

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

REINER VINICIUS PEROZZO

**Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa:
para além do *PAM-L2***

Porto Alegre

2017

REINER VINICIUS PEROZZO

**Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa:
para além do *PAM-L2***

Tese de Doutorado em Letras, vinculada à área Estudos da Linguagem e à linha de pesquisa Linguística Aplicada, apresentada como requisito parcial para a obtenção do título de Doutor pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

Porto Alegre

2017

CIP - Catalogação na Publicação

Perozzo, Reiner Vinicius

Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa: para além do PAM-L2 / Reiner Vinicius Perozzo. -- 2017.

225 f.

Orientador: Ubiratã Kickhöfel Alves.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Instituto de Letras, Programa de Pós-Graduação em Letras, Porto Alegre, BR-RS, 2017.

1. Cognição. 2. Gesto acústico-articulatório. 3. Realismo indireto. 4. Percepção fônica não nativa. I. Alves, Ubiratã Kickhöfel, orient. II. Título.

A presente Tese de Doutorado em Letras, de autoria de **Reiner Vinicius Perozzo**, sob o título ***Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa: para além do PAM-L2***, foi apresentada como requisito parcial para a obtenção do título de Doutor pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul. O trabalho encontra-se vinculado à área Estudos da Linguagem e à linha de pesquisa Linguística Aplicada, tendo sido devidamente defendido e aprovado no dia 31 de março de 2017, pela banca examinadora abaixo elencada.

Prof. Dr. Ubiratã Kickhöfel Alves (orientador)

Universidade Federal do Rio Grande do Sul

Prof.^a Dr.^a Adelaide Hercília Pescatori Silva

Universidade Federal do Paraná

Prof.^a Dr.^a Maria Lúcia de Castro Gomes

Universidade Tecnológica Federal do Paraná

Prof.^a Dr.^a Maity Simone Guerreiro Siqueira

Universidade Federal do Rio Grande do Sul

AGRADECIMENTOS

Agradeço à Capes pelo auxílio financeiro a mim concedido na forma de bolsa de Doutorado, ao longo dos últimos quatro anos.

Agradeço à minha família pelo apoio pleno e constante, e por ter sido basilar em todo o processo.

Agradeço ao professor Ubiratã Kickhöfel Alves pelos seis últimos anos de convivência, por ser um orientador exemplar e zeloso, e por ter me dado a oportunidade de crescer academicamente. O percurso entre a nossa primeira reunião e a defesa desta tese não poderia ter sido mais prodigioso e feliz. Agradeço, também, pelos incontáveis momentos de discussão e pelos inúmeros encontros de orientação. Sobretudo, agradeço pela grande amizade e pelo incansável companheirismo.

Agradeço aos colegas que estiveram ao meu lado nesta jornada. Em especial, agradeço ao Jorge Silva, à Natália Brambatti Guzzo, à Athany Gutierrez, à Juliana Feiden, à Ana Paula Scholl, à Leticia Pereyron e ao Felipe Kupske. Agradeço, também, às colegas mais recentes, Jeniffer Albuquerque, Gabi Gonçalves e Bruna de Los Santos, pelas conversas sempre muito frutíferas.

Agradeço à professora Ingrid Finger pelo carinho e pelas valiosas contribuições dadas a este trabalho.

Agradeço à professora Ana Beatriz Fontes pelos virtuosos momentos dentro e fora da universidade.

Agradeço à professora Adelaide Silva por ter aceitado ler e avaliar este trabalho, pela fundamental participação na construção deste texto, pelas contribuições substanciais à tese e por ser sempre tão cordial e prestativa.

Agradeço às professoras Maity Siqueira e Maria Lúcia de Castro Gomes por terem aceitado ler e avaliar este trabalho, e também pelas suas importantes contribuições.

Agradeço aos professores Elisa Battisti, Gisela Collischonn (*in memoriam*), Luiz Carlos Schwindt, Claudia Brescancini e Leda Bisol, dos quais tive o privilégio de ser aluno.

Agradeço, também, àqueles que não foram explicitamente mencionados, mas que, de uma forma ou outra, contribuíram para esta tese.

RESUMO

A presente tese se ocupa da percepção fônica de línguas não nativas e tem como objetivo repensar as premissas básicas do *PERCEPTUAL ASSIMILATION MODEL-L2* [*PAM-L2* (BEST; TYLER, 2007)] no que diz respeito aos seus eixos cognitivo, fônico e filosófico. De acordo com os proponentes do modelo, (i) a percepção da fala não nativa dispensa mecanismos cognitivos no que se refere a representações mentais ou processos inferenciais; (ii) a unidade de análise do evento perceptual, em termos de fala, é o gesto articulatório; e (iii) o acesso às informações disponíveis no mundo é direto, garantido pela atuação dos sentidos como nossos próprios sistemas perceptuais. Julgamos que tais premissas são limitadas e incoerentes com o objeto de investigação dos autores e argumentamos, portanto, que os eixos cognitivo, fônico e filosófico do modelo devam ser vislumbrados sobre um ponto de vista alternativo. Quanto ao primeiro eixo, defendemos que o evento perceptual seja concebido essencialmente como um fenômeno cognitivo, criado e gerenciado pelo encéfalo, que envolve abstrações, representações mentais e inferências acerca dos objetos do mundo. Em relação ao segundo eixo, julgamos que o tratamento acústico-articulatório (ALBANO, 2001) à unidade gestual seja mais apropriado à percepção de elementos fônicos não nativos, diferindo do tratamento articulatório (BROWMAN; GOLDSTEIN, 1989, 1992) que reside originalmente no modelo. No que diz respeito ao terceiro eixo, adotamos a posição realista indireta (JACKSON, 1977, 2010) como sendo aquela que abarca de modo mais adequado a percepção das unidades fônicas não nativas, em detrimento do realismo direto (J. GIBSON, 1966, 1986). Decorrentes de nosso refinamento teórico, questões adicionais à percepção fônica não nativa são endereçadas, as quais concernem à falsa dicotomia *L2* vs. *L1*, à influência do sistema grafêmico sobre a percepção fônica e à formação de novas categorias fônicas a serviço do idioma-alvo. Discutimos, também, os aspectos metodológicos de alguns estudos perceptuais, assim como suas implicações para uma nova caracterização do modelo, e ponderamos sobre o delineamento, o tipo de conhecimento de base testado (língua materna ou não nativa) e o objetivo de cada tarefa perceptual a ser empregada em ambiente laboratorial.

Palavras-chave: cognição; gesto acústico-articulatório; percepção fônica não nativa; realismo indireto.

ABSTRACT

This Doctoral Dissertation addresses the phonic perception of non-native languages and aims to rethink the central tenets of the PERCEPTUAL ASSIMILATION MODEL-L2 [PAM-L2 (BEST; TYLER, 2007)] with respect to its cognitive, phonic and philosophical spheres. According to the proponents of the model, (i) the perception of non-native speech disregards any cognitive mechanisms related to mental representations or inferential processes; (ii) the unit for analyzing perceptual events, in terms of speech, corresponds to the articulatory gesture; and (iii) we have direct access to the information available in the world, since our senses act as our own perceptual systems. We argue that these premises are limited and inconsistent with the research object of the authors and, thus, we assume that the cognitive, phonic and philosophical spheres of the model should be glimpsed from an alternative point of view. As for the first sphere, we argue that the perceptual event is essentially a cognitive phenomenon, created and managed by the brain, which involves abstractions, mental representations and inferences about the objects of the world. Regarding the second sphere, we state that an acoustic-articulatory treatment (ALBANO, 2001) to the gestural unit is more appropriate to the perception of non-native phonic elements, differing from the articulatory treatment (BROWMAN; GOLDSTEIN, 1989, 1992) that is originally conveyed by the model. With respect to the third sphere, opposing direct realism (J. GIBSON, 1966, 1986), we adopt the indirect realist position (JACKSON, 1977, 2010) as the one that encompasses the perception of non-native phonic units more adequately. Due to our theoretical refinement, we address additional issues to the phonic perception, which concern the false L₂ vs. L_E dichotomy, the impact of graphical systems on phonic perception, and the formation of new phonic categories at the service of the target language. We discuss methodological aspects of perceptual studies, as well as their implications for a new characterization of the model, and we also consider the design, the type of knowledge tested (native or nonnative language) and the purpose of each perceptual task to be used in a laboratory environment.

Keywords: *acoustic-articulatory gesture; cognition; indirect realism; nonnative phonic perception.*

LISTA DE FIGURAS

Figura 1	Variáveis do trato vocal e articuladores segundo a fonologia articulatória	40
Figura 2	Pauta gestual da palavra inglesa “SPAN”	41
Figura 3	Possíveis posições gestuais em um conglomerado	135
Figura 4	Pautas gestuais dos plurais -ãos, -ães e -ões segundo distinções morfológicas do singular em dialetos chiantes	136
Figura 5	Representação esquemática do evento perceptual fônico segundo nossa proposta nesta tese	164

LISTA DE QUADROS

Quadro 1	Previsões de assimilação segundo o <i>PAM</i>	46
Quadro 2	Relação de trabalhos acerca da percepção da fala não nativa	106
Quadro 3	Resumo das tarefas perceptuais e suas especificidades	185
Quadro 4	Periódicos consultados para a presente tese	224

LISTA DE SÍMBOLOS, SIGLAS E ABREVIATURAS

ABX	Elemento A, elemento B e elemento X (formato de uma tarefa de discriminação)
ADC	Assimilação a duas categorias
AL₂	Aquisição de segunda língua
ALE	Aquisição de língua estrangeira
ANC	Ambos não categorizáveis
ASP	<i>AUTOMATIC SELECTIVE PERCEPTION</i>
AUC	Assimilação a uma categoria
AX	Elemento A, elemento X (formato de uma tarefa de discriminação)
AXB	Elemento A, elemento X e elemento B (formato de uma tarefa de discriminação)
Capes	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDs	<i>COMPACT DISCS</i>
Crít.	Crítico
CV	Sequência de consoante mais vogal
CVC	Sequência de consoante mais vogal mais consoante
DAC	Diferença de adequação à categoria
Est.	Estreito
EUA	Estados Unidos da América
F₁	Primeiro formante ou primeira faixa de formantes
F₂	Segundo formante ou segunda faixa de formantes
HDs	<i>HARD DISCS</i>
IPA	<i>INTERNATIONAL PHONETIC ALPHABET</i>
L₁	Língua materna ou língua nativa
L₂	Segunda língua
L₂LP	<i>SECOND LANGUAGE PERCEPTION MODEL</i>
LE	Língua estrangeira
LNN	Língua não nativa
NCC	Não categorizado vs. categorizado
PAM	<i>PERCEPTUAL ASSIMILATION MODEL</i>
PAM-L₂	<i>PERCEPTUAL ASSIMILATION MODEL OF SECOND LANGUAGE SPEECH LEARNING</i>
Pl.	Plural
PUCRS	Pontifícia Universidade Católica do Rio Grande do Sul
SLM	<i>SPEECH LEARNING MODEL</i>
SSDs	<i>SOLID-STATE DRIVES</i>
UCPel	Universidade Católica de Pelotas
UECE	Universidade Estadual do Ceará
UFBA	Universidade Federal da Bahia
UFPB	Universidade Federal da Paraíba
UFPE	Universidade Federal de Pernambuco
UFPR	Universidade Federal do Paraná

UFRGS	Universidade Federal do Rio Grande do Sul
UFRJ	Universidade Federal do Rio de Janeiro
UFSC	Universidade Federal de Santa Catarina
UFSM	Universidade Federal de Santa Maria
UNB	Universidade de Brasília
UNESP	Universidade Estadual Paulista Júlio de Mesquita Filho
USP	Universidade de São Paulo
VOT	<i>VOICE ONSET TIME</i>
XAB	Elemento X, elemento A e elemento B (formato de uma tarefa de discriminação)

SUMÁRIO

1 INTRODUÇÃO	14
2 PERCEPÇÃO DA FALA NÃO NATIVA	22
2.1 Percepção da fala não nativa: O <i>PERCEPTUAL ASSIMILATION MODEL</i>	22
2.1.1 A abordagem ecológica de James Gibson e a teoria de percepção da fala de Carol Fowler	23
2.1.2 O gesto articulatório enquanto elemento perceptual da fala não nativa	37
2.1.3 Percepção da fala não nativa como condição assimilatória	43
2.2 Assimilação perceptual em uma segunda língua: O <i>PAM-L2</i>	48
2.3 Resumo do capítulo	51
3 ESTUDOS EM PERCEPÇÃO DE SONS DE LÍNGUAS NÃO NATIVAS	53
3.1 Artigos científicos com enfoque nas previsões assimilatórias do <i>PAM-L2</i>	57
3.1.1 Gilichinskaya e Strange (2010)	57
3.1.2 Bundgaard-Nielsen, Best e Tyler (2011b)	58
3.1.3 Escudero e Vasiliev (2011)	60
3.1.4 Escudero e Williams (2011)	61
3.1.5 Strange, Hisagi, Akahane-Yamada e Kubo (2011)	62
3.1.6 Antoniou, Best e Tyler (2012)	64
3.1.7 Bohn e Best (2012)	65
3.1.8 Darcy e Krüger (2012)	67
3.1.9 Antoniou, Best e Tyler (2013)	69
3.1.10 McKelvie-Sebileau e Davis (2014)	70
3.1.11 Tyler, Best, Faber e Levitt (2014)	71
3.2 Artigos científicos sem enfoque nas previsões assimilatórias do <i>PAM-L2</i>	73
3.2.1 Best e Hallé (2010)	73
3.2.2 Escudero e Chládková (2010)	75
3.2.3 Bundgaard-Nielsen, Best e Tyler (2011a)	77
3.2.4 Escudero, Simon e Mitterer (2012)	79
3.2.5 Escudero e Williams (2012)	80
3.2.6 Fabra e Romero (2012)	82
3.2.7 Shafiro, Levy, Khamis-Dakwar e Kharkhurin (2012)	83
3.2.8 Patihis, Oh e Mogilner (2013)	85
3.2.9 Dufour, Kriegel, Alleesab e Nguyen (2014)	87
3.2.10 Elvin, Escudero e Vasiliev (2014)	89
3.2.11 Escudero, Sisinni e Grimaldi (2014)	90
3.2.12 Stölten, Abrahamsson e Hyltenstam (2014)	92
3.2.13 Williams e Escudero (2014)	94
3.2.14 Yang e Fox (2014)	95
3.2.15 Alves e Zimmer (2015)	97

3.2.16 Holliday (2016)	99
3.3 Trabalhos brasileiros de pós-graduação com enfoque nas previsões assimilatórias do <i>PAM-L2</i>	101
3.3.1 Albuquerque (2012)	102
3.3.2 Perozzo (2013)	104
3.4 Considerações acerca dos estudos apresentados	105
3.5 Resumo do capítulo	115
4 PARA ALÉM DO PAM-L2: UMA NOVA FUNDAMENTAÇÃO TEÓRICA	117
4.1 A percepção como evento cognitivo e seu embasamento nas neurociências	118
4.2 A unidade gestual da percepção fônica	128
4.3 Realismo perceptual e seu alcance representacional	144
4.4 Questões pertinentes ao modelamento da percepção fônica de línguas não nativas	153
4.4.1 Segunda língua vs. língua estrangeira: uma fronteira inexistente	153
4.4.2 Influência do sistema grafêmico na percepção fônica de línguas não nativas	156
4.4.3 Formação de novas categorias: o que segue a partir das assimilações	159
4.5 Um modelo alternativo para a percepção fônica de línguas não nativas?	162
4.6 Resumo do capítulo	165
5 PERCEPÇÃO FÔNICA E PROCEDIMENTOS EMPÍRICOS DE TESTAGEM	167
5.1 O realismo direto e a arquitetura empírica da pesquisa no âmbito da percepção de sons linguísticos	168
5.2 Discussão metodológica dos estudos perceptuais e suas implicações para uma nova caracterização do modelo	173
5.2.1 Tarefa de assimilação perceptual	175
5.2.2 Tarefa de identificação	177
5.2.3 Tarefa de discriminação	179
5.2.4 Ponderações teórico-metodológicas em relação a estudos retratados no capítulo	186
5.3 Implementação das tarefas perceptuais fônicas não nativas à luz das considerações desta tese	189
5.3.1 Premissas metodológicas essenciais quanto à percepção fônica não nativa com base em nossa proposta	189
5.3.2 Operacionalizando aspectos particulares da percepção fônica não nativa	191
5.3.3 Tarefas laboratoriais indispensáveis à nossa proposta de percepção fônica não nativa	194
5.4 Resumo do capítulo	200
6 CONSIDERAÇÕES FINAIS	203
REFERÊNCIAS	207
APÊNDICES	224

1 INTRODUÇÃO

A percepção dos sons da fala, especialmente no que diz respeito àqueles que veiculam contrastes fonológicos, é um tema ligeiramente recente em termos de investigação científica, conforme assinala Nishida (2012). Seu início remonta aos anos de 1950 nos Laboratórios Haskins, nos Estados Unidos, quando pesquisadores como Pierre Delattre, Alvin Liberman e Franklin Cooper se dedicavam à experimentação acústica de sequências sonoras (DELATTRE; LIBERMAN; COOPER, 1955). Os primeiros testes perceptuais contavam com unidades auditivas que constituíam sílabas cv¹ do inglês e, através de uma maquinaria física [ferramentas que convertiam espectrogramas (desenhados a mão) em sinal acústico], os estudiosos alteravam as configurações formânticas de tais sílabas, com o intuito de promover distinções acústicas cujos resultados desencadeariam contrastes referentes a ponto de articulação para fones oclusivos orais.

Podemos dizer que, em um primeiro momento dos estudos em percepção da fala, os pesquisadores da área estavam voltados a explicar tal processo com base em informações de natureza acústica. Destacam-se, nessa etapa, os trabalhos de Pisoni (1973, 1974), Cole e Scott (1974) e Kuhl e Miller (1975), amplamente influenciados pelo paradigma estruturalista, através do qual se concebia a percepção da fala como sendo, basicamente, uma questão de audição, isto é, perceber a fala significava OUVIR a fala. O tratamento da percepção da fala no que compete a essa primeira escola se conceitua nos moldes psicoacústicos, em que os primitivos perceptuais são pistas acústicas apreendidas indiretamente através de processamento cognitivo e representações mentais. Além disso, de acordo com essa visão, as informações que os infantes² inicialmente percebem são gerais (não somente linguísticas) e a experiência com a língua desencadeia a formação de protótipos, propriedades e modelos (BEST, 1995).

Na medida em que as pesquisas sobre percepção da fala evoluíam, aumentavam, proporcionalmente, segundo indica Nishida (2012), as indagações acerca de quais eram os substratos das informações sonoras que promoviam as oposições lexicais, despertando um conflito entre a sua base acústica – vigente até então – e a sua base articulatória, a qual começava a se impor, principalmente, através da TEORIA MOTORA DA PERCEPÇÃO DA FALA [*MOTOR THEORY OF SPEECH PERCEPTION*

¹ Sílabas formadas pela sequência CONSOANTE-VOGAL.

² Termo referente às crianças que ainda não falam, ou seja, que se encontram em etapa anterior à aquisição da linguagem.

(LIBERMAN; MATTINGLY, 1985)]. Essa teoria assume a articulação como responsável pela percepção da fala, e essa sendo veiculada na cadeia sonora, deve ser mediada por uma estrutura linguística para ser percebida (e produzida). Segundo discorre Best (1995), a perspectiva motora para a percepção da fala prevê como primitivo de análise o gesto articulatório³ *PRETENDIDO*, derivado de comandos neuromotores, e estipula, também, o acesso indireto às representações mentais para tais gestos. A teoria postula que as informações que os infantes inicialmente percebem são, de fato, linguísticas e, no que concerne à experiência com a língua, o *INPUT* fonético nativo sintoniza o módulo da fala.

Outra escola que opera com o gesto articulatório (concebido sob outra natureza) e que configura um terceiro momento nos estudos em percepção da fala é a *TEORIA REALISTA DIRETA DA PERCEPÇÃO DA FALA* [*DIRECT REALIST THEORY OF SPEECH PERCEPTION* (FOWLER, 1986, 1996)]. Contudo, sublinhamos que o gesto articulatório na visão realista direta de Fowler (1986, 1996) é encarado como um objeto *REAL* (e não um gesto articulatório pretendido, que alude a uma representação mental). Quanto à proposta desenvolvida pela estudiosa, enfatizamos que “os gestos do trato vocal possuem propriedades invariantes próprias e seriam eles mesmos (e não suas estruturas de controles neurais, tal como sugere a *TM*) os componentes fonológicos de uma sentença” (NISHIDA, 2012, p. 40, acréscimo do autor). Assim como Fowler (1986, 1996), Best (1995) entende que o construto realista direto da percepção da fala se apoia sobre os gestos articulatórios distais (e reais) como seu primitivo, apreendidos diretamente através da extração de informações também distais, com base em sistemas perceptuais integrados, e da atividade exploratória ativa dos eventos de fala. Derivada da perspectiva ecológica da percepção (GIBSON, 1966, 1986) e vinculada a uma concepção dinâmica de língua, a teoria de Fowler defende que as informações que os infantes primeiramente percebem são gerais⁴, estando, nesse sentido, em conformidade com o que preconizaria a vertente psicoacústica⁵; porém, diferentemente do concebido na escola anterior, a experiência com a língua prevê a extração de invariantes gestuais nativos, acessados diretamente, a qual forma o conhecimento dos padrões sonoros da língua em questão (BEST, 1995).

As discussões envolvendo as bases da percepção da fala, no entanto, parecem estar longe de se esgotarem: podemos observar argumentos e contra-argumentos em favor de uma ou outra teoria.

³ Unidade de análise fonológica circunscrita no âmbito da fonologia gestual.

⁴ O termo “GERAIS”, nesse contexto, implica que não se trata de informações específicas da fala.

⁵ Ressaltamos, porém, que, na linha psicoacústica, a natureza das informações perceptuais é de cunho ACÚSTICO-AUDITÓRIO, não de cunho GESTUAL.

Este é o caso, por exemplo, da oposição entre John Ohala e Carol Fowler, relatada por Nishida (2012). Em busca de um substrato para a percepção de sons linguísticos, de um lado, Ohala (1996) levanta evidências fonológicas baseadas em propriedades acústicas do sinal de fala, e, de outro, Fowler (1996)⁶ enumera constatações articulatórias a partir da percepção e da produção da fala. Todavia, não temos como julgar qual das posições é mais virtuosa em seu empreendimento teórico, uma vez que os autores em questão partem de primitivos distintos, segundo ressalta Nishida (2012).

Devemos apontar que, ao longo de quase meio século de investigação sobre a percepção da fala, os estudos na área se preocuparam em elucidar a origem e o processo do fenômeno perceptual, lançando luz sobre a necessidade de se modelar, formalmente, como os indivíduos, no desenvolvimento de sua língua materna (L_1), percebem os elementos sonoros capazes de revelar as propriedades que marcam fronteiras entre os sons da fala. Tal modelagem, impulsionada pela averiguação dos primitivos concernentes à percepção da fala em L_1 , se projetou, subsequentemente, a observações em línguas não nativas, implicando novos questionamentos e desafios.

Desde as duas últimas décadas, temos notado a predominância de três modelos perceptuais da fala não nativa⁷ no que se refere às pesquisas nacionais e internacionais: o MODELO DE APRENDIZAGEM DA FALA [*SPEECH LEARNING MODEL, SLM* (FLEGE, 1995)] e o MODELO DE ASSIMILAÇÃO PERCEPTUAL [*PERCEPTUAL ASSIMILATION MODEL, PAM* (BEST, 1995)], bem como sua extensão, voltada primordialmente para a percepção de elementos sonoros referentes à segunda língua (L_2), designada MODELO DE ASSIMILAÇÃO PERCEPTUAL DA APRENDIZAGEM DA FALA EM SEGUNDA LÍNGUA [*PERCEPTUAL ASSIMILATION MODEL OF SECOND LANGUAGE SPEECH LEARNING, PAM-L2* (BEST; TYLER, 2007)]⁸. De acordo com a explanação de Perozzo e Alves (2016), o SLM engloba tanto a percepção quanto a produção de sons não nativos, habilidades condicionadas pela experiência linguística do aprendiz, além de variáveis associadas à idade de aprendizagem do

⁶ Exemplificam-se, aqui, os estudos do ano de 1996, especificamente, por terem sido publicados em uma mesma edição de um periódico científico, representando, portanto, o confronto direto entre ambos os pesquisadores. No entanto, cabe lembrar que o embate teórico iniciou anos antes das referidas publicações, a exemplo da proposta de Fowler (1986). Para uma apreciação mais elaborada sobre tal embate, recomendamos a leitura de Nishida (2012).

⁷ Os modelos conhecidos como MODELO DE PERCEPÇÃO DE SEGUNDA LÍNGUA {[*SECOND LANGUAGE PERCEPTION MODEL (L2LP)*], proposto por Escudero e Boersma (2004) e Escudero (2005, 2009)} e PERCEPÇÃO SELETIVA AUTOMÁTICA {[*AUTOMATIC SELECTIVE PERCEPTION (ASP)*], desenvolvido por Strange (2010)}, ainda que tenham ganhado certa difusão nos estudos em percepção de sons linguísticos, se mostram bastante recentes. Por isso, não os incluímos como parte integrante do rol de modelos perceptuais que são utilizados desde a década de 90.

⁸ A partir deste ponto, quando fizermos menção a tais modelos perceptuais, iremos suprimir sua referência autoral para fins de delimitação. Portanto, 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.

sistema-alvo de sons e ao período de chegada à comunidade geográfica em que o idioma-alvo é falado. O *PAM*, por sua vez, abrange os denominados “*NAÏVE LISTENERS*”, ou seja, ouvintes sem experiência com um idioma não materno, e pressupõe que a L_1 influencia a percepção da fala não nativa, além de enumerar as possibilidades de assimilação perceptual de elementos sonoros que não fazem parte do sistema materno dos ouvintes (PEROZZO; ALVES, 2016). A proposta do *PAM-L2* (que, por sua vez, é uma amplificação do *PAM*) incide sobre aprendizes de uma L_2 que estão adquirindo o sistema-alvo de sons, e seu princípio fundamental é o de que a aprendizagem perceptual é definida por princípios diferentes daqueles da L_1 .

Na presente tese, optamos por problematizar especificamente o *PAM* e o *PAM-L2*, de modo a reunir constatações que defendemos ser essenciais no tocante à percepção de sons linguísticos, mas que escapam a tais modelos. Embora um tratamento mais pleno sobre o *PAM-L2*, foco de nosso empreendimento teórico, seja feito no segundo capítulo desta tese, avaliamos ser primordial descrever, rapidamente, o escopo de tal modelo, a fim de que possamos prenunciar seus pontos mais fundamentais, os quais desencadeiam toda uma problematização que justifica nosso trabalho doutoral. Em linhas gerais, o *PAM-L2* procura mapear a relação entre elementos sonoros da L_2 frente aos da L_1 , e o faz com base em previsões acerca de como determinado contraste sonoro na L_2 , inexistente na sua L_1 , pode vir a ser percebido pelo aprendiz. Para cumprir esse propósito, o modelo parte de uma teoria filosófica com implicações psicológicas (o realismo direto perceptual), de uma unidade sonora de caráter distintivo (o gesto articulatório) e da noção de ambiente (a ecologia da percepção), além de direcionar sua proposta a aprendizes de uma L_2 – estabelecendo, de modo explícito, uma oposição diametral entre aprendizes de L_2 e aprendizes de L_1 ⁹. Assim, faz-se imprescindível esclarecer o porquê de, neste trabalho, colocarmos em evidência os modelos *PAM* e *PAM-L2* em detrimento do *SLM*, já que este também tem sido referenciado em estudos sobre percepção de sons não nativos. Três são as razões fundamentais para tratarmos do *PAM* e do *PAM-L2*, e não do *SLM*, na presente análise, descritas a seguir.

A primeira decorre do fato de que, embora o *PAM-L2* configure um modelo de PERCEPÇÃO DE SONS LINGUÍSTICOS, circunstância que fortemente o inclinaria a ser inserido nos construtos sobre aquisição de L_2 e bilinguismo sob uma vertente cognitiva, sua proposta ainda carece de uma abordagem que

⁹ A dicotomia L_2 - L_1 advém dos proponentes do *PAM-L2* e em nada representa a nossa visão sobre o desenvolvimento do conhecimento linguístico não nativo. Tal assunção será endereçada na seção 4.4.1 deste trabalho.

leve em conta ASPECTOS E PRESSUPOSTOS COGNITIVISTAS. O plano cognitivo não corresponde à essência de tal modelo propositalmente, pois seus proponentes seguem um referencial teórico que não atribui ao evento perceptual uma propriedade cognitiva. Como partimos do pressuposto de que qualquer experiência no âmbito da percepção seja necessariamente atrelada a questões cognitivas, é elementar que essa problematização seja considerada em nossas discussões neste trabalho. O *SLM*, diferentemente, evidencia algumas características encontradas no rol cognitivista (por exemplo, abstração, acuidade de produção, idade de aprendizagem, processos pré-atencionais, etc.).

A segunda se relaciona ao fato de que tanto o *PAM* como o *PAM-L2* estão ancorados ao primitivo fônico com o qual compactuamos, a saber, a UNIDADE GESTUAL. Contudo, o tratamento teórico fornecido a tal unidade advém de um substrato articulatório, delineado por Fowler (1986, 1996) e Browman e Goldstein (1989, 1992), e o tratamento teórico que julgamos ser mais pertinente à unidade gestual, conforme será proposto e defendido nesta tese, é aquele empreendido por Albano (2001), que tem caráter acústico-articulatório. O *SLM*, por outro lado, não parece possuir ligação explícita a uma determinada unidade de análise, mas presumimos que abarque pistas acústicas.

A terceira diz respeito à vinculação do *PAM* e do *PAM-L2* a uma TEORIA FILOSÓFICA, que, conseqüentemente, tem IMPLICAÇÕES PSICOLÓGICAS, expressando um compromisso metateórico de extrema relevância. A arena de discussões trazida nesta tese em termos de alocação filosófica diz respeito justamente à noção de termos acesso direto aos objetos reais disponíveis no mundo, conforme expresso em tais modelos, ou termos acesso indireto a eles. O *SLM*, por sua vez, não menciona se filiar a uma teoria filosófica e, tampouco, parece ter a preocupação de se sustentar filosoficamente (de maneira explícita).

Nossa argumentação nesta tese é a de que os entes da tríade teórica (aspectos cognitivos, primitivo fônico e postura filosófica) que se instauram no *PAM* e no *PAM-L2* parecem não estar em consonância entre si, seja porque seu apelo à cognição seja indevidamente minimizado, seja porque o tratamento de sua unidade sonora apresente limitações, acabando, também, por ser resgatada com dificuldades a partir do modelo, seja porque sua base filosófica não se aplica à linguagem, conforme propomos problematizar no quarto capítulo deste trabalho.

Ao considerarmos que há uma crescente utilização do *PAM-L2* nos trabalhos que investigam a percepção da fala em *L2/LE*, e que o modelo apresenta problemas em termos de seus eixos basilares¹⁰ (aspectos cognitivo, fônico e filosófico), a necessidade de discutir seus fundamentos se configura

¹⁰ A serem evidenciados no quarto capítulo desta tese.

como uma medida emergencial para os estudos em interfonologia, o que motiva nosso empreendimento teórico na tentativa de elucidar seus pontos passivos de uma discussão mais aprofundada. Sob esta perspectiva, afirmamos ser primordial trazer à tona, ao longo desta tese, alguns fatos intrigantes a respeito do modelo, os quais se alocam no núcleo de sua proposta e, necessariamente, decorrem da perspectiva conceitual de seu predecessor, o *PAM*.

Com base no que acabamos de expor, iremos nos debruçar sobre um questionamento (que corresponde à nossa grande questão de pesquisa) mais do que relevante para estudos interfonológicos que adotam o *PAM-L2* como fundamentação teórica – tanto para a condução da pesquisa em si, quanto para a análise e a interpretação dos resultados que se obtêm: a partir dos construtos que subjazem ao *PAM-L2*, QUAIS SÃO OS COMPONENTES TEÓRICOS NECESSÁRIOS PARA QUE, A PROPÓSITO DA PERCEPÇÃO FÔNICA NÃO NATIVA, O MODELO ATINJA MAIOR CONSONÂNCIA QUANTO ÀS SUAS BASES COGNITIVA, FÔNICA E FILOSÓFICA?

Claramente, nossa questão de pesquisa se relaciona intimamente ao objetivo geral desta tese, que é o de REPENSAR PREMISSAS DO *PAM-L2* NO QUE DIZ RESPEITO ÀS ESFERAS COGNITIVA, FÔNICA E FILOSÓFICA DO MODELO A PROPÓSITO DA PERCEPÇÃO FÔNICA DE UMA LÍNGUA NÃO NATIVA (LNN)¹¹, a fim de que possamos tratar o fenômeno perceptual referente às unidades fônicas com a devida consistência em termos de suas esferas cognitiva, fônica e filosófica. Para isso, enumeramos os seguintes objetivos específicos:

- a) verificar a exaustividade das previsões de assimilação veiculadas pelo *PAM-L2*;
- b) discutir e sugerir explicitamente a necessidade de recorrer a um aparato cognitivo no que concerne à percepção de aspectos fônicos de línguas não nativas;
- c) questionar o primitivo de análise do *PAM-L2*, sugerindo um tratamento acústico-articulatório a tal unidade;
- d) problematizar a alocação filosófica do *PAM-L2*, apontando uma perspectiva realista indireta ao evento perceptual;
- e) discorrer sobre questões pertinentes ao modelamento da percepção fônica em línguas não nativas, tais quais a dicotomia entre *L2* e língua estrangeira (*LE*), o impacto da grafia e a formação de novas categorias, de modo coerente à discussão lançada a partir dos objetivos específicos supracitados;

¹¹ POR LÍNGUA NÃO NATIVA NOS REFERIMOS A QUALQUER IDIOMA QUE NÃO O MATERNO, SEJA ELE DE CONHECIMENTO DO OUVINTE OU NÃO (OU SEJA, O OUVINTE PODE APRESENTAR ALTA OU BAIXA PROFICIÊNCIA EM TAL IDIOMA). BEST (1995), DIFERENTEMENTE, UTILIZA O MESMO TERMO SOMENTE PARA VEICULAR A IDEIA DE UMA LÍNGUA DESCONHECIDA AO OUVINTE.

- f) ponderar sobre quais experimentos perceptuais podem ser mais apropriados para verificar como a categorização sonora ocorre diante desta nova caracterização do modelo.

A fim de contemplar tais objetivos, julgamos indispensável percorrer os próximos cinco capítulos desta tese, apresentados, de forma breve, a seguir.

O capítulo 2 inicia com uma apresentação do *PAM* e levanta tópicos como sua fundamentação filosófica (e implicações psicológicas), sua postura linguística em termos de unidade sonora, e suas previsões quanto aos padrões de assimilação perceptual a que determinados sons da fala podem ser submetidos. Em seguida, passamos a uma apreciação do *PAM-L2*, que corresponde a uma adaptação do *PAM* ao cenário de percepção de sons de uma *L2* (o qual importa integralmente a bagagem teórica do *PAM*), em que mencionamos as alterações realizadas no modelo original para que tal propósito fosse cumprido.

A serviço do objetivo (a) está o capítulo 3, o qual se ocupa de uma relação entre diversos estudos, em grande parte internacionais, que se utilizaram do *PAM-L2* em diversos casos. Para cada trabalho apresentado, colocamos em evidência qual é o seu primitivo de análise, questionamos se o modelo consegue abarcar o componente teórico-metodológico empregado em tais estudos, verificamos se faz referência às previsões estabelecidas pelo modelo, e indicamos qual seria o padrão assimilatório do *PAM-L2* que poderia explicar os dados encontrados. Percorrer cada estudo nos permite analisar se aquelas previsões que o *PAM-L2* elenca se fazem exaustivas quanto às categorizações reportadas pelos diversos autores.

O capítulo 4 incorpora os objetivos (b), (c), (d), e (e), e se mostra, assim, como uma reflexão teórica acerca dos preceitos de base do *PAM-L2*. Desse modo, esse capítulo aponta a pertinência do expediente cognitivo para a percepção fônica, endereçando o objetivo (b); propõe o tratamento teórico aos moldes de Albano (2001), em detrimento do de Browman e Goldstein (1989, 1992), para a unidade gestual, cumprindo com o objetivo (c); problematiza a perspectiva filosófica do *PAM-L2*, sugerindo uma abordagem realista indireta ao evento perceptual, satisfazendo o objetivo (d); discorre sobre pontos relevantes no que diz respeito à percepção fônica não nativa, como os conceitos de *L2* e *LE*, influência da grafia e formação de novas categorias, contemplando o objetivo (e).

O capítulo 5, que vai ao encontro do objetivo (f), se concretiza a partir de uma discussão sobre os procedimentos empíricos de testagem, no sentido de averiguar quais são as medidas metodológicas mais eficazes para que a percepção fônica, de acordo com os preceitos discutidos no capítulo anterior, seja adequadamente aferida. Para tanto, os mesmos estudos reportados no capítulo

3 serão revisitados e, a partir deles, discorreremos sobre métodos experimentais que podem se mostrar efetivos para determinar o processo de assimilação perceptual de elementos sonoros não nativos. Frisamos que a realização de um experimento perceptual não se configura como sendo o intuito desse capítulo – diferentemente, ao final desse capítulo, procuramos apontar, com base em estudos já realizados, quais são as opções procedimentais mais apropriadas para a verificação das categorizações estipuladas para os elementos fônicos não nativos.

Como fechamento da presente tese, o sexto capítulo apresenta as conclusões do trabalho, bem como as suas limitações e seus direcionamentos futuros. Chamamos atenção para o fato de que a organização dos capítulos segundo a presente maneira pretende guiar o leitor em uma linha temporal de argumentação, em que o conteúdo de cada capítulo evolui na medida em que a tese se desenvolve e, ao seu final, reste claro o nosso empreendimento teórico. Desse modo, iniciaremos a tese em si com a apresentação dos modelos *PAM* e *PAM-L2*, relacionaremos os estudos de percepção de sons linguísticos que utilizam o *PAM-L2* (e, também, estudos que não o utilizam), discutiremos as incongruências do modelo acerca de seus eixos cognitivo, fônico e filosófico, ao mesmo tempo em que alvitaremos um construto alternativo, e, por último, com base nos estudos levantados, sugeriremos procedimentos empíricos de testagem acerca das categorias fônicas dos aprendizes.

2 PERCEPÇÃO DA FALA NÃO NATIVA

O trabalho de Catherine Best, na área da percepção da fala em seu nível translinguístico, é um dos mais influentes que conhecemos. Além de apresentar um embasamento teórico bastante robusto no tocante à filosofia, conforme expresso no próprio título de sua proposta de 1995, conhecida como *UMA VISÃO REALISTA DIRETA DA PERCEPÇÃO DA FALA TRANSLINGUÍSTICA (A DIRECT REALIST VIEW OF CROSS-LANGUAGE SPEECH PERCEPTION)*, a autora se utiliza de um primitivo fonológico específico e faz previsões acerca das possibilidades de discriminação dos sons linguísticos. Seu modelo de percepção da fala não nativa, o *MODELO DE ASSIMILAÇÃO PERCEPTUAL (PERCEPTUAL ASSIMILATION MODEL)*, referido por *PAM*, está inserido em seu estudo de 1995 e serve de apoio para o desenvolvimento de seu modelo sucessor, o *MODELO DE ASSIMILAÇÃO PERCEPTUAL PARA A APRENDIZAGEM DA FALA EM SEGUNDA LÍNGUA (PERCEPTUAL ASSIMILATION MODEL OF SECOND LANGUAGE SPEECH LEARNING)*, conhecido como *PAM-L2*, cunhado em colaboração com Michael Tyler em 2007.

O presente capítulo se configura como parte de nosso referencial teórico para as discussões que virão nos próximos capítulos. Julgamos pertinente resenhar o *PAM* em um primeiro momento, já que sua fundamentação é, praticamente, toda transposta à sua subsequente amplificação (o *PAM-L2*), servindo-lhe de base teórica. Desse modo, abordaremos, neste capítulo, o modelo *PAM*, passando por sua escola filosófica, seu primitivo de análise e suas previsões de assimilação¹², e, então, nos direcionaremos ao *PAM-L2*, de modo a tocar em aspectos relevantes de sua proposta, que o singularizam do modelo de 1995.

2.1 Percepção da fala não nativa: o *Perceptual Assimilation Model*

Em função de ser uma pesquisadora vinculada aos Laboratórios Haskins (Connecticut, EUA), Catherine Best incorpora a seu modelo perceptual da fala não nativa teorias advindas dos trabalhos de seus colegas, o que a faz estar em paralelismo teórico com os demais cientistas da fala alocados na instituição. Desse modo, em *UMA VISÃO REALISTA DIRETA DA PERCEPÇÃO DA FALA TRANSLINGUÍSTICA (BEST, 1995)*,

¹² O eixo referente à cognição não é mencionado nesse ordenamento, e julgamos importante explicar o porquê. No texto seminal, de 1995, supostos aspectos cognitivos aparecem coadunados à escola filosófica que atua como carro-chefe do modelo, e nunca são apresentados de maneira explícita. Mais adiante, na proposta de 2007, apesar da ausência de uma caracterização explícita da esfera cognitiva, seus efeitos podem ser sentidos a partir de algumas tomadas de decisão estipuladas pelos autores, tais como a referente à dicotomia *L2-LE*. Contudo, em virtude de defendermos a falta de consonância entre tal concepção cognitiva, apresentada implicitamente, e a teoria de base filosófica que a sustenta, instauramos o segundo objetivo [objetivo (b)] desta tese, a ser abordado no quarto capítulo.

insere-se o realismo direto perceptual a partir da TEORIA REALISTA DIRETA DA PERCEPÇÃO DA FALA, desenvolvida por Fowler (1986), e se anexa, também, a FONOLOGIA ARTICULATÓRIA, cunhada por Browman e Goldstein (1986, 1989).

Com o intuito de tratar dos eventos perceptuais da fala não nativa, Best (1995) segue, assim, duas direções ao longo de sua exposição no que compete ao PAM. A primeira tem a ver com a relação existente entre o realismo direto e a percepção da fala, em que a autora se debruça sobre a ecologia da percepção. A segunda diz respeito ao modelo realista direto da percepção da fala translinguística, voltada a ouvintes sem experiência com a língua-alvo, em que a estudiosa explicita o primitivo fonológico que adota e elenca predições quanto à discriminação sonora, as quais são potencialmente aplicáveis aos inventários fônicos das línguas do mundo. Por uma questão organizacional, versaremos sobre sua teoria filosófica (a qual tem implicações psicológicas) na seção 2.1.1, sobre sua teoria fonológica na seção 2.1.2 e sobre as predições que sustentam o modelo na seção 2.1.3, a seguir.

2.1.1 A abordagem ecológica de James Gibson e a teoria de percepção da fala de Carol Fowler

Referências constantes assinaladas por Best (1995) no que diz respeito à sua posição filosófica são as obras OS SENTIDOS CONSIDERADOS COMO SISTEMAS PERCEPTUAIS (*THE SENSES CONSIDERED AS PERCEPTUAL SYSTEMS*) e A ABORDAGEM ECOLÓGICA À PERCEPÇÃO VISUAL (*THE ECOLOGICAL APPROACH TO VISUAL PERCEPTION*), ambas de autoria de James Gibson, que datam de 1966 e 1979¹³, respectivamente. James Gibson foi um psicólogo experimentalista norte-americano que dedicou suas pesquisas majoritariamente à ótica, especialmente no que concerne à percepção visual¹⁴.

Muitas das observações que James Gibson faz em sua obra de 1986 já estão presentes no texto de 1966, o que significa que seus argumentos e suas constatações não só se mantiveram, mas também foram mais “polidos” e especificados ao sentido da visão. Entretanto, o que está mais explícito em OS SENTIDOS CONSIDERADOS COMO SISTEMAS PERCEPTUAIS (1966) é pontualmente o motivo do tratamento que o autor confere à percepção e as suas considerações sobre os sistemas de comunicação humana.

¹³ Quando nos referirmos à obra A ABORDAGEM ECOLÓGICA À PERCEPÇÃO VISUAL, estaremos utilizando o texto de 1986 em vez daquele de 1979, pois só tivemos acesso ao livro que corresponde à segunda metade da década de 80. No entanto, por se tratar apenas de uma edição subsequente, acreditamos que as informações veiculadas se mantenham igualmente integrais na publicação mais recente.

¹⁴ Discussões pertinentes acerca do construto gibsoniano aplicado ao PAM-L2 são realizadas em Alves e Silva (2016), leitura que recomendamos e de que trataremos no quarto capítulo desta tese.

J. Gibson (1966) inicia sua exposição afirmando que a tradição científica presume que os sentidos sejam canais sensoriais e que SENTIR algo pode corresponder a DETECTAR OU a TER UMA SENSAÇÃO. Para a sua teorização, o psicólogo entende que está se referindo ao primeiro significado; portanto, SENTIR se relaciona a DETECTAR. Paralelamente, o pesquisador chama de PERCEPÇÃO toda e qualquer concepção ou crença que a natureza produz através dos sentidos. Sua proposta perceptual tem como suporte três premissas relacionadas entre si. A PRIMEIRA delas é a de que os sentidos humanos podem obter informação sobre os objetos do mundo sem a intervenção de um processo intelectual uma vez que tais sentidos venham a OPERAR como sistemas perceptuais. A SEGUNDA postula que o cérebro apenas governa os órgãos perceptuais, mas não ORGANIZA o input sensorial e tampouco PROCESSA os dados ou as informações (p. 5) – fica clara a perspectiva mais fisiológica da percepção, em detrimento de uma perspectiva mais cognitiva para o evento perceptual. A TERCEIRA premissa tem a ver com o comportamento das espécies, que, orientado à exploração do meio em que vivem e interagem, só faz sentido se for analisado via ECOLOGIA¹⁵, e não através das leis da física.

Podemos verificar nessas três premissas um direcionamento do observador/percebedor às informações disponíveis em seu entorno, ou seja, há claramente uma noção de que a percepção presume uma AÇÃO no meio em que se insere o observador/percebedor. Assim, nossos sentidos passariam de elementos passivos a elementos ativos, atuando como sistemas PARA a percepção. Esses sistemas perceptuais seriam tão poderosos a ponto de dispensarem uma intervenção cognitiva (ou intelectual) no dado a ser percebido e estariam amparados por propriedades, intrínsecas ao indivíduo, de buscar ativamente as informações circundantes. Nas palavras do próprio J. Gibson (1966, p. 47), somos conduzidos a entender que:

Iremos conceber os sentidos externos de uma nova maneira, como ativos ao invés de passivos, como sistemas ao invés de canais, e como inter-relacionados ao invés de mutuamente exclusivos. Se eles funcionam para obter informação, não simplesmente para provocar sensações, tal função deve ser denotada por um termo diferente. Eles serão chamados, aqui, de sistemas perceptuais¹⁶ (GIBSON, J., 1966, p. 47).

¹⁵ Para J. Gibson (1966), o conceito de “ECOLOGIA” corresponde ao ambiente SIGNIFICATIVO, aquele que oferece condições para a manutenção da espécie, que exprime as inter-relações dos seres que habitam tal ambiente.

¹⁶ No original: “*We shall have to conceive the external senses in a new way, as active rather than passive, as systems rather than channels, and as interrelated rather than mutually exclusive. If they function to pick up information, not simply to arouse sensations, this function should be denoted by a different term. They will here be called perceptual systems*” (GIBSON, J., 1966, p. 47).

Depende da integração dessas premissas o fundamento que subjaz ao construto de James Gibson: para podermos sobreviver neste planeta e manter a nossa espécie em constante evolução, foi preciso que nossos sentidos atuassem como sistemas de percepção, os quais se orientaram apropriadamente à extração de informações no ambiente¹⁷ (GIBSON, J., 1966, p. 58). Assim, o motivo substancial do tratamento que James Gibson confere à percepção, com base em suas três premissas, reside na LIGAÇÃO DIRETA entre o ambiente e os sentidos enquanto sistemas perceptuais.

Apesar de defender a maneira direta como apreendemos as informações disponíveis no mundo, J. Gibson (1966, 1986) não se rotula como um realista direto (talvez porque isso não o preocupe). A título de síntese, frisamos que, para J. Gibson (1966), as informações visuais, auditivas, táteis, olfativas, gustativas, etc., apresentam estruturas invariantes que podem ser registradas pelos sistemas perceptuais sem a mediação de processos cognitivos ou inferências. Para além disso, esse cenário todo que abrange a descrição de James Gibson só se justifica porque o estudioso toma como ponto de partida a luta pela sobrevivência e a manutenção dos animais que habitam o planeta. Independentemente de a percepção direta ser aplicável, de fato, à realidade em termos de sobrevivência e manutenção da espécie, pelo menos no plano da linguagem, James Gibson assegura a impertinência de uma abordagem que presuma a percepção direta.

Faz-se interessante notar que, durante sua exposição sobre a linguagem, a argumentação de J. Gibson (1966) abrange o tema como um todo e, em seguida, vai se afunilando até chegar às unidades sonoras da comunicação. O estudioso entende que a linguagem corresponde a um conjunto de SÍMBOLOS que carregam os significados dos objetos que fazem parte do ambiente comum a todos os indivíduos. Tais símbolos dão conta de fazer com que os homens pensem sobre as mesmas coisas, tenham conceitos em comum e verifiquem conjuntamente esses conceitos.

J. Gibson (1966), ainda no tocante à linguagem, menciona que o significado SIMBÓLICO é diferente do significado PERCEPTUAL na medida em que a relação de um estímulo perceptual com sua fonte no ambiente é de um tipo, enquanto que a relação de um símbolo com o seu referente é de outro. Para o psicólogo, a primeira depende da ECOLOGIA DA ESTIMULAÇÃO e a segunda depende de uma COMUNIDADE LINGÜÍSTICA, que é uma invenção única da espécie humana. Ele acrescenta que, no primeiro caso, a relação entre as partes é intrínseca, como uma projeção, e, no segundo caso, a relação entre as partes é extrínseca, como um acordo social cujas convenções devem ser aprendidas. Ademais,

¹⁷ “AMBIENTE”, segundo J. Gibson (1966), refere-se ao espaço que circunda os percebedores.

segundo o estudioso, a linguagem é um código cultural, tradicional e arbitrário, mas, por outro lado, a conexão entre os estímulos e suas fontes não são.

Em termos sonoros, J. Gibson (1966) aponta que a apreensão dos conceitos é designada por um processo de dois estágios, que envolve tanto a DISCRIMINAÇÃO das articulações vocais como a APRENDIZAGEM do que elas significam. Conforme o autor sublinha, “os sons acústicos da fala especificam as consoantes, vogais, sílabas e palavras; as partes da fala, por sua vez, especificam algo mais¹⁸” (p. 91), deixando claro que ambas as classes de especificação devem ser aprendidas pela criança na etapa de aquisição da linguagem, mas os tipos de aprendizagem necessários são diferentes. Vemos, nitidamente, que J. Gibson (1966) coloca em evidência tanto a forma como a função dos itens da linguagem, indicando que a informação referente a eles se apresenta de modo simbólico e codificado.

Adiante em sua ilustração, o psicólogo afirma que os fonemas¹⁹ da fala humana podem ser especificados como respostas e como estímulos. Eles são respostas quando se referem aos movimentos articulatórios vocais no aparelho fonador, e são estímulos quando representam os padrões acústicos no meio, como transições de ondas e espectros. Em seu entendimento, J. Gibson (1966) salienta que os sons da fala especificam as suas fontes assim como os sons naturais (sons que não os da fala), além de serem estímulos potenciais (também como os sons naturais); porém, o fato de eles serem respostas AO MESMO TEMPO em que são estímulos faz deles realidades de um GRUPO SOCIAL, e não, meramente, realidades do ambiente terrestre.

Conforme expusemos anteriormente, várias ideias de J. Gibson (1966) persistiram desde a publicação de OS SENTIDOS CONSIDERADOS COMO SISTEMAS PERCEPTUAIS e foram aprimoradas em sua obra subsequente, A ABORDAGEM ECOLÓGICA À PERCEPÇÃO VISUAL (1986). Basicamente, ela se configura como um tratado sobre a percepção visual (em detrimento de outros sentidos humanos), em que o autor se preocupa em descrever a realidade circundante ao sujeito percebedor, utilizando-se de uma nomenclatura própria e procurando evidenciar supostas lacunas nas teorias óticas mais tradicionais.

