

O uso de pistas visuais na identificação das nasais inglesas em final de sílaba por aprendizes brasileiros

Denise Cristina Kluge¹, Mara Sílvia Reis², Rosana Denise Koerich³

^{1, 2, 3} – Universidade Federal de Santa Catarina (UFSC), Departamento de Comunicação e Expressão, Pós-Graduação em Letras Inglês e Literatura Correspondente

klugedenise@hotmail.com, marasreis@hotmail.com rkoerich@hotmail.com

Abstract. *Brazilian Portuguese (BP) and English have different realizations of the bilabial and alveolar nasal consonants in syllable-final position (coda). Whereas in English these nasals are articulatorily and auditorily distinctive, in BP the previous vowel receives the nasalization and the consonant is deleted. One of the most salient accent markers in the BP learners of English pronunciation is such nasalization and deletion. In the line of studies exploring the relationship between foreign language sounds perception and production, and considering that human speech is a multimodal function (Rosenblum,2005), this study investigated whether learners rely on phonetic information contained in visual cues when identifying the nasals /m/ and /n/ in syllable-final position. The data were collected with ten intermediate learners of the Extracurricular Course at UFSC. Perception was assessed by means of an Identification Test with three test conditions: (i) **audio only**, in which participants had to identify the final nasal of a word they heard; (ii) **video only**, in which participants had to identify the final nasal of a word presented through video only, without audio; and (iii) **audio-video**, in which the participants had to identify the final nasal of a word presented through video and audio. It was hypothesized that the combination of visual with auditory cues would favor the accurate identification of the nasal consonant in syllable-final position. As regards the phonological context, the three test conditions considered high, mid and low previous vowels. The results supported the hypothesis and revealed that the accurate identification of final /m/ and /n/ seemed to be influenced by the previous vowel with low vowels favoring and high vowels disfavoring identification.*

Key words: visual cues, nasal consonants, perception

Resumo. *A produção das nasais bilabial e alveolar em posição final de sílaba (coda) difere entre o português brasileiro (PB) e o inglês. Enquanto em inglês elas são articulatória e auditivamente distintas, em PB as nasais não são pronunciadas. A vogal antecedente recebe o traço nasal de modo que, via de regra, não há distinção entre as nasais. Um dos traços mais salientes de sotaque na fala de brasileiros aprendizes de inglês é tal nasalização e apagamento da consoante. Na linha de estudos que exploram a relação entre percepção e produção no aprendizado de sons da língua estrangeira e, considerando que a fala humana é uma função multimodal (Rosenblum,2005), este estudo investigou a utilização, por brasileiros aprendizes de inglês, de*

*informações fonéticas contidas em pistas visuais na percepção das consoantes nasais /m/ e /n/ do inglês em posição de coda. Os dados foram coletados com dez alunos do nível intermediário de inglês do Curso Extracurricular de Línguas da UFSC. A percepção foi aferida através de um teste de identificação no qual as consoantes nasais foram apresentadas em posição final de sílaba em três condições diferentes: (a) **apenas áudio**, no qual os participantes somente ouviram palavras e foram solicitados a identificar a nasal produzida; (b) **apenas vídeo**, no qual os participantes tiveram acesso ao estímulo visual, exclusivamente, e foram solicitados a identificar a nasal produzida, e (c) **áudio-vídeo**, no qual os participantes tiveram acesso aos dois tipos de estímulo e foram solicitados a identificar a nasal produzida. Quanto ao contexto fonológico antecedendo as nasais-alvo, as vogais anteriores alta, média e baixa foram selecionadas por apresentarem maior possibilidade de pares mínimos do que as demais vogais. Os resultados corroboraram a hipótese de que a combinação dos dois tipos de pistas – acústicas e visuais – favoreceria a identificação das nasais inglesas em posição final de sílaba, demonstrando que os participantes tiveram melhor desempenho nas condições **áudio-vídeo** e **apenas vídeo**, do que na condição **apenas áudio**. Em relação ao contexto fonológico, os resultados indicaram que a vogal anterior baixa parece favorecer a identificação das nasais /m/ e /n/ em posição final de sílaba, enquanto a vogal alta parece desfavorecer tal identificação.*

