da sua análise:

TABELA 2: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE PLOSIVA BILABIAL SURDA PARA AMERICANOS

PONTO DE ARTICULAÇÃO E PORÇÃO DE VOT PRESERVADA	SURDO	SONORO	ERRO
/p/ 100%	<b>93,75%</b> 60	<b>0%</b> 0	<b>6,25%</b> 4
/p/75%	<b>85,93%</b> 55	<b>1,56%</b> 1	<b>12,5%</b> 8
/p/ 50%	<b>64,06%</b> 41	<b>25%</b> 16	<b>10,93%</b> 7
/p/ 25%	<b>85,93%</b> 55	<b>1,56%</b> 1	<b>12,5%</b> 8
/p/ 0%	<b>48,43%</b> 31	<b>14,06%</b> 9	<b>37,5%</b> 24

Quanto aos índices de identificação para /p/, apresentados na tabela acima, pode-se verificar o seguinte:

Percebe-se que o grupo de americanos teve uma tendência de queda de atribuição de *status surdo* à medida que a manipulação foi aumentando, mas essa regularidade se deu, apenas, nos três primeiros graus de manipulação [(/p/ 100% *surdo* (93,75%) > /p/ 75% *surdo* (85,93%) > /p/ 50% *surdo* (64,06%)]. Apesar de ser possível verificar essa tendência de queda a partir dos dados descritivos, o teste estatístico de Friedman não apontou haver diferenças significativas quando comparadas, entre si, as porções de VOT consideradas *surdas* pelos americanos ( $X^2(2) = 6.200, p = 0.260$ ).

Quando deparados com a bilabial surda com apenas 25% da sua aspiração mantida, ao contrário do que poderia ser esperado, caso os americanos estivessem seguindo a pista VOT como primordial na identificação de consoantes, houve um grande índice de atribuição do *status surdo* (85,93%), mesmo com um corte quase total da aspiração, e um baixo índice de atribuição de *status sonoro* (1,56%). Considerando a manipulação total feita em tal consoante, encontra-se a distribuição entre as três possibilidades (*surdo, sonoro, erro*) de maneira mais equilibrada (*surdo*: 48,43%; *sonoro*: 14,06%; *erro*: 37,5%), assim como no corte de 50% na bilabial surda. Os índices indicados para o VOT de padrão zero artificial atribuído para /p/ dão indícios de que a falta de aspiração possa vir a confundir a identificação dos americanos, já que esta manipulação teve leve grau de atribuição de *status sonoro* (14,06%), porém, um considerável grau de *erro* (37,5%).

É importante verificar também que, ao mesmo tempo em que, pelo menos nas primeiras manipulações, a atribuição de *status* como *surdo* parecia aumentar de forma inversamente proporcional à manipulação, o *status* de atribuição para as consoantes como *sonoro* tendeu a aumentar para esses mesmos três graus [(/p/ 50% *sonoro* (25%) > /p/ 75% *sonoro* (1,56%) > /p/ 100% *sonoro*(0%)]. Esse aumento, entretanto, também não apresentou diferença estatística significativa, quando comparadas as fatias de VOT consideradas como *sonora* pelos americanos. Novamente, esse resultado foi verificado a partir da utilização de um teste Friedman, que comparou as cinco manipulações, agora consideradas *sonoras*, entre si ( $X^2(2) = 5.200, p = 0.460$ ).

Finalmente, quando as fatias de VOT que foram enquadrados na categoria *erro* foram comparadas entre si, a partir de um teste Friedman, diferenças

significativas não foram encontradas ( $X^2 (2) = 2.200, p = 0.360$ ). Em suma, apesar dos diferentes graus de manipulação, não foram encontradas, nos dados dos americanos, diferenças significativas nos índices referentes à plosiva bilabial, para nenhum dos três tipos de resposta (surda, sonora, ou equivoco referente ao ponto de articulação).

Seguem, agora, os dados e análises referentes à alveolar surda, levando-se em consideração os índices de identificação para essa consoante.

TABELA 3: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE ALVEOLAR SURDA PARA AMERICANOS

PONTO DE ARTICULAÇÃO E PORÇÃO DE VOT PRESERVADA	SURDO	SONORO	ERRO
/t/ 100%	<b>96,87%</b> 62	<b>0%</b> 0	<b>3,12%</b> 2
/t/ 75%	<b>50%</b> 32	<b>25%</b> 16	<b>25%</b> 16
/t/ 50%	<b>65,62%</b> 42	<b>7,81%</b> 5	<b>26,56%</b> 17
/t/ 25%	<b>71,87%</b> 46	<b>3,12%</b> 2	<b>25%</b> 16
/t/ 0%	<b>32,81%</b> 21	<b>29,68%</b> 19	<b>37,5%</b> 24

Considerando-se, então, a alveolar surda e suas manipulações, os índices não parecem seguir uma relação regular entre padrão de vozeamento e grau de manipulação. Ao contrário da bilabial surda anteriormente apresentada, a consoante alveolar não parece exibir uma regularidade nos índices a ela atribuídos. Como esperado, levando-se em conta a premissa inicial de que os americanos seguiriam o VOT como pista acústica majoritária na distinção entre segmentos surdos e sonoros, a atribuição do *status* fonológico para a produção fonética de 100% da aspiração foi

considerada como *surda* (96,87%). Houve, mesmo que com índice pequeno, uma parcela de informantes que consideraram o VOT genuíno como sendo *sonoro* (3,12%), mas não houve, porém, nenhuma resposta errada. Assim sendo, mesmo que o padrão de vozeamento tenha cambiado levemente, a identificação do ponto de articulação foi mantida, uma vez que tal ponto foi identificado como sendo o mesmo por todos os participantes. A retirada do primeiro quarto de aspiração resultou em índices ainda altos para o *status surdo* (50%), embora os índices de *sonoro* (25%) e *erro* (25%) tenham se equivalido. Isso pode se dever, talvez, pelo fato de um corte de 25% da aspiração ainda manter alguma duração significativa de VOT para alguns informantes, enquanto para outros essa aspiração não fosse mais o suficiente para ser considerada como pertencente a uma consoante surda, enquanto, para ainda outros, causasse confusão, já que parece que não havia aspiração suficiente para considerá-la surda e nem insuficiente o bastante para considerá-la como sonora. Comparando-se os graus de 50% e 25%, entretanto, encontram-se os índices mais inesperados, já que a atribuição do *status surdo* cresce (de 50% para a manipulação de apenas 25% da aspiração mantida para um índice de 65,62% para a manipulação intermediária).

Finalmente, observando-se os índices atribuídos à manipulação total da alveolar surda, têm-se indícios de que os americanos, talvez, não estejam seguindo apenas a pista VOT como primordial na identificação de tal consoante: percebe-se que, ao contrário da bilabial surda /p/, a alveolar surda /t/ tem distribuição bastante regular se comparada à mesma manipulação total da primeira consoante citada (surdo: 32,81%; sonoro: 29,68%; erro: 37,5%). O índice de consideração para *sonoro* aumenta (de 3,12% para 29,68%), como o esperado, entretanto, a atribuição para *surda* não diminui (de 25% para 37,5%). O índice de erro se equivale aos outros dois padrões de vozeamento (surdo: 32,81%, sonoro: 29,68%, erro: 37,5%). O teste de Friedman também não apresentou haver diferenças significativas quando comparados, entre si, as respostas atribuídas para *surdo* ( $X^2(2) = 7.200$ ,  $p = 0.440$ ) *sonoro* ( $X^2(2) = 6.200$ ,  $p = 0.340$ ) e *erro* ( $X^2(2) = 7.400$ ,  $p = 0.540$ ).

Em suma, no que diz respeito às respostas dos americanos, resultados estatisticamente significativos não foram encontrados em função dos diferentes graus de manipulação aos quais foram submetidos os participantes. Uma possível

explicação para tais resultados pode ser atribuída a características fonético-articulatórias do ponto de articulação desta consoante em si, cuja duração do VOT estaria na metade da média de aspiração das três consoantes comparadas (/k/ > /t/ > /p/) (LISKER e ABRAMSON, 1964)<sup>15</sup>.

Serão apresentados, no que segue, os índices de identificação para a velar surda, conforme tabela a seguir.