J. Gibson (1986) assume, conforme o próprio nome do livro sugere, uma referência ecológica à percepção visual e, como tal, o cenário de suas ponderações abrange a relação de mutualidade entre animal e ambiente. Para o estudioso, todos os organismos capazes de perceber e de se comportar de

¹⁸ No original: “*The acoustic sounds of speech specify the consonants, vowels, syllables, and words of speech; the parts of speech in turn specify something else*” (J. GIBSON, 1966, p. 91).

¹⁹ Embora possa haver certa imprecisão por parte de J. Gibson (1966) quanto à terminologia mais apropriada em teoria fonológica, optamos por manter as palavras que o autor utilizou em seu texto, de modo a não descaracterizar sua análise. A olhos mais cartesianos, a sentença “fonemas da fala” poderia soar perturbadora, já que, tradicionalmente, o correlato do fonema que pertence à fala seria o fone.

determinada maneira são entendidos como ANIMAIS e o espaço que os circunda é chamado de AMBIENTE. Outros animais, plantas, objetos e organismos, sejam eles animados ou inanimados, também fazem parte das adjacências de qualquer animal e são, portanto, parte do ambiente. Assim, a distinção entre os componentes de determinado ambiente se faz imprescindível para que os animais possam se manter e evoluir. O conceito de PERCEPÇÃO para J. Gibson (1986) passa por alguns desdobramentos ao longo de sua obra, conforme será elencado ainda neste capítulo, mas, de maneira geral, para o psicólogo, “percepção não é a resposta a um estímulo, mas, sim, um ato de extração de informações²⁰” (GIBSON, J., 1986, p. 56-57).

A noção ecológica de J. Gibson (1986) presume que nenhum animal poderia existir sem haver um ambiente circundante – da mesma forma, a existência de um ambiente implica, necessariamente, a existência de um animal a ser circundado. O psicólogo adverte que o fato de um animal ser um organismo que percebe e se comporta no ambiente não significa que ele perceba o mundo e se comporte no espaço nos mesmos termos que a física descreve. É justamente esta ressalva feita pelo psicólogo que lhe permite estabelecer uma diferença entre o ambiente habitado pelo animal e o mundo físico. Segundo os preceitos de J. Gibson (1986), o mundo da física abrange absolutamente tudo sobre o que conhecemos, como, por exemplo, átomos, objetos terrestres e galáxias. Todavia, um ambiente não pode compreender partículas atômicas e tampouco objetos astronômicos medidos em anos-luz. J. Gibson (1986) afirma que a dimensão em que um ambiente existe é intermediária, medida em milímetros e metros. Dessa maneira, o tamanho do menor animal reflete-se na fração de um milímetro e o tamanho do maior animal está contido em alguns metros (a mesma relação pode ser estendida também para a massa dos organismos vivos).

No que tange às unidades do ambiente, J. Gibson (1986) continua sua ilustração mencionando que o ambiente dos animais e, por conseguinte, dos homens, se estrutura em diferentes níveis de tamanhos. No nível do quilômetro, encontram-se morros e montanhas, no nível do metro, estão os penhascos e árvores e, no nível do milímetro, residem os cristais e as partículas do solo. De acordo com J. Gibson (1986), todos esses componentes são unidades presentes no ambiente terrestre e são chamadas, em síntese, de formas ou configurações do nosso mundo familiar. Cabe acrescentar que, na visão ecológica defendida pelo pesquisador, atribui-se o nome de ANINHAMENTO (*NESTING*) à inserção de unidades menores em unidades maiores, algo extremamente comum quando pensamos em nível

²⁰ No original: “*Perception is not a response to a stimulus but an act of information pickup*” (GIBSON, J., 1986, p. 56-57).

de ambiente. Assim, a título de exemplificação, árvores estariam aninhadas em cânions e cânions estariam aninhados em montanhas. J. Gibson (1986) esclarece que o aninhamento teria todas as condições para constituir uma hierarquia, exceto pelo fato de que tal recursividade observada não é, sob qualquer hipótese, categórica, mas repleta de transições e sobreposições.

Paralelamente ao que descrevemos, J. Gibson (1986) elenca, quanto à geologia, que o solo é organizado em vários níveis métricos e suas unidades aninham-se umas dentro das outras, tendendo a se repetir ao longo de toda a superfície terrestre. Observamos que os “grãos de areia tendem a ser do mesmo tamanho em qualquer lugar, assim como seixos e pedras. Lâminas de grama são mais ou menos similares entre si, bem como touceiras de capim e arbustos²¹” (GIBSON, J., 1986, p. 10). Nesse aspecto, J. Gibson (1986) aponta que:

É óbvio que tais unidades naturais não são perfeitamente uniformes como pedras de calçada cortadas pelo homem. Contudo, mesmo que sua repetição não seja regular metricamente, ela é regular estocasticamente, ou seja, de maneira probabilística. Em resumo, as unidades que compõem o espaço terrestre não se alteram na medida em que nos movemos para uma ou outra direção. Elas tendem a ser uniformemente espaçadas, e se forem distribuídas, tendem a ser uniformemente distribuídas²² (GIBSON, J., 1986, p. 10).

Algo a que devemos prestar extrema atenção na abordagem de J. Gibson (1986) é o fato de que, sob o prisma ecológico vislumbrado pelo estudioso, as unidades que escolhemos para descrever o ambiente dependem largamente do nível do ambiente que desejamos descrever. Para o psicólogo, devemos nos preocupar com o nicho²³ dos animais e dos homens, cujos sistemas perceptuais são incapazes de detectar átomos ou galáxias. Por conseguinte, devemos considerar como apreendemos as mesmas coisas que nossos ancestrais humanos apreenderam antes mesmo de terem conhecimento sobre átomos e galáxias. Resta claro, portanto, que J. Gibson (1986) PREOCUPA-SE COM A PERCEPÇÃO DIRETA E NÃO COM A PERCEPÇÃO INDIRETA (A QUAL É OBTIDA ATRAVÉS DE MICROSCÓPIOS, TELESCÓPIOS, FOTOGRAFIAS E

²¹ No original: “*Grains of sand tend to be of the same size everywhere, and so do pebbles and rocks. Blades of grass are all more or less similar to one another, and so are clumps of grass and bushes*” (GIBSON, J., 1986, p. 10).

²² No original: “*These natural units are not, of course, perfectly uniform like the man-made tiles of a pavement. Nevertheless, even if their repetition is not metrically regular, it is stochastically regular, that is to say, regular in a probabilistic way. In short, the component units of the ground do not get smaller as one goes north, for instance. They tend to be evenly spaced; and if they are scattered, they tend to be evenly scattered*” (GIBSON, J., 1986, p. 10).

²³ Segundo J. Gibson (1986, p. 128), NICHU não deve ser confundido com HABITAT. O primeiro se refere a como o animal vive e o segundo refere a onde ele vive.

FIGURAS), E PREOCUPA-SE MUITO MENOS COM O TIPO DE APREENSÃO DE CONHECIMENTO OBTIDO ATRAVÉS DA FALA E DA ESCRITA²⁴. Eis o que consideramos ser a PRIMEIRA acepção do termo “direto” na obra de J. Gibson (1986), que implica a percepção de determinado objeto sem mediação ou intervenção de qualquer aparato intermediário, independentemente de sua natureza.

A descrição da escala temporal de um ambiente também é realizada por J. Gibson (1986), que interpreta fenômenos relativos ao tempo como sendo eventos ambientais. Sua visão é a de que as mudanças percebidas no mundo, aquelas das quais dependem os atos de comportamento, não são extremamente lentas e nem extremamente rápidas, uma vez que os observadores humanos não conseguem perceber a erosão de uma montanha, mas, por outro lado, podem detectar a queda de uma pedra. J. Gibson (1986) pressupõe que nós não percebemos o tempo, mas, sim, processos, mudanças e sequências, e tais eventos seriam nossa ferramenta para compreender a passagem do tempo. As unidades naturais do ambiente terrestre e dos eventos não devem ser confundidas com as unidades métricas de espaço e tempo, e, novamente, a ideia de aninhamento se faz presente em termos temporais, pois eventos encontram-se dentro de eventos.

Como James Gibson precisa reconceber os componentes do mundo físico para engendrar seu aporte ecológico à percepção, o psicólogo também trata das formas existentes em um ambiente. J. Gibson (1986) declara que a forma de um ambiente é tanto permanente em alguns aspectos como mutável em outros. Assim, por exemplo, uma cozinha é relativamente permanente quanto à forma do piso, das paredes e do teto, mas a disposição espacial da mobília pode mudar. O quanto uma forma do ambiente pode persistir depende, ao que refere J. Gibson (1986), do tipo de substância que a compõe e de sua rigidez ou resistência à deformação. Podemos verificar que, normalmente, em um ambiente, algumas formas (e objetos) persistem e outras não, o que significa dizer que um ambiente manifesta tanto características invariantes como variantes. De acordo com J. Gibson (1986), um espaço totalmente invariante e imóvel seria totalmente rígido, deixando de se configurar como um ambiente, pois não haveria animais e tampouco plantas. O oposto disso, ou seja, um espaço mutável quanto a todas as suas partes e formas, em que houvesse completa variação, também não poderia ser caracterizado como um ambiente, já que não possuiria invariantes (PEROZZO; ALVES, 2016).

²⁴ Em sua obra de 1966, James Gibson procura tratar com mais detalhes o processo de percepção e apreensão de conhecimentos advindos por meio da fala e da escrita, diferente do que ocorre na obra de 1986, em que o estudioso é mais incisivo em não prover muitos esclarecimentos nesse mérito.

Quando nos questionamos acerca das propriedades que formam o ambiente, J. Gibson (1986) faz menção a três componentes: meio, substâncias e superfície. Para o pesquisador, o meio que habitamos é homogêneo e se constitui essencialmente de elementos sólidos (terra), líquidos (água) e gasosos (ar), além de permitir nossa respiração, locomoção, visão e detecção de vibrações. Nesse contexto, em função dos efeitos da gravidade, o eixo absoluto de referência é o vertical. No âmbito das substâncias²⁵, as quais diferem primordialmente em termos de composição, rigidez, viscosidade, elasticidade e plasticidade, seus representantes seriam elementos como pedra, terra, lama, argila, madeira, óleos, minerais, metais e tecidos vegetais e animais. Ademais, J. Gibson (1986) esclarece que as substâncias provocam diferentes efeitos bioquímicos, fisiológicos e comportamentais nos animais, sendo que algumas são nutritivas, outras carecem de valor nutritivo e outras, ainda, podem ser tóxicas. As superfícies, por sua vez, separam as substâncias do meio e são o *LOCUS* em que a maior parte da ação acontece, em que há absorção ou reflexão da luz, reações químicas, transmissão de vibrações das substâncias ao meio, entre outros fenômenos.

Um ambiente só se concretiza enquanto tal se for significativo para o indivíduo e, segundo determina J. Gibson (1986), o mundo da realidade ecológica, ao contrário do mundo da realidade física, consiste em coisas significativas. Desse modo, a natureza oferece aos animais um espaço terrestre com características próprias, abrigo, água, fogo, objetos, ferramentas e, também, acesso a outros animais. De posse de cada um desses itens, o animal é capaz de se manter e evoluir em seu ambiente, relacionando-se mutuamente com seu entorno e com diversos outros animais que o cercam. J. Gibson (1986) declara que, embora seja verdade que dois indivíduos não possam estar no mesmo lugar ao mesmo tempo, é possível que qualquer indivíduo esteja em todos os lugares disponíveis e todos os indivíduos estejam em um mesmo lugar em tempos distintos, garantindo a *MESMA CHANCE* de explorar o ambiente a todos os seus habitantes. Nesse sentido, “o ambiente circunda todos os observadores do mesmo modo que circunda um único observador²⁶” (GIBSON, J., 1986, p. 43).

Em se tratando especificamente da percepção visual, J. Gibson (1986) advoga que a informação disponível para a percepção está contida na luz e, assim, a luz no ponto de observação deve ser diferente (fundamentalmente quanto à intensidade) em diferentes direções para que ela

²⁵ Conforme explica J. Gibson (1986), embora para os seres aquáticos a água seja um *MEIO*, para os humanos, ela é classificada como uma *SUBSTÂNCIA*, pois não a habitamos.

²⁶ No original: “*the environment surrounds all observers in the same way that it surrounds a single observer*” (GIBSON, J., 1986, p. 43).

possa veicular qualquer informação. Hipoteticamente, se o ar presente no ambiente fosse preenchido com uma névoa densa que impedisse a luz de se acomodar sobre as superfícies, então não existiria luz ambiente estruturada – o ar seria, portanto, translúcido, mas não transparente. Disso decorreria que existiria radiação em qualquer ponto de observação; entretanto, sem diferenças de intensidade em diferentes direções, sem transições ou gradações de intensidade, não haveria estrutura nos feixes de luz. J. Gibson (1986) defende que é justamente a luz estruturada que nos permite perceber o ambiente ao nosso entorno, bem como esclarece que a luz ambiente²⁷ torna disponíveis as informações sobre as superfícies refletoras.

Conforme já citamos anteriormente, J. Gibson (1986) preconiza que A PRESENÇA DE INFORMAÇÃO ACERCA DE UM ESTÍMULO NÃO CAUSA A PERCEPÇÃO, POIS ESTA NÃO É UMA RESPOSTA A UM ESTÍMULO, MAS DEVE SER INTERPRETADA COMO UM ATO DE EXTRAIR INFORMAÇÃO. As informações óticas, por exemplo, têm a ver com mensagens, signos e sinais, e podem ser orais, escritas ou pictóricas. A título de ilustração, uma imagem sólida ou uma escultura veicula informações e as disponibiliza a quem olha. Interessantemente, o estudioso demonstra que:

Figuras e esculturas podem ser exibidas e, portanto, contêm informação e as tornam disponíveis a qualquer indivíduo que as observe. Elas, no entanto, são como as palavras faladas ou escritas de uma língua, criadas pelo homem. Elas fornecem informação que, como aquela transmitida pelas palavras, é mediada pela percepção do primeiro observador. Elas não permitem experiência em primeira mão, somente experiência em segunda mão²⁸ (GIBSON, J., 1986, p. 63).

O excerto que acabamos de citar se faz extremamente relevante na medida em que se refere ao que julgamos ser a SEGUNDA acepção do termo “direto” no tratado gibsoniano: A APREENSÃO DE CONHECIMENTO QUE NÃO SEJA A REPRODUÇÃO DA APREENSÃO DE CONHECIMENTO ALHEIA. Em outras palavras, se o indivíduo B percebe o objeto X em virtude de o indivíduo A ter percebido X previamente e ter-lhe atribuído significação, então A PERCEPÇÃO DE B ACERCA DO OBJETO X NÃO PODE SER DIRETA. Em consonância com os

²⁷ Para J. Gibson (1986), os conceitos de LUZ AMBIENTE e LUZ RADIANTE são mantidos em distinção. LUZ AMBIENTE é toda luz que resulta da iluminação, converge para um ponto de observação, consiste de um conjunto de ângulos sólidos com um ápice comum, é diferente em diferentes direções, possui estrutura, não se propaga, depende de um ambiente de superfícies e pode ser informação. LUZ RADIANTE é toda luz que causa iluminação, diverge a partir de uma fonte, consiste de densos feixes de raios, não é diferente em diferentes direções, não possui estrutura, se propaga, vem dos átomos e volta para os átomos e é energia.

²⁸ No original: “*Pictures and sculptures are apt to be displayed, and thus they contain information and make it available for anyone who looks. They nevertheless are, like the spoken and written words of language, man-made. They provide information that, like the information conveyed by words, is mediated by the perception of the first observer. They do not permit firsthand experience – only experience at second hand*” (GIBSON, J., 1986, p. 63).

preceitos de sua obra de 1966, vale destacar que, de acordo com as próprias palavras de J. Gibson (1986), sua abordagem ecológica não se presta à apreensão de conhecimento linguístico, seja ele veiculado através da fala ou da escrita.

Um conceito exponente que permeia as ideias de J. Gibson (1986) é o das CONCESSÕES (*AFFORDANCES*). Para o psicólogo, o ambiente fornece “subsídios” a seus habitantes para que, além de se manter e evoluir, eles possam apreender conhecimento. Portanto, as concessões do ambiente são aquilo que é oferecido ao animal, o que lhe é fornecido ou aquilo com que ele é equipado para as suas ações, e se referem tanto ao ambiente quanto aos animais, implicando, por excelência, uma relação de complementaridade. O que J. Gibson (1986) sugere como percepção, consoante à sua perspectiva ecológica, é PERCEBER A COMPOSIÇÃO E A CONFIGURAÇÃO DAS SUPERFÍCIES QUE CONSTITUEM O QUE ELAS CONCEDEM. No entender de J. Gibson (1986), outro aspecto essencial em relação às concessões do ambiente reside no fato de que elas podem ser, em algum nível, físicas, reais e objetivas, assim como podem ser fenomênicas, mentais e subjetivas, ou ainda, nem objetivas nem subjetivas. Esse parece ser um quebra-cabeças gibsoniano que incide na eliminação de dicotomias como OBJETIVO-SUBJETIVO, FÍSICO-PSÍQUICO, REAL-MENTAL, e pode funcionar como um escudo para críticas de ordem filosófica e psicológica acerca da natureza do conhecimento humano.

J. Gibson (1986) indica que uma determinada concessão aponta para duas direções: para o ambiente e para o observador, o que revalida a noção de complementaridade entre estas duas entidades. Além disso, a informação para especificar uma concessão também aponta para o ambiente e para o observador à medida que a especificação das utilidades do ambiente é acompanhada pela especificação do próprio observador (por exemplo, pernas, braços, mãos e boca). Salientamos que, independentemente de qual for, a concessão não muda conforme mudam as necessidades do observador – o observador pode ou não perceber ou responder à concessão (a partir de suas necessidades), mas essa, sendo invariante, está sempre disponível para ser percebida. Nesse quesito, a percepção de uma concessão é um processo que envolve a detecção de um objeto ecológico rico em valor, e nada tem a ver com a constatação de um objeto físico livre de valor, ao qual o significado é adicionado.

Já que tocamos no mérito de o observador acessar certa concessão, julgamos apropriado trazer a esta exposição uma situação em que ocorrem erros perceptuais. Para um realista direto como James Gibson, a percepção errônea resulta da extração errônea da informação, assim como a percepção acurada resulta da extração acurada da informação. J. Gibson (1986) comenta que um

adulto pode perceber erroneamente as concessões de uma porta de vidro ao tentar passar por ela, resultando no choque com a barreira e desencadeando um possível ferimento. Para o estudioso, a concessão responsável pela colisão não foi especificada (ou suficientemente especificada) pela matriz ótica, ocasionando a substituição de vidro por ar. Em suma, caso o indivíduo incorra em um erro perceptual, a responsabilidade por tal ato e qualquer problema a ele relacionado é inteiramente do percebedor – assim, criticar o realismo direto através da noção do erro não configuraria um julgamento legítimo, pois o erro não paira sobre o ambiente, e, sim, sobre o percebedor.

Lembramos que, de acordo com J. Gibson (1986), no que diz respeito à visão, as propriedades mais elementares do ambiente que constituem uma concessão qualquer são especificadas através da estrutura da luz ambiente e, como consequência, a concessão em si é especificada na luz ambiente. Tal afirmação dá suporte ao que tomamos ser a TERCEIRA acepção do termo “direto” no tratado gibsoniano: “AS CONCESSÕES BÁSICAS DO AMBIENTE SÃO PERCEPTÍVEIS E, GERALMENTE, SÃO PERCEPTÍVEIS DIRETAMENTE, SEM UMA QUANTIDADE EXCESSIVA DE APRENDIZAGEM²⁹” (GIBSON, J., 1986, p. 143). Torna-se evidente, portanto, que eventos cognitivos relacionados a inferências e a representações mentais não estão no escopo do caráter direto da percepção. Vale salientar que, se analisarmos atentamente tais acepções do termo “direto” na obra de J. Gibson (1986), iremos observar que elas se encontram em estreita ligação com as premissas endereçadas em seu texto de 1966, corroborando o fato de que as ideias do psicólogo se mantiveram ao longo dos anos.

Tanto a ecologia da percepção quanto o realismo direto são desenvolvidos em conjunção no construto teórico gibsoniano e podem, em princípio, se coadunar sem efeitos danosos para a consistência interna da proposta de James Gibson. O trabalho do psicólogo estampa uma íntima relação com as noções de EVOLUÇÃO na biologia e mapeia o processo da percepção visual como sendo um fenômeno em que o observador deve detectar as concessões do ambiente ao seu entorno, fazendo-o através de um acesso REAL e DIRETO à configuração e à composição das superfícies que constituem tais concessões, garantido pelo seu sistema perceptual ativo. A face EVOLUTIVA do empreendimento perceptual tem a ver com a capacidade de percebermos as unidades ao nosso redor da mesma maneira que nossos ancestrais humanos o fizeram, utilizando-nos das mesmas concessões (já que elas são invariantes), e a face REALISTA DIRETA do empreendimento perceptual se relaciona a três acepções: (a) independe de um aparato entre o percebedor e o objeto a ser percebido; (b) despreza a

²⁹ No original: “the basic affordances of the environment are perceivable and are usually perceivable directly, without an excessive amount of learning” (GIBSON, J., 1986, p. 143).

percepção em segunda mão, ou seja, o resultado da percepção de outro indivíduo como fonte primária para a percepção; e (c) descompromete-se com eventos cognitivos que subjazam a inferências e representações mentais acerca dos objetos do mundo.

Todo esse empreendimento teorizado por James Gibson acerca da abordagem ecológica à percepção visual é adaptado por Carol Fowler e serve de base para a sua teoria da percepção da fala nativa, designada *TEORIA REALISTA DIRETA DA PERCEPÇÃO DA FALA (DIRECT REALIST THEORY OF SPEECH PERCEPTION)*, que culmina em 1986. Nesse período, Carol Fowler faz uma releitura da obra de James Gibson e estende o construto gibsoniano a uma unidade da fala, o gesto articulatório. Anos mais tarde, em 1995, Catherine Best leva a proposta de Carol Fowler para o cenário da percepção da fala não nativa, reaproveitando as noções gibsonianas que já tinham sido previamente adaptadas por Carol Fowler, e lança o *MODELO DE ASSIMILAÇÃO PERCEPTUAL (PERCEPTUAL ASSIMILATION MODEL)*. No entanto, antes de passar às considerações de Catherine Best sobre a percepção da fala não nativa aos moldes da abordagem de James Gibson e de Carol Fowler, julgamos importante resgatar alguns postulados de Fowler (1986) no que se refere à sua teoria perceptual da fala.

Embora J. Gibson (1986) confira alguns eventos humanos à cognição, atributo mental pelo qual o psicólogo não se interessa, definitivamente, para Fowler (1986), a percepção da fala parece não estar atrelada à atividade cognitiva, sendo um fenômeno apreendido de modo direto, contrariamente aos postulados de J. Gibson (1966, 1986). Essa antinomia reflete o quão audaz se mostra Fowler (1986) ao afirmar que a percepção da fala é direta, a ponto de contradizer a máxima de J. Gibson (1966, 1986), que explicita o fato de qualquer evento perceptual relacionado à linguagem *NÃO SER DIRETO*. Para Fowler (1986), o sinal acústico possui todas as características de um meio informacional, no sentido de que se estrutura a partir das atividades do trato vocal e pode transmitir suas propriedades a um sistema perceptual auditório. Perceber um enunciado, segundo Fowler (1986), portanto, envolve a *PERCEPÇÃO DAS DIVERSAS CONCESSÕES DE UM EVENTO CONVERSACIONAL*.

No entendimento de Fowler (1986), a percepção de uma oclusiva bilabial, por exemplo, corresponde à extração de gestos apropriadamente coordenados a partir do trato vocal do falante, assim como a percepção acerca da aproximação de uma bola corresponde à extração da informação visual a partir da matriz ótica presente no meio. Vale destacar, no entanto, que, no estágio em que Fowler (1986) desenvolveu sua teoria, a psicóloga enfrentava alguns questionamentos acerca da relação entre uma unidade sonora e sua entrada lexical, indagação que impõe dificuldades para uma visão direta da percepção da fala. É nesse tocante que a pesquisadora parece estabelecer

vigorosamente uma cisão entre eventos de fala e eventos linguísticos, e, sobre estes, Fowler (1986) admite que não tem muito a discorrer – conforme se torna explícito nas próprias palavras da estudiosa: “Eu tenho muito pouco a oferecer sobre eventos linguísticos seguindo uma perspectiva de eventos [...] e aquilo que eu realmente tenho a dizer, eu considero, de fato, bastante incipiente³⁰” (FOWLER, 1986, p. 23).

A posição de Fowler (1986), portanto, é a de que TANTO A ATIVIDADE VOCAL ESTRUTURADA FONETICAMENTE COMO A INFORMAÇÃO LINGUÍSTICA SÃO PERCEBIDAS DIRETAMENTE A PARTIR DA EXTRAÇÃO DE PROPRIEDADES INVARIANTES DISPONÍVEIS NO SINAL ACÚSTICO, embora a origem de tal informação seja diferente. A pesquisadora defende que a informação básica para a estrutura fonética é fornecida por relações coordenadas entre os articuladores, enquanto que aquela para a mensagem linguística é fornecida por restrições que refletem, por exemplo, o contexto cultural envolvido na elocução.

Captado o essencial sobre as ideias de Fowler sobre a percepção da fala enquanto um evento direto, podemos tratar do ponto de vista de Best (1995) sobre a percepção da fala não nativa. Seguindo os postulados de Fowler (1986), Best (1995) advoga em favor do GESTO ARTICULATÓRIO³¹ como sendo o primitivo da percepção, o qual deve ser encarado enquanto um objeto real produzido pelo trato articulatório e detectado diretamente na fala. Segundo a pesquisadora, a informação gestual não é construída com base na análise de simples propriedades acústicas; entretanto, a onda sonora configura o meio que veicula informação sobre os gestos do trato, além de ser modelada por eles. No plano realista direto adotado por Best (1995), largamente baseado em Fowler (1986), os bebês percebem inicialmente apenas informação não linguística na cadeia da fala, análoga à natureza dos demais sons ambientes. Ao se desenvolver, a criança passa a descobrir correspondências entre invariantes gestuais de alta ordem³² e elementos de função linguística específicos da língua que a cerca. Esta sintonização perceptual se concretiza a partir da detecção de invariantes gestuais de alta ordem, próprios de cada idioma, que se encontram disponíveis na informação multimodal da fala.

A percepção dos gestos articulatórios na fala ocorre através de um sistema perceptual geral integrado, preparado para captar informações sobre eventos articulatórios distais (FOWLER, 1986).

³⁰ No original: “*I have very little to offer concerning an event perspective on linguistic events [...] and what I do have to say, I consider very tentative indeed*” (FOWLER, 1986, p. 23).

³¹ Na próxima seção, 2.1.2, realizaremos um apanhado mais detalhado acerca do gesto articulatório. Neste momento, nos referimos a tal unidade somente a título de informação.

³² Invariante gestual de alta ordem concerne ao ponto estável do sistema fônico, que a criança em fase de aquisição alcança/internaliza em razão da repetição de padrões de movimentos coordenados [definição sugerida pela professora Adelaide Silva, professora do Departamento de Literatura e Linguística da Universidade Federal do Paraná (UFPR), durante sua arguição na defesa desta tese].

Dito de outra maneira, a informação gestual detectada vem a especificar o evento distal³³ que produziu o som, incluindo a estrutura de objetos e superfícies em jogo. Best (1995), a partir de Fowler (1986) e J. Gibson (1966, 1986), explana que as informações sobre objetos e eventos distais são verídicas e detectadas diretamente a partir do fluxo de estimulação que emana do universo do percebedor e de sua exploração ativa. Best (1995) compactua, portanto, com a ideia de que o erro perceptual, ou seja, a falibilidade, não estaria presente no objeto a ser percebido, mas no ato da percepção, atribuindo exclusivamente ao percebedor o insucesso na detecção de qualquer concessão fornecida pela estrutura das superfícies a serem contempladas.

Diferentemente dos postulados presentes nas tradições estruturalista e gerativista, o expediente fonético e o expediente fonológico, segundo a teoria de Fowler (1986) e, por conseguinte, segundo o modelo perceptual de Best (1995), fazem parte de um único domínio e compartilham da unidade gestual. Nesse domínio, Best (1995) afirma que existe um NÍVEL MAIS ALTO (*HIGHER LEVEL*), responsável pela organização sonora de uma língua, e um NÍVEL MAIS BAIXO (*LOWER LEVEL*)³⁴, responsável pela gradiência fonética veiculada na fala. De acordo com tal configuração, os detalhes fonéticos e a estrutura fonológica encontram-se em paralelo, e a cadeia da fala é potencialmente capaz de oportunizar um rico fluxo de informações diretamente apreendidas pelos percebedores, ao contrário de lhes oferecer um conjunto empobrecido de pistas acústicas ou traços distintivos. Dado esse cenário, a autora chama a atenção para a importância das informações sonoras na constituição dos inventários, alegando que “a abordagem realista direta da fonologia gestual assume um domínio gestual comum tanto para os detalhes fonéticos como para a estrutura fonológica, em que as constelações de detalhes gestuais específicos de língua SÃO os elementos da língua³⁵” (BEST, 1995, p. 183). Em outras palavras, na abordagem de Best (1995), fonética e fonologia se encontram e estão ligadas a partir de uma unidade comum, o gesto articulatório.

Muito embora a proposta de Best (1995) não seja voltada para a percepção da fala nativa, a autora recupera, novamente, as ideias de Fowler (1986), e discorre sobre a aquisição da L1. Assim, o *PAM* se apoia sobre a ideia de que a aprendizagem perceptual ocasiona a descoberta de características distintivas cruciais, isto é, as diferenças mais notáveis entre os objetos e eventos que se mostram

³³ Conforme aponta Fernández (2013), o ESTÍMULO DISTAL, OU EVENTO DISTAL, se refere ao objeto do mundo exterior sobre o qual se exerce a função perceptual.

³⁴ Essa taxionomia será problematizada no capítulo 4 desta tese.

³⁵ No original: “*the direct realist approach of gestural phonology assumes a common gestural domain for both phonetic details and phonological structure, in which the constellations of language-specific gestural details are the phonological elements of the language*” (BEST, 1995, p. 183).

importantes para o percebedor – quaisquer informações que não forem interpretadas como relevantes tendem a não ser captadas. Consequentemente, a aprendizagem perceptual compreende a exploração de invariantes no fluxo da estimulação que venham a revelar propriedades funcionais e estruturais acerca dos objetos e eventos.

No entendimento de Best (1995), é frequente que os invariantes de alta ordem emergjam a partir dos invariantes de baixa ordem, o que reforça ainda mais o compromisso fonético para com o estabelecimento de padrões fonológicos. A organização de eventos articulatórios distais, de acordo com a autora, é o reflexo dos invariantes específicos de língua. Contudo, uma grande parcela desses invariantes (ou todos eles) se encontra além do alcance perceptual dos infantes, no sentido de que estes precisam descobrir como simples gestos articulatórios são empregados em constelações gestuais de alta ordem por falantes mais velhos.

A linha de raciocínio que Best (1995) percorre visa a acoplar uma teoria filosófica da percepção (com implicações psicológicas) que parte de pressupostos ecológicos à percepção visual com uma unidade que serve ao sistema linguístico. Para tanto, a pesquisadora se vale do construto teórico referente ao REALISMO DIRETO e da ECOLOGIA DA PERCEPÇÃO preconizados por James Gibson, além de se utilizar, em grande escala, dos postulados de Carol Fowler e de um primitivo de análise desenvolvido por Catherine Browman e Louis Goldstein, a saber, o gesto articulatório. Procuramos, nesta seção, fazer um apanhado da abordagem ecológica do experimentalista americano, a qual se sustenta sobre os pilares do realismo direto, e ilustrar como ela se conecta com a percepção de sons não nativos na visão de Catherine Best, sem, necessariamente, propor uma crítica a tais postulados, o que será feito posteriormente. A próxima seção ancora-se na apresentação da unidade gestual que figura como primitivo de análise no trabalho de Best (1995), o gesto articulatório, conforme mencionamos previamente, e em como tal unidade, segundo a autora, se relaciona à percepção da fala não nativa.

2.12 O gesto articulatório enquanto elemento perceptual da fala não nativa

Na seção anterior, fomos apresentados à noção de que a evolução e a manutenção dos animais são amplamente dependentes da detecção de concessões que o ambiente oferece. No âmbito linguístico, especialmente no que diz respeito à fonologia de uma língua, “a sintonização do

percebedor à fala nativa³⁶ promove a descoberta de invariantes de alta ordem que especificam as constelações gestuais das quais o inventário fonológico nativo é composto³⁷ (BEST, 1995, p. 185). Best (1995) afirma que, no momento em que tais percebedores detectam os invariantes de alta ordem, a informação então apreendida se torna reduzida – portanto, mais compacta – em comparação ao extenso leque de características gestuais de baixa ordem que inicialmente captavam, antes de reconhecerem os padrões de coordenação entre eles. É precisamente sobre os invariantes da fala, ou seja, os gestos articulatórios e suas constelações, que nos propomos a versar na presente seção.

Conforme aponta Silva (2014), existem, segundo Fowler (1980), dois grandes grupos de modelos no domínio da teoria fonológica atualmente – os de tempo extrínseco e aqueles de tempo intrínseco. A distinção básica entre eles reside no fato de que os modelos de tempo extrínseco admitem que a variável *TEMPO* seja externa à representação fonológica e acabe por ser implementada na fala, ordenada a partir das sequências das unidades em jogo. Diferentemente, os modelos de tempo intrínseco não só admitem tal variável no tratamento do dado fônico, mas também a tomam como elemento constituinte e fundamental no que tange às representações fonológicas.

O primitivo de análise adotado pelo modelo perceptual de Catherine Best é o *GESTO ARTICULATÓRIO*, unidade real e abstrata da fala. Tal unidade se vincula à classe dos modelos fonológicos de tempo intrínseco e, mais especificamente, é o constituinte básico da fonologia gestual³⁸ (BROWMAN; GOLDSTEIN, 1989, 1992). Cabe ressaltar que Fowler (1986, 1989) insere o referencial gestual na visão realista direta, estratégia de que se utiliza Best (1995) para especificar o gesto articulatório no *PAM* (e, posteriormente, no *PAM-L2*). Uma definição bastante geral acerca dos gestos articulatórios é a de que eles são “eventos que se manifestam durante a produção da fala e cujas consequências podem ser observadas nos movimentos dos articuladores da fala³⁹” (BROWMAN, GOLDSTEIN, 1992, p. 156).

Ao que compreendem Browman e Goldstein (1992), a tarefa da fonologia articulatória é tentar descrever as unidades lexicais a propósito de tais eventos, bem como as suas inter-relações.

³⁶ Essa possível sintonização do percebedor à fala nativa é uma noção que, inevitavelmente, se choca com a base filosófica que Best (1995) segue, o realismo direto. Tal incoerência será tratada no quarto capítulo desta tese.

³⁷ No original: “*perceivers’ attunement to native speech entails their discovery of the higher-order invariants that specify the gestural constellations of which the native phonological inventory is comprised*” (BEST, 1995, p. 185).

³⁸ Utilizaremos intercambiavelmente os termos “*FONOLOGIA GESTUAL*” e “*FONOLOGIA ARTICULATÓRIA*” para nos referirmos ao mesmo modelo fonológico.

³⁹ No original: “*events that unfold during speech production and whose consequences can be observed in the movements of the speech articulators*” (BROWMAN; GOLDSTEIN, 1992, p. 156).

Os gestos seriam, portanto, o primitivo básico para a distintividade linguística e figurariam como unidades de ação articulatória. Cabe a ressalva de que, para os autores, sendo os gestos articulatórios primitivos fonológicos, eles não correspondem a segmentos e, tampouco, a traços, embora possam lembrar estas outras unidades. Nesse sentido, recuperando os postulados dos estudiosos acerca dos gestos, concluímos que eles sejam “unidades de ação discretas pré-linguísticas que se fazem inerentes à maturação de uma criança em desenvolvimento e que, portanto, podem ser tomados como elementos do sistema fonológico no curso de sua evolução⁴⁰” (BROWMAN; GOLDSTEIN, 1989, p. 69).

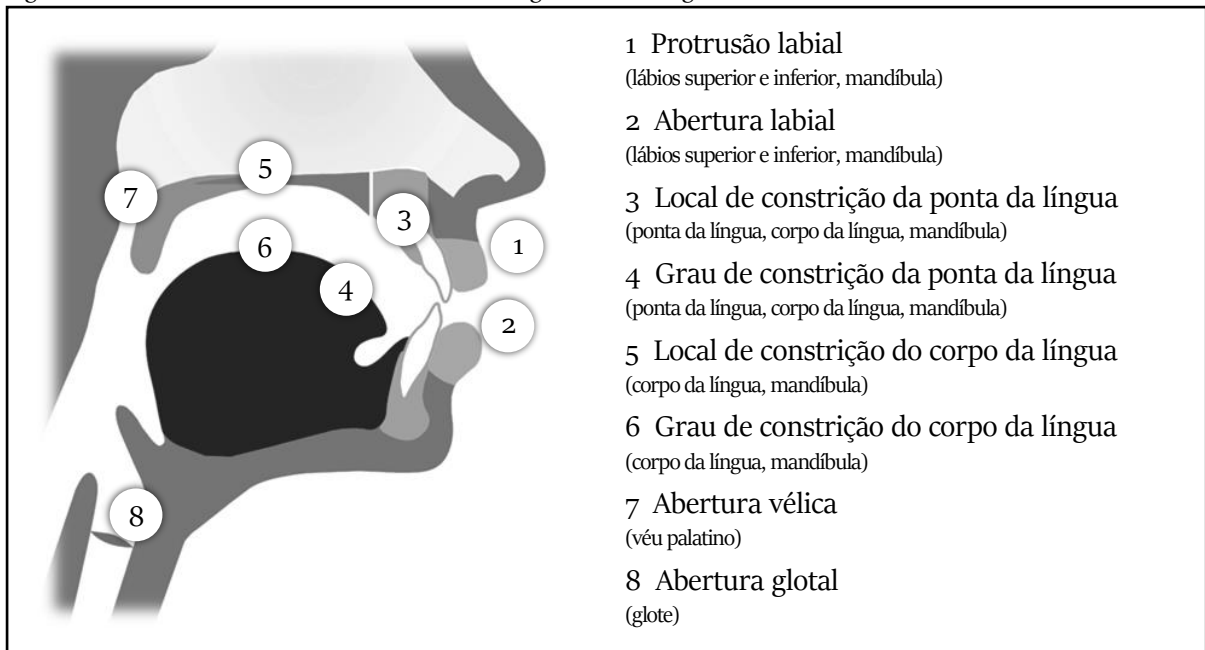
Para Browman e Goldstein (1989, 1992), os eventos articulatórios se relacionam à formação e à liberação de constrictões ao longo do trato vocal e se inserem no modelo de dinâmica da tarefa (SALTZMAN; KELSO, 1983; SALTZMAN, 1985). Vale lembrar que vários tipos de atividades articulatórias em coordenação, além da fala, se utilizam desse modelo, uma vez que ele dá conta da dinamicidade envolvida nos movimentos. Assim, os gestos articulatórios são designados a partir das variáveis do trato vocal, que especificam as dimensões das constrictões que nele ocorrem, e os articuladores se sistematizam em uma estrutura coordenativa, impactando a formação e a liberação dessas constrictões. Observemos a Figura 1, que ilustra as variáveis do trato vocal e seus articuladores.

Conforme podemos verificar a partir da Figura 1, a seguir, o gesto articulatório pode ser caracterizado com base em um conjunto de variáveis relacionadas, e a produção de sons e sílabas decorre da coordenação dos gestos⁴¹. No entendimento de Browman e Goldstein (1992), cada gesto designa uma constrictão local que é especificada em relação aos lábios, à ponta da língua, ao corpo da língua, ao véu palatino e à glote. Os autores fazem menção a quatro elementos constitutivos dos gestos articulatórios: GRAU DE CONSTRIÇÃO, LOCAL DE CONSTRIÇÃO, FORMATO DA CONSTRIÇÃO e RIGIDEZ. Cada um desses elementos desempenha papel fundamental no evento articulatório e pode ser especificado com base em um número de descritores.

⁴⁰ No original: “*prelinguistic discrete units of action that are inherent in the maturation of a developing child and that therefore can be harnessed as elements of a phonological system in the course of development*” (BROWMAN; GOLDSTEIN, 1989, p. 69).

⁴¹ Para Browman e Goldstein (1993), um conjunto de gestos coordenados entre si através de faseamento (propriedade que diz respeito ao fato de um gesto incidir sobre outro) corresponde a uma CONSTELAÇÃO GESTUAL.

Figura 1 – Variáveis do trato vocal e articuladores segundo a fonologia articulatória



Fonte: adaptada de Browman e Goldstein (1992).

O GRAU DE CONSTRIÇÃO corresponde a uma caracterização articulatória dos gestos que se relaciona ao nível de abertura dos articuladores envolvidos em determinada tarefa de produção – é análogo ao conceito tradicional de MODO DE ARTICULAÇÃO. Esse parâmetro pode ser descrito em função dos descritores: [fechado], para OCLUSIVAS; [crítico], para FRICATIVAS; [estreito] para APROXIMANTES e VOGAIS ALTAS; e [médio] e [largo] para VOGAIS MÉDIAS e VOGAIS BAIXAS.

O LOCAL DE CONSTRIÇÃO se refere ao ponto (na parede superior ou posterior do trato vocal) em que certa constrição ocorre – é análogo ao conceito tradicional de PONTO DE ARTICULAÇÃO. Tal parâmetro pode ser descrito em função dos descritores: [protruso], [labial] e [dental], para a variável LÁBIO; [labial], [dental], [alveolar], [pós-alveolar] e [palatal], para a variável PONTA DA LÍNGUA; [palatal], [velar], [uvular] e [faringal], para a variável CORPO DA LÍNGUA; e [faringal], para a variável RAIZ DA LÍNGUA.

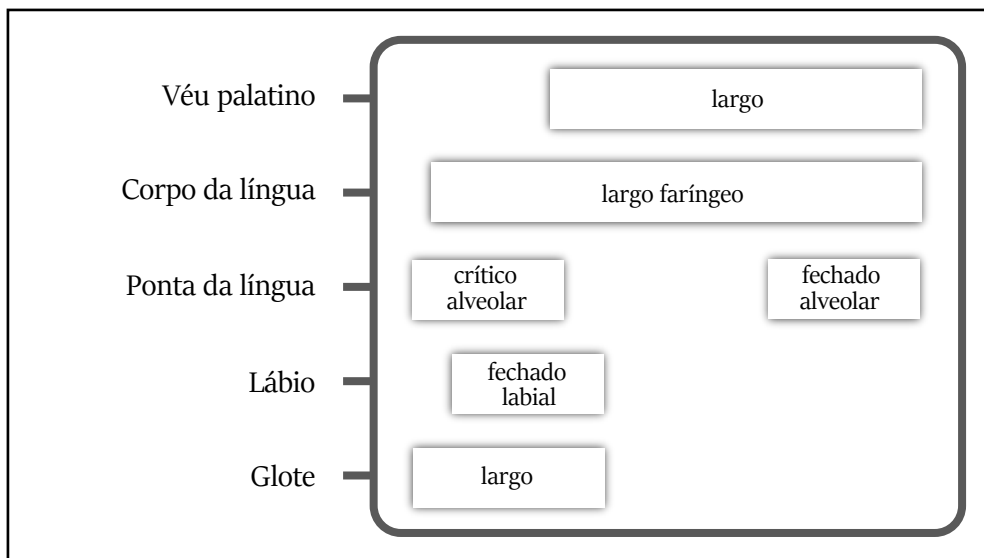
O FORMATO DA CONSTRIÇÃO⁴² implica a configuração física da constrição e está intimamente associado às variáveis PONTA DE LÍNGUA e CORPO DE LÍNGUA. Exemplos em que o formato da constrição se faz pertinente são os casos que envolvem as articulações apicais ou laminais, bem como aqueles em que há contato sublaminal na formação de oclusivas retroflexas. Consoantes laterais e róticas operam necessariamente com esse parâmetro.

⁴² O formato da constrição não foi incorporado originalmente ao modelo em jogo. Conforme pontua Silva (2002), se tal elemento tivesse sido contemplado inicialmente na fonologia gestual, tal paradigma teria sido mais exitoso no tratamento das consoantes laterais e róticas.

A RIGIDEZ está ligada à ideia de tensão empregada por um gesto articulatório e varia de acordo com o grau de constricção, o acento tônico e a taxa de elocução. Browman e Goldstein (1989) ressaltam que tal parâmetro pode ser a chave para explicar as diferenças entre as aproximantes e vogais, além de distinções entre vibrantes e tepes.

Conforme observamos anteriormente, uma das vantagens trazidas pela fonologia articulatória em comparação com outros modelos fonológicos é, segundo Silva (2014), a adição da variável TEMPO na composição dos gestos articulatórios e, portanto, a incorporação de tal variável à representação do falante. Ao levar em conta que os gestos articulatórios são unidades discretas com duração intrínseca, somos capazes de verificar as sobreposições e a sintonia dos gestos, aspectos de extrema relevância no tocante à gradiência e a fenômenos fônicos variáveis (BROWMAN; GOLDSTEIN, 1989, 1992). A título de ilustração, a Figura 2, a seguir, expressa a pauta gestual da palavra inglesa “SPAN”.

Figura 2 – Pauta gestual da palavra inglesa “SPAN”



Fonte: adaptada de Browman e Goldstein (1992).

Ao que pudemos verificar a partir da Figura 2, os gestos articulatórios podem ser formalizados por meio de pautas gestuais, que concernem a uma representação articulatória bidimensional, na qual o tempo de ativação dos gestos se aloca no eixo x e a magnitude⁴³ dos gestos

⁴³ Segundo observamos em Silva (2002), o termo “MAGNITUDE” diz respeito à ampliação ou à redução de determinado gesto fônico.

se encontra no eixo y. Quanto à palavra “*SPAN*”, temos, portanto, cinco variáveis do trato vocal cooperando para a sua produção:

- a) VÉU PALATINO, variável etiquetada com o descritor *LARGO*, indicando o abaixamento do palato mole e implicando, portanto, fluxo de ar na cavidade nasal;
- b) CORPO DA LÍNGUA, variável etiquetada com os descritores *LARGO* e *FARÍNGEO*, indicando a produção de uma vogal aberta e com constrição próxima à faringe⁴⁴;
- c) PONTA DA LÍNGUA, variável etiquetada, primeiramente, com os descritores *CRÍTICO* e *ALVEOLAR*, indicando que o fluxo de ar oral passa com fricção pelos alvéolos, e, então, com os descritores *FECHADO* e *ALVEOLAR*, indicando que a porção de ar oral proveniente da consoante nasal é bloqueada nos alvéolos;
- d) LÁBIO, variável etiquetada com os descritores *FECHADO* e *LABIAL*, indicando que ambos os lábios bloqueiam o fluxo egressivo de ar oral;
- e) GLOTE, variável etiquetada com o descritor *LARGO*, indicando que não há vibração de pregas vocais no momento da produção.

Percebemos, conforme a Figura 2, a sobreposição dos gestos ao longo do tempo, revelando que um gesto pode iniciar mesmo havendo, simultaneamente, a ativação de outro gesto. No quesito da pauta gestual observada, o gesto de ponta da língua inicia ao mesmo tempo em que inicia o gesto glote, assinalando a formação da fricativa alveolar surda. Em seguida, o gesto de lábio é ativado e se sobrepõe aos gestos de glote e ponta da língua, também ativados, caracterizando a oclusão bilabial. Nesse intervalo de tempo, o gesto de corpo da língua permanece ativado, por se tratar de um articulador presente em toda a tarefa de produção, e se estende ao longo da pauta, sinalizando a articulação da vogal aberta. Sobrepondo-se ao gesto de corpo da língua está o gesto de véu palatino, responsável pelo abaixamento do palato mole, que representa a nasalidade incidida sobre a vogal e percorre toda a sua trajetória. Por fim, há novamente a ativação do gesto de ponta da língua, de modo a ocasionar a formação da oclusão alveolar nasal. A pauta gestual endereçada capta, portanto, as trajetórias contínuas realizadas pela ativação dos diferentes gestos articulatórios que compõem a palavra inglesa “*SPAN*”.

Embora Best (1995) assumia teoricamente o gesto articulatório como seu primitivo de análise, sua explanação do modelo é majoritariamente conduzida a partir da utilização de *SEGMENTOS* fonéticos

⁴⁴ Uma vogal fechada anterior, por exemplo, seria etiquetada com os descritores *ESTREITO* e *PALATAL* (WOOD, 1982).

e fonológicos, ou seja, outra unidade de análise. Apesar de invocar um argumento *AD POPULUM*, a estudiosa procura se justificar a esse respeito, elencando que:

As línguas diferem mais fundamentalmente na orquestração de gestos simples das constelações gestuais. Tais estruturas fonológicas certamente não se restringem somente ao nível segmental da descrição linguística, mas cobrem outros níveis fonológicos, incluindo sílabas e seus constituintes, unidades rítmicas, etc. Porém, de maneira geral, a pesquisa em percepção translinguística tem se concentrado em comparações e contrastes segmentais; então, eu também focarei em segmentos e contrastes segmentais nesta discussão⁴⁵ (BEST, 1995, p. 192).

Em relação ao que apresentamos nesta seção (gesto articulatório) e na anterior (realismo direto e abordagem ecológica da percepção), destacamos que nossas considerações tiveram somente o objetivo de trazer à tona os princípios basilares de tais construtos, de modo a instrumentalizar o leitor sobre o cerne de cada um deles e sua associação ao *PAM-L2*. Salientamos que a aplicação desses dois temas no *PAM* e no *PAM-L2* será devidamente problematizada no capítulo 4 desta tese, em que serão confrontadas com teorias e abordagens alternativas. Passemos, então, à próxima seção, que tem por finalidade explicitar as relações de assimilação que regem o modelo perceptual de Best (1995).

2.1.3 Percepção da fala não nativa como condição assimilatória

Não restam dúvidas de que Catherine Best é suficientemente habilidosa no tratamento da percepção da fala não nativa, uma vez que empreende um trabalho louvável em termos de referencial teórico. Best (1995) se debruça sobre a percepção de sons não nativos amparada pelo construto de Fowler (1986, 1989), o qual conta com uma teoria filosófica (em interface com o prisma psicológico) da percepção, advinda e modificada a partir de J. Gibson (1966, 1986), e se vale de um primitivo fônico que provém de um modelo fonológico (BROWMAN; GOLDSTEIN, 1989, 1992).

De posse desse ferramental e a partir de predições de assimilação perceptual, Best (1995) estipula como *OUVINTES SEM EXPERIÊNCIA COM A FALA NÃO NATIVA* relacionam os elementos sonoros desta ao seu inventário materno, vindo a cunhar seu modelo de percepção da fala não nativa, O MODELO DE ASSIMILAÇÃO

⁴⁵ No original: “*Languages differ most dramatically in their assembly of simple gestures into gestural constellations. Such phonological structures are certainly not constrained to just the segmental level of linguistic description, but extend to other phonological levels including syllables and their constituents, rhythmic units, etc. To an overwhelming extent, however, cross-language research on speech perception has focused on segmental contrasts and comparisons, so I likewise focus on segments and segmental contrasts in this discussion*” (BEST, 1995, p. 192).

PERCEPTUAL (*PERCEPTUAL ASSIMILATION MODEL*), comumente referido como *PAM*. É no tocante a tal modelo que, nesta seção, objetivamos tecer considerações.

