

## Uma discussão acerca da aplicação do *Perceptual Assimilation Model-L2* à percepção fônica de língua estrangeira: questões de pesquisa e desafios teóricos

Discussing the application of the *Perceptual Assimilation Model-L2* to foreign language speech perception: research enquiries and theoretical challenges

Reiner Vinicius Perozzo\*  
Ubiratã Kickhöfel Alves\*\*

**RESUMO:** Este trabalho tem o objetivo de apontar possíveis alterações ao *Perceptual Assimilation Model-L2* [*PAM-L2* (BEST; TYLER, 2007)], modelo de percepção da fala em segunda língua, para que o mesmo dê conta da percepção fônica em contexto de línguas estrangeiras. Para tanto, apresentamos o *PAM-L2* e suas diretrizes, visitamos alguns estudos nacionais que se utilizaram do modelo e problematizamos suas bases fonológica e filosófica. Neste sentido, advogamos pela instauração do gesto acústico-articulatório concebido por Albano (2001) como unidade perceptual e defendemos o realismo indireto como teoria filosófica para acomodar a percepção fônica no cenário estrangeiro da aprendizagem perceptual. Concluímos nossa exposição levantando alguns pontos teóricos e metodológicos que merecem atenção quando tratamos da percepção fônica em língua estrangeira, fundamentais ao modelo em sua totalidade.

**PALAVRAS-CHAVE:** Ecologia da percepção. Gesto articulatório. *PAM-L2*. Percepção fônica estrangeira.

**ABSTRACT:** This article aims to suggest some changes to the *Perceptual Assimilation Model-L2* (*PAM-L2* (BEST; TYLER, 2007)), a model of second language speech perception, so that it may account for speech perception of foreign languages. In order to do so, at first we present *PAM-L2* and its guidelines, then we list some research studies carried out in Brazil that have utilized the model, and finally we problematize its phonological and philosophical bases. In this regard, we claim for the use of the articulatory gesture conceived by Albano (2001) as the perceptual unit of speech and we suggest that indirect realism may be suitable as a philosophical theory to encompass speech perception in foreign settings of perceptual learning. We conclude our discussion raising some other relevant theoretical and methodological issues, fundamental to the model as a whole.

**KEYWORDS:** Articulatory gesture. Foreign speech perception. *PAM-L2*. Perceptual ecology.

### 1. Introdução

Os estudos de percepção dos sons da fala, especialmente no que diz respeito àqueles que veiculam contrastes fonológicos, tiveram seu início nos Laboratórios *Haskins*, nos Estados

\* Mestre em Letras. Instituto de Letras. UFRGS.

\*\* Doutor em Letras. Instituto de Letras. UFRGS.

Unidos, por volta de 1950 (NISHIDA, 2012). Estudiosos como Alvin Liberman, Pierre Delattre e Franklin Cooper (DELATTRE; LIBERMAN; COOPER, 1955) se debruçavam sobre experimentos que continham sequências sonoras representativas de sílabas CV do inglês e, através de métodos físicos, alteravam suas configurações formânticas, a fim de promover diferenças acústicas cujos resultados implicavam distinções de ponto de articulação para segmentos oclusivos.

Com a evolução das pesquisas sobre percepção fônica, cresciam, também, as discussões sobre quais eram as bases das informações sonoras que refletiam as diferenças lexicais, causando uma oposição entre a sua natureza acústica e a sua natureza articulatória, conforme indica Nishida (2012). Neste cenário, encontram-se, de um lado, evidências fonológicas com base em propriedades acústicas do sinal de fala, observadas por Ohala (1996), e, de outro, fatos articulatórios veiculados através da percepção e da produção da fala, sustentados por Fowler (1996)<sup>1</sup>. Além disso, aliada à busca de compreender as propriedades que marcariam fronteiras perceptuais dos sons da fala, havia a necessidade de modelar formalmente como os indivíduos, ao adquirir seu idioma materno, percebiam auditivamente seus elementos sonoros. Tal modelagem, alavancada pela busca dos primitivos referentes à percepção da fala em língua materna, se estendeu, posteriormente, a fenômenos em línguas não nativas.

Ao longo dos últimos anos, verificamos a preponderância de três modelos perceptuais da fala não nativa no âmbito das pesquisas nacionais: o *Speech Learning Model* [*SLM*, Modelo de Aprendizagem da Fala (FLEGE, 1995)], o *Perceptual Assimilation Model* [*PAM*, Modelo de Assimilação Perceptual (BEST, 1995)] e o *Perceptual Assimilation Model of Second Language Speech Learning* [*PAM-L2*, Modelo de Assimilação Perceptual da Aprendizagem da Fala em Segunda Língua (BEST; TYLER, 2007)]<sup>2</sup>. O *SLM* trata tanto da percepção como da produção de sons não nativos, habilidades balizadas pela experiência linguística do aprendiz, além de variáveis relacionadas à idade de aprendizagem do sistema sonoro alvo e ao período de chegada à comunidade geográfica em que o idioma alvo é falado. O *PAM* volta-se aos chamados “*naïve listeners*”, isto é, ouvintes sem experiência com outro idioma, e parte do

---

<sup>1</sup> Citam-se, aqui, os trabalhos do ano de 1996, especialmente, por terem sido publicados no mesmo periódico científico, representando, portanto, o confronto direto entre ambos os estudiosos. Todavia, cabe lembrar que tal embate teórico iniciou anos antes da referida publicação, a exemplo da proposta de Fowler (1986).

<sup>2</sup> A partir deste ponto, no presente texto, quando mencionarmos os modelos perceptuais elencados, iremos suprimir sua referência autoral para fins de delimitação. Esclarecemos, portanto, que, ao tratarmos do *SLM*, do *PAM* e do *PAM-L2*, estamos nos referindo a Flege (1995) para o primeiro, a Best (1995) para o segundo e a Best e Tyler (2007) para o último.

pressuposto de que a língua materna influencia a percepção da fala não nativa, além de evidenciar as possibilidades de assimilação perceptual de elementos sonoros que não fazem parte do sistema materno dos ouvintes. O foco do *PAM-L2* (que é, por sua vez, uma extensão do *PAM*) recai sobre aprendizes de uma segunda língua (L2) que estão adquirindo o sistema fônico alvo, e seu postulado fundamental é o de que a aprendizagem perceptual é determinada por princípios diferentes daqueles do idioma materno.

No presente artigo, daremos especial atenção ao *PAM-L2* em virtude de sua crescente utilização nos trabalhos em contexto de L2, de seu compromisso com o contexto da percepção de sons não nativos, de seu robusto referencial teórico (tanto no que se refere ao entendimento filosófico da percepção da fala como no que tange à adoção clara de um primitivo fônico, o qual se mostra condizente com nossa perspectiva de aquisição de linguagem) e de suas predições quanto às possibilidades de discriminação categórica acerca das unidades sonoras veiculadas na fala. Entendemos, nesta direção, que seja necessário trazer à tona alguns pontos intrigantes a seu respeito, os quais se instanciam no âmago de sua proposta e, necessariamente, advêm da postura teórica de seu predecessor, o *PAM*. Por conseguinte, trataremos de um questionamento pertinente aos estudos que adotam o *PAM-L2* como fundamentação teórica, seja para a condução da pesquisa em si, seja para a análise e a interpretação dos resultados que obtêm, a saber: *quais são as modificações formais necessárias no modelo para sua aplicação a contextos de percepção fônica de uma língua estrangeira (em detrimento de L2)?*

Para responder a tal questionamento, julgamos relevante resenhar o modelo de percepção fônica que estamos discutindo, o *PAM-L2*, elencar alguns estudos que o utilizaram no âmbito nacional e, por fim, problematizar suas assunções teóricas na medida em que o contexto de percepção em língua estrangeira (LE) se instaura.

## **2. Percepção fônica de elementos sonoros não nativos: o *PAM-L2***

Primeiramente, podemos dizer que o *PAM-L2* (assim como o *PAM*, seu predecessor) advoga pelo gesto articulatório como primitivo fônico e pelo realismo direto como base filosófica. No arcabouço teórico de Best (1995) e de Best e Tyler (2007), tanto a fonética como a fonologia são fundamentadas no domínio dos gestos articulatórios, ou seja, existe um domínio gestual comum tanto para o material fonético como para a representação fonológica dos sons da fala. Duas constatações importantes acerca do modelo são as seguintes: a primeira delas é a de que *fonética* e *fonologia* não são interpretadas como dois domínios informacionais

separados, e a segunda prevê que propriedades fonéticas se encontram alocadas na estrutura fonológica.

Em 2007, Catherine Best e Michael Tyler lançaram uma versão expandida do *PAM*, a qual se ocupa da percepção de sons não nativos por aprendizes de L2. A esta versão estendida do modelo foi atribuído o nome de *PAM-L2*. A tarefa básica do *PAM-L2* é estabelecer que o aprendizado perceptual da L2 seja determinado por princípios não nativos de percepção de fala, considerando elementos comuns e complementares entre ouvintes sem experiência com a segunda língua e aprendizes de tal idioma. O modelo ainda se volta para a relação entre o aprendizado de uma língua e sua informação fonética e fonológica, e sugere brevemente como a experiência linguística – sendo monolíngue ou multilíngue – influencia a percepção fônica.

Na concepção de Best e Tyler (2007), os usuários da L2 estão sempre aprendendo o sistema alvo, sendo que existem diferenças qualitativas entre aprendizes que estão ativamente adquirindo a L2 e aprendizes em processo mais estável. Sob a ótica de Antoniou, Tyler e Best (2012), o principal foco do *PAM-L2* são os aprendizes que estão constantemente aprendendo a L2, além de fatores fundamentais para a formação das categorias fonético-fonológicas desse sistema, como a idade com que o aprendiz inicia sua aquisição, o tempo de residência no país da L2 e o uso da L2 (PEROZZO, 2013).

Destacamos que, no *PAM-L2*, os ouvintes estão sujeitos a perceber diferenças linguísticas tanto em relação ao detalhe fonético de caráter contrastivo (em termos de distintividade entre dois itens lexicais), como também no que tange aos aspectos variáveis (não distintivos) dentro das categorias das vogais e das consoantes. Convém, ainda, ressaltar que os ouvintes podem mostrar sensibilidade perceptual quanto ao aspecto fônico variável tanto na fala nativa como na fala em L2. Portanto, a percepção de elementos fônicos variáveis não só é prevista pelo *PAM-L2*, como também é consistente com seus princípios.