TABELA 4: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE VELAR SURDA PARA AMERICANOS

PONTO DE ARTICULAÇÃO E PORÇÃO DE VOT PRESERVADA	SURDO	SONORO	ERRO
/k/ 100%	<b>96,87%</b> 62	<b>0%</b> 0	<b>3,12%</b> 2
/k/ 75%	<b>95,31%</b> 61	<b>1,56%</b> 1	<b>3,12%</b> 2
/k/ 50%	<b>64,06%</b> 41	<b>0%</b> 0	<b>35,93%</b> 23
/k/ 25%	<b>48,43%</b> 31	<b>23,43%</b> 15	<b>28,12%</b> 18
/k/ 0%	<b>34,37%</b> 22	<b>39,06%</b> 25	<b>26,56%</b> 17

Partindo-se, então, para a análise da velar surda /k/, vê-se que, de acordo com o esperado, o estímulo genuíno tem altíssimo grau de atribuição de *status surdo* (96,87%), e nenhum índice para *sonoro*, embora haja um índice pequeno para erro (3,12%). A manipulação de 25% da aspiração (preservação de 75% do VOT original) ainda é suficiente para que a maioria dos americanos considere tal consoante manipulada como surda (50%), enquanto a outra metade se divide entre sonoro (25%) e erro (25%). Ao mesmo tempo, verifica-se um leve aumento nos

<sup>15</sup> Os autores propõem, no texto de 1964, as seguintes médias de duração para cada um dos pontos a seguir: /p/ média de 58ms, /t/ média de 70ms e /k/ média de 80ms. Quanto mais anterior for a consoante, maior será a duração da sua aspiração.

índices de *status sonoro* e de erro. A manipulação de 50%, novamente, apresenta comportamento irregular, porém, possivelmente explicável: há um índice considerável para atribuição de *status surdo* (64,04%), porém, há, também, um índice relativamente alto para erro (35,93%). Novamente, é possível que ainda haja uma aspiração consideravelmente alta para que a consoante seja considerada surda, enquanto, para outros informantes, essa aspiração não seja suficiente, causando confusão no momento da identificação e levando, portanto, à troca na identificação de ponto de articulação. Indo ao encontro da ideia de que quanto maior fosse a manipulação, menor seria a atribuição do *status surdo* para as consoantes modificadas, o corte feito para se chegar a apenas 25% da aspiração da coronal surda apresentou índices altos para o padrão *surdo* (48,43%), porém, dentro de todas as manipulações, tem a terceira posição na atribuição do *status sonoro* (perdendo, apenas, como esperado, para os VOTs que possuem maiores pedaços de aspiração mantidos: 100% de aspiração mantida: 96,87% e 75% de aspiração mantida: 95,31%). O índice de erro, em relação à manipulação de 50%, aumenta (35,93%), ao ponto que o padrão *sonoro* diminui (0%), o que fortaleceria a possibilidade de que uma aspiração cortada pela metade pudesse ser confusa para os participantes, que não conseguiriam encontrar espaço no seu inventário fonológico para uma manipulação “no meio do caminho”. Quanto à manipulação total de /k/, verifica-se que essa consoante foi a que mais obteve trocas nas identificações de vozeamento, se comparadas às manipulações semelhantes feitas em /p/ e /t/: foi nessa consoante, com sua aspiração total retirada, que mais se obteve atribuição do *status sonoro* (39,06%), embora ainda houvesse um número considerável de participantes que consideraram tal consoante como surda (34,37%), e houvesse um número alto para erro também.

Corroborando os índices acima explicitados pelos dados descritivos, o teste estatístico de Friedman encontrou diferenças significativas para todas as possíveis respostas atribuídas à consoante velar: *surdo* ( $X^2(2) = 7.500$ ,  $p = 0.003$ ), *sonoro* ( $X^2(2) = 7.500$ ,  $p = 0.002$ ) e *erro* ( $X^2(2) = 7.500$ ,  $p = 0.004$ ). Como o teste de Friedman compara apenas as porções de cada atribuição de resposta entre si, sem indicar onde a diferença significativa pode ser encontrada, os dados foram submetidos a uma segunda etapa de testes estatísticos nesta etapa, um teste de Wilcoxon Z

comparou, de forma pareada, os graus de manipulação e o *status* a cada um atribuído. Em outras palavras, através desses testes de Wilcoxon de caráter ad-hoc, pode-se explicitar onde, mais especificamente, essas diferenças se fazem presentes.

A partir da aplicação dos testes ad-hoc de Wilcoxon, para a atribuição de *status surdo*, foram encontradas diferenças significativas para mais de uma comparação entre as manipulações comparadas. Primeiramente, ao que parece, a comparação entre o padrão zero de VOT e as manipulações de, no mínimo, 50% da aspiração, são as que se fazem mais evidentes no que tange a diferenças significativas, ou seja, parece haver um limiar mínimo de metade da aspiração total de /k/, comparada à aspiração de padrão zero, para que haja influência na escolha de *status surdo* para a consoante velar (0 artificial x 50% retirado:  $Z = -2,994$ ;  $p = 0.003$ ; 0 artificial x 75% preservado:  $Z = -3,213$ ;  $p = 0.001$  e 0 artificial x 100% preservado:  $Z = -3,213$ ;  $p = 0.001$ ). Como as diferenças encontradas em maior número se dão com comparações feitas entre o zero artificial com, no mínimo, 50% da aspiração mantida, parece haver indícios de que o limiar fonético para que a consoante velar /k/ seja considerada surda se encontre por volta de metade do seu valor de aspiração preservado. Partindo para as comparações outras que apresentaram valores significativos, há indícios de que, não apenas seria necessária a preservação de, no mínimo, metade da aspiração, mas sim, um mínimo de 50% de diferença entre os graus de aspiração comparados, já que foram encontradas diferenças na comparação de 25% de aspiração de /k/ mantida contra 75% de aspiração mantida ( $Z = -2,877$ ;  $p = 0.004$ ) e, também, na comparação de 25% de aspiração preservada contra 100% de aspiração preservada ( $Z = -2,877$ ;  $p = 0.004$ ).

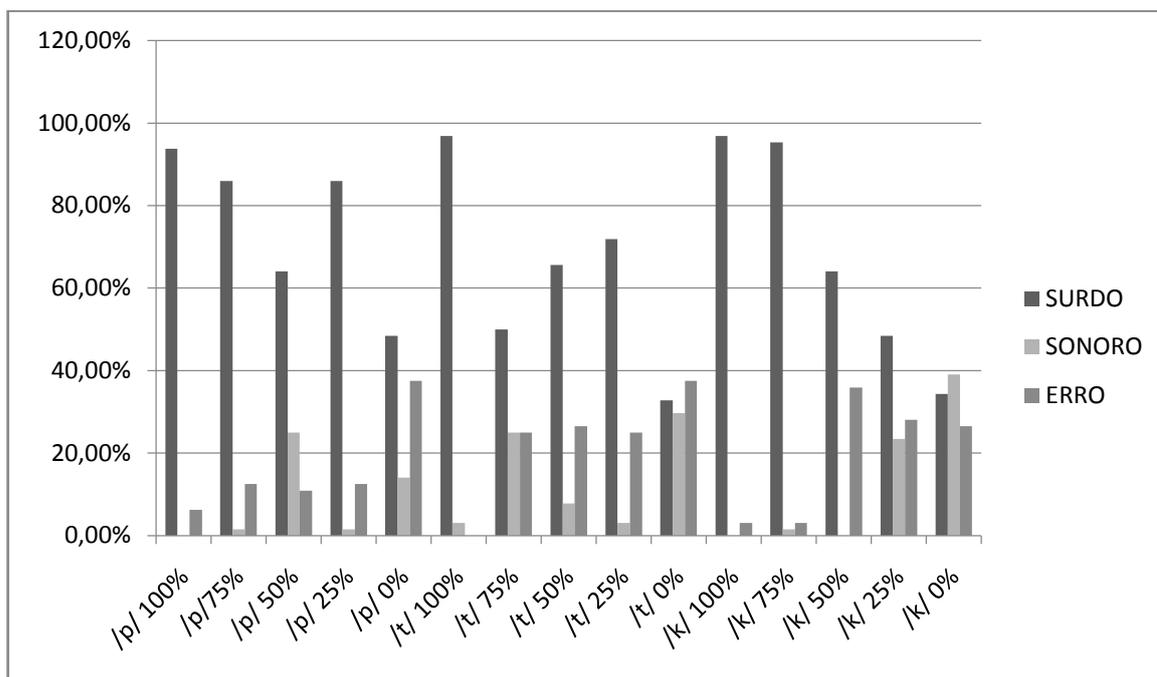
Esse limiar de 50% para /k/ também parece ser encontrado nas comparações pareadas feitas para as respostas atribuídas ao *status sonoro*, já que foram apontados três valores significativos pelo teste Wilcoxon, e todos tinham como comparação o VOT de padrão zero, respectivamente comparado com 50% ( $Z = -3,002$ ;  $p = 0.003$ ), 75% ( $Z = -3,002$ ;  $p = 0.003$ ) e 100% ( $Z = -3,002$ ;  $p = 0.003$ ) de aspiração mantidas. Dessa forma, parece haver, novamente, um limiar de 50% de aspiração mantida para que /k/ seja, de certa forma, identificada com vozeamento diferente, do padrão de VOT zero.

