

UMA VISÃO DINÂMICA DA PRODUÇÃO DA FALA EM L2: O CASO DA DESSONORIZAÇÃO TERMINAL

Márcia C. ZIMMER

Universidade Católica de Pelotas

Ubiratã Kickhöfel ALVES

Universidade Federal do Rio Grande do Sul

RESUMO

Este artigo apresenta um apanhado teórico das diferentes abordagens sobre a aprendizagem fonético-fonológica de segunda língua, destacando a concepção dinâmica para o estudo da produção e da percepção de fala em L2. Para isso, o processo de interfonologia chamado Dessonorização Terminal é revisitado nessas abordagens. Por fim, novos dados empíricos sobre esse processo são apresentados, com um estudo entre 15 falantes adultos brasileiros aprendizes de inglês e 5 falantes nativos do inglês americano. A descrição e discussão dos resultados oferecem evidências para uma visão dinâmica da aprendizagem da fala em L2.

ABSTRACT

This article presents an overview of different accounts of L2 phonetic-phonological learning, underscoring the dynamical approach to the study of L2 speech production and perception. For that purpose, the interphonology process of Terminal Devoicing in English is revisited in these accounts. New empirical data on this process are presented through a study involving 15 adult Brazilian speakers learners of English, and 5 North-American participants. The description and discussion of the findings offer evidence for a dynamic view of second language speech learning.

PALAVRAS-CHAVE

Abordagem dinâmica. Dessonorização Terminal. Fonologia da interlíngua.

KEY WORDS

Dynamical approaches. Interlanguage phonology. L2 speech. Terminal Devoicing

Introdução

As pessoas afirmam que as crianças aprendem uma língua estrangeira mais rápido e melhor do que os adultos. É senso comum acreditar que um aprendiz adulto de língua estrangeira só aprende a falar com pouco sotaque ou sem sotaque se viver por algum tempo no país onde a língua é falada. Seria mesmo assim? Ao abordar essas questões, é instigante proceder a alguns questionamentos. O que é o sotaque na L2, afinal de contas? Que evidências há hoje para reforçar ou para refutar essa concepção sobre a produção de fala numa língua estrangeira? Já que se trata de um processo de aprendizagem, não seria o caso de se levar em conta a variável tempo na produção de fala em L2? Indo um pouco além da visão clássica, este estudo se propõe a argumentar que o tempo relativo da unidade dinâmica – o gesto – e seu faseamento são elementos constitutivos da produção e percepção da fala, tanto na L1 como na L2, na medida em que explica aspectos deixados de fora por outras abordagens.

Os estudos na área de interfonologia desenvolvidos em nosso país não são novidade. De fato, há muito já se discute questões referentes ao sotaque em língua estrangeira, sobre a sua relação com a idade de início de estudo da segunda língua e sobre o papel do sistema da língua materna na aprendizagem deste novo sistema linguístico. Ainda que não se caracterize

como um assunto novo, a questão do sotaque parece evoluir com o próprio avanço dos estudos psicolinguísticos, acompanhando a evolução dos modelos teóricos e do modo de conceber e fazer fonologia, sem desvinculá-la da fonética.

Neste artigo, propomos perpassar essas diferentes concepções, sempre focalizando o dado de aquisição fonológica de L2. Através desse apanhado histórico, acreditamos que conseguiremos demonstrar, mais claramente, a pertinência aqui defendida de uma concepção dinâmica do dado de aquisição de segunda língua.

Ao rever historicamente o dado de aquisição de L2, começemos pela caracterização de tais fenômenos à luz de uma concepção estática, baseada em primitivos de tempo extrínseco¹ (FOWLER, 1980) da aquisição fonológica.

1. Abordagens Tradicionais

Nesta seção, revisaremos as abordagens calcadas em modelos clássicos de Fonologia, que embasaram muitos dos estudos de interfonologia desde a década de 60. Ao caracterizar brevemente esses modelos, apontaremos, sobretudo, os aspectos destoantes da concepção dinâmica que embasa o presente trabalho, concepção essa que será mais plenamente apresentada na quarta seção.

Conforme descrevem RODRIGUES *et al.* (2008), as unidades de análise caracterizadas por uma natureza estática são aquelas que representam binariamente a presença/ausência de um contraste fônico. Nesse sentido, os traços distintivos

¹ FOWLER (1980) foi a primeira autora a chamar os fonemas e segmentos de primitivos de tempo extrínseco, pois esses primitivos não carregam informação relativa ao tempo, sendo caracterizados como primitivos estáticos. Ao discorrer sobre teorias fonológicas tradicionais, a autora as chama de teorias de tempo extrínseco.

correspondem a unidades de análise de natureza estática por excelência, uma vez que descrevem os segmentos a partir da presença ou ausência de determinadas propriedades. Assim, as abordagens tradicionais caracterizam a interlíngua em termos de processos que envolvem apagamentos, inserções e substituições plenas de segmentos ou, até mesmo, de unidades menores, tais como traços. Tais processos caracterizam-se, indiscutivelmente, por um caráter categórico, em que nuances são desconsideradas. As nuances em questão vem sendo relegadas ao domínio da Fonética, a qual, apesar de fornecer insumos para a Fonologia e para os estudos cognitivos, mostra-se distanciada de um modelo que visa a formalizar o sistema fonológico, enquanto Teoria da Gramática. Desses modelos estáticos, frisemos, portanto, um descomprometimento com a natureza psicomotora ou corporificada da cognição. Essa natureza pode ser vista como a ponte que vai ligar o dado de produção, em seu caráter dinâmico, com a cognição e, por conseguinte, à noção de gramática fônica dinâmica.

1.1. A noção de marcação nos estudos de produção oral em língua estrangeira

Antes mesmo de mencionar a tradição gerativista, devemos nos deter, primeiramente, em um construto teórico nascido no Círculo Linguístico de Praga, sendo proposto primeiramente por TRUBETZKOY (1939) e JAKOBSON (1941): a noção de marcação. Essa noção, além de perpassar os diversos modelos fonológicos de descrição linguística e influenciar grande parte dos estudos de aquisição fonológica da língua materna, constitui, também, um conceito influente ao longo de toda a história dos estudos de aquisição fonológica de L2.

ECKMAN (1996: 198) define marcação da seguinte forma:

“Se a presença de uma estrutura p em uma língua implica a presença de alguma outra estrutura q, mas a presença de q em alguma língua não implica a presença de p, então p é marcado em relação a q, e q não é marcado em relação a p”. Verificamos, assim, o caráter binário da oposição marcado/não-marcado, em que um membro da oposição é privilegiado com relação a outro (*cf.* ECKMAN, 2004). Como exemplo, no que concerne ao vozeamento dos segmentos plosivos, os segmentos contendo o traço [+voz] seriam tratados como marcados, ao passo que os que carregam o traço [- voz] corresponderiam à contraparte não-marcada.

No cenário de estudos de aquisição de segunda língua, a noção de marcação tem sido, por muito tempo, privilegiada. Há, de fato, duas hipóteses pautadas nesse conceito, ambas postuladas por Eckman: a *Markedness Differential Hypothesis* (ECKMAN, 1977) ou Hipótese da Marcação Diferencial, doravante HDM, e a *Structural Conformity Hypothesis* (ECKMAN, 1996) ou Hipótese da Conformidade Estrutural, HCE.

A Hipótese de Marcação Diferencial caracteriza-se como uma extensão da Hipótese de Análise Contrastiva (*cf.* ECKMAN, 2004), pois postula que a noção de marcação de caráter tipológico deve ser incorporada à Análise Contrastiva para que se possa verificar o grau de dificuldade na aquisição de uma dada estrutura da L2. De acordo com a HDM, as áreas em que um aprendiz apresentará dificuldades podem ser previstas da seguinte maneira: 1) as áreas da língua-alvo que diferem da língua nativa, sendo mais marcadas do que a língua nativa, serão mais difíceis; 2) o grau relativo de dificuldade das áreas de diferença da língua-alvo que são mais marcadas do que a língua nativa corresponderá ao grau relativo de marcação; 3) as áreas da língua-alvo que diferirem da língua materna, mas que não forem mais marcadas do que esta, não serão mais difíceis (ECKMAN,

1977: 321).

Dessa forma, as premissas básicas dessa hipótese visam a prever o grau de dificuldade com que certas estruturas da língua estrangeira serão aprendidas. Essas previsões são feitas somente com base em diferenças e/ou semelhanças, em termos de marcação, entre as estruturas das línguas-fonte e alvo, sem que sejam levados em conta quaisquer fatores psicolinguísticos relacionados ao falante e, por conseguinte, aos processos cognitivos inerentes à aprendizagem humana.

Através da HMD, a noção de marcação é implementada pela Análise Contrastiva. Por sua vez, a Hipótese da Conformidade Estrutural (HCE) prevê que as generalizações universais que se mostram em línguas primárias mantêm-se, também, nas interlínguas (ECKMAN, 1991: 24). É interessante frisar que uma das principais consequências dessa hipótese é a possibilidade de ocorrência de padrões que não necessariamente existam na L1 nem tampouco na L2, mas que sejam decorrentes de uma tendência de prioridade de estruturas não-marcadas.

A concepção dinâmica através da qual vemos a aquisição de segunda língua nos leva a questionar o que realmente se quer dizer com o termo 'marcação'. Inegavelmente, tal termo tem servido como um guarda-chuva para uma diversidade de sentidos diferentes relacionados aos fenômenos linguísticos, passando por aspectos como dificuldade articulatória, frequência de ocorrência de determinada estrutura, dentro outros aspectos que deveriam ser analisados sob sua individualidade, não sob um único rótulo.

Em outras palavras, nos estudos linguísticos, a noção de marcação adquiriu um status inquestionável, além de um extenso rol de utilidades, sendo usada tanto no estruturalismo, como no gerativismo e nos estudos de aquisição de linguagem, nos estudos de línguas crioulas, dentre muitos outros. Contudo,

sua definição é ambígua e evitada de tautologias que não contribuem para sua compreensão, nem em termos de processo fonológico formal, nem em termos de processo psicolinguístico.

Dado o caráter genérico e um tanto quanto vago assumido pelo conceito de marcação, HASPELMATH (2006) aponta a existência de no mínimo doze diferentes sentidos para o termo, agrupando-os em quatro classes maiores: marcação como complexidade, como dificuldade, como anormalidade e como correlação multidimensional. Conforme as palavras do autor, o termo ‘marcação’ desenvolveu-se “com uma multiplicidade de sentidos amplamente divergentes dos quais os linguistas que o usam não se mostram conscientes” (p. 27). ZIMMER, SILVEIRA & ALVES (2009), com base nas considerações do autor, propõem uma substituição do termo pela consideração de aspectos como a experiência linguística, associada à capacidade humana de perceber características salientes do *input*, guiada pelos aspectos de frequência de certos segmentos, bem como da complexidade dos gestos e constelações gestuais. Assim, tal experiência é o produto da capacidade cognitiva geral do aprendiz de fazer associações a partir dos dados disponíveis no *input* da L1 e da L2.