O *PAM-L2* prevê que os fones da L2 são assimilados, primeiramente, em categorias da L1 (língua materna) já existentes, e, então, tornam-se estabelecidos como novas categorias. Tal processo ocorre, em um primeiro momento, no nível fonético, e, na medida em que o vocabulário da L2 se expande, os aprendizes se adequam à fonologia do idioma-alvo, possibilitando que os fones sejam discriminados com base em diferenças de categorias que são lexicalmente relevantes (ANTONIOU; TYLER; BEST, 2012).

Algo interessante e que merece atenção é o fato de haver no *PAM-L2* a existência de uma categoria fonológica e uma categoria fonética, as quais se alinham aos aspectos funcionais

no sistema, assumidos pelos gestos fônicos. A categoria fonológica “diz respeito à informação da fala que é relevante para diferenças lexicais mínimas em uma determinada língua”, e a categoria fonética representa as “relações gestuais invariantes que são sistemática e potencialmente perceptíveis aos ouvintes com percepção mais fina, como alofones posicionais ou diferentes realizações de uma categoria fonológica entre dialetos ou línguas” (BEST; TYLER, 2007, p. 25). Assim sendo, a estrutura fonética serve de base para a organização fonológica, sendo que fonética e fonologia fazem parte de um único sistema, o qual possui um nível mais baixo e um nível mais alto, mas se instanciam, representacionalmente, através do gesto articulatório. Desta maneira, o modelo reconhece tanto a sensibilidade aos detalhes gradientes e físicos do nível fonético como a categorização fonológica, de caráter abstrato, referente à percepção de distinções lexicais, aspectos esses que são abarcados pelo primitivo gestual.

Segundo Best e Tyler (2007), o *PAM-L2* se sustenta sobre quatro possíveis casos de contrastes dos sons da L2 que os aprendizes percebem como segmentos da fala:

- (a) Frente a um contraste entre duas categorias fonológicas da L2, o aprendiz percebe uma delas como sendo equivalente a determinada categoria fonológica da L1. Quanto à outra categoria a ser assimilada, é possível que ela seja alocada de maneira adequada no sistema da L1, ou seja, ela representa um bom exemplar para outra categoria da L1, ou, diferentemente, é possível que ela não seja categorizada. Dessa forma, o aprendiz praticamente não teria dificuldades em discriminar, minimamente, os segmentos-alvo que contrastam em pares mínimos.
- (b) Frente a um contraste entre duas categorias fonológicas da L2, o aprendiz percebe as duas como pertencentes à mesma categoria fonológica da L1. No entanto, uma das categorias é percebida como mais desviante do que a outra. Espera-se, portanto, que os aprendizes sejam capazes de discriminar relativamente bem estes dois fones da L2, mas não tão bem como a discriminação que poderia haver caso eles conseguissem classificar os segmentos em categorias diferentes.
- (c) Frente a um contraste entre duas categorias fonológicas da L2, o aprendiz percebe as duas como pertencentes à mesma categoria fonológica da L1. Todavia, ambas são classificadas como exemplares igualmente bons ou igualmente desviantes daquela categoria. O aprendiz terá problemas ao discriminar estes dois fones, que seriam

assimilados tanto fonética como fonologicamente a uma única categoria da L1, e duas palavras da L2 minimamente contrastantes seriam percebidas como homófonas.

- (d) Se o ouvinte não conseguir encontrar uma categoria da L1 para alocar dois segmentos contrastantes da L2, então não haverá assimilação fonológica da L2 para a L1. Neste sentido, o ouvinte reconhece uma mistura de similaridades entre os segmentos e não é capaz de categorizá-los devidamente.

Best (1995) e Best e Tyler (2007), a fim de situar filosoficamente seus modelos de percepção da fala, lançam mão de uma teoria perceptual defendida por James Gibson (um psicólogo experimentalista americano), denominada *Realismo Direto*. Basicamente, o livro de Gibson (1986), *The Ecological Approach to Visual Perception*, é um tratado sobre a percepção visual (em detrimento de outros sentidos), e o autor se preocupa em descrever a realidade circundante ao percebedor, utilizando-se de uma nomenclatura própria e procurando evidenciar supostas lacunas nas teorias mais tradicionais da ótica.

Nas primeiras páginas de sua obra, Gibson (1986) deixa claro que seu objeto de estudo refere-se ao nível ecológico, ao habitat dos animais e dos homens, cujos sistemas perceptuais não são capazes de detectar extremos, como átomos ou galáxias. Para o psicólogo, devemos considerar como percebemos o ambiente, ou seja, como apreendemos as mesmas coisas que nossos ancestrais humanos apreenderam antes mesmo de saberem sobre partículas atômicas e sistemas planetários. Por conseguinte, Gibson (1986) preocupa-se com a percepção direta e não com aquela indireta, a qual é obtida através de microscópios, telescópios, fotografias e figuras, e muito menos com o tipo de apreensão (de conhecimento) obtida pela fala e pela escrita (p. 10). Esta é a primeira acepção do termo “direto” na obra de Gibson, que *não prevê qualquer aparato entre o percebedor e o ambiente*. A sua segunda acepção diz respeito ao fato de os percebedores não necessitarem de uma quantidade excessiva de aprendizagem (p. 143), ou seja, *eventos cognitivos que se relacionam a inferências e a representações mentais são ignorados na perspectiva direta da percepção*.

Um conceito notável na obra de Gibson (1986) é o de “*affordances*”, ou concessões<sup>3</sup>, que se refere, *grosso modo*, àquilo que o meio nos concede, ou seja, nos oferece. Para o

---

<sup>3</sup> Tradução nossa.

estudioso, uma concessão é invariante e se faz constante desde o início da evolução animal, além de implicar a complementaridade entre animal e ambiente. O fogo, por exemplo, concede calor em uma noite fria, mas também oferece a possibilidade de alguém se queimar; um objeto que se aproxima concede a possibilidade de colisão. Nessa direção, segundo Gibson (1986), para nossos ancestrais, um coelho que se aproximasse concederia a possibilidade de se alimentarem, enquanto que um tigre que se aproximasse concederia a possibilidade de serem devorados. Conforme explica o psicólogo, as concessões não são, assim, boas ou ruins, positivas ou negativas, elas existem e se prestam às ações que permeiam o animal e seu ambiente.

Gibson (1986) assume que nós não percebemos o tempo, mas, sim, processos, mudanças e sequências. Segundo o estudioso, os eventos (episódios alinhados ao tempo) encontram-se dentro de eventos e formas encontram-se dentro de formas. As unidades naturais do ambiente terrestre e dos eventos não devem ser confundidas com as unidades métricas de espaço e tempo, que são, por excelência, convencionalizadas a partir de uma construção humana. Ao se aproximar de sua defesa à ecologia da percepção, Gibson (1986) elenca que o ambiente normalmente manifesta coisas que persistem e coisas que não persistem, além de pontuar que algumas características são invariantes e outras são variantes. Um espaço totalmente invariante, imutável nas suas partes e imóvel, seria completamente rígido e, obviamente, deixaria de ser um ambiente (de fato, não haveria nem animais e nem plantas). Por outro lado, “um espaço inteiramente variante e mutável quanto a todas as suas partes não seria um ambiente” (GIBSON, 1986, p. 14), uma vez que elementos de um ambiente precisam se perpetuar ao longo das cadeias temporal e espacial.

A essência de um ambiente ecológico, segundo os preceitos gibsonianos, é a de que ele circunda um indivíduo e, em um ambiente, há diversos pontos de observação. Estes, por sua vez, são constituídos pelos caminhos disponíveis para locomoção em um meio, e todos os habitantes de um ambiente têm igualmente a chance de explorá-lo. Nesse sentido, o ambiente circunda todos os observadores do mesmo modo que circunda um único observador. É justamente a ideia de um sujeito exploratório, capaz de acessar os objetos reais do ambiente, que dá sustentação aos modelos *PAM* (BEST, 1995) e *PAM-L2* (BEST; TYLER, 2007). Segundo Best (1995), a estimulação do ambiente fornece recursos informacionais diretos, ricos e confiáveis quanto ao mundo, sendo completamente dispensáveis os mecanismos cognitivos de representação mental e inferência. Apoiando-se sobre o aporte gibsoniano, Best (1995)

entende que a aprendizagem perceptual envolve uma sintonização do percebedor com o sistema-alvo de sons para a detecção de invariantes de alta ordem (as concessões), implicando diferenças funcionais no componente fonológico em desenvolvimento.

Em linhas gerais, segundo a visão de Best (1995) acerca da abordagem gibsoniana, “o percebedor apreende o objeto perceptual diretamente e não meramente um representante ou ‘substituto’ a partir do qual o objeto deva ser inferido” (BEST, 1995, p. 173). Sob o prisma de Best (1995) e de Best e Tyler (2007), perceber é extrair informação direta do mundo acerca de seus componentes – objetos, pessoas, relações, etc. – no tempo e no espaço, sem a mediação de um conhecimento inato ou associações mentais adquiridas. Nesse quesito, a pesquisadora faz a seguinte colocação:

Mecanismos cognitivos especiais para lidar com representações mentais e inferências indiretas não são necessários, porque os estímulos fornecem uma fonte de informação direta, rica e confiável sobre o mundo, e porque os percebedores se ocupam de seus sistemas perceptuais integrados na exploração ativa do mundo. Nesta visão, a aprendizagem perceptual envolve um maior ajuste para se detectarem invariantes de alta ordem disponíveis na cadeia de informação, em vez de mudanças em representações mentais e processos inferenciais. (BEST, 1995, p. 175)

No entanto, não são somente Best (1995) e Best e Tyler (2007) que se apropriam do construto gibsoniano para designar o que eles entendem por percepção<sup>4</sup>. Antes mesmo de os modelos *PAM* e *PAM-L2* de fato existirem, Fowler (1986) já fazia uma adaptação da teoria de Gibson (1986) para a sua teoria de percepção da fala nativa. Ou seja, verificamos a transposição de um objeto *somente visual* para um objeto que é visual, mas, em larga medida, também é *acústico*, de modo que estes dois meios remetam ao aparato articulatório. Ressaltamos, nesse âmbito, que a Fonologia Articulatória (ou Fonologia Gestual) de Browman e Goldstein (1989, 1992) e de Fowler (1986, 1996) prevê o aporte articulatório para a concepção da unidade gestual – além disso, o posicionamento dos referidos autores vai ao encontro de uma abordagem multimodal para a fala, em que variáveis de diversas naturezas (acústica, visual, etc.) atuam para a apropriação do sistema fônico. De acordo com Fowler (1996), sua teoria de percepção é uma teoria em que os gestos fonológicos são as ações públicas do trato vocal que causam a

---

<sup>4</sup> Best (1995) afirma explicitamente que, em função de ela seguir Gibson (1986), inferências e representações mentais são desnecessárias ao evento perceptual. No entanto, a autora prevê a existência de segmentos fonológicos e fonéticos, que estão intimamente ligados à noção representacional. Além disso, sua referência massiva ao segmento, dentro de uma proposta gestual, sugere certa inconsistência com seu posicionamento.

estrutura (elemento que condiciona a percepção) nos sinais acústicos da fala. Logo, a perspectiva adotada por Fowler (1986) – e, por conseguinte, a de Best (1995), principalmente – já é uma releitura da abordagem ecológica de Gibson (1986), adaptada ao âmbito dos sons da fala.