Diferenças significativas também foram encontradas na consoante velar quando comparados os índices de *erro* para essa consoante e, novamente, as comparações que apresentaram diferenças tinham o padrão de VOT zero com índices mais altos de aspiração mantidos. Nesse caso, o limiar de 50% não foi mantido como mínimo, o que pode indicar, levando-se em consideração as comparações feitas acima, que uma aspiração mínima de 50% de /k/ possa influenciar a percepção dos participantes americanos dessa pesquisa apenas no que tange ao padrão de vozeamento da consoante, mas não seja, ainda, o suficiente para que configure uma troca de ponto de articulação. O limiar para este caso, entretanto, parece ser um pouco maior, logo na fatia dos 75% de aspiração mantida, quando comparada à fatia de VOT de padrão zero ( $Z = -2,121$ ;  $p = 0.003$ ). O teste de Wilcoxon também apresentou diferença significativa na comparação feita para o índice de *erro* entre os dois extremos de aspiração: o mais manipulado – padrão zero artificial – *versus* o estímulo genuíno – com toda a aspiração mantida ( $Z = -2,121$ ;  $p = 0.003$ ).

Essas diferenças significativas nas atribuições dos stati de 'surdo' e 'sonoro' possivelmente possam ser explicadas, novamente, pelas características fonético-articulatórias dessa consoante: por ser /k/ uma consoante de alta aspiração, os informantes, provavelmente, “esperavam” uma aspiração longa. Como tal aspiração não ocorreu, essa situação pode ter feito com que o estímulo por eles ouvido tenha sido considerado sonoro, ou, ainda, como outra consoante de outro ponto de articulação com aspiração, como /p/, por exemplo; de fato, uma redução na aspiração de um /k/ pode ser, em termos de duração absoluta de VOT em milissegundos, semelhante ou equivalente à uma aspiração de um /p/ normalmente aspirado.

Como fechamento dos dados acima apresentados em tabelas e brevemente descritos, segue, abaixo, um gráfico ilustrativo dos índices de identificação por consoante e suas respectivas manipulações, de acordo com o ponto de articulação de cada uma delas:

GRÁFICO 1 – ÍNDICES DE IDENTIFICAÇÃO POR PONTO DE ARTICULAÇÃO AMERICANOS



Os dados apresentados indicam duas fortes tendências no que tange à identificação de consoantes surdas por parte dos americanos: há indícios de que não somente o VOT seja a única pista seguida pelos falantes nativos de língua inglesa no momento da tarefa de identificação aqui proposta, mas sim, há a possibilidade de ação conjunta de outras pistas. Isso fica sugerido pela não troca de vozeamento *surdo/sonoro* à medida que o corte de VOT aumenta. Ainda assim, os dados de /k/ parecem sugerir um caráter prioritário para o VOT, frente às outras pistas acústicas. Cabe mencionar essa manipulação crescente parece influenciar os informantes de outra forma, o que aumenta o índice de erro por parte dos americanos. Essa troca de identificação do ponto de articulação atribuído pode reforçar a ideia de que pistas outras trabalham em conjunto na identificação dos segmentos manipulados.

## 5.2 Descrição Geral dos Dados de Brasileiros

A seguir, serão verificados, de forma geral, como o grupo de brasileiros se comportou perante os estímulos, levando-se em consideração o ponto de articulação

das consoantes e seus graus de manipulação, considerando-os, ainda, como um grupo só, sem separá-los por nível de proficiência, análise essa que será feita no final deste capítulo.

Primeiramente, serão apresentados os dados de identificação da bilabial surda, conforme tabela abaixo:

TABELA 5: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE PLOSIVA BILABIAL SURDA PARA BRASILEIROS

PONTO DE ARTICULAÇÃO E PORÇÃO DE VOT PRESERVADA	SURDO	SONORO	ERRO
/p/ 100%	<b>79,16%</b> 38	<b>12,5%</b> 6	<b>8,33%</b> 4
/p/75%	<b>45,83%</b> 22	<b>41,66%</b> 20	<b>12,5%</b> 6
/p/ 50%	<b>70,83%</b> 34	<b>12,5%</b> 6	<b>16,66%</b> 8
/p/ 25%	<b>6,25%</b> 3	<b>47,91%</b> 23	<b>45,83%</b> 22
/p/ 0%	<b>27,08%</b> 13	<b>72,91%</b> 35	<b>0%</b> 0

Considerando-se, então, a bilabial surda, percebe-se um comportamento irregular vindo dos brasileiros: como esperado, a atribuição de *status surdo* à consoante não diminui ao mesmo passo em que a manipulação aumentou. Verificando-se os quatro primeiros cortes (incluindo o estímulo não manipulado), entretanto, percebe-se que o nível de erro aumenta (100%: 8,33; < 75%: 12,5; < 50%: 45,83%). O teste estatístico de Friedman não apontou diferenças significativas quando comparados os graus de manipulação para *surdo* entre si ( $X^2(2) = 7.200$ ,  $p = 0.310$ ). É possível, portanto, que, à medida que a manipulação aumente, aumente

a dificuldade de se enquadrar fonologicamente a percepção referente ao padrão de vozeamento, e às vezes, ao ponto de articulação dessas consoantes, fazendo, assim, com que o grau de erro aumente, embora essa dificuldade não tenha sido grande o suficiente para que se tornasse estatisticamente significativa ( $X^2 (2) = 6.400$ ,  $p = 0.320$ ). É importante salientar, então, que, quando deparados com a ausência quase total de aspiração, o índice de erro inexistente, dando lugar a uma atribuição de *status surdo* menor (27,08%) e uma grande tendência de atribuição do *status sonoro* (72,91%) atribuído ao VOT de padrão zero, o que seria mais esperado do grupo de americanos, que, supostamente, considerariam o VOT como pista primordial. Apesar de haver essa tendência relevante no *status sonoro*, não foram encontrados indícios na estatística de que hajam diferenças significativas quando comparadas respostas de todos os índices para *sonora* entre si.

Serão apresentados, na tabela 6, os dados referentes à alveolar surda para os brasileiros:

TABELA 6: ÍNDICES DE IDENTIFICAÇÃO PARA A PLOSIVA ALVEOLAR SURDA PARA BRASILEIROS

/t/ 100%	<b>97,91%</b> 47	<b>0%</b> 0	<b>2,08%</b> 1
/t/ 75%	<b>97,91%</b> 47	<b>0%</b> 0	<b>2,08%</b> 1
/t/ 50%	<b>25%</b> 12	<b>45,83%</b> 22	<b>29,16%</b> 14
/t/ 25%	<b>33,33%</b> 16	<b>22,91%</b> 11	<b>43,75%</b> 21
/t/ 0%	<b>20,83%</b> 10	<b>47,91%</b> 23	<b>31,25%</b> 15

Levando-se em consideração agora a alveolar surda /t/, percebe-se um comportamento ainda irregular quando relacionados atribuição de padrão de vozeamento e grau de manipulação. Como esperado para o grupo de brasileiros, que, provavelmente, não considerariam o VOT uma pista tão importante quanto os americanos, percebe-se que a manipulação primeira, de retirada de apenas 25% da

aspiração (o que resulta em 75% da aspiração original preservada), em nada afeta a percepção dos participantes, que continuam atribuindo o *status* de *surdo* para ambos os casos (atribuição de *status surdo* de 97,91% tanto para 100% quanto para 75% da aspiração mantida). Essa atribuição do *status surdo* diminui bruscamente, entretanto, na manipulação de metade da aspiração (25%), dando lugar a um alto índice de atribuição de *status sonoro* (45,83%) e grau de erro (29,16%). Esses graus oscilam de maneira semelhante até a manipulação chegar na sua aspiração completamente cortada: apesar de se esperar que os brasileiros não sigam o VOT como pista primordial na identificação de consoantes, parece existir, nesse grupo de brasileiros estudados, alguma fina característica presente nesse VOT manipulado que, ao ser cortado, fez com que os padrões de vozeamento para /t/ não necessariamente se trocassem de *surdo* para *sonoro*, mas sim, possibilitem uma maior distribuição entre as três possibilidades de resposta (*surdo*, *sonoro*, erro referente ao ponto de articulação).

É importante retomar o fato de que, em termos de ponto de articulação, esta consoante está no meio das três consoantes que são produzidas seguidas de aspiração e que sua manipulação, talvez por isso, não apresente uma mudança de vozeamento equilibrada. Se, por um lado, um pequeno corte de um /k/ possa não acarretar em uma mudança de padrão de vozeamento em questão de percepção de som (muito provavelmente pelo fato de o /k/ apresentar um VOT bastante longo, de modo que, ainda que manipulado, possa vir a soar como aspirado), por outro, o corte de um /t/, cujo valor de VOT já está, por natureza, “no meio do caminho”, pode acabar por distribuir a percepção dos participantes, que não apresentam comportamento semelhante como grupo. A observação possibilitada pela estatística descritiva não permite que seja atribuída nenhuma analogia no que tange grau de manipulação e padrão de vozeamento em relação à consoante /t/ por parte dos informantes brasileiros e essa inferência é corroborada pela estatística inferencial, que não revelou haver diferenças significativas, a partir de um teste de Friedman, para nenhuma das atribuições de respostas com os índices comparados entre si, tendo por base cada uma das possibilidades de respostas atribuídas a partir das manipulações da alveolar -> *surdo*:  $X^2(2) = 4.200$ ,  $p = 0.540$ ; *sonoro*:  $X^2(2) = 3.200$ ,  $p = 0.440$ ; *erro*:  $X^2(2) = 5.200$ ,  $p = 0.640$ . Em outras palavras, no que diz respeito à

consoante alveolar, não há diferenças significativas, nas respostas atribuídas pelos aprendizes, em função do grau de manipulação da consoante.