Palavras-chave: pistas visuais; consoantes nasais; percepção

1. Introdução

Conforme a afirmação de Rosenblum (2005, p.51), torna-se cada vez mais claro que a fala humana é uma função multimodal, geralmente apreendida tanto por meio visual (leitura labial), como por meio auditivo. O sinal de fala possui múltiplas pistas acústicas para os traços fonéticos que, muitas vezes, trazem informações redundantes. Tais redundâncias contribuem tanto para que os ouvintes sem problemas auditivos identifiquem a fala em muitos contextos em sua língua materna (L1), como, por exemplo, em situações em que há perda da qualidade do sinal pelo barulho (Sumby e Pollack, 1954), como também para que os ouvintes com problemas auditivos (Grant e Seitz, 1998). Grant e Seitz (1998, p. 2438) definem a expressão “benefício Audio Vídeo” (*AV benefit*) como a quantidade de benefício resultante da combinação das pistas visual e auditiva, e que, nestes termos, tem sido usada para descrever as vantagens de uma apresentação áudio-visual. Entretanto, o benefício AV pode depender do peso perceptual das pistas visual e auditiva (Hazan, Sennema, Faulkner, Ortega-Llebaria, Iba e Chung, 2006, p. 1741).

Uma forma de avaliar o peso perceptual é o efeito McGurk (McGurk e McDonald, 1976). Este efeito ocorre quando as pistas visual e acústica divergem. Por exemplo, um grupo de participantes do estudo identificou como /d/ um estímulo visual /g/ e auditivo /b/. Hazan, Sennema, Iba e Faulkner (2005, p. 362) reportam que este

resultado perceptual indica que a informação contida em cada modalidade (auditiva e visual) pode estar integrada antes que o som seja categorizado pelo ouvinte.

A pesquisa de Hazan et al. (2006) sobre o uso de pistas visuais na percepção de contrastes não-nativos é referência direta no presente estudo, pois investiga a percepção de consoantes e apresenta uma descrição detalhada do método utilizado para acessar a percepção. A pesquisa investigou o efeito de pistas visuais através da percepção de dois contrastes fonêmicos do inglês que se diferenciam pela saliência de seus gestos articulatórios: as consoantes labial (/b/-/p/) e labiodental (/v/) que apresentam elevado contraste distintivo (Experimento 1), e as consoantes /r/ e /l/ que apresentam um contraste distintivo menos visual (Experimento 2). Os dois experimentos foram rodados com aprendizes espanhóis e japoneses de inglês, além de um grupo de falantes nativos de inglês. A percepção foi testada em três condições diferentes: a) **apenas áudio** (A), (b) **apenas vídeo** (V), e (c) **áudio-vídeo** (AV). Os resultados do Experimento 1 indicaram que os dois grupos de aprendizes de inglês obtiveram escore maior na condição AV do que na condição A para o contraste amplamente distintivo, indicando, portanto, evidência do benefício áudio-visual. Quanto ao Experimento 2, os resultados mostraram que nenhum dos grupos de aprendizes apresentou evidência do benefício áudio-visual para o contraste distintivo menos visual.

Com base na performance dos dois grupos de aprendizes e do grupo de falantes nativos de inglês, os autores afirmam que a saliência visual tem impacto na percepção das pistas visuais dos contrastes consonantais tanto na língua nativa como na não-nativa, uma vez que os falantes nativos de inglês e os aprendizes de inglês obtiveram um escore muito menor na condição **apenas vídeo** para o contraste menos saliente (/l/-/r/) do que para o contraste mais saliente (labial/labiodental). Os autores também afirmam que o contraste mais saliente apresentou uma percepção quase perfeita na condição **apenas vídeo** para o grupo de falantes nativos e para alguns espanhóis aprendizes de inglês.

Baseando-se nesses estudos, a presente investigação partiu do pressuposto que brasileiros aprendizes de inglês se beneficiem da apresentação **áudio-vídeo** (AV) uma vez que o contraste das consoantes nasais labial/alveolar em posição final de sílaba é visualmente distintivo no inglês. Aprendizes brasileiros de inglês geralmente apresentam dificuldades com as consoantes nasais inglesas em posição final de sílaba devido às diferenças fonológicas das duas línguas em questão.