Aos moldes de Gibson (1986) e de Fowler (1986), Best (1995) e Best e Tyler (2007) notam que a relação estabelecida entre o sujeito percebedor e o ambiente se torna a condição necessária e suficiente para que o conhecimento, que ocorre via percepção, seja instaurado a partir da detecção das concessões que o ambiente fornece. Tal constatação leva os autores, portanto, a fazer a ressalva de que o empreendimento teórico que realizam opera primordialmente em situações de percepção de L2 em detrimento de LE. Para os autores, o cenário de LE encontra-se nitidamente abaixo do ideal no que se refere à ecologia da aprendizagem de uma língua, pois oferece um contexto empobrecido para a aprendizagem da língua alvo. Além disso, ocorre em um ambiente em que a L1 é muito influente sobre a língua alvo, a qual, geralmente, não se estende para fora da sala de aula. Nesse sentido, o quadro em que se insere a LE frequentemente emprega instrução formal e informações gramaticais, enquanto que as situações de conversação são muito minimizadas. Os autores também mencionam que os momentos de aprendizagem de uma LE constituem um espaço em que a língua alvo é frequentemente veiculada por professores com sotaque materno, apresentando um modelo variável ou incorreto dos detalhes fonéticos a serem adquiridos. Um ambiente de aprendizagem de L2, para Best e Tyler (2007), não contemplaria tais características e seria suficientemente robusto para garantir a aprendizagem perceptual da língua alvo.

Convém deixar claro que, conforme os próprios Best e Tyler (2007) salientam, a percepção fônica em um contexto em que a língua alvo é estrangeira – e não uma segunda língua<sup>5</sup> – ainda necessita de muita investigação, e o modelo por eles desenvolvido aplica-se fundamentalmente ao cenário de L2. Esta restrição não somente é explicitada pelos autores, mas também se caracteriza como condição imprescindível para a adoção do modelo. Oportunamente, a próxima seção irá tratar dos estudos realizados no Brasil que se utilizaram do aporte teórico do *PAM-L2*, ainda que para explicar a percepção fônica em contexto estrangeiro.

---

<sup>5</sup> Diferenças entre L2 e LE serão explicitadas na seção 3.

### 3. A utilização do *PAM-L2* em contextos de percepção fônica de LE

Os estudos sobre percepção dos sons de L2 no Brasil envolvem, principalmente, o inglês como língua alvo e se utilizam, basicamente, dos modelos *SLM*, *PAM*, e *PAM-L2*. Na alçada do *SLM*, apontamos os estudos de Koerich (2002), Kluge (2004, 2009), Rauber (2006), Reis (2006) e Moore (2008). No que tange ao *PAM*, destacamos a pesquisa de Reis (2010)<sup>6</sup> e, quanto aos trabalhos que utilizaram o *PAM-L2*, elencamos os de Bettoni-Techio (2008), Gutierrez e Battisti (2012), Perozzo (2013), Feiden, Alves e Finger (2014) e Feiden *et al.* (2016, no prelo).

Os estudos acima citados foram encontrados a partir de dois diferentes mecanismos de busca. O primeiro deles, de caráter documental, caracterizou-se por um levantamento realizado por Silveira (2010), a partir do qual selecionamos trabalhos que envolviam a utilização dos três modelos mencionados para explicar ou situar os resultados obtidos. Como tal repertório contou com a descrição de estudos conduzidos até 2009, procedemos ao exame dos currículos Lattes daqueles pesquisadores que, além de terem sido contemplados no levantamento de Silveira (2010), haviam começado a orientar trabalhos, em nível de pós-graduação, que também tivessem o compromisso de tratar seus dados à luz dos modelos perceptuais referidos. No entanto, não fomos bem sucedidos nesta segunda parte, pois não constavam, na lista de trabalhos orientados, quaisquer estudos que tivessem relação com os modelos perceptuais em que tínhamos interesse. Dessa forma, passamos ao segundo mecanismo de busca, de cunho exploratório, em que rastreamos estudos de pesquisadores que tiveram contato com nosso grupo de pesquisa, através de disciplinas oferecidas no Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul ou por meio de trabalhos realizados em conjunto.

Focaremos, nesta seção, apenas nos trabalhos que se utilizaram do *PAM-L2* enquanto referencial teórico para a percepção de sons não nativos e, também, que se valeram de suas previsões de assimilação para interpretar os resultados obtidos. Nossa linha de argumentação tem por finalidade apresentar cinco investigações interfonológicas que fazem uso do *PAM-L2* para compará-las entre si e evidenciar suas características comuns. Veremos que, apesar de ser um construto teórico voltado à L2, o modelo está sendo utilizado em pesquisas cujo foco de investigação é a percepção fônica de uma LE, o que exige dele futuras adaptações em nível estrutural, tarefa essa cujo primeiro passo damos no presente artigo. Além disso, resenharemos

---

<sup>6</sup> Reis (2010) menciona que os resultados encontrados em sua pesquisa foram analisados também sob a ótica do *PAM-L2* (BEST; TYLER, 2007).

os estudos mencionados porque eles serão, também, motivação para justificar a proposta de alterações ao *PAM-L2*, o que será feito ainda nesta seção.

O primeiro estudo a ser citado é o de Bettoni-Techio (2008), o qual buscou compreender como brasileiros viriam a perceber e produzir<sup>7</sup> a fricativa alveolar /s/ presente em [s]-clusters em posição de ataque inicial relativos a palavras do inglês, quando tal segmento era precedido por palavras cuja consoante em coda era vozeada e não vozeada. Além disso, a pesquisadora se interessava em verificar os efeitos de treinamento na percepção e na produção de tais sequências sonoras. Sua amostra de sujeitos foi composta por 23 alunos de inglês, adultos e crianças com minimamente 200 horas de instrução prévia no idioma, e a eles foram administradas tarefas de identificação, discriminação, fala semiespontânea e leitura. Os resultados analisados apontaram maiores índices de acuidade na identificação, na discriminação e na produção das sequências sonoras, tanto para as formas testadas quanto para novas formas. Adicionalmente, altos índices de acuidade nas tarefas foram observados na sessão de retenção, que ocorreu oito meses após o treinamento.

O segundo trabalho de que tratamos é o de Gutierrez e Battisti (2012), cujo objetivo era analisar a percepção da consoante nasal velar do inglês em posição de coda silábica final por aprendizes brasileiros do idioma. Para tanto, participaram da pesquisa 9 acadêmicos de uma instituição de ensino superior no sul do Brasil, com idade entre 19 e 25 anos, que estudavam a língua inglesa há aproximadamente 6 anos. Os alunos foram testados quanto ao seu nível de proficiência em inglês e, segundo o *Oxford Placement Test* (ALLAN, 2004), foram classificados como aprendizes de nível avançado. Para a condução do estudo, as pesquisadoras aplicaram aos acadêmicos um teste de discriminação categórica envolvendo as palavras alvo. As autoras apontam que os resultados obtidos indicam um bom nível de acuidade na discriminação perceptual entre a consoante nasal velar e a consoante nasal alveolar. Neste sentido, as pesquisadoras ressaltam que, na maioria das respostas fornecidas pelos alunos, a consoante nasal alveolar foi percebida, de fato, como sendo alveolar, e a nasal velar percebida, de fato, como velar.

O terceiro estudo que mencionamos é o de Perozzo (2013), que investigou como brasileiros percebiam o ponto de articulação de oclusivas não vozeadas sem soltura audível (a partir da informação coarticulatória com a vogal precedente) em posição de coda final de sílabas

---

<sup>7</sup> O modelo de Best e Tyler (2007), ao contrário do de Flege (1995), é um modelo perceptual, não de percepção-produção. Extensões à produção, assim como fez Bettoni-Techio (2008), ficam a critério da autora.

da língua inglesa. A pesquisa contou com 17 acadêmicos do curso de graduação em Letras de uma universidade no sul do Brasil, com nível de proficiência estipulado entre básico e intermediário, de acordo com o *Oxford Placement Test* (ALLAN, 2004). Para medir a acurácia quanto ao ponto de articulação das consoantes propostas, foram aplicados um teste de identificação e um teste de discriminação. Também com o interesse de verificar se a instrução explícita sobre o fenômeno da não soltura de oclusivas teria papel sobre a acuidade na percepção dos pontos de articulação das consoantes finais, os acadêmicos foram alocados em dois grupos, experimental e de controle. Os resultados sugeriram maiores índices de acuidade perceptual quanto aos pontos de articulação labial e velar. Além disso, o grupo experimental obteve maiores índices de acuidade quando comparado ao grupo controle em ambas as tarefas, e os acadêmicos de nível intermediário apresentaram maiores índices de acuidade do que os acadêmicos de nível básico.

O quarto estudo a que nos remetemos é o de Feiden, Alves e Finger (2014), cujo objetivo foi constatar os efeitos da anterioridade e da altura da língua sobre a percepção de argentinos quanto às vogais médias-altas e médias-baixas do português brasileiro. Os participantes da pesquisa foram 16 falantes de espanhol rioplatense, com idade entre 16 e 18 anos, estudantes de ensino médio que cursavam uma disciplina de língua portuguesa há, pelo menos, 2 anos em uma escola na Argentina<sup>8</sup>. Os alunos foram submetidos a um teste de identificação perceptual contendo pares mínimos do português, os quais eram dissilábicos e apresentavam padrão acentual trocaico. Os resultados mostraram que, tanto nas vogais anteriores quanto nas posteriores, os participantes obtiveram índice mais elevado de acerto nos itens que apresentavam vogais médias altas em comparação com os itens que apresentavam vogais médias baixas. Como conclusão, os pesquisadores observaram que os participantes tiveram dificuldades na identificação das vogais médias do português, porém advertem que os referidos aprendizes se encontram em um estágio de formação das novas categorias fonético-fonológicas da língua alvo, apesar da dificuldade em delimitar as fronteiras acústico-articulatórias que caracterizam tais categorias.