Serão apresentados a seguir, na tabela 7, os dados de identificação para a consoante velar surda:

TABELA 7: ÍNDICES DE IDENTIFICAÇÃO PARA A CONSOANTE PLOSIVA VELAR SURDA PARA BRASILEIROS

/k/ 100%	<b>100%</b> 48	<b>0%</b> 0	<b>0%</b> 0
/k/ 75%	<b>100%</b> 48	<b>0%</b> 0	<b>0%</b> 0
/k/ 50%	<b>62,5%</b> 30	<b>25%</b> 12	<b>12,5%</b> 6
/k/ 25%	<b>50%</b> 24	<b>39,58%</b> 19	<b>10,41%</b> 5
/k/ 0%	<b>8,33%</b> 4	<b>45,83%</b> 22	<b>45,83%</b> 22

Analisando-se, então, a consoante velar /k/, começa-se a ser observada uma pequena regularidade na relação entre grau de manipulação e atribuição de padrão de vozeamento, mas somente no que tange aos padrões *surdo* e *sonoro*, estando o erro à parte dessa relação. Percebe-se que a atribuição do *status surdo* é igualmente alta nos estímulos que mantêm sua total ou quase total duração de aspiração mantida (100% de *status surdo* tanto para o estímulo genuíno quanto para o estímulo com 75% de sua aspiração mantida). Assim sendo, quando deparados com os estímulos que contenham 100% e 75% de sua aspiração cortada, os brasileiros oscilam entre considerar tais consoantes como *sonoro* (39,58% para 25% de aspiração mantida e 45,83% no padrão zero artificial) ou recaem em erro (índice de 45,83% em ambas as últimas duas fatias de manipulação). Quando essa manipulação aumenta, entretanto, parece haver, ao mesmo tempo, uma queda de atribuição do *status surdo* às consoantes, da mesma forma que ocorreu com os participantes americanos. A diferença aqui a ser explicitada se dá na crescente

atribuição do *status sonoro* em relação ao aumento do grau de manipulação, e um índice irregular de erro. Esse aumento da atribuição do caráter sonoro seria mais esperada pelo grupo de americanos do que pelo de brasileiros. É importante, também, verificar que o índice de erro se iguala ao de *sonoro* no caso da manipulação total da aspiração em /k/ (45,82% em ambos os casos).

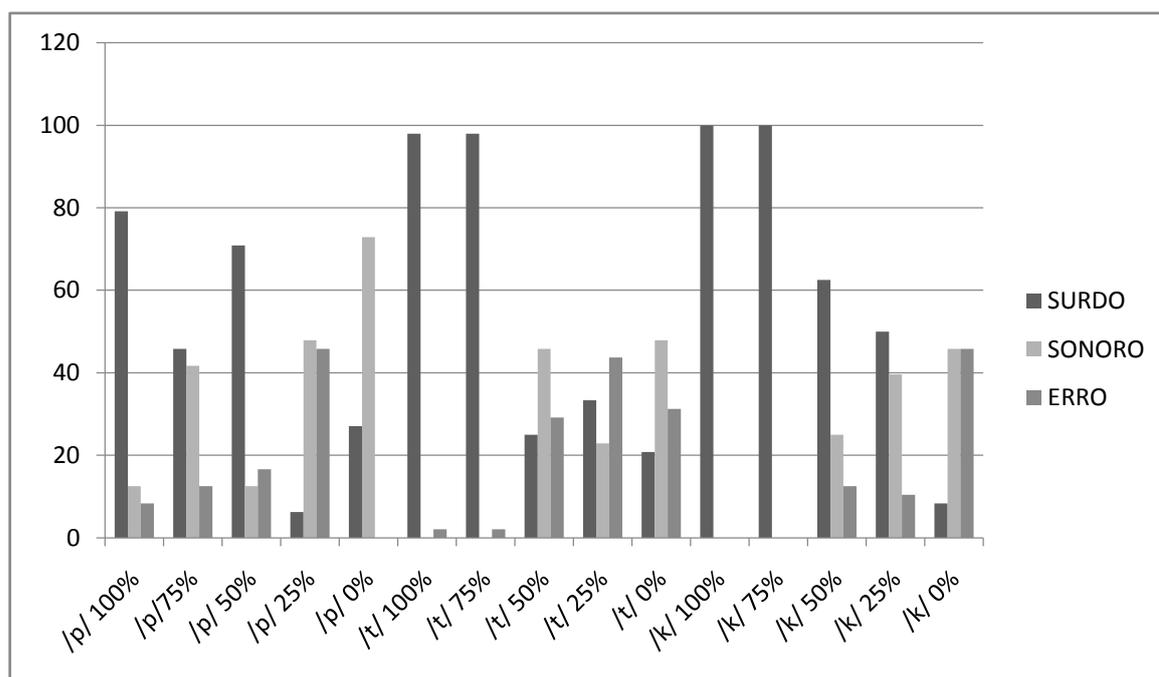
Os testes estatísticos corroboram a afirmação do parágrafo anterior sobre a particularidade adquirida pela atribuição do *erro* para a consoante velar: quando comparados entre si todos os graus de manipulação da consoante /k/ para os brasileiros, em cada uma das respostas tomadas isoladamente (surdo, sonoro ou erro), apenas o índice de *erro* apresentou resultado significativo, obtido através de um teste Friedman: ( $X^2(2) = 5.500, p = 0.002$ ). Em outras palavras, no que diz respeito à atribuição aos *status* de surdo ou de sonoro, os diferentes graus de manipulação não implicaram diferenças significativas para os aprendizes brasileiros, ao contrário do que se verificou com os americanos. Conforme já afirmado, diferenças significativas foram encontradas, apenas, entre os diferentes índices de erro atribuídos em cada grau de manipulação.

Para se saber com precisão entre que graus de manipulação essa diferença significativa, referentes aos índices de erro, se fez presente, todas as porções de manipulação foram, então, comparadas entre si, de forma pareada, a partir de testes ad hoc de Wilcoxon, com correção de Bonferroni. Parece haver, novamente, alguma relação entre significância estatística e grau mínimo de manipulação de 50%. Para o grupo de brasileiros deste estudo, três comparações feitas a partir do teste de Wilcoxon apresentaram valores significativos para *erro*. O padrão zero artificial apareceu em duas dessas comparações: uma delas quando comparado ao estímulo genuíno, sem aspiração (0 artificial x 100% de aspiração mantida:  $Z = -2,899; p = 0.001$ ) e, também, quando comparado à primeira fatia de manipulação, que mantém 75% da aspiração preservada (0 artificial x 75% de aspiração mantida:  $Z = -2,534; p = 0.002$ ). Com uma diferença marginalmente significativa, o limiar de 50% de aspiração também configurou em valores significativos para o grupo de brasileiros, agora apenas em uma comparação, baseada nas respostas que obtiveram atribuição a *erro*, quando comparado à aspiração original de 100% preservada (50% de aspiração mantida x 100% de aspiração mantida:  $Z = -2,430; p = 0.004$ ).

É preciso atentar para a característica especial da consoante velar, que se mostrou, de certa forma, como a mais vulnerável às diferenças significativas em função do grau de manipulação em ambos os grupos de participantes, sejam eles os americanos (que obtiveram valores significativos em todos os índices de resposta – *surdo*, *sonoro* e *erro*) ou sejam eles os brasileiros (que apresentaram valores significativos apenas no índice de *erro*). No que diz respeito às diferenças significativas referentes às respostas com erro na identificação de ponto de articulação em função dos diferentes graus de manipulação, chama a atenção o fato de que os brasileiros, ao contrário dos americanos somente apresentaram diferenças significativas no que se refere às respostas "erradas", e não no que diz respeito à atribuição de vozeamento. Já que os brasileiros se mostram mais sensíveis ao erro, e não à troca de padrão de vozeamento, é possível que a explicação para tal fato seja encontrada na transição formântica das vogais adjacentes à consoante. Como já explicado no capítulo sobre o Referencial Teórico, o VOT tem seu valor acentuado pela presença de vogais altas que o seguem, e, por isso, as palavras-alvo deste estudo têm, seguidas das consoantes, sempre vogais altas. As vogais altas, por sua vez, apresentariam uma influência considerável sobre as transições formânticas, o que poderia, por exemplo, causar enganos por parte dos participantes em relação ao /k/, que seria, então, identificado como um /p/. Somado a esse fato, outro fato que precisa ser mencionado diz respeito à questão das transições formânticas da vogal seguinte, transições essas responsáveis pela identificação dos pontos de articulação das consoantes plosivas. Nesse sentido, “o segmento [k] é bastante resistente em se coarticular com a vogal anterior [i], pois isso implicaria dificuldades de percepção devido ao fato de que pode ser confundido com as demais consoantes anteriores” (FOWLER, 1994; PEROZZO, 2013).