1.2 Abordagens tradicionais: a assunção de primitivos de tempo extrínseco

O estudo da interfonologia já se mostrava presente nas abordagens tradicionais, como na Fonologia Gerativa de CHOMSKY & HALLE (1968). A caracterização do gerativismo no estudo da aquisição fonológica de segunda língua teve consequências tanto na formalização dos processos fonético-fonológicos de interlíngua quanto nas próprias concepções basilares acerca do processo de aquisição de linguagem.

No que diz respeito à formalização dos processos interfonológicos através de regras, com base na visão defendida pelo *The Sound Pattern of English* (CHOMSKY & HALLE, 1968), fundamental para o modelo gerativista é a caracterização dos segmentos como matrizes de traços. Sendo binários, os traços distintivos não apresentam caráter dinâmico; apresentam natureza “sim” ou “não”; assim, um segmento ou é [+voz] ou [-voz], [+alt] ou [-alt], sem que se levem em consideração, por exemplo, características coarticulatórias que em muito influenciam a natureza acústica do som a ser produzido. As regras fonológicas, nesse sentido, teriam a função de transformar estruturas mais marcadas em menos marcadas, modificando o valor dos traços distintivos dos segmentos para a chegada a um segmento mais simples.

Além disso, a produção de formas diferentes do falar nativo é vista, sob a perspectiva gerativista, como a aplicação de regras, que podem ou não atuar na L1. Vejamos, de maneira mais aplicada, a noção de marcação e o funcionamento das regras gerativas, para que possamos contrastá-la com a visão dinâmica advogada neste artigo. Faremos menção, aqui, ao estudo de MAJOR (1987), sobre o fenômeno de Dessonorização Terminal na Interlíngua Português-Inglês. Sob a caracterização acima, os segmentos [b, d, γ] são tidos como mais marcados, uma vez que exibem o valor positivo do traço [sonoro], referente ao vozeamento das plosivas. Uma vez que essas obstruintes não ocorrem em posição final de sílaba em nossa língua, o autor afirma que tais consoantes, na interlíngua português-inglês, sofrem um processo de desvozeamento. Assim, segundo o autor, a seguinte regra encontra-se ativa na Interfonologia Português-Inglês:

[-son] → [-voz] / ____ # (MAJOR, 1987: 114)

De acordo com a regra acima, sons vozeados passam a sons surdos em posição final de sílaba, o que se concretiza, formalmente, a partir da mudança do traço [+voz] para [-voz] na plosiva final. MAJOR afirma também que a regra em questão não se mostra operante nem na L2, tampouco na L1 dos aprendizes. Sob a visão do autor, dessa forma, há regras fonológicas que operam apenas no sistema de interlíngua do aprendiz, sem necessariamente estarem presentes na língua fonte ou na língua alvo.

Como dar conta do processo de aprendizagem dessa distinção em coda, através, apenas, de traços binários? Parece-nos que uma abordagem binária não se faz possível, pela característica estática dos traços, que não captam um caráter fundamental da produção dos sons: o componente temporal, ou seja, a temporalidade envolvida na produção (e percepção) da fala. Faz-se necessário, assim, um outro olhar, não somente para a caracterização formal do fenômeno denominado interlíngua, mas, também, para a sua própria descrição e discussão teórica. Por isso, voltaremos a abordar o processo de Dessonorização Terminal em outras seções deste capítulo, revisitando-o em outros modelos. Movemo-nos, dessa forma, na direção de modelos linguístico-perceptuais que, embora nem sempre sejam dinâmicos na base, trazem evidências sobre detalhes fonéticos que podem ser interpretadas à luz dessa posição.

2. Abordagens inovadoras: modelos linguístico-perceptuais de percepção/produção de L2

Com base nas seções precedentes, acreditamos ter já apresentado argumentos contra uma abordagem do processo de

aquisição fonológica de L2 como a aplicação de uma regra X, em que um valor dito mais marcado passa a menos marcado, sob uma perspectiva do tipo ‘sim’ ou ‘não’. Assim, ao destacar o caráter estático de uma concepção de aquisição como “aplicação de regras”, damos procedimento à caracterização da evolução dos estudos de aquisição fonológica de segunda língua.

Os modelos linguístico-perceptuais apresentam uma nova forma de pensar tanto no que diz respeito à perspectiva de concepção do processo de aquisição de L2 como na formalização dos fenômenos desenvolvimentais. As dificuldades pelas quais passam os aprendizes, assim, não são mais concebidas como resultado do efeito de um período crítico delimitado para a aquisição plena de segunda língua. Tampouco os processos de aquisição dos sons são vistos como a substituição de uma categoria estática por outra, sem que se privilegie o contexto de ocorrência dos sons e das influências de tais ambientes.

Desconsiderando a assumpção da existência de um período crítico, os modelos linguístico-perceptuais postulam que o famoso “sotaque” estrangeiro não deve ser visto como decorrente de um programa maturacional, ou seja, do fechamento de uma janela para o aprendizado a partir de uma determinada idade. Pelo contrário: esses modelos interpretam as dificuldades características da produção e percepção oral da L2 como advindas da maneira como os aprendizes percebem os sons; disso resulta o sotaque. Assim, ainda que a capacidade cognitiva do aprendiz se mantenha ao longo de toda a vida – com mudanças previsíveis em termos de plasticidade cerebral (SALTHOUSE, 2009) –, a experiência com a produção e a percepção da língua materna exerce efeitos perceptíveis sobre a aprendizagem da L2.

Assim, os modelos de FLEGE (1995, 2002, 2003) e de BEST

et al. (2001) e BEST & TYLER (2007) são pautados por uma premissa bastante importante na aquisição de L2: os sons da língua estrangeira são categorizados dentro do espectro de sons da L1, durante a percepção dos sons da língua-alvo.

Vejamos cada um dos modelos em questão com maiores detalhes. O modelo denominado *Speech Learning Model* – SLM (FLEGE, 1995, 2002, 2003) desenvolve-se a partir da idéia basilar de construção de categorias fonéticas dos sons da L2 a partir do inventário fonético da L1. Dessa forma, os sons da língua estrangeira que se mostram muito próximos daqueles da L1 caracterizam-se como mais difíceis de adquirir, uma vez que o aprendiz não estabelece uma categoria fonética nova para os sons em questão, pois os interpreta como idênticos aos que se encontram em seu sistema de L1. Por outro lado, sons menos similares foneticamente são mais prováveis de serem distinguidos pelo aprendiz, de modo que seja facilitada a formação de novas categorias e, dessa forma, venha a ser possibilitada a aprendizagem do som em questão.

É importante mencionar, nesse sentido, que a capacidade do aprendiz de estabelecer novas categorias de sons da L1, sob a perspectiva do SLM, mostra-se garantida ao longo de toda a vida do indivíduo, ainda que o grau de acuidade no estabelecimento dessas categorias venha a diminuir em função da perda gradual de plasticidade cerebral do aprendiz. Essa diminuição da plasticidade pode ser confundida com a idade de chegada no país de segunda língua, ou com a idade de início da aprendizagem da mesma. Dentre os diversos estudos conduzidos por Flege e colaboradores acerca dos efeitos da Idade de Chegada do Aprendiz, podemos mencionar o experimento de TSUKADA *et al.* (2005). Com base nos resultados de seus experimentos, os autores afirmam que, ainda que crianças sejam mais bem sucedidas do que os adultos (L1 –

coreano) ao aprenderem as propriedades fonéticas das vogais do inglês, esses aprendizes mirins, mesmo assim, apresentam índices menores de discriminação vocálica, se comparadas às crianças nativas do inglês. Há, assim, um destaque ao ambiente de imersão e experiência dos falantes com a L1 e a L2.

De acordo com o *Speech Learning Model* (FLEGE, 1995, 2002), à medida que a percepção dos fones da L1 vai se desenvolvendo ao longo da infância e da adolescência, a assimilação de aspectos fonéticos das consoantes e vogais da L2 se mostra mais provável. Quando determinados fones da L2 continuam a ser identificados como instâncias de fonemas e alofones da L1, a formação de novas categorias será bloqueada. Entretanto, Flege deixa claro que tais limitações na percepção categórica da fala em L2 se originam da experiência prévia do aprendiz com outras línguas, e não do resultado de um programa maturacional.

Um dos exemplos de achados interessantes de Flege, que ilustram a importância de detalhes fonéticos na aprendizagem da L2, é o caso da Dessonorização Terminal, que foi abordado dentro de uma perspectiva gerativista na seção anterior. No inglês, as plosivas que sofrem o processo de desvozeamento terminal não compartilham as mesmas propriedades acústicas dos segmentos surdos /p/, /t/ e /k/ (FLEGE *et al.*, 1987). Mesmo assim, os falantes nativos do inglês são capazes de distinguir [p] de [b] ou [t] de [d], em posição final. No que diz respeito a segmentos plosivos, o vozeamento na clusura (ou a sua ausência) é uma importante pista acústica no que diz respeito ao contraste entre /p/ e /b/. Os falantes nativos de inglês que participaram do estudo de FLEGE *et al.* (1987) mantiveram um intervalo de vozeamento na clusura de 60 a 120ms na produção de /b/ final, em intervalos de tempo total de clusuras que variavam de 100 a 140 milissegundos. O desvozeamento foi, portanto, parcial. Os autores concluem que

a duração do vozeamento na closura é influenciada por diversos fatores, dentre os quais pode ser mencionada a influência exercida pelo ponto de articulação do segmento. De fato, o vozeamento na closura tende a diminuir à medida que a língua recua. Isso é uma consequência do fato de que, quanto mais posterior o ponto de articulação, mais fácil se torna para que a pressão supraglotal, atrás da obstrução oral, aumente. Para que o vozeamento não cesse, a pressão supraglotal deve ser mantida abaixo do valor referente à pressão subglotal, sendo essa uma condição necessária para que o fluxo de ar que gera o vozeamento possa subir através da glote. De acordo com KINGSTON (2007), para que se possa manter o valor de pressão supraglotal superior ao de subglotal, os falantes tendem a “expandir ativamente a cavidade oral, bem como permitir que tal cavidade se expanda passivamente, através do relaxamento dos músculos que delimitam tal cavidade, durante a pronúncia dos segmentos vozeados”. Nas plosivas vozeadas, tanto a expansão ativa quanto a passiva tendem a ser mais difíceis. De fato, a cavidade oral não fica tão ativamente expandida quanto pode ocorrer na produção de plosivas bilabiais ou alveolares. Além disso, as plosivas velares, quando comparadas com as bilabiais e com as alveolares, possuem a menor área de tecido posterior à closura, fato esse que explica uma expansão de caráter passivo também menor.