---

<sup>8</sup> Embora tais aprendizes tenham sido testados no Brasil, ou seja, em um ambiente de LE, eles aprenderam português na Argentina. Entretanto, os participantes eram intercambistas aqui no país, e estudaram português durante a sua estada. Tal fato nos impede de classificá-los, com propriedade, como aprendizes de português como LE ou como L2. Entretanto, inclinamo-nos a referi-los como aprendizes de português como LE em função de terem iniciado o estudo do português em seu país de origem e não no contexto nacional, e por terem recebido um maior número de horas de instrução no cenário estrangeiro.

Por fim, o quinto trabalho que referimos é o de Feiden *et al.* (2016, no prelo), que se propôs a analisar os índices de discriminação categórica entre as vogais médias-altas e médias-baixas do português brasileiro por falantes nativos de espanhol argentino. Este estudo é semelhante ao de Feiden, Alves e Finger (2014) por incorporar os mesmos estímulos auditivos e integrar os mesmos participantes, porém difere drasticamente no que concerne à natureza da tarefa: Feiden, Alves e Finger (2014) propuseram um teste de identificação perceptual, enquanto que Feiden *et al.* (2016, no prelo) administraram um teste de discriminação categórica, já mencionado. Neste estudo, os pesquisadores observaram que os alunos testados se mostraram capazes de discriminar o segmento anterior médio-alto do anterior médio-baixo do português, mas o mesmo não ocorreu para o contraste de altura dos segmentos análogos posteriores, indicando clara conjunção de tais elementos em uma única categoria fônica.

Para além de empregarem o mesmo modelo perceptual à análise e à interpretação dos resultados e se enquadrarem em uma metodologia experimental, os trabalhos destacados compartilham de uma situação comum extremamente importante para o propósito deste artigo: *a amostra utilizada em todos eles se caracteriza por indivíduos que percebem elementos fônicos de uma LE*. Nesse mérito, notamos que a motivação para alocá-los sob o rótulo de aprendizes de língua *estrangeira* (em detrimento de aprendizes de *segunda* língua) resulta de fatores cognitivo-ambientais, referentes ao domínio de uso do idioma, e pedagógicos, os quais se encontram interconectados.

Como toda a atribuição de “rótulos” pode ser escorregadia (ainda mais no que tange a aspectos humanísticos), justificaremos, brevemente, com base em Gass e Selinker (2008), nossa opção por tratar os participantes das pesquisas enumeradas como percebedores de sons de *LE*. Segundo Gass e Selinker (2008), as situações de aprendizagem de *LE*, e, portanto, de percepção fônica de *LE*, se diferenciam daquelas referentes à *L2* na medida em que ocorrem no ambiente em que a *L1* é falada (como, por exemplo, brasileiros aprendendo inglês no Brasil). Pedagogicamente, para os autores, é bastante comum em tal cenário haver instrução formal acerca do idioma a se adquirir, enquanto que, em um ambiente de aprendizagem de *L2*, geralmente, não há apelo a meios instrucionais. Tal definição vai ao encontro das características elencadas por Best e Tyler acerca das diferenças entre a aprendizagem de uma *L2* e de uma *LE*, reiterando o contraste entre o tipo de exposição informacional no caso de aprendizes de *L2* e aprendizes de *LE*. Em nosso entendimento, se os aprendizes estão inseridos em um ambiente onde eles têm contato com a língua alvo em um amplo espectro de situações, a aprendizagem

perceptual da fala deverá, por hipótese, ser otimizada. Oportunidades mais limitadas de contato com o idioma alvo implicam menor participação em práticas situacionais que o envolvem, e disso decorre menor acesso a informações oriundas de comunicação autêntica.

Dadas essas condições, os estudos relatados estão, de fato, se utilizando do *PAM-L2* para dar sustentação às suas pesquisas, bem como para analisar, em termos de percepção, os fatos fônicos que encontram. Verificamos, pois, que tais estudos alocam-se no contexto de percepção fônica em LE, mas fazem uso de um modelo voltado para a percepção fônica de uma L2, e, conforme explanam os próprios autores, o *PAM-L2* não se presta para a percepção fônica no contexto de LE. Dessa forma, vislumbramos duas opções: abandonar o modelo, no caso da percepção fônica em LE, ou adaptá-lo. Optamos, todavia, pela última alternativa. Sugerimos, portanto, que seja elementar a necessidade de adaptar o *PAM-L2* ao contexto de percepção de LE (que configura o caso da grande maioria dos aprendizes de idiomas no cenário nacional), no sentido de levar em conta a realidade dos aprendizes de línguas estrangeiras que residem no Brasil ou em países em que não haja uma segunda língua que ofereça uma quantidade similar de *input* em comparação com o da L1. Tais alterações dizem respeito aos seguintes aspectos:

- A percepção e a aprendizagem perceptual de uma LE deve ser encarada como um processo *cognitivo*, o qual envolve o processamento da informação veiculada na fala, e não como a *apropriação de ações* sem relação com mecanismos inferenciais e representações mentais.
- A unidade fônica da percepção deve ser o *gesto acústico-articulatório* cunhado por Albano (2001), uma vez que tal primitivo se relaciona e interage com outras operações linguísticas que compõem a língua a ser compreendida, e não o *gesto articulatório* de Browman e Goldstein (1989, 1992) e de Fowler (1986, 1996), que não possui compromisso com outros níveis da gramática.
- A noção de *ambiente ecológico* deve ser desvinculada do contexto de percepção em sala de LE, em virtude de que tal espaço carece das condições necessárias para configurar um ambiente ecológico de fato.
- O realismo *indireto* perceptual (a ser tratado na seção 4.2), em detrimento de sua contraparte *direta*, deve ser a teoria filosófica que embasa a percepção dos sons de uma LE, não somente por conseguir captar as relações fônicas entre L1 e LE, mas também

por se tratar de uma unidade referente ao signo linguístico, que é construído histórica e socialmente, figurando como uma entidade convencionada.

É no tocante a tais reflexões que nos ocupamos da próxima seção, com vistas a ponderar sobre as alterações necessárias ao modelo para que ele dê conta, em alguma medida, da percepção fônica em LE.

#### **4. *PAM-L2* e sua demanda na LE: possíveis alterações no modelo**

Nosso intuito ao discorrer sobre o *PAM-L2*, nesta seção, é o de prover mais esclarecimentos sobre as possíveis alterações no modelo para que ele venha a abarcar os fatos perceptuais em contexto de percepção fônica estrangeira, de acordo com os aspectos motivadores que justificam tal exercício teórico, expostos na seção anterior. Destacamos que não seremos exaustivos quanto às sugestões, mas tentaremos contemplar seus aspectos mais fundamentais, em função de que nossas ponderações configuram um primeiro passo na tentativa de adaptação do modelo, com o objetivo de que ensejem discussões subsequentes acerca de nossa reflexão.

Consideramos que, se o *PAM-L2* se resumisse apenas às predições de assimilação perceptual descritas na seção 2, em primeira análise, não necessitaria mais do que uma reflexão teórica que resultasse em acréscimos a casos de assimilação não contemplados pelo modelo original, independentemente de diferenciarmos LE de L2. Entretanto, não é só de predições que o modelo é formado: há todo um embasamento teórico que o sustenta (no sentido de conceber o desenvolvimento de conhecimento linguístico), o qual pode ser discutido sobre o ponto de vista fonológico, em que o modelo assume como primitivo fônico o gesto articulatório, e sobre a perspectiva filosófica, que tem a ver com a abordagem ecológica da percepção e com o realismo direto de James Gibson. Problematicamos, assim, nas duas subseções a seguir, o aspecto fonológico da percepção em LE e a noção de ambiente ecológico e realismo direto que embasam o *PAM-L2*, e sugerimos aportes alternativos para tratar do fenômeno perceptual em contextos de LE.

##### **4.1. O gesto articulatório e a gramática fônica do aprendiz/percebador de LE**

A justificativa para realizarmos uma discussão sobre o aspecto fonológico da percepção em LE reside no fato de que, ao perceber os elementos sonoros da LE a que está exposto, o

percebedor/aprendiz os relaciona àqueles de sua L1 (FLEGE, 1995; BEST; TYLER, 2007). Em nossa visão, diferentemente do que concebem Best e Tyler (2007), este processo diz respeito a um empreendimento mental<sup>9</sup> e motor. Desse modo, quando o conhecimento fonológico materno e o sistema fonológico alvo estão em contato, a *habilidade mental* (portanto, representacional) de categorizar os novos sons para distinções lexicais e a *habilidade motora*<sup>10</sup> (portanto, física) do aprendiz de perceber/produzir rotinas articulatórias operam absolutamente em conjunto.

A título de esclarecimento, julgamos pertinente definir o que concebemos pelos termos *habilidade mental* e *representação mental*, pois se farão úteis ao longo de nossa exposição. *Habilidade mental* ou *atividade mental*, segundo Gazzaniga *et al.* (2012), tem a ver com a mente em ação, isto é, diz respeito ao processamento das informações. As experiências de percepção que um indivíduo detém enquanto interage com o mundo, assim como memórias, pensamentos e sentimentos, são exemplos da mente em ação. Portanto, a atividade ou a habilidade mental resulta de processos biológicos (as ações das células nervosas e suas reações químicas associadas) no cérebro. *Representação mental* refere-se a um símbolo cognitivo interno que corresponde à realidade externa e pode ser *analógica*, quando trazemos à mente a imagem de um objeto, por exemplo, ou *simbólica*, que exprime o conceito ou nosso conhecimento sobre o objeto e não sobre suas propriedades físicas<sup>11</sup> (GAZZANIGA *et al.*, 2012). Evocamos, portanto, os postulados de Gazzaniga *et al.* (2012) e, para efeitos linguísticos, acreditamos que as representações mentais dos indivíduos sejam adaptativas, complexas e não lineares (DE BOT; LOWIE; VERSPOOR, 2007; BECKNER; *et al.*, 2009), em oposição à noção estática, cartesiana e serial de representação mental, veiculada fortemente pela tradição gerativa.

No arcabouço do modelo fonológico conhecido como Fonologia Articulatória, conforme propõem Browman e Goldstein (1989, 1992), os gestos articulatórios vêm a ser as unidades básicas de contraste fonológico nas línguas, figurando como caracterizações abstratas de eventos articulatórios físicos de tempo intrínseco<sup>12</sup>. Para os estudiosos, gestos articulatórios

---

<sup>9</sup> Seguimos a definição de mente de Kosslyn e Koenig (1995, p. 4), de acordo com a qual “a mente é aquilo que o cérebro faz”. Dessa maneira, segundo postulam Gazzaniga *et al.* (2012, p. 5), “o cérebro físico habilita a mente”, e ambos são inseparáveis.