De acordo com Fujimura e Erickson (1997, p. 100), tipicamente, as consoantes nasais são distintas quanto ao ponto de articulação, como por exemplo, entre /m/ e /n/ no inglês. Entretanto, muitas línguas não apresentam distinção de ponto de articulação para essas consoantes nasais em posição final de sílaba, como é o caso do português brasileiro (PB).

A nasalização do PB tem sido considerada sob pontos de vista diferentes. Para o presente estudo serão considerados os pontos de vista hoje mais difundidos: (1) a vogal recebe a nasalização da consoante que lhe segue, e (2) foneticamente, as consoantes nasais /m/ e /n/ em posição final de sílaba não são realizadas (Cristófaros Silva, 1999; Mateus e d'Andrade, 2000; Câmara Jr., 1971) e, portanto não são distintivas. Devido a essas diferenças nas duas línguas, espera-se que os aprendizes brasileiros de inglês

apresentem dificuldades em perceber as consoantes nasais inglesas em posição final de sílaba.

Kluge, Rauber, Reis e Bion (2007) investigaram a produção e percepção das nasais /m/ e /n/ em final de sílaba em palavras inglesas por um grupo de vinte alunos brasileiros aprendendo inglês em nível pré-intermediário, considerando a influência do processo de nasalização do português brasileiro. A percepção foi aferida através dois testes de percepção em condição de **apenas áudio**: um teste de discriminação categórica (baseado em Flege, Munro e Fox, 1994), e um teste de identificação. Esses testes também foram aplicados a um grupo de três falantes nativos de inglês norte-americano, como medida de comparação. Nos dois testes, a vogal antecedente à nasal alvo foi investigada como variável de contexto fonológico. Os resultados mostraram que os dois testes de percepção foram difíceis para os participantes brasileiros, uma vez que identificaram corretamente as consoantes nasais em menos da metade dos itens dos testes. Os resultados indicaram, também, que a vogal antecedente exerceu certa influência sobre a performance dos participantes brasileiros assim como dos falantes nativos na percepção das nasais em posição final de sílaba. Os dois grupos demonstraram dificuldade em discriminar ou identificar a consoante nasal na coda no contexto das mesmas vogais. As vogais altas tenderam a desfavorecer a discriminação (corroborando os estudos de Kurowski e Blumstein, 1987, 1995), enquanto que as vogais baixas pareceram favorecê-la.

A literatura reporta um número bastante limitado de estudos a sobre a percepção das consoantes em posição final de sílaba por falantes brasileiros de inglês (Kluge et al., 2007) e não há estudos sobre o efeito das pistas visuais na identificação das consoantes nasais inglesas em posição final de sílaba, que tenhamos conhecimento. Atentando para essas limitações, os objetivos principais deste estudo são: (i) investigar qual condição (**apenas áudio**, **apenas vídeo** ou **áudio-vídeo**) favorece a identificação das consoantes nasais inglesas /m/ e /n/ em posição final de sílaba, e (ii) investigar se o contexto fonológico influencia a percepção das consoantes nasais inglesas em posição final de sílaba, considerando a vogal antecedente a consoante nasal como variável.

2. Método, materiais e procedimentos

2.1. Participantes

Os participantes pesquisados foram 10 alunos (7 mulheres e 3 homens) entre 18 e 25 anos de idade, cursando o nível intermediário do Curso Extracurricular de Inglês da Universidade Federal de Santa Catarina (UFSC). A fim de garantir homogeneidade com relação ao nível de proficiência dos participantes, foi feita uma seleção prévia com um grupo maior de alunos. Dois falantes de inglês (um nativo e um não nativo) avaliaram e classificaram a pronúncia dos alunos em uma escala de 1 a 5 em uma gravação de 1 minuto. Os alunos avaliados como grau 1 (não-nativo) e grau 5 (pronúncia próxima a de um falante nativo) foram excluídos do grupo final. Desta forma, os 10 participantes deste estudo obtiveram graus 2, 3 ou 4 na escala de avaliação. De acordo com um questionário sobre informações biográficas e perguntas referentes à experiência dos participantes com a língua inglesa, nenhum dos participantes havia estado em país de língua inglesa. Todos

os participantes reportaram estarem mais expostos ao inglês americano do que ao britânico. Seis, dos dez participantes reportaram nunca terem tido instruções formais sobre pronúncia, sons ou símbolos fonéticos do inglês. Todos os participantes disseram não ter problemas auditivos.