O desvozeamento fonético parcial também acontece em inglês em fricativas. SMITH (1997) organizou as produções de /z/, por parte de falantes nativos do inglês, em três categorias: /z/ plenamente desvozeado (0 a 25%); /z/ parcialmente desvozeado (25 a 90%); /z/ vozeado (90 a 100% de vozeamento). No mesmo estudo, também foram apontadas diferenças referentes à duração do próprio segmento fricativo, uma vez que a duração de /z/ se mostrou significativamente

mais curta do que a da sua contraparte sonora /s/. Diferenças na duração da vogal precedente também foram encontradas, de modo que as vogais que seguiam /z/ foram significativamente mais longas do que aquelas que precediam /s/. Assim, em vez da distinção binária entre plosivas vozeadas e desvozeadas, a autora reclassificou as plosivas em três grupos: 1) sonoras vozeadas, 2) sonoras desvozeadas, e 3) desvozeadas.

Os achados relatados acima sugerem a existência de contrastes mais sutis do que os propostos numa visão fonológica de oposições categóricas binárias estabelecidas a partir de traços, como é o caso do + ou - sonoro. De fato, uma abordagem dinâmica da produção e da percepção de fala em L2, como a Fonologia Gestual, dá conta de forma mais adequada dos achados de estudos de interlíngua como os da Dessonorização Terminal, que voltará a ser abordada na seção 4.

Outro modelo linguístico-perceptual é o *Perceptual Assimilation Model* - PAM (BEST, 1993, 1994, 1995). Esse modelo assume a visão gestual, postulando a formação de categorias na L2 a partir de semelhanças e diferenças articulatórias dos sons da L1. Na visão de Best, a noção de semelhança se mostra mais clara, uma vez que remete ao mapeamento articulatório dos gestos; assim, os gestos² da L1 exercem papel fundamental na assimilação, em termos gestuais, dos sons da L2.

Havendo tal “assimilação”, os gestos articulatórios da L2 serão percebidos como se fossem os da L1. Os graus de semelhança articulatória entre os sons determinam a relação entre os fones da L2 e as categorias fonológicas da L1, de modo

² Os gestos são unidades de ação e constrição definidas dinamicamente (BROWMAN & GOLDSTEIN, 1992, 1995). Um gesto consiste na orquestração dos movimentos de vários articuladores, como lábios e mandíbula, a fim de realizar uma tarefa linguisticamente significativa, como a constrição labial.

a caracterizar diferentes graus de assimilação³ entre os sistemas fonológicos das duas línguas. No *Perceptual Assimilation Model*, o grau de semelhança perceptual entre o estímulo da L2 e o fone prototípico, representado simbólicamente (fonologicamente) e fisicamente pelo inventário gestual da L1, mostra-se como um fator importante para o estabelecimento ou não de uma nova categoria perceptual específica para a L2. A transferência dos padrões da L1 para a L2 caracteriza-se, portanto, como decorrente da dificuldade de criação de novas categorias prototípicas para os sons da L2, que acarreta dificuldades em estabelecer a separação funcional das categorias de sons das duas línguas.

Entretanto, é importante mencionar que, ao invés de focalizar a percepção do aprendiz de L2, o Modelo de Assimilação Perceptual dá ênfase a monolíngues do tipo *naïve*⁴, fato esse que constitui uma importante diferença – embora geralmente não-notada pela maioria dos estudos da área – entre o modelo de Best e o SLM, de Flege. Embora os termos “percepção não-nativa” e “percepção da fala em L2” sejam geralmente tratados como intercambiáveis, tais termos estabelecem predições acerca de dois grupos de participantes radicalmente distintos: os ouvintes *naïve* possuem nenhum, ou pouquíssimo, contato com a fala da L2, enquanto os aprendizes de L2 se distinguem em função de sua experiência com a L2 (BEST & TYLER, 2007).

Em suma, ambos os modelos se desenvolvem sob a premissa de que a percepção é condição *sine qua non* para a

³ O termo ‘assimilação’, conforme proposto pela autora, deve ser visto como a identificação de um som da L2 dentro de uma categoria fonética pré-existente da L1, de modo a dificultar, no caso da aprendizagem de uma nova língua, a discriminação entre os sons das duas línguas.

⁴ Os falantes *naïve* são aqueles que, diferentemente dos aprendizes de uma segunda língua, desconhecem a língua a que estão sendo expostos nos experimentos.

produção, chegando mesmo a motivar, no nível segmental, os desvios de produção da fala em língua estrangeira. O grau de semelhança entre os sons da L1 e os da L2 são determinantes de tal percepção, de modo a influenciarem o processo de formação de novas categorias de sons da L2. Entretanto, ainda que ambos os modelos incorporem a noção de percepção ao processo de aquisição de segunda língua, o modelo de Flege ainda se desenvolve a partir de primitivos de tempo extrínseco, embora acresça dados fonéticos à representação fonológica, ao passo que o modelo de Best parte de uma base gestual.

Assim, podemos estabelecer um elo entre a concepção de aquisição de L2 expressa por Best e a noção de gesto como primitivo de tempo intrínseco, base da FAR - Fonologia Articulatória (BROWN & GOLDSTEIN, 1992) e da FAAR – Fonologia Acústico-Articulatória (ALBANO, 2001). Nosso próximo passo, então, é partir para uma proposta da aprendizagem da L2 baseada na visão dos Sistemas Dinâmicos.

3. Uma proposta dinâmica para a aquisição da L2

Como se sabe, todas as línguas codificam informação linguística na ordenação de eventos no fluxo da fala. O fato de que a sequência informacional é fundamental para a fala foi enfatizado por estudos recentes, que mostram que tanto bebês como adultos conseguem captar regularidades, probabilisticamente, na ordenação de eventos como a sílaba em fluxos de novos estímulos na fala (PELLUCHI *et al.*, 2009; SAFFRAN, 2003). Contudo, devemos destacar que a ordenação sequencial não é o único tipo de informação temporal relevante para a produção e a percepção da fala. A informação linguística também é codificada por padrões presentes na orquestração rítmica de eventos no fluxo da fala (KELLO, 2003: 619), ou seja,

nas regularidades observadas nas quantidades de tempo real que ocorrem na intrincada dinâmica dos eventos de fala, que podem ser sequenciais ou parcialmente sequenciais (coarticulação), levando a uma sobreposição distribuída na informação temporal.

Quando se olha para a orquestração de eventos de fala numa perspectiva dinâmica, a produção/percepção da fala em L2 transcende a aprendizagem fonológica ou mesmo linguística propriamente dita. É aí que se propõe uma abordagem multimodal para a aquisição da língua materna e da língua estrangeira. A multimodalidade da cognição se manifesta numa gama variada de processos e mecanismos corporificados⁵, mas o foco aqui é a linguagem e, dentro do domínio linguístico, destacamos a relação multimodal entre a produção e a percepção da L2. Evidências empíricas recentes sugerem a existência de um conjunto comum de mecanismos cerebrais que subjazem tanto às funções da percepção quanto da produção da fala (CORBALLIS, 2003, 2010; GENTILUCCI & CORBALLIS, 2006; RIZOLLATTI & CRAIGHERO, 2004; PLAUT & KELLO, 1999). A maior parte desses achados diz respeito à função dos neurônios-espelho na mediação da representação da percepção e produção do movimento manual e bucal, bem como à origem braquiomanual⁶ da linguagem, e indicam que o conhecimento e as

⁵ Consultar KLATZY, BEHRMAN & MACWHINNEY (2008) para um relato detalhado acerca da questão da corporificação – *embodiment* - em vários fenômenos cognitivos.

⁶ Uma proposta recente sobre a evolução da linguagem do modo fácio-bráquio-manual para o oral (ARBIB, 2006) postula que a função original de busca do significado simbólico a partir da ação manual exigiu, em algum momento da evolução, uma maneira de distinguir os movimentos semelhantes de face, mão ou membros. Assim, a combinação de ações vocais intrinsecamente distintas (mas sem sentido) entre membros de um conjunto de gestos icônicos manuais pode ter fornecido os meios necessários para desambiguar proto-palavras manuais bastante similares. Dessa forma, os gestos vocais tornaram-se o principal meio de distinção entre palavras, ao passo que os gestos manuais foram reduzidos em importância, de modo a desempenharem o papel que assumem na comunicação contemporânea (CORBALLIS,

habilidades aprendidas para a percepção da fala também são empregados na sua produção (e vice-versa). Portanto, se for levada em consideração a premissa de que uma maneira otimizada de compartilhar a aprendizagem em diferentes modalidades implica compartilhar também mecanismos responsáveis por essas habilidades, o estudo da produção e da percepção da L2 pode se beneficiar imensamente de uma abordagem gestual multimodal para a aquisição da segunda língua.

A hipótese dos mecanismos compartilhados auxiliou muito a busca pelos aspectos fundamentais relacionados ao ritmo, ao *timing* na aquisição, percepção e produção da fala (KELLO, 2003: 620). De acordo com essa hipótese, alguns dos mecanismos que processam a orquestração temporal dos eventos envolvidos na percepção também seriam responsáveis pelo *timing* de eventos ligados à produção da fala. Isso permite o acoplamento, em tempo real, entre diferentes domínios, em que a dinâmica de um sistema influencia o *timing* de outro. Conforme PORT (2009), os seres humanos frequentemente acoplam vários sistemas, como ocorre quando estão dançando ao ritmo de uma música. Nesse caso, de acordo com o autor, o sistema de percepção auditiva do sujeito é acoplado ao som ambiente, e o sistema motor grosso é acoplado tanto à audição como aos sons musicais.

HOMMEL & MILLIKEN (2007) também levantam a questão de como o sistema perceptual se comunica com os planos de ação. Contudo, conforme discorrem os autores, os perceptos são codificados como representações de características do *input*

2003). Em um estágio posterior, gestos vocais vieram a constituir o método principal de distinguir palavras (MCNEILL, 1992).

de estímulo, ao passo que os planos de ação são codificados como especificação motora. Como, então, conceber a produção e a percepção da fala de modo a dar conta do fato de que a maneira como a fala é percebida pode alterar o modo como a fala é produzida? Essa questão pode ser respondida a partir da premissa de que tanto falantes como ouvintes recebem informação multimodal relativa às formas linguísticas e usam essa informação em todas as modalidades informacionais possíveis:

O uso da informação recebida em todas as modalidades informacionais requer uma moeda comum usada em todas as modalidades sensoriais. A moeda comum poderia ser a representação supramodal de MELTZOFF & MOORE (1997, 1999) ou, mais diretamente, o objeto distal da percepção. Essa mesma moeda comum é necessária para que se compreenda como ocorre a ação perceptualmente guiada e como acontece a comunicação entre falantes e ouvintes (FOWLER 2004: 200).

