

# IDENTIFICAÇÃO E DISCRIMINAÇÃO DE OCLUSIVAS NÃO VOZEADAS SEM SOLTURA DE AR AUDÍVEL DO INGLÊS: DADOS DE APRENDIZES BRASILEIROS

IDENTIFICACIÓN Y DISCRIMINACIÓN DE OCLUSIVAS SORDAS SIN SOLTURA DE AIRE AUDIBLE DEL INGLÉS: DATOS DE ESTUDIANTES BRASILEÑOS

IDENTIFICATION AND DISCRIMINATION OF ENGLISH UNRELEASED VOICELESS STOPS: DATA FROM BRAZILIAN LEARNERS

Reiner Vinicius Perozzo\*

Ubiratã Kickhöfel Alves\*\*

Universidade Federal do Rio Grande do Sul – UFRGS, Porto Alegre, BR

**RESUMO:** Neste estudo, investiga-se a percepção de aprendizes brasileiros, residentes em Porto Alegre (RS) e região metropolitana, acerca das diferenças de ponto de articulação entre as oclusivas não vozeadas [p ], [t ] e [k ], sem soltura de ar audível em final de palavra, presentes em inglês (variedade americana). Trinta e dois acadêmicos do curso de Letras da UFRGS participaram da pesquisa, os quais apresentaram níveis básico e intermediário de proficiência em inglês. Para verificar a acuidade perceptual em relação ao ponto de articulação das consoantes, designaram-se duas tarefas psicolinguísticas contendo palavras de estrutura silábica CVC, equidistribuídas segundo três vogais anteriores. Os resultados sugeriram que: (a) os segmentos [p ] e [k ] são percebidos com mais acuidade do que o segmento [t ]; (b) houve maiores índices de acuidade quando a vogal nuclear era curta e/ou frouxa; e (c) o nível de proficiência dos participantes não é determinante para a acuidade perceptual das consoantes testadas.

**PALAVRAS-CHAVE:** consoantes sem soltura; oclusivas; percepção.

**RESUMEN:** En este estudio se aborda la percepción de estudiantes brasileños, residentes en Porto Alegre y en la región metropolitana, acerca de las diferencias en los puntos de articulación entre las oclusivas sordas [p ], [t ] y [k ], sin soltura de aire audible en final de palabra, presentes en inglés (variedad norteamericana). Treinta y dos académicos del curso de Letras de la UFRGS, con niveles básico e intermedio de dominio del idioma inglés, participaron de la investigación. Para verificar la corrección perceptiva en relación al punto de articulación de las consonantes, se designaron dos tareas psicolinguísticas que contenían palabras de la estructura silábica CVC, distribuidas conforme a las tres vocales anteriores. Los resultados sugieren que: (a) los segmentos [p ] e [k ] son percibidos con más corrección que el segmento [t ]; (b) hubo mayores índices de corrección cuando la vocal nuclear era corta o débil; y (c) el nivel de dominio de inglés de los participantes no es determinante para la calidad perceptiva de las consonantes testeadas.

**PALABRAS-CLAVE:** consonantes sin soltura; oclusivas; percepción.

**ABSTRACT:** This study investigates how Brazilian EFL learners, living in Porto Alegre (RS), perceive the differences in place of articulation of the American English unreleased voiceless stops [p ], [t ], and [k ], in word-final position. Thirty-two undergraduate students (leveled as having basic and intermediate proficiency in English), taking their English major at UFRGS, participated in the research study. In order to verify the perceptual accuracy regarding the place of articulation of the consonants, two psycholinguistic tasks containing CVC words (whose nuclear segments were represented by front vowels) were conducted. The results suggest that: (a) the segments [p ] and [k ] are perceived much more accurately than the segment [t ]; (b) there were higher levels of accuracy in the tasks when the nuclear segment was a short/lax vowel; and (c) the participants' level of proficiency is not crucial to the perceptual accuracy of the consonants which were tested.

**KEYWORDS:** auditory perception; English stops; unreleased consonants.

## 1 INTRODUÇÃO

Conforme aponta a literatura, uma das realizações mais comuns no que tange às consoantes oclusivas em posição final de palavra em língua inglesa é a não soltura de ar audível desses segmentos (SELKIRK, 1982; YAVAŞ, 2006; CELCE-MURCIA *et al.*, 2010; DAVIDSON, 2011). Por se tratar de um fenômeno em que a pista acústica referente à soltura é suprimida, pressupõe-se que tal manifestação fonética possa impor dificuldades à percepção acurada de determinados segmentos na cadeia da fala, em termos de identificação e

\* UFRGS/CAPEL. Email: [linguistica.reiner@gmail.com](mailto:linguistica.reiner@gmail.com).

\*\* UFRGS/CNPq. Email: [ukalves@gmail.com](mailto:ukalves@gmail.com).

discriminação, no que compete a falantes brasileiros que estudam inglês como língua estrangeira. Dessa forma, se o segmento consonantal final das palavras *sick*, *sip* e *sit*, por exemplo, for realizado sem soltura audível, aprendizes brasileiros de inglês podem vir a não compreendê-las como perceptualmente distintas, uma vez que a não soltura da consoante oclusiva possivelmente minimize o entendimento do item lexical a que se refere, colocando a inteligibilidade das palavras em risco (ABRAMSON; TINGSABADH, 1999).

Atentando a tais considerações, este estudo tem o objetivo de verificar a acuidade na percepção (em termos de identificação e discriminação), por aprendizes brasileiros, acerca dos pontos de articulação de oclusivas não vozeadas, sem soltura de ar audível, presentes em posição final de palavras do inglês. Mais especificamente, a presente pesquisa visa a (a) examinar se existem maiores índices de acuidade sobre uma consoante específica em final de palavra; (b) determinar se a vogal precedente às consoantes em questão exerce efeito sobre a acuidade na percepção do ponto de articulação do segmento final; e (c) constatar se o nível de proficiência dos participantes tem impacto sobre a percepção dos segmentos-alvo.

## 2 NÃO SOLTURA AUDÍVEL: ARQUITETURA TEÓRICA DO FENÔMENO

Esta seção apresenta brevemente algumas considerações teóricas acerca do fenômeno da não soltura audível de consoantes oclusivas em posição final de palavra e/ou sílaba em língua inglesa, alicerçadas em um viés acústico e, por vezes, articulatorio.

Faz-se oportuno iniciar o assunto deixando claro que a informação sobre o ponto de articulação de um segmento oclusivo encontra-se nos gestos de fechamento e abertura dos articuladores e é acusticamente revelada, sobretudo, através de transições nas frequências formânticas que partem da vogal adjacente a esta consoante oclusiva, precisamente antes e depois do momento de oclusão e também no espectro da liberação de ar (ABRAMSON; TINGSABADH, 1999).

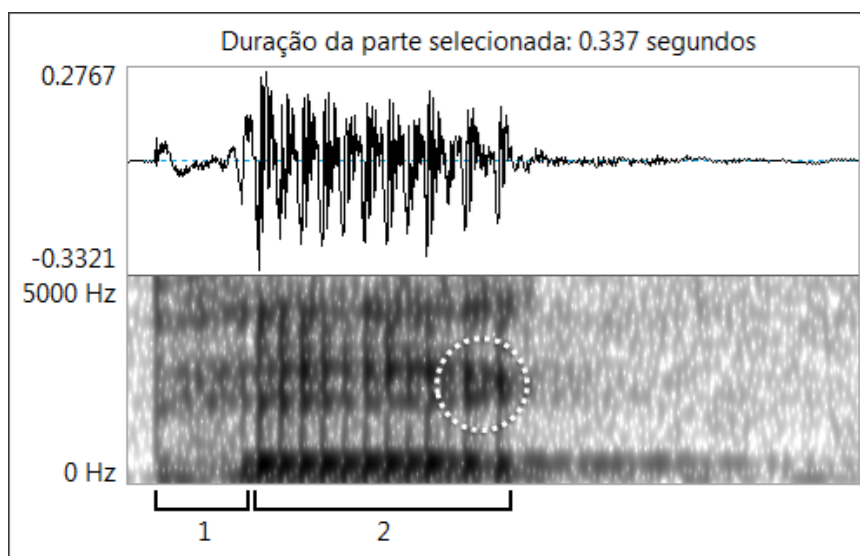
Segundo o que tem sido difundido na literatura, no inglês, as consoantes /p, b, t, d, k, g/ em posição de coda final podem ser variavelmente realizadas sem soltura de ar audível em contexto pré-pausal (SELKIRK, 1982; YAVAŞ, 2006; CELCE-MURCIA *et al.*, 2010; DAVIDSON, 2011). Tal manifestação fonética é desencadeada pela ausência de liberação do ar pulmonar egressivo após o gesto responsável pela oclusão do ponto de articulação da respectiva consoante. Desta forma, Lisker (1999) aponta que, quando há a não soltura do ar após a oclusão da consoante em coda final, a informação sobre o ponto de articulação do segmento consonantal é fornecida, principalmente, pelas alterações de frequência nos formantes da vogal precedente à consoante em coda.