Como forma ilustrativa e para fins de comparação holística, segue, abaixo, um gráfico, com a comparação de todos os pontos de articulação e suas múltiplas manipulações, atribuídas de seus índices de identificação por parte dos informantes brasileiros.

GRÁFICO 2 – ÍNDICES DE IDENTIFICAÇÃO POR PONTO DE ARTICULAÇÃO BRASILEIROS



De forma geral, pode-se observar que, quando comparados entre si de forma descritiva, os dados dos informantes americanos e dos brasileiros não parecem seguir uma regularidade no que se refere à relação entre grau de manipulação e atribuição de *status* de vozeamento para tais estímulos. Observam-se, entretanto, algumas tendências, tais como o fato de o índice de erro por parte dos brasileiros ser menor do que os dos americanos, de forma geral. Isso poderia indicar que, de fato, a manipulação do VOT influencia a percepção dos brasileiros de maneira menos direta, devido ao fato dessa pista acústica ser mais relevante para os falantes de língua inglesa do que para os brasileiros, já que os americanos, da mesma forma que os brasileiros, quando deparados com manipulações extremas do VOT, ainda tentam enquadrar a identificação dessa consoante de alguma forma, mesmo que os padrões de vozeamento não sejam respeitados, causando, assim, troca do ponto de articulação. Os americanos tendem a trocar mais o ponto de articulação quando deparados à manipulação do VOT, pois essa é uma pista de mais importância para este grupo do que para o brasileiro, fazendo com haja identificações equivocadas, inclusive ao ponto de articulação. Isso não ocorre tão fortemente com brasileiros, uma vez que, quando a manipulação leva os aprendizes a deixarem de marcar as consoantes como surdas, há uma tendência em preferirem a resposta de “sonora”

do que trocar o ponto, uma vez que a manipulação dessa pista não causa tantos equívocos, muito provavelmente por essa não ser uma das pistas acústicas mais atentadas pelos brasileiros.

Além disso, e de modo mais importante, cabe mencionar o fato de que diferenças estatisticamente significativas no que diz respeito às atribuições de *stati* surdo e sonoro, em função das diferentes manipulações de VOT, foram encontradas apenas entre falantes americanos. Ainda que os dados tenham sugerido que, tanto para americanos quanto para brasileiros, muitas pistas acústicas parecem estar agindo em conjunto na atribuição do status de sonoridade das consoantes plosivas iniciais, julga-se importante ressaltar o fato de a diferença estatística ser encontrada apenas entre americanos, o que nos leva a sugerir que, de certa forma, conforme esperado e já sugerido por trabalhos anteriores (ALVES, SCHWARTZHAUPT & BARATZ, 2011), o VOT parece estar assumindo um caráter mais decisivo nas decisões de sonoridade entre americanos do que entre brasileiros.

#### **5.4 Descrição dos Dados de Brasileiros em relação à Proficiência**

Serão agora apresentados os dados referentes ao participantes brasileiros, porém atentando-se à subdivisão feita após a aplicação do teste de proficiência entre participante *intermediários* e *avançados*. Os dados serão apresentados a partir das consoantes analisadas e suas respectivas manipulações, comparadas com os níveis de atribuição de *status* de vozeamento por cada um dos níveis de proficiência.

Primeiramente, serão verificados os índices da consoante bilabial surda /p/ e suas manipulações, cujos valores podem ser consultados na tabela 8, logo a seguir. De maneira descritiva, é possível perceber que, embora as porcentagens possam parecer, de certa forma, não tão próximas, deve-se estar atento ao número de respostas de cada grupo sobre o total de tokens por atribuição de acerto (*surdo* e *sonoro*) e erro. A diferença que existe entre as porcentagens se deve ao número total de *tokens* por grupo, que muda devido ao diferente número de participantes existentes em cada um deles (7 avançados e 9 intermediários). Devido ao pequeno número de *tokens* resultantes dos testes, pequenas diferenças podem resultar em

margens um pouco maiores de porcentagem. Na tabela a seguir, apresentam-se os valores das porcentagens e dos tokens por grupo de participantes, de acordo com seu nível de proficiência:

TABELA 8 - Índices de Identificação para /k/ e suas Manipulações por Nível de Proficiência

Consoante e Aspiração Preservada	AVANÇADOS	INTERMEDIÁRIOS
/p/ 100% surdo	<b>80%</b> 24/30	<b>100%</b> 36/36
/p/ 100% sonoro	<b>6,66%</b> 2/30	<b>0%</b> 0/36
/p/ 100% erro	<b>13,33%</b> 4/30	<b>0%</b> 0/36
/p/ 75% surdo	<b>73,33%</b> 22/30	<b>77,77%</b> 28/36
/p/ 75% sonoro	<b>0%</b> 0/30	<b>0%</b> 0/36
/p/ 75% erro	<b>26,66%</b> 8/30	<b>22,22%</b> 8/36
/p/ 50% surdo	<b>33,33%</b> 10/30	<b>33,33%</b> 12/36
/p/ 50% sonoro	<b>33,33%</b> 10/30	<b>25%</b> 9/36
/p/ 50% erro	<b>33,33%</b> 10/30	<b>25%</b> 9/36
/p/ 25% surdo	<b>86,66%</b> 26/30	<b>94,44%</b> 34/36
/p/ 25% sonoro	<b>0%</b> 0/30	<b>0%</b> 0/36
/p/ 25% erro	<b>13,33%</b> 4/30	<b>5,55%</b> 2/36
/p/ 0art surdo	<b>33,33%</b> 10/30	<b>33,33%</b> 12/36
/p/ 0art sonoro	<b>26,66%</b> 8/30	<b>22,22%</b> 8/36
/p/ 0art erro	<b>40%</b> 12/30	<b>44,44%</b> 16/36

Para a verificação estatística acerca dos índices de *surdo*, *sonoro* e *erro* para a consoante plosiva surda /p/, testes Mann-Whitney foram rodados, para que se comparassem os índices para cada tipo de resposta entre os níveis de proficiência. Comparando-se os dois níveis no que diz respeito aos índices de *surdo*, *sonoro* e *erro*, tomados isoladamente para a consoante em questão, não foram encontradas

diferenças significativas entre os níveis de proficiência analisados: *surdo* (U = 50.000 p = 0.988), *sonoro* (U = 44.500 p = 0.877) e *erro* (U = 45.000 p = 0.614).

Ainda levando-se em consideração possíveis diferenças em função dos níveis de de proficiência dos aprendizes brasileiros, conforme apresentado pela tabela 9, logo a seguir, a tendência de similaridade de respostas atribuídas pelos participantes de ambos os níveis se mantém também para a consoante alveolar surda /t/. Novamente, infere-se, pela observação da tabela a seguir, que os grupos de proficiência, embora distintos, não apresentam comportamento substancialmente diferentes entre si:

TABELA 9 - Índices de Identificação para /t/ e suas Manipulações por Nível de Proficiência

Consoante e Aspiração Preservada	AVANÇADOS	INTERMEDIÁRIOS
/t/ 100% surdo	<b>100%</b> 30/30	<b>100%</b> 36/36
/t/ 100% sonoro	<b>0%</b> 0/30	<b>100%</b> 0/36
/t/ 100% erro	<b>0%</b> 0/30	<b>0%</b> 0/36
/t/ 75% surdo	<b>46,66%</b> 14/30	<b>50%</b> 18/36
/t/ 75% sonoro	<b>0%</b> 0/30	<b>1,11%</b> 4/36
/t/ 75% erro	<b>53,33%</b> 16/30	<b>38,88%</b> 14/36
/t/ 50% surdo	<b>53,33%</b> 16/30	<b>50%</b> 18/36
/t/ 50% sonoro	<b>13,33%</b> 4/30	<b>0%</b> 0/30
/t/ 50% erro	<b>33,33%</b> 10/30	<b>50%</b> 18/36
/t/ 25% surdo	<b>60%</b> 18/30	<b>47,22%</b> 17/36
/t/ 25% sonoro	<b>20%</b> 6/30	<b>19,4%</b> 7/36
/t/ 25% erro	<b>20%</b> 6/30	<b>33,33%</b> 12/36
/t/ 0art surdo	<b>53,33%</b> 16/30	<b>61,11%</b> 22/36
/t/ 0art sonoro	<b>6,66%</b> 2/30	<b>11,11%</b> 4/36
/t/ 0art erro	<b>40%</b> 12/30	<b>27,77%</b> 10/36

Novamente, testes estatísticos de Mann-Whiney foram aplicados a fins de comparação entre grupos de proficiência, para se verificarem possíveis diferenças estatisticamente significativas entre ambos os grupos, levando em conta os índices de resposta atribuídos a *surdo*, *sonoro* e *erro*, tomados individualmente. Da mesma forma ocorrida com /p/, a consoante /t/ não apresentou diferenças significativas ao ter seus índices de resposta comparados entre os níveis de proficiência pelo teste estatístico Mann-Whitney: *surdo* (U = 43.500 p = 0.180), *sonoro* (U = 43.550 p = 0.700) e *erro* (U = 45.000 p = 0.555).