Se essa moeda comum, de uso público (GOLDSTEIN & FOWLER, 2003; FOWLER & GALANTUCCI, 2005), relacionada aos mecanismos compartilhados que subjazem ao movimento sensório-motor do corpo cognitivo⁷, for o gesto, então a Teoria dos Sistemas Dinâmicos pode ser aplicada à emergência e mudança de eventos que se desenrolam no tempo, como é o caso da fala. Mas o que é exatamente um sistema dinâmico?

⁷ O uso do termo 'corpo cognitivo' deve-se à tentativa de referir, no jargão da dinâmica fônica, a continuidade entre a cognição corporal e mental sem refletir dicotomias já cristalizadas na literatura clássica, tanto psicolinguística como fonológica.

3.1. Sistemas dinâmicos

De acordo com VAN GELDER & PORT (1995), um sistema dinâmico pode ser definido como um conjunto de variáveis (como distâncias, velocidades, ativações, taxas de desenvolvimento e mudança, entre outras) que se alteram simultaneamente no transcorrer do tempo devido a influências mútuas entre as variáveis. Os modelos dinâmicos têm por objetivo especificar como e onde ocorrem as mudanças de estados de um sistema.

Há vários construtos importantes na Teoria dos Sistemas Dinâmicos (TSD). Em um sistema como o da produção oral, ao se identificarem as partes do sistema relevantes (como a posição da mandíbula, da língua e dos lábios), podem-se atribuir valores numéricos ao estado atual desses articuladores, e plotar, em um gráfico dimensional: 1) a posição, num eixo, para a mandíbula, 2) em outro eixo, a posição para a língua, e 3) um terceiro eixo, a posição do lábio inferior. Esse gráfico permite que se visualize o modo como cada um desses articuladores se movimentam, ou muda seu estado no transcorrer do tempo. Essa representação, que neste caso é tridimensional, é usualmente chamada de “espaço de estados” (BROWMAN & GOLDSTEIN, 1986, 1992; ELMAN, 1998).

Outro exemplo de um construto importante na TSD é o de “atrator”. Um atrator é um estado em direção ao qual, em condições normais, um sistema dinâmico tenderá a chegar ou se aproximar. Uma criança brincando num balanço de uma pracinha, por exemplo, constitui um sistema dinâmico com um atrator que tem o balanço e a criança na posição vertical, de descanso. O balanço, então, pode oscilar para frente e para trás se a criança é empurrada ou se balança com a ajuda de suas pernas, mas há uma força atratora que puxa a criança para trás em direção à posição de descanso. O objetivo de uma análise de

sistemas dinâmicos dessa situação seria descrever o comportamento do sistema usando equações matemáticas que podem nos informar como o estado do sistema (neste exemplo, a posição da criança num determinado momento) muda ao longo do tempo (ELMAN, 1998: 27).

Esses construtos atinentes aos sistemas dinâmicos são fundamentais para que compreendamos a produção e a percepção da fala em L2 sob uma perspectiva dinâmica dentro do arcabouço da Fonologia Gestual. Esse é o assunto da próxima seção.

3.2. A produção e a percepção da fala em L2 na abordagem da Fonologia Gestual

A Fonologia Gestual parte da premissa de que a fala pode ser decomponível em unidades de ação do trato vocal, os gestos. Essas unidades de contraste, ou fonológicas, são isomórficas às unidades contínuas, ou fonéticas (GOLDSTEIN *et al.*, 2006). Nessa abordagem, os gestos articulatórios constituem, simultaneamente, unidades de ação (codificando os movimentos articulatórios para a formação de constrictões no trato vocal) e de informação (codificando contraste). Como essas unidades informacionais e de ação incorporam o simbólico e o concreto, não há necessidade de tradução entre unidades de representação e sua execução como tarefas de fala. Num nível organizacional, as tarefas de constrictão são controladas, e não o movimento de articuladores individuais (SALTZMAN & MUNHALL, 1989). Esse controle é alcançado ativando e desativando pontos atratores no espaço da tarefa definido dentro do trato. A dinâmica de atratores pontuais constitui um tipo de sistema interessante para caracterizar o controle de movimentos discretos por várias características: 1) porque é possível associar um determinado

ponto no espaço da tarefa com o alvo do movimento discreto; 2) porque o alvo é atingido a despeito das condições iniciais ou das perturbações ao longo da trajetória. Esse comportamento resistente a perturbações é observado para gestos de fala tanto no domínio temporal como no espacial (BYRD, 2006).

A hipótese da existência de unidades gestuais de ação na produção da fala pode gerar várias previsões acerca dos modos em que um enunciado pode exibir sotaque durante a produção na língua estrangeira, através do tratamento dado a erros outros em língua materna. Assim, os *lapsus linguae*⁸ são vistos de maneira distinta da abordagem simbólica, que enxerga os erros como trocas ou substituições de unidades simbólicas num plano que é, depois da manipulação do símbolo, executado normalmente. As previsões de uma abordagem gestual dos lapsos de fala são as seguintes: se as unidades fonológicas atômicas são gestos com forças de ativação que variam no tempo, tais erros podem resultar de ativação (parcial ou completa) de uma unidade gestual num tempo inapropriado durante a produção. Além disso, há também gradientes decorrentes da magnitude dos gestos. Esses erros gradientes podem não ser representados em corpora de erros transcritos. Sugestões de que gradientes decorrentes de lapsos realmente ocorrem são encontrados em vários estudos sobre erros ou *slips* na língua materna, envolvendo diversos tipos de métodos: MOWREY & MCKAY (1990) utilizaram Eletrografia Magnética (EGM), FRISCH & WRIGHT (2002) usaram dados de fonética

⁸ Como exemplos de lapsos da língua, GOLDSTEIN *et al.* (2006, p. 388) fornecem a tentativa de produzir do sintagma 'coffee pot' como 'poffee cot' ou 'poffee pot', mas não como 'cottee poff', dado o fato de que as trocas dos segmentos envolvidos tendem a preservar a mesma posição na sílaba da expressão original. De fato, conforme também apontam os autores, consoantes são apenas trocadas por consoantes e vogais, por sua vez, apenas por vogais. Dessa forma, os lapsos de língua são sistemáticos em termos de ocorrência e de distribuição, sendo, portanto, um portal para o estudo das unidades atômicas da fala, tais como o gesto articulatório.

acústica e, mais recentemente, GOLDSTEIN *et al.* (2007) usaram dados de cinética articulatória. Nesse último estudo, Goldstein e colegas demonstraram que o tipo de lapso mais comum é a intrusão inadequada de um gesto durante a produção de outro gesto apropriado. As intrusões encontradas eram bastante variáveis e gradientes em suas magnitudes espaciais.

Algumas conclusões advindas dos achados encontrados em estudos sobre erros envolvendo a língua materna podem ser transpostas para o que provavelmente subjaz à produção oral na língua estrangeira. A partir da afirmação de GOLDSTEIN e colegas (2007) de que um ou mais gestos produzidos em relações temporais deslocadas não resultam em movimentos aleatórios, podemos depreender que, dependendo da maneira como uma pessoa produz uma palavra, sintagma ou enunciado em língua estrangeira, alguns interlocutores conseguem identificar corretamente sua língua materna. Segundo GOLDSTEIN e colegas (2007), as ações gestuais errôneas: 1) preservarão suas propriedades como unidades linguisticamente significantes, bem como 2) estarão diretamente relacionadas à estrutura gestual.

Ora, erros de produção oral em língua estrangeira, além de impedirem a compreensão do(s) interlocutor(es), parecem estar ligados, em maior ou menor grau (dependendo do nível de proficiência do falante), à estrutura gestual da língua materna, tanto no que diz respeito à produção de gestos individuais de construção, como em agrupamentos gestuais maiores, que podem ser sequências gestuais ou sobreposições (como é o caso da coarticulação) envolvendo acoplamento de vários gestos. Além disso, as ações gestuais errôneas são sensíveis a contextos dinâmicos e temporais mais abrangentes, como a taxa de elocução, e também à base rítmica inerente à língua, como a estrutura métrica e a alternância entre vogais e consoantes.

Os modelos de oscilação dos sistemas dinâmicos apresentam: modos preferidos de coordenação e estágios de transição característicos para esses modos. A sincronização, tanto em termos de faseamento como de desfaseamento dos gestos articulatórios, pode exercer efeitos sobre a percepção do aprendiz, e, dessa forma, moldar um sistema de L2 visto à luz do da L1. Além disso, a sincronização dos gestos também pode se mostrar dificultosa para a produção da sílaba da L2 – envolvendo a orquestração entre gestos de vogais e consoantes, tanto em termos de sequência como na sua coprodução. Por fim, as relações temporais de fase e antifase envolvidas na orquestração gestual da L1 podem, durante a produção de todo um enunciado na língua estrangeira, influenciar a produção dos padrões rítmicos da fala da L2, pois a relação entre o *timing* que orquestra a organização temporal das constelações gestuais na fala em L1 e em L2 parece ser o fator de maior peso no chamado ‘sotaque estrangeiro’.

Assim, podemos inferir que o acoplamento e o *entrainment*⁹ (carreamento) entre diferentes modos oscilatórios de sistemas de fala (o da L1 e o da L2, por exemplo), pode fundamentar a ideia de acento ou sotaque. Modos preferidos de coordenação oscilatória na produção gestual da L1 podem provocar desfaseamento entre gestos durante a produção da L2. Quando o aprendiz da L2 ainda tem pouca fluência, o sistema dinâmico

⁹ O *entrainment* ou carreamento refere-se ao modo como dois sistemas oscilatórios em interação (como dois pêndulos, por exemplo), que têm períodos diferentes quando funcionam independentemente, entram em fase, funcionando no mesmo período. Os dois osciladores podem entrar em sincronia, mas outras relações de fase também são possíveis.

da L1, dotado de todos os atratores característicos da língua materna, age sobre o sistema da L2, modificando a dinâmica da tarefa na produção da língua estrangeira ao enviesar os atratores característicos da L2 no espaço de estados articulatórios do falante não monolíngue, de modo a causar a fala com sotaque.