Conforme estabelecem Ladefoged e Johnson (2011), a diferença perceptual em relação às palavras inglesas *rap*, *rat* e *rack*, quando produzidas sem soltura de ar audível, reside na maneira como as vogais terminam, ou seja, na qualidade vocálica, pois as consoantes que se aloca antes e depois das vogais, em função de processos coarticulatórios, exercem influência sobre a articulação de tais segmentos vocálicos.

Cabe acrescentar que, ainda que as transições formânticas que partem da vogal em direção à consoante em coda possam vir a ser uma das principais pistas para as distinções entre os pontos de articulação, existem evidências de que a não soltura audível de ar da consoante final possa reduzir a inteligibilidade e a percepção acurada das palavras que passam por tal processo até mesmo em casos de falantes nativos de inglês (LISKER, 1999). Neste quesito, Abramson e Tingsabadh (1999) explanam que a percepção do ponto de articulação da consoante apresenta maiores índices de acuidade quando as transições formânticas do gesto de fechamento são, por sua vez, complementadas pela informação acústica da soltura de ar.

A maneira como ocorre a não soltura de ar de uma oclusiva final pode ser facilmente visualizada na Figura 1, a qual representa a palavra inglesa “*pick*” (cuja consoante final foi realizada sem soltura audível) no plano do oscilograma e do espectrograma:

Figura 1 – Visualização da palavra “pick”, realizada por um falante nativo de inglês americano, sem soltura de ar audível da consoante em coda [k]



Fonte: adaptada de Perozzo (2013).

Conforme ilustra a Figura 1, (a) o espaço referente ao número 1 indica o momento de abertura dos articuladores (no caso, os lábios para a realização da oclusiva [p]) e o tempo de aspiração que precede a vogal nuclear; e (b) o espaço referente ao número 2 indica a vogal [ɪ] e, especialmente no final do espaço da vogal, é possível ver uma região de formantes circunscrita na cor branca, assinalando o encontro de F2 e F3 – característico da articulação dorsal (LADEFOGED, 2005), evidenciando a presença da oclusiva.

É importante ressaltar, na Figura 1, que a articulação da consoante [k] está presente, mas o que não se observa (e também não se ouve) é a liberação audível da corrente de ar da explosão (*burst*). Em outras palavras, os formantes da porção final da vogal nuclear, os quais se encontram em adjacência à consoante seguinte [k], indicam a presença de uma coarticulação dorsal. Tal constatação reforça os postulados de Lisker (1999) e de Ladefoged e Johnson (2011), que atribuem à vogal nuclear o papel de carregar a informação acerca do ponto de articulação de sua consoante seguinte sem soltura audível.

Esperar-se-ia questionar a real necessidade de haver soltura para a percepção de oclusivas finais, já que a vogal nuclear é capaz de veicular as informações acústicas sobre o ponto de articulação da consoante contígua. Todavia, faz-se pertinente atentar para o fato de que, conforme explicitado anteriormente, a percepção acurada do ponto de articulação da consoante oclusiva é favorecida, prioritariamente, quando as transições formânticas referentes ao gesto de fechamento são subsidiadas pela informação acústica de soltura de ar (ABRAMSON; TINGSABADH, 1999). Desta forma, neste estudo, observou-se como se dão os graus de identificação e discriminação, por parte de aprendizes brasileiros de inglês, quando a informação adicional da soltura não se faz disponível.

### 3 METODOLOGIA

A presente seção inicia contemplando as questões de pesquisa, passa a uma descrição geral dos participantes da pesquisa, relata brevemente a organização das palavras que serviram como estímulo auditivo e, por fim, explana a confecção dos testes aplicados e seus procedimentos.

#### 3.1 Questões de pesquisa

- **Questão 1:** Há maiores níveis de acuidade na percepção de um ponto de articulação específico (bilabial [p], alveolar [t], velar [k]) em final de palavra?
- **Questão 2:** A vogal nuclear exerce influência sobre a percepção das consoantes finais?

- **Questão 3:** Aprendizes brasileiros de inglês de nível de proficiência intermediário atingem maiores níveis de acuidade nas tarefas de percepção, quando comparados aos estudantes de nível de proficiência básico?

### 3.2 Participantes

O estudo contou com a participação de 32 acadêmicos do primeiro semestre do curso de Letras (Licenciatura e Bacharelado), matriculados na disciplina Inglês I, turmas A e B, da Universidade Federal do Rio Grande do Sul. Como medida de categorização dos acadêmicos e com o intuito de prover informações de caráter linguístico e extralinguístico ao tratamento dos dados, os participantes foram convidados a preencher um questionário biográfico e acadêmico com base em Rauber (2006), Nobre-Oliveira (2007), Alves (2008), Bettoni-Techio (2008), Kluge (2009) e no *Language Experience and Proficiency Questionnaire* (MARIAN; BLUMENFELD; KAUSHANSKAYA, 2007). Todos os acadêmicos assinaram um termo de consentimento, indicando estar de acordo com os propósitos da pesquisa e dela fazerem parte. Cada participante recebeu um número de identificação que serviu de senha para acesso aos testes durante a coleta de dados, bem como atendeu à finalidade de manter em sigilo suas identidades.

Para que se pudesse melhor caracterizar a amostra e verificar possíveis efeitos do nível de proficiência sobre o fenômeno fonético analisado, os aprendizes foram submetidos ao *Oxford Placement Test* (ALLAN, 2004), teste de nivelamento em língua inglesa, que apontou duas categorias, a saber, básico e intermediário. No Quadro 1, a seguir, apresentam-se, de maneira resumida, as informações relevantes acerca dos participantes da pesquisa:

Quadro 1 – Dados biográficos dos participantes

<i>Nº de participantes</i>	<i>Nível</i>	<i>Sexo</i>	<i>Idade</i>
32	<i>Básico: 11</i> <i>Intermediário: 21</i>	<i>Masculino: 15</i> <i>Feminino: 17</i>	<i>Média: 20,2</i> <i>DP: 4,3</i>

Fonte: elaborado pelo autor.

Nota: Sinais convencionais utilizados:

Nº – número

DP – desvio padrão

### 3.3 Estímulos auditivos

As palavras utilizadas como estímulos auditivos foram gravadas em um estúdio profissional por três falantes nativos de inglês americano, provenientes da região oeste dos EUA, de gênero masculino, com idade entre 20 e 21 anos. O microfone utilizado era unidirecional, da marca Violet e modelo Wedge, com faixa de frequência de gravação de 20 Hz a 20000 Hz, sendo que sua taxa de amostragem era de 44100 Hz.

Utilizou-se o *software Macmillan English Dictionary for Advanced Learners* (versão em CD-ROM) para selecionar as palavras que figuraram como estímulo auditivo na pesquisa. As palavras foram cuidadosamente selecionadas para que contivessem precisamente uma consoante em posição de *onset* e uma consoante em posição de *coda*, satisfazendo, assim, a estrutura silábica CVC. Os segmentos vocálicos contidos nas palavras foram equidistribuídos entre as vogais anteriores [i], [ɪ] e [æ]<sup>1</sup>.