Duas são as observações fundamentais, segundo Best (1995), referentes a um modelo translinguístico de assimilação perceptual. A primeira delas é a definição do que são segmentos não nativos, e a segunda diz respeito a como esses segmentos são percebidos pelos ouvintes. De acordo com a autora, QUALQUER SEGMENTO NÃO NATIVO É AQUELE CUJOS ELEMENTOS GESTUAIS NÃO CORRESPONDEM PRECISAMENTE ÀS CONSTELAÇÕES GESTUAIS NATIVAS (o mesmo vale para faseamentos gestuais distintos). Quanto à percepção dos segmentos não nativos, a estudiosa pontua que ELES TENDEM A SER INTERPRETADOS DE ACORDO COM AS SUAS SIMILARIDADES OU DISCREPÂNCIAS EM RELAÇÃO ÀS CONSTELAÇÕES GESTUAIS NATIVAS QUE SE ALOCAM PRÓXIMAS A ELES NO ESPAÇO FONOLÓGICO DA LI. Nesse sentido, Best (1995) afirma que:

Em virtude de o domínio fonético universal e o espaço fonológico nativo serem definidos pelo formato do trato vocal e pelas características dinâmicas dos gestos articulatórios, estas propriedades distais fornecem as dimensões dentro das quais a similaridade é julgada. Para um falante nativo de uma língua que não tenha oclusivas dentais, mas possua oclusivas bilabiais, alveolares e velares, a constrição da ponta da língua relativa à oclusão dental ocorre diretamente mais perto, no espaço fonológico nativo, do ponto alveolar do que dos outros pontos, já que a articulação envolvida é a mesma e o local da constrição é mais semelhante do que aqueles das oclusivas bilabiais ou velares⁴⁶ (BEST, 1995, p. 193-194).

Best (1995) lança mão de dois termos extremamente relevantes para a sua argumentação, a saber, CATEGORIA e ASSIMILAÇÃO. Para a autora, o termo CATEGORIA denota a configuração prototípica de um determinado gesto ou constelação gestual, resgatando o seu respectivo espaço fonético-fonológico. Por exemplo, quando brasileiros sem experiência com a fala não nativa ouvem a fricativa interdental surda presente no inglês, eles podem vir a alocá-la na categoria nativa da fricativa labiodental surda, isto é, tendem a categorizá-la como um exemplar de [f] do português brasileiro.

O conceito de ASSIMILAÇÃO veiculado no trabalho de Best (1995) não tem relação com o processo fonológico homônimo, tratado em abordagens sobretudo gerativas, em que um determinado som compartilha propriedades de outro som adjacente (DE LACY, 2007; HAYES, 2009; CRISTÓFARO-SILVA, 2011; GUSSENHOVEN; JACOBS, 2011). A assimilação, segundo Best (1995), pode ser

⁴⁶ No original: “Because the universal phonetic domain and native phonological space are defined by the spatial layout of the vocal tract and the dynamic characteristics of articulatory gestures, those distal properties provide the dimensions within which similarity is judged. For a native listener of a language that has no dental stop but does have bilabial, alveolar, and velar stops, the tongue tip constriction of the dental stop is straightforwardly closer in native phonological space to the alveolar place than to the others, because the articulation involved is the same and the place of constriction is more similar than those of bilabial or velar stop” (BEST, 1995, p. 193-194).

entendida como a incorporação de um som qualquer da língua-alvo a determinada categoria da L_1 . A título de exemplificação, se ouvintes argentinos sem experiência com outro idioma percebem a vogal anterior média baixa do português como pertencente à categoria de [e] em seu espaço fonético-fonológico nativo, então tal vogal do português foi assimilada à categoria de [e] do espanhol e o contraste entre ambos os segmentos inexistente.

Além de detectar similaridades gestuais entre a L_1 e a L_{NN} , Best (1995) destaca que o ouvinte poderá, também, observar discrepâncias gestuais na língua-alvo quando uma constelação desta é comparada com a da sua L_1 . Os casos mais comuns de disparidade são aqueles em que o ouvinte percebe que determinado som da fala faz parte de uma língua, porém não consegue categorizá-lo em seu inventário materno. Paralelamente, a título de exemplificação, conforme aponta a própria Best (1995), norte-americanos percebem os sons cliques, presentes em algumas línguas africanas (LADEFOGED; MADDIESON, 1996), como sendo fones tão incomuns ou esdrúxulos que chegam a não ser interpretados como sons da fala.

No que concerne à percepção de segmentos não nativos, Best (1995) enumera três padrões de assimilação:

- a) ASSIMILAÇÃO A UMA CATEGORIA NATIVA, em que determinado segmento é claramente assimilado a uma categoria nativa, podendo figurar como um bom exemplar daquela categoria, como um exemplar aceitável (mas não ideal) de tal categoria, ou como um exemplar notavelmente desviante daquela categoria;
- b) ASSIMILAÇÃO COMO UM SOM NÃO CATEGORIZÁVEL DA FALA, em que determinado segmento é assimilado ao espaço fonológico nativo como uma constelação gestual da fala, mas não como um exemplar claro de qualquer categoria nativa;
- c) NÃO ASSIMILAÇÃO À FALA, em que determinado segmento não é, de maneira alguma, assimilado ao espaço fonológico nativo, e é ouvido como um som qualquer, mas não como um som da fala.

Esses padrões de assimilação servem de apoio para combinações de contrastes (pares de segmentos, por exemplo) quando elementos não nativos são percebidos, frente ao sistema da L_1 . Desse modo, Best (1995) elenca seis previsões de assimilação, as quais dispomos a seguir⁴⁷.

⁴⁷ Os exemplos de assimilação fornecidos foram criados pelo autor desta tese, são de cunho hipotético e possuem apenas caráter didático-descritivo.

Quadro 1 – Previsões de assimilação segundo o PAM (continua)

<i>TWO-CATEGORY ASSIMILATION</i> (ASSIMILAÇÃO A DUAS CATEGORIAS)	
Cada segmento não nativo é assimilado a uma categoria nativa diferente, e se espera que a discriminação seja excelente.	<p>[θ] inglês → [f] português</p> <p>[ð] inglês → [v] português</p>
<i>CATEGORY-GOODNESS DIFFERENCE</i> (DIFERENÇA DE ADEQUAÇÃO À CATEGORIA)	
Ambos os segmentos não nativos são assimilados a uma mesma categoria nativa, porém um deles é interpretado como bom exemplar da categoria e o outro é interpretado como exemplar mais desviante. Espera-se que a discriminação varie de moderada a muito boa, a depender da magnitude da discrepância entre os segmentos em questão.	<p>[ɪ] alemão ⊕ ↘</p> <p>[ʏ] alemão ⊖ ↙ [i] português</p>
<i>SINGLE-CATEGORY ASSIMILATION</i> (ASSIMILAÇÃO A UMA CATEGORIA)	
Ambos os segmentos não nativos são assimilados a uma mesma categoria nativa, porém ambos figuram como exemplares igualmente bons ou igualmente desviantes da categoria nativa. Espera-se que haja baixa discriminação entre os segmentos em jogo.	<p>[ɕ] russo ↘</p> <p>[ʂ] russo ↙ [ʃ] português</p>
<i>BOTH UNCATEGORIZABLE</i> (AMBOS NÃO CATEGORIZÁVEIS)	
Ambos os segmentos não nativos se acomodam ao espaço fonético-fonológico do ouvinte, mas não são alocados em nenhuma categoria nativa. Espera-se que a discriminação varie de baixa a muito boa, a depender da proximidade de cada um dos segmentos e das categorias nativas acionadas.	<p>[y] norueguês → [?] português</p> <p>[ɥ] norueguês → [?] português</p>
<i>UNCATEGORIZED VS. CATEGORIZED</i> (NÃO CATEGORIZADO VS. CATEGORIZADO)	
Um segmento não nativo é assimilado a uma categoria nativa, enquanto que o outro se acomoda no espaço fonético-fonológico do ouvinte, mas não cabe em nenhuma categoria nativa. Espera-se que a discriminação seja muito boa.	<p>[d] margi → [d] português</p> <p>[ɟ] margi → [?] português</p>

Quadro 1 – Previsões de assimilação segundo o *PAM* (continuação)

<i>NONASSIMILABLE</i> (NÃO ASSIMILÁVEIS)	
Nenhum dos dois segmentos não nativos se acomoda no espaço fonético-fonológico do ouvinte, sendo interpretados como sons não pertencentes à fala. Espera-se que a discriminação varie de boa a muito boa.	[+] nama → ☒ português
	[] nama → ☒ português

Fonte: adaptado de Best (1995).

Convém ressaltar que existe uma diferença entre determinado segmento ser *ACOMODADO* no espaço fonético-fonológico do ouvinte e ser *ASSIMILADO* a uma categoria nativa⁴⁸, conforme podemos observar no caso de segmentos *NÃO CATEGORIZADOS VS. CATEGORIZADOS*. Ser *ACOMODADO* no espaço fonético-fonológico do ouvinte implica simplesmente que o segmento é reconhecido como um som da fala, independentemente de ser associado a um segmento da *L1* do ouvinte. Ser *ASSIMILADO* a uma categoria nativa requer necessariamente que tal segmento se acomode no espaço fonético-fonológico do ouvinte, mas que, também, seja associado a uma ou mais categorias nativas.

As previsões de assimilação enumeradas, aliadas ao realismo direto de James Gibson, à teoria de percepção realista direta da fala de Carol Fowler e ao gesto articulatório de Catherine Browman e Louis Goldstein constituem o cerne do modelo perceptual de Catherine Best. Tais previsões funcionam como orientação para que a percepção sonora da fala translinguística seja hipotetizada, de maneira que, a partir de diferenças nas constelações gestuais, se possa prever a acuidade da discriminação quanto aos contrastes não nativos⁴⁹.

Best (1995) admite que existem ainda diversos aspectos a serem trabalhados no *PAM*, sendo um deles a sua aplicação a aprendizes de uma *L2*. Tal questão é, de fato, levada a sério e, anos mais tarde, o *PAM*, que era voltado exclusivamente a ouvintes sem experiência com outros idiomas, é estendido para dar conta da percepção da fala em *L2*. Esta nova versão, que conta com a coautoria de Michael Tyler, é chamada de *PAM-L2* e será apresentada na próxima seção.

⁴⁸ Embora Best (1995) não explicita claramente a diferença entre esses dois termos, a autora menciona que um segmento pode “cair” no espaço fonético, o que chamamos de um segmento *ACOMODADO*, mas pode não ser *ASSIMILADO* a uma categoria.

⁴⁹ É justamente a associação existente entre a assimilação a um conhecimento prévio e o realismo direto que iremos problematizar no quarto capítulo desta tese.

2.2 Assimilação perceptual em uma segunda língua: o *PAM-L2*

Devido a pontos comuns e complementares existentes na percepção da fala não nativa e na fala em L_2 , o *PAM-L2* surge, em 2007, como um modelo que visa a dar conta das relações perceptuais tanto no que compete a ouvintes sem experiência com uma L_{NN} quanto no que diz respeito a aprendizes de uma L_2 . Nesta versão expandida do *PAM*, Catherine Best e Michael Tyler cooperaram de modo a instaurar que a aprendizagem perceptual da L_2 é determinada por princípios não nativos da percepção da fala⁵⁰.

A primeira tarefa de Best e Tyler (2007) é definir a população a quem se dirigem, a saber, ouvintes não nativos e aprendizes de L_2 :

[...] definimos ouvintes não nativos mais apropriadamente do que fizemos no passado. Para nós, eles são monolíngues funcionais, que não estão ativamente aprendendo ou utilizando uma L_2 , além de serem linguisticamente ingênuos à língua-alvo que eventualmente é testada. Em comparação, aprendizes de L_2 são pessoas que estão engajadas no processo de aprendizagem de uma L_2 para que possam atingir objetivos funcionais e comunicativos, ou seja, não se encontram meramente em um contexto escolar por questões de satisfação ou em função de requisitos acadêmicos⁵¹ (BEST; TYLER, 2007, p. 16).

No que se refere a como esses dois grandes grupos se comportam frente à percepção de sons não nativos, Best e Tyler (2007) evidenciam que a percepção de contrastes entre sons por ouvintes sem experiência com o idioma-alvo é sistematicamente influenciada por vieses fonotáticos, padrões coarticulatórios e variações fonéticas ou alofônicas. Interessantemente, segundo os pesquisadores, “a percepção não está confinada a diferenças que sejam relevantes para contrastes fonológicos⁵²” (BEST; TYLER, 2007, p. 18), uma vez que monolíngues adultos exibem sensibilidade à variação fonética tanto na fala nativa como na fala não nativa. Aprendizes de L_2 , embora compartilhem de semelhanças quanto a habilidades perceptuais com ouvintes sem qualquer experiência com o idioma-alvo, diferem desses na medida em que são mais proficientes. Dessa maneira, aprendizes de L_2 já possuem um

⁵⁰ Toda a noção de realismo direto perceptual, de ecologia da percepção e de gesto articulatório que estava presente no *PAM* é trazida ao construto do *PAM-L2*.

⁵¹ No original: “[...] we define nonnative listeners more strictly than has often been done in the past. For us, they are functional monolinguals, i.e., not actively learning or using an L_2 , and are linguistically naive to the target language of the test stimuli. By comparison, L_2 learners are people who are in the process of actively learning an L_2 to achieve functional, communicative goals, that is, not merely in a classroom for satisfaction of educational requirements” (BEST, TYLER, 2007, p. 16).

⁵² No original: “perception is not confined to differences that are relevant to native phonological contrasts” (BEST; TYLER, 2007, p.18).

inventário lexical da língua-alvo (o que favorece a aprendizagem perceptual dos novos segmentos), bem como já compreendem razoavelmente sua estrutura gramatical.

Aqueles aprendizes que possuem mais familiaridade com a L_2 , segundo Best e Tyler (2007), identificam e discriminam determinados contrastes na língua-alvo com mais acuidade do que aqueles aprendizes menos experientes, porém, atingem menor acuidade se comparados a falantes nativos da L_2 em questão. A partir de evidências empíricas, os estudiosos apontam que, quanto maior for o uso da L_2 , mais hábil o aprendiz se torna, não somente no que tange à percepção dos sons, mas também no que concerne à sua produção⁵³, corroborando a importância da experiência com a L_2 para com a aprendizagem perceptual e produtiva.

Chamamos atenção para o fato de que Best e Tyler (2007) esclarecem que seu modelo se volta para aprendizes de L_2 , e não para aprendizes de L_1 . Tal cisão ocorre em virtude de que, segundo os pesquisadores, o contexto de L_1 situa-se claramente aquém das condições ideais quanto à ecologia envolvida na aprendizagem de um idioma, pois se caracteriza como um ambiente “EMPOBRECIDO” [nas palavras dos próprios autores (p. 19)] para a aprendizagem da língua-alvo. Os pesquisadores complementam que a aprendizagem de L_1 acontece em um cenário em que a L_1 exerce grande influência sobre o idioma-alvo, o qual, de maneira muito frequente, não avança para além da sala de aula. Adicionalmente, segundo os referidos autores, no quadro em que se estabelece a L_1 , as oportunidades de conversação são bastante reduzidas, o cenário se resume quase exclusivamente ao emprego de instrução formal, e há grande priorização de conteúdos gramaticais⁵⁴. Na mesma direção, Best e Tyler (2007) referem que os momentos de aprendizagem de uma L_1 se dão em um meio em que a língua-alvo comumente advém de professores com sotaque de sua L_1 , fornecendo um modelo “VARIÁVEL” ou “INCORRETO” [também nas palavras dos próprios autores (p. 19)] quanto aos detalhes fonéticos a serem aprendidos. Por conseguinte, o cenário de aprendizagem de L_2 seria, conforme Best e Tyler (2007), suficientemente adequado para assegurar a aprendizagem perceptual da língua-alvo⁵⁵.

Julgamos conveniente ressaltar que, de acordo com Best e Tyler (2007), investigações acerca da percepção de aspectos sonoros em um cenário de L_1 (e não de L_2) ainda se fazem incipientes, e seu

⁵³ Ainda que seja admitida esta relação de percepção-produção, o *PAM-L2* não é um modelo de produção, ao contrário do *SLM*, que conecta diretamente percepção e produção.

⁵⁴ Nesta passagem, notamos um tom de crítica à pedagogia de L_1 , mas devemos lembrar que um ambiente de aprendizagem de L_1 não se resume exclusivamente a contextos instrucionais.

⁵⁵ Estas afirmações dos autores, acerca da possível distinção entre L_2 e L_1 , serão discutidas criticamente no capítulo 4 desta tese.

modelo volta-se essencialmente ao contexto de L_2 . Esse condicionante não somente é evidenciado pelos autores, mas também figura como premissa substancial para que o modelo seja adotado.

No que concerne à percepção de elementos sonoros, o *PAM-L2* propõe que os fones da L_2 sejam assimilados, num primeiro momento, em categorias da L_1 já existentes, e, então, tornem-se estabelecidos como novas categorias. Esse processo ocorre, segundo Best e Tyler (2007), no nível fonético primeiramente (*LOWER-ORDER*), e, conforme o inventário lexical da L_2 aumenta, os aprendizes se alinham ao núcleo fonológico (*HIGHER-ORDER*) do idioma-alvo, tornando possível que os fones sejam discriminados a partir de diferenças de categorias que são lexicalmente importantes (ANTONIOU; TYLER; BEST, 2012).

Ao que concebem Best e Tyler (2007), 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 (PEROZZO; ALVES, 2016). Dessa forma, no *PAM-L2*, podemos verificar a existência de uma CATEGORIA FONOLÓGICA e uma CATEGORIA FONÉTICA, as quais se coadunam em prol dos aspectos funcionais no sistema, sustentadas pelos gestos articulatórios. Conforme já apontamos, a CATEGORIA FONOLÓGICA se relaciona à informação da fala que é imprescindível para que diferenças lexicais mínimas em uma língua sejam estabelecidas, e a CATEGORIA FONÉTICA se refere aos eventos gestuais invariantes que são perceptíveis de modo sistemático e potencial aos ouvintes com maior habilidade perceptual, como, por exemplo, alofones posicionais ou diferentes realizações de certa categoria fonológica entre línguas ou dialetos.

Ao estender o *PAM* ao cenário de percepção em L_2 , Best e Tyler (2007) recuperam as seis previsões de assimilação presentes no modelo de 1995 e as ajustam. Como os autores estão interessados somente em casos de assimilação perceptual, a primeira manobra a ser feita é a exclusão da última previsão de assimilação do *PAM*, ou seja, a previsão NÃO ASSIMILÁVEIS está fora do escopo do *PAM-L2*. A segunda medida a ser tomada é combinar duas previsões advindas do *PAM*, a ASSIMILAÇÃO A DUAS CATEGORIAS e a NÃO CATEGORIZADO VS. CATEGORIZADO. Se pensarmos que as discrepâncias nas configurações gestuais são responsáveis pela discriminação, então a adjunção dessas duas previsões não inviabiliza a acuidade derivada da oposição na L_2 . Como resultado, o *PAM-L2* visa a dar conta da discriminação de contrastes da L_2 considerando apenas quatro das seis previsões disponíveis em seu modelo predecessor, a saber: ASSIMILAÇÃO A DUAS CATEGORIAS (em conjunção com NÃO CATEGORIZADO VS. CATEGORIZADO); DIFERENÇA DE ADEQUAÇÃO À CATEGORIA; ASSIMILAÇÃO A UMA CATEGORIA; e AMBOS NÃO CATEGORIZÁVEIS.

Mais do que prever relações assimilatórias entre contrastes gestuais da L_2 para a L_1 , O *PAM-L2* constitui uma tentativa de dar ao fenômeno perceptual a oportunidade de servir como elemento primordial na formação da fonologia da L_2 . Vemos esse passo como sendo digno de grande mérito, uma vez que a instanciação de uma gramática (seja ela fônica ou de outra ordem) só pode se concretizar via mecanismos perceptuais. Este capítulo serviu de ferramental teórico, de cunho principalmente descritivo, para que os três eixos fundamentais do *PAM* e do *PAM-L2* fossem apresentados: sua posição cognitiva (implícita), sua filiação linguística em termos de unidade de análise e sua teoria filosófica. Assim, não foi nosso objetivo apontar, essencialmente, o que consideramos ser os problemas dos modelos, mas, sim, instrumentalizar o leitor para que ele seja capaz de compreender a razão e a natureza das críticas que serão feitas a ambos os trabalhos no capítulo 4 desta tese.

2.3 Resumo do capítulo

Este capítulo tratou da apreciação do modelo *PAM* e de seu sucessor, *PAM-L2*. Expusemos que os dois construtos fundamentais para a elaboração do *PAM* foram, de um lado, a noção de ecologia e de realismo direto de James Gibson (com ênfase em suas obras de 1966 e de 1986) e, de outro lado, a teoria de percepção da fala de Carol Fowler, que tem por primitivo o gesto articulatório e que, de certa forma, pressupõe aspectos cognitivos, ainda que indireta e implicitamente (como, por exemplo, a seleção dos invariantes e o julgamento desses em detrimento de outros). J. Gibson (1966, 1986) defende que a percepção dos objetos do mundo se dá de maneira direta, já que os nossos sentidos atuam como sistemas orientados à extração de informações, de cunho invariante, disponíveis no ambiente. Disso decorre a hipótese de que não precisamos apelar a inferências, a ideias ou a representações simbólicas para conseguirmos perceber o que nos é oferecido pelo nosso entorno. Além disso, o psicólogo advoga a favor de que uma percepção direta dispensa a mediação ou a intervenção de aparatos intermediários, como telescópios ou fotografias, e também estabelece que a percepção “em segunda mão” não configura uma experiência direta porque é o resultado da percepção do outro.

Na teoria de Fowler (1986), os gestos articulatórios são percebidos diretamente (sem a necessidade de mecanismos cognitivos) através do meio acústico, o qual se faz suficientemente capaz de veicular as características invariantes expressas por tais unidades e detectadas pelos sistemas de

percepção. Sublinhamos a advertência de J. Gibson (1966, 1986) quanto aos sons da fala não poderem ser apreendidos diretamente porque envolvem um código permeado por simbolismo e cuja estrutura é socialmente acordada.

A partir desse arcabouço teórico, Catherine Best assume que seu modelo (*PAM*), voltado para a percepção de sons da fala não nativa, dá conta de predizer com sucesso as possíveis categorizações de elementos sonoros com base em diferenças nas constelações gestuais que são acessadas diretamente pelos percebedores/ouvintes sem contato prévio com o idioma a que são expostos. O cerne de suas previsões compreende padrões de assimilação perceptual, que respondem pelos gestos articulatórios não nativos serem ou não categorizados no espaço fonético-fonológico do sistema materno.

De modo a estender seu programa de pesquisa à percepção de uma L_2 , Catherine Best conta com Michael Tyler e, conjuntamente, fazem alguns ajustes ao modelo, o qual passa, então, a ser chamado de *PAM-L2*. Nessa versão, os autores estipulam aspectos comuns e complementares entre ouvintes sem experiência com idiomas não nativos e aprendizes de uma L_2 , além de reconfigurarem as previsões de assimilação perceptual, passando de seis para quatro padrões preditivos. Distinções entre L_2 e L_1 também são apontadas no *PAM-L2*, e seus proponentes afirmam que ao modelo compete apenas o contexto de aprendizagem de L_2 , pois o ambiente de L_1 é deficiente em diversos aspectos, especialmente no que tange à qualidade do *INPUT*.

Passaremos, neste momento, ao terceiro capítulo desta tese, o qual tem por finalidade abordar uma série de estudos que investigaram a percepção de sons não nativos [indo, portanto, ao encontro do objetivo (a) desta tese]. Algumas das pesquisas se utilizaram do *PAM-L2* para se sustentar teoricamente e/ou para interpretar seus respectivos achados, enquanto que outras prescindiram do modelo tanto como referencial teórico quanto como método para vislumbrar os resultados obtidos. Tal levantamento é de fundamental importância para verificarmos se existem, de fato, mais possibilidades de assimilação que não foram contempladas pelo *PAM-L2*, e também para, posteriormente, fornecer insumo acerca da elaboração de tarefas perceptuais que sejam capazes de testar empiricamente as previsões advindas do modelo.

3 ESTUDOS EM PERCEPÇÃO DE SONS DE LÍNGUAS NÃO NATIVAS

O *PAM-L2* é um modelo perceptual de grande abrangência nos estudos interfonológicos. Esta constatação serve de apoio para desenvolvermos esta tese, no sentido de que ela se justifica, em grande medida, em função de o modelo cobrir preditivamente uma variada gama de fenômenos fônicos.

O presente capítulo está ligado ao objetivo (a) desta tese e propõe elencar trabalhos que versam sobre a percepção de sons de línguas não nativas, conduzidos tanto em nível nacional como internacional. O levantamento bibliográfico a ser delineado tem o propósito de averiguar se as previsões de assimilação perceptual sugeridas por Best e Tyler (2007) se fazem exaustivas, considerando-se um grande rol de estudos. Além disso, as pesquisas relatadas também serão úteis para que possamos contemplar o objetivo (f) deste trabalho, que se vincula à discussão de procedimentos empíricos para a testagem das categorizações fônicas elencadas no *PAM-L2*, a ser apresentada no capítulo 5.

A condução do levantamento bibliográfico ilustrado neste capítulo obedeceu a critérios que residem no expediente do CONTEÚDO e do GÊNERO TEXTUAL dos trabalhos. Do ponto de vista do conteúdo, consideramos em nossa revisão somente os trabalhos que levaram em conta a PERCEPÇÃO DE ELEMENTOS FÔNICOS POR POPULAÇÕES NÃO NATIVAS do idioma em questão. Adicionalmente, enumeramos os estudos que se debruçaram sobre a percepção fônica apenas no PLANO SEGMENTAL; portanto, não foram analisados aqueles trabalhos envolvendo a percepção de constituintes suprasegmentais (como sílaba, palavra, etc.), de fenômenos prosódicos e de propriedades relacionadas a acento, tom e entonação. Finalmente, tampouco mencionamos estudos cuja natureza se voltou para condições adversas de percepção, como, por exemplo, situações de barulho ou ruído.

Do ponto de vista do gênero textual, selecionamos modalidades específicas de textos, a saber, ARTIGOS CIENTÍFICOS, DISSERTAÇÕES e TESES, como sendo as fontes de insumo específicas para a revisão da literatura sobre percepção de sons não nativos. No que se refere aos artigos científicos, salientamos que esses são mais expressivos, em termos quantitativos, do que as dissertações e as teses.

Para o levantamento dos artigos científicos, realizado de outubro de 2015 a abril de 2016, selecionamos aqueles registrados no Portal de Periódicos da Coordenação de Aperfeiçoamento de

Pessoal de Nível Superior (Capes)⁵⁶ sob os índices A1 e A2 do Qualis (2014), publicados nos últimos 7 anos (desde 2010) – espelhando o critério temporal utilizado na busca das dissertações e teses (triênio 2010-2012 e quadriênio 2013-2016, de acordo com a avaliação dos cursos de pós-graduação da Capes). Além disso, também elegemos os artigos publicados nos periódicos internacionais em que constam as pesquisas de Catherine Best e Michael Tyler, disponíveis em suas páginas eletrônicas pessoais⁵⁷, uma vez que estivessem relacionadas ao nosso objeto de estudo e respeitassem as condições de conteúdo que estabelecemos. Ao total, procuramos por artigos científicos em 54 periódicos, dos quais 12 eram periódicos nacionais e 42, internacionais⁵⁸.

Na página eletrônica de cada periódico, selecionamos os anos das publicações e, em cada ano, realizamos uma busca também por título. Assim, foram observados todos os títulos de todas as publicações no respectivo ano. Quando o título não se fazia informativo, analisávamos o seu resumo e, eventualmente, quando este não era satisfatoriamente esclarecedor, examinávamos todo o texto. Ressaltamos que nossa opção pela busca por títulos nos periódicos, em detrimento da busca por palavras-chave, decorre do fato de que nem sempre as palavras-chave são fiéis à essência do trabalho por que respondem. Por conseguinte, dados os critérios de seleção a que aderimos, chegamos a 27 artigos (1 artigo nacional e 26 artigos internacionais) relevantes para a nossa discussão, os quais serão reportados em ordem cronológica nesta seção.

Sublinhamos que nem todos os artigos científicos a serem reportados se utilizaram dos preceitos do *PAM-L2* como fundamentação teórica ou como meio de interpretar seus achados. Assim, no que compete aos estudos que tomaram por base as previsões assimilatórias do *PAM-L2*, salientaremos os casos de assimilação perceptual verificados (informados pelos respectivos autores) e, no que concerne aos estudos que não remontam às previsões assimilatórias do modelo, procuraremos averiguar se os seus resultados e as suas interpretações – em termos dos padrões de

⁵⁶ O Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) é uma biblioteca virtual que reúne e disponibiliza, a instituições de ensino e pesquisa no Brasil, trabalhos referentes à produção científica nacional e internacional. O repositório conta com um acervo de mais de 37 mil títulos com texto completo, 128 bases referenciais, 11 bases dedicadas exclusivamente a patentes, além de livros, enciclopédias e obras de referência, normas técnicas, estatísticas e conteúdo audiovisual (texto adaptado da página eletrônica do Portal de Periódicos da Capes). O Portal pode ser acessado através do link <http://www.periodicos.capes.gov.br>.

⁵⁷ A página eletrônica pessoal de Catherine Best pode ser encontrada através do link https://www.westernsydney.edu.au/staff_profiles/uws_profiles/professor_catherine_best, e a de Michael Tyler pode ser acessada por meio do link https://www.westernsydney.edu.au/staff_profiles/uws_profiles/associate_professor_michael_tyler.

⁵⁸ A lista de periódicos consultados pode ser encontrada na seção de apêndices desta tese.

categorização – poderiam se alinhar à perspectiva da proposta de Best e Tyler (2007). Nesse último caso, tomaremos como parâmetros de interpretação a análise e a configuração dos dados referentes aos índices de acuidade ou aos escores obtidos pelos participantes nas tarefas de percepção empregadas, de maneira que seus resultados possam ser atribuídos a previsões de assimilação específicas.

Conforme mencionamos no capítulo introdutório desta tese, é indispensável referir, em cada trabalho, o primitivo de análise adotado e verificar se os dados podem ser tratados via *PAM-L2*, bem como constatar se as previsões estabelecidas pelo modelo dão conta de abarcar os resultados encontrados. Ademais, analisar cada um dos artigos científicos nos permite contemplar se as previsões elencadas pelo *PAM-L2* se mostram exaustivas no que concerne às categorizações reportadas. Dessa forma, a apresentação de cada artigo estará, quase unanimemente, condicionada à seguinte estrutura: (a) primeiramente, será disposto o objeto de investigação da pesquisa; (b) em seguida, apontaremos os objetivos e/ou as hipóteses do trabalho; (c) então, discorreremos sobre a metodologia utilizada; e (d) por último, trataremos dos resultados encontrados e sua respectiva discussão.

Em relação às teses e às dissertações, o procedimento de busca por trabalhos nacionais que trataram do tema foi baseado na metodologia da pesquisa de Alves (2015). Com o objetivo de realizar um mapeamento da área de aquisição fonético-fonológica de *L2/LE*, o referido autor realizou um levantamento de trabalhos de mestrado e doutorado de 14 universidades brasileiras⁵⁹ durante o triênio 2010-2012.

Além de nos valermos do levantamento de Alves (2015), replicamos o procedimento de busca adotado pelo pesquisador de modo a abarcar, nos mesmos programas pesquisados, o quadriênio 2013-2016, interstício esse também estipulado pela Capes. Assim, para podermos formar um corpo de trabalhos a serem resenhados, nosso percurso de análise se deu tanto a partir da enumeração de Alves (2015) como da verificação cabível ao período que se seguiu (o quadriênio 2013-2016), abrangendo, portanto, dois períodos temporais em sequência.

⁵⁹ Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS); Universidade Católica de Pelotas (UCPel); Universidade de Brasília (UNB); Universidade de São Paulo (USP); Universidade Estadual do Ceará (UECE); Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP); Universidade Federal da Bahia (UFBA); Universidade Federal da Paraíba (UFPB); Universidade Federal de Pernambuco (UFPE); Universidade Federal de Santa Catarina (UFSC); Universidade Federal de Santa Maria (UFSM); Universidade Federal do Paraná (UFPR); Universidade Federal do Rio de Janeiro (UFRJ); e Universidade Federal do Rio Grande do Sul (UFRGS).

Procedimentalmente, acessamos as páginas eletrônicas dos programas de pós-graduação em Letras e/ou Linguística das universidades pesquisadas por Alves (2015) e tivemos acesso aos trabalhos desenvolvidos em cada instituição. Com base nos critérios temporais e de conteúdo mencionados, procuramos pelas dissertações e teses através de seus títulos e, quando esses não eram suficientemente informativos, passávamos à análise de seu resumo. Este primeiro levantamento resultou em 7 dissertações de mestrado e 3 teses de doutorado.

Por questões de delimitação de espaço, no que concerne aos trabalhos brasileiros de pós-graduação, iremos apresentar somente aquelas pesquisas que vislumbraram seus resultados explicitamente à luz das previsões assimilatórias do *PAM-L2*. Entendemos que, se seguissemos o paralelismo estrutural da apresentação feita em relação aos artigos científicos (que serão endereçados mesmo se seus resultados não tiverem enfoque nas previsões assimilatórias do *PAM-L2*), teríamos um capítulo demasiadamente extenso, ainda mais por se tratar de gêneros acadêmicos que preveem maior riqueza de detalhes nas informações, diferentemente de artigos científicos, que tendem a ser mais compactos. Dessa forma, o objetivo que propomos cumprir [objetivo (a)] a partir deste capítulo será satisfeito com base na contribuição majoritária dos resultados provenientes dos artigos científicos.

Para fins organizacionais, as duas próximas seções irão relacionar os artigos científicos revisados, sendo que, em uma delas (3.1), haverá a apresentação dos artigos que tiveram enfoque explícito nas previsões de assimilação do referido modelo e, em outra (3.2), apresentaremos os artigos que não tiveram enfoque em tais previsões de assimilação. A seção seguinte (3.3) irá contemplar os trabalhos brasileiros de pós-graduação que vislumbraram seus resultados à luz dos padrões assimilatórios do *PAM-L2*. A penúltima última seção (3.4), que se utilizará fundamentalmente das seções prévias, tratará de explicitar nossas considerações sobre os estudos como forma de oferecer respostas ao objetivo do capítulo, relacionado à verificação da exaustividade das previsões de assimilação veiculadas no *PAM-L2*. Ao final, apresentaremos o resumo do capítulo (seção 3.5).

Antes de darmos início à seção 3.1, frisamos que os mesmos estudos reportados neste capítulo serão discutidos sob o enfoque procedimental no capítulo 5 desta tese, de modo a fornecer insumos metodológicos para a verificação de procedimentos empíricos de testagem acerca das categorias fônicas que podem ser estabelecidas pelos ouvintes não nativos.

3.1 Artigos científicos com enfoque nas previsões assimilatórias do *PAML2*

Os estudos a serem relatados nesta seção levaram em consideração, de forma explícita, as previsões de assimilação veiculadas no *PAM-L2*, e serão apresentados individualmente em curtas subseções.

3.1.1 Gilichinskaya e Strange (2010)

A pesquisa de Gilichinskaya e Strange (2010) avaliou a similaridade perceptual das vogais do inglês americano em relação às vogais do russo por ouvintes russos monolíngues. O objetivo do trabalho era constatar como os ouvintes russos assimilariam algumas vogais do inglês americano e, então, discorrer sobre a influência da similaridade acústica entre os sistemas quanto à categorização dos segmentos e elencar quais deles se mostram mais difíceis em termos de assimilação.

A investigação tomou forma com base em duas questões de pesquisa. A primeira propôs verificar quão acuradamente a similaridade acústica entre as vogais russas e inglesas previa a sua assimilação perceptual pelos ouvintes russos, e a segunda procurou explorar quais contrastes figuravam como sendo os mais difíceis perceptualmente para os mesmos ouvintes.

Participaram do estudo 19 falantes nativos de russo com idade média de 36 anos. Os participantes eram pesquisadores, técnicos e/ou alunos de pós-graduação pertencentes a um laboratório de criologia do solo, com sede na Universidade de Pushchino em Moscou. A maioria deles havia estudado inglês no ensino médio e outros cursaram disciplinas voltadas à língua inglesa durante a faculdade. Aqueles que tiveram contato com o inglês fora da Rússia o fizeram por um período de tempo inferior a 8 meses e, segundo os autores, a experiência linguística desses aprendizes em específico não viria a comprometer os resultados da investigação.

Os estímulos auditivos utilizados na pesquisa corresponderam a 8 vogais do inglês americano (/i:, I, ε, æ:, α:, Λ, U, u:/), gravadas por um falante da variedade nova-iorquina, as quais foram extraídas da sequência dissilábica /Vpə/. Tais vogais compuseram a tarefa de identificação aplicada aos participantes, que contava com a apresentação auditiva de um segmento mais uma escala de adequabilidade de 9 pontos (1 = diferente da vogal em russo; 9 = similar à vogal em russo), segundo a qual o participante comparava a similaridade do segmento ouvido com aquele referente ao da sua L₁. Havia 4 blocos de estímulos e cada um apresentava 24 estímulos (8 vogais x 3 repetições). A lista

de respostas exibiu 9 opções de vogais possíveis, [i, i̇, εa, a, jε, ε, o, u, uo], registradas de acordo com a ortografia russa padrão.

Os resultados encontrados indicaram que as vogais inglesas [i:, α:, u:] foram as mais consistentemente assimiladas às categorias vocálicas russas [i, a:, u], enquanto que as vogais inglesas [ε, æ:, ʌ] apresentaram as respostas menos consistentes no que compete à assimilação, mas, ainda assim, tais vogais foram associadas às categorias vocálicas russas [a, εa]. As vogais inglesas [ɪ, ʊ] foram assimiladas às categorias russas [ε, o], respectivamente, com um nível intermediário de consistência. Em aspectos gerais, os autores postulam que, com exceção da vogal inglesa [ε], a similaridade acústica, definida por padrões espectrais na região central do segmento, correspondeu satisfatoriamente com as assimilações perceptuais realizadas pelos ouvintes russos.

No que concerne à segunda questão de pesquisa proposta, os estudiosos apontam que as sobreposições acerca das categorizações feitas pelos russos foram avaliadas na modalidade intrassujeito. Assim, os segmentos ingleses [ɪ, ε] foram consistentemente assimilados ao segmento russo [ε] por 9 participantes; da mesma forma, os segmentos ingleses [α:, ʌ] foram consistentemente assimilados ao segmento russo [a] por 9 participantes. Devido aos altos índices de adequabilidade de tais contrastes frente aos segmentos russos (o que significa uma neutralização das distinções entre as vogais-alvo), os autores, em termos do *PAM*, julgaram se tratar de casos de ASSIMILAÇÃO A UMA CATEGORIA; no que tange aos contrastes com níveis moderados de adequabilidade, os pesquisadores sugerem que sejam casos do padrão NÃO CATEGORIZADO VS. CATEGORIZADO⁶⁰; e quanto aos casos em que os contrastes apresentavam baixos níveis de adequabilidade, ou seja, quando as distinções entre as vogais inglesas e russas eram plenamente mapeáveis, os estudiosos indicam instâncias de ASSIMILAÇÃO A DUAS CATEGORIAS, o que representa o estado ótimo da diferenciação de contrastes.

3.12 Bundgaard-Nielsen, Best e Tyler (2011b)

Bundgaard-Nielsen, Best e Tyler (2011b) verificaram a percepção das vogais do inglês australiano por falantes nativos de japonês, com o intuito de vislumbrar a assimilação perceptual das vogais do inglês às categorias nativas dos japoneses.

⁶⁰ No que diz respeito ao *PAM-L2*, o padrão NÃO CATEGORIZADO VS. CATEGORIZADO se incorpora ao padrão ASSIMILAÇÃO A DUAS CATEGORIAS, conforme apontado no segundo capítulo desta tese.

Participaram da pesquisa 31 estudantes universitários (23 mulheres e 8 homens), com idade média de 23 anos, os quais estavam aprendendo inglês na Austrália havia menos de três meses, mas tinham estudado inglês previamente no Japão durante aproximadamente nove anos, instruídos por professores cuja L_1 era o japonês.

Os participantes foram testados longitudinalmente (em duas etapas) a partir de uma tarefa⁶¹ de identificação perceptual e de uma tarefa de discriminação categórica, cujos estímulos eram palavras dissilábicas, gravadas por falantes nativos de inglês australiano, as quais se estruturavam de acordo com a sequência /hVba/ e apresentavam tonicidade na primeira sílaba.

Na tarefa de identificação perceptual, os participantes ouviram 324 palavras e deveriam identificar a vogal da primeira sílaba aos moldes de suas vogais nativas, além de registrar o quão próxima da sua L_1 era a vogal ouvida no estímulo da língua alvo, através de uma escala de 1 (exemplar ruim) a 7 (exemplar excelente). As hipóteses para a tarefa de identificação eram as de que (1) as vogais curtas do inglês australiano (/ɪ, e, ɐ, ɔ, u/) seriam assimiladas perceptualmente às contrapartes vocálicas monomoraicas⁶² do japonês (/i, e, a, o, u/); (2) as vogais longas do inglês australiano (/i:, e:, ɐ:, o:, ɯ:/) seriam assimiladas perceptualmente às contrapartes vocálicas bimoraicas do japonês (/i:, e:, a:, o:, u:/); e (3) os ditongos do inglês australiano, tais como /eɪ, æ, oɪ/, seriam assimilados perceptualmente aos ditongos bimoraicos correspondentes do japonês, como /ei, ai, oi/.

Na tarefa de discriminação categórica, a qual seguiu o padrão AXB ⁶³ e contou com as mesmas 324 palavras agrupadas em tríades, os participantes deveriam marcar em qual posição estava o

⁶¹ Independentemente da taxionomia utilizada pelos autores para se referirem aos instrumentos de coleta de dados empregados em seus estudos, seguimos Primi, Nascimento e Souza (2004) no que compete à diferença entre *TESTE* e *TAREFA*. Para os autores, *TESTES* propõem *TAREFAS* específicas para observar a manifestação do comportamento. Assim, tarefas de identificação e discriminação, por exemplo, encontram-se dentro do âmbito geral referente aos testes perceptuais.

⁶² Os termos “*MONOMORAICAS*” e “*BIMORAICAS*” se referem, como os nomes sugerem, ao conceito de *MORA*. Em teoria fonológica, a unidade *mora* diz respeito ao peso (ou à quantidade) de uma sílaba atribuído aos segmentos que a compõem. No trabalho de Hayes (1989), uma das análises *moraicas* mais difundidas, o autor classifica *monotongos* como *monomoraicos* (uma *mora*) e *vogais longas*, assim como *ditongos*, como *bimoraicos* (duas *moras*).

⁶³ Conforme será explicitado no quinto capítulo desta tese, o padrão AXB corresponde a um formato de tarefa de discriminação em que três elementos compõem determinado estímulo, sendo que *A* e *B* devem ser elementos distintos e *X* deve ser igual a *A* ou a *B*. Nesse formato, a questão apresentada ao participante geralmente é “*O elemento que ocupa a segunda posição no estímulo é igual ao primeiro elemento ou é igual ao último elemento da tríade?*”. Outro padrão similar de tarefa de discriminação é o do tipo ABX , em que, novamente, *A* e *B* são elementos distintos e *X* corresponde a *A* ou a *B*. Diferentemente, nesse formato, a questão usualmente apresentada ao participante é “*O último elemento do estímulo é igual ao primeiro elemento ou é igual ao segundo elemento da tríade?*”. Na versão mais simplificada do formato

segmento incompatível com os demais na sequência apresentada. Os segmentos utilizados na tarefa de discriminação categórica corresponderam a quatro contrastes, a saber, /i:/-/ɪə/, /i:/-/ɪ/, /ɜ:/-/ʌ:/ e /əʊ-/ /o:/. As hipóteses para a tarefa de discriminação eram as de que (1) a discriminação do contraste /i:/-/ɪə/ seria razoável e exibiria o padrão de ASSIMILAÇÃO A UMA CATEGORIA; (2) a discriminação do contraste /i:/-/ɪ/ seria excelente e exibiria o padrão de ASSIMILAÇÃO A DUAS CATEGORIAS; (3) a discriminação do contraste /ɜ:/-/ʌ:/ poderia variar de boa a excelente e exibiria o padrão de ASSIMILAÇÃO A DUAS CATEGORIAS; e (4) a discriminação do contraste /əʊ-/ /o:/ deveria ser moderada e também exibiria o padrão de ASSIMILAÇÃO A DUAS CATEGORIAS. Em ambas as tarefas, os estímulos foram apresentados em um microcomputador através do qual os participantes poderiam fornecer suas respostas.

Comparando-se as duas etapas de testagem (com intervalo de seis a oito meses), os autores encontraram uma associação positiva entre o tamanho do inventário lexical e a percepção vocálica na língua alvo. Os resultados do estudo corroboram as hipóteses enumeradas e, conforme podemos observar, as previsões de assimilação elencadas no *PAM-L2* dão conta dos resultados encontrados, em que os padrões em questão são (1) ASSIMILAÇÃO A UMA CATEGORIA, em que dois fones da língua alvo são assimilados a uma única categoria da L_1 ; e (2) ASSIMILAÇÃO A DUAS CATEGORIAS, em que dois fones da língua alvo são assimilados a duas categorias distintas da L_1 .

3.1.3 Escudero e Vasiliev (2011)

Escudero e Vasiliev (2011) debruçaram-se sobre a assimilação perceptual de similaridades acústicas entre o inglês e o francês canadenses por ouvintes peruanos monolíngues cuja L_1 é o espanhol. O objetivo do estudo foi o de verificar se os ouvintes peruanos eram capazes de categorizar auditivamente as vogais [ɛ] e [æ] produzidas por falantes de inglês e francês canadenses.

Os autores partem do princípio de que semelhanças acústicas entre segmentos implicam semelhanças perceptuais em relação aos sons da fala, além de conceberem que a percepção auditiva de sons nativos ou não nativos deve se relacionar a propriedades acústicas dos sons na L_1 . Assim, os pesquisadores examinaram se as diferenças acústicas entre os segmentos [ɛ] e [æ], produzidos tanto

AxB , a Ax , o elemento x pode ser igual ou diferente do elemento A , e a questão apresentada ao participante comumente é “Os elementos do estímulo são iguais ou diferentes?”.

por canadenses falantes de inglês como por canadenses falantes de francês, levariam os ouvintes peruanos a diferentes mapeamentos perceptuais (assimilações).

Participaram do estudo 64 ouvintes monolíngues de espanhol peruano, com idade média de 24 anos, os quais foram convidados a realizar uma tarefa de identificação perceptual, composta por 30 estímulos que compreendiam as vogais [ɛ] e [æ] inseridas nas sílabas /bVs/, /bVk/, /bVg/, /sVk/ e /vVst/. Os participantes foram informados de que eles ouviriam produções de vogais do espanhol articuladas em fala rápida (em vez de vogais não nativas), as quais deveriam ser categorizadas segundo uma das cinco vogais espanholas /i, e, a, o, u/, representadas através de grafemas na tela do computador.

As hipóteses para a tarefa de identificação, considerando-se propriedades acústicas, eram as de que (1) o contraste /ɛ/-/æ/ produzido em inglês canadense seria assimilado à categoria de /a/ em espanhol peruano, enquanto que o contraste /ɛ/-/æ/ produzido em francês canadense seria assimilado ao contraste /e/-/a/ em espanhol peruano; e (2) as diferenças acústicas entre as vogais /ɛ/-/æ/, produzidas tanto por falantes de inglês como por falantes de francês, resultariam em diferentes mapeamentos perceptuais em função de seu contexto consonantal.

Os resultados apontaram que os ouvintes peruanos seguiram, de fato, os padrões de assimilação previstos pelos autores, o que vem a corroborar a primeira hipótese. Em relação à segunda hipótese, a qual foi parcialmente corroborada, os resultados apontaram que a categoria francesa /ɛ/ foi majoritariamente assimilada à categoria espanhola /e/ quando presente no contexto /bVg/, enquanto que foi majoritariamente assimilada à categoria espanhola /a/ quando presente nos outros quatro contextos (/bVs/, /bVk/, /sVk/ e /vVst/). Os autores do estudo evidenciam, portanto, que a percepção dos ouvintes peruanos quanto ao contraste /ɛ/-/æ/ pode ser alocada em casos de (1) ASSIMILAÇÃO A UMA CATEGORIA, assim como em casos de (2) ASSIMILAÇÃO A DUAS CATEGORIAS, estando de acordo, conseqüentemente, com as previsões de assimilação estabelecidas pelo PAM-L2.

3.14 Escudero e Williams (2011)

Escudero e Williams (2011) abordaram a percepção das vogais do holandês por falantes nativos de espanhol peruano, com o objetivo de mapear a assimilação perceptual das vogais holandesas às categorias nativas dos peruanos.

Um grupo de 40 falantes monolíngues de espanhol peruano (20 mulheres e 20 homens), com idade média de 24 anos, participaram do estudo. Esse, por sua vez, compreendeu uma tarefa de identificação de 240 produções de vogais holandesas, representando, assim, 20 instâncias de cada um dos segmentos (/i, y, ɪ, ʏ, ø, e, ε, a, ɑ, ɔ, o, u/). Todas as vogais foram produzidas por 20 falantes da variedade padrão e se encontravam em posição interconsonantal (/sVs/), da qual foram retiradas, de modo a evitar possíveis efeitos coarticulatórios. Os participantes foram informados de que eles ouviriam produções de vogais do espanhol articuladas em fala rápida, em vez de vogais não nativas.

As hipóteses para a tarefa de identificação, considerando-se semelhanças acústicas, eram as de que (1) as vogais holandesas /i, y, ε, a, u/ seriam classificadas como as contrapartes espanholas /i, i, a, o, u/; (2) as vogais holandesas /ɪ, ɑ, ɔ/ seriam categorizadas como as respectivas vogais espanholas /i-e/, /a-o/ e /o-u/; (3) as vogais holandesas /ʏ/ e /ø/, embora ocupem posições mais centrais no espaço acústico espanhol, deveriam ser classificadas como /e/ ou /ei-eu/; e (4) as vogais holandesas /e/ e /o/ deveriam ser categorizadas como os ditongos espanhóis /ei/ ou /ou/. Para a realização da tarefa, os 240 estímulos foram apresentados aleatoriamente em um microcomputador e os participantes deveriam selecionar a vogal pronunciada com base no sistema grafêmico espanhol (contendo 5 vogais e 4 ditongos).

As hipóteses do estudo são corroboradas a partir de seus resultados e, embora as previsões de assimilação sejam, por excelência, reflexo de tarefas de discriminação (o que não foi o caso desta pesquisa), os autores fazem menção a tais padrões de assimilação na medida em que se valem de uma matriz de confusão, em que podemos observar o direcionamento das categorizações. Nesse sentido, os autores, que têm familiaridade com o construto do *PAM-L2*, concluem que a percepção dos sons testados apresentou (1) diversos casos de ASSIMILAÇÃO A UMA CATEGORIA, assim como (2) alguns casos de ASSIMILAÇÃO A DUAS CATEGORIAS, satisfazendo, portanto, parte das previsões de assimilação designadas pelo *PAM-L2*.

3.1.5 Strange, Hisagi, Akahane-Yamada e Kubo (2011)

Strange *ET AL.* (2011) trataram da percepção das vogais do inglês americano por falantes nativos de japonês, com o intuito de compreender a assimilação perceptual das vogais do inglês às categorias nativas dos japoneses.

Participaram do estudo 21 falantes nativos de japonês (11 mulheres e 10 homens), com idade média de 22 anos, os quais tinham pouca exposição a situações interacionais em inglês e poucas oportunidades de instrução formal do idioma. Os participantes foram testados em duas tarefas, uma de discriminação categórica e outra de identificação (chamada pelos autores de tarefa de assimilação perceptual), cujos estímulos eram palavras dissilábicas, gravadas por um falante nativo de inglês americano, estruturadas na sequência /hVbə/ e com tonicidade na primeira sílaba.

A tarefa de discriminação categórica seguiu o padrão AXB e objetivou contrastar 13 pares de vogais: 4 pares de controle ([i:/ɛ, u:/ʌ, æ:/ɪ, ɑ:/ʊ]) e 9 pares experimentais ([i:/ɪ, u:/ʊ, ɪ/ɛ, ʊ/ʌ, ɛ/ɑ:, ʌ/ɑ:, ɪ/ʊ, ɛ/ʌ, æ:/ɑ:]). As três palavras de cada tríade foram organizadas em 4 combinações e resultaram em 12 estímulos para cada contraste, sendo que havia 7 blocos de 26 estímulos, totalizando 182 estímulos. Vale lembrar que os pares de controle foram usados como instâncias de possível assimilação a duas categorias (dois sons da língua alvo sendo assimilados em duas categorias distintas da L₁), enquanto que os pares experimentais foram usados como instâncias de possível assimilação a uma categoria (dois sons da língua alvo sendo assimilados em apenas uma categoria da L₁).