Finalmente, apresenta-se a tabela 10, relativa aos dados de /k/ coletados com os participantes brasileiros, divididos pelo número de *tokens* de acordo com cada um dos níveis de proficiência (*avançados* e *intermediários*):

TABELA 10 - Índices de Identificação para /k/ e suas Manipulações por Nível de Proficiência

Consoante e Aspiração Preservada	AVANÇADOS	INTERMEDIÁRIOS
/k/ 100% surdo	<b>86,66%</b> 26/30	<b>100%</b> 36/36
/k/ 100% sonoro	<b>6,66%</b> 2/30	<b>0%</b> 0/36
/k/ 100% erro	<b>6,66%</b> 2/30	<b>0%</b> 0/36
/k/ 75% surdo	<b>86,66%</b> 26/30	<b>100%</b> 36/36
/k/ 75% sonoro	<b>0%</b> 0/30	<b>0%</b> 0/36
/k/ 75% erro	<b>6,66%</b> 2/30	<b>0%</b> 0/36
/k/ 50% surdo	<b>80%</b> 24/30	<b>77,77%</b> 28/36
/k/ 50% sonoro	<b>0%</b> 0/30	<b>0%</b> 0/36
/k/ 50% erro	<b>20%</b> 6/30	<b>22,22%</b> 8/36
/k/ 25% surdo	<b>53,33%</b> 16/30	<b>44,44%</b> 16/36
/k/ 25% sonoro	<b>20%</b> 6/30	<b>22,22%</b> 8/36
/k/ 25% erro	<b>26,66%</b> 8/30	<b>33,33%</b> 12/36
/k/ 0art surdo	<b>33,33%</b>	<b>27,77%</b>

	10/30	10/36
/k/ Oart sonoro	<b>53,33%</b> 16/30	<b>52,77%</b> 19/36
/k/ Oart erro	<b>6,66%</b> 2/30	<b>22,22%</b> 8/36

Conforme apresentado na tabela 10, os índices atribuídos a cada um dos graus de manipulação e suas possíveis atribuições de *status* de vozeamento parecem seguir a tendência das consoantes anteriormente comparadas entre os grupos, e parecem não apresentar diferenças significativas quando comparados os níveis de proficiência deste estudo. Testes estatísticos de Mann-Whitney compararam cada uma dessas atribuições (surdo, sonoro e erro) entre os grupos, e confirma a visível similaridade entre os grupos, que não apresentam comportamento diferenciado em nenhuma classificação, seja ela de *surdo* ( $U = 44.000$   $p = 0.888$ ), de *sonoro* ( $U = 45.000$   $p = 0.555$ ) ou de *erro* ( $U = 50.000$   $p = 0.400$ ).

Todas as comparações já apresentadas vão ao encontro dos resultados de estudos anteriores, como o de Alves e Motta (2013), que demonstram que, mesmo que aprendizes brasileiros tenham nível de proficiência elevado, eles ainda possuem um “ouvido” nativo do português brasileiro, ou seja, níveis mais elevados de proficiência em uma segunda língua não parecem, mesmo assim, tomar o VOT como pista prioritária para a distinção entre consoantes surdas e sonoras. O presente estudo não contou com dois grupos de proficiência bastante distantes entre si, como seria o caso se fossem apresentados dois grupos distintos, onde um deles fosse avançado e o outro fosse básico, por exemplo. Essa comparação entre extremos, porém, ainda não parece ser suficiente a ponto de se verificar comportamentos particularmente diferenciados entre si, já que essa comparação também não apresentou diferenças no estudo de Alves e Motta (2013), por exemplo.

## 5.5 Discussão dos Resultados

Como forma de organizar a presente sessão de Discussão dos Resultados, serão reproduzidas aqui as hipóteses propostas no início deste trabalho, com suas análises específicas.

Como hipótese primeira, que dizia respeito aos participantes americanos, inferiu-se que, em cada um dos pontos de articulação das consoantes analisadas neste estudo, seriam encontradas diferenças significativas para os índices de resposta possíveis ('surdo', 'sonoro' e 'erro'), já que, teoricamente, os americanos seguiriam, primordialmente, a pista acústica do VOT e, dessa forma, sendo esta pista manipulada, a percepção dos americanos seria influenciada pela manipulação:

**H1:** Considerando-se as respostas dos sujeitos americanos em cada um dos pontos de articulação individualmente, haverá diferenças significativas em cada um dos índices de resposta ("surda", "sonora" e "erro") a partir dos diferentes graus de manipulação do intervalo de VOT. **Hipótese parcialmente corroborada.**

A hipótese 1 foi parcialmente corroborada pois um teste de Friedman, que comparou todas as porções de VOT por ponto de articulação, indicou haver diferenças significativas em cada um dos índices de resposta (*surdo*, *sonoro* e *erro*) apenas para a consoante velar surda (*surdo*:  $X^2(2) = 7.500$ ,  $p = 0.003$ ; *sonoro*  $X^2(2) = 7.500$ ,  $p = 0.003$  e *erro*  $X^2(2) = 7.500$ ,  $p = 0.004$ ), mas não foram encontradas diferenças para as consoantes bilabial surda /p/ (*surdo*:  $X^2(2) = 6.200$ ,  $p = 0.260$ ; *sonoro*:  $X^2(2) = 5.200$ ,  $p = 0.460$ ; *erro*:  $X^2(2) = 2.200$ ,  $p = 0.360$ ) e alveolar surda /t/ (*surdo*:  $X^2(2) = 7.200$ ,  $p = 0.440$ ; *sonoro*:  $X^2(2) = 6.200$ ,  $p = 0.340$  e *erro*:  $X^2(2) = 7.400$ ,  $p = 0.540$ ).

Para este estudo, já que o intuito principal deste trabalho se volta para a tentativa de se encontrar um possível limiar fonético-fonológico para a identificação das consoantes analisadas de acordo com suas possíveis atribuições de *status* de vozeamento, após as diferenças significativas apontadas para /k/, realizou-se, então, uma nova rodada de testes estatísticos *ad-hoc*, que compararam todas as fatias de vozeamento entre si, a fim de se encontrar em quais comparações seriam apontadas diferenças significativas. Para esse segundo fim, aplicaram-se testes de comparação em pares Wilcoxon. Cabe mencionar, novamente, que estes testes, por seu caráter *ad hoc*, foram apenas rodados para as fatias de manipulação de /k/, pois foi a única consoante que apresentou valor estatisticamente significativo das três consoantes analisadas.

Os testes Wilcoxon realizados apresentaram com maior especificidade onde as diferenças de /k/ foram encontradas em cada uma das atribuições de vozeamento possíveis:

Nas comparações para /k/ considerado como *surdo* pelos participantes, as comparações se mostraram diferentemente significativas nos graus de manipulação apresentados abaixo:

- 0 artificial *versus* 50% de manipulação mantida ( $Z = -2,994$   $p = 0.003$ );
- 0 artificial *versus* 75% de manipulação mantida ( $Z = -3,213$   $p = 0.001$ );
- 0 artificial *versus* 100% de manipulação mantida ( $Z = -3,213$   $p = 0.001$ );
- 25% de manipulação mantida *versus* 75% de manipulação mantida ( $Z = -2,877$   $p = 0.004$ );
- 25% de manipulação mantida *versus* 100% de manipulação mantida ( $Z = -2,877$   $p = 0.004$ ).

Nas comparações para /k/ considerado como *sonoro* pelos participantes, as comparações se mostraram diferentemente significativas nos graus de manipulação apresentados abaixo:

- 0 artificial *versus* 50% de manipulação mantida ( $Z = -3,002$   $p = 0.003$ );
- 0 artificial *versus* 75% de manipulação mantida ( $Z = -3,002$   $p = 0.003$ );
- 0 artificial *versus* 100% de manipulação mantida ( $Z = -3,002$   $p = 0.003$ ).

E, finalmente, nas comparações para /k/ consideradas como *erro*, quando os participantes identificaram a consoante velar surda como outra que não [k] ou [g], as comparações se mostraram diferentemente significativas nos graus de manipulação apresentados abaixo:

- 0 artificial *versus* 75% de manipulação mantida ( $Z = -2,121$   $p = 0.003$ );
- 0 artificial *versus* 100% de manipulação mantida ( $Z = -2,121$   $p = 0.003$ );

Infere-se, pelos resultados apresentados acima, que se existe um limiar fonético que seja capaz de influenciar a percepção dos participantes no que se refere à pista acústica de VOT, este limiar parece estar por volta da metade da aspiração total para a consoante /k/. Supõe-se que a consoante velar surda apresente este *status* diferenciado para os americanos, diferentemente das outras

consoantes, pelos motivos já apresentados na seção 5.3, que remetem a características acústico-articulatórias pertencentes à natureza da consoante em questão.