2.2. Material

Os dados desta pesquisa foram coletados em dezembro de 2006 na Universidade Federal de Santa Catarina. Três instrumentos foram desenvolvidos para a coleta e análise de dados: dois questionários, sendo um para acessar as informações biográficas dos participantes, e outro para acessar as impressões e opiniões dos participantes considerando a duração do teste e o grau de dificuldade de cada uma das três condições do teste de identificação, e um teste de identificação conforme detalhado abaixo.

2.3. Teste de Identificação

O teste deste estudo foi baseado no design do teste de identificação descrito no estudo de Hazan et al. (2006) que foi desenvolvido para acessar a sensibilidade de aprendizes de uma segunda língua às informações fonéticas contidas nas pistas visuais na identificação um contraste fonêmico não-nativo. O teste continha dois blocos com 18 palavras monossílabas (CVC) para cada uma das três condições (A, AV, V). As palavras-alvo foram três pares mínimos contrastando /m/ e /n/ na coda precedidas pela vogal alta /ɪ/ (*Tim/tin*), média /ɛ/ (*gem/gen*) ou baixa /æ/ (*cam/can*).

Para a construção do teste, um falante nativo masculino de inglês americano, com treinamento fonético, gravou os estímulos orais e visuais do teste. A gravação dos estímulos visuais foi feita em uma sala com isolamento acústico utilizando uma câmera digital Canon Elura 40-MC. Para que a boca do falante ficasse em evidência durante a gravação de cada item, ela foi colocada em uma moldura. A câmera foi programada pra gravar em modo estéreo a 44,1 kHz, 16 bits. O vídeo foi gravado com resolução digital de 720 x 480 pixels. Após a gravação, o vídeo com o som foi importado para um Apple Macintosh PowerBook G4 através do programa de edição de vídeo iMovie. Neste programa, o vídeo foi editado a fim de isolar e separar cada item criando um vídeo clipe para cada item alvo. Os videoclipes foram editados de forma que a pronúncia de cada palavra iniciasse e terminasse com uma expressão facial neutra.

Os itens do teste foram os mesmos para cada uma das três condições testadas, sendo que na condição **apenas áudio** ou na condição **apenas vídeo**, as pistas visuais e auditivas foram removidas, respectivamente. Para a condição **apenas áudio**, uma cópia do projeto original foi aberta e o áudio foi extraído do vídeo e exportado para um arquivo separado com a resolução original (estéreo, 16 bit, 44.1 kHz). Os arquivos contendo apenas o áudio foram editados em arquivos separados com a ajuda do software Bias do programa Peak DV e, posteriormente, os arquivos para cada uma das palavras foram salvos em um arquivo AIFF, primeiramente, e depois convertidos para MP3 com o programa iTunes do Apple. Para a condição somente vídeo, outra cópia do projeto original foi aberta no iMovie para a remoção do áudio. As palavras foram exportadas individualmente como arquivos AVI com resolução de 320 x 240 pixel. Para a condição **áudio-vídeo**, uma terceira cópia do projeto original foi aberta no iMovie e o videoclipe correspondente a cada palavra foi exportado como arquivo AVI com resolução de 320 x

240 pixel e áudio com resolução estéreo de 16 bit, 44.1 kHz. Desta forma, as seis palavras monossilábicas com /m/ ou /n/ na coda foram apresentadas em três condições: (a) **apenas áudio** (A), onde os participantes ouviram a pronúncia de uma palavra e identificaram a consoante nasal final, (b) **áudio-vídeo** (AV), onde os participantes viram e ouviram a pronúncia da palavra alvo e identificaram a consoante nasal final, e (c) **apenas vídeo** (V), onde os participantes apenas viram o movimento labial durante a pronúncia de cada palavra e identificaram a consoante nasal final.