Ainda no que tange à escolha dos segmentos consonantais, cabe salientar que a posição de *coda* silábica das palavras foi ocupada unicamente por consoantes oclusivas, a saber: oclusiva bilabial [p], oclusiva alveolar [t] e oclusiva velar [k]. Foram utilizadas oclusivas finais não vozeadas em virtude de que, de acordo com a ferramenta computacional utilizada, havia maior quantidade de itens lexicais terminados em oclusivas não vozeadas do que em oclusivas vozeadas. Tal escolha foi decisiva para que todos os segmentos que

<sup>1</sup>Inicialmente, o estudo iria contar tanto com vogais anteriores como com vogais posteriores para ocupar o núcleo das sílabas. No entanto, apenas uma vogal posterior, a saber, [ɑ] conseguiu satisfazer a condição de trio mínimo – dessa maneira, haveria três vogais anteriores ([i, ɪ, æ]) e uma vogal posterior ([ɑ]), a qual poderia, eventualmente, comprometer de alguma forma os dados, por ser posterior. Logo, como não houve palavras suficientes que possuíam vogais posteriores para formar os trios mínimos, optou-se por conduzir o estudo somente com os segmentos vocálicos anteriores, mantendo-se, assim, [i], [ɪ] e [æ].

compuseram os estímulos auditivos estivessem uniformemente distribuídos, ou seja, as mesmas consoantes em posição de *onset*, as mesmas vogais em posição nuclear e as mesmas consoantes em posição de coda (como em *beep*, *beat*, *beak*, formando um trio mínimo, por exemplo). Observe-se no Quadro 2, a seguir, a distribuição segmental das palavras selecionadas:

Quadro 2 – Distribuição segmental das palavras selecionadas.

	[p]	[t]	[k]
[i]	<i>beep</i>	<i>beat</i>	<i>beak</i>
	<i>weep</i>	<i>wheat</i>	<i>weak</i>
	<i>seep</i>	<i>seat</i>	<i>seek</i>
[ɪ]	<i>pip</i>	<i>pit</i>	<i>pick</i>
	<i>lip</i>	<i>lit</i>	<i>lick</i>
	<i>sip</i>	<i>sit</i>	<i>sick</i>
[æ]	<i>bap</i>	<i>bat</i>	<i>back</i>
	<i>map</i>	<i>mat</i>	<i>mac</i>
	<i>sap</i>	<i>sat</i>	<i>sack</i>

Fonte: Perozzo (2013).

### 3.4 Tarefas de percepção

Com o intuito de verificar a percepção de plosivas não vozeadas, sem soltura de ar audível, em posição de coda simples final presentes no inglês (L2) pelos aprendizes brasileiros, designaram-se duas tarefas psicolinguísticas: (a) Tarefa de identificação perceptual e (b) Tarefa de discriminação categórica, seguindo o formato ABX (LIBERMAN *et al.*, 1957).

As tarefas foram confeccionadas através do software Adobe Flash Professional CS5® (2012) e disponibilizadas via web. Durante a realização das tarefas, as respostas dos participantes eram armazenadas em uma planilha do Google Docs® (2012), onde apareciam codificadas, e posteriormente eram convertidas para o software Microsoft Excel® (2010).

Os aprendizes realizaram as tarefas em horário diurno no Laboratório de Línguas da Universidade, e as atividades foram executadas durante 25 minutos em média. A sala estava silenciosa e, durante a realização das questões, todos os participantes utilizavam fones de ouvido da marca Sony®, modelo MDR-XD100/B, com faixa de frequência de 12 Hz a 22000 Hz, de tiara ajustável para maior conforto.

Nas próximas subseções, as tarefas utilizadas na pesquisa são descritas em maiores detalhes.

#### 3.4.1 Tarefa de identificação perceptual (TIP)

A tarefa de identificação perceptual, doravante Tarefa A, contou com 81 questões e, correspondentemente, 81 palavras (27 *types* multiplicados por três repetições correspondentes aos três falantes nativos) que apareciam aleatoriamente para cada participante.

Os participantes não eram autorizados a ouvir o estímulo auditivo mais do que uma vez e tampouco houve tempo limite de resposta. Nesta tarefa, os alunos deveriam ouvir uma palavra e responder qual era a sua consoante final, ou seja, identificar o segmento, conforme as opções /p/, /t/ ou /k/.

Durante a programação do teste, tomou-se o cuidado de habilitar as opções de resposta somente depois de os participantes terem ouvido o estímulo, evitando, possivelmente, que eles marcassem qualquer opção de resposta sem terem, ao menos, prestado atenção ao que escutaram. Somente após selecionar a resposta julgada correta, eles poderiam responder à próxima questão. É oportuno destacar que, depois de selecionar

uma resposta e avançar para a palavra seguinte, não era possível retroceder e modificar a alternativa previamente eleita. Após a questão 40, havia um intervalo de 5 minutos para que os participantes pudessem descansar.

Os aprendizes levaram, em média, 5,51 min (DP=1,08 min) para concluir a realização da tarefa, desconsiderando o tempo de intervalo e possíveis pausas entre as questões.

### 3.4.2 Tarefa de discriminação categórica (TDC)

A tarefa de discriminação categórica, doravante Tarefa B, seguiu o formato ABX (LIBERMAN *et al.*, 1957), era constituída de 135 questões e, correspondentemente, 135 sequências de três palavras (tríades) que apareciam aleatoriamente aos participantes. Diferentemente da Tarefa A, em que cada estímulo auditivo corresponde a uma única palavra, na Tarefa B, cada estímulo auditivo corresponde a uma sequência de três palavras, divididas, neste estudo, em *conjuntos ternários* e *catch trials*. Em um conjunto ternário, duas palavras são iguais e uma é diferente (*beep-beak-beep*), enquanto que, em um *catch trial*, as três palavras são idênticas (*beep-beep-beep*).

As combinações de palavras possíveis para formar um conjunto ternário derivam de uma *tríade* (sequência de três palavras que compartilham o mesmo segmento em posição de *onset* e a mesma vogal nuclear, diferindo apenas quanto à consoante em posição de coda: *beep-beat-beak*), a qual designa quais são as palavras que irão compor seus respectivos arranjos. A título de exemplificação, o raciocínio para a confecção dos estímulos da Tarefa B acontece da seguinte maneira: (a) cada tríade, por exemplo, *beep-beat-beak*, pode dar origem a: (i) 12 combinações diferentes de palavras (conjuntos ternários): *beep-beat-beep*; *beep-beat-beat*; *beep-beak-beep*; *beep-beak-beak*; *beat-beep-beep*; *beat-beep-beat*; *beat-beak-beat*; *beat-beak-beak*; *beak-beep-beep*; *beak-beep-beak*; *beak-beat-beat*; *beak-beat-beak*; e (ii) 3 combinações iguais de palavras (*catch trials*): *beep-beep-beep*; *beat-beat-beat*; *beak-beak-beak*; (b) multiplicando o número de combinações diferentes que cada tríade pode ter pelo número de tríades totais, a saber, 9, resultam 108 combinações diferentes – ou seja, 108 conjuntos ternários; (c) multiplicando o número de combinações iguais que cada tríade pode ter (3) pelo número de tríades totais, a saber, 9, resultam 27 combinações iguais – isto é, 27 *catch trials*; e (d) somando-se os 108 conjuntos ternários aos 27 *catch trials*, resulta um total de 135 sequências de palavras, as quais constituíram a tarefa B.

A tarefa B requiritava que os participantes respondessem se a consoante final da última palavra da sequência era (a) igual à consoante final da primeira palavra; (b) igual à consoante final da segunda palavra; ou (c) igual à consoante final das demais palavras. O intervalo de tempo entre cada palavra em um conjunto ternário/ *catch trial* era de 500 ms.

Após as questões 45 e 90, os alunos poderiam fazer um intervalo de 5 minutos (caso preferissem), garantindo dois intervalos à tarefa por se tratar de um número maior de questões. Para concluir esta etapa da testagem, os participantes demoraram, em média, 12,58 min (DP<sup>2</sup>=1,18 min) – desconsiderando o tempo de intervalo sugerido e possíveis pausas aleatórias que eles possam ter feito, pois esse registro indica a duração dos estímulos auditivos e o tempo que os participantes levaram para se dirigir à questão seguinte.

## 4 RESULTADOS E DISCUSSÃO

Esta seção dedica-se à análise dos dados, realizada através do *software Statistical Package for Social Studies*<sup>®</sup> [SPSS (2009), versão 18], e à discussão dos resultados obtidos no presente estudo. A análise será dividida em três partes, de modo a contemplar o escopo de cada questão de pesquisa, previamente apresentadas na Seção 3.1.