<sup>10</sup> O termo *motor* está sendo usado aqui como sinônimo de *mecânico* e não está vinculado aos preceitos da Teoria Motora da Percepção da Fala (LIBERMAN; MATTINGLY, 1985).

<sup>11</sup> Para fins de argumentação, seguiremos a contraparte *simbólica* da representação mental.

<sup>12</sup> Por tempo intrínseco, entenda-se a incorporação da dimensão *tempo* na especificação fonológica de um segmento. Para uma explicação mais elaborada acerca de uma unidade sonora com tempo intrínseco, recomendamos a leitura de Fowler (1980).

e organizações gestuais podem captar informações tanto categóricas quanto gradientes na cadeia da fala.

Goldstein e Fowler (2003), trabalhando no desenvolvimento da Fonologia Articulatória, consideram que as formas linguísticas são os meios que as línguas oferecem para que a comunicação entre as pessoas se faça possível e, por isso, advogam que tais formas são tipos de ações públicas e não exclusivamente categorias mentais, isoladas da realidade física, como apresenta a maioria das teorias fonológicas, de percepção e de produção. Os pesquisadores entendem que as unidades fonológicas são abstratas com respeito às variáveis articulatórias e acústicas que são tipicamente medidas, mas não tão abstratas a ponto de ignorarem a área vocal e alocarem-se exclusivamente na mente. Interessantemente, sob a perspectiva gestual, “as unidades fonológicas mais básicas devem ser discretas e recombináveis” (GOLDSTEIN; FOWLER, 2003, p. 189), além de deverem se utilizar de uma moeda comum entre percepção e produção.

Algo que merece particular atenção é o fato de a Fonologia Articulatória não ter sido criada com a pretensão de explicar o processo de aquisição ou desenvolvimento do sistema sonoro de uma L2 ou LE, repousando exclusivamente sobre a L1. Conforme podemos observar nos textos iniciais do modelo (BROWMAN; GOLDSTEIN, 1989, 1992), a Fonologia Articulatória se reserva a elucidar relações gestuais veiculadas na fala e que têm um compromisso com distinções lexicais, mas nada afirma explicitamente sobre a *adição* de um novo sistema àquele já existente:

O objetivo de uma gramática fonológica é dar conta do conhecimento (implícito) *nativo* dos falantes com relação à estrutura fonológica e às regularidades em uma determinada língua, incluindo o inventário de unidades lexicalmente contrastivas, restrições quanto a formas fonológicas e alternâncias sistemáticas no que se refere a formas lexicais resultantes de combinação morfológica e adjunção em contextos prosódicos específicos. [(GOLDSTEIN; FOWLER, 2003, p. 170) grifo nosso]

Contudo, o modelo pode ser utilizado para ilustrar a percepção e a produção da fala em L2 ou LE (segundo faz, por exemplo, o *PAM-L2*), uma vez que as informações sonoras linguísticas se apoiam sobre unidades gestuais. Neste sentido, podemos fazer um paralelo entre o desenvolvimento fonológico estrangeiro e aquele que ocorre em estágios iniciais da L1, levando em conta os postulados de Goldstein e Fowler (2003). De acordo com os pesquisadores, para que a criança associe seu trato vocal e os movimentos de sua face àqueles do modelo (um

adulto), a representação da fala deve refletir o que é comum entre a ação do adulto percebida visualmente e a ação gerada pela criança, a qual é percebida proprioceptivamente. De igual maneira, acreditamos que, em um ambiente de L2 ou LE, o aprendiz (que é percebedor e produtor de sons da fala) deve ter como representação o reflexo daquilo que é percebido visual e acusticamente com base no modelo e daquilo que é percebido via propriocepção pelo próprio aprendiz.

Concordamos com Browman e Goldstein (1989, 1992) acerca das faces abstrata e concreta do gesto articulatório e concordamos, também, com Goldstein e Fowler (2003) sobre a necessidade de uma teoria fonológica levar em conta aspectos mais concretos da cadeia sonora, como a gradiência apresentada nos sons da fala. Compartilhamos, também, da postura de Albano (2001) no que compete ao fato de as variáveis do trato vocal, responsáveis por definir determinado gesto articulatório, estarem duplamente relacionadas aos articuladores: para a autora, na esfera abstrata, as variáveis encontram-se funcionalmente unidas sob o articulador envolvido mais diretamente com as constrições que executam, e, na esfera concreta, tais variáveis oportunizam o movimento do articulador (ou seja, a realização de uma tarefa mecânica) em questão.

O que nos parece muito oportuno na descrição gestual delineada por Albano (2001) é o seu compromisso de relacionar operações gestuais a operações linguísticas outras que ocorrem na gramática como um todo. Tal conexão fica clara quando a pesquisadora explana que o realinhamento e o redimensionamento dos gestos articulatórios não somente dão conta da lexicalização dialetal ou estilística daqueles processos que, em princípio, seriam esporádicos na fala corrente, mas também atuam como o meio através do qual as regularidades fônicas adentram “níveis mais profundos da gramática” (ALBANO, 2001, p. 102). Evidências de cunho morfofonológico (entre outras) trazidas por Albano (2001) compreendem parte dos exemplos que dão subsídio a tal constatação. Vemos, portanto, ao longo da exposição feita por Albano (2001), a tentativa de prever relações entre unidades de natureza fônica e outras unidades gramaticais, como o morfema, no caso da formação do plural em alguns itens lexicais do português, e como o sintagma, nos casos de sândi externo.

Embora Goldstein e Fowler (2003) admitam que haja alguma abstração no que se refere ao gesto articulatório, sua argumentação é conduzida, em boa parte, a salientar com enorme robustez o caráter físico de tal unidade fônica. É nesse sentido que julgamos mais conveniente a visão de Albano (2001), em que as faces abstrata e física do gesto articulatório parecem estar

mais equilibradas, além de tal unidade estar conectada e interagir com outros aspectos da gramática, também de cunho abstrato. Ora, em nosso ponto de vista, ao contrário do que referem os construtos de Gibson (1986), intriga imaginar que, em um ambiente de percepção de sons de uma LE, o percebedor tenha que se valer com tamanho afinco de ações puramente<sup>13</sup>. Destacamos que, de maneira nenhuma, estamos deixando de lado a contraparte física do gesto articulatório (afinal, ela é de fundamental importância na percepção e na produção da fala), mas nos cabe acrescentar que, frente ao sistema linguístico da L1, que já está formado e internalizado, é indispensável que o percebedor atue com base em inferências e representações mentais<sup>14</sup>.

Em consonância com Flege (1995) e Best (1995), parece-nos irrefutável a afirmação de que os indivíduos, enquanto percebedores de sons não nativos, recorram ao seu conhecimento prévio da L1 no momento em que decidem se determinado som da fala corresponde ou não a uma categoria nativa. No entanto, chamamos a atenção para o fato de que, no momento em que o percebedor reconhece distinções lexicais veiculadas através da discriminação categórica – atribuindo, assim, significado linguístico à massa sonora a que tem acesso –, ele está resgatando o inventário lexical da língua alvo, cumprindo um exercício *cognitivo* por excelência. Perceber diferenças sonoras em uma LE, em nosso entendimento, não consiste apenas em se apropriar de uma tarefa articulatória, mas compreende a relação cognitiva entre unidades sonoras prévias (L1) e novas (LE), cujo resultado pode ser visto em rotinas motoras<sup>15</sup>.

Finalmente, reiteramos que o gesto articulatório deva ser o primitivo sonoro que permeia a percepção de novos elementos fônicos. Além disso, com base em Albano (2001), acreditamos que, de fato, existe uma gramática fônica, abstrata e representacional, a qual opera com o gesto articulatório. Este, por sua vez, corresponde a uma unidade de contrastes lexicais que se apresenta tanto de maneira abstrata (mental) como de maneira física (motora). A gramática fônica do aprendiz/percebedor de LE, em nossa visão, se configura, pelo menos no que compete à discriminação categórica, como um espaço de constantes relações entre unidades sonoras

---

<sup>13</sup> Prenunciamos um dos motivos por que discordamos da adoção da abordagem realista direta à percepção sonora em LE. Outros argumentos serão fornecidos na seção 4.2.

<sup>14</sup> Este argumento também será utilizado na seção 4.2 para trazer à baila o construto do realismo indireto, uma perspectiva antagônica ao realismo direto.

<sup>15</sup> Explicitamos, portanto, o fato de que os ouvintes percebem os sons da LE com base em seu inventário fônico da L1, e aproveitamos para destacar, neste âmbito, que o percebedor/aprendiz deve aprender a temporalizar os gestos articulatórios que já possui, habilidade referente à orquestração destes gestos (SILVA, 2014; ZIMMER; ALVES, 2012).

prévias e novas, as quais podem ser mais ou menos estáveis a depender do domínio de uso da LE.

#### 4.2 A percepção fônica não nativa e os limites de sua ecologia

Acreditamos que a fala, em seu sentido tradicional, é parte constitutiva da língua. Desse modo, operações fonéticas, fonológicas, morfológicas e sintáticas, bem como relações semânticas e pragmáticas, estão alocadas sob o domínio da língua e estão em constante interação. Sendo a língua, portanto, um bem cultural que funciona como instrumento de comunicação convencionado e situado social e historicamente, e que envolve alto grau de abstração, ela não está no escopo de uma teoria como a de Gibson (1986), conforme o próprio autor salienta, a qual se ancora no realismo direto. Fowler (1986, 1996) e Goldstein e Fowler (2003) parecem, de certa forma, ignorar a advertência feita por Gibson (1986) acerca de sua teoria não ser aplicável ao conhecimento linguístico, veiculado pela fala ou pela escrita (p. 10)<sup>16</sup>.

Imaginemos um contexto de percepção de LE, como é o caso da sala de aula, por exemplo<sup>17</sup>. Os percebedores/aprendizes têm contato com apenas uma porção selecionada do idioma alvo, não sendo capazes de experienciá-lo em sua totalidade. A parte sonora do idioma geralmente é acessada pelos percebedores/aprendizes através da mediação da fala do professor ou de gravações em áudio e/ou vídeo, sendo, portanto, indireta. Lembramos que, para Gibson (1986), a percepção direta implica a ausência de um aparato entre o percebedor e o objeto, portanto, recursos tecnológicos audiovisuais não se encaixam em uma visão direta da percepção. Ademais, a fala do outro, seja a do professor ou a dos colegas, tampouco pode ser considerada como sendo perceptualmente direta. Isso se dá em função de que o realismo direto despreza a percepção em *segunda mão* (utilizando o termo gibsoniano), ou seja, o resultado da percepção de outro indivíduo como fonte primária para a percepção.