Para garantir que os estímulos orais e visuais contendo as consoantes nasais fossem claros e representativos da língua inglesa, os mesmos foram digitalizados e ordenados aleatoriamente para que um falante nativo de inglês americano pudesse ouvi-los e/ou vê-los e identificar a consoante nasal final de cada palavra. Apenas os estímulos que foram consistentemente identificados (100%) pelo falante nativo foram utilizados na elaboração do teste de percepção. Este procedimento foi baseado no método utilizado em um estudo de Harnsberger (2000, p. 766). Cada condição (A, AV e V) apresentava dois blocos de 18 itens, totalizando 108 itens para cada participante. Cada palavra alvo foi repetida três vezes em cada bloco.

O teste foi aplicado em um laptop através do programa Power Point e a ordem dos itens de cada bloco em cada uma das três condições foi aleatorizada para minimizar o efeito de prática do teste. Seguindo o estudo de Hazan et al. (2006), duas ordens de apresentação foram utilizadas: A, AV, V ou AV, A, V, sendo que estas duas ordens de apresentação foram contrabalanceadas. De acordo com os autores, a condição **apenas vídeo** foi sempre apresentada por último, pois é a condição considerada mais difícil para os participantes.

A tarefa dos participantes consistia em identificar a consoante nasal final que ouviam, viam ou ouviam-viam em cada uma das três condições circulando uma das consoantes nasais (/m/ ou /n/) em uma folha de respostas. Um teste de familiarização contendo 8 itens para cada uma das três condições também foi desenvolvido.

2.4. Procedimento

Os dados foram coletados individualmente em sessões de aproximadamente 20 minutos. As instruções foram dadas em inglês e em português para assegurar que os participantes entendessem completamente a tarefa. Primeiramente, o participante respondeu ao questionário sobre as informações biográficas. Em seguida, os testes foram aplicados, sendo que metade dos participantes seguiu ordem: (1) condição **apenas áudio**, (2) questionário sobre a condição **apenas áudio**, (3) condição **áudio-vídeo**, (4) questionário sobre a condição **áudio-vídeo**, (5) condição **apenas vídeo**, (6) questionário sobre a condição **apenas vídeo**; e a outra metade seguiu a ordem: (1) condição **áudio-vídeo**, (2) questionário sobre a condição **áudio-vídeo**, (3) condição **apenas áudio**, (4) questionário sobre a condição **apenas áudio**, (5) condição **apenas vídeo**, (6) questionário sobre a condição **apenas vídeo**.

Quanto aos dados, os resultados foram estatisticamente analisados usando o programa SPSS para Windows – versão 11.0. Como os dados não estavam normalmente distribuídos devido ao limitado número de participantes, foram utilizados os testes

estatísticos não paramétricos: (1) Friedman, que equivale à análise de variância de medidas repetidas (*repeated-measure ANOVA*), e (2) Wilcoxon, um teste post-hoc. O nível de significância alpha (p) foi estabelecido em .05.

4. Resultados e discussão

Em relação ao primeiro objetivo do estudo, a tabela 1 mostra o número de identificações corretas e as porcentagens correspondentes para cada um dos participantes. As porcentagens foram baseadas no total de 18 respostas para cada uma das consoantes nasais em cada uma das condições para cada participante, totalizando 180 respostas para cada condição.

Partic.	A /m/	AV /m/	V /m/	A /n/	AV /n/	V /n/
1	18(100%)	18(100%)	18(100%)	17(94.44%)	18(100%)	18(100%)
2	12(66.67%)	18(100%)	18(100%)	13(72.22%)	18(100%)	17(94.44%)
3	15(83.33%)	17(94.44%)	18(100%)	17(94.44%)	18(100%)	17(94.44%)
4	18(100%)	16(88.89%)	13(72.22%)	8(44.44%)	16(88.89%)	11(61.11%)
5	16(88.89%)	17(94.44%)	18(100%)	16(88.89%)	18(100%)	18(100%)
6	17(94.44%)	18(100%)	18(100%)	15(83.33%)	18(100%)	18(100%)
7	14(77.78%)	17(94.44%)	16(88.89%)	12(66.67%)	17(94.44%)	16(88.89%)
8	12(66.67%)	17(94.44%)	14(77.78%)	13(72.22%)	17(94.44%)	15(83.33%)
9	13(72.22%)	15(83.33%)	15(83.33%)	14(77.78%)	17(94.44%)	10(55.55%)
10	12(66.67%)	17(94.44%)	18(100%)	15(83.33%)	18(100%)	14(77.78%)
Total	147 (81.67%)	170 (94.44%)	166 (92.22%)	140 (77.77%)	175 (97.22%)	154 (85.55%)