A hipótese segunda versava sobre as diferenças possivelmente existentes nas respostas por parte do grupo de participantes brasileiros, quando comparados todos os índices de manipulação, em cada uma das atribuições de respostas possíveis ('surdo', 'sonoro', 'erro'):

**H2:** Considerando-se as respostas dos sujeitos brasileiros, em cada um dos pontos de articulação individualmente, não haverá diferenças significativas em cada um dos índices de resposta ('surda', 'sonora' e 'erro') a partir dos diferentes graus de manipulação do intervalo de VOT. **Hipótese parcialmente corroborada.**

Tal qual ocorreu na **H1**, a **H2** foi corroborada por apresentar diferença significativa no índice de *erro* da consoante velar surda /k/ ( $X^2(2) = 5.500$ ,  $p = 0.002$ ), porém parcialmente, por não apresentar valores significantemente estatísticos para os os índices de *surda* ( $X^2(2) = 7.500$ ,  $p = 0.320$ ) e *sonora* ( $X^2(2) = 3.200$ ,  $p = 0.440$ ) referentes a esse ponto (ao contrário do que havíamos visto nas atribuições de sonoridade para a velar por parte de americanos), e nem mesmo para as outras consoantes (/p/: *surdo*:  $X^2(2) = 7.200$ ,  $p = 0.310$  *sonoro*:  $X^2(2) = 5.200$ ,  $p = 0.460$  *erro*:  $X^2(2) = 6.400$ ,  $p = 0.320$  /t/: *surdo*:  $X^2(2) = 4.200$ ,  $p = 0.540$  *sonoro*:  $X^2(2) = 7.500$ ,  $p = 0.440$  *erro*:  $X^2(2) = 5.200$ ,  $p = 0.640$ ).

Novamente, tendo sido encontrado o valor significativo para o índice de *erro* para /k/ para os participantes brasileiros, após o teste de Friedman, partiu-se para a comparação entre todas as fatias de manipulação, a fim de se verificar a possibilidade de um limiar que deixaria os brasileiros mais suscetíveis ao *erro* referente ao ponto de articulação. Após feitas as comparações entre os índices obtidos pelos brasileiros, seguem, abaixo, os valores que se apresentaram estatisticamente significativos:

- Zero artificial *versus* 75% de manipulação mantida ( $Z = -2,534$   $p = 0.002$ );
- Zero artificial *versus* 100% de manipulação mantida ( $Z = -2,899$   $p = 0.001$ );

- 50% de manipulação mantida *versus* 100% de manipulação mantida ( $Z = -2,430$   $p = 0.004$ ).

Pode-se supor, pelos resultados estatísticos obtidos, que a chance de o grupo de brasileiros fazer uma escolha equivocada de ponto de articulação se mostra mais evidente quando há, no mínimo, uma diferença de aspiração de 50% em comparação com outra. Novamente, este limiar de 50% se mostra importante, agora no grupo de brasileiros. É importante atentar para o fato de que quanto menor a significância do resultado, menor também parece ser a distância entre os valores manipulados comparados entre si, já que os valores mais extremos comparados que obtiveram resultado estatisticamente significativo são a comparação entre o zero artificial e, respectivamente, 75% e 100% de aspiração mantida, enquanto o resultado de significância marginal reside numa comparação de valores mais estreitos.

Em suma, cabe ressaltar, mais uma vez, que, considerando as respostas 'surda' e 'sonora' por parte dos aprendizes brasileiros, independentemente do ponto de articulação, não houve diferenças significativas em função do grau de manipulação de VOT. Conforme demonstrado, a única diferença significativa verificada foi encontrada nos graus de erro. Tal fato, que se mostra diferente do que foi verificado para as respostas atribuídas pelos participantes para /k/, nos leva a sugerir que, conforme já apontado em estudos anteriores, a pista de VOT parece apresentar um caráter menos prioritário, para os brasileiros, nas distinções entre segmentos surdos e sonoros.

Considerando-se que o grupo de participantes brasileiros era heterogêneo, foi necessário aplicar um teste de nivelamento com os participantes, a fim de que se equiparassem informantes de níveis de proficiência diferentes para, então, podê-los comparar com os participantes americanos. A hipótese terceira tinha por objetivo verificar, portanto, se o nível de proficiência seria um fator determinante para a comparação intragrupo de participantes brasileiros:

**H3:** Considerando-se cada um dos pontos de articulação individualmente, não haverá diferenças significativas nos índices de resposta "surda", "sonora" e "erro" entre os dois níveis de proficiência dos aprendizes. **Hipótese corroborada.**

A **H3** foi considerada corroborada pois os testes Mann-Whitney, utilizados para se compararem os níveis entre si, não apresentaram valores significativos em nenhuma das comparações, conforme evidenciado abaixo:

- Bilabial – surda (U = 50.000 p = 0.988);
- Bilabial – sonora (U = 44.500 p = 0.877);
- Bilabial – erro (U = 45.000 p = 0.614);
- Alveolar – surda (U = 43.500 p = 0.180);
- Alveolar – sonora (U = 43.550 p = 0.700);
- Alveolar – erro (U = 50.500 p = 0.333);
- Velar – Surda (U = 44.000 p = 0.888);
- Velar – Sonora (U = 45.000 p = 0.555);
- Velar – Erro (U = 50.000 p = 0.400).

As comparações não significativas entre brasileiros demonstram que o nível de proficiência não parece ser um fator determinante no momento de identificar as consoantes com VOT manipulado. Poder-se-ia pensar que quanto maior o nível de proficiência em segunda língua, maior seria, também, a afinação de percepção perante a pistas acústicas a serem seguidas pelos falantes nativos da segunda língua, mas os resultados parecem estar de acordo com a teoria de Best e Tyler (2007), ou seja, pode-se inferir que os brasileiros ainda estejam adequando e identificando de maneira similar todas as consoantes da língua inglesa comparadas neste estudo, não se mostrando diferentes de acordo com o nível de proficiência. Um “ouvido” de falante brasileiro continua, apesar de seu nível de proficiência, um “ouvido” de falante brasileiro. Estes resultados também parecem ir ao encontro dos já apontados em estudos anteriores (cf. ALVES & MOTTA, 2013), em que foi demonstrado que, independentemente do nível de proficiência dos aprendizes brasileiros, o VOT ainda não se mostra como a pista acústica prioritariamente seguida para as distinções entre as plosivas surdas e sonoras do inglês.

## 6 Considerações Finais

Através deste estudo, propôs-se verificar a importância da pista acústica VOT nas distinções de sonoridade em plosivas iniciais da língua inglesa, por parte de americanos e brasileiros aprendizes de inglês. Sendo o VOT uma pista atribuída, em termos de duração, de maneira diferente em ambas as línguas, a verificação de como a sua manipulação afetaria dois grupos de L1s diferentes (onde cada um tem um sistema fonológico que reagiria de forma diferente perante ao *Voice Onset Time*), poderia contribuir para com os estudos de aquisição de segunda língua. Como apresentado no capítulo sobre o referencial teórico, o *Voice Onset Time* seria uma pista acústica primordial na identificação de consoantes surdas para os falantes nativos de língua inglesa, enquanto, provavelmente, para os brasileiros, essa pista seria a intensidade atribuída ao *burst*.

Estudos anteriores, já mencionados diversas vezes neste trabalho, mostraram que a manipulação de pistas acústicas (em especial, do VOT) pode ser um instrumento de testagem útil no que diz respeito à verificação do caráter prioritário ou não de pistas acústicas específicas nas distinções funcionais dos sistemas linguísticos. No caso do *Voice Onset Time*, é interessante a verificação de como a manipulação viria a influenciar essa percepção, pois a sua duração varia entre as línguas. Considerando-se o inglês, em que o VOT das consoantes se mostra bastante longo, cabe investigar como línguas que apresentam plosivas surdas com um padrão de VOT zero, como o português, fazem uso dessa pista acústica para a distinção entre segmentos surdos e sonoros da língua-alvo..

A intenção primeira deste trabalho foi verificar, a partir das “fatias” manipuladas de VOT, se haveria um limiar específico ou próximo a um possível ponto de corte de aspiração, valor esse a partir do qual os participantes viriam a modificar, prioritariamente, sua identificação dos segmentos como surdos ou sonoros. Apesar de este estudo não conseguir atestar qual seria, exatamente, este limiar, foi possível verificar que o *Voice Onset Time* ainda parece ser uma pista primordial para os americanos, já que este grupo não somente obteve graus de erro maiores perante à manipulação do que o grupo dos informantes brasileiros, mas,

também, apresentou, no caso de /k/ diferenças significativas nos graus de atribuição de sonoridade em função dos diferentes índices de manipulação dos sons.