Na tarefa de assimilação perceptual, os participantes ouviram 24 estímulos (3 instâncias multiplicadas por 8 vogais) organizados em 6 blocos, e escolhiam, dentre 16 opções [hV(V)], a resposta que lhes parecia mais perceptualmente adequada segundo o seu inventário materno – após a escolha da alternativa, o estímulo era apresentado novamente, seguido de uma escala de 9 pontos, em que deveriam classificar o quão próxima era a vogal que eles ouviam daquela existente em japonês (1 = estrangeira; 9 = japonesa). Apesar de os autores não terem estabelecido uma hipótese a ser testada, podemos depreender uma questão norteadora em relação à pesquisa, que procurou explorar como ocorreram as categorizações dos sons linguísticos não nativos em termos do sistema fônico materno dos japoneses.

Os resultados da tarefa de discriminação apontaram nitidamente a (1) ASSIMILAÇÃO A DUAS CATEGORIAS, no caso dos itens de controle, e diversas sobreposições, no caso dos itens experimentais, revelando (2) ASSIMILAÇÃO A UMA CATEGORIA. Quanto à tarefa de identificação, os resultados são bastante difusos, mas os autores sugerem que as vogais mais consistentemente assimiladas às categorias japonesas foram [ɪ, ɑ:, ʌ]. De maneira geral, segundo os autores, os resultados do estudo demonstram que algumas vogais americanas são muito difíceis, em termos de diferenciação

perceptual, para os ouvintes japoneses, mesmo quando os estímulos são simples e a tarefa é relativamente fácil.

3.1.6 Antoniou, Best e Tyler (2012)

Antoniou, Best e Tyler (2012) pesquisaram os efeitos do modo linguístico (monolíngue vs. bilíngue) na percepção de distinções de vozeamento entre oclusivas por falantes monolíngues e bilíngues de grego e inglês.

A hipótese do estudo era a de que, se os falantes bilíngues desenvolveram categorias fonéticas específicas de língua para a distinção de vozeamento em oclusivas, então devem ser observados efeitos do modo linguístico em tarefas que requeiram julgamentos fonológicos acerca dos sons ouvidos. Dessa forma, os autores conduziram experimentos que objetivavam examinar como o modo linguístico impacta a categorização e os índices de adequação a categorias de oclusivas distintas quanto ao vozeamento, bem como a discriminação de contrastes referentes a tais padrões de sonoridade.

Participaram da pesquisa 20 bilíngues inglês-grego (idade média de 26 anos, 10 homens e 10 mulheres), 20 bilíngues grego-inglês (idade média de 25 anos, 10 homens e 10 mulheres), sendo que ambos os grupos eram provenientes da comunidade grecoaustraliana de Sidnei, 20 monolíngues australianos (idade média de 23 anos, 10 homens e 10 mulheres) e 20 monolíngues gregos (idade média de 26 anos, 10 homens e 10 mulheres), residentes em Atenas.

Dois experimentos foram conduzidos, ambos consistindo em uma tarefa de categorização e em uma tarefa de discriminação categórica. O primeiro experimento tinha como intuito verificar a percepção do contraste de vozeamento em início de sílaba e contou com os estímulos /pa, ta, ba, da/, os quais foram gravados por falantes nativos de grego e de inglês. Na tarefa de categorização, que contou com 64 estímulos, os participantes deveriam ouvir as sílabas isoladamente e selecionar a opção, em registro grafêmico, que correspondesse ao som ouvido (no caso, a respectiva oclusiva). Imediatamente após a escolha, o estímulo era novamente apresentado e o participante deveria classificar, em uma escala de 7 pontos (1 = muito estranho; 7 = excelente), o quão prototípico era o som ouvido. Na tarefa de discriminação categórica, que seguiu o formato AXB, os participantes deveriam ouvir as tríades e indicar se o estímulo medial era igual ao estímulo inicial ou igual ao estímulo final – nesta tarefa, foram utilizadas 128 tríades com estímulos contrastantes.

O segundo experimento tinha como finalidade investigar a percepção do contraste de vozeamento em posição intervocálica, ou seja, em sílabas mediais, e contou com os estímulos /a'pa, a'ta, a'ba, a'da/, os quais também foram gravados por falantes nativos de grego e de inglês. O *DESIGN* do segundo experimento, assim como seus procedimentos metodológicos, seguiram a mesma configuração do primeiro experimento, variando, apenas, no que concerne aos estímulos empregados (sílabas do tipo /vcv/).

Os resultados provenientes das tarefas de identificação indicam que os bilíngues não sobrepõem categorias fonológicas, sendo acurados quanto à percepção dos contrastes apresentados, evidenciando a alta capacidade de os bilíngues atentarem para informações fonológicas abstratas de alta ordem. Esta constatação resta óbvia a partir de suas categorizações, já que os participantes elegeram fones gregos e ingleses como cognatos fonológicos [*VOT* zero do inglês (/b/) paralelo ao *VOT* negativo do grego (/b/) e *VOT* positivo do inglês (/p/) paralelo ao *VOT* zero do grego (/p/)].

Segundo os pesquisadores, tais observações são compatíveis com a hipótese do *PAM-L2* de que os bilíngues desenvolveram categorias fonéticas separadas para o inglês (*L2*), e que estas estão ligadas às categorias fonéticas da *L1* através de um nível mais abstrato, o das unidades fonológicas /p, t, b, d/. No entanto, o fato de serem bilíngues não os diferiu de ambos os grupos de monolíngues nas tarefas de discriminação, sugerindo que o bilinguismo, nesse caso, não teve um papel fundamental para o estabelecimento dos contrastes de vozeamento.

Em linhas gerais, os níveis de acuidade na identificação dos contrastes por parte dos monolíngues ficaram abaixo dos níveis de acuidade dos bilíngues, refletindo sobreposições de categorias fonéticas e fonológicas. Conforme apontam os autores, os bilíngues, pelo menos no que tange à categorização dos segmentos, apresentam, segundo o *PAM-L2*, o padrão (1) ASSIMILAÇÃO A DUAS CATEGORIAS, e os monolíngues oscilam entre os padrões referentes à (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA e (3) ASSIMILAÇÃO A UMA CATEGORIA. Como conclusão, Antoniou, Best e Tyler (2012) destacam que, no que se refere aos bilíngues, tais achados servem de evidência para a integração de ambos os sistemas linguísticos em um espaço fonético comum, influenciado pelo ambiente da sua língua dominante.

3.1.7 Bohn e Best (2012)

Bohn e Best (2012) investigaram a percepção das consoantes aproximantes do inglês americano por falantes nativos de dinamarquês e de alemão. O principal objetivo do estudo era o de

elucidar com mais propriedade as relações perceptuais entre restrições fonológicas e influências fonéticas nativas no que concerne às aproximantes inglesas.

A pesquisa contou com três experimentos, compostos por uma tarefa de identificação perceptual e uma tarefa de discriminação categórica, idênticos quanto ao seu conteúdo metodológico, mas distintos quanto às amostras testadas. Os estímulos utilizados foram as aproximantes /r, l, w, j/ presentes no inglês americano, inseridas no contexto /Aak/, gerando as sequências /rak/, /lak/, /wak/ e /jak/, as quais variaram em um contínuo acústico de 10 opções (que correspondiam a alterações formânticas do segmento inicial da sílaba em questão, a fim de gerar impressões acústicas referentes a gestos articulatorios distintos, mas dentro da categoria das aproximantes).

Na tarefa de identificação, composta por 200 estímulos, cada qual (e seu respectivo contínuo) era apresentado auditivamente aos participantes, que deveriam eleger sua resposta com base em duas opções. Na tarefa de discriminação, que seguiu o padrão AXB e foi composta por 140 tríades, os autores testaram três contrastes, a saber, /rak-lak/, /wak-jak/ e /wak-rak/ (e seus respectivos contínuos).

O primeiro experimento foi aplicado a 18 falantes monolíngues de dinamarquês (3 homens e 15 mulheres), com idade média de 23 anos, e tinha como hipótese a proposição de que os índices de acuidade na percepção auditiva dos estímulos seriam inferiores aos índices apresentados por falantes norte-americanos cuja L₁ era o inglês, uma vez consideradas as diferenças referentes à quantidade de consoantes aproximantes existentes em ambas as línguas. Os resultados obtidos no primeiro experimento não corroboraram a hipótese prevista, já que os participantes dinamarqueses não diferiram dos americanos quanto à identificação e à discriminação das sequências testadas.

O segundo experimento objetivou comparar a identificação e a discriminação dos contrastes mencionados entre dinamarqueses não experientes em inglês (L_E), ou seja, sem conhecimento funcional do idioma, e dinamarqueses experientes em inglês (L_E), isto é, competentes e fluentes, a fim de verificar o papel da experiência linguística na apreensão do contraste entre as aproximantes, especialmente a distinção entre /w/ e /r/. A hipótese era a de que falantes competentes e fluentes teriam maior acuidade na percepção dos contrastes. Para tanto, 15 participantes (9 homens e 6 mulheres), com idade média de 27 anos, compuseram o grupo dos falantes não experientes, e outros 15 participantes (5 homens e 10 mulheres), com idade média de 24 anos, foram testados. Os resultados corroboraram parcialmente a hipótese prevista, sendo que apenas para a percepção do

contraste /w-r/ houve diferenças significativas entre os grupos, sugerindo a influência da experiência linguística na diferenciação das sequências apresentadas.

O terceiro experimento contou com 18 falantes monolíngues de alemão (8 homens e 10 mulheres), com idade média de 21 anos, e também tinha como hipótese a proposição de que os índices de acuidade na percepção auditiva das sequências seriam inferiores aos índices apresentados por falantes norte-americanos. Contrariamente à hipótese, os resultados não corroboraram a previsão, pois os participantes alemães tampouco diferiram dos americanos quanto à percepção das sequências testadas. Comparando-se os participantes monolíngues alemães aos dinamarqueses, não foram observadas diferenças estatisticamente significativas no que compete à percepção dos contrastes estimados. Interessantemente, em diversos casos, os julgamentos em relação aos contrastes apresentavam maior acuidade por falantes não nativos de inglês do que pelos próprios norte-americanos.

Em resumo, o desempenho dos falantes dinamarqueses no que se refere à identificação de /w-r/ e /r-l/ foi muito próximo ao dos norte-americanos, e seu nível de discriminação de /w-j/ foi ainda mais alto se comparados aos falantes monolíngues de inglês (independentemente de sua experiência linguística com o inglês). No caso dos alemães, a identificação dos contrastes foi praticamente categórica, mas a discriminação dos contrastes /w-r/ e /r-l/ foi descritivamente menor quando tais participantes são comparados aos norte-americanos e aos dinamarqueses. Em termos do *PAM-L2*, os autores organizaram o mapeamento da percepção dos grupos testados como sendo, basicamente, casos de ASSIMILAÇÃO A DUAS CATEGORIAS, o que implica altos níveis de acuidade perceptual dos contrastes testados.

3.18 Darcy e Krüger (2012)

Darcy e Krüger (2012) analisaram como crianças turcas percebiam auditivamente as vogais do alemão⁶⁴. A motivação para realizar o estudo com crianças se deu em função de que, segundo os autores, a influência da *L1* sobre a aquisição fonológica de outro idioma parece ser reduzida, ou menos impactante, em crianças.

⁶⁴ Neste estudo, também foi aplicada uma tarefa de produção. No entanto, por não fazerem parte de nosso escopo, referências aos dados de produção serão omitidas.

Participaram da pesquisa 28 crianças (8 meninos e 20 meninas), com idade média de 11 anos, os quais foram divididos em dois grupos: grupo experimental, formado por 14 bilíngues falantes de turco e alemão (5 meninos e 9 meninas), e grupo-controle, constituído por 14 monolíngues falantes de alemão (3 meninos e 11 meninas). Todos os participantes bilíngues pertenciam a uma comunidade turca na área de Berlin e estavam expostos ao alemão por um período médio de 3 anos.

Os estímulos auditivos utilizados foram 10 sílabas do tipo *cvc* variáveis quanto à vogal nuclear, a saber, [pa:p], [ka:k], [pi:p], [ki:k], [pɪp], [kɪk], [pe:p], [ke:k], [pɛp], [kɛk], inicialmente alocadas em sentenças veículos e, então, segmentadas. O teste de percepção consistiu de uma tarefa de discriminação categórica, composta por 48 estímulos, cujo objetivo era testar os contrastes [a:-i:], [i:-ɪ], [e:-ɛ] e [i:-e:], em que os participantes deveriam ouvir a tríade apresentada e indicar qual era a posição em que se encontrava o elemento destoante (*ODDITY DISCRIMINATION TASK*), clicando sobre uma de três imagens na tela do computador.

Levando em consideração estudos prévios, os autores elaboraram quatro hipóteses, cada qual referente a um contraste em alemão a ser mapeado pelos participantes turcos. Para o contraste [i:-e:], os pesquisadores previram dificuldades perceptuais, gerando um caso de *ASSIMILAÇÃO A UMA CATEGORIA*. Para o contraste [i:-ɪ], os autores previram alguma dificuldade perceptual, mas não muito acentuada, envolvendo um caso de *DIFERENÇA DE ADEQUAÇÃO À CATEGORIA*. Para o contraste [e:-ɛ], os estudiosos previram não haver dificuldades perceptuais, possivelmente implicando uma *ASSIMILAÇÃO A DUAS CATEGORIAS*. Finalmente, para o contraste [a:-i:], que funcionou como controle, Darcy e Krüger (2012) previram *total ASSIMILAÇÃO A DUAS CATEGORIAS*. Além disso, previu-se também que haveria diferenças entre falantes monolíngues de alemão e bilíngues falantes de turco e alemão quanto à percepção dos contrastes.

Os resultados obtidos foram ao encontro das hipóteses levantadas, corroborando-as, sendo que, para os participantes bilíngues, o contraste perceptualmente mais difícil foi [i:-e:], seguido de [i:-ɪ]. Por sua vez, os contrastes [a:-i:] e [e:-ɛ] foram mantidos em altos níveis de acuidade. Ademais, os autores mencionam que, muito embora os participantes bilíngues tenham sido expostos ao alemão desde bastante novos, eles diferiram quanto à percepção dos contrastes testados se comparados aos participantes monolíngues, indicando possível influência do sistema fonológico materno sobre a apropriação do sistema alvo.

3.1.9 Antoniou, Best e Tyler (2013)

Antoniou, Best e Tyler (2013) examinaram a percepção sonora das oclusivas do ma'di por falantes monolíngues e bilíngues de grego e inglês. Com base no princípio de que o bilinguismo pode vir a aumentar ou inibir o desempenho perceptual de sons da fala, a depender do contraste não nativo e do modo linguístico do bilíngue (no caso do presente estudo, inglês-grego ou grego-inglês), os autores conduziram experimentos que visavam a investigar como o modo linguístico impacta a categorização e os índices de adequação a categorias, bem como a discriminação de contrastes não nativos referentes a padrões de vozeamento.

No estudo, comparou-se o desempenho perceptual entre monolíngues, entre bilíngues e entre monolíngues e bilíngues. Participaram da pesquisa 20 bilíngues inglês-grego (10 homens e 10 mulheres), 20 bilíngues grego-inglês (10 homens e 10 mulheres), sendo que ambos os grupos eram provenientes da comunidade grecoaustraliana de Sidnei, 25 monolíngues australianos (13 homens e 12 mulheres) e 20 monolíngues gregos (10 homens e 10 mulheres), residentes em Atenas.

Os estímulos utilizados na pesquisa foram pseudopalavras do ma'di, a saber, /ba, pa, ^mba, ɸa, da, ta, ⁿda, d'a/, as quais foram produzidas em uma frase veículo e, posteriormente, extraídas através de ferramental acústico. A testagem consistiu em (1) uma tarefa de identificação, em que os segmentos bilabiais /ba, pa, ^mba, ɸa/ e coronais /da, ta, ⁿda, d'a/ foram apresentados em posição inicial da sílaba /Ca/, e em (2) uma tarefa de discriminação, em que três tipos de contrastes em início de sílaba foram utilizados: oclusivas pré-vozeadas /ba, da/ vs. oclusivas não vozeadas e não aspiradas /pa, ta/; oclusivas pré-vozeadas /ba, da/ vs. oclusivas pré-nasalizadas /^mba, ⁿda/; e oclusivas pré-vozeadas /ba, da/ vs. oclusivas implosivas /ɸa, d'a/.

Para a realização da tarefa de identificação, os participantes ouviam o estímulo e escolhiam, dentre 8 rótulos, qual representava o segmento ouvido. Quanto à condução da tarefa de discriminação, a qual seguiu o padrão AXB, os participantes ouviram 96 estímulos (16 tríades para 6 contrastes) e deveriam indicar se o segmento consonantal da sílaba medial era igual ao segmento consonantal da sílaba inicial ou ao da sílaba final.

As hipóteses gerais da pesquisa se dividem de acordo com a amostra de que dispõem os autores, ligando-se aos monolíngues e aos bilíngues. Dado o inventário consonantal dos monolíngues australianos, esperava-se que (1) o contraste /ba, da/-/pa, ta/ fosse o de DIFERENÇA DE ADEQUAÇÃO À

CATEGORIA; (2) o contraste /ba, da/-/ᵐba, ᵐda/ fosse o de ASSIMILAÇÃO A DUAS CATEGORIAS⁶⁵; e (3) o contraste /ba, da/-/ᶠa, ᶠa/ fosse o de ASSIMILAÇÃO A UMA CATEGORIA. Considerando-se o inventário consonantal dos monolíngues gregos, esperava-se que (1) o contraste /ba, da/-/pa, ta/ fosse o de ASSIMILAÇÃO A DUAS CATEGORIAS; (2) o contraste /ba, da/-/ᵐba, ᵐda/ fosse o de ASSIMILAÇÃO A UMA CATEGORIA; e (3) o contraste /ba, da/-/ᶠa, ᶠa/ fosse o de ASSIMILAÇÃO A UMA CATEGORIA. No que se refere aos bilíngues, levando em conta a sua experiência linguística, os autores previram quatro possibilidades de desempenho, segundo as quais o julgamento dos bilíngues em relação à percepção dos sons em jogo poderia refletir (1) uma fusão do conhecimento fonológico de ambas as línguas; (2) a persistência da L₁; (3) a sobreposição da língua dominante; e (4) uma separação do conhecimento fonológico de ambas as línguas.

De maneira geral, os resultados, tanto em termos de identificação quanto em termos de discriminação, apontaram, haver para os participantes, diversos padrões de assimilação, visíveis através de previsões como (1) ASSIMILAÇÃO A DUAS CATEGORIAS, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, (3) AMBOS NÃO CATEGORIZÁVEIS e (4) ASSIMILAÇÃO A UMA CATEGORIA. Ademais, os monolíngues diferiram quanto à percepção dos segmentos testados em função de seu inventário nativo, sendo que cada qual privilegiou diferentes aspectos em relação aos estímulos ouvidos; bilíngues não diferiram quanto à percepção dos segmentos testados; e bilíngues e monolíngues obtiveram índices de acuidade diferentes a depender de sua experiência linguística.

3.1.10 McKelvie-Sebileau e Davis (2014)

McKelvie-Sebileau e Davis (2014) pesquisaram a discriminação de contrastes de *VOT* do tailandês por parte de monolíngues falantes de inglês e de bilíngues falantes de inglês e francês⁶⁶. O tailandês é uma língua que contrasta três níveis de *VOT*, havendo, portanto, distinção entre oclusivas com *VOT* negativo ([ba]), oclusivas com *VOT* zero ([pa]) e oclusivas com *VOT* positivo ([p^ha]). Falantes nativos de inglês tendem a diferenciar *VOT* zero de *VOT* positivo, enquanto que falantes nativos de

⁶⁵ Os autores registram tal contraste como sendo do tipo *UC*, *UNCATEGORIZED-CATEGORIZED*, utilizado em referência ao *PAM* (Best, 1995), por se tratar de uma amostra cujos participantes não têm experiência com a L₂. No *PAM-L2* (Best; Tyler, 2007), conforme explicado no capítulo anterior, o contraste *UC* passa a ser alocado sob o mesmo rótulo do *TC*, *TWO-CATEGORY ASSIMILATION*, por isso fazemos menção a ele.

⁶⁶ O estudo também contou com tarefas de produção. Contudo, por não ser o foco da presente revisão de literatura, referências aos dados de produção serão omitidas.

francês tendem a contrastar *VOT* negativo de *VOT* zero, daí o triplo contraste em tailandês ser interessante para as duas línguas em questão no estudo.

Participaram da pesquisa 43 falantes monolíngues de inglês (14 homens e 29 mulheres), 25 falantes bilíngues de inglês e francês (7 homens e 18 mulheres) e 28 falantes bilíngues de francês e inglês (16 homens e 12 mulheres). A tarefa de discriminação realizada seguiu o padrão AX, contou com 100 pares envolvendo os contrastes [pa-ba], [pa-p^ha] e [ba-p^ha], e as respostas fornecidas poderiam ser apenas “igual” ou “diferente”, a depender se ambos os estímulos apresentados eram idênticos ou distintos.

Foram utilizados, na tarefa de discriminação, 100 pares de sílabas (sendo que 60 deles apresentavam contraste de *VOT* e 40 não), e os participantes não foram informados acerca de qual língua estava em teste. A hipótese do estudo era a de que a experiência com a L₂ facilitaria a discriminação de seus contrastes funcionais; portanto, os participantes poderiam, em princípio, discriminar entre os três padrões de *VOT* com que teriam contato.

Os resultados sugerem o contraste [pa-ba] como sendo o mais difícil de discriminar por todos os grupos, seguido do contraste [pa-p^ha] e do contraste [ba-p^ha]. Além disso, falantes bilíngues de inglês e francês e falantes monolíngues de inglês não diferiram quanto à discriminação dos três contrastes, indicando que, nesse caso, a experiência linguística não teve papel relevante para a percepção de um padrão negativo de *VOT* (o que seria minimamente esperado, uma vez que a L₂ dos falantes bilíngues de inglês é o francês, a qual apresenta *VOT* negativo). Os falantes bilíngues de francês e inglês tiveram melhor desempenho do que os outros dois grupos quando apresentados ao contraste [pa-ba], indicando a transferência do conhecimento fonológico da L₁ para a língua alvo. Adicionalmente, os autores acomodam seus achados às previsões desenvolvidas no PAM-L₂ e ilustram os casos de (1) ASSIMILAÇÃO A UMA CATEGORIA, (2) ASSIMILAÇÃO A DUAS CATEGORIAS e, também, (3) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA.

3.1.11 Tyler, Best, Faber e Levitt (2014)

O estudo de Tyler *ET AL.* (2014) se voltou para a assimilação perceptual e a discriminação de contrastes fônicos não nativos. Os autores examinaram a percepção auditiva de falantes monolíngues

de inglês americano quanto à distinção de segmentos vocálicos pertencentes às línguas norueguesa, tailandesa e francesa, a fim de verificar as categorizações previstas pelo *PAM*⁶⁷.

Os autores tentaram responder a duas questões de pesquisa. A primeira delas indagava de que maneira os ouvintes americanos passam a categorizar os contrastes vocálicos não nativos, e a segunda propunha verificar em que medida as categorizações estabelecidas pelos participantes se relacionavam com as previsões assimilatórias do modelo de Best (1995).

Participaram da pesquisa 13 estudantes universitários, falantes nativos de inglês americano, que desconheciam idiomas cujos sistemas sonoros contivessem os contrastes vocálicos testados. Os estímulos auditivos utilizados compreenderam 2 contrastes vocálicos em norueguês, /ki/-/ky/ e /ki/-/kɥ/, 1 contraste vocálico em tailandês, /bu/-/bɤ/, e 3 contrastes vocálicos em francês, /dø/-/dœ/, /sy/-/sø/ e /bo/-/bõ/, respectivamente. Tais sequências foram organizadas de modo a satisfazer o padrão *AXB* de uma das tarefas perceptuais empregadas na pesquisa, em que os estímulos no formato *AAB* e *ABB* partiam da vogal mais periférica para a menos periférica, e os estímulos no formato *BBA* e *BAA* partiam da vogal menos periférica para a mais periférica⁶⁸.

Duas foram as tarefas aplicadas aos participantes, uma de discriminação categórica e outra de identificação perceptual. A primeira era composta por 288 tríades, das quais 48 representavam o(s) contraste(s) vocálico(s) a serem testados, apresentadas aleatoriamente aos participantes. A segunda consistia em 240 estímulos, em que cada sílaba a ser testada era exibida aleatoriamente 20 vezes (sendo 18 estímulos computados e 2 estímulos para a familiarização dos participantes com a tarefa). Os participantes ouviam o mesmo estímulo duas vezes: na primeira, ouviam e circulavam uma palavra-chave inglesa correspondente à vogal ouvida (no caso das vogais orais, as palavras eram ‘HEED’, ‘HID’, ‘AID’, ‘ED’, ‘AD’, ‘ODD’, ‘AWED’, ‘HOED’, ‘HOOD’, ‘DUD’, ‘FOOD’, e ‘HEARD’, e, no caso das vogais nasalizadas, as palavras eram ‘END’, ‘DONNED’ e ‘OWNED’) e, na segunda, avaliavam a similaridade da vogal com a da palavra-chave (1 = diferente; 5 = idêntica).

No que concerne aos resultados da tarefa de identificação, os autores mencionam que houve altos níveis de variabilidade interindividual quanto aos padrões de assimilação das vogais não nativas

⁶⁷ Os autores também procuraram testar o modelo de Polka e Bohn (2003, 2011), conhecido como *MODELO DA VOGAL REFERENTE NATURAL (NATURAL REFERENT VOWEL FRAMEWORK)*. Não pretendemos trazer à tona essa discussão por questões de delimitação, mas, basicamente, o modelo prediz que as vogais mais periféricas agem como âncoras perceptuais para guiar o desenvolvimento das categorias vocálicas nativas.

⁶⁸ Em norueguês, a vogal em /ki/ é mais periférica do que as vogais em /ky/ e /kɥ/; em tailandês, a vogal em /bu/ é mais periférica do que a vogal em /bɤ/; e em francês, a vogal em /dø/ é mais periférica do que a vogal em /dœ/, a vogal em /sy/ é mais periférica do que a vogal em /sø/, e a vogal em /bo/ é mais periférica do que a vogal em /bõ/.

e relatam que a maioria dos segmentos foi assimilada como (1) ASSIMILAÇÃO A DUAS CATEGORIAS, (2) NÃO CATEGORIZADO VS. CATEGORIZADO, ou (3) AMBOS NÃO CATEGORIZÁVEIS. O padrão ASSIMILAÇÃO A UMA CATEGORIA foi encontrado somente para as vogais das sílabas /ki/ e /ky/ do norueguês. Em relação aos resultados da tarefa de discriminação, os pesquisadores afirmam que todos os contrastes, com exceção do par /ki/-/ky/, foram discriminados com altíssimos índices de acuidade. Além disso, os padrões que se aplicaram aos resultados provenientes da tarefa de identificação foram retomados para os resultados advindos da tarefa de discriminação.

3.2 Artigos científicos sem enfoque nas previsões assimilatórias do *PAM-L2*

Diferentemente dos artigos científicos apresentados na seção anterior, os estudos a serem elencados nesta seção não fazem referência explícita às previsões de assimilação⁶⁹ designadas no *PAM-L2*, e também serão apresentados individualmente. Cabe retomar que os parâmetros de interpretação dos dados que utilizaremos para atribuir previsões de assimilação específicas estão relacionados aos índices de acuidade ou aos escores obtidos pelos participantes nos experimentos empregados pelos pesquisadores.

3.2.1 Best e Hallé (2010)

Best e Hallé (2010) investigaram como falantes nativos de inglês americano e de francês europeu percebiam, tanto no que se refere à discriminação categórica como no que tange à identificação perceptual, o contraste de vozeamento em obstruintes iniciais existente em hebraico, /dla/-/tla/, zulu, /ɬa/-/tʰa/, e tlingit, /dʒa/-/tʰa/.

Diferentemente de outros estudos de Catherine Best e colaboradores, a linha de raciocínio que Best e Hallé (2010) seguem neste artigo remete a uma comparação entre uma fonologia com base em traços distintivos (aos moldes da geometria de traços) e uma fonologia com base em gestos articulatórios para o tratamento dado ao fenômeno fônico em questão. Assim, o trabalho central dos pesquisadores não é o de estabelecer hipóteses experimentais e testá-las estatisticamente, mas, sim, o de apresentar ao leitor evidências que se inclinam para uma perspectiva gestual. Todavia, testes

⁶⁹ Embora não seja feita referência explícita às previsões de assimilação do *PAM-L2*, os trabalhos elencados mencionam, explicitamente, aspectos teóricos de tal modelo.

estatísticos comparativos foram realizados, e é sobre eles que nos voltaremos para tecer considerações acerca das categorizações sonoras que podem decorrer dos resultados obtidos.

Atuaram como participantes da pesquisa 35 estudantes universitários, divididos em dois grupos de acordo com sua L1. Desse modo, um grupo foi composto por 19 falantes nativos de inglês americano e outro grupo contou com 16 falantes nativos de francês europeu. Segundo os autores, nenhum participante reportou ter tido contato prévio com as línguas testadas. Em termos de experimentação, os acadêmicos foram convidados a realizar duas tarefas de percepção auditiva, uma de discriminação categórica e uma de identificação perceptual, sendo que ambas apresentavam como estímulos as sílabas /dlɑ/-/tla/, /ʒɑ/-/ʔɑ/ e /dʒɑ/-/tʃɑ/, respectivamente pertencentes ao hebraico, ao zulu e ao tlingit, gravados por falantes nativos de tais idiomas e do gênero masculino. A tarefa de discriminação categórica (formato AXB) foi composta de 180 tríades, sendo 60 delas para cada contraste de vozeamento, e a tarefa de identificação perceptual continha 90 estímulos, sendo 30 deles referentes a cada língua-alvo, em que os participantes ouviam cada estímulo e o transcreviam utilizando o teclado do computador.

Os resultados encontrados a partir da tarefa de discriminação categórica indicaram não haver diferenças estatisticamente significativas entre os ouvintes americanos e os ouvintes franceses quanto à acuidade na percepção do contraste de vozeamento das línguas hebraica e zulu, sendo que ambos os grupos obtiveram altos níveis de acurácia ao desempenhar a tarefa. Por outro lado, os ouvintes americanos tiveram índices significativamente mais elevados do que os franceses na percepção do contraste de vozeamento existente em tlingit, exibindo vantagem em relação aos participantes francófonos.

Segundo os autores, haveria duas razões plausíveis para dar conta de explicar os maiores índices de acuidade dos ouvintes americanos: a primeira corresponde ao fato de as africadas serem estrangeiras ao sistema fônico do francês e apresentarem uma organização gestual incomum se vislumbrada frente às demais constelações de gestos que operam na fonologia francesa, o que importaria dificuldades perceptuais aos participantes francófonos; a segunda tem a ver com a constatação de que o contraste de vozeamento nas africadas do tlingit é foneticamente comparável ao das oclusivas do inglês (incluindo as africadas), mas distinto do contraste de vozeamento nas oclusivas do francês.

Em relação à tarefa de identificação perceptual, as respostas foram, em geral, congruentes com o padrão encontrado na tarefa de discriminação categórica. Nesse sentido, os pesquisadores

fazem três observações. A primeira é a de que as fricativas laterais do zulu foram frequentemente interpretadas pelos participantes como *ONSETS* complexos com um componente lateral (para os ouvintes americanos) ou com um componente fricativo (para os ouvintes franceses); por outro lado, tanto os ouvintes americanos como os ouvintes franceses tenderam a interpretar as africadas do tlingit, mas não os *ONSETS* complexos do hebraico, como oclusivas (e não como *ONSETS* complexos, por exemplo). A segunda tem a ver com o fato de que o ponto coronal das africadas do tlingit e das oclusivas dos *ONSETS* complexos do hebraico foi frequentemente percebido como velar. A terceira, por fim, estabelece que houve uma assimetria de ponto de articulação quanto ao vozeamento dos pares hebraicos, em que as transcrições que apresentaram um segmento não vozeado estavam ligadas ao ponto velar, enquanto que as transcrições que apresentaram um segmento vozeado estavam ligadas ao ponto coronal.

Apesar de o estudo tangenciar a discussão referente às previsões de assimilação difundidas pelo *PAM* (uma vez que o propósito dos autores do artigo é problematizar o fenômeno fônico em questão à luz de uma fonologia baseada em gestos articulatorios em detrimento de uma fonologia baseada em traços distintivos), parece-nos que os resultados obtidos a partir da tarefa de discriminação categórica podem ser tratados sob o expediente dos padrões assimilatórios descritos por Best (1995) e Best e Tyler (2007). Dessa maneira, restam nítidos casos de (1) ASSIMILAÇÃO A DUAS CATEGORIAS, no caso da percepção dos americanos e dos franceses quanto aos contrastes em hebraico e zulu, e somente no caso dos americanos quanto aos contrastes em tlingit; e (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, no caso da percepção dos franceses quanto aos contrastes em tlingit, cujos resultados se mostraram bastante dispersos.

3.2.2 Escudero e Chládková (2010)

Escudero e Chládková, em sua pesquisa de 2010, objetivaram comparar as categorizações sonoras feitas por falantes de espanhol peruano quanto às vogais do inglês americano e do inglês britânico. O objeto de investigação do estudo, portanto, se ocupava da tentativa de constatar quais eram os padrões de assimilação perceptual apresentados pelos ouvintes peruanos.

Duas hipóteses regeram o trabalho. A primeira delas era a de que os participantes iriam perceber as vogais do inglês americano e do inglês britânico como instâncias de vogais nativas se elas apresentassem valores de F_1 e F_2 próximos aos valores das vogais do espanhol peruano. A segunda,

relacionada à primeira, estipulava que a categorização de determinada vogal não nativa ao sistema sonoro do espanhol peruano poderia variar dependendo do dialeto de inglês exibido – a título de ilustração, quando produzido aos moldes da variedade britânica, o segmento /æ/ poderia ser categorizado como o /a/ do espanhol, enquanto que, quando produzido aos moldes da variedade americana, o mesmo segmento poderia ser categorizado como o /e/ do espanhol.

Os participantes de pesquisa foram 40 universitários peruanos (20 homens e 20 mulheres, com idade média de 24 anos), falantes nativos de espanhol e com pouquíssima experiência em relação à língua inglesa. Os acadêmicos foram convidados a realizar uma tarefa de identificação perceptual, a qual contou com 9 vogais comuns aos inventários fônicos das variedades americana e britânica de inglês (/i, ɪ, ε, a, ɑ, ʌ, ɔ, u, u/) e com 5 vogais pertencentes ao espanhol (/i, e, a, o, u/), todas sintetizadas a partir de estímulos naturais. A tarefa implicava que os participantes identificassem 205 estímulos auditivos (10 repetições para as 18 vogais inglesas mais 5 repetições para as 5 vogais espanholas) com base nos segmentos do espanhol, apresentados em forma de grafemas na tela do computador. Cabe salientar que os pesquisadores informaram os participantes de que as vogais a serem ouvidas haviam sido extraídas da fala contínua em espanhol e, por vezes, poderiam parecer exemplares mais desviantes daqueles segmentos que lhes eram familiares, além de instruí-los a escolher uma das 5 categorias apresentadas, mesmo quando não estavam certos de sua resposta.

Os resultados encontrados indicaram que não houve diferença estatisticamente significativa quanto à comparação das vogais /i/ e /u/ nos dois dialetos em jogo, indicando que, independentemente de pertencerem ao inglês americano ou ao inglês britânico, tais vogais foram assimiladas biunivocamente às categorias correspondentes no espanhol. No entanto, para as 7 demais vogais, as categorizações dos ouvintes foram impactadas pelo dialeto a que estavam expostos, revelando a associação de duas, três ou quatro vogais estrangeiras a uma mesma categoria nativa. Por exemplo, ao segmento /ɑ/ do inglês britânico foram associados os segmentos /a/ e /o/ do espanhol, e ao segmento /ɪ/ da mesma variedade foram associados os segmentos /i/, /e/ e /u/ do espanhol. De maneira geral, houve mais categorizações inesperadas para os segmentos referentes à variedade britânica do que para aqueles relativos ao dialeto americano. A interpretação dos resultados evidencia, portanto, a corroboração das duas hipóteses levantadas pelos autores.

Ainda que os pesquisadores não façam referência ao estabelecimento de categorias aos moldes do *PAM-L2*, presumimos que as previsões de tal modelo dão conta de abarcar os resultados encontrados e os padrões emergentes poderiam vir a ser o de (1) ASSIMILAÇÃO A UMA CATEGORIA e (2)

DIFERENÇA DE ADEQUAÇÃO À CATEGORIA para a maioria dos casos, além de (3) ASSIMILAÇÃO A DUAS CATEGORIAS NO caso dos segmentos /i/ e /u/, pois são os únicos que não se mostram associados a quaisquer outras vogais testadas que não as correspondentes em espanhol.

3.2.3 Bundgaard-Nielsen, Best e Tyler (2011a)

Bundgaard-Nielsen, Best e Tyler (2011a) investigaram o impacto do tamanho do vocabulário na percepção de vogais da L₂, partindo do pressuposto de que DETER UM MAIOR VOCABULÁRIO NA L₂ CONDUZ A UM PROCESSO DE REFONOLOGIZAÇÃO⁷⁰ no caso de aprendizes adultos de um outro idioma. O estudo, portanto, examinou as similaridades e as diferenças na percepção de vogais⁷¹ e ditongos do inglês australiano (/i:, I, e, e:, ɜ:, æ, ɐ, ɛ:, o:, ʊ, ɥ:, ɔ, ɪə, æI, ae, oI, əɥ, æɔ/) por dois grupos de ouvintes japoneses, usuários nativos de um sistema vocálico com apenas 5 vogais (/i, e, a, o, u/ e suas contrapartes longas), distintos quanto ao tamanho de seu vocabulário na L₂.

A hipótese que embasou a pesquisa foi a de que os aprendizes japoneses seriam altamente sensíveis a diferenças de duração (dado o sistema moraico do japonês), bem como a algumas distinções espectrais acerca do inventário vocálico do inglês australiano. Especificamente, os autores lançaram mão das seguintes previsões: os monotongos curtos do inglês australiano seriam percebidos como vogais monomoraicas do japonês; os monotongos longos do inglês australiano seriam percebidos como vogais bimoraicas do japonês; e os ditongos do inglês australiano seriam percebidos como combinações de vogais bimoraicas do japonês.

Participaram do estudo 11 falantes nativos de japonês (3 homens e 8 mulheres, com idade média de 26 anos), residentes em Sidnei há menos de 12 semanas, os quais haviam tido instrução formal da língua inglesa no Japão durante 7 anos em média, frequentemente realizada por professores japoneses. Os participantes realizaram uma tarefa de identificação perceptual, a qual consistiu de 108 estímulos auditivos, representados por vogais e ditongos do inglês americano e retirados de palavras produzidas de acordo com a estrutura silábica /hVbɐ/.

⁷⁰ Para os autores, o processo de REFONOLOGIZAÇÃO significa o estabelecimento de uma fonologia da L₂ através da modificação ou adição de elementos ao sistema fonológico da L₁ existente.

⁷¹ Os autores ressaltam que todos os monotongos do inglês australiano se caracterizam por algum grau de mudança espectral durante a sua produção e, normalmente, são distinguíveis com base em F₁ e F₂, exceto pelos pares /i:-I/ e /ɛ:-ɐ/, que se diferenciam, principalmente, em termos de duração.

Basicamente, a tarefa pressupunha que os participantes ouvissem individualmente as vogais e os ditongos do inglês australiano e selecionassem a opção que correspondia àquele(s) segmento(s) segundo um código de símbolos em japonês. Além disso, os participantes deveriam indicar em uma escala o quanto o segmento ouvido se parecia com a contraparte japonesa (1 = exemplar muito desviante; 7 = exemplar excelente). Ao final da testagem, os participantes responderam a uma tarefa de tamanho de vocabulário, composto por múltiplas opções, que, segundo Bundgaard-Nielsen, Best e Tyler (2011a), foi capaz de avaliar a amplitude do seu conhecimento lexical.

De modo geral, considerando-se ambos os grupos do estudo, 13 das 18 vogais inglesas foram assimiladas consistentemente a categorias vocálicas do japonês, a saber, /i:, ɪ, e, e:, ɐ, ɛ:, ʊ, ʉ:, ɔ, ɪə, æɪ, œ, oɪ/. As 5 vogais inglesas remanescentes, /ɜ:, æ, o:, æɔ, əʉ/, foram NÃO CATEGORIZADAS, ou seja, não foram identificadas como qualquer categoria vocálica do japonês. Os resultados encontrados dão suporte ao pressuposto inicial dos autores, o qual se referia ao fato de que um maior vocabulário na L₂ conduziria a um processo de refonologização, uma vez que os aprendizes com mais vocabulário não só foram mais consistentes na identificação dos fones da L₂ em relação aos da L₁, mas também em termos de quantas categorias alternativas da L₁ eles selecionaram para cada vogal da L₂.

Quanto à hipótese levantada pelos autores acerca das diferenças de duração e comportamento espectral das vogais na identificação das categorias vocálicas australianas, os pesquisadores reportam que os aprendizes japoneses utilizaram sistematicamente as categorias maternas referentes a elementos monomoraicos e bimoraicos para identificar vogais curtas e frouxas e vogais longas e tensas do inglês australiano, corroborando a previsão do estudo. Nesse sentido, os autores acrescentam que tais resultados podem indicar também, segundo a literatura, que a informação relativa à duração exerce um papel fundamental na percepção de fones não nativos, especialmente quando as propriedades espectrais de dado segmento podem ser ambíguas ao ouvinte.

Bundgaard-Nielsen, Best e Tyler (2011a) não mencionam os padrões de assimilação encontrados, mesmo se tratando de um estudo que teria potencial para isso. Julgamos que a ausência dessas informações seja proposital, já que os pesquisadores se voltam para outro aspecto basilar do *PAM-L₂*, que é o conhecimento lexical como potencializador dos contrastes não nativos. Assim, os autores reiteram que os achados da pesquisa estão alinhados à predição do *PAM-L₂* de que a aquisição vocabular da L₂ guia o aprendiz em direção à refonologização precoce em ambientes conversacionais de imersão. Não há dúvida, contudo, de que os resultados encontrados podem ser vislumbrados

frente às previsões assimilatórias que o modelo estabelece, e, nesse aspecto, sugerimos que os casos de assimilação em jogo seriam, basicamente, (1) ASSIMILAÇÃO A DUAS CATEGORIAS e (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, a depender da consistência das opções de categorização e da acuidade na identificação dos segmentos testados.

3.2.4 Escudero, Simon e Mitterer (2012)

Escudero, Simon e Mitterer (2012) examinaram o papel da variação dialetal na percepção das vogais do inglês, partindo do pressuposto de que semelhanças acústicas entre segmentos desencadeiam semelhanças perceptuais em relação aos sons da fala. Desse modo, os autores compararam o desempenho entre falantes de holandês pertencentes a duas variedades, setentrional e flamenga (holandeses e belgas, portanto), em tarefas de identificação e discriminação das vogais /ɛ/ e /æ/ presentes em inglês britânico.

Os participantes da pesquisa foram 42 falantes de holandês, divididos em dois grupos: holandeses (4 homens e 17 mulheres), com idade média de 22 anos, e belgas (5 homens e 16 mulheres), com idade média de 21 anos, sendo que os integrantes de ambos os grupos estavam iniciando seu contato formal com a língua inglesa e nunca haviam frequentado ou residido em qualquer país onde se falasse inglês.

O estudo contou com duas tarefas de identificação perceptual, chamadas de tarefas de categorização. Na primeira delas, a de categorização do holandês, os participantes deveriam ouvir 60 pseudopalavras, seguindo a estrutura /hVk/, cujo elemento nuclear variava acusticamente entre /ɛ/ e /æ/, e indicar, dentre as doze opções mostradas (uma para cada grafema holandês), qual vogal do holandês o estímulo acústico representava – os participantes não foram informados acerca do fato de as vogais não serem nativas. Os informantes deveriam, também, indicar em uma escala de 7 pontos, o quão fidedigno era o exemplar ouvido (1 = péssimo exemplar; 7 = ótimo exemplar) em termos de sua L1.

Na segunda tarefa, referente à categorização do inglês, os participantes deveriam ouvir 60 vogais (variando acusticamente entre /ɛ/ e /æ/), retiradas de sílabas do tipo CVC, e indicar qual vogal do inglês o estímulo acústico representava – eles poderiam escolher a vogal desejada com base em palavras reais que serviam de modelo para a categorização, como, por exemplo, a palavra ‘CAT’, contendo o segmento /æ/, a palavra ‘CALL’, contendo o segmento /ɔ:/, a palavra ‘MUG’, contendo a

vogal /ʌ/, etc. Os pesquisadores trataram ambas as tarefas como uma questão exploratória (de modo a verificar quais seriam as opções mais eleitas pelos participantes tarefas), em detrimento de haver propriamente uma hipótese.

Os resultados da primeira tarefa indicam que (1) o segmento /ɛ/ do inglês foi categorizado como o segmento /ɛ/ do holandês, embora esse mapeamento tenha sido mais frequente para os belgas do que para os holandeses, e (2) o segmento /æ/ do inglês foi categorizado como os segmentos /ɛ/ e /ɑ/ do holandês por ambos os grupos.

Em relação à segunda tarefa, os resultados sugerem que os participantes não diferem quanto à classificação do segmento /ɛ/, porém, diferem quanto à classificação do segmento /æ/: os participantes holandeses tenderam a classificar /æ/ como /ɑ/ e os participantes belgas tenderam a classificar /æ/ justamente como /æ/.

Os autores mencionam que, com base nos resultados obtidos, podem-se vislumbrar importantes diferenças dialetais entre as variedades consideradas no que diz respeito à categorização de /ɛ/ e /æ/, refletindo o fato de que diferentes dialetos não só influenciam a percepção de um sistema sonoro, mas também impactam a sua produção. Ainda que os pesquisadores não façam referência ao estabelecimento de categorias aos moldes do *PAM-L2*, entendemos que as previsões de tal modelo dão conta de abarcar os resultados encontrados e os padrões emergentes poderiam vir a ser o de (1) ASSIMILAÇÃO A DUAS CATEGORIAS e o de (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA.

3.2.5 Escudero e Williams (2012)

Escudero e Williams (2012) investigaram a influência de diferenças dialetais na percepção de vogais estrangeiras. Para tanto, os pesquisadores compararam o desempenho entre falantes de espanhol peruano e falantes de espanhol ibérico em tarefas de identificação e discriminação de vogais do holandês. Novamente, para os autores, semelhanças acústicas entre segmentos desencadeiam semelhanças perceptuais em relação aos sons da fala.

Os participantes da pesquisa foram 51 falantes de espanhol peruano e 48 falantes de espanhol ibérico, sendo que os sujeitos de ambos os grupos eram aprendizes de holandês e estavam residindo na Holanda no momento da testagem.

A tarefa de identificação perceptual contou com 240 estímulos correspondentes às 12 vogais do holandês (/ɑ, a, ɛ, e, ø, i, í, o, u, o, y, y/), retirados de sílabas de estrutura /sVs/, em que os

participantes deveriam ouvir e escolher um grafema ou uma dupla de grafemas (<a, aa, e, ee, eu, i, ie, o, oe, oo, u, uu>) que representasse tais segmentos.

A tarefa de discriminação categórica compreendeu 200 estímulos auditivos, contou apenas com cinco contrastes vocálicos, a saber, /a-ɑ/, /i-ɪ/, /y-ʏ/, /i-y/ e /ɪ-y/, e seguiu o formato X_{AB} . Interessantemente, no que tange à tarefa de discriminação, enquanto o elemento x era um estímulo natural, os elementos A e B eram estímulos sintéticos com base em amostras sonoras reais.

Para ambas as tarefas, tinha-se como hipótese a proposição de que, tanto para os falantes de espanhol peruano como para os falantes de espanhol ibérico, a percepção das vogais holandesas acusticamente dissimilares às do espanhol seria muito semelhante entre ambos os dialetos.

Os resultados da tarefa de identificação apontaram que os falantes de espanhol ibérico tiveram maiores índices de acuidade na classificação das vogais holandesas /ɔ, ɑ, a, e, ε, ʏ/ quando comparados aos falantes de espanhol peruano.

Quanto à tarefa de discriminação, os autores mencionam que, de maneira geral, os aprendizes tiveram maior dificuldade perceptual em relação ao contraste holandês /a-ɑ/, seguido pelo par /i-ɪ/, e, mais especificamente, os aprendizes espanhóis tiveram maiores índices de acuidade na discriminação dos pares mais difíceis (/a-ɑ/ e /i-ɪ/, por exemplo) quando comparados aos peruanos.

Comparando o inventário vocálico dos dois dialetos de espanhol em jogo ao inventário vocálico do holandês, os autores concluem que o desempenho mais acurado dos espanhóis sobre os peruanos tem a ver com a proximidade relativa entre as vogais do espanhol ibérico e as vogais do holandês, o que garantiria maiores índices de acuidade perceptual, uma vez que, para Escudero e Williams (2012), semelhanças acústicas se refletem em semelhanças perceptuais.

Muito embora os autores não façam referência ao estabelecimento de categorias aos moldes do *PAM-L2*, entendemos que as previsões de tal modelo dão conta de abarcar os resultados encontrados, e os padrões mais sobressalentes poderiam vir a ser o de (1) ASSIMILAÇÃO A UMA CATEGORIA, o de (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA e o de (3) ASSIMILAÇÃO A DUAS CATEGORIAS.

3.2.6 Fabra e Romero (2012)

Fabra e Romero (2012) analisaram a percepção de vogais do inglês por falantes nativos de catalão⁷². O objetivo do trabalho era o de verificar se diferentes níveis de proficiência em inglês tornariam mais acurada a percepção de contrastes vocálicos estrangeiros.

Participaram da pesquisa 34 falantes nativos de catalão, com idade média de 22 anos, os quais foram divididos em três grupos relativos ao nível de proficiência que apresentaram em língua inglesa. Desse modo, 12 participantes foram alocados no nível básico, 12 no nível intermediário e 10 no nível avançado.

Os estímulos auditivos utilizados consistiam de sílabas do tipo /sV/ (derivadas de sílabas de estrutura /sVt/, modificadas acusticamente para evitar a presença da oclusiva alveolar), em que o elemento vocálico poderia ser uma das sete vogais catalãs, /i, e, ε, a, ə, o, u/, ou uma das seguintes vogais inglesas: /i, ɪ, ε, æ, ʌ, ɑ, ʊ, u/. Os estímulos foram combinados em tríades para que os pesquisadores pudessem examinar a discriminação categórica (AXB) de 11 contrastes vocálicos em duas tarefas de percepção separadas.

No primeiro experimento, o qual contou com 112 estímulos, os autores testaram 7 contrastes que incluíam uma vogal do catalão e uma vogal do inglês na mesma tríade, a saber, /a-æ/, /a-ɑ/, /ε-ε/, /a-ʌ/, /i-i/, /i-ɪ/ e /u-u/, com o objetivo de examinar se os participantes discriminariam elementos nativos de elementos não nativos.

No segundo experimento, composto por 64 estímulos, os pesquisadores testaram 4 contrastes vocálicos apenas do inglês americano, a saber, /ɑ-ʌ/, /ε-æ/, /i-ɪ/ e /u-ʊ/, com o intuito de verificar se os participantes conseguiriam discriminar apenas entre segmentos não nativos. Considerando a quantidade de vogais presentes nos sistemas fonéticos do catalão e do inglês americano, os autores lançam mão da hipótese de que alguns contrastes envolvendo a L₁ e a L_E e alguns contrastes envolvendo somente a L_E seriam de difícil discriminação categórica para os catalães, especialmente para aprendizes com baixa proficiência em inglês.

Os resultados do primeiro experimento revelaram não haver diferenças entre os participantes em relação à discriminação dos pares /a-æ/, /a-ɑ/, /ε-ε/ e /a-ʌ/ (isto é, o nível de proficiência dos catalães não impactou o seu desempenho na tarefa). Porém, foram observadas

⁷² Esta é outra pesquisa que contou com tarefa de produção e, por não se tratar de nosso objeto de investigação, referências a tais dados serão desconsideradas.

diferenças moderadas entre os grupos no que diz respeito aos contrastes /i-i/, /i-I/ e /u-u/, em que os falantes com nível de proficiência mais alto tiveram maior acuidade, em nível descritivo, na discriminação das sequências. Disso decorre o fato de os participantes ainda demonstrarem dificuldades em relação à percepção de sons estrangeiros, uma vez que não os diferenciam, em sua totalidade, dos sons nativos (especialmente nos três casos apontados).