Tabela 1. Número de identificação correta das consoantes nasais na coda no teste de identificação por participante (total de 18 itens para cada consoante em cada condição do teste)

Considerando a performance dos participantes como um grupo, os resultados mostram que a porcentagem de respostas corretas na identificação das consoantes nasais /m/ e /n/ em posição de coda aumentou gradualmente da condição **apenas áudio** (A) para a condição **apenas vídeo** (V) e, então, para a condição **áudio-vídeo** (AV). Diferentemente dos resultados encontrados por Hazan et al. (2006) que mostraram que a condição **apenas vídeo** era a mais difícil; os resultados deste estudo indicam que a condição **apenas vídeo** revelou certo grau de dificuldade uma vez que os participantes identificaram corretamente a consoante nasal-alvo em 320 dos 360 itens, mas que, esta condição, não foi a mais difícil neste estudo.

O teste estatístico Friedman revela efeito significativo para condição de teste na identificação do /m/ ($X^2(2, N=10) = 8.909, p = .012$). O teste Wilcoxon foi, então, rodado a fim de verificar quais os pares de condições eram estatisticamente significativos na identificação da nasal bilabial na coda, obtendo-se os seguintes resultados: (a) diferença não significativa para os pares 'A x V' ($Z = -1.854, p = .064$) e 'AV x V' ($Z =$

- .649, $p = .516$); (b) diferença significativa para o par 'AV x A' ($Z = -2.203$, $p = .028$). Desta forma, os resultados indicam que a condição **áudio-vídeo** favoreceu a correta identificação da consoante nasal /m/ quando comparada com a condição **apenas áudio**.

Quanto à identificação da consoante nasal /n/, o teste Friedman também mostrou um efeito estatisticamente significativo para condição de teste ($X^2 (2, N=10) = 14.389$, $p = .001$). O teste post-hoc, Wilcoxon, levou aos seguintes resultados: (a) diferença não significativa para os pares 'A x V' ($Z = -1.550$, $p = .121$) e 'AV x V' ($Z = -2.384$, $p = .017$); e (b) diferença significativa para o par 'AV x A' ($Z = -2.814$, $p = .005$). Os resultados indicam que, assim como na identificação do /m/, a condição AV parece favorecer a correta identificação da consoante nasal /n/ quando comparada com a condição A. Os resultados encontrados neste estudo para a identificação de ambas consoantes nasais em posição final de sílaba corroboram os resultados encontrados por Hazan et al. (2006), e parecem indicar a importância que o input visual pode ter na percepção de contrastes visualmente distintos.

Os resultados deste estudo também estão de acordo com as impressões dos participantes (acessadas por meio de um questionário, como foi explicado na seção 2.2.) em relação à dificuldade que eles tiveram em identificar as consoantes nasais em cada uma das condições do teste. A maioria dos participantes (6 de 10) revelou que a condição AV foi a mais fácil para identificar a coda nasal inglesa, porque nesta condição eles podiam ver o movimento da boca facilitando a identificação da consoante que eles ouviram. Quanto à condição mais difícil, a maioria dos participantes (8) revelou que a condição **apenas áudio** foi a mais difícil, devido ao fato de que as palavras não estavam contextualizadas e eles não tinham recurso algum que pudesse ajudá-los a decidir que consoante nasal eles ouviam. Quando perguntados, no questionário, se prestavam atenção no movimento labial quando falam com outra pessoa tanto em PB como em inglês, seis participantes revelaram que eles prestam atenção no PB em situações de barulho excessivo ou quando querem prestar mais atenção ao que está sendo dito, e 7 reportaram que prestam atenção nos movimentos labiais quando falam inglês com alguém pessoalmente ou quando querem entender palavras específicas.