Entre falantes de nível mais elevado de proficiência na fala em L2 pode ocorrer a deriva gestual, muito bem documentada por SANCIER & FOWLER (1997). Quando o falante proficiente vive num país no exterior por algum tempo, em regime de imersão linguística, seu sistema de produção de fala e, por conseguinte, a maneira de orquestrar as variáveis do trato começa a entrar em *entrainment* com o sistema da L2. Assim, pode ocorrer que, ao falar a língua materna, ele produza a deriva (*drift*) gestual para além da orquestração gestual de alguns sons e/ou do ritmo da L1 em alguns momentos, uma vez que sua produção gestual está com seus atratores configurados de acordo com a orquestração gestual que se aproxima mais dos padrões da L2, devido à imersão ao longo do tempo.

Chegamos, assim, a uma caracterização da aquisição de segunda língua que, além de reconhecer a relação fundamental entre percepção e produção da fala, através de uma moeda comum aos dois domínios – o gesto –, consegue explicar a aquisição fonológica de modo a caracterizá-la como um processo dinâmico. Esse construto permite a explicação dos efeitos da experiência linguística (na produção da L1 ou da L2) sem que seja necessário recorrer a mecanismos maturacionais e a noções como período crítico. A visão dinâmica proporciona, também, uma nova noção de erro ou sotaque na L2 que vai ao encontro das propriedades do primitivo da Fonologia Gestual, o gesto, que é dotado de tempo intrínseco. Sob essa concepção, a caracterização do processo de aquisição da língua estrangeira não se resume a uma perspectiva de presença ou ausência de um

traço ou segmento, mas, sim, da sincronização gestual.

Essa observação a respeito dos ajustes temporais dentro da sílaba remete-nos, mais uma vez, ao estudo da Dessonorização Terminal, já abordado, neste artigo, dentro de uma visão clássica (1.2), vindo a ser abordado, na seção que segue, numa perspectiva dinâmica.

4. O desvozeamento terminal de obstruintes sonoras na interlíngua de aprendizes brasileiros de inglês: uma abordagem dinâmica

Conforme já vimos na seção 1.2, o Processo de Desvozeamento Terminal, sob a regra apresentada na seção em questão, implica a perda do traço [+voz], através de uma regra de substituição, pelo traço [-voz]. Sob tal perspectiva de substituição plena, expressa através da caracterização de MAJOR (1987), há, na interlíngua português-inglês, uma neutralização total, no que diz respeito à distinção entre obstruintes surdas e sonoras finais.

ZIMMER & ALVES (2007, 2008) revisitaram o processo de Desvozeamento Terminal (doravante DT) na Interlíngua Português-Ingês e, sob a perspectiva aqui advogada, considerando a influência de mais de uma pista acústica para a caracterização da distinção fonológica funcional entre unidades acústicas, verificaram a impossibilidade da caracterização de tal fenômeno de forma discreta. Em suas investigações, ZIMMER & ALVES (2008) questionaram, sobretudo, se o processo de DT implicava neutralização total de vozeamento, de modo a permitir sua caracterização em termos categóricos. Ao considerarem o papel de múltiplas pistas acústicas na caracterização de um fenômeno fonético-fonológico, os autores questionaram, ainda, se a duração da vogal precedente ao

segmento vozeado pode ser usada, assim como é por parte de falantes nativos do inglês, como um indicador do contraste entre pares como ‘cap’ e ‘cab’, por exemplo. No estudo de 2008, que contou respectivamente com oito participantes brasileiros e 3 estrangeiros, os autores relataram dois achados principais: 1) o grupo de participantes brasileiros produziu uma percentagem de vozeamento na clausura que não demonstra neutralização de contraste entre plosivas surdas e sonoras na língua inglesa; 2) os participantes brasileiros produziram vogais mais longas antes de plosivas sonoras do que aquelas produzidas antes de plosivas surdas; contudo, a duração das vogais antecedendo plosivas surdas foi maior no grupo brasileiro comparado ao norte-americano. Esses achados serão retomados mais adiante, na comparação com os dados do estudo relatado a seguir.

Considerando a noção dinâmica discutida no presente artigo, reportamos aqui um novo estudo, que expande e aprofunda os resultados dos estudos de ZIMMER & ALVES (2007, 2008). Este novo estudo contou com 15 participantes, falantes brasileiros aprendizes de inglês, sendo 7 com proficiência intermediário-superior e 8 de proficiência avançada, conforme apontado pelo teste de nivelamento TOEIC (*Test of English for International Communication*). A investigação contou, ainda, com dados de 5 falantes nativos do inglês norte-americano, provenientes da região nordeste dos Estados Unidos, cujas produções, obtidas na cidade de Amherst-Massachusetts, possibilitaram a verificação do fenômeno no falar nativo.

O instrumento de coleta de dados consistiu de uma tarefa de leitura de frases, contendo as seguintes palavras-alvo: (a) seis itens lexicais encerrados por plosivas vozeadas: *bob, cab, bad, pad, rag, bag*; (b) duas palavras encerradas por fricativas vozeadas, cuja grafia poderia incitar o processo de transferência grafo-fônico-

fonológica: *bobs, falls*; (c) oito pares mínimos ou análogos, distintos quanto ao vozeamento, de (a) e (b), encerradas por plosivas surdas, para possibilitar a comparação entre os pares mínimos: *bop, cap, bat, pat, rack, back*.

As palavras acima descritas foram apresentadas em duas diferentes frases-veículo, visando à verificação dos efeitos do contexto seguinte ao segmento obstruinte final: (a) *Say _____* (contexto anterior a pausa); (b) *Say _____ again* (contexto anterior a vogal). Cada palavra foi repetida 5 vezes em cada contexto. As frases foram apresentadas, aleatoriamente, através de *slides* confeccionados no software power-point, na tela de computadores do tipo laptop (modelo Vaio PCG-V505ECP, para as gravações no Brasil, e modelo Dell Inspiron 1505, para as gravações realizadas com os falantes nativos, nos Estados Unidos). Os dados dos participantes brasileiros foram coletados em uma cabine de isolamento acústico, no Campus do Curso de Jornalismo da universidade de um dos pesquisadores. As gravações dos dados dos sujeitos americanos foram realizadas em uma sala silenciosa, na residência de um dos pesquisadores, através do software Adobe Audition 2.0. Os dados foram digitalizados a uma frequência de 44.1 kHz.

As palavras-alvo foram segmentadas com o software Audacity 1.2.6 e, posteriormente, submetidas à análise acústica através do software Praat – version 4.6.02 (BOERSMA & WEENINK, 2007). Foram realizadas as medições dos seguintes parâmetros acústicos: (1) duração da closura dos segmentos plosivos; (2) percentagem de vozeamento na closura; (3) duração da vogal precedente.

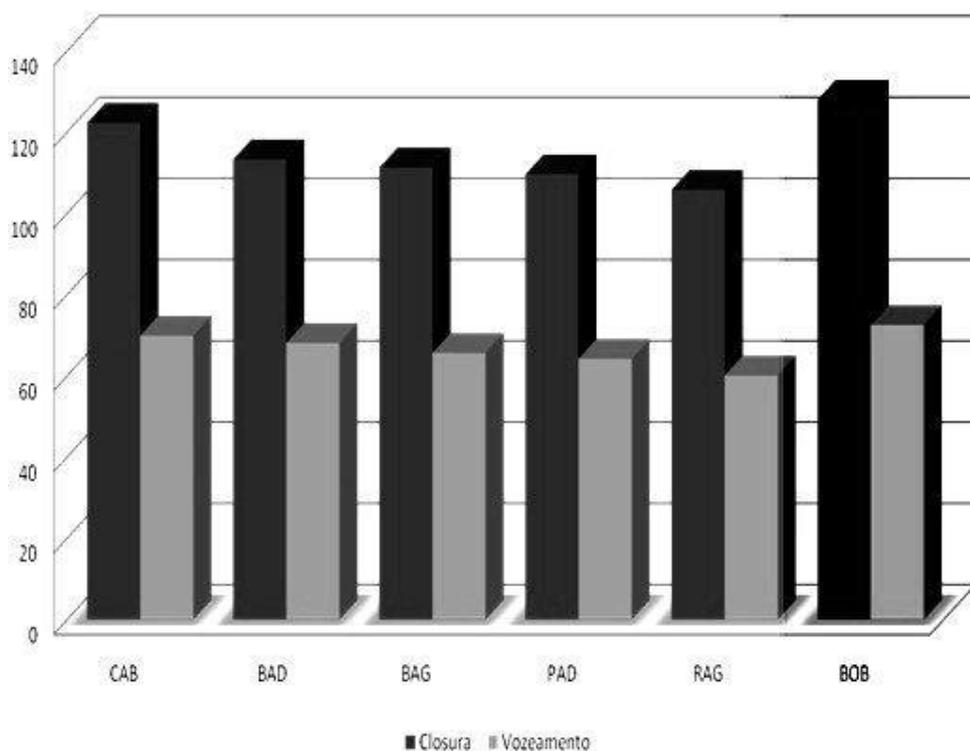
Para a descrição dos resultados, focalizaremos, separadamente, os dois contextos do estudo. Iniciaremos pela descrição dos dados produzidos anteriormente ao contexto de

pausa. Na seção 5.2, discutiremos o contexto pré- vocálico.

4.1. Contexto anterior à pausa

No que diz respeito ao contexto anterior à pausa (“Say _____.”), o Gráfico 1 apresenta a relação entre o tempo total de clusura e a duração de vozeamento, por parte dos aprendizes brasileiros, nas palavras encerradas por segmentos plosivos.

GRÁFICO 1 - Intervalos de clusura e de vozeamento nas obstruintes vozeadas em contexto anterior à pausa, produzidas pelos aprendizes brasileiros



O Gráfico 1 deixa evidente que, em todas as palavras-alvo, há, sim, um intervalo de tempo com vozeamento na closura, intervalo de tempo esse que chega, no caso de certas palavras, a ultrapassar mais de 50% do tempo total de closura. Ao ressaltar a constância desse intervalo de vozeamento, podemos aventar a hipótese de que os aprendizes estejam utilizando uma espécie de estratégia de desvozeamento da porção final da closura para dessonorizar o estouro, o que poderia mascarar uma possível vogal epentética seguinte, advinda da não-aquisição dos padrões silábicos da língua-alvo. Fica clara, desde o presente momento, a incompatibilidade de uma caracterização de neutralização plena de vozeamento, conforme era advogado através da regra apresentada em 1.2.

Frente aos dados apresentados no Gráfico 1, é natural indagar se o vozeamento parcial não poderia ser afetado, também, pela própria duração da closura do segmento plosivo. Foi feita, assim, uma comparação entre os intervalos de closura produzidos pelos aprendizes brasileiros e aqueles produzidos pelos falantes nativos de inglês, que é apresentada na Tabela 1.