Quanto ao segundo experimento, os resultados evidenciam que os pares /i-I/ e /u-U/ tiveram índices de acuidade mais elevados do que os pares /ɑ-Λ/ e /ɛ-æ/ (ou seja, os participantes conseguem estabelecer mais facilmente as distinções perceptuais entre os pares /i-I/ e /u-U/ se comparados aos pares /ɑ-Λ/ e /ɛ-æ/).

De maneira geral, em ambos os experimentos, não foram encontradas diferenças estatisticamente significativas relativas ao nível de proficiência dos aprendizes, o que sugere que ser pouco ou muito proficiente em inglês, no caso dos catalães pesquisados, não reflete maiores índices de acuidade discriminatória.

Embora os autores façam associações teóricas com o *PAM-L2*, eles não mencionam quais poderiam ser os padrões de assimilação previstos pelo modelo. Logo, defendemos que tais padrões são potencialmente cabíveis ao tratamento dos resultados, sobretudo os casos de (1) ASSIMILAÇÃO A UMA CATEGORIA, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA e, em algumas instâncias, (3) ASSIMILAÇÃO A DUAS CATEGORIAS.

3.2.7 Shafiro, Levy, Khamis-Dakwar e Kharkhurin (2012)

O estudo de Shafiro *ET AL.* (2012) propôs examinar a percepção de vogais e consoantes do inglês americano através de matrizes de confusão, em que as respostas dos participantes são cruzadas de acordo com o estímulo apresentado. Os autores elaboraram uma tarefa de identificação categórica e compararam duas amostras, um grupo de bilíngues falantes de árabe e inglês (cuja *L1* era o árabe) e um grupo de falantes monolíngues dos dialetos ingleses falados nos Emirados Árabes Unidos.

Os objetivos do trabalho, os quais foram, posteriormente, testados como hipóteses, eram os seguintes: (i) verificar os índices de acuidade na identificação dos segmentos; (ii) avaliar a relativa dificuldade perceptual e padrões de erros de classificação quanto aos segmentos; e (iii) examinar os efeitos potenciais de diferentes contextos vocálicos na percepção das consoantes.

Participaram da pesquisa 42 estudantes universitários com idade média de 20 anos, sendo que 25 alunos (20 mulheres e 5 homens) compuseram o grupo cuja *L1* era o árabe e 17 alunos (12

mulheres e 5 homens) formaram o grupo cuja *L1* era o inglês. A tarefa dos participantes era identificar 12 vogais apresentadas no contexto /hVd/ e 20 consoantes em três contextos vocálicos (/iCi/, /aCa/, /uCu/), ambos os grupos de segmentos pertencentes à variedade norte-americana, diferente daquela a que os participantes estão expostos. Conforme apontam os autores, os dialetos ingleses existentes nos Emirados Árabes Unidos possuem características mais similares ao inglês britânico padrão do que ao inglês norte-americano. No que se refere à testagem das vogais, os participantes ouviram 120 estímulos (12 vogais produzidas por 5 homens e por 5 mulheres), e quanto à identificação das consoantes, os participantes ouviram 360 estímulos (20 consoantes produzidas por 1 homem e 2 mulheres, em 3 contextos vocálicos diferentes, apresentados aleatoriamente por 2 vezes).

Os resultados da pesquisa, quanto ao primeiro objetivo proposto, apontaram que os participantes de ambos os grupos conseguem identificar com altos índices de acuidade os segmentos investigados, independentemente de sua *L1*. No entanto, as vogais /ɑ/, /ɔ/ e /æ/ e a consoante /ð/ impuseram maior dificuldade perceptual para ambos os grupos. Em relação ao segundo objetivo, os pesquisadores ilustram que os padrões de erros que se fizeram presentes para a identificação das vogais e das consoantes foram extremamente consistentes entre os dois grupos testados, apesar da variabilidade dialetal referente ao árabe e ao inglês falados pelos participantes. Ainda, a maioria das confusões de vogais e consoantes do estudo deram respaldo à proposição de que aqueles sons que não possuem uma contraparte nas línguas ou dialetos nativos dos participantes acarretaram dificuldades perceptuais. Por fim, no que diz respeito ao terceiro objetivo, os contextos vocálicos /iCi/ e /aCa/ desencadearam confusões perceptuais especialmente sobre a consoante /ð/—enquanto os participantes obtiveram altos índices de acuidade na identificação da fricativa interdental sonora quando essa se apresentava no contexto /uðu/, o mesmo não ocorreu nos contextos /iði/ e /aða/. As confusões mais frequentes foram aquelas que envolveram a substituição de /ð/ por /z/ e /v/ entre as vogais anteriores alta e baixa.

Os resultados encontrados não foram vislumbrados à luz dos padrões assimilatórios do *PAM-L2*, e tampouco foi feito qualquer paralelo entre os achados do estudo e o estabelecimento de categorias segundo o referido modelo. Todavia, defendemos que as previsões do *PAM-L2* dão conta de abarcar tais resultados e, nesse aspecto, os padrões que possivelmente emergiriam seriam: (1) ASSIMILAÇÃO A DUAS CATEGORIAS na maioria dos casos examinados, e tanto (2) ASSIMILAÇÃO A UMA CATEGORIA quanto (3) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA no caso dos baixos índices de acuidade perceptual,

refletidos, por exemplo, na identificação das vogais as vogais /ɑ/, /ɔ/ e /æ/ e da consoante /ð/, especialmente quando essa se encontrava entre /i/ e /a/.

3.2.8 Patihis, Oh e Mogilner (2013)

Patihis, Oh e Mogilner (2013) investigaram a percepção das consoantes oclusivas do coreano por monolíngues, bilíngues e trlíngues, sendo que nenhum dos participantes tinha conhecimento do sistema sonoro coreano, portanto, não tinham experiência linguística com o idioma. O coreano, interessante, contrasta as oclusivas de acordo com três parâmetros: *LENIS* (/p/), *TENSAS*⁷³ (/p*/) e *ASPIRADAS* (/p^h/), padrão pouco frequente nas línguas do mundo já estudadas, que, em sua maioria, operam com dois parâmetros (por exemplo, vozeadas vs. não vozeadas, aspiradas vs. não aspiradas), conforme pontuam Ladefoged e Maddieson (1996).

O estudo, que tinha como instrumento uma tarefa perceptual, se dispunha a (i) verificar se a quantidade de línguas que os participantes dominavam iria se correlacionar positivamente a altos índices de discriminação no que se refere às oclusivas do coreano; e (ii) examinar se, ao serem comparados os grupos de participantes que integravam a pesquisa, haveria diferenças estatisticamente significativas quanto à acuidade discriminatória acerca das oclusivas.

Quatro grupos participaram da pesquisa, sendo um grupo-controle e três grupos experimentais. O primeiro grupo (de controle) era composto por 14 falantes monolíngues de inglês (11 mulheres e 3 homens), com idade média de 19 anos. O segundo grupo era composto por 14 falantes bilíngues de espanhol e inglês (7 mulheres e 7 homens), com idade média de 19 anos. O terceiro grupo, formado por 14 pessoas (11 mulheres e 3 homens), com idade média de 20 anos, compreendia falantes bilíngues de armênio e inglês. O quarto e último grupo consistiu, também, de 14 participantes (11 mulheres e 3 homens), com idade média de 27 anos, e se caracterizava por falantes trlíngues. Em relação aos bilíngues, a maioria adquiriu o inglês após ter adquirido sua L₁;

⁷³ Segundo Ladefoged e Maddieson (1996), o contraste segmental *LENIS* vs. *FORTIS* [OU *FROUXO* vs. *TENSO* (e respectivas variantes) para algumas explicações] tem a ver com diferenças relativas à aspiração, ao vozeamento, ou à duração (além de outros aspectos articulatórios e acústicos), que variam a depender da análise da língua em questão. Por exemplo, na tradição germânica, os termos “*LENIS*” e “*FORTIS*” se mostram mais fortuitos para descrever as diferenças entre o grau de *VOT* das oclusivas, ou seja, estão relacionados à *ASPIRAÇÃO*. Na tradição românica, por outro lado, as diferenças entre [p] e [b] podem ser captadas de modo mais apropriado se forem tratadas à luz do *VOZEAMENTO*. Especificamente, no que se refere aos termos fonéticos em questão, optamos por reportá-los, neste trecho da tese, segundo a própria nomenclatura utilizada pelos autores que conduziram o estudo, os quais deixam escapar o paralelismo *LENIS/FORTIS* ou *FROUXO/TENSO* e acabam por utilizar uma versão híbrida, *LENIS/TENSO*.

todavia, houve alguns casos em que ambas as línguas foram adquiridas simultaneamente e pouquíssimos casos em que o inglês foi adquirido primeiro. No que compete aos trilíngues, todos falavam inglês (mas não como L_1), mas as outras línguas que eles dominavam eram bastante diversas entre si (persa, hebraico, ucraniano, hindi, etc.). Os ouvintes foram convidados a realizar uma tarefa de discriminação categórica (no formato ABX), composta por 108 trios mínimos com estrutura silábica CVC , contemplando os três pontos de articulação respectivamente.

Para dar conta de cumprir o primeiro objetivo do trabalho, que indagava se o acúmulo de línguas estava positivamente relacionado a altos índices de acuidade discriminatória em relação ao contraste em coreano, os pesquisadores procederam com testes de correlação, os quais revelaram não haver correlação entre essas duas variáveis. Em outras palavras, dominar um maior número de línguas não corresponde a discriminar o contraste entre as oclusivas coreanas.

Quanto ao segundo objetivo (que propunha analisar se haveria diferenças significativas entre os grupos quanto à acuidade discriminatória acerca dos segmentos-alvo), os estudiosos conduziram testes de efeito, os quais indicaram resultados interessantes. O primeiro deles é o de que os grupos de monolíngues, de bilíngues armênio-inglês e de trilíngues obtiveram maiores índices de acuidade na discriminação do contraste consonantal coreano quando comparados ao grupo de bilíngues espanhol-inglês. Disso decorre a constatação sugerida pelos autores de que aqueles participantes fluentes em línguas que fazem a distinção entre oclusivas aspiradas e não aspiradas tiveram um melhor desempenho na acuidade discriminatória das oclusivas do coreano do que aqueles participantes fluentes em línguas que não fazem tal distinção.

Tal afirmação também é corroborada pela segunda análise realizada pelos pesquisadores, que revelam o segundo resultado previamente referido: ao excluir os participantes fluentes em línguas que aspiram oclusivas, os autores não encontraram diferenças significativas entre os participantes quanto à discriminação do contraste coreano, o que indica que, em larga medida, a acuidade discriminatória proposta na tarefa está condicionada pelo compartilhamento de propriedades fonéticas similares entre as línguas em jogo.

Apesar de os resultados encontrados não serem tratados sob a perspectiva do *PAM-L2*, ignorando qualquer referência ao estabelecimento de categorias segundo o modelo perceptual, presumimos que as previsões de tal modelo dão conta de abarcar os achados do estudo. Logo, os padrões emergentes poderiam a vir a ser o de (1) ASSIMILAÇÃO A UMA CATEGORIA e (2) DIFERENÇA DE ADEQUAÇÃO

À CATEGORIA no caso do grupo de bilíngues espanhol-inglês; e (3) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, bem como (4) ASSIMILAÇÃO A DUAS CATEGORIAS, no caso dos demais grupos testados.

3.2.9 Dufour, Kriegel, Alleesaib e Nguyen (2014)

Dufour *ET AL.* (2014) pesquisaram como falantes nativos de Maurício, cuja L₁ é o crioulo de Maurício (formado com base no francês) e cuja L₂ é o francês, percebiam o contraste entre /s/ e /ʃ/ presente em francês padrão. A escolha por esse contraste se justifica na medida em que tanto /s/ como /ʃ/ são produzidos como [s] no dialeto crioulo em questão; logo, tanto a palavra ‘*SAC*’ (que, em francês padrão, é pronunciada [sak]) como a palavra ‘*CHAQUE*’ (que, em francês padrão, é pronunciada [ʃak]) são pronunciadas como [sak] no crioulo de Maurício. Entretanto, como os participantes testados eram bilíngues (aprenderam a L₂, francês, antes dos 7 anos), eles estão expostos à variedade francesa, que estabelece a diferença entre /s/ e /ʃ/, e também estão expostos ao crioulo, que neutraliza tal contraste.

Os autores não propuseram hipóteses explícitas (embora realizem testes estatísticos) para o estudo, mas elencaram duas questões de pesquisa. A primeira objetivava verificar se os bilíngues, falantes do crioulo de Maurício e de francês, iriam categorizar os segmentos /s/ e /ʃ/ assim como o fazem os falantes monolíngues de francês. A segunda pretendia averiguar se os bilíngues apresentariam diferenças quanto à categorização de /s/ e /ʃ/ se fossem informados sobre a camada socioeconômica a que pertencia o locutor que gravou os estímulos. Nessa segunda questão de pesquisa, presumia-se que os falantes bilíngues nativos de Maurício alocados em altas camadas sociais teriam mais chances de produzir o contraste entre a fricativa alveolar e a fricativa pós-alveolar, em função de estarem vinculados a uma posição de maior prestígio frente à sociedade.

Participaram do estudo falantes monolíngues de francês e falantes bilíngues de crioulo e francês. Os falantes monolíngues, que compreenderam 15 participantes, fizeram parte do grupo-controle (com idade média de 22 anos), enquanto que os falantes bilíngues fizeram parte de dois grupos experimentais (cada qual com 15 participantes). Um dos grupos experimentais (cujos participantes tinham, em média, 21 anos) não recebeu qualquer informação sobre o nível socioeconômico do locutor que gravou os estímulos. Por sua vez, o outro grupo (cujos participantes tinham, em média, 22 anos) ficou sabendo que o locutor dos estímulos era o “Sr. Mathieu, funcionário do Banco MCB em Port-Louis” (novamente, por ocupar um cargo tido como de prestígio, a ideia a

ser transmitida aos participantes era a de que Sr. Mathieu deveria se utilizar de uma variedade mais próxima ao dialeto padrão, contrastando, portanto, /s/ e /ʃ/ em sua produção).

A testagem consistiu de uma tarefa de identificação, em que os participantes deveriam classificar o segmento inicial da sílaba /Cak/ como /s/ ou /ʃ/. Os estímulos utilizados foram alterados acusticamente em termos da porção referente ao ruído fricativo, sendo que foi gerada uma escala de 11 pontos que abrangia desde um /s/ prototípico, de um lado, até um /ʃ/ prototípico, de outro. Assim, originaram-se 11 palavras, diferentes quanto à característica fricativa, que lembravam mais /s/ ou mais /ʃ/. Cada item da escala foi apresentado 9 vezes a cada participante, resultando em 99 estímulos a serem ouvidos por cada um deles. Cabe pontuar que todos os estímulos utilizados derivaram de dois itens, ‘SAC’ e ‘CHAQUE’, que foram produzidos em condições naturais por um homem falante nativo de Port Louis, bilíngue de crioulo e francês.

Os resultados obtidos sugeriram não haver diferenças estatisticamente significativas entre o desempenho dos bilíngues e dos monolíngues quanto à categorização dos segmentos em questão. Isso significa que, independentemente de serem bilíngues ou monolíngues, o nível de acuidade na identificação de /s/ e /ʃ/ foi alto para ambos os grupos. No que se refere ao fato de haver diferenças entre os bilíngues a depender do fornecimento da informação sobre a posição socioeconômica do locutor, os pesquisadores encontraram diferenças estatisticamente significativas entre os bilíngues justamente nos pontos mais centrais da escala, em que a distinção entre /s/ e /ʃ/ é bastante sutil. Nesse aspecto, os bilíngues que foram informados sobre a posição socioeconômica do locutor julgaram haver mais instâncias de /ʃ/ do que de /s/. Para os autores, isso revela uma forte ligação entre a capacidade de produção de um segmento pós-alveolar e a camada social a que pertence o locutor, a ponto de haver um efeito no julgamento acerca do local de constrição da fricativa.

Embora Dufour *ET AL.* (2014) não façam referência às previsões de assimilação enumeradas pelo *PAM-L2*, supomos que tais padrões acomodam os resultados obtidos. Assim, dado o alto índice de acuidade perceptual na categorização de /s/ e /ʃ/ por ambos os grupos, sugerimos que os resultados, de maneira geral, se enquadram no padrão (1) ASSIMILAÇÃO A DUAS CATEGORIAS. Contudo, especialmente no caso da diferença entre os bilíngues quanto à classificação dos segmentos mais centrais da escala, que implicaram divergências no julgamento das fricativas, presumimos que o padrão mais consistente seja o de (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA.

3.2.10 Elvin, Escudero e Vasiliev (2014)

Elvin, Escudero e Vasiliev (2014) examinaram a percepção das vogais orais presentes em português brasileiro (/i, e, ε, a, ɔ, o, u/) por ouvintes espanhóis e australianos. Além de investigarem os índices de discriminação categórica resultantes das tarefas de percepção auditiva, os pesquisadores propuseram testar o poder explanatório de duas hipóteses preditoras de acuidade perceptual, a saber, tamanho do inventário vocálico e similaridade acústica.

Conforme hipotetizam os autores, se o tamanho do inventário vocálico for um bom preditor da acuidade perceptual, ouvintes australianos devem apresentar maiores índices de acurácia na discriminação dos contrastes vocálicos do português, uma vez que a probabilidade de os australianos possuírem vogais foneticamente similares às do português é mais alta em função de o inglês australiano dispor de um sistema vocálico amplo. Por outro lado, se a similaridade acústica determinar o sucesso na discriminação vocálica não nativa, ouvintes espanhóis devem apresentar maiores índices de acurácia na discriminação dos contrastes vocálicos do português, dado o fato de que as medidas acústicas (F_1 e F_2) entre as vogais dos sistemas português e espanhol são muito próximas.

Participaram do estudo dois grupos experimentais: 16 ouvintes monolíngues de inglês australiano e 15 ouvintes monolíngues de espanhol ibérico, sendo que a média de idade dos participantes, em ambos os grupos, era de 25 anos. Os ouvintes foram convidados a realizar uma tarefa de discriminação categórica (de acordo com o formato XAB), composta por 6 partes isoladas, sendo que cada qual objetivava contrastar um dos 6 pares vocálicos a seguir: /a-ɔ/, /a-ε/, /e-i/, /o-u/, /e-ε/ e /o-ɔ/. Os estímulos que ocuparam a posição X eram itens de fala natural (advindos do contexto /fVfe/, retirado de uma frase-veículo), e os estímulos ocupando as posições A e B eram itens de fala sintetizada⁷⁴.

Os resultados apontaram que os ouvintes espanhóis apresentaram maiores índices de acuidade na discriminação dos contrastes testados, com exceção dos pares /o-u/ e /e-ε/ – em relação a esses dois pares, não houve diferença estatisticamente significativa entre espanhóis e australianos.

⁷⁴ Segundo os autores, foram utilizados estímulos sintéticos com valores médios das vogais do português para que os ouvintes fizessem seus julgamentos baseados em uma comparação entre itens individuais e valores prototípicos. Para os estudiosos, isso resulta em uma tarefa de discriminação categórica, em que se espera que os ouvintes fundamentem sua decisão sobre A e B serem mais similares a X com base em diferenças fonêmicas, ao invés de diferenças acústicas.

Além disso, os resultados também sugeriram que os pares /e-i/ e /o-u/ apresentaram níveis de acuidade discriminatória significativamente menores do que os demais pares, indicando que eles são os mais difíceis de serem diferenciados. Diferentemente, o par /a-ε/ foi o que apresentou os maiores índices de acuidade discriminatória.

Elvin, Escudero e Vasiliev (2014) ressaltam que o fato de os ouvintes espanhóis obterem maior acuidade perceptual, mesmo não possuindo determinados contrastes presentes tanto em português como em inglês, indica que as vogais que estão acusticamente mais próximas às vogais-alvo sem ativação de categorias competitivas são mais facilmente discriminadas. Nesse aspecto, ouvintes australianos podem usar todas as suas categorias vocálicas que estão mais próximas acusticamente da vogal-alvo do português brasileiro, e isso pode impor problemas de julgamento, já que há uma multiplicidade de opções de resposta possíveis, resultando em um baixo desempenho perceptual (como refletem os resultados do estudo). O melhor desempenho dos espanhóis na discriminação categórica dos pares testados, portanto, é uma evidência que, segundo os estudiosos, corrobora a hipótese de que a similaridade acústica (em detrimento do tamanho do inventário vocálico) é o preditor mais confiável para a acuidade perceptual de contrastes sonoros não nativos.

Apesar de os autores não mostrarem os resultados quantificados separadamente para cada grupo, eles mencionam que o nível de dificuldade na discriminação categórica dos contrastes apresentados pelos participantes (de ambos os grupos) foi o seguinte: /o-u/ e /e-i/, então /a-ɔ/, /e-ε/, /o-ɔ/, e, por fim, /a-ε/. Ainda que os pesquisadores não façam referência ao estabelecimento de categorias aos moldes do *PAM-L2*, presumimos que as previsões de tal modelo dão conta de abarcar os resultados encontrados e os padrões emergentes poderiam vir a ser o de (1) ASSIMILAÇÃO A UMA CATEGORIA no caso dos contrastes /o-u/ e /e-i/, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA no caso dos contrastes /a-ɔ/, /e-ε/ e /o-ɔ/, e (3) ASSIMILAÇÃO A DUAS CATEGORIAS no caso do contraste /a-ε/.

3.2.11 Escudero, Sisinni e Grimaldi (2014)

Escudero, Sisinni e Grimaldi (2014) investigaram o efeito do inventário vocálico e de propriedades acústicas na percepção de vogais inglesas. Basicamente, os autores estavam interessados em verificar como aprendizes italianos (pertencentes ao dialeto de Salento, cujo sistema vocálico prescinde dos segmentos médios-baixos [ε] e [ɔ]) identificavam e discriminavam as vogais do inglês, mais especificamente da variedade britânica meridional.

O estudo contou com 30 participantes do gênero masculino, sendo 12 monolíngues provenientes de Salento (Itália), que atuaram no grupo experimental, e 18 monolíngues britânicos, que configuraram o grupo-controle. Os estímulos auditivos utilizados nas tarefas de identificação e discriminação foram as vogais /i:, ɪ, ε, æ, ʌ, ɑ:, ɒ, ɔ:, ɜ:, ʊ, u:/, retiradas da posição nuclear de sílabas CVC referentes a palavras reais do dialeto inglês em questão.

A tarefa de identificação foi composta por 108 estímulos apresentados aleatoriamente e, para a sua realização, os participantes italianos não sabiam que se tratava de segmentos ingleses (imediatamente após selecionar a vogal desejada, os participantes deveriam classificá-la de 1 a 5, em que 1 correspondia ao segmento totalmente diferente daquele da L1 e 5 correspondia ao segmento totalmente idêntico àquele da L1).

Em relação à tarefa de discriminação, composta por 144 estímulos, o objetivo era testar os contrastes /i:-ɪ/, /ɒ-ɔ:/, /æ-ɑ:/, /æ-ʌ/, /ɑ:-ʌ/, /ε-ɜ:/, /ɒ-ɜ:/, /u:-ʊ/ e /i:-u:/, na qual os participantes deveriam ouvir a tríade apresentada e indicar qual era a posição em que se encontrava o elemento destoante (*ODDITY DISCRIMINATION TASK*), através do toque nos dígitos 1, 2 ou 3 do teclado, respectivamente. Com base na comparação acústica entre as vogais italianas e as vogais inglesas, as hipóteses para ambas as tarefas eram as de que (1) as vogais inglesas /i:, ε, æ, ɒ, u:/ seriam classificadas de acordo com a sua contraparte acústica italiana mais próxima, /i, e, a, o, u/; (2) as vogais inglesas /ɪ, ɑ:, ʌ, ɜ:, ɔ:, ʊ/ seriam classificadas segundo duas ou mais contrapartes acústicas italianas mais próximas: /ɪ/ seria classificada como /i/ ou /e/; /ɑ:/ e /ʌ/ seriam classificadas como /a/ e /o/; /ɜ:/ seria classificada como /e/, /o/ ou /a/; e /ɔ:/ e /ʊ/ seriam classificadas como /o/ e /u/.

Conforme os resultados obtidos na tarefa de identificação, diversas vogais inglesas foram categorizadas segundo as previsões estabelecidas nas hipóteses do trabalho, especialmente aquelas designadas na primeira hipótese. Os resultados da tarefa de discriminação se alinham aos de identificação e revelam que alguns contrastes vocálicos impõem maiores dificuldades aos ouvintes italianos, especialmente em função de seu estágio inicial de desenvolvimento no idioma estrangeiro. Nesse quesito, os contrastes /ɑ:-ʌ/ e /ɒ-ɔ:/ implicaram maior dificuldade para a categorização perceptual dos ouvintes italianos.

Os pesquisadores responsáveis pelo estudo chamam atenção para o pressuposto de que as diferenças acústicas entre os segmentos testados levam a diferenças perceptuais dos sons no âmbito

não nativo e de que a variabilidade do sinal acústico pode resultar em categorizações errôneas no mapeamento interfonológico.

Ainda que os autores não conectem seus achados às previsões estabelecidas pelo *PAM-L2*, estas parecem, mais uma vez, comportar os resultados. Nesse sentido, sugerimos que os padrões do tipo (1) ASSIMILAÇÃO A DUAS CATEGORIAS, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA e (3) ASSIMILAÇÃO A UMA CATEGORIA são aplicáveis aos dados veiculados no estudo.

3.2.12 Stölten, Abrahamsson e Hyltenstam (2014)

O objeto de investigação da pesquisa de Stölten, Abrahamsson e Hyltenstam (2014) tinha a ver com os efeitos da idade de aquisição sobre os padrões de *VOT* do sueco. Mais especificamente, o estudo verificou como falantes nativos de espanhol identificavam as oclusivas suecas /p, b, t, d, k, g/. Assim como ocorre em inglês, as oclusivas /p, t, k/ e /b, d, g/ em sueco diferem em termos de *VOT* positivo e *VOT* zero, respectivamente, e a percepção auditiva de tal contraste, no caso de ouvintes não nativos, pode ser influenciada por questões referentes à idade com que tais ouvintes iniciaram a aquisição do sistema-alvo.

Aliando a capacidade perceptual dos seres humanos ao construto psicolinguístico relacionado à discussão sobre períodos críticos para a aquisição de línguas, os autores lançaram mão de duas hipóteses. A primeira era a de que fatores como (i) início da aquisição e (ii) *STATUS* de falante nativo de sueco ou aprendiz de sueco como *L2* influenciam a localização das fronteiras entre as categorias sonoras. Neste âmbito, os falantes nativos de sueco apresentariam maiores índices de acuidade perceptual quanto ao *VOT* das oclusivas em questão se comparados aos bilíngues tardios. Por sua vez, aos bilíngues precoces caberia o meio termo: ficariam entre os falantes nativos e os bilíngues tardios no que tange aos níveis de acurácia na percepção dos segmentos testados. A segunda hipótese, intimamente relacionada à primeira, estabelecia que pouquíssimos aprendizes tardios de sueco (*L2*) e a maioria dos aprendizes precoces iriam exibir índices de acuidade perceptual semelhantes aos dos falantes nativos⁷⁵.

Como amostra da pesquisa, participaram 41 falantes nativos de espanhol (9 homens e 32 mulheres, idade média de 32 anos) cuja *L2* era o sueco, os quais foram divididos em 2 grupos: (i)

⁷⁵ Os autores descrevem uma terceira hipótese, mas ela tem a ver com dados de produção, e, como já foi mencionado, tais dados não serão cobertos nesta apresentação.

aprendizes precoces, que aprenderam sueco entre 1 e 11 anos de idade, e (ii) aprendizes tardios, que aprenderam sueco entre 13 e 19 anos de idade. Vale destacar que esses 41 falantes bilíngues tinham, em média, 25 anos de residência em Estocolmo e adquiriram a variedade de sueco falada nessa região. Como grupo-controle, recorreu-se a 15 falantes nativos de sueco, com idade média de 30 anos, os quais realizaram a mesma tarefa aplicada aos bilíngues.

O teste perceptual empregado compreendia uma tarefa de identificação categórica que contou com manipulação acústica dos valores referentes ao *VOT* das oclusivas em sueco, sendo que as alterações derivaram de estímulos naturais gravados por uma pesquisadora de nacionalidade sueca com treinamento em fonética. Os estímulos utilizados foram as palavras /par/, /bar/, /tal/, /dal/, /kal/ e /gal/, que variaram em 31 pontos de um contínuo com intervalo de 5ms, cujas extremidades eram -60ms e 90ms – pontos suficientes para acomodar fonologicamente o contraste de *VOT* existente entre as oclusivas do espanhol e do sueco. Designaram-se, portanto, 31 estímulos para identificação de cada ponto de articulação (bilabial, alveolar e velar), apresentados aleatoriamente aos participantes, totalizando uma tarefa de 93 estímulos.

Os resultados são bastante interessantes. Embora os participantes tenham obtido altos índices de acuidade perceptual quanto às oclusivas testadas (100% no caso dos falantes nativos, 94% no caso dos bilíngues precoces e 93% no caso dos bilíngues tardios), os quais não diferiram estatisticamente, notou-se que os falantes nativos de sueco mudaram a localização das fronteiras fonológicas em pontos onde o *VOT* apresentava médias mais altas de duração (isto é, contrastavam oclusivas surdas e sonoras na medida em que aumentava a duração existente na liberação do ar pulmônico), enquanto que os bilíngues mudaram a localização das fronteiras fonológicas de maneira mais difusa e espaçada.

Apesar de haver variabilidade no ponto de corte entre uma oclusiva surda e uma oclusiva sonora por parte dos bilíngues, os pesquisadores observaram que havia mais consistência na segmentação das categorias segundo as respostas fornecidas pelos bilíngues precoces do que pelos bilíngues tardios. Os pesquisadores relatam, adicionalmente, que somente dois bilíngues tardios apresentaram níveis de identificação perceptual do tipo nativo e, no caso dos bilíngues precoces, cerca de 70% dos participantes apresentaram índices de identificação perceptual no nível dos falantes nativos. Tais achados confirmam tanto a primeira como a segunda hipótese elencadas para a investigação.

Novamente, os achados não foram tratados segundo os padrões de assimilação sugeridos pelo *PAM-L2*, além de nenhuma relação ter sido feita entre os resultados do estudo e as previsões veiculadas em tal modelo perceptual. Todavia, presumimos que os padrões assimilatórios do *PAM-L2* dão conta de acomodar os resultados obtidos e, possivelmente, os padrões relevantes seriam: (1) ASSIMILAÇÃO A DUAS CATEGORIAS de um modo geral, dados os altíssimos índices de acuidade verificados, e (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, no caso dos julgamentos mais difusos observados somente na região central do contínuo acústico.

3.2.13 Williams e Escudero (2014)

Williams e Escudero (2014) trataram, novamente, das influências dialetais na percepção de vogais de um sistema estrangeiro. Os autores pesquisaram, então, a percepção de vogais do holandês por falantes nativos de inglês britânico, pertencentes aos dialetos do norte e do sul da Inglaterra.

Os participantes provenientes do norte do país totalizaram 20 sujeitos (idade média de 20 anos, 10 homens e 10 mulheres) e os participantes do sul do país também totalizaram 20 sujeitos (idade média de 20 anos, 10 homens e 10 mulheres). A hipótese do estudo era a de que similaridades acústicas entre vogais nativas e não nativas são capazes de prognosticar a percepção vocálica não nativa e, também, de que o dialeto materno influencia, em alguma medida, a percepção de contrastes vocálicos não nativos.

Serviram de estímulo para o trabalho 300 palavras holandesas, inseridas no contexto /fVf/, cujo elemento vocálico nuclear variou de acordo com os segmentos /i, y, ɪ, ʏ, ε, a:, α, ɔ, u, e, ø, ei, œy, ʌu/. Foi administrada aos participantes uma tarefa de categorização de múltipla escolha, em que eles deveriam ouvir o estímulo e julgar, segundo a representação ortográfica dos segmentos vocálicos ingleses /i:, ɪ, ε, ɜ:, a:, α:, ʌ, ɒ, ɔ:, ʊ, u:, eɪ, əʊ, aʊ, aɪ, ɔɪ/, qual fone estrangeiro se relacionava à categoria nativa – uma vez que lhes foi informado que eles ouviriam sons linguísticos retirados da fala corrente produzida por ingleses.

Os resultados obtidos revelam que os ingleses apresentaram altos índices de acuidade para a classificação das vogais, assim como para a categorização dos ditongos, e apontam para o fato de que, ao ter familiaridade com o outro dialeto na *L1*, os ouvintes se mostram ainda mais habilidosos no que concerne ao seu julgamento acerca das vogais. Para os autores, a similaridade acústica entre os

segmentos pode, sim, prever a assimilação de vogais nativas às categorias da L_1 , e as variedades linguísticas potencializam a percepção da fala estrangeira.

Os autores não fazem menção aos padrões de assimilação estabelecidos pelo *PAM-L2*, mas, ao que parece, o modelo daria conta, de fato, de sustentar os dados encontrados pelos pesquisadores. Além disso, sugerimos que os padrões de assimilação envolvidos no que compete a essa pesquisa em específico sejam (1) ASSIMILAÇÃO A DUAS CATEGORIAS e, em alguns poucos casos, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA.

3.2.14 Yang e Fox (2014)

Yang e Fox (2014) debruçaram-se sobre a percepção de vogais inglesas por falantes bilíngues de chinês e inglês, falantes monolíngues de chinês e falantes monolíngues de inglês (variedade americana). O objetivo do trabalho era o de verificar como a L_1 e a experiência com a L_2 influenciam a organização espacial das vogais, em termos de percepção, dos ouvintes chineses bilíngues e monolíngues.

Os autores contavam com duas hipóteses. A primeira era a de que os ouvintes chineses, ao ouvir as vogais do inglês, iriam perceber os segmentos periféricos /i, a, u/ como sendo os mais distantes no espaço perceptual, dadas as suas distinções acústicas (F_1 e F_2). A segunda hipótese previa que os ouvintes chineses, especialmente os que tinham menos experiência com o inglês, tenderiam a assimilar as vogais inglesas menos periféricas (que estão mais próximas entre si no espaço acústico) a categorias vocálicas similares na L_1 , uma vez que os dialetos chineses falados pelos participantes possuem inventários vocálicos menores do que a variedade americana do inglês a que estão expostos. Ainda em relação à segunda hipótese, os pesquisadores presumiam que, no que diz respeito aos ouvintes chineses com maior experiência com o inglês, a imersão nessa língua modificaria gradualmente a sua habilidade perceptual, refletindo uma acuidade perceptual muito próxima à dos falantes nativos de inglês americano.

Os participantes da pesquisa foram distribuídos em três grupos, sendo dois experimentais e um de controle. Os dois primeiros, experimentais, foram compostos por 11 ouvintes monolíngues de chinês (idade média de 53 anos), com pouquíssima ou nenhuma exposição ao inglês, residentes na China, e 10 ouvintes bilíngues de chinês e inglês (idade média de 27 anos), todos residentes em Ohio (EUA) e alunos de cursos de graduação, que, antes de migrar para os EUA tinham tido, em média, 10

anos de estudo formal do idioma na China. O grupo-controle consistiu em 10 ouvintes monolíngues americanos (idade média de 22 anos), falantes nativos do dialeto de Ohio, e seus participantes também eram alunos de cursos de graduação no momento da testagem.

Para a realização da pesquisa, os pesquisadores aplicaram uma tarefa de discriminação categórica no formato *AX* (porém, chamaram-no de “teste de identificação de similaridade”), contendo 312 estímulos. Como os autores estavam interessados em examinar a percepção das categorias vocálicas inglesas, os segmentos /i, I, e, ε, æ, u, ʊ, o, ɔ, a, ɔI, aI, aʊ/ encontravam-se como elementos nucleares em sequências do tipo /hVd/. Assim, havia 13 elementos vocálicos distribuídos em 12 pares no formato *AB* e em 12 pares no formato *BA*, contrastantes entre si. Os participantes deveriam, então, ouvir o par mínimo apresentado e eleger, a partir de uma escala de dissimilaridade de 9 pontos (1 = muito similar; 9 = muito diferente), o quão semelhantes ou distintas eram as sílabas ouvidas – caso desejassem, os participantes poderiam repetir os estímulos até três vezes.

Com base na análise feita pelos pesquisadores, todos os participantes se valeram de distinções bidimensionais como anterior/posterior e alta/baixa para classificar as vogais que ouviram. Quanto aos ouvintes monolíngues de inglês, o terceiro parâmetro que auxiliou na percepção dos contrastes vocálicos foi o *OFFSET* alto-anterior, componente final de elementos nucleares como /ɔI/ e /aI/, relacionado, acusticamente, a mudanças espectrais. Quanto aos ouvintes bilíngues, similarmente ao que ocorreu com os monolíngues de inglês, o terceiro parâmetro que auxiliou na percepção dos contrastes vocálicos foi o *STATUS* monotongo/ditongo do elemento nuclear. Quanto aos ouvintes monolíngues de chinês, o terceiro parâmetro que auxiliou na percepção dos contrastes vocálicos foi a separação das vogais altas (anterior e posterior) das demais. Além disso, ouvintes monolíngues de chinês tenderam menos a julgar as vogais como distintas quando comparados aos ouvintes monolíngues de inglês, derivando daí um agrupamento das vogais inglesas no espaço acústico perceptual dos chineses e evidenciando um afastamento das vogais periféricas, corroborando a primeira hipótese levantada pelos estudiosos.

Conforme estabelecido na segunda hipótese apresentada pelos autores, ouvintes monolíngues de chinês apresentaram maiores índices de assimilação perceptual das vogais da *L2* ao seu sistema de *L1* se comparados aos ouvintes bilíngues, fato corroborado por meio da análise estatística realizada. Yang e Fox (2014) salientam que, inclusive, os ouvintes monolíngues de chinês tenderam a assimilar as vogais da *L2* em grandes aglomerados, correspondendo às categorias

vocálicas chinesas, enquanto que os bilíngues perceberam os contrastes vocálicos de maneira aproximada à dos ouvintes monolíngues de inglês.

Como conclusão, os autores apontam que, de fato, há grande influência da experiência linguística no julgamento de similaridade ou dissimilaridade entre as vogais, como ilustraram os índices de discriminação dos bilíngues. Ademais, segundo explanam os pesquisadores, as diferenças existentes entre ouvintes monolíngues de inglês e de chinês quanto à acuidade perceptual dos itens-alvo reforça o papel da L_1 na percepção auditiva translinguística.

Em nosso entendimento, o *PAM-L2* é um modelo plenamente cabível no que concerne ao tratamento dos resultados em termos de suas previsões de assimilação. Assim, julgamos que os possíveis casos de assimilação a emergir seriam (1) ASSIMILAÇÃO A DUAS CATEGORIAS e (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA no que se refere aos julgamentos perceptuais dos ouvintes bilíngues e, no que diz respeito aos julgamentos perceptuais dos ouvintes monolíngues de chinês, os padrões de assimilação seriam (3) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, novamente, (4) ASSIMILAÇÃO A UMA CATEGORIA, para a maioria das respostas fornecidas, e, em alguns casos, (5) ASSIMILAÇÃO A DUAS CATEGORIAS, especialmente no quesito da separação entre vogais altas e o restante do inventário vocálico.

3.2.15 Alves e Zimmer (2015)

Alves e Zimmer (2015) analisaram como diferentes padrões de *VOT* do inglês americano eram percebidos por aprendizes brasileiros do idioma⁷⁶. Os pesquisadores partiram do pressuposto de que os aprendizes com baixa proficiência não recorreriam primordialmente ao *VOT* para distinções funcionais, enquanto que aprendizes com alta proficiência o fariam.

Participaram da pesquisa 34 sujeitos, divididos em dois grupos, cada qual correspondente a um determinado nível de proficiência: 24 participantes foram alocados no nível básico e 10 participantes compuseram o nível avançado. Os estímulos utilizados foram produções reais das palavras inglesas ‘*PIT*’, ‘*BIT*’, ‘*TICK*’, ‘*DICK*’, ‘*KILL*’ e ‘*GILL*’, gravadas por falantes norte-americanos.

Em se tratando das palavras iniciadas por oclusivas não vozeadas, atestou-se nos estímulos utilizados, somente, a presença de *VOT* positivo; no que diz respeito às palavras iniciadas por oclusivas vozeadas, atestaram-se *VOT* negativo e *VOT* zero. Além disso, incluiu-se outro tipo de padrão de *VOT*,

⁷⁶ Este é outro trabalho em que foram administradas tarefas de percepção e de produção. Porém, em função dos objetivos desta tese, serão desconsideradas referências aos dados de produção.

chamado de *vOT* zero manipulado, que resultou de uma modificação acústica: os pesquisadores selecionaram apenas as palavras contendo *vOT* positivo e retiraram a porção referente à aspiração (retirada essa que aproximaria um padrão de *vOT* positivo a um padrão de *vOT* zero), de modo a suprimir tal pista acústica, mas mantendo as outras pistas acústicas (como a frequência fundamental da vogal seguinte à oclusiva ou a intensidade da liberação do ar pulmônico) que contribuem para a caracterização de uma oclusiva não vozeada em inglês.

Os autores aplicaram aos participantes uma tarefa de identificação perceptual, contendo 48 estímulos, e uma tarefa de discriminação categórica, que seguiu o formato *AXB* e contou com 45 tríades (36 delas envolvendo o contraste de *vOT* e 9 delas compostas por elementos idênticos, os chamados *CATCH TRIALS*).

Quanto à tarefa de identificação, os estudiosos tinham como hipótese a falta de diferença estatística entre os dois grupos em relação aos padrões de *vOT* negativo, *vOT* positivo e *vOT* zero; no entanto, os autores previam diferença estatística entre ambos os grupos em relação ao padrão de *vOT* zero manipulado. Quanto à tarefa de discriminação, os pesquisadores tinham como hipótese a falta de diferença estatística entre os dois grupos em relação à percepção do contraste *vOT* negativo-*vOT* positivo; todavia, os autores previam diferenças estatísticas entre ambos os grupos em relação à percepção dos contrastes *vOT* negativo-*vOT* zero manipulado e *vOT* positivo-*vOT* zero manipulado.

Os resultados obtidos corroboram parcialmente as hipóteses elencadas pelos pesquisadores. Na tarefa de identificação, não houve diferenças estatísticas entre os participantes em qualquer um dos quatro padrões de *vOT*, sendo que a diferença entre nível de proficiência na classificação dos estímulos contendo *vOT* zero manipulado, hipotetizada pelos pesquisadores, não foi corroborada. Na tarefa de discriminação, tampouco houve diferenças estatísticas entre os participantes em qualquer um dos três contrastes de *vOT*, sendo que a diferença entre nível de proficiência na discriminação dos contrastes *vOT* negativo-*vOT* zero manipulado e *vOT* positivo-*vOT* zero manipulado não foi observada, contrariando a hipótese dos pesquisadores.

De maneira geral, tanto para a identificação como para a discriminação, os participantes não diferiram, em termos de proficiência, quanto à percepção dos padrões de *vOT* investigados. Em outras palavras, os resultados sugerem que o nível de proficiência da amostra testada não tem impacto na percepção dos padrões referidos, evidenciando que o *vOT* não se mostra como a pista fundamentalmente seguida pelos aprendizes na percepção dos padrões apresentados. Os autores,

ainda, ressaltam a necessidade de instrução explícita sobre o fenômeno para que ele passe a servir como uma pista relevante na manutenção do contraste de vozeamento em inglês.

Embora façam referência a aspectos teóricos do *PAM-L2*, Alves e Zimmer (2015) não relacionam seus resultados aos padrões de assimilação previstos pelo modelo de Best e Tyler (2007). Supomos que os dados gerados poderiam ser tratados à luz do *PAM-L2* e os padrões mais recorrentes seriam os de (1) ASSIMILAÇÃO A DUAS CATEGORIAS, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, e (3) ASSIMILAÇÃO A UMA CATEGORIA (este especialmente para os resultados da discriminação do contraste *VOT* positivo-*VOT* zero manipulado, no caso dos aprendizes de nível avançado).

3.2.16 Holliday (2016)

A pesquisa de Holliday (2016) propôs examinar a assimilação perceptual e a discriminação dos segmentos /s^h/ e /s^{*}/ do coreano por ouvintes chineses com diferentes níveis de proficiência na língua testada. Além de versar sobre o referido processo assimilatório, o autor parte da premissa de que, em muitos casos, a experiência com uma L₂ pode vir a minimizar a discriminação de contrastes fonológicos não nativos, seja pela atenção a outras pistas (contextuais, por exemplo) que não acústicas ou articatórias, seja pelo conhecimento linguístico resultante da experiência com o idioma não nativo.

Com base em estudos anteriores, envolvendo amostras e contrastes semelhantes ao trabalho em questão, o autor lançou mão de três hipóteses. A primeira previa que ouvintes monolíngues chineses não apresentariam acuidade na discriminação do contraste coreano /s^h/-/s^{*}/ diante de qualquer contexto vocálico. A segunda estabelecia que tanto monolíngues chineses como chineses aprendizes de coreano apresentariam acuidade na discriminação do contraste coreano /s^h/-/s^{*}/. Por fim, a terceira estipulava que ouvintes monolíngues chineses apresentariam acuidade na discriminação de /s^h/ e /s^{*}/, mas os aprendizes chineses de coreano não apresentariam acuidade na discriminação de tais segmentos.

Participaram do estudo 69 estudantes universitários que frequentavam cursos de graduação em Bloomington e em Seul, divididos em 4 grupos. O primeiro grupo era composto por 18 ouvintes chineses monolíngues (13 mulheres e 5 homens), com idade média de 20 anos. O segundo grupo era composto por 15 ouvintes chineses (9 mulheres e 6 homens), aprendizes iniciantes de coreano, com idade média de 24 anos. O terceiro grupo era composto por 17 ouvintes chineses (11 mulheres e 6

homens), aprendizes avançados de coreano, com idade média de 24 anos. O quarto grupo era composto por 19 ouvintes monolíngues coreanos (12 mulheres e 7 homens), também com idade média de 24 anos.

Os estímulos auditivos utilizados na primeira tarefa eram sílabas do tipo *cv*, em que os segmentos /s^h, s*, tɕ, tɕ^h, tɕ*, t, t^h, t*/ ocupavam a posição consonantal e os segmentos /i, a, u/ ocupavam a posição vocálica. Aqueles utilizados na segunda tarefa eram caracterizados por tríades de sílabas *cv*, as quais eram compostas por contrastes envolvendo (a) dois elementos fricativos: /s^h/ vs. /s*/; (b) um elemento fricativo e um elemento africado: /s^h/ ou /s*/ vs. /tɕ/, /tɕ^h/ ou /tɕ*/; (c) dois elementos africados: /tɕ/ vs. /tɕ^h/, /tɕ*/ vs. /tɕ/ ou /tɕ^h/ vs. /tɕ*/; e (d) dois elementos oclusivos, que atuaram como distratores⁷⁷. Os elementos vocálicos eram representados pelos segmentos /i, a, u/. Por fim, os estímulos auditivos utilizados na terceira tarefa eram também sílabas do tipo *cv*, em que os segmentos /s^h/ ou /s*/ ocupavam a posição consonantal, e os segmentos /a, ʌ, o, u, i, i/ ocupavam a posição vocálica.

O primeiro teste correspondeu a uma tarefa de assimilação perceptual⁷⁸ e contou com 156 estímulos, gravados por 4 locutoras nativas de coreano. Os participantes ouviam cada estímulo e selecionavam qual consoante lembrava a sua contraparte em mandarim. Para tanto, eram exibidas 11 opções de respostas na tela, cada qual representando, em mandarim, o segmento ouvido (3 fricativas, 6 africadas e 2 oclusivas). Após a seleção, os participantes elegiam o quão similar era o estímulo ouvido à categoria consonantal do mandarim. O segundo teste tinha a ver com uma tarefa de discriminação categórica no formato *ODDITY DISCRIMINATION TASK*, em que os participantes deveriam ouvir cada tríade apresentada e indicar qual era a posição em que se encontrava o elemento destoante. Das 108 tríades veiculadas nessa tarefa, 60 foram formadas com base em múltiplos falantes e 40 foram configuradas a partir de um só falante. A terceira tarefa se referia a uma tarefa de identificação perceptual⁷⁹, composta por 60 estímulos, gravados por 5 falantes nativos de coreano. Nessa tarefa, os participantes ouviam o estímulo e o relacionavam a uma representação ortográfica em coreano.

⁷⁷ O autor não menciona quais eram os segmentos oclusivos utilizados.

⁷⁸ Segundo o autor, os falantes nativos de coreano não realizaram a tarefa de assimilação perceptual porque essa presumia a existência de categorias fonológicas nativas do mandarim.

⁷⁹ De acordo com o pesquisador, os falantes monolíngues de mandarim não realizaram a tarefa de identificação perceptual, uma vez que essa requeria conhecimento prévio de coreano.

Ao interpretar os resultados considerando as três tarefas de percepção aplicadas aos participantes, Holliday (2016) descreve que as evidências encontradas se mostram favoráveis à terceira hipótese levantada. Nesse sentido, dentre os falantes nativos de mandarim, a assimilação perceptual das fricativas sibilantes do coreano difere entre aqueles que não têm experiência com o coreano e aqueles que são aprendizes dessa língua. Tais diferenças na assimilação perceptual predizem a capacidade discriminatória desses grupos de ouvintes, as quais resultam no fato de os ouvintes monolíngues chineses apresentarem maiores índices de acuidade quanto à discriminação das fricativas coreanas do que os aprendizes básicos de coreano. Ainda, mesmo com aproximadamente três anos de experiência com o coreano, os aprendizes avançados tampouco apresentaram maior acuidade perceptual acerca das fricativas testadas, uma vez comparados aos monolíngues chineses, os quais nunca haviam estudado coreano.

Muito embora o autor não faça referência ao estabelecimento de categorias com base nos padrões assimilatórios do *PAM* ou do *PAM-L2*, presumimos que as previsões desses modelos dão conta de abarcar os resultados encontrados, cujos padrões de categorização poderiam a vir a ser (1) ASSIMILAÇÃO A UMA CATEGORIA, no caso dos baixos índices de acuidade, (2) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, no caso dos índices moderados de acuidade, e (3) ASSIMILAÇÃO A DUAS CATEGORIAS, no caso dos altos níveis de acuidade.

3.3 Trabalhos brasileiros de pós-graduação com enfoque nas previsões assimilatórias do *PAM-L2*

Em nosso levantamento de trabalhos brasileiros em nível de pós-graduação referentes ao período de 2010 a 2016, encontramos um número razoável de dissertações de mestrado e teses de doutorado que tinham como objeto de estudo a percepção de sons linguísticos não nativos e que satisfaziam nossos critérios fundamentais de busca. Nesse contexto, endereçamos os estudos de Kuhn (2011), Albuquerque (2012), Perozzo (2013), Silva (2014), Elpo (2015), Moreira (2015), Schwartzhaupt (2015) e Oliveira (2016), no que diz respeito a dissertações de mestrado, e as pesquisas de Souza (2012), Machry da Silva (2014) e Santos (2014), que compreendem teses de doutorado.

Com exceção dos trabalhos de Albuquerque (2012) e Perozzo (2013), nenhum estudo vislumbrou seus achados sob o prisma dos padrões de assimilação perceptual indicados no *PAM-L2*, ainda que tenham mencionado o modelo em seu referencial teórico, principalmente, para evidenciar

o papel da experiência linguística no desenvolvimento da percepção de sons linguísticos não nativos⁸⁰. Notamos que, apesar de os trabalhos de pós-graduação citados resenharem o *PAM* e o *PAM-L2*, sobretudo quando se trata de discorrer sobre o papel da experiência linguística sobre a percepção, tais estudos não necessariamente assumem o *PAM* ou o *PAM-L2*. Ademais, alguns estudos, inclusive, resenham outros modelos perceptuais, como o *SLM*, sem aderir a um ou a outro construto. Em outras palavras, parece não ser uma preocupação dos trabalhos mencionados se comprometer com um modelo de percepção específico.

Conforme argumentamos no início deste capítulo, por questões de delimitação de espaço, ao tratar dos trabalhos brasileiros de pós-graduação que tocaram no quesito da percepção de sons linguísticos, faremos menção somente aos estudos que vislumbraram explicitamente seus resultados sob o enfoque das previsões assimilatórias do *PAM-L2*, os quais correspondem somente às dissertações de mestrado de Albuquerque (2012) e Perozzo (2013). É importante salientar que tais estudos não foram planejados de modo a seguir os preceitos do *PAM-L2* do início ao fim, mas oportunizaram uma apresentação de seus resultados segundo uma interpretação dos padrões assimilatórios do *PAM-L2*.