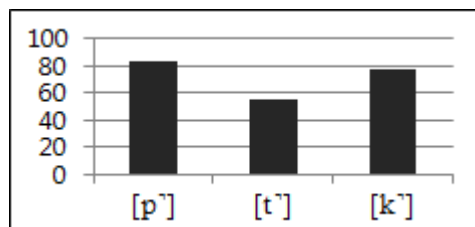
<sup>2</sup> DP=desvio padrão.

#### 4.1 Percepção de um ponto de articulação específico

A percepção que os aprendizes tiveram acerca de um ponto específico de articulação foi verificada em termos de identificação e de discriminação, correspondentemente através das Tarefas A e B.

Na Tarefa A, os níveis de acuidade foram de 83%<sup>3</sup> para [p], 55% para [t] e 77% para [k], conforme apresenta o Gráfico 1.

Gráfico 1 – Percentual de acuidade na identificação das consoantes finais sem soltura de ar audível



Fonte: elaborado pelo autor.

O teste de Friedman realizado indicou haver diferenças estatisticamente significativas, no que compete à acuidade perceptual, entre as três consoantes ( $\chi^2(2)=43,216$ ;  $p<0,001$ ). Testes de Wilcoxon, com correção de Bonferroni, sugeriram diferenças estatisticamente significativas entre [p] e [t] ( $Z=-4,868$ ;  $p<0,001$ ) e entre [t] e [k] ( $Z=-4,899$ ;  $p<0,001$ ).

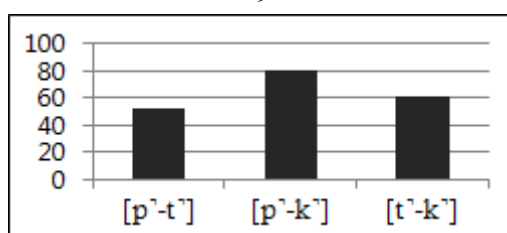
De acordo com tais resultados, parece que os aprendizes têm maior facilidade de identificar [p] e [k], enquanto a identificação de [t] ainda permanece mais difícil. Tal constatação pode se dar em função de que maiores transições formânticas carregam informações acústicas mais robustas (HUME, 1999), o que faria com que as consoantes labiais e velares tivessem preponderância sobre as coronais, uma vez que seus gestos demoram mais tempo para serem formados.

Jun (1995) afirma que as transições formânticas de [t] são caracterizadas como “fracas” se comparadas àquelas dos segmentos [p] e [k]. Além disso, o gesto da língua que daria conta da produção do segmento [t] é bastante rápido, implicando curtas pistas de transições formânticas (Hume *et al.*, 1999).

Com base em suas pesquisas relacionadas a pontos de articulação através de tecnologia eletropalatográfica, Reis e Espesser (2006) deixam claro que as regiões alveolar e velar se encontram em espaços completamente diferentes. Incorporando a região labial à análise, pode-se estimar que tal postulado dê origem à hipótese articulatória de que, quanto maior a distância entre os pontos de articulação, maiores os índices de acuidade na percepção dos pontos de articulação. Desta forma, seria mais fácil perceber diferenças acústico-articulatórias entre os segmentos [p] e [k] do que entre os segmentos [p] e [t], ou mesmo, [t] e [k]. Neste sentido, Hume *et al.* (1999), ao se referir aos estudos de Jun (1995), Côte (1997), Boersma (1998) e Hume (1999), reforça que maiores saliências perceptuais são desencadeadas por pistas acústicas mais robustas, baseadas em informações articulatórias mais densas.

Na Tarefa B, os níveis de acuidade na discriminação das consoantes finais foram de 52% para conjuntos ternários compostos por [p - t], 80% para aqueles compostos por [p - k] e 62% para aqueles compostos por [t - k], como evidencia o Gráfico 2.

Gráfico 2 – Percentual de acuidade na discriminação das consoantes finais sem soltura de ar audível



Fonte: elaborado pelo autor.

<sup>3</sup> Em todos os gráficos apresentados, o eixo y representa a mediana percentual de acuidade na tarefa em questão.

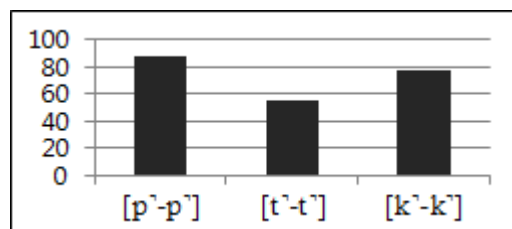
O teste de Friedman realizado indicou haver diferenças estatisticamente significativas entre as consoantes finais pertencentes aos conjuntos ternários ( $\chi^2(2)=48,778$ ;  $p<0,001$ ). Testes de Wilcoxon, com correção de Bonferroni, sugeriram diferenças estatisticamente significativas entre as palavras derivadas dos três conjuntos ternários, respectivamente entre os contrastes [p -t ] e [p -k ] ( $Z=-4,915$ ;  $p<0,001$ ), [p -t ] e [t -k ] ( $Z=-4,801$ ;  $p<0,001$ ), e também entre [t -k ] e [p -k ] ( $Z=-4,663$ ;  $p<0,001$ ).

Novamente, os índices de acuidade na Tarefa B não foram suficientemente altos naqueles estímulos que apresentavam a oclusiva alveolar sem soltura juntamente com outra oclusiva. Parece razoável que aquela caracterização acústica e articulatória do segmento [t], explicada previamente, se mantém no tocante à discriminação. Sob este enfoque, é cabível retomar a assunção de Jun (1995), em que o autor propõe um ranking universal de saliência perceptual para oclusivas sem soltura, de acordo com o qual as oclusivas dorsais dominam as labiais que, por sua vez, dominam as coronais.

Ao se verificarem os dados acima expressos, o Modelo de Assimilação Perceptual voltado para segunda língua, proposto por Best e Tyler (2007) oferece uma explicação que pode vir a dar conta dos baixos índices de acuidade em relação ao segmento [t]. Os autores estabelecem que os sons-alvo mais prováveis de ser adquiridos são aqueles foneticamente mais distantes dos da L1 ou das categorias de sons da L2 já fortemente estabelecidos. Assim, se comparada aos segmentos [p] e [k], a oclusiva alveolar em uma tarefa de discriminação representaria uma consoante de localização intermediária no trato vocal, tornando sua percepção mais difícil. Considerando-se este argumento, pode ter sido o caso de os aprendizes ainda não estarem aptos a estabelecer uma categoria fonológica individual para [t], dada sua instanciação articulatória intermediária e, como resultado, este segmento poderia ser percebido ora como [p] e ora como [k].

No que concerne aos *catch trials* (seqüências de três palavras com consoantes finais idênticas) na Tarefa B, os níveis de discriminação dos aprendizes foram de 88% para [p -p ], 55% para [t -t ] e 77% para [k -k ], como demonstrado no Gráfico 3.

Gráfico 3 – Percentual de acuidade na discriminação das consoantes finais idênticas sem soltura de ar audível



Fonte: elaborado pelo autor.

O teste de Friedman realizado indicou diferenças estatisticamente significativas entre as três categorias de consoantes ( $\chi^2(2)=39,982$ ;  $p<0,001$ ). Testes de Wilcoxon, com correção de Bonferroni, sugeriram diferenças estatisticamente significativas entre os estímulos terminados em [p -p ] e [t -t ] ( $Z=-4,564$ ;  $p<0,001$ ) e entre os estímulos terminados em [t -t ] e [k -k ] ( $Z=-4,567$ ;  $p<0,001$ ).

Em consonância com os resultados dos outros testes, aqueles estímulos contendo oclusivas alveolares são os que atingem níveis mais baixos de acuidade. Mais uma vez, parece que as propriedades gestuais de [t] impõem algum tipo de dificuldade perceptual em relação aos julgamentos dos participantes. Logo, fatores como (a) pouca informação formântica, (b) rapidez na produção do gesto e (c) ponto de articulação intermediário ainda são levados em conta para explicar a baixa acuidade perceptual de [t]. Por outro lado, os estímulos contendo oclusivas labiais e velares parecem ser bem perceptíveis, considerando sua informação acústica robusta, seu maior tempo de produção e, também, suas transições formânticas mais agudas. Tais resultados referentes aos *catch trials* são, de certa forma, surpreendentes pois, dado seu objetivo, os aprendizes deveriam ter atingido níveis altíssimos de acuidade. Isso pode revelar a real dificuldade perceptual que os aprendizes encontraram em discriminar entre as consoantes finais, mesmo sendo idênticas.