TABELA 1 - Tempos de closura e de vozeamento nas obstruintes vozeadas em contexto anterior à pausa nos grupos de brasileiros e americanos

Palavras	Grupo	N	Médias	DP	Teste t (sig bicaudal)
CAB - closura	BR	15	122,27	25,20	2,97
	USA	5	87,66	8,31	(p=0,008*)
CAB - vozeamento	BR	15	69,89	24,69	-0,26
	USA	5	72,98	15,19	(p=0,797)
PAD - closura	BR	15	109,63	29,59	3,13
	USA	5	67,32	3,43	(p=0,006*)
PAD - vozeamento	BR	15	64,19	26,34	-0,25
	USA	5	67,22	3,52	(p=0,804)
BAG - closura	BR	15	111,27	43,84	1,73
	USA	5	76,53	10,55	(p=0,102)
BAG - vozeamento	BR	15	65,61	28,55	1,36
	USA	5	45,93	25,75	(p=0,004*)
BOB - closura	BR	15	127,98	24,12	3,27
	USA	5	90,26	14,51	(p=0,004*)
BOB - vozeamento	BR	15	72,5	25,67	-0,67
	USA	5	81,28	24,87	(p=0,513)
RAG - closura	BR	15	105,81	21,92	2,21
	USA	5	83,45	6,91	(p=0,040*)
RAG - vozeamento	BR	15	60,04	27,48	-0,29
	USA	5	63,79	11,07	(p=0,773)
BAD - closura	BR	15	113,29	31,20	3,59
	USA	5	61,20	12,00	(p=0,002*)
BAD - vozeamento	BR	15	68,11	32,77	0,40
	USA	5	62,03	7,33	(p=0,691)

A observação da Tabela 1 indica também que os tempos de closura produzidos por falantes brasileiros são, no caso de todas as palavras pesquisadas, mais longos do que aqueles apresentados por falantes americanos. Percebe-se que, à exceção de 'bag', houve diferença significativa de duração da closura produzida por brasileiros e americanos em todas as outras palavras terminadas em plosivas vozeadas, a saber: 'cab', 'pad', 'bob', 'rag', 'bad'.

Além de apresentar os tempos de closura produzidos por falantes nativos e não-nativos, a tabela acima apresenta, ainda, os índices de vozeamento produzidos pelos falantes de ambos os grupos. Através da comparação das taxas de vozeamento entre americanos e brasileiros, vemos que mesmo o falar nativo não apresenta intervalos de vozeamento muito superiores àqueles produzidos pelos aprendizes de L2, pois não foi apurada diferença significativa de tempo de vozeamento entre o grupo de brasileiros e de norte-americanos em nenhuma das palavras. De fato, conforme podemos observar no caso das palavras 'bad' e 'bag', a média de tempo de vozeamento, realizada por brasileiros, é superior à encontrada nos dados dos informantes americanos. Fica evidente nos dados, portanto, que a dessonorização parcial ocorre, também, no próprio falar nativo inglês.

Os dados apontados na Tabela 1 são altamente elucidativos, pois indicam que, ao invés de uma dessonorização no grupo de brasileiros, o que ocorre é uma produção extralonga da closura. Isso parece indicar a inversão da lógica presente na noção de neutralização clássica: não há 'perda' ou 'neutralização' de vozeamento na produção de brasileiros de níveis mais avançados de proficiência, uma vez que a quantidade de vozeamento é equivalente nos dois grupos, mas há, sim, uma compensação, um alongamento do tempo da closura.

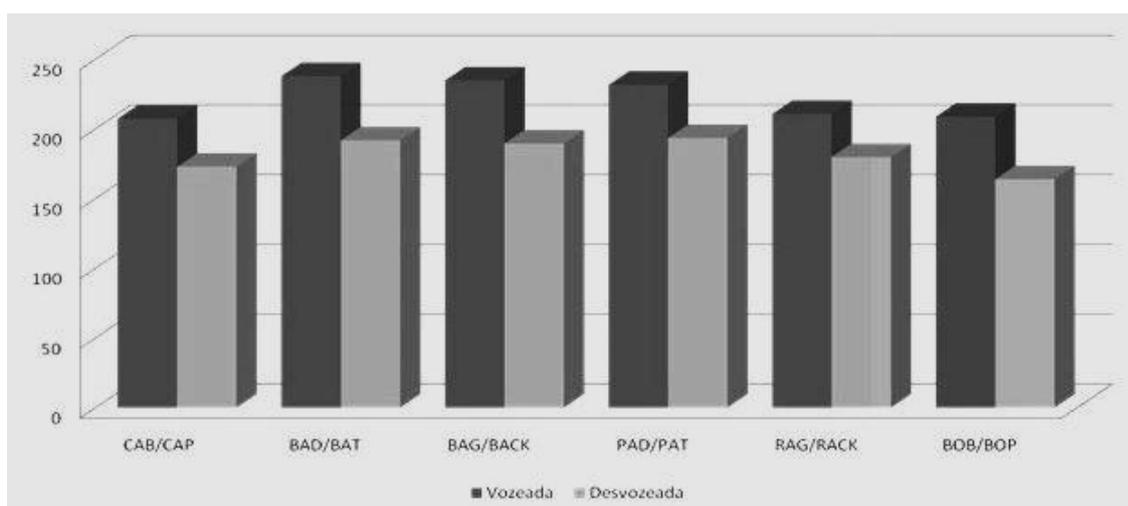
Esses dados vão, também, ao encontro da noção de aquisição fonológica de L2 como ajuste da fala ao *timing* da língua-alvo. Como os segmentos plosivos não ocorrem em posição final de palavra no PB, há uma desestabilização rítmica na produção dos aprendizes; o tempo exagerado de duração da clusura pode ser uma estratégia acústica usada para compensar a não-produção de um segmento vocálico final. Tal estratégia compensatória demonstra o caráter interlinguístico da produção dos aprendizes: se, por um lado, vogais finais já não são produzidas na sua fala em L2, por outro, o segmento plosivo ainda não apresenta o *timing* daquele encontrado nas formas-alvo. Tal diferença de *timing* pode exercer, sim, influências no grau de funcionalidade da distinção surdo/sonoro: com um tempo de clusura maior, a relação tempo de clusura/tempo de vozeamento diminui, de modo que a percentagem de vozeamento na clusura seja reduzida e, dessa forma, haja o risco de os segmentos em questão virem a ser percebidos como surdos. Frente a tal constatação, é ressaltada, mais uma vez, a incompatibilidade de uma caracterização da produção dos segmentos obstruintes finais /β/, /δ/ e /γ/ sob um rótulo binário, descritos, erroneamente, como se fossem equivalentes às contrapartes surdas /p/, /t/ e /k/.

Como a dessonorização parcial, que ocorre na interlíngua português-ínglês, pode também ser verificada na produção nativa, para se falar em perda ou não de oposição funcional entre segmentos surdos e sonoros, muito mais do que a verificação da presença ou ausência de vozeamento na clusura, é necessário analisar a ação das diversas pistas acústicas utilizadas pelos aprendizes para o estabelecimento da distinção.

Ao considerar a necessidade de observação de diversas pistas acústicas para a observação da diferença funcional entre pares mínimos tais como ‘cap’ e ‘cab’, medimos também a

duração da vogal que antecede o segmento plosivo. Conforme aponta a literatura (LADEFOGED, 1993; KINGSTON & DIEHL, 1994), a duração da vogal revela-se como fato importante na distinção entre os segmentos surdos e sonoros por falantes do inglês, fato esse já discutido neste artigo. O Gráfico 2, a seguir, apresenta os resultados da verificação das durações dos segmentos vocálicos que antecedem segmentos surdos e sonoros, na interlíngua português-inglês.

GRÁFICO 2 - Duração das vogais antecedentes às oclusivas vozeadas e desvozeadas em contexto anterior a pausas, produzidas pelos informantes brasileiros

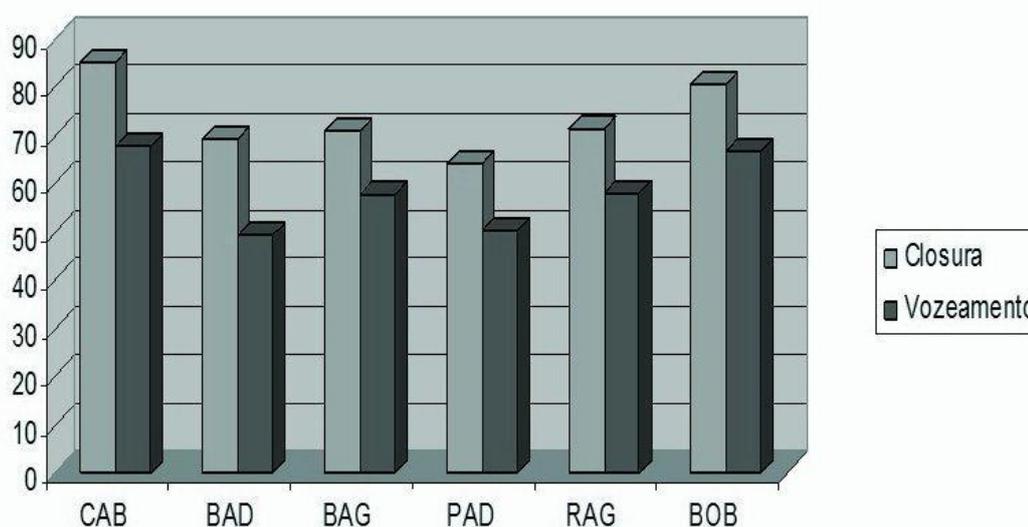


Verificamos, através do gráfico acima, que a duração da vogal que antecede os segmentos /β/, /δ, /γ/ é sempre maior do que a encontrada nas vogais seguidas por /p/, /t/ e /k/, tendência essa que caracteriza o falar nativo. Dessa forma, parece que os sujeitos se encontram em vias de aquisição, também, da distinção demarcada por essa pista acústica. Constatamos, assim, argumentos adicionais para uma perspectiva que veja o *timing* dos segmentos, a dinâmica de produção de sons, sem caracterizar categoricamente a distinção como “adquirida” *versus* “não-adquirida”. No que segue, passamos à discussão do contexto anterior a vogais.

4.2. Contexto anterior à vogal

No que diz respeito ao contexto anterior a vogais (“Say _____ again.”), os resultados condizem com a perspectiva dinâmica advogada ao longo deste texto. O Gráfico 3 apresenta a relação entre tempo de clusura e tempo de vozeamento, nas palavras produzidas pelos aprendizes brasileiros de inglês.