Em relação ao segundo objetivo do estudo, a Tabela 2 mostra o número e a porcentagem de identificação correta das consoantes nasais inglesas em posição final de sílaba no contexto de três vogais anteriores: alta (/ɪ/), média (/ɛ/) e baixa (/æ/). Os resultados estão baseados na análise das três condições do teste de identificação.

Vogal Anterior	No. Respostas /m/	No. Respostas Corretas /m/	No. Respostas /n/	No. Respostas Corretas /n/
/ɪ/	180	147 (81.67%)	180	154 (85.56%)
/ɛ/	180	169 (93.89%)	180	148 (82.22%)
/æ/	180	165 (91.67%)	180	170 (94.44%)
Total	540	481 (89.07%)	540	472 (87.44%)

Tabela 2. Resultado da percepção das consoantes nasais na coda em relação vogal antecedente à consoante

A Tabela 2 mostra que a consoante nasal /m/ foi menos frequentemente identificada no contexto da vogal anterior alta /ɪ/. Os resultados também mostram que os participantes tiveram desempenho similar nos contextos das vogais média e baixa. O teste Friedman confirma que a vogal antecedente teve efeito na identificação da consoante nasal /m/ ($X^2(2, N=10) = 9.267, p = .01$). Rodado o teste Wilcoxon para verificar quais diferenças entre os pares de vogais anteriores eram estatisticamente significativas na identificação da consoante nasal bilabial, encontrou-se os seguintes resultados: (a) diferença não significativa para o par ‘vogal baixa - /æ/’ x ‘vogal média - /ɛ/’ ($Z = -.755, p = .450$), e (b) diferença significativa para os pares ‘vogal alta - /ɪ/’ x ‘vogal média - /ɛ/’ ($Z = -2.536, p = .011$) e ‘vogal alta - /ɪ/’ x ‘vogal baixa - /æ/’ ($Z = -2.047, p = .041$). Desta forma, pode-se dizer que, os resultados do estudo indicam que a vogal anterior alta /ɪ/ parece desfavorecer a identificação da consoante nasal inglesa /m/, resultados que corroboram os encontrados por Kurowski e Blumstein (1987, 1995) e Kluge et al. (2007).

No que diz respeito à identificação da consoante nasal /n/ na coda, a Tabela 2 mostra que esta foi mais frequentemente identificada no contexto da vogal anterior baixa /æ/. Os resultados mostram também que os aprendizes brasileiros de inglês tiveram uma performance semelhante no contexto das vogais anteriores média e alta. O teste Friedman confirma haver efeito da vogal antecedente na identificação da consoante nasal /n/ ($X^2(2, N=10) = 9.455, p = 0.009$) e os testes Wilcoxon produziram os seguintes resultados: (a) diferença não significativa para o par ‘vogal média - /ɛ/’ x ‘vogal alta - /ɪ/’ ($Z = -.862, p = .389$), e (b) diferença estatisticamente significativa para os pares ‘vogal alta - /ɪ/’ x ‘vogal baixa - /æ/’ ($Z = -2.198, p = .028$) e ‘vogal baixa - /æ/’ x ‘vogal média - /ɛ/’ ($Z = -2.552, p = .011$).

Portanto, os resultados indicam que a vogal anterior baixa /æ/ parece favorecer a identificação da consoante nasal alveolar em posição final de sílaba. Este resultado corrobora os encontrados por Kluge et al. (2007), entretanto não foi encontrado efeito da vogal anterior baixa na identificação da consoante nasal inglesa /m/.

4. Conclusões, limitações e sugestões para pesquisas futuras

A hipótese principal deste estudo propunha que a condição **áudio-vídeo** (AV) favorecesse a identificação das consoantes nasais inglesas /m/ e /n/ em final de sílaba por aprendizes brasileiros de inglês.

Os resultados encontrados suportaram esta hipótese indicando que os participantes se beneficiaram da condição **áudio-vídeo** (AV), mas somente quando comparada à condição **apenas áudio** (A) na identificação de ambas as consoantes.

Concluindo, os aprendizes brasileiros de inglês deste estudo parecem ter se beneficiado da apresentação **áudio-vídeo** (AV), como discutido por Grant e Seitz

(1998), na identificação das consoantes nasais /m/ e /n/ em posição final de sílaba do inglês, uma vez que o contraste investigado é visualmente distintivo.