3.3.1 Albuquerque (2012)

Albuquerque (2012) verificou a percepção acerca da distinção de vozeamento em segmentos oclusivos, alocados em posição final de sílaba, correspondente a palavras do inglês americano por falantes nativos de português brasileiro. Fazia parte de seu objetivo, também, averiguar a influência do ponto de articulação e do nível de proficiência dos aprendizes sobre a percepção do fenômeno em jogo.

Participaram da pesquisa 60 estudantes universitários (32 mulheres e 28 homens), com idade média de 30 anos, os quais eram alunos de língua inglesa em um curso oferecido pela universidade em que ocorreu a pesquisa (Universidade Federal do Paraná). Os acadêmicos, pertencentes ao grupo experimental, foram divididos em três níveis de proficiência em língua inglesa, básico, pré-intermediário e intermediário, cada qual consistindo em 20 participantes. Assegurou-se que os aprendizes de nível básico estivessem cursando, pelo menos, o segundo semestre do curso de

⁸⁰ No que se refere à tese de Santos (2014), não se aplicam as previsões de assimilação do *PAM-L2* em função de a autora não considerar que as previsões de tal modelo possam ser suficientes para o fenômeno que ela está tratando. Iremos discorrer a respeito dessa constatação na seção 3.4.

língua inglesa, o que garantiria um conhecimento minimamente satisfatório acerca do contraste de vozeamento em final de palavras do inglês. Outros 8 alunos do curso de idiomas oferecido pela universidade fizeram parte da pesquisa, porém alocados no grupo-controle, pois se tratava de falantes nativos de língua inglesa, advindos do estado da Carolina do Norte. Tais alunos (3 mulheres e 5 homens) tinham idade média de 29 anos e permaneceram no Brasil por cerca de um mês.

Os participantes foram testados a partir de uma tarefa de identificação e de uma tarefa de discriminação (formato *AX*), cujos estímulos eram palavras monossilábicas, gravadas por um falante nativo de inglês americano proveniente, também, da Carolina do Norte. Desse modo, as palavras utilizadas como estímulo foram ‘CAP’, ‘CAB’, ‘BAT’, ‘BAD’, ‘BACK’ e ‘BAG’, sendo que todas passaram por algum tipo de manipulação acústica, já que um dos objetivos do experimento era verificar se havia uma pista acústica que fosse mais saliente (que auxiliasse na percepção das oclusivas surdas e sonoras alocadas em final de palavra). Logo, as manipulações acústicas residiram no âmbito da duração da vogal que antecedia a oclusiva final e do vozeamento dessa.

No que se refere à interpretação dos resultados, Albuquerque (2012) menciona que os aprendizes brasileiros discriminaram e identificaram com maior acuidade o contraste surdo *vs.* sonoro para as bilabiais a partir da presença de vozeamento do que a partir da duração da vogal antecedente preservada. No que se refere às oclusivas alveolares e velares, o resultado foi contrário: a duração da vogal antecedente preservada repercutiu em maiores índices de acuidade para a distinção do vozeamento final. A pesquisadora chama atenção para o fato de que é possível, de fato, observar a interação das variáveis “ponto de articulação” e “parâmetro acústico”. Em termos de diferenças de acuidade em função dos diferentes níveis de proficiência, a autora relata não ter encontrado diferenças significativas entre os grupos, o que parece sugerir que o nível de proficiência não é determinante para a acuidade no contraste do vozeamento final.

Ao vislumbrar os achados à luz do *PAM-L2*, a autora menciona que a percepção dos participantes quanto ao contraste de vozeamento pode ser estabelecida por meio dos padrões assimilatórios do *PAM-L2* e refletem, especialmente, casos como (1) ASSIMILAÇÃO A DUAS CATEGORIAS, considerando-se que a discriminação e identificação no que respeita à manipulação da duração da vogal alcançaram, na maioria dos dados, um percentual de acerto bastante próximo ao dos dados que não passaram por manipulação; e (2) ASSIMILAÇÃO A UMA CATEGORIA, no caso das manipulações envolvendo vozeamento, especialmente para alveolares e velares, pois os dados parecem não ter tido uma distribuição tão regular. Nesse quesito, a pesquisadora chama atenção para o fato de que a

situação em jogo reflete que ambos os membros do contraste (surdo e sonoro) são percebidos como exemplares igualmente desviantes ou igualmente aceitáveis da categoria da L1.

3.3.2 Perozzo (2013)

Perozzo (2013) investigou a percepção auditiva de consoantes oclusivas não vozeadas do inglês, realizadas sem soltura de ar audível em posição final de palavra, por aprendizes brasileiros pertencentes a dois níveis de proficiência (básico e intermediário) em língua inglesa. Os objetivos do trabalho residiram em (a) verificar qual ponto de articulação (bilabial, alveolar, velar) favorecia a percepção dos contrastes e analisar se as vogais mediais [i, ɪ, æ] exerciam efeito sobre a percepção das consoantes em posição final; (b) examinar se a instrução explícita sobre o fenômeno fonético em questão era efetiva para a percepção acurada das diferenças de ponto de articulação; e (c) constatar se o nível de proficiência dos participantes impactava a percepção dos referidos contrastes.

A pesquisa, ao longo de suas etapas, contou com 17 acadêmicos (8 mulheres e 9 homens) do primeiro semestre do curso de graduação em Letras da Universidade federal do Rio Grande do Sul, com idade média de 20 anos, matriculados em duas turmas de língua inglesa referentes ao primeiro semestre da grade curricular. Dos acadêmicos em questão, 5 apresentavam nível básico de proficiência na língua inglesa, enquanto que 12 apresentavam nível intermediário. Além disso, 10 acadêmicos fizeram parte do grupo experimental (que recebeu instrução explícita sobre o fenômeno fonético analisado) e 7 acadêmicos fizeram parte do grupo-controle (que não receberam qualquer tipo de instrução sobre o tema).

O estudo foi conduzido com base em duas tarefas perceptuais, uma de discriminação categórica (formato ABX) e outra de identificação, que empregavam a apresentação auditiva de palavras formadas pelas consoantes iniciais [p, b, s, m, l, w], pelas vogais nucleares [i, ɪ, æ] e pelas consoantes finais [p^ɹ, t^ɹ, k^ɹ]. Os estímulos foram gravados por três falantes nativos de inglês americano, provenientes da região oeste dos Estados Unidos, que tinham, em média, 21 anos de idade e estavam residindo no Brasil há menos de dois anos.

Através dos experimentos utilizados, o autor menciona que (a) os pontos de articulação labial e velar foram os que obtiveram maiores índices de acuidade, e as vogais [ɪ] e [æ] conferiram maior acuidade na percepção dos pontos de articulação em questão; (b) o grupo experimental obteve maiores índices de acuidade quando comparado ao grupo-controle (o que parece revelar um efeito

positivo da instrução explícita); e (c) os acadêmicos pertencentes ao nível intermediário de proficiência obtiveram maiores índices de acuidade do que os pertencentes ao nível básico. Ao tratar dos achados segundo os padrões de assimilação do *PAM-L2*, Perozzo (2013) ressalta os casos de (1) ASSIMILAÇÃO A UMA CATEGORIA, nos casos de baixa acuidade perceptual envolvendo o ponto alveolar e os contrastes cuja vogal nuclear era [i]; (2) ASSIMILAÇÃO A DUAS CATEGORIAS, nos casos envolvendo altos índices de acuidade, a exemplo do ponto bilabial e de contrastes cujas vogais nucleares eram [ɪ] e [æ]; e (3) DIFERENÇA DE ADEQUAÇÃO À CATEGORIA, especialmente, nos casos envolvendo a capacidade perceptual do ponto velar e da vogal [ɪ] e [i] em posição nuclear.

Observamos, com base nas duas dissertações examinadas, a de Albuquerque (2012) e a de Perozzo (2013), que os resultados obtidos, além de serem interpretados segundo suas teorias de base, também foram vislumbrados pelos pesquisadores segundo as previsões assimilatórias sugeridas por Best e Tyler (2007), mesmo que de maneira relativamente incipiente.

Neste momento, acabamos de cobrir diferentes estudos que tiveram como objeto de estudo a percepção de sons linguísticos de línguas não nativas. No que se refere aos artigos resenhados, observamos que parte deles vislumbraram seus resultados segundo os padrões assimilatórios propostos no *PAM-L2* (seção 3.1), e parte não (seção 3.2). As duas dissertações apresentadas na seção 3.3, diferentemente daquelas que foram somente citadas, se empenharam em acomodar seus resultados sob os padrões assimilatórios elencados em Best e Tyler (2007). Com base nesse levantamento, passamos à seção 3.4, que estabelece considerações sobre tais estudos na medida em que propõe fornecer respostas ao primeiro objetivo enumerado nesta tese, o de verificar a exaustividade das previsões de assimilação empreendidas no *PAM-L2*.

3.4 Considerações acerca dos estudos apresentados

Ao relatar os estudos trazidos nesta seção, conforme mencionado no início do capítulo, pretendemos verificar se os trabalhos referentes à percepção dos sons da fala estrangeira, reportados em periódicos internacionais e nacionais, e também por meio de dissertações, se utilizam das predições do *PAM-L2*. De modo mais importante, procuramos verificar se o *PAM-L2*, no que concerne às suas previsões de assimilação, daria conta de abarcar os fenômenos investigados e os resultados obtidos.

De modo a ficar mais sintetizado visualmente ao leitor, retomaremos, no Quadro 2, a seguir, os artigos e as dissertações em formato resumido (e em ordem cronológica), elencando quatro características importantes de cada trabalho. Na primeira coluna, indicamos qual é o trabalho que estará sendo analisado; na segunda coluna, mencionamos qual é a unidade de análise adotada no respectivo estudo; na terceira coluna, averiguamos se o *PAM-L2* consegue acomodar o objeto de investigação de cada estudo, em termos de interpretação de seus resultados; na quarta coluna, apontamos se o estudo faz referência explícita às previsões assimilatórias estabelecidas pelo *PAM-L2*; e na quinta coluna, elencamos o padrão assimilatório correspondente aos achados dos estudos – no caso daquelas pesquisas que não fazem referência explícita às previsões assimilatórias, sugerimos qual seria o possível padrão de assimilação do *PAM-L2* (ou possível conjunto de padrões assimilatórios) que viria a explicar os dados encontrados. A partir deste levantamento, poderemos tecer algumas considerações iniciais sobre a adoção ao *PAM-L2* e sobre como tal modelo se relaciona com as pesquisas conduzidas.

Quadro 2 – Relação de trabalhos acerca da percepção da fala não nativa (continua)

Trabalho	Unidade de análise	Acomodação no <i>PAM-L2</i>	Previsões do <i>PAM-L2</i>	Padrão assimilatório do <i>PAM-L2</i>
Best e Hallé (2010)	Gesto articulatório	Sim	Não faz referência	- DAC - AUC
Escudero e Chládková (2010)	Pista acústica	Sim	Não faz referência	- ADC - DAC - AUC
Gilichinskaya e Strange (2010)	Pista acústica	Sim	Faz referência	- ADC - AUC - NCC
Bundgaard-Nielsen, Best e Tyler (2011a)	Segmento	Sim	Não faz referência	- ADC - AUC
Bundgaard-Nielsen, Best e Tyler (2011b)	Segmento	Sim	Faz referência	- ADC - DAC - AUC
Escudero e Vasiliev (2011)	Pista acústica	Sim	Faz referência	- ADC - AUC
Escudero e Williams (2011)	Pista acústica	Sim	Faz referência	- ADC - AUC
Strange <i>ET AL.</i> (2011)	Segmento	Sim	Faz referência	- ADC - AUC
Albuquerque (2012)	Pista acústica	Sim	Faz referência	- ADC - AUC

Quadro 2 – Relação de trabalhos acerca da percepção da fala não nativa (continuação)

Trabalho	Unidade de análise	Acomodação no PAM-L2	Previsões do PAM-L2	Padrão assimilatório do PAM-L2
Antoniou, Best e Tyler (2012)	Segmento	Sim	Faz referência	- ADC - DAC - AUC
Bohn e Best (2012)	Segmento	Sim	Faz referência	- ADC
Darcy e Krüger (2012)	Segmento	Sim	Faz referência	- ADC - DAC - AUC
Escudero e Williams (2012)	Pista acústica	Sim	Não faz referência	- ADC - DAC - AUC
Escudero, Simon e Mitterer (2012)	Pista acústica	Sim	Não faz referência	- ADC - DAC
Fabra e Romero (2012)	Segmento	Sim	Não faz referência	- ADC - DAC - AUC
Shafiro <i>ET AL.</i> (2012)	Segmento	Sim	Não faz referência	- ADC - DAC - AUC
Antoniou, Best e Tyler (2013)	Segmento	Sim	Faz referência	- ADC - DAC - AUC - ANC
Patihis, Oh e Mogilner (2013)	Segmento	Sim	Não faz referência	- ADC - DAC - AUC
Perozzo (2013)	Pista acústica	Sim	Faz referência	- ADC - DAC - AUC
Dufour <i>ET AL.</i> (2014)	Segmento	Sim	Não faz referência	- ADC - DAC
Elvin, Escudero e Vasiliev (2014)	Pista acústica	Sim	Não faz referência	- ADC - DAC - AUC
Escudero, Sisinni e Grimaldi (2014)	Pista acústica	Sim	Não faz referência	- ADC - DAC - AUC
McKelvie-Sebileau e Davis (2014)	Pista acústica	Sim	Faz referência	- ADC - DAC - AUC
Tyler <i>ET AL.</i> (2014)	Segmento	Sim	Faz referência	- ADC - ANC - NCC
Stölten, Abrahamsson e Hyltenstam (2014)	Pista acústica	Sim	Não faz referência	- ADC - DAC

Quadro 2 – Relação de trabalhos acerca da percepção da fala não nativa (continuação)

Trabalho	Unidade de análise	Acomodação no <i>PAM-L2</i>	Previsões do <i>PAM-L2</i>	Padrão assimilatório do <i>PAM-L2</i>
Williams e Escudero (2014)	Pista acústica	Sim	Não faz referência	- ADC - DAC
Yang e Fox (2014)	Pista acústica	Sim	Não faz referência	- ADC - DAC - AUC
Alves e Zimmer (2015)	Pista acústica	Sim	Não faz referência	- ADC - DAC - AUC
Holliday (2016)	Segmento	Sim	Não faz referência	- ADC - DAC - AUC

Fonte: elaborado pelo autor.

Nota: Sinais convencionais utilizados:

ADC – assimilação a duas categorias

DAC – diferença de adequação à categoria

AUC – assimilação a uma categoria

ANC – ambos não categorizáveis

NCC – não categorizado vs. categorizado

Ao que podemos observar a partir do Quadro 2, notamos que as unidades de análise são variadas, englobando o GESTO ARTICULATÓRIO, O SEGMENTO e a PISTA ACÚSTICA. Ressaltamos que, em nenhum dos estudos, os autores defendem explicitamente qual é a sua unidade de análise, mas, por vezes, mencionam sua designação ao longo do texto. Vale acentuar que, curiosamente, o primitivo GESTO ARTICULATÓRIO – unidade fônica basilar que constitui o primitivo do *PAM-L2* – aparece somente em uma das pesquisas, a de Best e Hallé (2010), e não com foco nos padrões preditivos. Em linhas gerais, tal unidade praticamente não é referenciada nos estudos relacionados, diferentemente do segmento e da pista acústica. Isso nos causa grande estranhamento, uma vez que, segundo os postulados do *PAM* e do *PAM-L2*, são as diferenças gestuais, e não segmentais, que exprimem as relações entre as previsões de assimilação.

Em um referencial teórico aos moldes do *PAM* e do *PAM-L2*, o não cumprimento de uma diretriz básica, a saber, a concepção de que é o gesto articulatório a unidade real e abstrata da fala, implica uma grande fragilização ao modelo. Se nem nas pesquisas de que participam Catherine Best e Michael Tyler, podemos encontrar coerência epistemológica entre unidade de análise e referencial teórico, então parece que um dos seus eixos basilares não está sendo levado a sério. Deve restar claro que não estamos criticando a adoção a um ou a outro termo em função do mérito de seus primitivos. De fato, nosso pedido de alerta vai na direção de haver paralelismo entre teoria de base e unidade de análise.

Certamente, não podemos esperar que haja uma descrição gestual de todas as rotinas articulatórias que empreendemos durante a fala (assim como ocorre na tabela de sons da fala do *IPA*), até porque não se trata de uma descrição trivial, mas é importante que, mesmo com limitações procedimentais, procuremos informar de que ponto estamos partindo e sob que tipo de primitivo de análise estamos atuando.

Uma situação que também nos intriga é o fato de haver menção ao termo *SEGMENTO* em diversas pesquisas resenhadas. Essa designação pode ser usada tanto para refletir os pressupostos de uma fonologia clássica, estando para a noção de *SEGMENTO FONOLÓGICO*, como para se referir a unidades da fonética articulatória, por exemplo, estando para a noção de *SEGMENTO FONÉTICO*. Isso gera uma insegurança teórica para que possamos compreender o plano de fundo de tais pesquisas, que, muitas vezes, não se faz claro. Tal impasse fica evidente a partir das ponderações da própria Best (1995), que, no texto de lançamento do *PAM*, afirma que se vale do termo *SEGMENTO*, mesmo concebendo que os contrastes sonoros nas línguas sejam mantidos por diferenças de configurações gestuais.

Em relação aos estudos que se valem de *PISTAS ACÚSTICAS* em termos de unidade, parece-nos que pode haver uma motivação fonética que talvez venha a justificar a adoção de tal propriedade. Faz-se possível que, para os autores que se utilizam desse arcabouço acústico, seria através de aspectos acústicos que alcançaríamos a unidade gestual, aproximando a informação acústica à informação gestual⁸¹. Deixamos claro, todavia, que nossa ponderação se trata de uma especulação, e pode não refletir fielmente a intenção dos autores (até porque eles não esclarecem suas razões ao se utilizar do parâmetro acústico). No que tange às pesquisas que se utilizam do *segmento*, hipotetizamos que a escolha por essa unidade ocorra em virtude de ser amplamente difundida, o que viria a acarretar certa simplicidade de entendimento aos leitores, já que a noção de *segmento* é bastante familiar àqueles que se debruçam sobre a percepção da fala.

Quanto à acomodação desses estudos no arcabouço procedimental do *PAM-L2*, ou seja, quanto à possibilidade de o objeto de investigação de tais estudos ser tratado à luz de um modelo perceptual como o *PAM-L2*, parece-nos possível que o modelo de Best e Tyler (2007) dê conta de acolher empiricamente os fatos fônicos encontrados⁸². Sob nosso ponto de vista, não deve haver impedimentos, portanto, para que os dados encontrados em cada um dos trabalhos sejam abarcados

⁸¹ Pode vir a ser o caso de a adoção ao gesto segundo um tratamento acústico-articulatório reduzir em parte esse problema, conforme será tratado no capítulo 4 desta tese.

⁸² Referimo-nos, aqui, ao emprego das previsões de assimilação.

pelo *PAM-L2*. A grande razão para isso decorre do fato de que, em algum nível, todas as pesquisas estipulam uma comparação entre dois sistemas (ou subsistemas) fônicos distintos, gerando verificações de caráter dedutivo, as quais podem ser tranquilamente delineadas no âmbito do *PAM-L2*, uma vez que o modelo é essencialmente preditivo no que compete à discriminação de pares de unidades da fala estrangeira.

Alguns estudos reportados não fazem referência explícita às previsões estabelecidas pelo *PAM-L2*. Isso significa que, embora os autores dos respectivos trabalhos estejam tratando da categorização de sons da fala, eles não alinham seus achados aos padrões de assimilação que o *PAM-L2* institui. O que ocorre em alguns casos, como no estudo de Alves e Zimmer (2015), é haver a EXPLICITAÇÃO de aspectos teóricos veiculados pelo modelo, mas os dados obtidos não serem interpretados segundo os padrões de assimilação – em vez disso, os autores conduzem sua discussão mais com o propósito de apontar o papel de fatores acústico-articulatórios na apropriação de padrões sonoros do que conectar os resultados às previsões, caso a caso, oferecidas pelo *PAM-L2*. Em nosso entendimento, não parece haver inconsistência nesta postura, já que, além de se sustentar sobre padrões assimilatórios, o *PAM-L2* promove uma excelente problematização acerca da interação do sistema fônico da língua-fonte e do da língua-alvo.

Segundo o que supomos anteriormente, parece que os dados advindos de tais estudos podem, de fato, ser tratados com base nos padrões de assimilação definidos pelo *PAM-L2*. Esta averiguação fica evidente nos casos em que os trabalhos são regidos pelas previsões de assimilação, uma vez que os próprios autores indicam qual seria o padrão assimilatório (ou conjunto de padrões assimilatórios) em jogo nos casos de discriminação. Contudo, em relação aos estudos que não fazem referência explícita às previsões de assimilação do *PAM-L2*, concluímos, em primeira análise, que a quantificação dos dados apresentados e o seu direcionamento frente à percepção dos ouvintes compreendem argumento empírico suficiente para conectarmos tais resultados às previsões de assimilação. Muito embora esse caminho seja de cunho indutivo, já que somente após os resultados somos capazes de atribuir cada caso de discriminação a um padrão assimilatório, ele parece ser efetivo para afirmarmos, ainda que com todas as ressalvas lógicas, que as previsões de assimilação não só conseguem abarcar os dados encontrados, como também se mostraram exaustivas (pelo menos para os estudos reportados nesta seção) no que concerne ao tratamento dos índices de acuidade e erro advindos dos testes perceptuais utilizados.

Cabe trazer à tona, neste momento, uma problemática de bastante vigor teórico no que concerne à exaustividade das previsões assimilatórias do *PAM-L2*, fomentada por Santos (2014), em sua tese de doutorado, e no trabalho posterior, de Santos e Rauber (2016). Tais trabalhos, ao discorrer sobre o *PAM-L2*, mencionam que as previsões de assimilação estipuladas por esse modelo se reservam aos casos em que o inventário fônico da língua-alvo é mais complexo (numeroso) do que o da *L1*. Para respaldar sua afirmação, as pesquisadoras recorrem aos postulados de Escudero e Boersma (2002), em que os autores comentam sobre um contraste binário na *L2* ser assimilável a três categorias na *L1*, padrão assimilatório que os estudiosos chamam de ASSIMILAÇÃO A MÚLTIPLAS CATEGORIAS (*MULTIPLE-CATEGORY ASSIMILATION*). Essa situação envolveria a percepção de um contraste binário da *L2* como sendo assimilável a três categorias da *L1*. Conforme Santos (2014) relata, esse padrão oferece sua contribuição na medida em que traz evidências acerca da existência de múltiplas categorias passíveis de assimilação em uma *L1* com maior inventário vocálico a partir de um sistema de *L2* com menor número dessas categorias.

Ainda que a ponderação de Santos (2014) seja extremamente oportuna e pertinente, julgamos que não viria a ser desabonadora ao *PAM-L2*, uma vez que esse novo padrão, assimilação a múltiplas categorias, poderia ser reinterpretado à luz de padrões já existentes no modelo de Best e Tyler (2007). Nesse caso, se [e] ou [i] do espanhol é interpretado ora como [e], ora como [i], e ora como [i] do holandês, então temos que uma dada categoria do espanhol corresponde a outra dada categoria do holandês, configurando uma circunstância de assimilação a uma categoria propriamente dita. Também, pode ser o caso de esse padrão chamado de “assimilação a múltiplas categorias” estar relacionado a uma diferença de adequação à categoria. Assim, se [e] ou [i] do espanhol (*L2*) é interpretado ora como [e], ora como [i], e ora como [i] do holandês, mas com diferenças de adequabilidade (o que responderia pela variabilidade nas decisões dos participantes), então temos que uma dada categoria do espanhol corresponde, com diferentes graus de adequabilidade, a outra dada categoria do holandês, configurando uma circunstância de diferença de adequação à categoria.

Devemos chamar atenção, no entanto, para o fato de que soa um pouco curioso que, em uma tarefa de discriminação, sejam observadas respostas dessa natureza. As respostas envolvidas nas tarefas de discriminação são do tipo “o último elemento ouvido é igual ao primeiro”, ou “apenas o primeiro elemento é diferente”. Daí surgem, minimamente, duas possibilidades: (a) ou se trata de contraste binário na *L1*, mas cada categoria do contraste é apresentada ao participante em isolado, como em uma tarefa de identificação ou assimilação, e, dadas todas as respostas possíveis, ele marque

três opções da L_2 – isto é, não se trata de uma tarefa de discriminação; e (b) ou se trata de um contraste binário na L_1 , apresentado sob a forma de uma tarefa de discriminação, mas, durante a análise dos resultados, os pesquisadores encontraram a conexão desse contraste na L_1 com três categorias distintas na L_2 . Diante dessas ponderações, então, em (a), não se pode afirmar que um contraste binário na L_2 é apresentado ao participante, mas que apenas uma categoria vocálica da L_2 lhe é apresentada; e, em (b), há de ressaltar que as três categorias da L_1 a que os participantes conectaram o contraste binário da L_2 não foram escolhidas simultaneamente. Em qualquer um dos casos, conforme deve ter ficado claro, não precisamos ir além das previsões de assimilação encontradas em Best e Tyler (2007).

Sobre o impacto do tamanho do inventário-alvo nas assimilações às categorias do inventário-fonte, Santos (2014) fornece o exemplo relatado por Escudero e Boersma (2002), em que holandeses aprendizes de espanhol perceberam as vogais anteriores do espanhol, [i] e [e], a partir das categorias nativas [i, ɪ, e]. Nesse caso, duas categorias da L_2 , com menor sistema vocálico, foram assimiladas a três categorias da L_1 , com maior sistema vocálico, diferindo do que se tem encontrado na literatura em termos de resultados, já que os estudos geralmente utilizam uma língua-alvo com maior número de categorias e uma língua-fonte com menor número delas, o que restringira as previsões de assimilação. Nesse quesito, fazemos duas ressalvas. A primeira, de cunho articulatório, é a de que quanto maior for o número de categorias vocálicas nativas, mais opções de assimilação existem, já que uma grande quantidade de categorias vocálicas nativas implica pequenas regiões acústico-articulatórias⁸³, resultando que suas fronteiras sejam menos espaçadas e os que os julgamentos perceptuais acerca das assimilações sejam mais permissivos. A segunda, de cunho epistemológico, é a de que os padrões de assimilação do *PAM-L2* são designados como *PREVISÕES*, visto que se trata de um modelo com propósitos dedutivos, e não indutivos. Ao se partir das premissas, ou seja, das previsões de assimilação para estipular as direções das categorias, o tamanho dos inventários em jogo torna-se indiferente, pois sempre haverá a possibilidade de uma categoria não nativa ser assimilada a duas categorias não nativa. Por exemplo, a categoria vocálica [e] do espanhol (L_2) pode ser assimilada à [e] ou à [ɪ] do holandês (L_1), assim como essa mesma categoria espanhola pode ser assimilada a outras categorias holandesas – par a par, todas as possibilidades na relação L_{NN-L_1} são contempladas. Se o caminho inverso for feito (indutivo), ou seja, se a partir dos resultados encontrados se desejar

⁸³ Afinal de contas, o contraste de cinco categorias vocálicas em um dado idioma ocupa, *GROSSO MODO*, a mesma área no trato vocal referente ao contraste de quinze categorias vocálicas em outro idioma.

estipular previsões de assimilação, então as opções de assimilação podem vir a ser bem numerosas. Além de não garantir a verdade lógica que reside no modelo, o percurso indutivo se torna bastante oneroso a qualquer modelo preditivo, pois pode incitar tentativas de criação de padrões do tipo ASSIMILAÇÃO A N CATEGORIAS.

De toda forma, conforme os resultados observados nos estudos apresentados ao longo do capítulo, a percepção dos ouvintes quanto aos contrastes testados pode ser alocada minimamente em um dos padrões assimilatórios previstos pelo *PAM-L2*, o que faz do modelo um excelente preditor de assimilação perceptual de distinções sonoras não nativas⁸⁴.

Apesar de não contemplado no quadro resumitivo apresentado neste capítulo, existe ainda um ponto digno de discussão que se evidencia em alguns dos estudos aqui relacionados: o fato de, em alguns trabalhos, o *PAM-L2* estar sendo utilizado para explicar a categorização de sons em um contexto que seria de *L1*, em detrimento de um contexto de *L2*. Ou seja, em algumas pesquisas, parece que a comunidade científica ignora, de alguma maneira, as diretrizes (ainda que questionáveis) de Best e Tyler (2007) acerca de o *PAM-L2* se voltar para a percepção em *L2* e, não, em *L1*.

Alguns trabalhos contam apenas com monolíngues [como é o caso de Escudero e Vasiliev (2011), Escudero e Williams (2011), Bohn e Best (2012), Williams e Escudero (2014)] em tarefas de percepção de um idioma não nativo. Nesses estudos, percebemos maior referência dos autores ao *PAM* do que ao *PAM-L2*, mas não julgamos que tais pesquisas sejam inadequadas ao nosso levantamento bibliográfico, uma vez que as mesmas predições que constam no *PAM* também aparecem no *PAM-L2* – e um dos objetivos desta tese é verificar, justamente, se as predições se fazem exaustivas aos dados encontrados.

Outros trabalhos constituem-se de falantes monolíngues e bilíngues [como os de Antoniou, Best e Tyler (2013) e McKelvie-Sebileau e Davis (2014)], mas o idioma testado nas tarefas de percepção foge ao conhecimento dos participantes. Ou seja, mesmo havendo participantes bilíngues, a língua a que eles eram expostos perceptualmente lhes era desconhecida (ma'di e tailandês). Novamente, notamos maior evidência do *PAM* em relação ao *PAM-L2*, mas fatores desenvolvimentais, que necessitam de uma abordagem via *PAM-L2*, são recuperados pelos autores.

⁸⁴ Vale lembrar, entretanto, que o modelo não se concentra na formação de novas categorias, o que é de extrema importância quando concebemos o desenvolvimento fônico de uma *L2*. Trata-se, portanto, de um aspecto lacunar no *PAM-L2*, de modo que teceremos considerações sobre esse tema no capítulo 4.

O estudo de Bundgaard-Nielsen, Best e Tyler (2011b) se configura como um caso bastante peculiar. Os autores testaram, na Austrália, japoneses que residiam no país por um período menor do que três meses – o que, em primeira instância, corresponde a um contexto de *L2*. No entanto, os participantes, antes de residirem na Austrália, tiveram contato com a língua inglesa no Japão somente através de instrução explícita – o que reflete (numa visão tradicional) um ambiente de *LE*. Dessa forma, embora os autores designem sua amostra como aprendizes de inglês como *L2* (por estarem imersos no cenário em que o inglês é falado como idioma nativo), devemos atentar para o fato de que os mesmos já estudavam inglês em seu país de origem durante um tempo médio de 9 anos⁸⁵.

As pesquisas de Escudero, Simon e Mitterer (2012), Fabra e Romero (2012), Escudero, Sisinni e Grimaldi (2014) e Alves e Zimmer (2015) testam a percepção sonora de aprendizes de *LE*, mas não fazem referência explícita às previsões assimilatórias do *PAM-L2*. Embora mencionem, em algum momento, os postulados do *PAM-L2*, a amostra utilizada em tais estudos diz respeito a um cenário de *LE*, distanciando-se, em algum nível, do escopo do *PAM-L2*, que preconiza um ambiente de *L2*. Vale apontar que a pesquisa de Strange *ET AL.* (2011) é a única que utiliza uma amostra de aprendizes de inglês como *LE* e, ao mesmo tempo, faz referência explícita às previsões de assimilação do *PAM*, o que soa intrigante, pois tal modelo se presta para ouvintes sem experiência com o idioma-alvo.

Destacamos que, de acordo com nossa análise, as únicas pesquisas que contaram com participantes aprendizes de uma *L2* foram as de Antoniou, Best e Tyler (2012), Darcy e Krüger (2012) e Escudero e Williams (2012). É bem verdade que, nos dois primeiros estudos, também se utilizaram falantes monolíngues, mas esses vieram a ser designados como grupos-controle, pois o foco estava nos bilíngues, isto é, em como a experiência linguística impacta a percepção de contrastes em uma *L2*. Tais estudos apresentam, também, referência explícita às previsões do *PAM-L2*. No trabalho de Escudero e Williams (2012), embora não haja referência explícita aos padrões assimilatórios do *PAM-L2*, a amostra é toda constituída de peruanos e espanhóis aprendizes de holandês como *L2*, indo ao encontro do que instaura o modelo (aprendizes de *L2*), caso suas previsões tivessem sido endereçadas.

Das ponderações que fizemos sobre os trabalhos aqui apresentados, a que mais nos intriga diz respeito ao fato de que, salvo algumas poucas exceções, a referência ao gesto articulatório como unidade sonora da percepção fônica é quase inexistente nos estudos reportados (e isso também é

⁸⁵ Todavia, esse fato não parece ser relevante para Best e Tyler (2007), uma vez que não consideram como realmente efetivas as experiências linguísticas em *LE*. Essa questão será problematizada no quarto capítulo desta tese.

observado nos textos seminais dos próprios *PAM* e *PAM-L2*). Nesse sentido, verificamos que, nem nas pesquisas em que a própria Catherine Best é coautora, o primitivo gestual se mostra operante. É extremamente relevante ter em mente que, a partir do momento em que um modelo perceptual fundamentado sobre os preceitos da fonologia gestual é adotado, *DEVER-SE-IA FICAR COM TODA A BAGAGEM QUE ELE TRAZ*. Vale reiterar que as próprias previsões de assimilação elencadas por Best e Tyler (2007) já se alheiam ao gesto, ou seja, os autores não operacionalizam tal unidade em um de seus momentos principais, que seria a proposição das previsões de assimilação. Em outras palavras, parece-nos, no mínimo, questionável e contrassensual se basear em um modelo como o *PAM-L2* e prescindir do gesto articulatório como primitivo de análise, dado que o vínculo do modelo à fonologia articulatória é um de seus atributos substanciais (novamente, retomamos que, no texto de 1995, Catherine Best explicita que passa a utilizar uma unidade que corresponde ao segmento). A omissão ou a substituição da unidade gestual acarreta não só um choque epistemológico a diversas pesquisas aqui relacionadas, mas também evidencia a falta de compromisso em assumir o modelo na sua totalidade.

Utilizar-se do *PAM-L2* parece ser muito atraente quando os inventários sonoros nativos e não nativos são postos lado a lado. Contudo, valer-se de tal modelo deveria ir muito além do jogo lógico que envolve a comparação de sistemas fônicos no momento em que são evidenciadas as previsões de assimilação perceptual: tem-se que estar alinhado à sua vertente filosófica (com implicações psicológicas) acerca de como percebemos e temos acesso à realidade que nos circunda, e se deve compactuar com seu construto fonológico, que advém necessariamente de uma base gestual. Incongruências teóricas⁸⁶ à parte, em cumprimento ao objetivo (a) desta tese, sugerimos que o *PAM-L2*, em termos de suas previsões de assimilação, seja um modelo suficiente e exaustivo para o tratamento dos dados referentes à assimilação perceptual da fala não nativa, ainda que seu substrato teórico seja alvo de indagações.

3.5 Resumo do capítulo

O escopo deste capítulo correspondeu ao elenco de artigos internacionais e nacionais, bem como de dissertações realizadas no Brasil, que abordaram a percepção de sons linguísticos não

⁸⁶ Tais incongruências, que podem ser encontradas nos próprios textos de 1995 e 2007, serão problematizadas no capítulo 4 e compreendem grande parte da motivação para que esta tese seja proposta, visando a elencar alternativas a essas questões problemáticas.

nativos no período entre 2010 e 2016. Relacionamos 13 trabalhos (11 artigos científicos e 2 dissertações de mestrado) que vislumbraram seus achados à luz do *PAM-L2*, em termos das previsões de assimilação do modelo, e apresentamos 16 estudos que não se comprometeram em focar tais padrões (porém, mencionaram o modelo). Independentemente de se valerem do construto do *PAM-L2*, quando os resultados obtidos foram conjecturados sob o prisma das previsões de assimilação dele decorrentes, em todos os estudos encontramos associações com seus padrões preditivos, o que garante ao *PAM-L2* exaustividade nesse mérito.

Examinando os trabalhos, observamos o emprego de diferentes unidades de análise, tais como *PISTA ACÚSTICA* e *SEGMENTO*, e de diferentes casos de assimilação perceptual, sendo *ASSIMILAÇÃO A DUAS CATEGORIAS*, *ASSIMILAÇÃO A UMA CATEGORIA* e *DIFERENÇA DE ADEQUAÇÃO À CATEGORIA* os padrões mais frequentes. Ao sintetizar as pesquisas, notamos algo que nos intrigou: apenas um estudo contou com a unidade gestual de análise, e a referência a esse primitivo não foi em termos procedimentais (seu encaixe nos padrões de assimilação), mas em termos teóricos – comparar um fenômeno sonoro via análise gestual e via geometria de traços. Isso reforça a hipótese de que, embora o *PAM-L2* disponha de um primitivo próprio, o comprometimento com sua unidade gestual aparentemente não está sendo manifestado pela comunidade científica que se utiliza do modelo e, tampouco, pelos seus proponentes.

Lembramos que os estudos listados neste capítulo serão retomados no capítulo 5 para uma discussão acerca dos procedimentos empíricos de testagem das categorizações fônicas. Neste momento, passaremos ao capítulo 4, que se ocupará da problematização dos eixos basilares do *PAM-L2*.

4 PARA ALÉM DO *PAM-L2*: UMA NOVA FUNDAMENTAÇÃO TEÓRICA

Nosso objetivo ao discorrer sobre o *PAM-L2*, neste capítulo, é o de problematizá-lo com o intuito de prover esclarecimentos sobre as alterações pelas quais defendemos que o modelo deva passar para que possa atingir consistência interna e coerência. Conforme salientamos no capítulo introdutório, há algumas incongruências teóricas, que residem na base no *PAM-L2*, que devem ser repensadas quando o assunto é a percepção fônica de línguas não nativas. Reconhecemos que, evidentemente, não seremos exaustivos quanto às sugestões de modificação, mas tentaremos contemplar seus aspectos mais fundamentais, em função de que nossas ponderações configuram um primeiro passo na tentativa de reconcepção do modelo, de modo que discussões vindouras sejam ensejadas a partir de nossa reflexão.

Se o *PAM-L2* se resumisse apenas às predições de assimilação perceptual descritas no capítulo 2, em primeira análise, consideramos que seriam poucos os ajustes necessários ao modelo. Por exemplo, bastaria que essas previsões fossem vislumbradas à luz de um primitivo gestual. Entretanto, não é só de previsões que o modelo é formado: de antemão, podemos dizer que há todo um embasamento teórico que o sustenta, o qual pode e deve ser discutido sobre o ponto de vista FONOLÓGICO, em que o modelo assume como primitivo fônico o gesto articulatorio, e sobre o ponto de vista FILOSÓFICO (com implicações psicológicas), que imprime o realismo direto de James Gibson e sua abordagem ecológica à percepção. Igualmente importante, o modelo deve ser fundamentado quanto ao expediente COGNITIVO, discussão que não só escapa à sua essência de sua teorização, mas se mostra irrelevante⁸⁷ para os proponentes do *PAM-L2*.

Serão contemplados, neste capítulo, os objetivos (b), (c), (d), e (e), explicitados na introdução da tese. A seção 4.1 [que se reserva ao objetivo (b), o de discutir e sugerir explicitamente a necessidade de recorrer a um aparato cognitivo no que concerne à percepção de aspectos fônicos de línguas não nativas] pretende tratar o evento perceptual à luz de um enfoque cognitivista, de acordo com o qual

⁸⁷ Tal proposição fica clara ao observarmos este excerto de Best (1995, p. 179): “[...] *the direct realist perspective [...] rejects the notion that neural and/or cognitive mechanisms are needed to interpret/process/decode/draw inferences from the speech signal*” (“[...] a perspectiva realista direta [...] rejeita a noção de que mecanismos neurais e/ou cognitivos são necessários para interpretar/processar/decodificar/extrair inferências a partir do sinal da fala”). Na mesma direção, encontramos este outro trecho de Best (1995, p. 175): “*Special cognitive mechanisms for handling mental representations and indirect inferencing are not needed, because the flow of stimulation provides a rich and reliable source of direct information about the world [...]*” (“Mecanismos cognitivos especiais para gerenciar representações mentais e inferências indiretas não são necessários, pois o fluxo de estimulação fornece uma fonte rica e confiável de informação direta sobre o mundo [...]).

representações mentais se fazem cruciais para a apropriação de conhecimento. Na seção 4.2 [em que se contempla o objetivo (c), o de questionar o primitivo de análise do *PAM-L2*], traremos argumentos para a adoção da unidade gestual segundo o tratamento de Albano (2001), em detrimento do tratamento de Browman e Goldstein (1989, 1992) e de Fowler (1986). A seção 4.3 [que remete ao objetivo (d), o de problematizar a alocação filosófica do *PAM-L2*] apresenta a abordagem realista indireta para o fenômeno perceptual fônico, opondo-se à perspectiva original do *PAM-L2*. Por fim, de maneira a cumprir o objetivo (e), o de discorrer sobre questões pertinentes ao modelamento da percepção fônica em línguas não nativas, a seção 4.4 ocupa-se de ponderações adicionais referentes ao modelo tradicional, tais como a discussão sobre a suposta dicotomia *L2* vs. *LE*, a influência grafêmica e a formação de novas categorias no que compete à percepção fônica não nativa.

Os itens aqui elencados retratam características fundamentais de uma nova configuração do *PAM-L2*, no que diz respeito à percepção fônica de línguas não nativas, e poderão servir de base para um novo modelo independente futuro, cujo ponto de partida seja o *PAM-L2*. A comparação dessa nova caracterização do modelo com o próprio *PAM-L2* é, portanto, inevitável. Assim, poderemos passar de um modelo que (a) DISPENSA REPRESENTAÇÕES MENTAIS; (b) SE UTILIZA DO GESTO ARTICULATÓRIO SEGUNDO O TRATAMENTO DE BROWMAN E GOLDSTEIN (1989, 1992) E DE FOWLER (1986); (c) É REALISTA DIRETO E SE FUNDAMENTA SOBRE UMA ABORDAGEM ECOLÓGICA para um modelo que (a') DEMANDA NECESSARIAMENTE REPRESENTAÇÕES MENTAIS; (b') SE UTILIZA DO GESTO ARTICULATÓRIO EM CONSONÂNCIA COM O TRATAMENTO DE ALBANO (2001); (c') É REALISTA INDIRETO. Em decorrência dessas modificações fundamentais, deixaríamos de ter um modelo que (d) CONSIDERA APENAS A *L2*; (e) NÃO PONDERA SOBRE OS ASPECTOS GRAFÊMICOS DA PERCEPÇÃO; e (f) SÓ PREVÊ CASOS DE ASSIMILAÇÃO para vislumbrar uma proposta que (d') CONSIDERA TANTO A *L2* COMO A *LE* (EM TERMOS TRADICIONAIS); (e') PROBLEMATIZA OS ASPECTOS GRAFÊMICOS DA PERCEPÇÃO; e (f') DISCORRE SOBRE CASOS DE ASSIMILAÇÃO E DE FORMAÇÃO DE NOVAS CATEGORIAS.

4.1 A percepção como evento cognitivo e seu embasamento nas neurociências

Afirmar que um evento perceptual qualquer que experienciamos corresponde a um fenômeno cognitivo não é uma proposição consensual. Isso vai depender justamente do fenômeno a que nos voltamos e do aporte teórico que concebemos. Para J. Gibson (1966, 1986), como relatamos no segundo capítulo desta tese, no que compete à percepção visual, a informação do estímulo se encontra estruturada na luz ambiente e é capaz de veicular todas as propriedades invariantes necessárias para que tenhamos acesso aos objetos do mundo sem que seja preciso recorrer a

processos mentais ou inferências. O psicólogo, no entanto, se faz bastante cauteloso ao destacar que os fenômenos que descreve não se relacionam à linguagem, já que a apreensão de tal conhecimento, cultural e socialmente convencionado em sua essência, não pode ser compreendida como direta (aos moldes da percepção visual). Assim, temos que a percepção visual pode, segundo J. Gibson (1966, 1986) dispensar o aparato cognitivo, mas o mesmo não é verdadeiro para percepção do conhecimento linguístico.

Também, conforme mencionamos previamente, Fowler (1986), em virtude de seguir J. Gibson (1966, 1986), entende que muitos eventos perceptuais prescindem de uma tarefa cognitiva. Nesse caso, suas ponderações dizem respeito exclusivamente ao seu objeto de investigação, a fala. Todavia, de modo menos prudente do que J. Gibson (1966, 1986), a pesquisadora preconiza que a fala é captada, sim, sem a mediação de processos cognitivos: seu argumento é o de que o sinal acústico, sendo disponibilizado em um meio informacional estruturado pelas atividades vocais, transmite, de fato, suas propriedades diretamente a um sistema de percepção. Por conseguinte, Best (1995) e Best e Tyler (2007) advogam a favor do mesmo posicionamento e levam o construto de Fowler (1986) à percepção dos sons de línguas não nativas⁸⁸.

Em consonância com J. Gibson (1966, 1986), mas diferentemente de Fowler (1986), Best (1995) e Best e Tyler (2007), julgamos que, no âmbito do conhecimento linguístico, uma postura realista direta não pode e não tem a contribuir. Nesse quesito, estabelecemos que o aparato cognitivo não só se faz presente, mas é de fundamental importância para que acessemos e apreendamos as informações veiculadas na linguagem. Para tanto, respaldamo-nos nas explicações fornecidas pelas neurociências acerca do expediente cognitivo da percepção, e as tomaremos como evidência para postular que a percepção do material linguístico deve ser primordialmente dependente de mecanismos mentais.

⁸⁸ Essa postura fica clara a partir do seguinte excerto: *“This model is founded on a direct realist approach to perception, in which articulatory gestures are assumed to be the perceptual primitives for speech perception, including the perception of non-native speech. It is further assumed that evidence of those gestures is available in the speech signal and is directly picked up by listeners, without need for recourse to innate knowledge of the vocal tract or to indirect cognitive processing of raw acoustic information”* (BEST, 1995, p. 199). “Este modelo se sustenta sobre a abordagem realista direta da percepção, em que se presume que os gestos articulatórios sejam os primitivos perceptuais para a percepção da fala, incluindo a percepção da fala não nativa. Presume-se também que evidências de tais gestos encontram-se disponíveis no sinal da fala e são diretamente captadas pelos ouvintes, sem necessidade de recorrer a um conhecimento inato do aparato vocal ou a um processamento cognitivo indireto da informação acústica bruta”.

Nosso embasamento nas neurociências, particularmente na neurociência cognitiva (MILLER; VANDOME; MCBREWSTER, 2009; KANDEL *ET AL.*, 2014; KEMMERER, 2014) e na psicologia cognitiva (BRAISBY; GELLATLY, 2005; QUINLAN; DYSON, 2010; GAZZANIGA *ET AL.*, 2012; EYSENCK, 2015; STERNBERG, R.; STERNBERG, K., 2016) para endereçar os fenômenos perceptuais humanos reside na ATUAÇÃO INQUESTIONÁVEL DO CÉREBRO, bem como de seu funcionamento, nas atividades que executamos em nossa vida cotidiana. Essas áreas têm avançado teórica e experimentalmente no que tange à cognição humana e oferecem evidências contundentes acerca de temas complexos que circunscrevem a percepção e os processos mentais referentes à linguagem.

Para iniciar nossa exposição favorável ao expediente cognitivo do evento perceptual, recorremos a um excerto do próprio J. Gibson (1966), em que o autor trata de algumas das relações entre o percebedor e as concessões que o ambiente fornece:

Em resumo, o observador humano APRENDE a detectar o que chamamos de valores ou significados das coisas, percebendo suas características distintivas, colocando-as em categorias e subcategorias, notando suas similaridades e diferenças e mesmo estudando-as para seus próprios objetivos, independentemente de aprender o que fazer com elas. Toda essa discriminação, que é maravilhosa, deve ser baseada inteiramente na educação de sua ATENÇÃO às sutilezas da informação invariante do estímulo⁸⁹ (GIBSON, J., 1966, p. 285, grifo nosso).

J. Gibson (1966) destaca que o observador aloca características distintivas dos objetos do mundo (e, por conseguinte, de seus valores e significados) em CATEGORIAS E SUBCATEGORIAS. Concordamos com o estudioso no quesito de que categorias e subcategorias existem; no entanto, é preciso atentar para o fato de que, segundo o psicólogo, o processo de detecção da informação não é permeado por categorias já existentes: primeiramente, ocorre a detecção direta das informações e, posteriormente, suas características são categorizadas. Em outras palavras, quando o percebedor olha para os objetos do mundo, ele extrai diretamente (sem processos cognitivos ou inferenciais) as informações referentes aos dados e, somente depois, possivelmente através de processos que escapam à posição realista direta, tais dados são alocados em categorias. É precisamente nesse mérito que divergimos de J. Gibson (1966).

⁸⁹ No original: “In short, the human observer learns to detect what have been called the values or meanings of things, perceiving their distinctive features, putting them into categories and subcategories, noticing their similarities and differences and even studying them for their own sakes, apart from learning what to do about them. All this discrimination, wonderful to say, has to be based entirely on the education of his attention to the subtleties of invariant stimulus information” (GIBSON, J., 1966, p. 285).

Advogamos a favor da posição de que, no próprio evento de olharmos para os objetos do mundo, estão presentes as categorias que operam sobre as informações que chegam até nós. Dito de outra forma, em nossa perspectiva, as categorias de que dispomos se ocupam da observação dos objetos do mundo, permeiam-na e envolvem-nos, diferentemente da posição de J. Gibson (1966), em que tais categorias não fazem parte do processo de percepção, o qual se dá através da detecção direta das propriedades invariantes do estímulo. Julgamos primordial ressaltar que, em momento nenhum, insinuamos dizer que as categorias (sejam quais forem) de que dispomos sejam inatas. Tampouco presumimos que o evento perceptual incorra em alguma espécie de reducionismo (já que são nossas categorias prévias que operam na percepção e, aparentemente, preconizariam perceber o que “cabe” nessas categorias), pois defendemos que tal fenômeno ocorra de maneira iterativa, na medida em que os dados são observados. Assim, qualquer objeto que não se adeque a uma categoria prévia tem potencialidades de fomentar a criação de uma nova categoria. Portanto, entendemos que, uma vez que somos seres dotados de um aparato cognitivo, temos a capacidade de racionalizar a informação que a nós se apresenta (BORGHI, 2005; BRAISBY; GELLATLY, 2005; MILDNER, 2008; GAZZANIGA *ET AL.*, 2012; AMARAL, 2014; KANDEL *ET AL.*, 2014; ORMROD, 2015)⁹⁰. O argumento que defendemos encontra respaldo na explicação que as neurociências oferecem ao fenômeno perceptual e também se sustenta filosoficamente, conforme veremos mais adiante.

Outro possível alvo de questionamentos relativo à abordagem de J. Gibson (1966) é a utilização de termos que, juntos, em alguma medida, carregam acepções ligadas a processos mentais ou intelectuais: APRENDIZAGEM e ATENÇÃO. Podemos dizer que soa, no mínimo, intrigante que o estudioso recorra a tais construtos na sua exposição sobre a detecção dos dados, já que a relação entre percebedor e estímulo se baseia na extração direta de elementos invariantes, os quais dariam conta de causar a experiência perceptual sem compromisso com aspectos cognitivos, conforme o próprio autor, em suas obras de maior repercussão, faz questão de frisar. No entanto, James Gibson não responde sozinho pela tarefa de tocar na questão da aprendizagem perceptual, mas o faz em conjunto com sua exponencial interlocutora acadêmica e esposa: a também psicóloga Eleanor Gibson. O tema foi tratado pelo casal Gibson, especialmente, em um artigo publicado na metade do século passado

⁹⁰ Conforme deverá ficar claro na seção 4.2, que trata da unidade gestual da percepção fônica, o que Albano (2001) pontua sobre o fato de aprendermos os padrões fônicos executando-os é garantido através da coadunação de elementos externos (rotinas articulatórias, que estão para os objetos do mundo) e internos (gramática fônica, que está para o conhecimento internalizado). Isso só é possível em função de nossa capacidade de racionalizar a informação e fazer uso de nosso conhecimento prévio. Através dessa postura, atingimos coerência entre o expediente cognitivo e o fônico.