GRÁFICO 3 - Intervalos de tempo de clusura e de vozeamento dentro da clusura, em segmentos em contexto pré-vocálico, produzidos pelos informantes brasileiros



Observamos, no Gráfico 3, que a percentagem de vozeamento na closura, no contexto anterior à vogal, mostra-se bastante superior àquela encontrada frente a plosivas que antecedem pausas. Deve ser ressaltado, novamente, o fato de que todas as palavras-alvo apresentavam a estrutura CVC (ex: 'cab'). Podemos aventar a hipótese de que, em função da ausência de pausa fonética entre a plosiva final e a vogal seguinte, a soltura da consoante ocorre sobre o segmento vocálico, comportando-se, em termos silábicos, como se fosse o *onset* de uma nova sílaba. Seria atenuada, assim, a dificuldade de produção dos segmentos em posição anterior à pausa.

Na análise das medidas de duração da closura nesse contexto, observamos que, quando seguidas por uma vogal, as obstruintes produzidas por brasileiros apresentavam uma duração de closura significativamente menor, como observamos na Tabela 2.

TABELA 2 - Comparação das durações de closura nos contextos anterior a pausa e pré-vocálico no grupo brasileiro

Palavras	N	Médias	DP	Teste t (sig bicaudal)
CAB_pausa	15	122,27	25,20	4,77
CAB_V		85,50	22,51	(p=0,000*)
PAD_pausa	15	109,63	29,59	5,66
PAD_V		64,30	21,80	(p=0,000*)
BAG_pausa	15	111,27	43,84	3,77
BAG_V		70,92	15,18	(p=0,002*)
BOB_pausa	15	127,98	24,12	6,29
BOB_V		80,58	13,99	(p=0,000*)
RAG_pausa	15	105,90	22,75	3,89
RAG_V		71,59	18,90	(p=0,002*)
BAD_pausa	15	115,89	30,63	3,68
BAD_V		69,33	30,78	(p=0,003*)

Novamente, é confirmada a menor dificuldade de produção dos segmentos plosivos quando seguidos, sem pausas, por uma vogal. A duração da closura deixa de ser exageradamente longa, uma vez que a sequência gestual a ser produzida é equivalente a uma que ocorre no PB – ou seja, obstruinte seguida de vogal.

Com relação à duração da vogal que antecede segmentos surdos e sonoros, no contexto em que a plosiva antecede um segmento vocálico, verificamos, também neste contexto, uma duração maior frente aos membros vozeados dos pares

mínimos, conforme evidencia a Tabela 3.

TABELA 3 - Comparação das durações dos segmentos vocálicos que antecedem as plosivas surdas e sonoras seguidas de vogal, produzidos pelos informantes brasileiros

Palavras	N	Médias	DP	Teste t (sig bicaudal)
CAB	15	147,64	28,71	5,95
CAP		124,07	25,29	(p=0,000*)
PAD	15	161,02	32,85	3,36
PAT		141,86	18,02	(p=0,002*)
BAG	15	168,18	27,10	5,86
BACK		140,90	21,47	(p=0,000*)
BAD	15	150,44	23,70	1,97
BAT		139,79	27,10	(p=0,034*)
RAG	15	161,58	37,41	2,03
RACK		146,90	33,00	(p=0,032*)
BOB	15	170,49	30,23	3,19
BOP		152,36	20,06	(p=0,003*)

A Tabela 3 traz a diferença, sempre significativa, na comparação entre os tempos de vogais que antecedem consoantes surdas e sonoras nos três pares de oclusivas que compõem a amostra. Isso evidencia o fato de que a duração da vogal antecedente às plosivas já está sendo percebida como pista e utilizada na distinção entre oclusivas surdas e sonoras pelos aprendizes brasileiros.

Os dados acima ganham ainda mais força quando se

comparam a duração das vogais que precedem obstruintes vozeadas e desvozeadas entre os grupos de brasileiros e americanos, exibidos na tabela 4.

TABELA 4 - Comparação das durações dos segmentos vocálicos que antecedem as plosivas surdas e sonoras seguidas de vogal, produzidos pelos informantes brasileiros e americanos

Palavras	Grupo	N	Médias	DP	Teste t (sig bicaudal)
CAB	BR	15	207,34	36,10	-0,70
	USA	5	218,95	11,76	(p=0,496)
CAP	BR	15	172,92	37,65	1,43
	USA	5	146,56	27,07	(p=0,168)
PAD	BR	15	231,60	44,38	-0,90
	USA	5	251,10	32,28	(p=0,381)
PAT	BR	15	193,54	46,64	1,21
	USA	5	157,81	83,58	(p=0,240)
BAG	BR	15	234,78	43,99	-3,16
	USA	5	307,91	47,55	(p=0,005*)
BACK	BR	15	189,60	40,40	-0,87
	USA	5	206,83	29,12	(p=0,394)
BOB	BR	15	208,81	35,48	-2,94
	USA	5	258,52	20,83	(p=0,009*)
BOP	BR	15	163,98	31,64	-1,57
	USA	5	188,47	24,79	(p=0,134)
RAG	BR	15	210,77	35,14	-2,68
	USA	5	255,09	16,90	(p=0,015*)
RACK	BR	15	179,92	29,86	-0,03
	USA	5	180,27	13,53	(p=0,980)
BAD	BR	15	237,85	40,01	-2,71
	USA	5	296,29	47,09	(p=0,014*)
BAT	BR	15	191,86	36,47	-1,77
	USA	5	224,16	30,87	(p=0,093)

Como podemos observar na Tabela 4, a comparação entre os dois grupos em relação à duração de vogais que antecedem as plosivas indica dois fatos interessantes: 1) que não houve diferença significativa entre os dois grupos quanto à duração de vogais antecedentes a plosivas surdas; 2) que, dentre as seis palavras terminadas em plosivas sonoras, houve diferença significativa entre os dois grupos em três delas: ‘bag’, ‘bob’ e ‘bad’.

Resumindo, observamos que os participantes brasileiros produziram vogais mais longas antes de plosivas sonoras do que aquelas produzidas antes de plosivas surdas, conforme observado por ZIMMER & ALVES (2008). Entretanto, diferentemente do estudo de 2008, a duração das vogais antecedendo plosivas surdas não foi significativamente maior no grupo brasileiro comparado ao norte-americano. Além disso, observamos, neste estudo, que o grupo brasileiro produziu durações vocálicas significativamente maiores em quatro das seis palavras terminadas em oclusivas sonoras. Concluímos, assim, que os aprendizes brasileiros já fazem a distinção entre plosivas surdas e sonoras em posição final, no que tange à duração vocálica. Portanto, a duração da vogal é uma pista acústica em aquisição, embora em algumas plosivas sonoras ainda haja algumas diferenças de duração vocálica exagerada em relação aos norte-americanos.

4.3. Dessonorização terminal: considerações finais

Ao concluirmos a discussão acerca do processo de Desvozeamento Terminal na interlíngua português-inglês, verificamos, em função dos dados apresentados, que os aprendizes se encontram em vias de aquisição das duas pistas acústicas pesquisadas: o vozeamento na closura e a duração da

vogal. Ainda que os índices encontrados nos dados não equivalham, ainda, ao falar nativo, reduzir as produções dos aprendizes a uma simples caracterização de “dessonorização”, ou neutralização plena de oposição, mostra-se como uma visão por demais simplista do processo de aquisição. De fato, a oposição funcional entre surdas e sonoras parece ser, efetivamente, mantida pelos aprendizes brasileiros de inglês.

Contudo, parece-nos que o principal fator responsável pelo sotaque na produção de fala em L2 está relacionado ao *timing* que as constelações gestuais podem assumir na segunda língua, alterando relações de faseamento entre gestos e, por conseguinte, produzindo uma estrutura prosódica diferente daquela que é característica da língua-alvo. Essa diferença pode-se dar em vários graus, podendo ser bastante diferente, quando os gestos ainda estão entrincheirados no *timing* da L1, ou levemente diferente, quando as relações de faseamento entre os gestos e acoplamento entre constelações gestuais já se aproximam da L2. Isso fica claro ao verificarmos a duração das closuras dos segmentos plosivos produzidos pelos aprendizes brasileiros de inglês. De fato, frente à dificuldade de produção de segmentos plosivos anteriores a pausas, os aprendizes tendem a alongar a closura de suas obstruintes, possivelmente como estratégia compensatória para a ausência do segmento vocálico que tende a ser produzido, após a consoante, em sua L1. Cabe ao aprendiz, portanto, adquirir este novo *timing*, para que a relação entre duração da closura e vozeamento diminua e, por conseguinte, o próprio sotaque estrangeiro venha a ser atenuado. Além disso, ainda que tais aprendizes já distingam, através de suas produções, vogais que antecedem plosivas surdas e sonoras, tal distinção pode ser, também, mais realçada, o que também implica um aperfeiçoamento do *timing* da L2. Ainda que o aprendiz não tenha atingido o *timing* do falar nativo, mostra-se

indiscutível, entretanto, o fato de que a distinção entre os pares mínimos encerrados por segmentos obstruintes surdos e sonoros é parcialmente implementada.

Conclusões

Ao discutirmos tanto a abordagem estática quanto a fonologia dinâmica ao longo deste capítulo, defendemos a relevância de considerar o dado de aquisição fonológica de L2 a partir de uma perspectiva dinâmica. Assim, ao concluirmos este trabalho, queremos deixar claro que tal concepção não somente se mostra mais pertinente para o entendimento do sistema fônico da língua do aprendiz, mas, também, reflete toda uma concepção de linguagem que vai além do âmbito do componente fonético-fonológico.

Os dados de Desvozeamento Terminal, apresentados ao longo deste trabalho, evidenciam a pertinência da visão dinâmica aqui defendida. Verificamos, ao longo do texto, que a trajetória de aprendizagem do aprendiz, no que diz respeito à aquisição de tal aspecto, deve ser considerada sob uma perspectiva que vá além de uma descrição binária dos dados (desvozeamento ‘sim’ ou ‘não’). Ao lidarmos com uma visão dinâmica de aquisição, por outro lado, conseguimos expressar a gradiência presente no sistema linguístico em desenvolvimento, gradiência essa que se mostra como uma das principais características de qualquer forma de aquisição do conhecimento. Assim, ao considerarmos a aquisição fonológica de L2 por uma ótica dinâmica, conseguimos expressar, com maior riqueza, o quão perto ou longe a língua do aprendiz se encontra do sistema-fonte e do sistema-alvo.

Finalmente, mais do que discutir a trajetória da aquisição fonológica do aprendiz de L2, é necessário expressar o caráter

epistemológico da proposta. Nesse sentido, é necessário salientar que o ato de adotar uma perspectiva estática ou dinâmica deve implicar muito mais do que uma escolha ingênua do modo como os dados serão descritos e, posteriormente, do modelo utilizado para a análise. Mais do que isso, é preciso ter em mente que cada uma das perspectivas em questão revela uma diferente concepção, por parte do pesquisador, acerca do como é formado o conhecimento linguístico, visto que cada um desses posicionamentos remete a uma concepção de língua diferenciada. Assim, ao termos defendido, ao longo deste artigo, uma concepção dinâmica, esperamos ter apresentado uma concepção de aprendizagem de L2 que não se mostra isolada dos outros domínios do conhecimento, ao considerarmos um sistema sempre sujeito a alterações em virtude do meio em que se encontra e dos indivíduos que dele fazem uso.