---

<sup>16</sup> Para Gibson (1986), a luz estruturada fornece informação sobre o ambiente visual. Para Fowler (1986, 1996) e Goldstein e Fowler (2003), analogamente, o gesto articulatório é especificado pelo sinal acústico da fala (p. 183), ou seja, o ar acusticamente estruturado fornece informação sobre o ambiente auditivo. Todavia, mantemos certa reserva quanto a tal postulado de Goldstein e Fowler (2003), pois questionamos a comensurabilidade entre a informação veiculada visualmente e aquela veiculada acusticamente, uma vez que os aparatos visual e auditivo operam de maneira distinta no que compete à detecção dos estímulos que circundam o percebedor.

<sup>17</sup> Referimo-nos ao contexto instrucional para exemplificar o ambiente perceptual de LE apenas para fins ilustrativos. Convém destacar que, para Best e Tyler (2007), independentemente de ser instrucional, qualquer ambiente perceptual de LE é deficitário quando comparado ao de L2.

As palavras e os diálogos que constam nos manuais didáticos são gravados por locutores treinados, os quais modalizam sua voz para atingir diversos efeitos entoacionais. Aquela informação que soa como surpresa ou animação é propositalmente criada e não reflete, necessariamente, a emoção que o locutor está sentindo no momento da elocução. Ademais, as gravações ocorrem em estúdios profissionais com tratamento acústico, visando a impedir que o sinal acústico seja mascarado por qualquer barulho simultâneo à fala, o que, de fato, não acontece em situações reais de comunicação. Enfatizamos, assim, que o uso das gravações não segue os preceitos ecológicos gibsonianos e tampouco resulta em uma percepção direta, o que, a uma primeira vista, viria a corroborar a previsão de Best e Tyler (2007) de que, para a percepção em contexto de LE, seu modelo não se aplica.

Obviamente, não podemos ser ingênuos a ponto de reduzir as limitações do ambiente perceptual de LE a uma pedagogia de LE problemática: mesmo se dispuséssemos da mais adequada das pedagogias, ela não seria suficiente para garantir as concessões que um ambiente de L2 *teoricamente* garante. Reiteramos que nossa argumentação não objetiva criticar a pedagogia de línguas estrangeiras desenvolvida no país, pois falantes de LE altamente competentes, oriundos de instrução formal explícita<sup>18</sup>, abundam nas diversas interações comunicativas.

Inicialmente, sem considerar o que expusemos até aqui, poderíamos pensar que o contexto de sala de aula, no que tange à percepção fônica especialmente, seria um ambiente ecológico. Tal constatação seria baseada no fato de que a sala de aula configura o espaço que cerca o percebedor/aprendiz e lhe fornece os subsídios necessários para que ele perceba elementos de seu entorno e esteja apto a desempenhar tarefas. Essa condição estaria de acordo com a visão de Gibson (1986), uma vez que, para o psicólogo, a característica fundamental de um ambiente é a de que ele circunda o indivíduo. No entanto, uma leitura mais cuidadosa da obra de Gibson (1986) nos fornece subsídios para constatarmos que, de fato, o cenário de percepção e aprendizagem de LE não seja ecológico, e apresentaremos duas razões para não seguirmos essa taxonomia.

O primeiro motivo para não rotularmos a percepção de LE em sala de aula como ecológica, em termos gibsonianos, ancora-se na proposição de que um ambiente ecológico, nos moldes de Gibson (1986), é muito amplo e está ligado à exploração e à sobrevivência. Gibson

---

<sup>18</sup> Presumimos que esta mesma condição pode gerar questionamentos pertinentes, inclusive, para contextos de aprendizagem de L2, conforme mencionamos na conclusão deste trabalho.

(1986) declara que, para que o percebedor obtenha sucesso na detecção das concessões, é primordial que o ambiente seja significativo. Um ambiente significativo consiste de uma área territorial, abrigo, água, objetos, ferramentas e outros animais, além de estar em mútua relação com o percebedor e de fornecer a ele tudo de que precisa para se desenvolver. Bem sabemos que a sala de aula de LE não oferece todos os mecanismos e ferramentas para que o aprendiz/percebedor desenvolva seu conhecimento linguístico e perceptual, e isso se torna um problema para uma abordagem ecológica da percepção. O segundo motivo para não tratarmos o contexto da percepção em LE como relacionado a um ambiente ecológico reside na constatação de que a sala de aula, isto é, o entorno do percebedor, faz parte de um cenário idealizado: tal ambiente foi pensado para fins didáticos e construído para prover algumas condições de percepção<sup>19</sup>. Em outras palavras, o espaço de sala de aula de LE é uma criação e uma adaptação de um contexto muito maior do que é capaz de expressar se comparado ao contexto de imersão *idealmente* propiciado pelo cenário de L2, e não estava disponível naturalmente ao aprendiz/percebedor desde sua gênese.

Sugerimos, portanto, que o termo *ambiente ecológico* seja desvinculado daquilo que concebemos ser o espaço referente à sala de aula de LE, local em que ocorre em larga medida a percepção fônica estrangeira, uma vez que tal cenário não cumpre os pré-requisitos básicos do que entendemos por ambiente ecológico, conforme explicitado no parágrafo anterior.

No tocante ao acesso à realidade que envolve o percebedor de LE, propomos que a forma mais adequada de realismo seja a *indireta* ao invés de direta. Porém, antes de defendermos o realismo indireto, julgamos apropriado definir o que é o realismo em si. Segundo informa o *Blackwell Dictionary of Western Philosophy* (BUNNIN; YU, 2004), os diferentes tipos de realismo dizem respeito à existência objetiva de vários objetos e propriedades, tais como o mundo externo, objetos matemáticos, universais, entidades teóricas, relações causais e propriedades estéticas e morais. A ideia central do realismo é a de que os objetos existem no universo, independentemente de o percebedor saber ou acreditar em sua existência. Em outras palavras, conforme ilustra Dancy (1985), o realismo na teoria da percepção – realismo perceptual – representa o ponto de vista de que os “objetos que apreendemos são capazes de

---

<sup>19</sup> Trazemos esta discussão sobre a percepção fônica no contexto de sala de aula apenas como referência a um cenário mais prototípico de um ambiente de LE. Todavia, ressaltamos que o contexto de LE é muito mais amplo do que o espaço da sala de aula, não se reduzindo, portanto, a somente este contexto.

existir e geralmente existem de fato, e retêm pelo menos algumas das propriedades que apreendemos terem, mesmo quando são despercebidos” (p. 182).

Essencialmente no que compete à natureza da percepção humana, em um enfoque realista, há uma longa discussão que impõe uma questão-problema aos filósofos da percepção: percebemos os objetos do mundo de maneira direta ou indireta? Para Gibson (1986), Fowler (1986), Best (1995) e Best e Tyler (2007), não há dúvida: a percepção humana é realista direta. Para os filósofos da percepção, o enredo não é tão simples assim. Dancy (1985) afirma que o contraste entre o direto e o indireto é bastante escorregadio e, por vezes, difícil de estabelecer com firmeza. Mound (2003) explana que a controvérsia reside no fato de percebermos os objetos do mundo diretamente ou por meio de elementos intermediários, como ideias, imagens, impressões, sensações ou dados dos sentidos.

De modo bastante interessante, Dancy (1985) comenta que a disputa entre o realista direto e o indireto se estabelece a partir da possibilidade de estarmos ou não diretamente conscientes acerca da existência e da natureza dos objetos. Sob esta perspectiva, o autor expõe que:

Ambos, como realistas, concordam que os objetos físicos que vemos e tocamos são capazes de existir e retêm algumas das suas propriedades, (mesmo) quando não apreendidos. Mas o realista indireto afirma que nunca estamos diretamente conscientes de objetos físicos; estamos apenas indiretamente conscientes deles, em virtude de uma consciência direta de um objeto intermediário (variadamente descrito como uma ideia, dado dos sentidos, percepto ou aparência). O realista direto nega esta afirmação. [(DANCY, 1985, p. 183) observações parentéticas nossas]

O realismo indireto prevê que o mundo externo realmente existe, mas nossa percepção em relação a ele é mediada pela percepção de objetos intermediários subjetivos, como, por exemplo, sensações (BROWN, 2009). Conforme Dancy (1985), a vertente indireta do realismo presume que nunca estamos diretamente conscientes dos objetos físicos, em função de estarmos diretamente conscientes de um objeto intermediário, que podem ser ideias, imagens, impressões, sensações ou dados dos sentidos, segundo, também, aponta Mound (2003). A possibilidade de haver algo entre o percebido e o objeto a ser percebido vai completamente ao encontro da hipótese de que *a percepção de elementos sonoros de uma LE é filtrada pelo conhecimento fônico da L1*, conjectura em favor da qual nos posicionamos.

Seguimos a perspectiva realista indireta (JACKSON, 1977, 2010), portanto, em função de três premissas elementares. A primeira delas, conforme problematizamos nesta seção, é mais transparente quanto à conotação do termo *indireto* e tem a ver com a constatação de que, no contexto de LE, a percepção e a aquisição fônicas ocorrem de maneira mediada, seja pela fala do professor, seja pelo ferramental tecnológico disponível ao aprendiz. A segunda premissa em favor da abordagem indireta, advinda das ponderações feitas na seção anterior, reside na assunção de que o conhecimento fônico da L1 baliza a percepção de elementos sonoros estrangeiros, condicionando as suas categorizações (segundo evidencia a literatura em aquisição fonológica não nativa). Por fim, a terceira premissa que motiva a noção indireta do realismo se apoia sobre a interação da unidade gestual, de natureza acústico-articulatória, com outros âmbitos da gramática, refletindo-se, inclusive, no estabelecimento da relação significante-significado.