(GIBSON, J.; GIBSON, E., 1955). Há de se deixar claro, todavia, que grande parte do construto referente à aprendizagem perceptual na via ecológica e realista direta é empreendida majoritariamente por Eleanor Gibson (1963, 1969, 1970, 1982, 1984) e, eventualmente, pela estudiosa em colaboração com outros pesquisadores (GIBSON, E.; RADER, 1979; GIBSON, E.; PICK, 2000).

Seguindo os pressupostos inicialmente incitados por James Gibson, o casal Gibson se reporta à aprendizagem perceptual afirmando que esta corresponde ao fato de o percebedor estar, progressivamente, em contato mais próximo com o ambiente, configurando um *APERFEIÇOAMENTO (IMPROVEMENT)* da percepção (GIBSON, J.; GIBSON, E., 1955). Essa prática, segundo E. Gibson (1963)⁹¹, visa a reduzir as generalizações sobre o estímulo, aumentar a precisão da discriminação de variáveis que se encontram presentes, de fato, na estimulação e, além disso, detectar fatores relevantes ou propriedades distintas que não haviam sido caracterizadas até então. Vale lembrar que a proposição que subjaz a tal perspectiva se apoia muito mais nos atributos do estímulo do que no aparato racional do percebedor.

Ainda no que respeita à aprendizagem perceptual, porém, agora em termos da linguagem falada e escrita, E. Gibson e J. Gibson (1972) enfatizam que “o observador humano aprende a extrair informação dos constituintes da fala e da escrita, mas essa informação é diferente⁹²” (p. 507), no sentido de que não é a mesma que advém diretamente do ambiente. Para os psicólogos, a aprendizagem do mundo a partir da fala (e, posteriormente, da escrita) tem a ver com um processo muito mais complexo do que aquele tido em primeira mão (ou diretamente). Tal posição ratifica os postulados de J. Gibson (1966, 1986) apresentados no segundo capítulo desta tese. De qualquer modo, especulamos que recorrer a questões concernentes à atenção e à aprendizagem fora de um aporte cognitivista pode ser limitante, dada a excelência das pesquisas teórica e experimental que são conduzidas na área e dado o acúmulo de conhecimento que esta tem congregado ao longo das últimas décadas – especialmente, após o início da revolução cognitiva em 1950⁹³.

Ressalvas gibsonianas à parte, destacamos nossa inclinação ao construto de Eric Kandel, neurocientista austríaco naturalizado norte-americano. O estudioso explana que a percepção se

⁹¹ Segundo explicam Adolph e Kretch (2015), a teoria de Eleanor Gibson é consistente com a abordagem de James Gibson, mas não é idêntica a ela. As autoras apontam que, embora eles fossem um casal e compartilhassem de muitos argumentos e ideias sobre a aprendizagem perceptual, escreveram apenas cinco artigos juntos.

⁹² No original: “a human observer learns to extract information from the constituents of spoken and written language, but this is information of a quite different sort” (GIBSON, E.; GIBSON, J., 1972, p. 507).

⁹³ Miller (2003) traça um percurso histórico bastante interessante no que compete à revolução cognitiva.

refere essencialmente a um processo de *CONSTRUÇÃO COGNITIVA* que depende não somente do estímulo externo, mas também do aparato mental do sujeito que experiencia o evento perceptual, ou seja, é amplamente dependente da organização de sistemas sensoriais e motores no encéfalo (KANDEL, 2014). Nessa perspectiva, Gazzaniga *ET AL.* (2012) pontuam que a percepção que temos do mundo não funciona como uma câmera ou como um gravador de áudio, que captam de maneira fiel e passiva as propriedades existentes nos estímulos com que nos deparamos. De acordo com os pesquisadores, o que nós vemos, ouvimos, degustamos, tocamos ou cheiramos é resultado de *PROCESSOS CEREBRAIS QUE CONSTROEM AS EXPERIÊNCIAS PERCEPTUAIS*. Resta notório, portanto, que o evento de perceber os objetos do mundo só pode se concretizar se houver mecanismos cognitivos que operam na construção da percepção.

Gazzaniga *ET AL.* (2012) afirmam que o encéfalo não processa os estímulos brutos (em termos de visão ou de qualquer outro sentido humano), implicando que eles sejam necessariamente traduzidos em sinais químicos e elétricos para que o cérebro possa interpretá-los. Assim, as diferentes propriedades do mundo físico são codificadas ou traduzidas por diferentes padrões de impulsos neurais, fenômeno designado como *CODIFICAÇÃO SENSORIAL* (GAZZANIGA; *ET AL.*, 2012; GARDNER; JOHNSON, 2014). A codificação sensorial inicia com a transdução, em que os receptores sensoriais – neurônios especializados – produzem impulsos neurais quando recebem estimulação física ou química. A informação em jogo é, então, transmitida ao cérebro sob a forma de impulsos neurais. Paralelamente, Gardner e Johnson (2014) esclarecem que o sistema nervoso extrai somente determinados fragmentos de informação de cada estímulo, vindo a ignorar outros, e, a título de ilustração, destacam que ondas eletromagnéticas de distintas frequências são *RECEBIDAS*, mas *cores* são *VISTAS*. Esse exemplo denota que a mesma informação que está no mundo sob a forma de propriedades físicas, objetivas, também está em nosso aparato cognitivo sob uma forma mental, subjetiva, já analisada. Em outras palavras, existe toda uma codificação do estímulo para que ele possa ser interpretado à luz de um evento perceptual:

Recebem-se ondas de pressão de objetos que vibram em diferentes frequências, mas se ouvem sons, palavras e músicas. Compostos químicos são encontrados flutuando no ar ou na água, mas são experimentados como odores e sabores. Cores, tons, odores e sabores são criações mentais, construídas pelo encéfalo a partir da experiência sensorial. Eles não existem como tal fora do encéfalo (GARDNER; JOHNSON, 2014, p. 401).

Com vistas ao que expusemos, Gardner e Johnson (2014), ao tocar no quesito da percepção visual, por exemplo, evidenciam que ela se trata de uma CRIAÇÃO do cérebro, o que vai ao encontro dos postulados de Kandel (2014) e de Gazzaniga *ET AL.* (2012). Os autores salientam que, no momento em que damos procedência ao evento perceptual relacionado à visão, tal processo se vale dos sinais de entrada extraídos da imagem sobre a retina e os toma como base. Todavia, Gardner e Johnson (2014) advertem que aquilo que é “visto com os olhos da mente” se expande para além do que se encontra nos sinais de entrada. Faz-se oportuno enfatizar que para haver a percepção visual, “o encéfalo utiliza a informação que extraiu previamente como base para conjecturas fundamentadas em dados – inferências da percepção acerca do estado do mundo” (GARDNER; JOHNSON, 2014, p. 392).

Gilbert (2014) sustenta que a percepção visual requer uma interação entre retina, núcleos talâmicos e várias regiões do córtex cerebral. Os limites da visão são definidos pela retina, em que há o enquadro de pequenos detalhes, a discriminação de movimentos minúsculos e a habilidade de detectar contrastes finos e diferenças sutis em termos do comprimento de onda concernente à luz refletida. Tais informações, provenientes de cenas complexas, são adquiridas pelo córtex visual, o qual as analisa, separando-as de seu plano de fundo. O pesquisador afirma que esse processo presume tanto a análise de propriedades locais (orientação, direção, cor, etc.) como a integração dessas propriedades em todo o espaço de visão.

Nosso objetivo, ao discorrer sobre as ponderações de neurocientistas e psicólogos cognitivos acerca da percepção (eventualmente, acerca da percepção visual), não é o de estabelecer um tratado fisiológico sobre ela, mas o de caracterizar, em linhas gerais, o evento perceptual como sendo operacionalizado por mecanismos cerebrais de ordem cognitiva. Assim, a ideia que resume nossa elucubração, e que nos difere da abordagem gibsoniana à percepção, é a de que os sistemas perceptuais são construídos para realizar inferências acerca do mundo, uma vez que, conforme argumentam Gardner e Johnson (2014), contrariamente ao entendimento intuitivo (baseado na experiência pessoal), as percepções não são cópias diretas do mundo à volta do indivíduo. Nessa perspectiva, conforme esclarece Amaral (2014), a percepção é entendida como uma ABSTRAÇÃO, e não como uma réplica da realidade.

Ao que os autores apresentados indicam, o evento perceptual é fruto de operações cognitivas subsidiadas pelo encéfalo. O objeto que existe no mundo externo é construído em nosso aparato mental e, passa, portanto, a ter um correlato interno. Nesse âmbito, recorreremos a dois conceitos extremamente importantes, o de HABILIDADE MENTAL e o de REPRESENTAÇÃO MENTAL, os quais se farão úteis

ao longo de nossa exposição. *HABILIDADE MENTAL* OU *ATIVIDADE MENTAL*, de acordo com Gazzaniga *ET AL.* (2012), tem a ver com a mente⁹⁴ em ação, ou seja, relaciona-se 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. Por conseguinte, 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. A representação mental pode ser analógica, quando, por exemplo, trazemos à mente a imagem de um objeto, ou simbólica, que exprime o conceito ou nosso conhecimento sobre o objeto e não sobre suas propriedades físicas (GAZZANIGA *ET AL.*, 2012).

Referimos anteriormente que, no momento em que experienciamos um evento perceptual, já se encontram presentes categorias que operam sobre as informações que acessamos. Paralelamente, defendemos que tais categorias exercem influência sobre a percepção dos objetos do mundo. Essas duas afirmações encontram sustentação nas ponderações de Gardner e Johnson (2014), em que os estudiosos ilustram que o objeto ou o fenômeno que percebemos é sempre uma combinação do *ESTÍMULO* em si e da *MEMÓRIA* que ele evoca. Daí colocamos em evidência a noção de que somos capazes de racionalizar as informações que chegam até nós, pois operamos mentalmente sobre o dado perceptual que se expõe, interpretando-o “no contexto de nossas experiências passadas e de objetivos atuais” (GARDNER; JOHNSON, 2014, p. 413). Convém acrescentar que a ação cognitiva de atuar sobre as informações dispostas no ambiente, analisar e refletir acerca de suas formas e funções, antecipar os tipos de demandas possíveis e ativar conhecimentos e habilidades prévias para a execução de novas tarefas configura uma capacidade cognitiva altamente complexa (MCLAUGHLIN, 1987).

Chamamos atenção para o fato de que operar mentalmente sobre as informações que acessamos, racionalizá-las, não implica afirmar que os estímulos sejam pobres ou deficientes – designa tão somente a nossa capacidade cognitiva de processar e interpretar tais informações de acordo com nosso conhecimento de mundo. Defendemos, assim, que o ambiente em nosso entorno é capaz de fornecer uma vasta fonte de informação multimodal sobre os objetos e os fatos do mundo,

⁹⁴ Seguimos a definição de mente de Kosslyn e Koenig, de acordo com a qual “a mente é aquilo que o cérebro faz” (1995, p. 4). Dessa maneira, segundo postulam Gazzaniga *ET AL.* (2012, p. 5), “o cérebro físico habilita a mente”, e ambos são inseparáveis.

mas que esses só fazem sentido se vislumbrados sob a ótica de nossas experiências prévias, as quais, como se relata na literatura, impregnam o evento perceptual.

Até este momento, concebemos que, ao perceber determinado objeto, estamos nos valendo do estímulo a que ele se refere e também do que já experienciamos previamente (GARDNER; JOHNSON, 2014). Resumidamente, tal objeto é construído pelo encéfalo (KANDEL, 2014) e se torna representado mentalmente em nosso aparato cognitivo sob a forma de um conceito, de uma imagem, de uma ação ou de um significado, por exemplo. Ou seja, temos uma representação mental, que é interna, de um objeto do mundo externo (GAZZANIGA *ET AL.*, 2012). Quando uma criança entra em contato com um animal a ela desconhecido, como um elefante, por exemplo, sua percepção do mamífero pode ser a de que se trata de um ser vivo acinzentado de grande porte. Segundo Schacter e Wagner (2014), na medida em que a criança tem mais oportunidades de observar elefantes, aquela representação mental prévia passa a se ajustar, incorporando outros fragmentos de informações: esses mamíferos vivem em determinado ambiente, manifestam um padrão comum de comportamento, emitem sons distintos e podem participar de apresentações circenses. Essa explanação confere à representação mental grande flexibilidade e demarca sua propriedade ADAPTATIVA e COMPLEXA, uma vez que o conceito se mostra bastante elástico e se encontra permeado por inúmeras características e variáveis.

Julgamos que o aporte teórico a que nos remetemos e a que recorremos até o momento, a propósito de haver representações mentais criadas a partir de nossas experiências perceptuais, tenha sido suficiente para tornar contundente a atuação do aparato cognitivo no processo de percepção. A partir do que discorremos sobre as representações mentais, no sentido de serem adaptativas e complexas, acentuamos que tais características podem ser atribuídas também às representações que formamos no que concerne à linguagem (conhecimento linguístico). É exatamente nesse mérito que, recentemente, teóricos da área (DE BOT; LOWIE; VERSPOOR, 2007; LARSEN-FREEMAN, 1997; LARSEN-FREEMAN; CAMERON, 2008; BECKNER; BYBEE, 2009; BECKNER; *ET AL.*, 2009; CORNISH; TAMARIZ; KIRBY, 2009; ELLIS; LARSEN-FREEMAN, 2009; SCHOENEMANN, 2009) propuseram desafiar o paradigma⁹⁵ e impulsionaram o tratamento do estudo da linguagem via teoria dos SISTEMAS ADAPTATIVOS COMPLEXOS [*COMPLEX ADAPTIVE SYSTEMS*] (NICOLIS; PRIGOGINE, 1989; BAK; CHEN, 1991; LEWIN,

⁹⁵ Referências paradigmáticas relacionadas compreendem, entre outros, os trabalhos de Chomsky e Halle (1968), Chomsky (1972) e Kiparsky (1982).

1993; HOLLAND, 1995, 2006; KAUFFMAN, 1995; CILLIERS, 1998; ADAMI, 2002)], comumente referida através da sigla *CAS*.

Beckner *ET AL.* (2009) explicam que, ao interpretar a língua como um *CAS*, presumimos a adaptabilidade do sistema e a combinação de diversos agentes que interagem. Os autores citam os participantes de uma determinada comunidade de fala como sendo os múltiplos agentes de um *CAS*, mas podemos também supor que as próprias unidades da língua sejam múltiplos agentes – unidades que respondem pela sua estrutura sonora, pela formação de suas palavras e pela organização dessas palavras em sentenças, por exemplo. No tocante à adaptabilidade, os pesquisadores mencionam a influência das interações passadas sobre o comportamento dos falantes, no sentido de que interações atuais e prévias conjuntamente oportunizam o comportamento futuro. Novamente, podemos inferir que as representações mentais acerca do conhecimento linguístico também sejam adaptativas na medida em que vão se incorporando a elas fragmentos de informações ao longo das experiências perceptuais que administramos. Os estudiosos destacam que as estruturas linguísticas (e, evidentemente, suas representações mentais) “emergem a partir de padrões inter-relacionados de experiência, interação social e mecanismos cognitivos” (BECKNER; *ET AL.* 2009, p. 2).

Nossa apreciação é a de que as descrições feitas por neurocientistas e psicólogos cognitivos acerca da configuração das representações mentais podem se alinhar apropriadamente com os pressupostos da teoria dos *CAS* no âmbito do conhecimento linguístico⁹⁶. Nessas circunstâncias, as unidades fonológicas, morfológicas e sintáticas das línguas, assim como suas relações semânticas e pragmáticas, podem ser instauradas representacionalmente através das experiências perceptuais prévias e novas, considerando informações de que já dispomos e informações atualizadas em conjunto com seu significado e contexto de uso⁹⁷.

Dada a exposição que fizemos até então, sublinhamos a coerência epistemológica que atingimos quando associamos o construto neurocognitivo e psicológico, apresentado anteriormente, ao construto linguístico acerca de como vislumbramos as representações concernentes aos símbolos da linguagem: a partir dos eventos perceptuais, podemos formar representações mentais que são

⁹⁶ Frisamos que o enfoque dado pela teoria dos *CAS* ao estudo das línguas é demarcado por aquilo que chamamos de ABORDAGENS LINGÜÍSTICAS BASEADAS NO USO (*USAGE-BASED APPROACHES TO LANGUAGE*), as quais fazem parte de um movimento que se opõe à hipótese inatista veiculada na tradição gerativa da linguística.

⁹⁷ Conforme deverá ficar claro na próxima seção, a unidade gestual que adotamos também prevê compromisso com aspectos simbólicos e físicos referentes às informações da fala.

racionalizadas por nosso aparato cognitivo e que se moldam ao longo de nossas experiências e interações com os objetos do mundo, os quais são abundantes em suas informações.

Podemos dizer, portanto, que o primeiro ponto em que divergimos de Best (1995) e de Best e Tyler (2007) corresponde à natureza do evento perceptual. Para os pesquisadores, conforme relatamos no segundo capítulo desta tese, a percepção é um processo que ocorre diretamente entre os sentidos (tomados como sistemas perceptuais) e o fluxo de informação que emana do objeto a ser percebido, sem pressupor uma operação cognitiva. Nossa posição, diferentemente, estabelece que a percepção se faz, por excelência, enquanto um processo mediado pela cognição, em que os objetos externos que estão disponíveis no mundo (ricos em suas propriedades) são construídos internamente pelo encéfalo (GAZZANIGA *ET AL.*, 2012; KANDEL, 2014). Vale realçar que, de forma nenhuma, queremos dar a entender que o mundo real, aquele “fora de nossas cabeças”, inexistente, conforme ficará claro na seção 4.3.

Nossa preocupação nesta seção foi traçar um entendimento sobre a percepção como um evento essencialmente cognitivo, bem como estipular a configuração de nossas representações mentais acerca dos objetos do mundo. A próxima seção irá endereçar com especificidade os aspectos fônicos das representações linguísticas de que dispomos, além de sugerir como esses se relacionam com os demais entes do conhecimento linguístico sob um viés cognitivista.

4.2 A unidade gestual da percepção fônica

A partir do que expusemos até o momento, deve ter ficado evidente a nossa concepção de que as representações mentais linguísticas se instauram a partir da íntima conexão entre um PROCESSO COGNITIVO GERAL DE RACIONALIZAÇÃO DO CONHECIMENTO e a RIQUEZA DAS INFORMAÇÕES QUE SE ENCONTRAM DISPOSTAS NO MUNDO. Assim, a informação da fala a que temos acesso, por exemplo, é vislumbrada segundo o nosso aparato cognitivo (o qual está permeado por aspectos subjetivos e atribui valores aos objetos do mundo circundante), ao mesmo tempo em que é capaz de conferir à às nossas representações toda a variabilidade e a complexidade que reside nos objetos ao nosso entorno.

Argumentamos, pois, a favor da existência de um conhecimento mental quanto a unidades fônicas das línguas (LUND, 2001; LECLERCQ; ZIMMERMANN, 2005; STYLES, 2005; KRAMER; WIEGMANN; KIRLIK, 2007; MILDNER, 2008), e, com base na literatura em neurociências e psicologia cognitiva a que fizemos referência ao longo da seção 4.1, podemos afirmar com

propriedade que a nossa relação com o ambiente influencia as nossas representações mentais. Em outras palavras, decorrente desse posicionamento é a proposição de que a experiência com os objetos do mundo (aqui, de natureza linguística) exerce influência sobre as representações mentais que se encontram armazenadas em nosso aparato cognitivo, o que confere ao conhecimento linguístico uma via de mão dupla: atuamos sobre os dados linguísticos que acessamos e eles atuam sobre nós.

Para os propósitos desta tese, defendemos a posição de que os fatos sonoros que ocorrem nas línguas se refletem a partir de uma UNIDADE GESTUAL (BROWMAN; GOLDSTEIN, 1989, 1992, 1995; SALTZMAN; MUNHALL, 1989; FOWLER; SALTZMAN, 1993; FOWLER, 2007; ALBANO, 2001, 2007; GOLDSTEIN; FOWLER, 2003; GOLDSTEIN; BYRD; SALTZMAN, 2006; SILVA, 2002). É essencialmente sobre a caracterização gestual das representações linguísticas que esta seção propõe discorrer.

A unidade sonora a que iremos nos remeter nesta seção é, então, de caráter GESTUAL. Explicitamos, na seção 2.1.2, sua teorização sob o prisma articulatório, segundo o que concebem Catherine Browman, Louis Goldstein e Carol Fowler. Nesta seção, também faremos um apanhado sobre essa unidade, mas à luz de um tratamento acústico-articulatório, ao que concebe Eleonora Albano em sua obra de 2001, intitulada O GESTO E SUAS BORDAS: ESBOÇO DE FONOLOGIA ACÚSTICO-ARTICULATÓRIA DO PORTUGUÊS BRASILEIRO. De antemão, sublinhamos que não se trata de concepções diferentes do que sejam gestos: o aporte articulatório e o aporte acústico-articulatório são especificidades delineadas pelos autores no intento de descrever os fatos sonoros das línguas a que se voltam, encargo que impõe consequências para a sua investigação empírica (informação verbal)⁹⁸.

Fonética e fonologia já eram vistas, no estruturalismo, como entidades distintas, conforme linguistas tais quais Edward Sapir e Nicolai Trubetzkoy. Mais recentemente, no período gerativista, ambas continuam sendo tratadas como dois entes separados. Nesse cenário, Albano (2001) afirma que esses dois campos de estudo poderiam restar separados se essa distinção se resumisse ao método de investigação que empregam. A fonética, por um lado, vislumbra o som linguístico enquanto uma realidade física, respondendo por uma ciência natural. A fonologia, por outro lado, endereça o som linguístico nas vias de uma realidade semiológica, cabível a uma ciência humana.

Essa cisão entre fonética e fonologia, segundo a autora, foi potencialmente motivada pela ideia da distintividade dos sons da fala, os quais foram atrelados ao conceito de fonema. Dada a

⁹⁸ Esclarecimento realizado pela professora Adelaide Silva, em um curso que ministrou sobre fonologia gestual em 2015.

grande elaboração dessa unidade fonológica discreta em território europeu e norte-americano na primeira metade do século passado (estruturalismo linguístico), convergindo em uma proposta aclamada pelos teóricos da área, contribuições acerca das gradiências fonéticas corresponderiam a uma ameaça à luta pelo aspecto contrastivo veiculado pela noção fonêmica. Tal “bravata” representaria aceitar uma variação contínua nas distinções lexicais, cuja implicação seria a de que um significante alicerçado sobre diferenças potencialmente infinitas jamais poderia estabelecer contrastes de significado. Nesse mérito, a pesquisadora elenca que:

Os fatos coletados pelo afã documentador dos foneticistas falavam a favor dos fonólogos. O esforço empreendido pela ASSOCIAÇÃO FONÉTICA INTERNACIONAL, desde a sua fundação em 1886, para normatizar a transcrição fonética através de um alfabeto universal, revelou que as categorias de vogais e consoantes existentes nas línguas do mundo não eram tantas nem tão díspares, limitando-se a umas poucas centenas, mesmo quando incluídos os sons mais raros. Considerando-se o número de categorias possíveis se os contínuos acústicos envolvidos na fala fossem divisíveis em porções apenas discrimináveis, esses achados dão inequívoco respaldo à universalidade do caráter discreto das unidades fônicas (ALBANO, 2011, p. 13, grifo da autora).

Albano (2001) menciona que a necessidade de caracterizar unidades discretas e abstratas no plano empírico, juntamente com o fato de que um contraste fônico nem sempre preserva a mesma função entre contextos diferentes, acabou por instaurar uma unidade que impunha nebulosidade à separação entre fonética e fonologia: tal unidade era o ALOFONE (variante posicional). Nesse contexto, uma tarefa difícil seria a de ora separar o alofone do fonema e ora separar o alofone dos efeitos da articulação. Uma alternativa apostada por Chomsky e Halle (1968) foi a de tratar, primeiramente, o alofone como sendo alterações categóricas dos fonemas ou como sendo ajustes gradientes dessas categorias. Em um segundo momento, os alofones estariam suscetíveis ao controle motor voluntário, denominados alofones extrínsecos (como o padrão de *VOT* positivo encontrado nas oclusivas não vozeadas do inglês), ou estariam suscetíveis à dinâmica motora involuntária, denominados alofones intrínsecos (como a anteriorização de uma oclusiva velar diante de uma vogal anterior alta). Contudo, essa saída não foi exitosa em alguns aspectos e, em conjunto com outras evidências acerca do comportamento dos sons linguísticos em seus inventários nas mais diversas situações, poderia ter se tornado impraticável.

Para dar conta de abarcar a explicação de fenômenos relacionados, chega-se a um nível de abstração que pode sobrecarregar fortemente o rol das operações fonológicas, limitando-as em larga medida. Nesse sentido, Albano (2001, p. 27) adiciona que aqueles processos naturais a partir do plano

fonético podem “gerar distâncias indefinidamente grandes entre o nível subjacente e o nível superficial” quando aplicados de modo cumulativo à mesma forma, o que resulta em complexas questões sobre como delimitar a profundidade de uma derivação.

Posicionando-se contra uma gramática derivacional, Albano (2001) explica que uma gramática inteiramente rasa (sem derivações) é a única compatível com a proposição de que, na sincronia, existem somente processos fônicos “vivos”, ou seja, variações quantitativas da pronúncia que resultam dos mecanismos de sobreposição e redução da magnitude dos gestos. Para a autora, convincentes análises de fenômenos como o *SANDHI* externo prenunciam mudar a concepção tradicional de que um léxico com diversas variantes de uma dada forma não é mais do que um repositório idiossincrático. A estudiosa defende que só há dois modos como os sons de fala podem se encontrar em variação: aquele que decorre da elasticidade própria da mecânica de produção, capturando sobreposições e reduções em termos da magnitude dos gestos, e aquele relacionado a questões lexicais e/ou gramaticais, considerando o realinhamento e o redimensionamento dos gestos presentes no léxico.

Albano (2001) argumenta veementemente que fonética e fonologia devem ser comensuráveis. Essa comensurabilidade pode ser garantida através da adoção a uma unidade que venha a mediar *ROTINAS ARTICULATÓRIAS*, de ordem física, e *CONHECIMENTO INTERNALIZADO* no que concerne às relações entre tais rotinas, de ordem abstrata, que se instanciam em uma *GRAMÁTICA FÔNICA* – trata-se, portanto, da unidade gestual. Assim, se espera que o conhecimento internalizado advenha da tarefa efetiva de realizar a ação articulatória, culminando no que a estudiosa designa como “aprender fazendo” (p. 30). Ainda que não explicitada pela autora, a conjunção do expediente físico e do expediente abstrato na unidade de análise que a pesquisadora empreende é muito paralela aos pressupostos relacionados à posição cognitivista por nós adotada nesta tese e abordada no capítulo anterior, em que a formação de conhecimento também se dá pela integração desses dois expedientes. A própria noção de aprendizagem através da execução de uma tarefa, defendida por Albano (2001), se mostra em plena consonância com os postulados cognitivistas que defendemos, os quais implicam que as informações do mundo externo, ricas e multimodais em sua essência, sejam coadunadas com nosso aparato mental para que possamos desempenhar as mais diversas tarefas.

Complementarmente, no que diz respeito às variáveis do trato vocal, as quais são responsáveis por caracterizar determinado gesto acústico-articulatório⁹⁹, Albano (2001) entende que elas se encontram em dupla relação com os articuladores. No plano abstrato, as variáveis restam funcionalmente unidas sob o articulador envolvido mais diretamente com as constrictões que efetua, e, no plano concreto, tais variáveis oportunizam o movimento do articulador em questão, ou seja, facultam a realização de uma tarefa mecânica. Esse é um argumento que nos leva à ponderação de que fonética e fonologia não se encontram em domínios informacionais distintos, mas passam a fazer parte de um único expediente, o ÂMBITO FÔNICO, em que passam a ser indissociáveis. Em outras palavras, o que é físico e concreto passa a ter um correspondente mental e abstrato, situação que oportuniza a dependência recíproca entre os expedientes fonético e fonológico. Logo, o termo FÔNICO propõe conglobar fatos fonéticos e fatos fonológicos, implodindo a fronteira existente entre as duas áreas. Um fenômeno fonético é um fenômeno fonológico e vice-versa, e essa íntima relação se instancia a partir do gesto. Mais uma vez, as observações de Albano (2001) vão completamente ao encontro da perspectiva cognitivista que adotamos, a qual prevê que fatores mentais/abstratos e fatores motores/físicos devam permanecer integrados, em uma íntima relação de retroalimentação.

É importante notar que a autora não se mostra contra a concepção de fonologia, e tampouco sugere que a fonética deva existir em detrimento dela. Isso fica claro a partir de situações de duas naturezas, uma teórica e outra mais empírica, levantadas pela pesquisadora. No quesito teórico, como evidencia ao longo de sua obra, Albano (2001) advoga a favor do rompimento da fronteira existente entre fonética e fonologia, e sustenta seu posicionamento com argumentos históricos. No quesito prático, aquele observável, a autora afirma que “uma habilidade INQUESTIONAVELMENTE FONOLÓGICA é a de julgar se um logatoma é ou não uma palavra possível na língua” (ALBANO, 2001, p. 28, grifo nosso).

Albano (2001) avança teoricamente em suas constatações sobre a unidade gestual que adota e, ao se lançar no empreendimento de uma fonologia de cunho acústico-articulatório (no lugar de uma fonologia puramente articulatória, como aquela de Catherine Browman e Louis Goldstein), a estudiosa passa a revisar o modelo norte-americano e o adapta em termos de, pelo menos, três

⁹⁹ Fazemos menção ao termo “GESTO ACÚSTICO-ARTICULATÓRIO” para deixar claro que estamos nos baseando no tratamento (acústico-articulatório) que Albano (2001) fornece à unidade gestual em questão. Como deverá ficar claro a partir da leitura das próximas páginas, a inclusão da palavra “ACÚSTICO” ao referido termo não é arbitrária, mas, sim, motivada pelo importante papel que os parâmetros acústicos desempenham na designação de um gesto fônico. Desse modo, não apenas questões puramente articulatórias se mostram notáveis na unidade gestual, mas questões acústicas se revelam oportunas – daí, passamos de um gesto articulatório para um gesto acústico-articulatório.

circunstâncias, as quais permeiam seu trabalho. A primeira se relaciona ao acréscimo da noção DE ELO AUDITIVO E ACÚSTICO ao gesto articulatório, incorporando critérios quânticos e de dispersão adaptativa. A segunda tem por base a assunção de que os lugares de PROJEÇÃO SIMBÓLICA do gesto são as suas BORDAS. A terceira designa que o REALINHAMENTO e o REDIMENSIONAMENTO de gestos são os meios pelos quais as regularidades fônicas adentram níveis mais profundos da gramática. Tratemos, rapidamente, de cada uma delas.

Albano (2001) explana que existe forte relação entre os processos fônicos que se apresentam na fala e a mudança linguística lexicalizada. Tal relação serve de sustentação ao postulado de que a especificação do gesto compreende um elo auditivo que integra critérios quânticos e critérios de dispersão adaptativa. Em linhas gerais, para a estudiosa, no que concerne à teoria quântica das relações acústico-articulatórias, os inventários fônicos seriam selecionados conforme sua resistência à distorção. Assim, os elementos sonoros teriam uma estrutura acústica estável dentro de certa margem de variabilidade articulatória, ou uma estrutura auditiva estável dentro de uma certa gama de variabilidade acústica. Esses sons, tidos como os mais frequentes nas línguas, comportam as combinações mais salientes de propriedades primárias e secundárias, além de apresentar relações acústico-articulatórias ou acústico-auditivas designadas quânticas. A implicação de tais relações incide no fato de a variação do parâmetro articulatório configurar platôs e saltos no parâmetro acústico. No que se refere à teoria da dispersão adaptativa dos sistemas fônicos, Albano (2001) elenca que a habilidade de produzir e reconhecer gestos tem a ver como uma adaptação de ordem filogenética a restrições no âmbito motor e perceptual. Essas restrições preservam contraste suficiente, segundo as quais os gestos devem exibir distâncias articulatórias minimamente razoáveis entre si para que seus efeitos de distintividade na elocução sejam mantidos. Com base nessas ponderações, a autora traz argumentos que propiciam a inserção de PARÂMETROS ACÚSTICOS ao modelamento articulatório, e isso fica nítido com a noção de regiões acústico-articulatórias do trato vocal, proposta nos trabalhos de Johan Ohala, Kenneth Stevens e Björn Lindblom, da qual Albano (2001) se vale.

Regiões acústico-articulatórias diversas impõem conseqüências acústicas distintas, e essa afirmação nos outorga caracterizar um gesto como uma “manobra motora” capaz de produzir efeitos acústicos que respondem por diferentes regiões do aparato vocal e por diferentes tipos de constrição. Albano (2001, p. 98) ilustra que “o que confere unidade ao gesto de fechamento dental não é o articulador ativo ponta da língua, mas a amplitude baixa ou zero da oclusão acoplada ao caráter

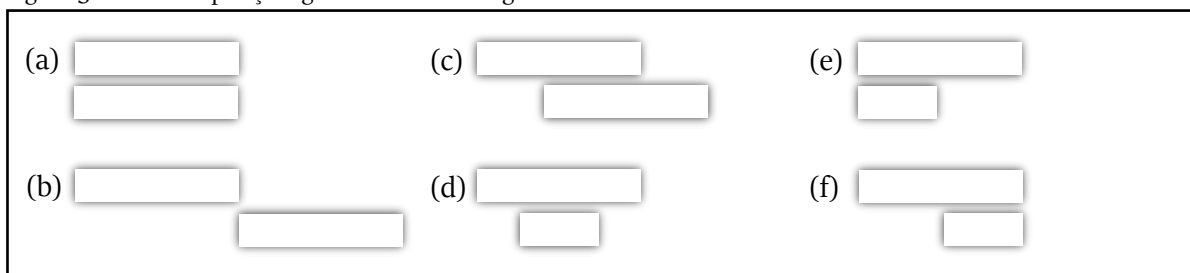
inequivocamente agudo das ressonâncias produzidas na região coronal”. Nesse sentido, Albano (2001) esclarece que, em algum ponto de suas trajetórias, duas variáveis do trato que se deslocam na mesma região do trato produzem um sinal acústico característico de um dado gesto, tornando-o distinto dos demais presentes no mesmo inventário. A propósito disso, a estudiosa aponta que:

Um certo deslizamento dos componentes gestuais é esperado se a definição do gesto se refere a regiões acústico-articulatórias discretas, capazes de garantir efeitos acústicos estáveis e, ao mesmo tempo, acomodar variações estilísticas, muitas delas desencadeadas por fatores prosódicos que antecipam ou adiam as manobras articulatórias que causam tais efeitos” (ALBANO, 2001, p. 101).

Albano (2001) comenta que é durante a infância que as pessoas atingem o ápice de sua capacidade de se deixar levar por atividades sensório-motoras repetitivas, as quais se instauram na fala e nos demais âmbitos da motricidade. Para a autora, a produção dos gestos está sujeita a diferenças individuais, mas seu reconhecimento está sujeito à maturação. Nessa via de mão dupla, faz-se oportuna a observação de que a especificação do gesto reside na identificação das REGIÕES NO TRATO, e não na identificação de articuladores e trajetórias articulatórias, o que faz com que essas regiões sejam determinadas em conformidade com a dinâmica do trato vocal, interpretado como um TUBO ACÚSTICO. Advém dessas ponderações o tratamento tanto acústico quanto articatório que Albano (2001) dispensa à unidade gestual, o gesto acústico-articulatório.

Albano (2001) defende que as BORDAS – início e fim, portanto – dos gestos são os lugares de sua projeção simbólica. Evidentemente, para que se faça referência simbólica à natureza dinâmica da unidade gestual, é necessário considerar um conjunto indeterminado de pontos entre seu início e seu fim (um intervalo). Desse modo, conforme explica a pesquisadora, os gestos podem restar completamente sobrepostos, ou seja, ter ambas as bordas alinhadas paralelamente, ou podem restar parcialmente sobrepostos, isto é, um pode iniciar e/ou terminar durante o curso do outro, em que há o alinhamento de uma das bordas a um intervalo. Outra possibilidade, ainda, configura a situação em que os gestos estejam justapostos (um inicia depois do término de outro), alinhando uma borda final a uma borda inicial. Esses casos são representados na Figura 3 a seguir:

Figura 3 – Possíveis posições gestuais em um conglomerado



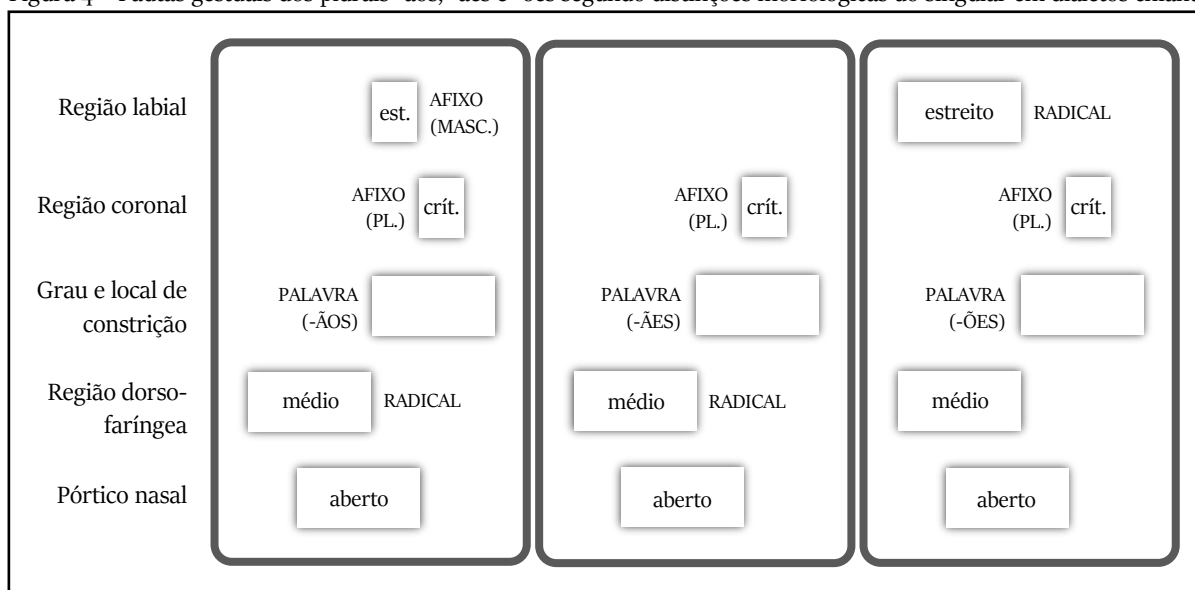
Fonte: adaptada de Albano (2001)

A partir do diagrama expresso na Figura 3, Albano (2001) salienta que as relações que realmente exigem um modelamento dinâmico são (c), (d), (e) e (f): aquelas que apresentam bordas alinhadas a intervalos. Por outro lado, (a) e (b) cabem adequadamente no tratamento realizado pelos modelos fonológicos tradicionais, já que esses operam com noções binárias no que tange ao tratamento teórico que oferecem (no caso, total sobreposição e total separação, em que se resumem fatos tidos como categóricos). Como exemplo das relações que impõem desafios ao paradigma gerativo e que se alinham a uma descrição gestual, a estudiosa elenca os casos referentes à africacão de [t] e [d], à iotização de vogal antes de [s] e à inserção de [i] em encontros consonantais, explorados em sua obra.

No âmbito do REALINHAMENTO e REDIMENSIONAMENTO dos gestos fônicos, Albano (2001) se fundamenta em fenômenos de SANDHI em português para estabelecer relações que incidem, por exemplo, no mérito da morfologia em português¹⁰⁰. Albano (2001) comenta que a concepção estática e sequencial das unidades fônicas geralmente se conecta a uma concepção de igual ordem no que se refere a unidades morfológicas. Nas próprias palavras da autora, a “não linearidade radical das relações entre os componentes das pautas gestuais permite uma maior flexibilidade no tratamento das fronteiras morfológicas” (ALBANO, 2001, p. 121). Nesse sentido, a apreciação realizada por Albano (2001) tem a ver com o plural de palavras terminadas por -ão em dialetos chiantes, as quais podem receber três análises gramaticais distintas para cada um dos casos de tal flexão, a saber, -ãos, -ães e -ões. Para a pesquisadora, os fatos diacrônicos referentes ao papel das semivogais podem ser facilmente expressos com base em pautas gestuais sinalizadas com as devidas categorias morfológicas, conforme podemos observar na Figura 4 a seguir:

¹⁰⁰ Trazemos, como ilustração, um exemplo de SANDHI interno, mas a autora também se reporta a processos de SANDHI externo.

Figura 4 – Pautas gestuais dos plurais -ãos, -ães e -ões segundo distinções morfológicas do singular em dialetos chiantes



Fonte: adaptada de Albano (2001).

Com base na Figura 4, Albano (2001) exemplifica suas análises morfológicas à luz da configuração gestual alimentada pelas variáveis em jogo. No primeiro caso, o do plural terminado em -ãos, temos que a semivogal corresponde a um afixo (marca de masculino) e, ao receber um sufixo fricativo, sobrepõe o gesto labial ao seu componente “iotizante”. No segundo caso, o que apresenta o plural terminado em -ães, a semivogal corresponde a um elemento paragógico e o radical se encerra no gesto dorso-faríngeo. Ao receber um sufixo fricativo, a vogal paragógica se perde, fazendo com que a transição “iotizante” do sufixo soe como uma semivogal palatal. No terceiro caso, o do plural terminado em -ões, a semivogal corresponde ao resultado de um desdobramento da vogal do radical, sendo que este termina no gesto labial. Há, então, a expansão do gesto labial, evitando o mascaramento da fronteira do radical pelo componente “iotizante” do sufixo, ao tornar paralelos os gestos labial e dorso-faríngeo com o deslizamento do gesto labial para a esquerda.

De acordo com Albano (2001), os exemplos morfológicos apresentados são bastante arrojados em sinalizar as vantagens de haver unidades fonológicas que sejam comensuráveis com unidades fonéticas, já que tais instâncias refletem o fato de a língua poder se utilizar de uma “elasticidade esquemática” no estabelecimento do léxico, apontando contrastes entre classes gramaticais e/ou lexicais. No que tange à relação entre léxico e gramática, a autora elenca que a “gramática fônica, isto é, o conjunto de restrições de boa formação que licencia as entradas lexicais” (p. 140), não se reserva a um domínio demarcado no léxico (pois não é um módulo no sentido computacional), mas se

encontra distribuída ao longo dele através de regularidades que têm altas, mas não unitárias probabilidades.

Endereçados os pontos mais essenciais da proposta de Albano (2001), avaliamos que o tratamento que a estudiosa presta à unidade gestual é mais apropriado e coerente do que o tratamento que Browman e Goldstein (1989, 1992, 1995) prestam a tal unidade. É bem verdade que a adoção a um primitivo gestual em um modelo de percepção fônica não é novidade (já que o *PAM* e o *PAM-L2* se utilizam do gesto articulatório como unidade da percepção), mas incorporar a unidade gestual de natureza ACÚSTICO-ARTICULATÓRIA a um modelo de percepção se mostra, de fato, como uma característica inovadora no quesito da explicação do evento perceptual.

Ressalvamos que, se não forem delineadas vantagens, a inovação, por si só, não se faz suficiente para uma nova concepção de tratamento teórico à unidade fônica da percepção. Assim, para que possamos captar o que está em jogo na relação entre os elementos fônicos das línguas na perspectiva perceptual, buscamos amparo nos preceitos de Albano (2001). Isso ocorre na medida em que seu trabalho congrega três fatores imprescindíveis a qualquer unidade que se encontre alocada no plano dos gestos fônicos e que pretenda responder pelo evento da percepção, relacionando a informação advinda do evento perceptual com a informação que reside na representação. O primeiro fator diz respeito à PROPORCIONALIDADE com que Albano (2001) relaciona a esfera abstrata, portanto, mental, do gesto acústico-articulatório com a sua esfera motora¹⁰¹, portanto, física¹⁰². O segundo fator reside na APROXIMAÇÃO do gesto acústico-articulatório a outras unidades que operam na gramática (conhecimento linguístico), como, por exemplo, o morfema. Por fim, o terceiro fator vincula-se à RELEVÂNCIA da informação acústica imbricada no primitivo gestual, a qual se estabelece em termos de ação (resultado da articulação) e de representação (parte da abstração).

Albano (2001) relaciona de maneira PROPORCIONAL a esfera abstrata (mental) do gesto acústico-articulatório com sua esfera física (motora), e apontamos essa simetria como sendo bastante

¹⁰¹ Utilizamos o termo “MOTOR” como sinônimo de MECÂNICO, em que configurações articulatórias de produção, assim como seu resultado acústico, operam em conjunto. A partir do uso de tal termo, não temos a intenção, portanto, de vinculá-lo ao arcabouço teórico da Teoria Motora da Percepção da Fala (LIBERMAN; MATTINGLY, 1985).

¹⁰² À primeira vista, pode parecer que esse argumento seja igualmente aplicável ao tratamento articulatório de Browman e Goldstein (1989, 1992) à unidade gestual, já que existe uma relação entre abstração e concretude tanto no plano articulatório como no plano acústico-articulatório – o que acabaria por neutralizar nosso argumento. Todavia, há que se atentar para uma possível diferença quanto à proporção em que abstração e concretude se instauram em ambos os tratamentos mencionados: Albano (2001) se mostra muito mais preocupada e atenta a fenômenos gramaticais em um nível simbólico do que Browman e Goldstein (1989, 1992).

congruente e salutar, já que a flexibilidade fonética deve ser comensurável com a organização fonológica das unidades fônicas. É bem verdade que Browman e Goldstein (1989, 1992) fazem referência às faces concreta e abstrata do gesto articulatório, alusão de que Albano (2001) também compartilha. Na mesma linha, Fowler (2003) adverte que se faz primordial que uma teoria fonológica leve em conta aspectos mais concretos da cadeia sonora, como por exemplo, a gradiência registrada nos sons linguísticos. Novamente, Albano (2001) alinha-se a esse posicionamento, bem como o defende. No entanto, cabe a ressalva de que, embora Goldstein e Fowler (2003) admitam que haja abstração no que se refere ao gesto articulatório, sua argumentação é conduzida, em boa parte, a salientar com robustez o caráter físico de tal unidade. É nesse mérito que julgamos mais conveniente a visão de Albano (2001), em que as faces abstrata e física do gesto articulatório parecem estar mais equilibradas, ou seja, consideramos que parece haver maior SIMBOLISMO na perspectiva de Albano (2001) do que na de Browman e Goldstein (1989, 1992). Ademais, por empreender o construto de uma GRAMÁTICA FÔNICA, resta evidente, por parte de Albano (2001), a preocupação, no nível simbólico, em estabelecer vínculos estreitos com um meio informacional abstrato da linguagem.

O segundo fator que reportamos ser apropriado no aporte gestual de Albano (2001) reflete a APROXIMAÇÃO entre o gesto acústico-articulatório e o morfema. Essa, definitivamente, não é uma tarefa trivial, e representa um avanço teórico que segue em duas direções. A primeira é tornar evidente que o gesto acústico-articulatório faz parte de um conhecimento abstrato da linguagem em função de ele “andar ao lado” de uma outra unidade indiscutivelmente gramatical e abstrata – o morfema. A segunda é explicitar em que instâncias ocorrem as relações entre gestos e demais entes gramaticais, resultando na preocupação analítica da autora em expressar as relações UNIDADE FÔNICA-UNIDADE MORFOLÓGICA, referentes ao realinhamento e ao redimensionamento dos gestos fônicos, conforme expusemos previamente.

Está aí outra diferença entre o tratamento teórico do gesto empreendido por Albano (2001) e aquele empreendido por Browman e Goldstein (1989, 1992) e Goldstein e Fowler (2003), distinção essa que confere maior poder explanatório à elucubração de Albano (2001). Em nossa visão, os estudiosos norte-americanos não se preocupam em relacionar o gesto a outras unidades gramaticais, o que pode servir de evidência para questionarmos o nível de abstração que o gesto assume para tais pesquisadores. No limite, questionamos se a falta de conexão entre gestos e outros entes da gramática designa um compromisso teórico apenas com a habilidade articulatória e não com as relações entre unidades gramaticais envolvendo aspectos fônicos, morfológicos, sintáticos, semânticos e

pragmáticos. Esse questionamento parece ser genuíno, uma vez que, conforme relatado no segundo capítulo desta tese, Fowler (1986) esclarece que teria “pouco a oferecer” teoricamente em termos de eventos linguísticos, ao contrário de eventos da fala, seu notável foco de investigação. Ademais, essa constatação reforça a nossa suposição de que a contraparte abstrata do gesto segundo Browman e Goldstein (1989, 1992), Fowler (1986) e Goldstein e Fowler (2003) seja bem menos simbólica, limitando tal unidade quando vislumbramos toda a informação que a língua carrega.

O terceiro fator que apontamos ser profícuo na abordagem de Albano (2001) e que se faz importante no campo da percepção fônica é a RELEVÂNCIA dos aspectos acústicos do gesto fônico. Para a autora, a incorporação de parâmetros acústicos na unidade gestual dá suporte à concepção de uma fonologia de base acústico-articulatória. Para nosso propósito nesta tese, o de verificar o processo de percepção das unidades fônicas, essa incorporação responde por duas considerações teóricas de extrema pertinência, além de servir como estratégia metodológica para abordar tal unidade.

A primeira consideração evidencia as faces ARTICULATÓRIA e ACÚSTICA dos fatos fônicos. Por mais que sejam atreladas a uma relação causal em produções naturais (e não sintéticas) – já que o sinal acústico da fala advém de constrições envolvendo as variáveis do trato vocal –, podem resultar em consequências distintas para interpretação auditiva da organização gestual das línguas. Nesse sentido, Albano (2001) relata como exemplo o caso de um falante de português brasileiro, pertencente a uma variedade não chiante, ao produzir as palavras “vítima” (proparoxítone) e “vitima” (paroxítone) na mesma frase veículo. Em sua análise, a autora afirma que, articulatoriamente, a configuração gestual das palavras é muito semelhante (excetuando-se as dimensões dos blocos que correspondem às vogais, responsáveis pela distinção acentual), mas, auditivamente, pode ser interpretada ou como uma africação ou como uma simples palatalização. Essa discrepância interpretativa/perceptual seria derivada de diferenças de amplitude existentes nas relações entre as consoantes e as vogais em questão, sendo que, na proparoxítone, há maior amplitude no ruído e, na paroxítone, há maior amplitude na vogal.

Uma ressalva substancial que diz respeito às propriedades acústicas da organização gestual no evento perceptual é a de que elas estão tanto para o lado motor, residindo nos parâmetros que derivam das manobras articulatórias, quanto para o lado abstrato, dimensionadas a partir de manifestações gradientes expressas na gramática fônica. Defendemos, assim, uma possível retroalimentação entre os lados físico e representacional da informação acústica contida no gesto, a qual encontra sustentação em termos de primitivo de análise e de orientação cognitiva, em que

representações novas (informação acústica advinda da experiência, empírica) são instanciadas com base em representações já consolidadas (informação acústica armazenada, dedutiva) – perspectiva a favor da qual nos posicionamos, como deve ter ficado claro a partir da primeira seção deste capítulo.

A segunda consideração, que, de certa maneira, pode ser deduzida da primeira, reside no fomento da característica MULTIMODAL da informação fônica, cuja explicitação não se faz redundante. A visualização da articulação de determinado gesto aliada à audição de suas propriedades acústicas, simultaneamente, permite que o percebedor se aproprie de modo multimodal do evento comunicativo (obviamente, em conjunto com expressões faciais, postura, movimento dos olhos, etc.). De acordo com Mariani (2009), diversos estudos em inteligibilidade relatam haver apreensão mais efetiva da informação no evento comunicativo quando pistas visuais e acústicas se fazem presentes, especialmente, em ambientes com ruído ou quando a mensagem a ser veiculada é complexa. Segundo o estudioso, essa constatação levou a pesquisas no campo da síntese e do reconhecimento da fala. Nesse quesito, destacamos os trabalhos de Le Goff e Benoit (1996) e Benoit, Massaro e Cohen (1997). Complementarmente, estudos que verificaram o uso das informações faciais no reconhecimento da fala [PETAJAN *ET AL.* (1988); GOLDSHEN (1997)] demonstraram que os participantes que tiveram acesso tanto à informação visual quanto à informação acústica obtiveram maiores índices de acuidade na tarefa se comparados àqueles que tiveram acesso somente à informação visual ou acústica¹⁰³.