#### 4.2 O papel da vogal nuclear na percepção das consoantes finais

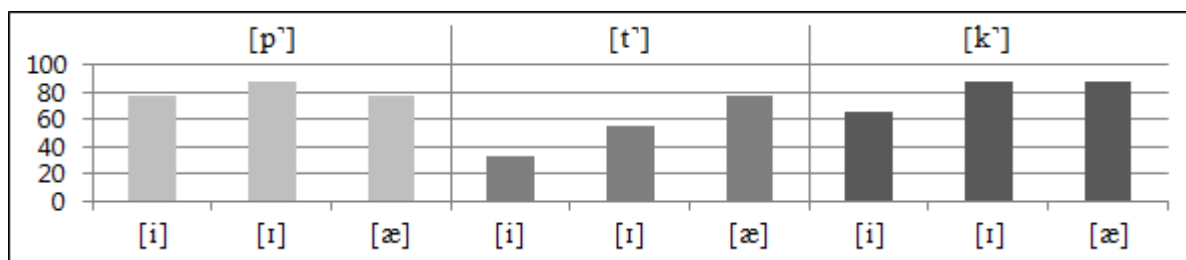
A influência da vogal nuclear sobre a percepção auditiva dos aprendizes em relação ao ponto de articulação das consoantes finais foi medida no que diz respeito tanto à identificação (Tarefa A) como à discriminação (Tarefa B).

Na Tarefa A, quando as palavras que terminavam em [p] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 77% de acuidade na discriminação da consoante final. Quando [I] figurava como vogal nuclear precedendo tal consoante, houve 88% de acuidade perceptual. Ainda, quando a vogal [æ] precedia [p], os níveis de acuidade perceptual na identificação da consoante final chegaram a 77%.

Quando as palavras que terminavam em [t] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 33% de acuidade na discriminação da consoante final. Quando [I] figurava como vogal nuclear precedendo a mesma consoante, houve 55% de acuidade perceptual. Ainda, quando a vogal [æ] precedia [t], os níveis de acuidade perceptual na identificação da consoante final chegaram a 77%.

Quando as palavras que terminavam em [k] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 66% de acuidade na discriminação da consoante final. Quando [I] figurava como vogal nuclear precedendo a mesma consoante, houve 88% de acuidade perceptual. Ainda, quando a vogal [æ] precedia [k], os níveis de acuidade perceptual na identificação da consoante final chegaram, também, a 88%. Tais resultados são mostrados no Gráfico 4.

Gráfico 4 – Percentual de acuidade na identificação das consoantes finais sem soltura de acordo com a vogal nuclear



Fonte: elaborado pelo autor.

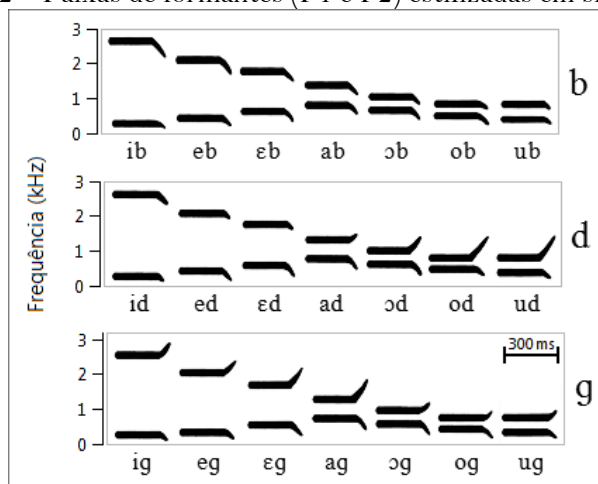
O teste de Friedman realizado indicou haver diferenças estatisticamente significativas entre as vogais precedendo [p] ( $\chi^2(2)=26,636$ ;  $p<0,001$ ). Testes de Wilcoxon, com correção de Bonferroni, sugeriram haver diferenças estatisticamente significativas entre as vogais [i] e [I] ( $Z=-3,426$ ;  $p=0,001$ ) e entre as vogais [I] e [æ] ( $Z=-4,224$ ;  $p<0,001$ ). Diferenças significativas também foram encontradas entre as três vogais precedendo [t] ( $\chi^2(2)=41,450$ ;  $p<0,001$ ), que foram verificadas nos testes de Wilcoxon (com correção de Bonferroni), correspondendo a [i] e [I] ( $Z=-4,104$ ;  $p<0,001$ ), [i] e [æ] ( $Z=-4,819$ ;  $p<0,001$ ), e também a [I] e [æ] ( $Z=-3,535$ ;  $p<0,001$ ). Além disso, diferenças significativas foram encontradas entre as vogais precedendo [k] ( $\chi^2(2)=23,145$ ;  $p<0,001$ ), que foram verificadas através de teste de Wilcoxon, correspondendo a [i] e [I] ( $Z=-3,895$ ;  $p<0,001$ ) e também a [I] e [æ] ( $Z=-3,939$ ;  $p<0,001$ ).

Diversas hipóteses podem ser encontradas, na literatura, na tentativa de explicar a relação entre a vogal nuclear e a consoante adjacente em termos de percepção. Householder (1956) defende que vogais frouxas propiciam a percepção acurada da consoante final. Lisker (1999) entende que os índices de acuidade na percepção de uma oclusiva final estejam diretamente relacionados ao status de ditongo ou monotongo do segmento nuclear. Para o pesquisador, os ditongos viriam a implicar menores índices de acuidade à percepção das oclusivas finais. Uma hipótese fonológica, proposta por van Oostendorp (1995) e Kang (2003), presume que as vogais frouxas propiciam maior acuidade perceptual ao segmento final porque as vogais frouxas em inglês necessitam de uma consoante em posição de coda.

Entretanto, a hipótese aqui adotada para dar conta do fenômeno foi proposta por Perozzo (2013) e se sustenta sobre o aspecto da transição formântica existente entre a vogal nuclear e a consoante final. Segundo esta hipótese, tem-se que: quando as faixas de formantes que partem da vogal em direção à consoante final apresentarem a transição formântica *característica* dos pontos de articulação labial, coronal e

dorsal, haverá maiores índices de acuidade quanto à percepção de tais pontos. Esta hipótese foi baseada na descrição das transições formânticas características de certos segmentos, definidas por Ladefoged (2005), e presume que a percepção acurada do ponto de articulação de uma consoante oclusiva final esteja relacionada à fidelidade máxima à transição formântica que ocorre da vogal nuclear em direção à consoante em coda. Portanto, quanto mais próxima for uma transição formântica do padrão esperado (característico), mais acurada será a percepção do segmento final em jogo. Para que o leitor possa compreender melhor o movimento das transições formânticas, apresenta-se, a seguir, um esquema (Figura 2) das faixas de formantes (vogal-consoante), adaptados de Delattre, Liberman e Cooper (1955):

Figura 2 – Faixas de formantes (F1 e F2) estilizadas em sílabas VC



Fonte: adaptada de Delattre, Liberman e Cooper (1955).

Um detalhe importante a ser considerado em relação a esta hipótese é a natureza da vogal, que age de maneira consistente sobre as transições formânticas. Além disso, de acordo com a hipótese aqui descrita, quando o respeito máximo às transições formânticas características que partem da vogal nuclear para a consoante em coda for insuficiente para a percepção acurada das consoantes finais, então será o caráter curto da vogal nuclear que irá garantir a acuidade quanto à percepção do ponto de articulação dos segmentos finais. Este “adendo” à hipótese apresentada em Perozzo (2013) surge a partir das constatações de Householder (1956) e de Lisker (1999). De um lado, Householder (1956) aponta as vogais frouxas como otimizadoras da percepção da consoante final; de outro, Lisker (1999) afirma que é o status de monotongo do segmento nuclear que irá propiciar maior acuidade perceptual à consoante final. Comparando os dois olhares (segmentos frouxos versus monotongais), a pista de que ambas as vogais compartilham é a *duração* – as vogais frouxas duram menos do que as vogais tensas e os monotongos duram menos do que os ditongos<sup>4</sup>. Daí o motivo pelo qual se propôs ser o caráter *curto* da vogal nuclear a garantir a acuidade perceptual das consoantes finais.