Cabe destacar que, em nossa concepção, a aprendizagem perceptual é um processo cognitivo por excelência, que envolve o processamento da informação da fala, assim como outras funções mentais – outro ponto que nos diferencia da perspectiva de Gibson (1986), Best (1995) e Best e Tyler (2007). Desse modo, seguimos Chamot *et al.* (1993), de acordo com o qual a aprendizagem, seja ela relacionada à habilidade linguística ou a outras habilidades humanas, implica um curso cognitivo *per se*, em que os indivíduos ativamente selecionam e organizam as informações que os circundam, além de relacioná-las ao seu conhecimento prévio. Desse modo, os aprendizes internalizam aquilo que consideram ser importante e utilizam as informações armazenadas. A ação cognitiva de *selecionar as informações linguísticas dispostas no ambiente, analisar suas formas e funções, refletir sobre suas produções, antecipar os tipos de demandas possíveis e ativar o conhecimento e as habilidades prévias para a execução de novas tarefas linguísticas* configura uma capacidade cognitiva complexa (MCLAUGHLIN, 1987).

Paralelamente, Brown, Roediger e McDaniel (2014) explicam que a aprendizagem se refere ao ato de adquirir conhecimento e habilidades, tornando-os disponíveis na memória para serem utilizados em oportunidades e problemas futuros, mecanismo que, por si só, já é capaz de abalar a perspectiva realista direta acerca da percepção da fala em LE. Ao que ilustra Mildner (2008), todo o tipo de aprendizagem se apoia sobre a *plasticidade*<sup>20</sup> do nosso cérebro,

---

<sup>20</sup> A plasticidade desenvolvimental compreende mudanças em conexões que ocorrem após o nascimento como consequência de interações com o meio. Tais mudanças tornam possível o processo de aprender e adquirir novas

contemplada na estrutura, no desenvolvimento e no funcionamento do sistema nervoso. Vinculamo-nos, portanto, a uma perspectiva cognitivista (e representacional) acerca da aprendizagem, diferente da posição de Gibson (1986).

A título de conclusão, não podemos esquecer que nós, seres humanos, temos a habilidade de racionalizar o conhecimento que chega até nós e de filtrar cognitivamente aquilo que nos é representativo, cabendo a este processo, também, recursos inferenciais e a instituição de uma representação mental acerca dos objetos percebidos. Adicionalmente, recorreremos à ideia da existência de um conhecimento mental quanto a unidades fônicas justamente pelo respaldo na literatura em psicologia e neuropsicologia (LUND, 2001; LECLERCQ; ZIMMERMANN, 2005; STYLES, 2005; KRAMER; WIEGMANN; KIRLIK, 2007; MILDNER, 2008; GAZZANIGA; HEATHERTON; HALPERN; HEINE, 2012), a qual faz referência a mecanismos de atenção e traz evidências a favor do armazenamento de informações sonoras na memória. Nesse âmbito, pensamos que uma abordagem da percepção fônica vinculada ao realismo indireto parece ser muito mais apropriada, além de dar conta, de modo mais elegante, das estipulações que fazemos sobre o mundo que nos cerca e, conseqüentemente, sobre o conhecimento dos padrões sonoros a que temos acesso.

## 5. Conclusão

Os pontos que levantamos neste artigo dizem respeito a possíveis modificações no modelo *PAM-L2* no momento em que este é utilizado para dar conta da percepção fônica de uma LE, especialmente nos contextos em que o idioma alvo não é amplamente utilizado fora do espaço da sala de aula. Tal situação corresponde largamente ao cenário da percepção do sistema sonoro do inglês, por exemplo, por falantes nativos do português brasileiro, que não estão inseridos em um ambiente onde a língua alvo é a principal ferramenta nas interações verbais.

Muito temos a considerar quando tentamos adaptar um modelo de percepção da fala em L2 para o cenário da percepção fônica em LE. No que concerne ao *PAM-L2*, especialmente, sugerimos as modificações a seguir:

---

experiências, mas também resultam desses processos e se fazem presentes desde os primórdios da existência humana. Outro tipo de plasticidade se refere à capacidade de recuperar funções cognitivas após danos cerebrais (MILDNER, 2008, p. 139).

- Ao contrário de Gibson (1986), Best (1995) e Best e Tyler (2007), que não preveem o uso de inferências e representações mentais para o evento perceptual, defendemos uma abordagem cognitiva da percepção e da aprendizagem perceptual, em que o aprendiz/percebedor opera tanto mentalmente, nas relações abstratas entre os sons da L1 e da LE, quanto fisicamente, nas rotinas articulatórias.
- Em detrimento da visão de Fowler (1986, 1986), clamamos pela adesão ao gesto articulatório na visão de Albano (2001), pois este se relaciona a outras unidades linguísticas, como morfológicas e sintáticas, sendo bastante conveniente para o tratamento da fala enquanto parte da gramática e, principalmente, pelo fato de tal unidade fônica percorrer o léxico dos usuários da língua.
- Sugerimos que o conceito de ambiente ecológico seja desvinculado das situações de percepção fônica em LE, especialmente nos contextos de sala de aula, uma vez que, neste cenário, a comunicação muitas vezes não é autêntica e o percebedor carece de uma exposição que seja, de fato, massiva à língua em uso efetivamente.
- Diferentemente de Gibson (1986), Best (1995) e Best e Tyler (2007), que concebem o realismo direto como forma de acesso aos objetos do mundo, prezamos pelo realismo indireto (JACKSON, 1977, 2010) como sendo o meio de acesso às unidades fônicas disponíveis ao percebedor. Optamos pelo realismo indireto perceptual porque ele dá conta do fato de que a L1 permeia as relações gestuais já internalizadas e aquelas que estão por adentrar a sua gramática fônica, além de reconhecer que existe sempre um aparato (seja mental ou físico) entre o percebedor e o objeto a ser percebido.

Reconhecemos, entretanto, que ainda se faz necessário incorporar ao modelo uma teoria que dê conta de mecanismos cognitivos como atenção e processamento do *input*, já que se faz indispensável que, para a percepção fônica em LE, o percebedor esteja atento aos contrastes finos da massa sonora. A atenção ao *input* fornece ao percebedor não somente a possibilidade de distinguir unidades fônicas, mas também de conectar tais informações ao seu léxico e, conseqüentemente, à sua gramática. Neste sentido, atentar ao *input* permite a seleção dos aspectos acústico-articulatórios primordiais para os contrastes funcionais na LE, uma vez que línguas distintas podem eleger diferentes aspectos acústico-articulatórios prioritários, essenciais no sistema alvo, para veicular oposições de significado (ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015).

Ainda no que diz respeito a uma maior elaboração em termos de processamento do *input*, é preciso, também, considerar a multimodalidade da informação linguística. Dentro dessa variada gama de estímulos de caráter multimodal, outro compromisso do modelo é explicar o impacto das relações grafo-fônico-fonológicas (ZIMMER; SILVEIRA; ALVES, 2009) sobre a percepção da fala, uma vez que o conhecimento da língua escrita pode causar assimilações perceptuais equivocadas, resultando em sobreposições de categorias fônicas no nível interfonológico.

Destacamos, igualmente, a necessidade de se elaborarem tarefas de percepção que se proponham a captar com mais precisão a maneira como os percebedores categorizam os sons da LE. Dessa forma, as predições estabelecidas pelo *PAM-L2* poderiam ser testadas empiricamente e, como consequência, poderíamos ter fortes indicadores das possibilidades de assimilação perceptual. Tal constatação revelaria importantes fatos sobre percepção de distinções sonoras e também sobre a organização abstrata das unidades da fala, implicando conhecimento acerca das relações gestuais no âmbito tipológico.

Enfatizamos que as fronteiras entre LE-L2 estão longe de serem bem delimitadas. Sempre haverá casos em que o aprendiz de um novo idioma estará inserido no país em que a língua alvo é falada, mas poderá vir a não se relacionar em diversas situações comunicativas com os falantes nativos locais. Por outro lado, sempre haverá aprendizes de um novo idioma que, mesmo não tendo a oportunidade de vivenciar a língua no país em que tal sistema é falado, irão se dedicar ao estudo da nova língua e procurar se cercar do *input* estrangeiro a partir de inúmeras ferramentas (seriados, vídeos, músicas, *podcasts*, etc.). Outrossim, o contexto de aprendizagem de L2 pode vir a ser, de fato, ecológico, mas nem sempre o é – assim como alguns contextos de aprendizagem de LE podem vir a ser ecológicos também. Neste trabalho, porém, pensamos em como adaptar as premissas da proposta original de Best e Tyler (2007) para contextos de percepção fônica em LE.

Em nosso entendimento, os pontos positivos do *PAM-L2* residem, principalmente, na filiação do modelo à unidade gestual e nas previsões de assimilação perceptual, uma vez que as categorizações feitas pelos ouvintes são, em grande medida, nelas baseadas. Além disso, como o *PAM-L2* prevê a adoção do gesto articulatório, as estipulações e os resultados advindos de pesquisas que assumem o modelo podem contribuir para o próprio construto da Fonologia Articulatória.

É nossa tarefa, no entanto, esclarecer que algumas instâncias primordiais do modelo articulatório não estão sendo seguidas (seja pela falta de compromisso com a unidade gestual postulada pelo *PAM-L2*, seja por concepções divergentes em relação ao paradigma realista direto ou pela própria dificuldade ao se pensar em um ambiente ecológico de fato). Isso deixa margem para a possibilidade de revisões, ou, até mesmo, da elaboração de um novo modelo (o que seria um objetivo a ser atingido em longo prazo, e que pode ter, em seu cerne, a presente reflexão), decorrente de contínuos reparos e ajustes teóricos que deem conta de retratar mais fielmente o cenário da percepção da fala em LE. Como resultado, estudos futuros (que tentem aprofundar as modificações tanto para LE como para a L2) podem conduzir o modelo para a noção de percepção indireta e para o gesto acústico-articulatório, bem como à equivalência referente aos contextos de LE e L2 quanto à percepção fônica.

### Referências bibliográficas

ALBANO, E. **O Gesto e suas Bordas**: Esboço de Fonologia Acústico-Articulatória do Português Brasileiro. Campinas: Mercado de Letras, 2001.

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

ALVES, U.; ZIMMER, M. Percepção e Produção dos Padrões de VOT do Inglês por Aprendizes Brasileiros: O Papel de Múltiplas Pistas Acústicas sob uma Perspectiva Dinâmica. **Alfa: Revista de Linguística**, V. 59, p. 157-180, 2015. **crossref** <http://dx.doi.org/10.1590/1981-5794-1502-7>

ANTONIOU, M.; TYLER, M.; BEST, C. Two Ways to Listen: Do L2-Dominant Bilinguals Perceive Stop Voicing According to Language Mode? **Journal of Phonetics**. V. 40, p. 582-594, 2012. **crossref** <http://dx.doi.org/10.1016/j.wocn.2012.05.005>

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M.; CROFT, W.; ELLIS, N.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language is a Complex Adaptive System: Position Paper. **Language Learning**, V. 59, p. 1-26, 2009. **crossref** <http://dx.doi.org/10.1111/j.1467-9922.2009.00533.x>

BEST, C. A Direct Realist View of Cross-language Speech Perception. In: STRANGE, W. (Ed.). **Speech Perception and Linguistic Experience**: Theoretical and Methodological Issues in Cross-Language Speech Research. Timonium: York Press, 1995. p. 167-200.