Como apresentado nos gráficos e explicações ao longo do trabalho, pôde-se perceber que, mesmo que de maneira irregular ou não sistemática, a manipulação do VOT em diversas gradações ainda pode levar a futuros estudos mais aprofundados, dada a gama de possibilidades a serem testadas e refletidas sobre a questão da percepção auditiva de elementos manipulados. Por ora, deve-se mencionar que os resultados deste estudo confirmam o caráter dinâmico da linguagem, uma vez que a grande variabilidade nos dados reforçam não haver pontos plenamente fixos ou limiares para as distinções entre sons. Além disso, os dados também evidenciam que, tanto para americanos quanto para brasileiros, múltiplas pistas acústicas parecem estar envolvidas nas distinções entre os segmentos plosivos surdos e sonoros do inglês. Ainda que não se possa negar o papel de múltiplas pistas acústicas para ambos os grupos, deve-se mencionar, outrossim, que os dados parecem sugerir que, dentre estas múltiplas pistas, o VOT parece assumir um caráter mais primordial para as distinções funcionais surdo vs. sonoro para os americanos do que para os aprendizes brasileiros de inglês. .

É importante ressaltar, em relação aos participantes brasileiros e seus dois níveis de proficiência, que este estudo apresentou dados que não indicariam diferenças significativas referentes ao status da pista VOT em função do grau de adiantamento em inglês, já que, quando deparados com diversas manipulações acústicas, os brasileiros tendem a errar menos que os americanos, o que poderia levar à inferência de que mesmo que a manipulação de VOT atue, de alguma forma, na identificação auditiva desses informantes, essa percepção seria menos prejudicada do que no grupo de americanos.

Não se pode, contudo, ignorar a hipótese de que, havendo a possibilidade de múltiplas pistas acústicas atuarem em conjunto no momento da percepção auditiva (de ambos os grupos), exista a probabilidade de uma pista outra, que não o VOT, esteja assumindo o caráter de prioritário para as distinções entre segmentos plosivos surdos e sonoros. Futuros estudos, que isolem outras pistas acústicas dos segmentos plosivos, se fazem necessários frente à tarefa de verificação da

possibilidade de haver, para os aprendizes brasileiros de inglês, uma pista prioritária para as distinções de sonoridade

Este trabalho teve caráter inegavelmente embrionário, e, por isso, algumas falhas podem ser ainda corrigidas em estudos futuros. Um exemplo de limitação a ser apresentado se refere aos valores brutos dos VOTs genuínos: eles variam muito de acordo com a locutora, como em *pit*, por exemplo, em que a locutora 1 apresenta VOT positivo com valor de 73ms, enquanto a locutora 2 apresenta uma produção de VOT positivo com duração de 124ms para a mesma palavra, por exemplo (veja-se Tabela 1). A tentativa de aprimoramento desta limitação seria a maior aproximação possível de valores de produção de falantes nativos condizentes com os que são encontrados na literatura da área. Dado o número limitado de estímulos disponíveis para este estudo, estímulos esses, conforme já mencionado na Metodologia, "herdados" de experimentos anteriores, tal correção pretende ser feita em estudos futuros.

Além disso, como exemplo de aprimoramento para novas aplicações e reflexões do instrumento utilizado, propõem-se as seguintes: (a) aumento no número de informantes, para que os dados se tornem mais robustos e consistentes; (b) verificação da possibilidade de haver uma espécie de "erro sistemático", ou seja, verificação de quais consoantes estão sendo escolhidas pelos participantes quando há um índice de erro grande para uma consoante específica; (c) implementação de uma escala de *likert* aos testes propostos; (d) desenvolvimento de um instrumento de discriminação utilizando os mesmos estímulos, a fim de se cruzar os dados e ver se, de alguma forma, eles se complementariam.

Espera-se que os resultados do presente trabalho tenham apontado para o caráter importante do VOT nas distinções de sonoridade, de modo a demonstrar, também, o papel que outras pistas, associadas a essa, podem ter nas distinções funcionais da língua-alvo. Através da alta gama de variabilidade encontrada nos nossos dados, espera-se, também, ter evidenciado o caráter variável e dinâmico da linguagem, caráter esse que se mostra bastante claro ao se discutir as características desenvolvimentais de um novo sistema linguístico.

## Referências

ALVES e MOTTA: Focusing on the right cue: Perception of *Voiceless* and *Voiced* stops in English by Brazilian learners. Artigo aceito para publicação na revista *Phrasis*. Bélgica, 2013.

ALVES, SCHWARTZHAUPT e MOTTA: Percepção e Produção das Plosivas Iniciais do Inglês (L2) por Aprendizes Brasileiros: Verificação dos Efeitos do *Voice Onset Time*. Trabalho Apresentado no IX CELSUL, 2012

ALVES, SCHWARTZHAUPT, BARATZ: Percepção e produção dos padrões de VOT do inglês (L2) por aprendizes brasileiros. In: FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R.; KESKE-SOARES, M. Estudos em Aquisição Fonológica – Volume 4, 3-4. Pelotas: Editora e Gráfica Universitária da UFPel, 2011.

ALVES, ZIMMER: The dynamics of perception and production of VOT patterns in English by Brazilian learners. In: MELLO, E.; PETTORINO, M.; RASO, T, 2012.

ASHBY: *Understanding Phonetics: Understanding Language Series*. Oxford University Press, 2011.

BANDEIRA e ZIMMER: A Transferência dos Padrões de VOT de Plosivas Surdas no Multilinguismo. *Revista Letras de Hoje*. EDIPUCRS. Disponível em: <http://revistaseletronicas.pucrs.br/fale/ojs/index.php/fale/article/view/7503/6577>, 2011.

BEST e TYLER: Nonnative and second-language speech perception: Commonalities and complementarities. In: BOHN, Ocke-Schwen; MUNRO, Murray J. *Language Experience in Second Language Speech Learning: Studies in honor of James Emil Flege*: 13-34. Amsterdam: John Benjamins, 2007.

BOERSMA e WEENINK: Praat – Doing Phonetics by Computer - Version 5.2.9 Disponível em [www.praat.org](http://www.praat.org), 2011.

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

ELLIS: *Analyzing Learner Language*. Oxford: Oxford University Press, 2005.

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

FOWLER, C. Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception & Psychophysics*. V. 55, n. 6, 1994.

GUSSENHOVEN e JACOBS: *Understanding Phonology: Terceira Edição – Understanding Language Series*. Oxford University Press, 2011.

LADEFOGED e MADDIESON: *The Sounds of the World's Languages*. Oxford Cambridge, MA: Blackwell, 1996.

LISKER e ABRAMSON: A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, 384-422, 1964.

MONTEIRO: *Sistemas Dinâmicos*. Livraria da Física. Terceira Edição, 2011.

MUNRO e DERWING: *The Foundations of Accent and Intelligibility in Pronunciation Research*: Cambridge University Press, 2011.

PEROZZO, R. V. Percepção de oclusivas não vozeadas sem soltura audível em codas finais do inglês (L2) por brasileiros: o papel do context fonético-fonológico, da instrução explícita e do nível de proficiência. *Dissertação de Mestrado*. Universidade Federal do Rio Grande do Sul, 2013.

POLLITT: The meaning of OOPT Scores. Disponível em [www.oxfordenglishtesting.com](http://www.oxfordenglishtesting.com): 2007.

PORT e VAN GELDER (eds): *Mind as Motion: Explorations in the Dynamics of Cognition*. The MIT Press, 1995.

PURPURA: The Oxford Online Placement Test: What does it measure and how? Disponível em [www.oxfordenglishtesting.com](http://www.oxfordenglishtesting.com), 2007.

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

SCHWARTZHAUPT, ALVES, FONTES: O VOT como pista suficiente para a distinção surdo/sonoro: dados de falantes do inglês americano. Trabalho apresentado no IV Seminário de Aquisição Fonológica. Universidade Federal de Santa Maria, 2013.

SCHWARTZHAUPT: Factors influencing *Voice Onset Time*: analyzing Brazilian Portuguese, English and Interlanguage data. Monografia (graduação). Universidade Federal do Rio Grande do Sul, 2012.

VEIGA-FRANÇA, Karoline: A aquisição das plosivas surdas do inglês por falantes do português brasileiro: implicações teóricas decorrentes de duas formas de descrição de dados. Dissertação de Mestrado. Universidade Católica de Pelotas, 2011.

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

ZIMMER e ALVES, U. K. Uma visão dinâmica da produção da fala em L2: o caso da dessonorização terminal. *Revista da ABRALIN*, n. 2: esp., 2012.

ZIMMER e ALVES: On the Status of Terminal Devoicing as an Interlanguage Process among Brazilian Learners of English. *Ilha do Desterro*, v. 55, p. 41-62, 2008.

ZIMMER, SILVEIRA, ALVES: Pronunciation instruction for Brazilians: bringing theory and practice together. Newcastle upon Tyne: Cambridge Scholars Publishing, 2009.

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