Quanto à influência do contexto fonológico, os resultados mostraram que a vogal antecedente parece influenciar a identificação das consoantes nasais inglesas seguintes. Os resultados também demonstraram que a vogal anterior baixa tende a favorecer a identificação da consoante alveolar na coda e que a vogal anterior alta tende a desfavorecer a identificação da consoante nasal bilabial em posição final de sílaba pelos participantes deste estudo.

Algumas limitações do estudo foram detectadas durante o processo de análise de dados. Uma das principais limitações é, com certeza, o número limitado de participantes aliado ao fato de não haver dados de falantes nativos de inglês, como medida de referência. Em termos de variáveis, o presente estudo investigou a identificação das consoantes nasais-alvo no contexto de apenas três vogais antecedentes às mesmas. Um estudo futuro poderia investigar a identificação das nasais no contexto de outras vogais antecedentes, bem como analisar a identificação destas no contexto de cada uma das vogais antecedentes em cada uma das três condições (A, V, e A-V) separadamente, já que neste estudo as três condições foram analisadas como um todo para a variável contexto fonológico antecedente.

6. Agradecimento

As autoras deste artigo gostariam de agradecer a CAPES por financiar esta pesquisa.

7. Referências

- CÂMARA JR, J.M. *Problemas da lingüística descritiva*. Petrópolis: Editora Vozes. 128p., 1971.
- CRISTÓFARO SILVA, T. *Fonética e fonologia do português: Roteiro de estudos e guia de exercícios*. São Paulo: Contexto. 254p., 1999.
- FLEGE, J. E.; MUNRO, M. J., & FOX, R. A. Auditory and Categorical Effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, 95(6), 3623-3641, 1994.
- FUJIMURA, O. & ERICKSON, D. Acoustic Phonetics. In. HARDCASTLE, W. J. & LAVER, J. (Eds), *The handbook of phonetic sciences*. Cambridge: Blackwell Publishers., p. 65-115, 1997.
- GRANT, K. W., AND SEITZ, P. F. Measures of auditory-visual integration in nonsense syllables and sentences. *Journal of the Acoustical Society of America*, 104, 2438-2450, 1998.
- HAZAN, V.; SENNEMA, A.; ORTEGA-LLEBARIAD, M.; IBA, M.; FAULKNER, A. Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, 47, 360-378, 2005.

- HAZAN, V.; SENNEMA, A.; FAULKNER, A.; ORTEGA-LLEBARIAD, M.; IBA, M.; CHUNG, H. The use of visual cues in the perception of non-native consonant contrasts. *Journal Acoustical Society of America*, 119 (3), 1740-1751, 2006.
- KLUGE, D. C.; RAUBER, A. S.; REIS, M. S.; BION, R.A.H. The relationship between perception and production of English nasal codas by Brazilian learners of English. In *Proceedings of Interspeech 2007*, 2297-2300, 2007.
- KUROWSKI, K. M. AND BLUMSTEIN, S. E. Acoustic properties for place of articulation in nasal consonants. *Journal of the Acoustical Society of America*, 81(6), 1917-1927, 1987.
- KUROWSKI, K. M. AND BLUMSTEIN, S. E. Acoustic properties for the perception of nasal consonants. In M. K. Huffman and R.A. Krakow (Eds.) *Nasals, Nasalization and the Velum* Academic Press, Inc., p. 197-222, 1995.
- MATEUS, M. H. M., & D´ANDRADE, E. *The phonology of Portuguese*. Oxford: Oxford University Press. 376p., 2000.
- MCGURK, H. AND MACDONALS, J. Hearing lips and seeing voices. *Nature*, 264, 746-748, 1976.
- ROSENBLUM, L.D. Primacy of multimodal Speech Perception. In D.B. Pisoni and E.R. Remez (Eds.) *The handbook of speech perception*. Blackwell Publishing. p.51-78, 2005.
- SUMBY, W. H., AND POLLACK, I. Visual contribution to speech intelligibility in noise, *Journal of the Acoustical Society of America*, 26, 212–215, 1954.