Referências

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

ARBIB, Michael A. **Action to Language via the Mirror Neuron System**. Cambridge: Cambridge University Press, 2006.

BEST, Catherine. **Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development**. In: BOYSSON-BARDIES, B. (ed.). *Developmental neurocognition: speech and face processing in the first year of life*. Dordrecht: Kluwer, 1993, p. 289-304.

_____. **The emergence of native-like phonological influence in infants:** A perceptual assimilation model. In: NUSSBAUM, H.; GOODMAN, J.; HOWARD, C. (ed.). *The transition from speech to spoken words: The development of speech perception*. Cambridge, MA: MIT Press, 1994, p. 167-224.

_____. **A direct realist view of cross-language speech perception.** In: STRANGE, W. (ed.). *Speech perception and linguistic experience: issues in cross-language research*. Timonium, MD: York Press, 1995, p. 171-204.

BEST, Catherine; McROBERTS, Gerald W.; GOODELL, Elizabeth. **Discrimination of non-native contrasts varying in perceptual assimilation to the listener's native phonological system.** *Journal of the Acoustical Society of America*, v. 109, n. 2, 2001, p. 775-794.

BEST, Catherine & TYLER, Michael D. **Nonnative and second-language speech perception: commonalities and complementarities.** In: BOHN, Ocke-Schwen; MUNRO, Murray J. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins, 2007, p. 13-34.

BROWMAN, Catherine P. & GOLDSTEIN, Louis. **Towards an articulatory phonology.** *Phonology Yearbook*, 3, 1986. 219-252.

_____. **Articulatory phonology:** an overview. *Phonetica*, 49, p. 155-180.

BYRD, Dany. **Relating prosody and dynamic events:** comments on the papers by Cho, Navas, and Smiljanic. *Papers*

in *Laboratory Phonology*, n. 8, 2006.

CHOMSKY, Noam & HALLE, Morris. **The Sound Pattern of English**. New York: Harper & How, 1968.

CORBALLIS, Michael C. **From hand to mouth: The gestural origins of language**. In: Christiansen, Morten; KIRBY, Simon (Eds). *Language Evolution: The States of the Art*. Oxford: Oxford University Press., 2003, p. 201-218.

_____. **Mirror neurons and the evolution of language**. *Brain & Language*, n. 112, v. 1, , 2010. p. 25-35.

ECKMAN, Fred R. **Markedness and the contrastive analysis hypothesis**. *Language Learning*, n. 27, 1977. p. 315-330.

_____. **The Structural Conformity Hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners**. *Studies in Second Language Acquisition*, v. 13, 1996. p. 23-41.

_____. **A functional-typological approach to second language acquisition theory**. In: RITCHIE, W. C.; BHATIA, T. K. (ed.). *Handbook of second language acquisition*. San Diego, CA: Academic Press, 1996, p. 195-211.

_____. **From phonemic differences to constraint rankings: Research on Second Language Phonology**. *Studies in Second Language Acquisition*, v. 26, 2004. p. 513-549.

ELMAN, Jeffrey. **Connectionism, Artificial Life, and Dynamical Systems: New approaches to old questions**. In: BECHTEL, W.; GRAHAM, G. (Eds.). *A Companion to Cognitive Science*. Oxford: Basil Blackwood, 1998.

FLEGE, James E. **Second Language Speech Learning: Theory, findings, and problems.** In: STRANGE, W. (ed.). *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press, 1995, p. 233-277.

_____. **Interactions between the native and second-language phonetic systems.** In: Burmeister, P., Pirske, T. e Rhode, A. (Ed.). *An integrated view of language development: papers in honor of Henning Wode.* Trier: Wissenschaftlicher Verlag, 2002, p. 217-243.

_____. **Assessing constraints on second-language segmental production and perception.** In: Meyer, A. e Schiller, N. *Phonetics and phonology in language comprehension and production: differences and similarities.* Berlin: Mouton, 2003.

FLEGE, James E.; MCCUTCHEON, Martin & SMITH, Steven. **The development of skills in producing word-final stops.** *Journal of the Acoustics Society of America*, v. 82, n. 2, 1987.

FOWLER, Carol A. **Coarticulation and theories of extrinsic timing control.** *Journal of Phonetics*, n. 8, 1980. 113-133.

_____. **Speech as a supramodal or amodal phenomenon.** In: CALVER, Gemma A.; SPENCE; STEIN, Barry E. (Eds). *The Handbook of Multisensory Processes.* MIT Press, 2004, p. 189-201.

FOWLER, Carol A. & GALANTUCCI, B. **The relation of Speech Perception and Speech production.** In: PISONI, David B.; REMEZ, Robert E. *The Handbook of Speech Perception.* Blackwell Publishers, 2005, p. 633-652.

FRISCH, Stefan & WRIGHT, Richard. **The phonetics of phonological speech errors:** an acoustic analysis of slips of the tongue. *Journal of Phonetics*, n. 30, 2002. p.139–162.

GENTILUCCI, Maurizio & CORBALLIS, Michael. **From manual gesture to speech:** a gradual transition. *Neuroscience and Biobehavioral Reviews*, n. 30, 2006. p. 949–960.

GOLDSTEIN, Louis & FOWLER, Carol. **Articulatory phonology:** A phonology for public language use. In: SCHILLER, N.O.; MEYER, A.S. (Eds.). *Phonetics and Phonology in Language Comprehension and Production*. Mouton de Gruyter, 2003, p. 159-207.

GOLDSTEIN, Louis; BYRD, Dany & SALTZMAN, Elliot. **The role of vocal tract gestural action units.** In: ARBIB, Michael (Ed). *Action to Language via the Mirror Neuron System*. Cambridge: Cambridge University Press, 2006, p. 215-249.

GOLDSTEIN, Louis; POUPLIER, Marianne; CHEN, Larissa, SALTZMAN, E. & BYRD, Dani. **Dynamic action units slip in speech production errors.** *Cognition*, n. 103, 2007. , p. 386-412.

HASPELMATH, Martin. **Against markedness (and what to replace it with).** *Journal of Linguistics*, v. 42, n. 1, 2006. p. 25-70.

HOMMEL, Bernhardt & MILLIKEN, B. **Special issue on integration in and across perception and action.** *Psychological Research*, n. 71, v. 1, 2007.

JAKOBSON, Roman. **Child language, aphasia, and**

phonological universals. The Hague: Mouton, 1941.

KELLO, Christian T. **Patterns of timing in the acquisition, perception, and production of speech.** *Journal of Phonetics*, 31, 2003. pp 619–626.

KINGSTON, John & DIEHL, Randy. **Phonetic knowledge.** *Language*, v. 70, n. 3, 1994. p. 419-454.

_____. **Introduction to Phonetic Theory.** Manuscrito, 2007.

LADEFOGED, Peter. **A course in Phonetics.** Harcourt Brace, 3a ed., 1993.

MAJOR, Roy C. **A Model for Interlanguage Phonology.** In: IOUP, Georgette; WEINBERGER, Steven H. *Interlanguage Phonology: The acquisition of a Second Language Sound System.* Cambridge: Newbury House Publishers, 1987, p. 101-123.

MCNEILL, David. **Hand and mind: what gestures reveal about thought.** Chicago: University of Chicago Press, 1992.

MOWREY, Richard A. & MACKAY, Ian R. **Phonological primitives: electromyographic speech error evidence.** *Journal of the Acoustical Society of America*, n. 88, v.3, 1990. p.1299–1312.

PELUCCHI, Bruna; HAY, Jessica F. & SAFFRAN, Jenny. **Statistical learning in a natural language by 8-month old infants.** *Child Development*, n. 80, 2009. p. 674-685.

PLAUT, David & KELLO, Christian. **The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach.** In: MACWHINNEY, Brian (Ed.). *The emergence of language.* Mahwah,

NJ: Erlbaum, 1999, p. 381-415.

PORT, Robert. **The dynamics of language.** In: MEYERS, Robert A. (Eds.). *The Encyclopedia of Complexity and Systems Science*, 2009, p. 2310-2323.

RIZZOLATTI, Giacomo & CRAIGHERO, L. **The mirror-neuron system.** *Annual Review of Neuroscience*, n. 2, 2004. p. 169-92.

RODRIGUES, Luciana Lessa; FREITAS, Maria Cláudia; ALBANO, Eleonora Cavalcante & BERTI, Larissa Cristina. **Acertos gradientes nos chamados erros de pronúncia.** *Revista Letras*, n. 36, 2008. p. 85-112.

SAFFRAN, Jenny. **Statistical language learning:** Mechanisms and constraints. *Current Directions in Psychological Science*, n. 12, 2003. p. 110-114.

SANCIER, Michele L. & FOWLER, Carol A. **Gestural drift in a bilingual speaker of Brazilian Portuguese and English.** *Journal of Phonetics*, n. 25, 1997. p. 421-436.

SMITH, Caroline L. **The devoicing of /z/ in American English: effects of local and prosodic context.** *Journal of Phonetics*, 25, 1997. p. 471-500.

TRUBETZKOY, Nikolai. **Principles of Phonology.** Paris: Klincksieck, 1941.

TSUKADA, Kimiko; BIRDSONG, David; BIALYSTOK, Ellen; MACK, Molly; SUNG, Hyekyung & FLEGE, James. **A developmental study of English vowel production and perception by native Korean adults and children.** *Journal of*

Phonetics, v. 33, 2005. p. 263-290.

VAN GELDER, Timothy & PORT, Robert. **It's about time.** An overview of the dynamical approach to cognition. In: PORT, Robert; Van GELDER, Timothy (Eds). *Mind as motion*. Cambridge: M.A.: MIT, 1995, p 1-43.

ZIMMER, Márcia Cristina & ALVES, Ubiratã Kickhöfel. **A desonorização terminal na aprendizagem da L2:** evidências do continuum fonética-fonologia. *Letras de Hoje*, v. 32, n.3, 2007. p. 56-68.

_____. **On the status of Terminal Devoicing as an interlanguage process among Brazilian Learners of English.** *Ilha do Desterro*, n. 55, 2008. p. 41-62.

ZIMMER, Márcia Cristina.; SILVEIRA, Rosane & ALVES, Ubiratã Kickhöfel. **Pronunciation Instruction for Brazilians:** Bringing Theory and Practice Together. Newcastle: Cambridge Scholars Publishing, 2009.