Considerando-se a hipótese previamente descrita, no que diz respeito ao segmento [p], quanto mais característica for a transição formântica da vogal para a consoante, maior será a acuidade na percepção, implicando diferenças significativas entre os segmentos vocálicos. Desse modo, tanto [ɪ] como [i] poderiam ser candidatos a facilitar a percepção das consoantes finais. A vogal [æ] não seria uma escolha potencial por apresentar uma transição formântica menos característica em direção à consoante em coda. Porém, entre [ɪ] e [i], dá-se preferência à vogal [ɪ] em função de ela apresentar menor duração.

No que tange ao segmento [t], de acordo com Ladefoged (2005), seu F2 caracteriza-se por movimentos sutis, portanto, espera-se que maiores índices de acuidade sejam encontrados quando a vogal precedente à consoante em coda apresentar menores transições formânticas, no caso, [æ], vogal que, dentre as anteriores, possui F2 mais baixo (o que foi o caso nos resultados deste estudo). Depois de se eleger a vogal [æ] como a de maior contribuição para a acuidade de [t], restam as duas outras vogais, [ɪ] e [i], cujo tamanho da transição formântica é semelhante (DELATTRE; LIBERMAN; COOPER, 1955). Então, o desempate foi dado por aquela que apresenta menor duração, [ɪ], a segunda melhor colocada em termos de percepção (55% de acuidade).

<sup>4</sup> Assume-se, neste trabalho, que a vogal [æ] seja frouxa e curta, tal como estabelecem Celce-Murcia et al. (2012).

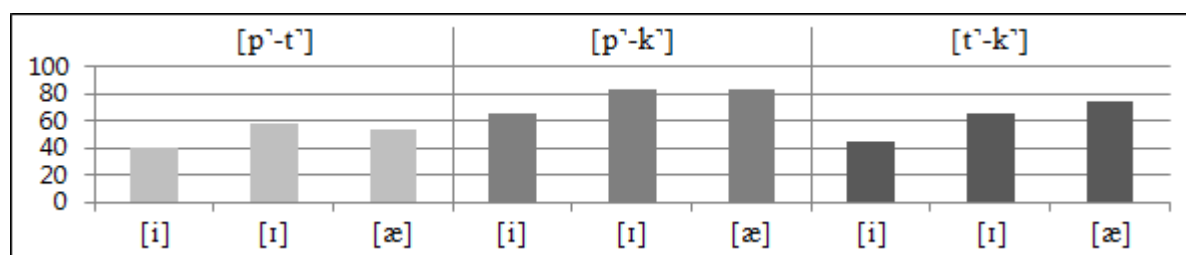
Em relação ao segmento [k], F2 tende a subir (LADEFOGED, 2005), portanto, espera-se que maiores índices de acuidade sejam encontrados quando a vogal precedente for [æ], uma vez que há maior fidelidade à transição formântica característica de [k]. Depois de se eleger a vogal [æ] como a de maior contribuição para a acuidade de [k], mais uma vez, restam as duas outras vogais, [ɪ] e [i], cuja transição formântica é muito semelhante (DELATTRE, LIBERMAN; COOPER, 1955). Portanto, o desempate poderá ser dado, novamente, por aquela vogal que apresentar menor duração, considerando-se, assim, que transição formântica e duração são pistas importantes para a percepção acurada da oclusiva sem soltura.

Na Tarefa B, quando as tríades que terminavam em [p -t] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 41% de acuidade na discriminação das consoantes finais. Quando [ɪ] era a vogal nuclear precedendo as mesmas consoantes, eles atingiram 58% de acuidade nas repostas. Além disso, quando a vogal [æ] precedia [p -t], os níveis de acuidade perceptual dos aprendizes foi de 54% na discriminação das consoantes finais.

Quando as tríades que terminavam em [p -k] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 66% de acuidade na discriminação das consoantes finais. Quando [ɪ] era a vogal nuclear precedendo as mesmas consoantes, eles atingiram 83% de acuidade nas repostas. Além disso, quando a vogal [æ] precedia [p -k], os níveis de acuidade perceptual dos aprendizes também foi de 83% na discriminação das consoantes finais.

Quando as tríades que terminavam em [t -k] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 45% de acuidade na discriminação das consoantes finais. Quando [ɪ] era a vogal nuclear precedendo as mesmas consoantes, eles atingiram 66% de acuidade nas repostas. Além disso, quando a vogal [æ] precedia [t -k], os níveis de acuidade perceptual dos aprendizes foi de 75% na discriminação das consoantes finais. O Gráfico 5 apresenta tais resultados.

Gráfico 5 – Percentual de acuidade na discriminação das consoantes finais sem soltura de ar audível de acordo com a vogal nuclear



Fonte: elaborado pelo autor.

O teste de Friedman realizado indicou haver diferenças estatisticamente significativas entre as vogais nos estímulos contendo as consoantes [p -t] ( $\chi^2(2)=12,932$ ;  $p=0,002$ ). Testes de Wilcoxon, com correção de Bonferroni, sugeriram haver diferenças significativas entre as vogais [i] e [ɪ] ( $Z=-3,100$ ;  $p=0,002$ ) e entre as vogais [i] e [æ] ( $Z=-3,006$ ;  $p=0,003$ ). Diferenças significativas também foram encontradas entre as vogais nos estímulos contendo as consoantes [p -k] ( $\chi^2(2)=24,268$ ;  $p<0,001$ ), que foram verificadas através de testes de Wilcoxon (com correção de Bonferroni), correspondendo às vogais [i] e [ɪ] ( $Z=-3,980$ ;  $p<0,001$ ) e também às vogais [i] e [æ] ( $Z=-3,824$ ;  $p<0,001$ ). Ainda, diferenças significativas foram encontradas entre as vogais nos estímulos contendo as consoantes [t -k] ( $\chi^2(2)=26,914$ ;  $p<0,001$ ), que foram verificadas através de testes de Wilcoxon (com correção de Bonferroni), correspondendo às vogais [i] e [ɪ] ( $Z=-3,199$ ;  $p=0,001$ ) e também às vogais [i] e [æ] ( $Z=-4,468$ ;  $p<0,001$ ).

Dados os resultados estatísticos para as vogais precedentes às palavras terminadas em [p -t], pode-se explicar a preferência por [ɪ] em detrimento de [i] com base no quesito de duração – já que as transições formânticas destas duas vogais frente a [p] são muito semelhantes, o status curto da vogal [ɪ] lhe daria preferência à percepção do ponto de articulação da consoante final. Possivelmente, a duração também tenha sido o critério para que a vogal [ɪ] implicasse maior acuidade na percepção do ponto coronal. Além disso, os baixos índices de acuidade de [i] também frente a [t] podem ser explicados com base em Delattre, Liberman e Cooper (1955), os quais observam que, quanto mais anterior for a vogal precedente à consoante [t], mais este segmento se parece, em termos de F2, com um [p], e quanto mais posterior for a vogal

precedente, mais ele se parece com um [k]. Desse modo, poderia haver a perda de um contraste em função da vogal nuclear [i], implicando menores índices de acuidade nesta vogal.

Para a percepção acurada dos estímulos cujas palavras terminavam nas codas [p - k], tem-se que tanto a vogal [i] como a vogal [æ] repercutem em maiores índices de acuidade. Novamente (a) a vogal [i] seria a primeira opção para a acuidade em [p] porque resulta em transições formânticas mais características do ponto labial, e a vogal [æ] seria a segunda opção para acuidade em [p] por ser mais curta do que a vogal [i]; e (b) a vogal [æ] seria a primeira opção para a acuidade em [k] porque resulta em transições formânticas mais características do ponto dorsal, e a vogal [i] seria a segunda opção para acuidade em [k] por ser mais curta do que a vogal [i]. Verificou-se que a vogal [i] continua a implicar índices mais baixos de acuidade, o que pode ser explicado pela possibilidade de, especialmente frente a esta vogal, a consoante [k] ser eventualmente percebida como [p], gerando dificuldades de discriminação (e possivelmente de identificação), conforme aponta Delattre (1958).