BEST, C.; TYLER, M. Nonnative and Second-Language Speech Perception: Commonalities and Complementarities. In: BOHN, O.; MUNRO, M. (Orgs.). **Language Experience in Second Language Speech Learning**: In honor of James Emil Flege. Filadélfia: John Benjamins Publishing Company, 2007, p. 13-34. **crossref** <http://dx.doi.org/10.1075/llt.17.07bes>

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.

BROWMAN, C.; GOLDSTEIN, L. Articulatory Gestures as Phonological Units. **Haskins Laboratories Status Report on Speech Research**. V. 100, p. 69-101, 1992.

BROWMAN, C.; GOLDSTEIN, L. Articulatory Phonology: An Overview. **Haskins Laboratories Status Report on Speech Research**. V. 112, p. 23-42, 1992. **crossref** <http://dx.doi.org/10.1159/000261913>

BROWN, D. Indirect Perceptual Realism and Demonstratives. **Philosophical Studies**. V. 145, p. 377-394, 2009. **crossref** <http://dx.doi.org/10.1007/s11098-008-9237-x>

BROWN, P.; ROEDIGER, H.; MCDANIEL, M. **Make It Stick: The Science of Successful Learning**. Cambridge: Harvard University Press, 2014. **crossref** <http://dx.doi.org/10.4159/9780674419377>

BUNNIN, N.; YU; J. **The Blackwell Dictionary of Western Philosophy**. Malden: Blackwell, 2004.

CHAMOT, A; BARNHARD, S.; EI-DINARY, P.; CARBONARO, G.; ROBBINS, J. **Methods for Teaching Learning Strategies in the Foreign Language Classroom and Assessment of Language Skills for Instruction: Final Report**. Washington: Georgetown University, 1993.

DANCY, J. **Epistemologia Contemporânea**. Cambridge: Harvard University Press, 1985.

DE BOT, K.; LOWIE, W.; VERSPOOR, M. A Dynamic Systems Theory Approach to Second Language Acquisition. **Bilingualism: Language and Cognition**, V. 10, n. 1, p. 7–21, 2007. **crossref** <http://dx.doi.org/10.1017/S1366728906002732>

DELATTRE, P.; LIBERMAN, A.; COOPER, F. Acoustic Loci and Transitional Cues for Consonants. **Journal of the Acoustical Society of America**, V. 27, n. 4, 1955, p. 769-773. **crossref** <http://dx.doi.org/10.1121/1.1908024>

FEIDEN, J.; ALVES, U.; FINGER, I. O Efeito da Anterioridade e da Altura na Identificação das Vogais Médias Altas e Médias Baixas do Português Brasileiro por Falantes de Espanhol. **Letras de Hoje**. V. 49, n. 1, p. 85-94, 2014. **crossref** <http://dx.doi.org/10.15448/1984-7726.2014.1.14715>

FEIDEN, J.; PEROZZO, R.; FINGER, I.; FONTES, A. Percepção de Vogais Médias Altas e Médias Baixas do Português Brasileiro por Falantes de Espanhol em Tarefa de Discriminação Categórica. In: ALVES, U. (Org.). **Aquisição Fonético-Fonológica de Língua Estrangeira: Investigações Rio-Grandenses e Argentinas em Discussão**. São Paulo: Pontes (no prelo).

FLEGE, J. Second Language Speech Learning: Theory, Findings, and Problems. In: STRANGE, W. (Ed.). **Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research**. Timonium: York Press, 1995. p. 233-272.

FOWLER, C. Coarticulation and Theories of Extrinsic Timing. **Haskins Laboratories Status Report on Speech Research**. V. 8, p. 113-133, 1980. **crossref** <http://dx.doi.org/10.3758/BF03211675>

FOWLER, C. An Event Approach to the Study of Speech Perception from a Direct-Realist Perspective. **Journal of Phonetics**, V. 14, p. 3-28, 1986.

FOWLER, C. Invariants, Specifiers, Cues: An Investigation of Locus Equations as Information for Place of Articulation. **Perception & Psychophysics**. V. 55, n. 6, 1994, 597-610.

FOWLER, C. Listeners Do Hear Sounds, not Tongues. **Journal of the Acoustical Society of America**, V. 99, p. 1730-1741, 1996. **crossref** <http://dx.doi.org/10.1121/1.415237>

GASS, S.; SELINKER, L. **Second Language Acquisition: An Introductory Course**. 3. ed. Nova Iorque: Routledge, 2008.

GAZZANIGA, M.; HEATHERTON, T.; HALPERN, D.; HEINE, S. **Psychological Science**. Nova Iorque: W.W. Norton & Company, 2012.

GIBSON, J. **The Ecological Approach to Visual Perception**. Nova Iorque: Psychology Press, 1986.

GOLDSTEIN, L.; FOWLER, C. Articulatory Phonology: A Phonology for Public Language Use. In: SCHILLER, N.; MEYER, A. (Eds.). **Phonetics and Phonology in Language Comprehension and Production**, Mouton de Gruyter, 2003. p. 159-207. **crossref** <http://dx.doi.org/10.1515/9783110895094.159>

GUTIERRES, A.; BATTISTI, E. A Variação da Nasal Velar por Aprendizes Brasileiros de Inglês. **Anais do X Encontro do CELSUL**, 2012.

JACKSON, F. **Perception**. Nova Iorque: Cambridge University Press, 1977.

JACKSON, F. Representative Realism. In: DANCY, J.; SOSA, E.; STEUP, M. (Eds.). **A Companion to Epistemology**. 2. ed. Malden: Blackwell, 2010. p. 702-705.

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.

KLUGE, D. **Perception and Production of Final Nasals by Brazilians Learners of English**. 2004. 175 f. Dissertação (Mestrado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2004.

KOERICH, R. **Perception and Production of Word-Final Vowel Epenthesis by Brazilian EFL Students**. 2002. 261 f. Tese (Doutorado em Letras)—Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2002.

KOSSLYN, S.; KOENIG, O. **Wet Mind: The New Cognitive Neuroscience**. New York: Free Press, 1995.

KRAMER, A.; WIEGMANN, D.; KIRLIK, A. **Attention: from Theory to Practice**. Nova Iorque: Oxford University Press, 2007.

LECLERCQ, M.; ZIMMERMANN, P. **Applied Neuropsychology of Attention: Theory, Diagnosis and Rehabilitation**. Londres: Psychology Press, 2005.

LIBERMAN, A. M.; MATTINGLY, I. G. The Motor Theory of Speech Perception Revised. In: **Cognition**, V. 21, p.1-36, 1985. **crossref** [http://dx.doi.org/10.1016/0010-0277\(85\)90021-6](http://dx.doi.org/10.1016/0010-0277(85)90021-6)

LUND, N. **Attention and Pattern Recognition**. Filadélfia: Routledge, 2001.

MCLAUGHLIN, B. **Theories of Second Language Learning**. Londres: Edward Arnold Press, 1987.

MILDNER, V. **The Cognitive Neuroscience of Human Communication**. Nova Iorque: Taylor & Francis Group, 2008.

MOORE, D. H. **The Perception of English Word-Final /l/ by Brazilian Learners**. 2008. 140 f. Dissertação (Mestrado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2008.

MOUND, B. **Perception**. Durham: Acumen, 2003.

NISHIDA, G. **Sobre Teorias de Percepção da Fala**. 2012. 174 f. Tese (Doutorado em Letras)—Faculdade de Letras. Universidade Federal do Paraná, Curitiba, 2012.

OHALA, J. Speech Perception is Hearing Sounds, not Tongues. **Journal of the Acoustical Society of America**, V. 99, p.1718-1725, 1996. **crossref** <http://dx.doi.org/10.1121/1.414696>

PEROZZO, R. V. **Percepção de Oclusivas Não Vozeadas sem Soltura Audível em Codas Finais do Inglês (L2) por Brasileiros: O Papel do Contexto Fonético-Fonológico, da Instrução Explícita e do Nível de Proficiência**. 2013. 191 f. Dissertação de Mestrado (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, M. **The Perception and Production of English Interdental Fricatives by Brazilian EFL Learners**. 2006. 186 f. Dissertação de Mestrado (Mestrado em Letras)—Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2006.

REIS, M. **The Perception and Production of the English Voiceless Interdental Fricative by Speakers of European French and Brazilian Portuguese**. 2010. 271 f. Tese (Doutorado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2010.

SCHWARTZHAUPT, B.; ALVES, U.; FONTES, A. The Role of L1 Knowledge on L2 Speech Perception: Investigating how Native Speakers and Brazilian Learners Categorize Different VOT Patterns in English. **Revista de Estudos da Linguagem**, V. 23, p. 311-334, 2015.

**crossref** <http://dx.doi.org/10.17851/2237-2083.23.2.311-334>

SILVA, A. A Variável Tempo nos Estudos em Aquisição. In: BRAWERMAN-ALBINI, A.; GOMES, M. (Orgs.). **O Jeitinho Brasileiro de Falar Inglês: Pesquisas sobre a Pronúncia do Inglês por Falantes Brasileiros**. São Paulo: Pontes, 2014, p. 33-47.

SILVEIRA, R. Uma Análise da Produção Acadêmica na Área da Interfonologia Português-Ínglês. In: RAUBER, A.; WATKINS, M.; SILVEIRA, R.; KOERICH, R. (Orgs.). **The Acquisition of Second Language Speech: Studies in Honor of Professor Barbara O. Baptista**. Florianópolis: Editora Insular, p. 3-19. 2010.

STYLES, E. **Attention, Perception and Memory: an Integrated Introduction**. Nova Iorque: Psychology Press, 2005.

ZIMMER, M.; ALVES, U. Uma Visão Dinâmica da Produção da Fala em L2: O Caso da Dessonorização Terminal. **Revista da ABRALIN**, V. 11, p. 221-272, 2012.

ZIMMER, M.; SILVEIRA, R.; ALVES, U. **Pronunciation Instruction for Brazilians: Bringing Theory and Practice Together**. New Castle upon Tyne, UK: Cambridge Scholars Publishing, 2009.

Artigo recebido em: 30.10.2015

Artigo aprovado em: 16.02.2016