Exatamente o mesmo padrão se repete para as palavras em um mesmo estímulo cujas consoantes finais são [t -k]: (a) a vogal [æ] seria a primeira opção para a acuidade em [t] porque resulta em transições formânticas mais características do ponto coronal, e a vogal [i] seria a segunda opção para acuidade em [t] por ser mais curta do que a vogal [i]; e (b) a vogal [æ] seria a primeira opção para a acuidade em [k] porque resulta em transições formânticas mais características do ponto dorsal, e a vogal [i] seria a segunda opção para acuidade em [k] por ser mais curta do que a vogal [i]. Os baixos índices de acuidade implicados pela vogal [i] nas palavras contendo as codas [t -k] poderiam ocorrer em função de dois fatores: (a) [t] poderia vir a ser confundido com um [p], uma vez que, quanto mais anterior a vogal precedente a [t], mais este segmento se parece acusticamente com um segmento labial, em termos de F2 (Delattre; Liberman; Cooper, 1955); e/ou (b) [k] poderia vir a ser confundido com um [p], uma vez que, por efeitos de coarticulação com a vogal [i], o ponto dorsal pode ser confundido perceptualmente com o ponto labial (DELATTRE, 1958).

Complementarmente, para explicar os baixos índices de acuidade discriminatória envolvendo os estímulos terminados em [p -t] e [t -k] à luz do modelo de Best e Tyler (2007), podem-se evocar dois casos de assimilação. O primeiro seria o caso conhecido como *category-goodness assimilation contrast*, o qual prevê que ambas as categorias fonológicas da L2 são percebidas como equivalente à mesma categoria fonológica da L1, só que uma é percebida como mais desviante do que a outra. Assim, quando os aprendizes não conseguem discriminar entre [p -t] e [t -k], pode ser que eles estejam percebendo o segmento [t] como [p] ou [k], respectivamente. Como a categoria fonológica que corresponde ao segmento [t] ainda pode estar em desenvolvimento, na tentativa de discriminar entre [p -t], os aprendizes associam a consoante labial da L2 a um excelente exemplar de /p/; a consoante alveolar da L2, por sua vez, também é associada à categoria fonológica de /p/, muito embora sendo classificada como um exemplar mais desviante da oclusiva labial. O mesmo ocorre na discriminação de [t -k], em que a consoante velar da L2 pode ser associada a um excelente exemplar de /k/, enquanto a oclusiva alveolar da L2 pode figurar como um exemplar mais desviante da oclusiva velar. O segundo caso possível de assimilação seria o *single-category L2 contrast assimilation*, em que ambas as categorias fonológicas da L2 são percebidas como equivalentes às da L1. Neste formato, os sons da L2 presentes no contraste [p -t] seriam exemplares igualmente bons da categoria fonológica de /p/ e, da mesma forma, os sons da L2 presentes no contraste [t -k] seriam exemplares igualmente bons da categoria fonológica de /k/.

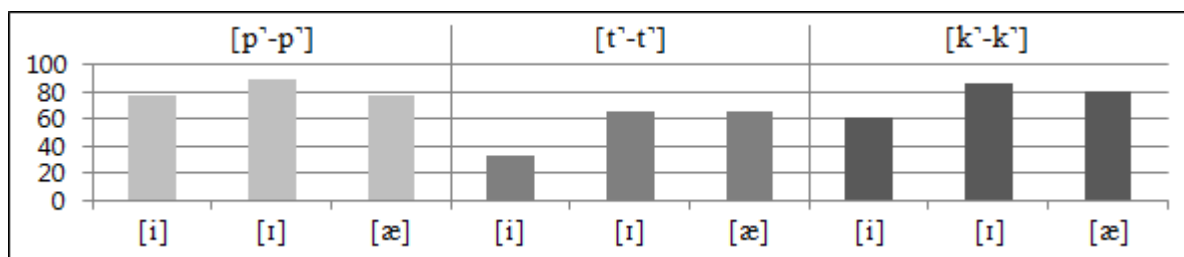
No que compete aos *catch trials* (sequências de três palavras com consoantes finais idênticas) ainda na Tarefa B, quando os estímulos que terminavam em [p -p] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 78% de acuidade na discriminação das consoantes finais. Quando a vogal nuclear era [I], eles atingiram 90% de acuidade perceptual. Ainda, quando a vogal nuclear era [æ], os níveis de acuidade na discriminação das consoantes finais foram 78%.

Quando os estímulos que terminavam em [t -t] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 33% de acuidade na discriminação das consoantes finais. Quando a vogal nuclear era [I], eles atingiram 66% de acuidade perceptual. Ainda, quando a vogal nuclear era [æ], os níveis de acuidade na discriminação das consoantes finais foram 66%.

Quando os estímulos que terminavam em [k -k] tinham como segmento nuclear a vogal [i], os aprendizes atingiram 62% de acuidade na discriminação das consoantes finais. Quando a vogal nuclear era

[ɪ], eles atingiram 87% de acuidade perceptual. Ainda, quando a vogal nuclear era [æ], os níveis de acuidade na discriminação das consoantes finais foram 80%. O Gráfico 6 apresenta tais resultados.

Gráfico 6 – Percentual de acuidade na discriminação das consoantes idênticas finais sem soltura de ar audível de acordo com a vogal nuclear



Fonte: elaborado pelo autor.

O teste de Friedman realizado indicou haver diferenças estatisticamente significativas entre as vogais precedendo [p -p] ( $\chi^2(2)=7,896$ ;  $p=0,019$ ). Testes de Wilcoxon, com correção de Bonferroni, sugeriram haver diferenças entre as vogais [i] e [ɪ] ( $Z=-2,676$ ;  $p=0,007$ ) e entre as vogais [ɪ] e [æ] ( $Z=-2,556$ ;  $p=0,011$ ). Diferenças significativas também foram encontradas entre as vogais precedendo [t -t] ( $\chi^2(2)=6,660$ ;  $p=0,036$ ), que foram verificadas através de testes de Wilcoxon (com correção de Bonferroni), correspondendo a [i] e [ɪ] ( $Z=-2,967$ ;  $p=0,003$ ) e a [i] e [æ] ( $Z=-2,137$ ;  $p=0,013$ ). Além disso, diferenças significativas também foram encontradas entre as vogais precedendo [k -k] ( $\chi^2(2)=11,318$ ;  $p=0,003$ ), as quais foram determinadas através de testes de Wilcoxon (com correção de Bonferroni), correspondendo a [i] e [ɪ] ( $Z=-3,245$ ;  $p=0,001$ ) e também a [i] e [æ] ( $Z=-2,659$ ;  $p=0,008$ ).

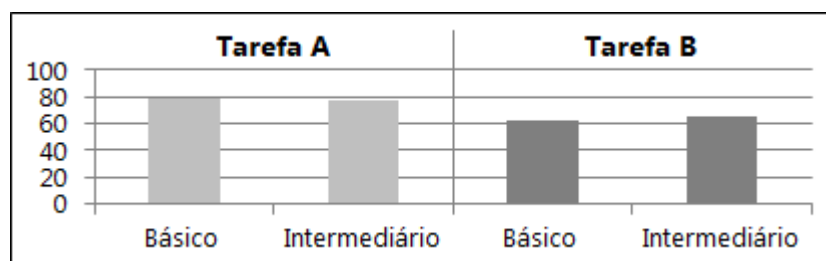
Observa-se que o padrão existente na percepção acurada da consoante final, o qual depende da configuração das transições formânticas e da duração da vogal, se mantém também para os *catch trials*. Tal resultado vem a corroborar a hipótese anteriormente prevista.

#### 4.3 A influência do nível de proficiência na percepção das consoantes finais

Para determinar se o nível de proficiência dos aprendizes teria alguma implicação para a percepção acurada das consoantes finais, compararam-se os níveis básico ( $n=11$ ) e intermediário ( $n=21$ ) em ambas as Tarefas A e B.

Na Tarefa A, os aprendizes de nível básico atingiram 79% de acuidade na identificação das consoantes finais, e os aprendizes de nível intermediário atingiram 77%. Na tarefa B, os aprendizes de nível básico tiveram 62% de acuidade na discriminação das consoantes finais, enquanto os aprendizes de nível intermediário tiveram 65% de acuidade. O Gráfico 7 apresenta a comparação dos dois grupos em ambas as Tarefas.

Gráfico 7 – Percentual de acuidade nas tarefas de percepção de acordo com nível de proficiência



Fonte: elaborado pelo autor.

No tocante à Tarefa A, o teste de Mann-Whitney realizado indicou não haver diferenças estatisticamente significativas entre os aprendizes de nível básico e intermediário em termos de acuidade na identificação dos segmentos finais ( $U=112,0$ ;  $p=0,889$ ). Na Tarefa B, respectivamente, outro teste de Mann-Whitney

realizado indicou não haver diferenças estatisticamente significativas entre os aprendizes de nível básico e intermediário em termos de acuidade na discriminação dos segmentos finais ( $U=98,5$ ;  $p=0,499$ ).

Como o nível de proficiência dos aprendizes não tem impacto sobre a percepção das oclusivas finais sem soltura de ar audível, uma conclusão razoável é a de que, desde os estágios iniciais do período de aquisição da L2, os aprendizes já atingem níveis consideráveis de acuidade perceptual, o que implica afirmar que eles já estão aptos a perceber detalhes fonéticos mais sutis. Assim, pode-se afirmar que, como o fenômeno abordado é estritamente fonético, a capacidade perceptual seja desencadeada por fatores essencialmente fisiológicos e cognitivos (percepção acústica e auditiva de um sinal físico por parte de um instrumental auditório – o ouvido, que percebe oscilações de frequências do som) do que fatores relacionados à proficiência em uma L2 específica.

## 5 CONSIDERAÇÕES FINAIS

Este estudo objetivou verificar a percepção de oclusivas finais sem soltura de ar audível presentes no inglês por aprendizes brasileiros. Além de uma discussão sobre o efeito do ponto de articulação das consoantes finais sobre a percepção, procurou-se tecer considerações importantes acerca das relações existentes entre vogal nuclear e consoante sem soltura em posição de coda. Tais observações foram extremamente relevantes para contribuir com o *background* teórico estabelecido até então no campo da Fonética Acústica e também permitiram reconsiderar o impacto que padrões assimilatórios de categorias fonológicas não nativas podem ter sobre um sistema fonológico.

Estudos vindouros com a utilização de pseudopalavras (em que as restrições fonotáticas sejam respeitadas e todos os segmentos vocálicos do inglês sejam considerados) e um número maior de participantes podem ser uma alternativa interessante para se obter resultados mais conclusivos acerca dos efeitos da vogal nuclear sobre as oclusivas finais sem soltura. Além disso, um estudo que conte com um número robusto de diferentes vogais seria fundamental para que a hipótese sobre a duração vocálica fosse posta em xeque (e talvez a tensão se mostrasse uma pista primordial), de modo a se afirmar com mais propriedade que a tendência de percepção das consoantes finais seguiria ou não o que foi previsto pela hipótese.

Ainda é tarefa de pesquisas futuras verificar a percepção das oclusivas finais em aprendizes com diferentes sistemas de L1, as quais poderiam garantir comparações tipológicas relevantes em termos de capacidade perceptual e também apontar aspectos comuns e complementares entre tais sistemas no que tange a propriedades acústicas mais finas.

## REFERÊNCIAS

ABRAMSON, A.; TINGSABADH, K. Thai Final Stops. *Phonetica*, v. 56, p. 111-122, 1999.

ALLAN, D. *Oxford Placement Test 1*. Oxford: Oxford University Press, 2004.

ALVES, U. A. *Aquisição das sequências finais de obstruintes do inglês (l2) por falantes do Sul do Brasil: análise via teoria da otimidade*. 2008. 337 f. Tese (Doutorado em Letras) - Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2008.

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

BETTONI-TECHIO, M. *Perceptual training and word initial /s/-clusters in Brazilian Portuguese/English interphonology*. 2008. 214 f. Tese (Doutorado em Letras)—Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2008.

BOERSMA, P. *The OCP in functional phonology*. Manuscript, University of Amsterdam, 1998. Disponível em <http://ruccs.rutgers.edu/roa.html>. Acesso em 17 abr. 2011.

CELCE-MURCIA, M.; BRINTON, D.; GOODWIN, J.; GRINER, B. *Teaching pronunciation: a course book and reference guide*. 2. ed. Hong Kong: Cambridge University Press, 2010.

CÔTÉ, M-H. *Phonetic salience and consonant cluster simplification*. In BRUENING, B; KANG, Y.; MCGINNIS, M. (Eds.). Cambridge, MA: MIT, 1997. p. 229-262.

DAVIDSON, L. Characteristics of stop releases in American English spontaneous speech. *Speech Communication*, v. 53, n. 8, p. 1042-1058, 2011.

DELATTRE, P. Acoustic cues in speech: first report. *Phonetica*, v. 2, p. 108-118, 1958.

\_\_\_\_\_. Unreleased velar plosives after back-rounded vowels. *Journal of the Acoustical Society of America*, v. 30, p. 581-582, 1958.

\_\_\_\_\_; LIBERMAN, A.; COOPER, F. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, v. 27, n. 4, p. 769-773, 1955.

HOUSEHOLDER, F. Unreleased PTK in American English. In: HALLE, M.; LUNT, H.; MCLEAN, H.; VAN SCHOONEVELD, C. (Eds.). *For Roman Jakobson*. Mouton: The Hague, 1956. p. 235-244.

HUME, E.; JOHNSON, K.; SEO, M.; TSERDANELIS, G. A cross-linguistic study of stop place perception. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 14., San Francisco. *Anais*. University of California, 1999. p. 2069-2072.

JUN, J. *Perceptual and articulatory factors in place assimilation: an optimality theoretic approach*. 1995. 184 f. Tese (Theoretical Linguistics) – Humanities, University of California, Los Angeles, 1995.

KANG, Y., Perceptual similarity in loanword adaptation: adaptation of English post-vocalic word-final stops to Korean. *Phonology*, v. 20, p. 219-273, 2003.

KLUGE, D. *Brazilian EFL learners' identification of word-final /m-n/: native/non-native realizations and effect of visual*. 2009. 189 f. Tese (Doutorado em Letras) – Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2009.

LADEFOGED, P.; JOHNSON, K. *A course in Phonetics*. 6. ed. Boston: Cengage Learning, 2011.

\_\_\_\_\_. *Vowels and consonants*. 2. ed. Oxford: Blackwell Publishing, 2005.

LIBERMAN, A.M.; HARRIS, K.S. et al. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, v. 54, p. 358-368, 1957.

LISKER, L. Perceiving Final Voiceless Stops without Release: Effects of Preceding Monophthongs versus Nonmonophthongs. *Phonetica*, n. 56, p. 44-55, 1999.

MARIAN, V; BLUMENFELD, H.; KAUSHANSKAYA, M. The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of speech, language, and hearing research*, v. 50, p. 940-967, 2007.

NOBRE-OLIVEIRA, D. *The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers*. 2007. 211 f. Tese (Doutorado em Letras) – Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2007.

PEROZZO, R. V. *Percepção de oclusivas não vozeadas sem soltura auditiva em codas finais do inglês (L2) por brasileiros: o papel do contexto fonético-fonológico, da instrução explícita e do nível de proficiência*. 2013. 191 f. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.

RAUBER, A. *Perception and production of English Vowels by Brazilian EFL Speakers*. 2006. 218 f. Tese (Doutorado em Letras) - Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2006.

REIS, C.; ESPESSER, R. Estudo eletropalatográfico de fones consonantais e vocálicos do português brasileiro. *Estudos da Língua(gem)*, n. 3, p. 181-204, 2006.

SELKIRK, E. O. The Syllable. In H. Hulst & N. Smith (Eds.), *The Structure of Phonological Representations*. Dordrecht: Foris Publications, 1982.

van OOSTENDORP, M. Incomplete devoicing in formal phonology. *Lingua*, 118(9), 2007. p. 1362-1374.

YAVAŞ, M. *Applied English Phonology*. Malden: Blackwell Publishers, 2006.

***Recebido em 17/10/2013. Aprovado em 27/04/2014.***