No campo metodológico, especificamente, podemos nos utilizar das informações advindas do plano acústico, mensuráveis em sua natureza, para tentar estabelecer relações gestuais, isto é, para tentar chegar, de alguma maneira, à unidade gestual, já que, em nosso país, o acesso ao manuseio de dados articulatórios (seja para fins de coleta, seja para fins de análise) se faz bastante restrito. A via acústica, seria, portanto, um mecanismo para compreender padrões expressos gestualmente, disponível em larga escala aos pesquisadores brasileiros.

Dado o conjunto de ponderações que elencamos nesta seção, acreditamos ter deixado claro em que medida o gesto acústico-articulatório se mostra vantajoso para que possamos empregá-lo como unidade fônica da percepção. Assim, parece-nos oportuno este momento para afirmarmos que, ao referenciar o processo perceptual relativo à unidade gestual das línguas não nativas, iremos fazê-lo sob o rótulo de PERCEPÇÃO FÔNICA, e não, simplesmente, PERCEPÇÃO DA FALA. A escolha por um termo em detrimento do outro consiste no fato de que, ao perceber os elementos fônicos das línguas, estamos

¹⁰³ Um fenômeno interessante no âmbito da colisão entre informações visuais e acústicas pode ser observado a partir do efeito McGurk [MCGURK; MACDONALD (1976)].

entrando em contato com informações que não se vinculam ao plano da fala enquanto articulação, desprovida de significado, mas que estão intimamente ligadas a um sistema de símbolos, dotado de um léxico e, assim, permeado por relações fônicas, morfológicas, sintáticas, semânticas e pragmáticas.

Reiteramos, portanto, que o gesto acústico-articulatório deva ser a unidade que permeia a percepção de novos elementos fônicos. Além disso, com base em Albano (2001), postulamos que, de fato, existe uma gramática fônica, abstrata e representacional, a qual opera com o gesto acústico-articulatório. Esse, 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 não nativo, em nossa visão, se configura, pelo menos no que tange à percepção, como um espaço de constantes relações entre unidades sonoras prévias e novas, as quais podem ser mais ou menos estáveis a depender do uso do idioma-alvo.

Há, ainda, um ponto relevante de ser abordado no que se refere ao *PAM-L2*, em que não poderíamos deixar de tocar, especialmente agora que cumprimos a tarefa de apresentar e discutir a unidade gestual sob os prismas articulatório (BROWMAN; GOLSTEIN, 1989, 1992) e acústico-articulatório (ALBANO, 2001). Esse ponto se refere ao fato de haver possíveis contradições na esfera fônica do *PAM-L2*, conferindo algumas incongruências ao argumento de base dos autores, visto que assumem os princípios de uma fonologia de cunho gestual.

São recorrentes no texto de 2007 compartimentações entre FONÉTICA e FONOLOGIA e, correspondentemente, entre elementos de BAIXO NÍVEL e ALTO NÍVEL. Em uma perspectiva gestual, fatos fonéticos e fonológicos fazem parte de um único domínio, sendo indissociáveis. A contraposição entre elementos de baixo nível e de alto nível também, eventualmente, parece retomar essa mesma antinomia entre o que é fonético e o que é fonológico. Ora aspectos de alto nível são designados como sendo morfológicos e sintáticos, ora são considerados fonológicos. Um excerto em particular que requer atenção é o seguinte:

Sugerimos que os eventos/objetos perceptuais que são relevantes para a aprendizagem da fala em *L2* não são MERAMENTE FONÉTICOS. Propriedades da fala que são relevantes no idioma são diferenciadas não somente no NÍVEL FONÉTICO, mas também no NÍVEL FONOLÓGICO DE ALTA ORDEM, bem como no NÍVEL GESTUAL DE BAIXA ORDEM. Diferenças entre a *L1* e a *L2* nos níveis gestual, fonético ou fonológico podem influenciar, em conjunto OU SEPARADAMENTE, as

habilidades discriminatórias do aprendiz de L₂, dependendo do contexto ou dos objetivos do percebedor¹⁰⁴ (BEST; TYLER, 2007, p. 25, grifo nosso).

Com base nesse trecho, encontramos algumas incongruências teóricas que precisam ser resolvidas. Primeiramente, a expressão “meramente fonéticos” pressupõe que os fatos fonéticos sejam mais simples e mais básicos do que outros fatos que ocorrem na língua. Ainda que os autores possam vir a não acreditar que os objetos ou eventos perceptuais sejam meramente fonéticos, o fato de se referirem ao termo em isolado já evidencia que há, por exemplo, uma outra instância do conhecimento linguístico que poderia ser somente fonológica.

Similar acepção é aquela que podemos inferir a partir da segunda expressão destacada, “nível fonético”. Ora, resta praticamente indubitável que os autores assumem um nível fonético separado de um nível fonológico – não fosse o referencial teórico de que comungam, isso não seria surpreendente; entretanto, sob uma perspectiva gestual, essa postura não é adequada. Best e Tyler (2007), na sequência, fazem menção não somente a um nível fonológico, mas a um nível fonológico “de alta ordem”. Se, em uma aproximação com a terminologia tradicional, a antinomia fonética vs. fonologia está para um fato de baixa ordem vs. um fato de alta ordem, intriga-nos o que poderia ser um fato fonológico em uma posição ainda mais alta. Há, no entanto, que se considerar que, para os autores, a utilização desses dois termos possa prever uma relação de sinonímia.

Os autores também estipulam a existência de um “nível gestual de baixa ordem”. Novamente, em uma concepção gestual, essa afirmação soa contrassensual, pois não estão previstas baixa ordem e alta ordem. Os elementos fônicos não só são tratados em um âmbito em que fonética e fonologia (em termos tradicionais) são inseparáveis, mas percorrem o léxico como um todo, conforme postula Albano (2001). Por último, Best e Tyler (2007) apontam que, separadamente ou em conjunto, os níveis gestual, fonético ou fonológico podem influenciar as habilidades discriminatórias do aprendiz de L₂. Além de se remeterem a níveis cuja divisão é teoricamente equivocada [tanto para Browman e Goldstein (1989, 1992) como para Albano (2001)], os autores sugerem que qualquer um dos níveis pode influenciar a discriminação, mas, segundo o que defendemos, tanto a assimilação como a

¹⁰⁴ No original: “We suggest that perceptual objects/events that are relevant to L₂ speech learning are not merely phonetic. Language-relevant speech properties are differentiated not only at the phonetic level but also at the higher-order phonological level, as well as at the lower-order gestural level. L₁-L₂ differences at a gestural, phonetic, or phonological level may each influence the L₂ learner’s discrimination abilities, separately or together, depending on the context or the perceiver’s goals” (BEST; TYLER, 2007, p. 25).

identificação também podem ser influenciadas por esses mesmos níveis (ainda que não sigamos os autores na estipulação de um nível fonético e um nível fonológico).

Tais aspectos, que consideramos ser deslizos, levam-nos a questionar se, de fato, Best e Tyler (2007) se apropriam do conceito de uma unidade gestual. Uma possibilidade é que, em seu entendimento, o gesto possa figurar somente enquanto mecanismo de articulação, sem compromisso com outras abstrações que a língua demanda. De qualquer modo, esclarecimentos são fundamentais para que se possa interpretar corretamente os pressupostos do modelo em seu expediente fônico.

Para encerrar esta seção, trazemos à baila um oportuno questionamento feito por Alves e Silva (2016) no que compete à percepção e ao armazenamento da informação fônica em termos do *PAM-L2*, o qual transcrevemos a seguir:

[...] considerando-se as possibilidades de assimilação previstas no *PAM-L2*, poderíamos pensar se estariam os indivíduos armazenando unidades gestuais abstratas na sua mente; se assim fosse, os casos de assimilação da língua-alvo a categorias da L_1 não seriam advindos de comparações entre o estímulo recém ouvido e o previamente armazenado, de modo que a representação abstrata prévia estaria exercendo efeito sobre os tipos de categorizações a serem realizadas como o novo estímulo de L_2 ? Sendo afirmativa a resposta para esse questionamento, de que forma tal perspectiva de aquisição seria distinta de outra de caráter mais indireto? (ALVES; SILVA, 2016, p. 114)

O referido questionamento demanda uma resposta dupla, uma em direção ao armazenamento das representações gestuais e sua comparação com o estímulo novo e outra em direção ao acesso às informações fônicas, tomando como base os efeitos da L_1 sobre a aquisição da língua-alvo. Antes de nos posicionarmos, entretanto, faremos alusão à resposta que Alves e Silva (2016) oferecem.

De maneira precatada, os autores exercitam possíveis justificativas ao posicionamento de Best e Tyler (2007) ao trazer considerações de uma perspectiva psicológica que esteja em consonância com a teoria de base filosófica que sustenta o *PAM-L2*. Alves e Silva (2016) remontam aos estudos de Eleanor Gibson acerca da aprendizagem da percepção, ou seja, como aprendemos a perceber. Segundo os autores, em sua tentativa de prover respostas à luz dos pressupostos do próprio *PAM-L2*, esse processo seria compreendido como um procedimento de diferenciação, não de criação. A palavra-chave aqui é a experiência com o ambiente, a partir da qual conseguiríamos diferenciar propriedades do nosso entorno que anteriormente seriam identificadas como um único elemento. Assim, a ideia de aprender a perceber resulta em criar novas categorias. O acúmulo de pautas gestuais, por exemplo, seria uma visão inapropriada da aprendizagem perceptual, que, de acordo

com os postulados de Eleanor Gibson, se instaura através da ação e não de uma representação mental.

Nossa resposta a tal questionamento, advinda do referencial teórico das neurociências e da psicologia cognitiva (e que, portanto, se coloca para além dos preceitos que regem o *PAM-L2*), segue a premissa de que as unidades gestuais seriam, sim, representadas (ou armazenadas) em nosso aparato cognitivo (ou mente). Elas se instanciam em nossa cognição através de nossa capacidade perceptual, que se ocupa de criar um símbolo gestual interno e mental (abstrato) para uma contraparte externa e concreta (física). Qualquer estímulo novo que se apresente será necessariamente instanciado cognitivamente a partir das informações que já ocupam nosso sistema representacional. Essa perspectiva, em primeira análise, não difere de uma de caráter indireto, em que os dados a que temos acesso são vislumbrados através de um ente intermediário (que se aloca entre o percebedor e o objeto a ser percebido).

O que deve ficar claro no que tange à 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 abordagem do casal Gibson e de Best e Tyler (2007). Nesse aspecto, 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.

Convenientemente, a próxima seção irá se ocupar do cenário filosófico que envolve o fenômeno perceptual das unidades fônicas das línguas não nativas e apresentará uma proposta alternativa aos pressupostos filosóficos (com implicações psicológicas) de James Gibson e, por conseguinte, de Carol Fowler, Catherine Best e Michael Tyler.

4.3 Realismo perceptual e seu alcance representacional

A presente seção discorre sobre o terreno filosófico em que se sustenta o evento perceptual. A teoria filosófica de percepção a que recorreremos deverá ser suficientemente potente para embasar o que concebemos como sendo o fenômeno de perceber os objetos do mundo e se encontra

perfeitamente alinhada com a nossa visão de cognição e de representação do conhecimento linguístico.

Iniciemos com uma proposição. Um sorriso, em primeira análise, parece ser veiculador do sentimento de felicidade ou de que algo se apresenta de maneira positiva. Arriscamos afirmar que um sorriso pode “falar” por si mesmo, aqui ou no Japão. Ao ver uma imagem de uma pessoa sorrindo, são grandes as chances de dizermos que ela está feliz ou que algo positivo aconteceu a ela, capaz de gerar um sorriso em sua face. Todavia, se tivermos nascido no oriente, é muito provável que não venhamos a associar o sorriso estampado na face de alguém com seu estado de positividade ou felicidade. O sorriso é objetivamente observável, mas o sentimento positivo ou de felicidade não necessariamente o acompanha.

Aprofundemos um pouco mais a questão. Masuda *ET AL.* (2008) pediram que universitários americanos e japoneses avaliassem duas ilustrações, sendo que em ambas havia cinco pessoas: uma no centro, em posição de destaque, e quatro em seu entorno, em posição mais periférica. Em uma das ilustrações, tanto o personagem central como os personagens adjacentes apareciam com um sorriso no rosto. Em outra, o personagem central aparecia com um sorriso no rosto, mas os personagens adjacentes aparentavam estar zangados ou tristes, sem qualquer expressão facial que indicasse felicidade ou positividade. Quando questionados se, em ambas as ilustrações, o personagem central estava feliz, universitários americanos responderam de maneira favorável, alegando que, sim, o personagem central aparentava estar feliz. Por outro lado, quando universitários japoneses responderam à mesma pergunta, eles somente afirmaram que o personagem central estaria feliz quando os personagens periféricos também estivessem felizes, ou seja, quando também apresentassem um sorriso no rosto¹⁰⁵.

Em grande parte da cultura oriental, aparentar felicidade ou tristeza não depende exclusivamente de um indivíduo, mas do comportamento daqueles que o cercam. Evidenciamos que, nesse caso, a ação de sorrir, ou seja, o dado disponível no mundo referente a um sorriso, não é percebida de forma transparente, não se encontra livre de interpretações ou igual para todos os percebedores. Em outras palavras, a detecção dessa informação é completamente distinta para ambos os grupos de universitários. O dado em jogo, aqui, é altamente dependente de um conhecimento convencionalizado socialmente, o qual incide na tomada de decisão sobre determinado

¹⁰⁵ Estudos de ordem similar são os de Imai, Kanero e Masuda (2016); Ito, Masuda e Hioki (2012); Masuda, Gonzalez, Kwan e Nisbett (2008).

estado ou situação. Dito de outra maneira, o dado é dependente do juízo de valor do observador, que, nesse caso, poderia escolher uma de duas opções – responder se o personagem estava feliz ou triste. Tal julgamento existe e é invocado no momento em que “olhamos” para os objetos do mundo. Assim, o contexto que nos circunda só fará sentido e será informativo acerca de um objeto ou estado se for vislumbrado por uma referência prévia, acordada em um meio social e dotada de significação.

Em que medida a situação aqui reportada se relaciona com a percepção das unidades fônicas das línguas não nativas deve ser uma das primeiras indagações a serem feitas para os propósitos desta tese. Vimos em seções anteriores que o *PAM-L2* é um modelo que pretende se sustentar fonicamente a partir do gesto articulatorio, uma unidade proveniente de um modelo fonológico, e se acomodar filosoficamente através de uma teoria perceptual de cunho realista direto. Assim como propusemos, na seção anterior, um tratamento alternativo à unidade fônica do modelo, oferecemos, nesta seção, uma outra visão quanto ao posicionamento filosófico que subjaz ao *PAM-L2*, e é nesse quesito que respondemos ao questionamento incitado no início deste parágrafo: o paralelo existente entre a situação-exemplo relatada aqui e a percepção fônica de línguas não nativas circunscreve-se na constatação de que o evento perceptual ocorre por vias *INDIRETAS*.

Tanto a postura de James Gibson quanto a nossa compartilham de uma característica: ambas são *REALISTAS*. Mas o que significa, afinal, ser realista? 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, propriedades estéticas e morais e outras mentes. A ideia central do realismo é a de que os objetos existem no universo, independentemente de o percebido 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 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, sob o plano 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 J. 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 provenientes dos sentidos.

De modo bastante interessante, Dancy (1985) comenta que a disputa entre o realista direto e o realista indireto se instaura a partir da possibilidade de acessarmos direta ou indiretamente os objetos a serem percebidos. Acerca de tal embate, 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 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).

O realismo indireto prevê que o mundo externo realmente existe, mas nossa percepção em relação a ele é mediada pela percepção de objetos mentais intermediários e subjetivos, como, por exemplo, sensações (BROWN, 2009). Ao que explana Dancy (1985), a vertente indireta do realismo presume que nunca acessamos diretamente os objetos físicos, em função de acessarmos diretamente um objeto intermediário, que pode dizer respeito a ideias, imagens, impressões, sensações ou dados dos sentidos, conforme, também, aponta Mound (2003). No realismo indireto, percebemos diretamente um item subjetivo e percebemos indiretamente um item objetivo; em uma relação causal, o OBJETO DISTAL (objetivo, público) é percebido INDIRETAMENTE em função da percepção DIRETA de um OBJETO PROXIMAL (subjetivo, privado). Somente a título de ilustração, podemos ressaltar que a possibilidade de haver algo entre o percebedor e o objeto a ser percebido vai completamente ao encontro da hipótese de que a percepção de elementos sonoros de uma LNN é filtrada pelo conhecimento fônico da L1 (o qual configura um dos diversos objetos intermediários possíveis), conjectura em favor da qual nos posicionamos.

John Yolton, filósofo americano, sinaliza, em um artigo datado de 1968, sua discordância em relação aos preceitos de James Gibson. Yolton (1968) afirma que J. Gibson (1966) propõe retirar o elemento sensorial da percepção¹⁰⁶, o qual se refere à informação do estímulo acerca de cores, tamanhos, formas, etc. No entanto, para Yolton (1968), a percepção pressupõe, inequivocamente, as

¹⁰⁶ Em função do fato de James Gibson tratar os próprios sentidos como sistemas perceptuais, conforme apresentamos no segundo capítulo desta tese.

sensações, e a informação veiculada das sensações para as percepções só se torna significativa para o ser humano se traduzida em termos cognitivos. Nessa direção, Yolton (1968) menciona que:

Eu não vejo como alguém poderia deduzir ou inferir que um objeto corresponde a um amigo ou a um inimigo, masculino ou feminino, etc., se tudo de que precisássemos fosse a informação referente ao estado dos nossos receptores. O tipo de informação que captamos diariamente através de nosso aparato sensorial e perceptual é aprendido, é o resultado da experiência prévia e das convenções a que estamos expostos em nossa sociedade e em nossa família. Tal aprendizagem e instrução não é realizada através das propriedades dos órgãos dos sentidos¹⁰⁷ (YOLTON, 1968, p. 405).

No entendimento de Yolton (1968), a concepção realista de James Gibson só faria sentido em um mundo inabitado por caça ou caçador ou por diferentes espécies, pois pressupõe e depende de qualidades sensórias orientadas aos objetos. Acreditamos que a observação de Yolton (1968) seja apropriada, uma vez que a apreensão de intenções, por exemplo, reside em um plano que escapa a propriedades invariantes transmitidas através do fluxo de energia, necessitando de um ferramental mais elaborado e mais rico (possivelmente de ordem indireta) para ser percebido.

Finalmente, o julgamento que Yolton (1968) faz acerca do realismo de James Gibson é o de que se trata de um realismo do senso comum. Se considerarmos que a noção de ambiente ecológico só se sustenta em função do “realismo direto” de J. Gibson (1966, 1986), e tal realismo, aos olhos de um filósofo, não passa de uma banalidade, é possível que a ecologia de James Gibson se encontre teoricamente fragilizada.

James Gibson propõe uma descrição ecológica do ambiente, reconfigurando, por exemplo, o campo da ótica sob o aspecto perceptual, porque essa lhe parece ser a maneira adequada de estabelecer as relações entre o percebedor e seu entorno. Em linhas gerais, como vislumbramos no segundo capítulo deste trabalho, ambiente e percebedor precisam estar em consonância o tempo todo para que o indivíduo possa captar diretamente as informações que emanam dos objetos do mundo, propriedades invariantes que são acessadas de modo direto pelo sujeito via seus sentidos que atuam como sistemas perceptuais. Para James Gibson, ecologia e percepção se fundem, já que as relações entre percebedor e ambiente são intrínsecas; logo, não é preciso haver aparatos intermediários entre

¹⁰⁷ No original: “I see no way in which one could deduce or otherwise infer that an object is friend or enemy, male or female, etc., if all we ever had to go on was information regarding the state of our receptors. The kind of information we daily pick up with our sensory and perceptual apparatus is learned, is a function of previous experience and of the teaching we are all exposed to in our society and family. Such learning and teaching is not done in terms of or through the properties of sense organs” (YOLTON, 1968, p. 405).

eles. Em uma perspectiva realista indireta, torna-se desnecessário discutir se o ambiente que nos circunda é ecológico ou não, porque nunca o acessamos diretamente. Nessa vertente, podemos igualmente utilizar, em sua totalidade, os recursos de que o ambiente dispõe, mas esse processo é dependente de um objeto intermediário que nos faculta avaliar e apreender o que é pertinente em determinada situação. Isso pode servir de argumento para casos extremos em que, mesmo em um ambiente pouco concessivo [com poucas concessões (ou *AFFORDANCES*, na terminologia gibsoniana)], o ser humano consegue sobreviver¹⁰⁸. Ele necessariamente precisa ir além das concessões diretamente apreendidas para que a manutenção de sua existência (ou de sua espécie) esteja, em algum nível, assegurada.

Ao abordar o tema da percepção, Edward Lowe, filósofo britânico cujo trabalho se concentrou, basicamente, em metafísica e filosofia da mente, se mostra favorável ao realismo indireto e, em 1981, reúne em um artigo quatro argumentos empíricos que respaldam o cunho indireto do evento perceptual¹⁰⁹. Tais argumentos correspondem ao do INTERVALO DE TEMPO, ao da TRANSMISSÃO, ao da ILUSÃO e ao da ALUCINAÇÃO. Vejamos, brevemente, cada um deles.

O argumento do INTERVALO DE TEMPO se estrutura, basicamente, a partir dos casos em que a percepção é bimodal, por exemplo, em que os objetos públicos são percebidos visual e auditivamente. Nesse cenário, Lowe (1981) salienta que existe grande discrepância entre as propriedades do objeto quando esse é percebido somente por uma modalidade ou por mais de uma ao mesmo tempo. O exemplo que o filósofo cita diz respeito à situação em que uma pessoa observa, com alguma distância, um homem martelando certo artefato: o observador só ouve as batidas instantes depois de vê-lo martelar, resultando em um intervalo de tempo que responde pela falta de sincronia entre a visão da pancada e a audição que dela decorre. Outro exemplo fornecido pelo teórico é o fato de podermos enxergar estrelas que já deixaram de existir; vemos apenas sua luz, seus raios, que demoram centenas ou milhares de anos para chegar à Terra.

O argumento da TRANSMISSÃO tem a ver com a percepção sendo dependente de um processo de transmissão entre o objeto e o cérebro do percebedor e tem íntima relação com o conceito de codificação sensorial, mencionada na primeira seção desta tese. Assim, no âmbito da visão, quando a energia eletromagnética chega ao olho humano, ocorre uma reação química em que a luz é

¹⁰⁸ Agradecemos à professora Adelaide Silva pela problematização levantada na etapa de qualificação desta tese, pois seus questionamentos se mostraram extremamente relevantes a esse respeito.

¹⁰⁹ Para argumentos adicionais contra o realismo direto, recomendamos o texto de Le Morvan (2004).

transformada em eletricidade e, assim, transmitida ao cérebro através dos nervos, que capta o impulso elétrico e forma as imagens que vemos. Segundo Lowe (1981), o argumento pode ser ilustrado pela observação de um holograma, em que as ondas eletromagnéticas que chegam aos olhos dos percebedores são reconstruídas de modo que correspondam àquelas que seriam observadas originalmente.

O terceiro argumento, o da ILUSÃO, corresponde aos casos em que uma miragem é supostamente vista no deserto, ou quando um bastão fino é imerso até a metade em um recipiente com água, resultando em uma imagem que interpretamos ser de um bastão parcialmente torto. Além disso, Lowe (1981) faz referência à ilusão de vermos, por exemplo, duas canetas quando, na verdade, seguramos apenas uma próxima aos nossos olhos e focamos em objetos distantes. O argumento da ALUCINAÇÃO é similar ao da ilusão, exceto pelo fato de que não há a percepção de um objeto público e, mesmo assim, a experiência perceptual é tida como verídica (por quem realiza a ação de perceber) e simula um evento em que há, de fato, um objeto público a ser apreendido.

Com base no que expusemos, defendemos que o realismo indireto (JACKSON, 1977, 2010) seja a teoria filosófica que acomoda, de modo mais propício e coesivo, os fatos linguísticos em termos perceptuais, especialmente, no que tange à percepção fônica não nativa. Essa constatação se dá em função de três premissas interconectadas que, em nossa compreensão, incidem no evento perceptual de unidades fônicas não nativas e encontram sustentação em uma abordagem indireta da percepção. Tais premissas são explicitadas a seguir e, então, discutidas.

- a) O inventário fônico da L₁ filtra a percepção das unidades fônicas não nativas.
- b) O evento perceptual é construído e dependente do cérebro.
- c) A linguagem e suas unidades compreendem um código de significação socialmente construído e convencionado.

Desde a obra *PRINCIPES DE PHONOLOGIE*, ensejada por Nicolai Trubetzkoy, temos conhecimento de que a aprendizagem (e, portanto, a percepção fônica) de um sistema linguístico não nativo é balizada pela L₁. Recentemente, esse fato tem sido modelado também por Hancin-Bhatt (1994), Best (1995), Flege (1995), Brown (2000), Best e Tyler (2007) e Strange (2010), os quais assumem que a L₁ “filtra” os elementos fônicos das línguas não nativas e corroboram tal assunção através de pesquisas experimentais. Por conseguinte, a categorização das unidades fônicas é condicionada pelas propriedades do sistema fônico materno. Essa constatação figura como evidência cabal de que o acesso aos gestos fônicos não nativos só pode ocorrer de maneira indireta, resultando na primeira

inadequação do realismo direto a um modelo de percepção fônica. Além disso, ao assumir que categorias fônicas da língua-alvo possam ser assimiladas a categorias da L_1 , Best (1995) e Best e Tyler (2007) estão obrigatoriamente vislumbrando tais associações com base em um sistema que é, por excelência, prévio, tornando sua postura incompatível com um realismo de cunho direto. Não há, portanto, como presumir que o acesso aos objetos distais não nativos seja direto quando se considera a atuação do filtro da L_1 sobre a percepção dos gestos fônicos.

Nesta tese, concebemos que o fenômeno da percepção corresponde a uma tarefa essencialmente cognitiva, conforme explanam Braisby e Gellatly (2005), Miller, Vandome e McBrewster (2009), Quinlan e Dyson (2010), Gazzaniga *ET AL.* (2012), Kandel *ET AL.* (2014), Kemmerer (2014), Eysenck (2015) e Sternberg, R. e Sternberg, K. (2016). Toda a informação externa que chega até nós, percebedores, passa pelo processo de codificação sensorial para que possa, então, ser interpretada pelo encéfalo. Ondas mecânicas e ondulações na pressão do ar são captadas por células do nosso aparato auditivo, mas só se traduzem em sinfonias após terem sido processadas pelo cérebro. O mesmo se aplica às imagens que observamos, conforme já expusemos. No evento perceptual, os estímulos distais são “trabalhados” por células do nosso corpo que são especializadas em converter dados brutos em elementos significativos, os quais são comparados com os demais elementos já representados em nossa cognição e dão base para inferências e testes de hipóteses (GARDNER; JOHNSON, 2014). É nesse mérito que o realismo indireto se coaduna com a postura cognitivista à percepção que adotamos.

Chamamos atenção para o fato de que, no momento em que o percebedor/aprendiz de uma L_{NN} reconhece distinções lexicais veiculadas através da discriminação – atribuindo, assim, significado linguístico à massa sonora a que tem acesso –, ele poderá resgatar o inventário lexical da língua-alvo, cumprindo um exercício cognitivo por excelência. Perceber diferenças sonoras em outro idioma, em nosso entendimento, não consiste apenas em se apropriar de uma tarefa acústico-articulatória, mas compreende a relação cognitiva entre unidades sonoras prévias e novas, cujo resultado pode ser visto em rotinas motoras.

Conforme o próprio J. Gibson (1966, 1986) declara, a apreensão do conhecimento linguístico, veiculado pela fala ou pela escrita, se encontra fora do escopo do realismo direto. Isso ocorre porque a linguagem é um sistema criado pelo homem, regido por convenções, determinado pelo comportamento e pela organização social instaurados histórica e localmente nas comunidades de fala e de prática. A linguagem se sustenta a partir de sua significação, a qual opera em um domínio

indispensavelmente simbólico. Perceber unidades fônicas das línguas não nativas, especialmente na medida em que se avança conjuntamente na apropriação de seu inventário lexical, implica perceber um conjunto de inter-relações gestuais, morfológicas¹¹⁰, sintáticas, semânticas, pragmáticas e discursivas. O realismo indireto, em nosso entendimento, capta de forma eficiente a percepção da linguagem tendo em vista que o conhecimento linguístico vai muito além de apenas movimentos articulatórios no trato vocal. Assim, no momento da percepção, o aglomerado de significação representado pela linguagem é veiculado, portanto, entre o percebedor e o produtor da mensagem.

Embora estejamos nos encaminhando para a conclusão da seção, é importante que façamos quatro ressalvas no que concerne a alguns posicionamentos de Best (1995) e de Best e Tyler (2007). Primeiramente, ao assumir que existe um elemento intermediário entre o percebedor e o objeto do mundo real, não estamos pressupondo, necessariamente, que tal elemento seja inato. Como relatamos, ele pode se referir a sensações ou a dados dos sentidos, por exemplo (DANCY, 1985). O ataque ao termo *INATO* utilizado por Best (1995, p. 174) pode ter sido motivado simplesmente como uma maneira de reagir ao paradigma gerativo, que prevê estruturas inatas para a linguagem. Portanto, tal acusação não se aplica ao realismo indireto. Em segundo lugar, a existência de elementos intermediários não pressupõe que os estímulos forneçam uma fonte pobre de informação sobre o mundo, apenas refreia uma apreensão direta da realidade. Os objetos continuam sendo *RICOS*, *INFORMATIVOS* e de *NATUREZA MULTIMODAL*, porém são acessados indiretamente. Ademais, em terceiro lugar, afirmar que as informações veiculadas pelos objetos são confiáveis é contar só parte da história: o que dizer sobre o erro perceptual? No realismo direto, o erro é tratado de maneira extremamente simplista e sua justificativa é a de que o conhecimento dos nossos estados sensoriais é falível e que, portanto, a consciência direta não deve ser infalível (DANCY, 1985). Conforme explica Dancy (1985), é pouco provável que o realismo direto ofereça uma explicação satisfatória para o erro perceptual sem cair no realismo indireto. O erro perceptual, segundo o aporte realista indireto, é fruto de uma incompatibilidade na relação existente entre o elemento subjetivo, percebido diretamente, e o elemento objetivo, percebido indiretamente. Por último, explorar ativamente o mundo não é uma habilidade proporcionada exclusivamente pela abordagem ecológica. É viável pensarmos que, mesmo em uma perspectiva indireta, somos capazes de explorar o mundo de modo ativo, afinal, isso é uma

¹¹⁰ Observações pertinentes sobre a relação entre realismo indireto e os pronomes demonstrativos são encontradas no texto de Brown (2009).

característica humana; novamente, a questão de o acesso aos objetos ao nosso redor ser indireto não compromete em nenhum aspecto o nosso sistema perceptual.

Estabelecidas as devidas considerações sobre as razões pelas quais julgamos que o realismo indireto deva embasar o fenômeno perceptual em termos das unidades fônicas não nativas, iremos nos direcionar a questões adicionais que devem ser contempladas em um modelamento da percepção de tais unidades, endereçadas na próxima seção.

4.4 Questões pertinentes ao modelamento da percepção fônica de línguas não nativas

Apesar de identificarmos os pontos discutidos nas três seções anteriores como sendo os eixos basilares da tese que propomos, sugerimos que outras questões, advindas das considerações tecidas em Best e Tyler (2007), também sejam pertinentes quando refletimos sobre os mecanismos de um modelamento do evento perceptual relativo a línguas não nativas. Desse modo, nosso intuito nesta seção é o de discutir sobre a suposta dicotomia L_2 vs. L_E , sobre a influência gráfica e sobre a formação de novas categorias no que compete à percepção fônica não nativa. Ressaltamos, no entanto, que não objetivamos oferecer respostas peremptórias a essas questões, mas julgamos indispensável que sejam encaradas como dignas de atenção.

4.4.1 Segunda língua vs língua estrangeira: uma fronteira inexistente

A abordagem ecológica em conjunto com a concepção filosófica do *PAM-L2* faz com que seus proponentes restrinjam conceitualmente os contextos de aplicação do modelo. A noção de ambiente é fundamental para a percepção fônica não nativa, segundo os autores, porque, quanto mais concessões estiverem disponíveis no entorno do percebedor, maiores são as chances de o evento perceptual ocorrer de maneira exitosa. Assim, Best e Tyler (2007) declaram que o ambiente de L_2 , em detrimento do ambiente de L_E , é o mais apropriado para o percebedor. Nesse contexto, em que o idioma-alvo é falado fora da sala de aula, em diversas situações e por diferentes sujeitos, as oportunidades de percepção são mais robustas, pois, idealmente, o percebedor faz uso de uma fonte praticamente inesgotável de insumo.

Segundo ressaltam Alves e Silva (2016), no que tange ao ambiente de L_E , Best e Tyler (2007) são categóricos em afirmar que as concessões são muito reduzidas. Best e Tyler (2007) destacam que,

de maneira geral, o contexto de *LE* oferece um cenário empobrecido de aprendizagem, pois conta com grande influência da *L1* e reduz as situações comunicativas de modo que essas não consigam se estender para além da sala de aula. Ademais, conforme já expresso no segundo capítulo, para Best e Tyler (2007), além de ser fortemente influenciado pelo sotaque de *L1* do professor, tal ambiente conta com o emprego extensivo da instrução formal acerca de aspectos lexicais e gramaticais, impedindo que a aprendizagem ocorra de maneira comunicativa, baseada em situações realísticas que circundam o cotidiano dos aprendizes e que requerem o uso efetivo do idioma-alvo. Nesse quesito, segundo os próprios estudiosos comentam, o contexto de *LE* apresenta um modelo variável ou “incorreto” (BEST; TYLER, 2007, p. 19) dos detalhes fonéticos da língua alvo.

Concebemos que, se a dicotomia *L2* vs. *LE* realmente procedesse, estaríamos seguros de que a devida aplicação do modelo *PAM-L2* corresponderia exclusivamente aos contextos de *L2*, assim como enfatizam Best e Tyler (2007). Todavia, ocorre que as fronteiras entre *L2* e *LE* estão longe de serem bem delimitadas, seja pelas oportunidades extremamente facilitadas de contato virtual com o idioma-alvo, seja por seus domínios de uso, que se estendem para além do que uma simples antinomia pode designar. 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 chance 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, como seriados, vídeos, músicas, *PODCASTS*, programas de rádio, etc.).

Um ambiente dito de *L2*, contudo, pode não favorecer necessariamente a percepção fônica, o que viria a neutralizar as diferenças mais fundamentais entre *L2* e *LE*. Nesse sentido, de nada adianta estar imerso em um contexto de *L2* e não *ATENTAR* para o *INPUT* fônico que se recebe, por exemplo. É bem possível que um indivíduo, tido como aprendiz de uma *LE*, atente muito mais para aspectos finos do *INPUT* fônico estrangeiro do que um aprendiz de *L2*, que pode atentar para aspectos morfológicos ou sintáticos, por exemplo, não se reservando à informação fônica. Uma análise mais cuidadosa acerca da percepção de *L2* vs. *LE* é suficiente para verificar que, no final das contas, tal dicotomia poderia ser válida em uma ínfima gama de situações.

Seguindo essa linha de argumentação, como deve ter ficado claro a partir de nossas considerações nas seções 4.1 e 4.3, o enfoque realista direto do *PAM-L2* não se sustenta nos próprios contextos de aquisição de *L2*, haja vista a noção de representações mentais e do acesso indireto aos

objetos do mundo. A partir do momento em que defendemos que o evento perceptual é uma tarefa cognitiva por excelência, que demanda inferências, comparações, processamento e atenção ao *INPUT*, foco explícito em estruturas linguísticas, ou seja, aspectos que fogem a uma vertente realista direta, as barreiras entre *L2* e *LE* absolutamente se rompem. Se o aprendiz de uma *LE* não atentar ou não receber insumo suficiente, são baixas as possibilidades de que ele adquira o sistema-alvo (ZIMMER; ALVES, 2006; ZIMMER; SILVEIRA; ALVES, 2009). Se o próprio aprendiz em contexto de *L2*, por quaisquer que sejam as razões, não tiver amplas oportunidades ou concessões, o mesmo irá acontecer.

Mais incisivas do que o argumento cognitivo levantado no parágrafo anterior, o qual encontra respaldo nas ponderações estabelecidas previamente, são, justamente, as proposições que o realismo indireto tem a nos oferecer em termos de um ambiente favorável à percepção. Conforme argumentamos na seção 4.3, o fato de o ambiente ser muito ou pouco concessivo, ou seja, prover em larga ou pequena escala as oportunidades necessárias e suficientes para que os sujeitos percebedores se utilizem diretamente de suas propriedades em prol do evento perceptual, é praticamente irrelevante, pois ele não é a fonte majoritária de possibilidades. Isso ocorre porque, de acordo com o realismo indireto, sempre existe um aparato intermediário alocado entre o percebedor e o mundo, que faculta ao percebedor utilizar as “ferramentas” que o meio oferece de acordo com sua necessidade e propósito. Logo, a noção dicotômica entre *L2* (ambiente muito concessivo) e *LE* (ambiente pouco concessivo) desaba em função de a própria noção de “ambiente dotado de oportunidades” perder seu sentido. É por essa razão que propomos, nesta tese, o termo *LÍNGUA NÃO NATIVA (LNN)*¹¹¹, pois abarca os contextos que, tradicionalmente, são denominados de *L2* e *LE*, mas que não fazem sentido sob a nossa ótica realista indireta e nossa concepção de cognição no que tange à interação percebedor-ambiente.

Alves e Silva (2016) fazem uma ponderação substancial no que se refere à dicotomia *L2* vs. *LE* assumida por Best e Tyler (2007). Os linguistas brasileiros sugerem que as considerações feitas pelos proponentes do modelo nesse mérito podem, na verdade, dizer respeito muito mais à noção de aprendizagem perceptual (Chamot *ET AL.*, 1993) do que às previsões de assimilação previstas no construto do *PAM-L2*. Em outras palavras, estariam em jogo questões referentes ao processo e não ao produto das assimilações ou das categorizações que respondem pelo evento perceptual. Assim, os estudiosos mencionam que “ainda que, realmente, a noção de aprendizagem perceptual *TALVEZ* precise

¹¹¹ O termo já havia sido mencionado anteriormente, mas, neste trecho específico do texto, encontra-se a justificativa realista indireta para que, de fato, o termo se instaure.

ser diferenciada entre ‘segunda língua’ e ‘língua estrangeira’, conforme sugerem os autores, julgamos que não há mecanismos formais ou necessidade de previsões diferenciadas para os ambientes de AL₂ ou ALE” (ALVES; SILVA, 2016, p. 122, grifo nosso).

Salientamos, portanto, que, ao desconsiderar os pressupostos do realismo direto, a abordagem ecológica passa a não fazer sentido, pois não milita a favor de um contexto específico mais adequado para a extração de concessões. Casos de sucesso na aquisição de uma LE (que estão longe de ser raros), se vislumbrados em consonância com os preceitos da ecologia de James Gibson, podem sugerir, inclusive, que a noção de ambiente ecológico seja apenas uma idealização, uma vez que, com poucas “ferramentas” (em alusão a concessões), o aprendiz do idioma-alvo pode apresentar pleno desenvolvimento da língua em questão.

4.4.2 Influência do sistema grafêmico na percepção fônica de línguas não nativas

Vestígios da fala são comumente encontrados em trabalhos escolares de alunos em fase de aquisição da escrita da L₁. Em algumas situações, aspectos fônicos permeiam, inclusive, a escrita de adultos que, em princípio, teriam passado por todo o processo de escolarização. Em português, observamos casos que vão desde a inserção de grafemas em palavras, como “muinto” para “muito” (MELLO *ET AL.*, 2014), a substituição de grafemas, como “sinhora” para “senhora” (MIRANDA, 2010), até casos de hipossegmentação, como “dinovo” para “de novo” (CUNHA, 2004) e hipersegmentação, como “em bora” para “embora” (CUNHA; MIRANDA, 2006).

O caminho inverso, foco de nossa discussão nesta seção, também ocorre. Diversos estudos têm apontado existir influência do sistema grafêmico sobre a oralidade, especialmente no que concerne à percepção fônica das línguas, seja na L₁ (ZIEGLER; FERRAND, 1998; PERRE; ZIEGLER, 2007; PEEREMAN; DUFOUR; BURT, 2009; PERRE; BERTRAND; ZIEGLER, 2011; PETROVA; GASKELL; FERRAND, 2011; RASTLE *ET AL.*, 2011), seja na L_{NN} (ZIMMER, 2004; BASSETTI, 2006, 2008; DETEY; NESPOULOUS, 2008; ZIMMER; SILVEIRA; ALVES, 2009; SHOWALTER; HAYES-HARB, 2013). No primeiro caso, o reconhecimento de rimas em palavras do tipo “TOAST/ROAST”, por exemplo, tende a ocorrer com menor tempo de reação do que em palavras do tipo “TOAST/GHOST”, evidenciando diferenças de processamento entre ambas as situações (SEIDENBERG; TANNENHAUS, 1979). No segundo caso, Tuinman e Cutler (2010) reportam que holandeses, ao realizar uma tarefa

de discriminação, mencionaram perceber sons da fala que, na verdade, não foram produzidos de maneira audível (especificamente, o [ɹ] “intrusivo” que ocorre variavelmente em dialetos britânicos).

Em relação ao âmbito experimental, diferentes metodologias são empregadas nos estudos que investigam o papel da grafia sobre a percepção fônica. As tarefas são variadas, compreendendo julgamentos de rimas (ZIEGLER; FERRAND, 1998; BASSETTI, 2006; PETROVA; GASKELL; FERRAND, 2011), relações entre figuras e suas correspondências grafêmicas (SHOWALTER; HAYES-HARB, 2013), nomeação de figuras (RASTLE *ET AL.*, 2011), decisão lexical (ZIEGLER; FERRAND, 1998; PERRE; ZIEGLER, 2007; PETROVA; GASKELL; FERRAND, 2011; RASTLE *ET AL.*, 2011; SHOWALTER; HAYES-HARB, 2013) e soletração (BASSETTI, 2006).

Um aspecto pertinente levantado por Zimmer, Silveira e Alves (2009) concerne às relações mais transparentes ou mais opacas entre som e letra. Os pesquisadores explanam que indivíduos cuja *L1* apresenta relações mais transparentes entre som e letra podem ter mais dificuldade de adquirir o sistema estrangeiro se esse exibir relações mais opacas. Assim, a título de exemplificação, um brasileiro (cuja *L1* tem relação mais transparente entre som e letra) que está aprendendo inglês (cujo inventário tem relação mais opaca entre som e letra) pode se deparar com itens lexicais que o façam confundir a maneira de pronunciá-los ou percebê-los.

Em termos do *PAM* e do *PAM-L2*, não são abordadas questões referentes à influência da escrita sobre a percepção fônica. Nesse sentido, Alves e Silva (2016) ponderam que uma possível justificativa para que os proponentes do modelo ignorem o papel do sistema grafêmico reside em sua consonância com os preceitos gibsonianos, os quais pressupõem que a língua escrita é uma criação humana e, conseqüentemente, escapa aos propósitos de seu realismo direto e sua abordagem ecológica. Por outro lado, ressaltamos que Best e Tyler (2007) não seguiram os pressupostos gibsonianos de que a língua oral tampouco cabe no realismo direto. O contexto ideal para a aplicação do *PAM-L2*, no âmbito da exclusão de uma transferência grafo-fônico-fonológica (ZIMMER; SILVEIRA; ALVES, 2009)¹¹², seria aquele em que os percebedores não conhecessem o código escrito, conforme salientam Alves e

¹¹² Por questões terminológicas e de paralelismo teórico, julgamos pertinente, nesta tese, substituir o termo “transferência grafo-fônico-fonológica” pelo termo “transferência grafo-fônica” apenas. O motivo de tal substituição reside no fato de seguirmos Albano (2001), que prevê a indissociabilidade entre fonética e fonologia, o que nos permite trabalhar com o termo PERCEPÇÃO FÔNICA. Assim, na vertente gestual, a justaposição dos dois últimos elementos do termo empregado por Zimmer, Silveira e Alves (2009) poderia vir a soar, eventualmente, como uma redundância, implementada, pelos autores do termo, para que esse viesse a ser empregado por pesquisadores de diferentes posturas teóricas (tanto aqueles que considerassem o nível FÔNICO quanto aqueles que considerassem um nível FONOLÓGICO, de natureza dicotômica).

Silva (2016). No entanto, como os próprios autores advertem, na prática, essa opção se faz bastante limitada.

Segundo comentam Rastle *ET AL.* (2011), os efeitos da ortografia na percepção da fala refletem a grande interação entre o registro escrito e o registro falado durante o processamento da linguagem, em que associações entre letra e som previamente aprendidas dão origem a um rápido e automático fluxo de ativação em ambas as direções (da letra para o som e do som para a letra) em termos de representação. O argumento dos autores coincide integralmente com o que defendemos nesta tese em termos de posição filosófica e de orientação cognitiva. No caso do realismo indireto, o conhecimento da escrita também pode funcionar (e geralmente funciona) como um elemento intermediário entre o conhecimento fônico internalizado e o sistema fônico não nativo, permeando as diversas possibilidades de assimilação e/ou categorização. No caso da cognição, é nítido que tanto a percepção fônica não nativa quanto o conhecimento fônico a ser formado a partir da experiência com o sistema não nativo se associam e se instanciam com base na gramática fônica materna, revelando a habilidade mental e simbólica de o percebedor lidar com as unidades fônicas a que tem acesso e captando essencialmente a noção de multimodalidade da informação (visual, acústica e gráfica) que chega até ele.

Se a relação ou a influência da grafia na percepção fônica é inegável, é possível prevê-la em um modelo perceptual de línguas não nativas? Somos levados a sugerir que, sim, é possível prevê-la; porém, a limitação parece ser, fundamentalmente, de ordem metodológica. A representação gráfica de uma língua pode ocorrer pelas vias de sistemas logográficos ou alfabéticos, por exemplo, sendo variável. No entanto, convém salientar que, independentemente da natureza da representação gráfica utilizada, todas elas compartilham de uma propriedade tipográfica essencial: um símbolo gráfico que corresponde ou engloba um elemento oral (seja este associado ao plano puramente fônico ou ao plano morfológico).

No momento da experimentação, em vez de se tentar barrar o acesso às formas escritas do objeto de investigação empírica, deveríamos estar subsidiados por um modelo que previsse *AD HOC* ou *POST HOC*, em algum nível, o papel ou a relação do sistema gráfico com as decisões dos participantes no plano fônico. Com vistas a tal propósito, Bassetti (2008) oferece esclarecimentos bastante oportunos. A pesquisadora menciona que, no expediente linguístico, fatores como TIPO DE SISTEMA DE ESCRITA e NÍVEL DE TRANSPARÊNCIA FONOLÓGICA [esse também levantado por Zimmer, Silveira e Alves (2009)] são cruciais para o tratamento da percepção fônica não nativa. No expediente cognitivo, Bassetti

(2008) pontua que fatores intimamente ligados à experiência dos aprendizes com a L_{NN}, COMO NÍVEL DE PROFICIÊNCIA E TEMPO DE ESTUDO, são indispensáveis se desejarmos ponderar sobre a influência da grafia na percepção fônica. Desse modo, espera-se que indivíduos mais proficientes na L_{NN} se valham menos de seu sistema grafêmico materno do que indivíduos menos proficientes, bem como se estima que indivíduos com mais tempo de estudo da L_{NN} reduzam seu apelo a aspectos grafêmicos da L₁. No expediente pedagógico, Bassetti (2008) elenca que fatores COMO CONTEXTO DE APRENDIZAGEM E MÉTODO DE ENSINO são substanciais no evento perceptual fônico. A estudiosa afirma que indivíduos que receberam instrução formal da L_{NN} podem ser mais propensos a se valer de transferências grafo-fônicas da L₁ para a L_{NN} do que indivíduos que aprenderam o idioma-alvo fora do contexto escolar. Paralelamente, materiais de ensino que empregam grande quantidade de informação escrita podem desencadear maior conexão entre aspectos grafêmicos e fônicos se comparados a materiais que se utilizam mais da oralidade.

A partir dos postulados de Bassetti (2008), reunimos algumas considerações que podem indicar uma alternativa para que consigamos captar a relação ou os efeitos do sistema grafêmico sobre a percepção fônica não nativa. Metodologicamente, um modelo de percepção fônica não nativa seria satisfatório ao contemplar a influência dessas variáveis sobre os resultados de seus testes perceptuais, oportunizando duas ferramentas empíricas: um questionário detalhado sobre a experiência linguística do aprendiz com a L_{NN}, de modo que os resultados obtidos nas tarefas fossem confrontados com as informações dispostas no instrumento de anamnese, e tarefas específicas que tratassem das correspondências grafo-fônicas vislumbradas pelo aprendiz, cujos resultados também pudessem ser interpretados em consonância com os das demais tarefas. Teoricamente, esse modelo deveria prever espaço para o tratamento de aspectos grafêmicos na análise dos resultados que dizem respeito às relações entre unidades fônicas da L₁ e da L_{NN}. Novamente, se retomarmos a noção de multimodalidade das informações acessadas pelos percebedores, nossa proposta de modelamento de percepção fônica não nativa pode vir a abrir esse espaço, pois conta com argumentos filosóficos e cognitivos que permitem e acomodam o caráter polivalente dos objetos do mundo.

4.4.3 Formação de novas categorias: o que segue a partir das assimilações

Mencionamos no segundo capítulo desta tese que o foco do modelo PAM são ouvintes sem qualquer experiência com o idioma-alvo. Assim, o modelo visa a estipular quais categorias da L₁

servem de base para que os novos elementos fônicos sejam a elas associados, resultando no que Best (1995) chama de assimilações a categorias nativas. No caso do *PAM-L2*, embora seja um modelo que possa levar em conta ouvintes que estão desenvolvendo a língua-alvo, isto é, aprendizes de uma L₂, os autores não se empenham em tentar descrever ou explicar de que modo ocorreria a formação de novas categorias, fenômeno possível de ocorrer longitudinalmente à medida que tais aprendizes acumulam experiência com a L₂¹¹³.

Avaliamos que um modelo de percepção fônica de línguas não nativas possa se valer tanto de casos de assimilação como de casos em que potencialmente haja a formação de novas categorias, distintas daquelas da L₁. Dito de outra forma, é conveniente que sejam endereçadas especificidades desse último caso, pois se trata de um passo adiante, de notável importância, no desenvolvimento de uma gramática fônica não nativa. Nesse quesito, recorreremos às considerações de Flege (1995):

A vasta gama de sons da L₂ pode, primeiramente, ser identificada em termos de um alofone posicional da L₁, mas, conforme os aprendizes de L₂ adquirem experiência no idioma-alvo, eles podem gradualmente distinguir a diferença fonética entre certos sons da L₂ e o som da L₁ mais próximo. Quando isso ocorre, uma representação da categoria fonética pode ser estabelecida para o novo som da L₂, representação essa que se torna INDEPENDENTE das outras que foram estabelecidas previamente para a L₁¹¹⁴ (FLEGE, 1995, p. 263, grifo nosso).

Salvaguardadas questões terminológicas, no que compete ao que é tradicionalmente visto como fonético ou como fonológico, e questões referentes à distinção entre os primitivos de ambas as propostas, julgamos que as ponderações de Flege (1995) sejam de extrema pertinência no que diz respeito a um modelo de percepção fônica de línguas não nativas, particularmente em sua aplicação a aprendizes que estão desenvolvendo o conhecimento linguístico veiculado na língua-alvo.

¹¹³ Bundgaard-Nielsen, Best e Tyler (2011a) afirmam que a percepção exitosa das unidades fônicas da L₂ prevê um aumento do número de palavras contrastantes na L₂, as quais, inicialmente, eram interpretadas como idênticas em função do sistema fonológico materno. Essa nova sintonização aos contrastes fônicos com base na otimização do léxico causaria, de acordo com os autores, um processo de refonologização. Cabe ressaltar que essa proposição parece ser satisfatória, mas, ainda que mencionada em Best e Tyler (2007), não é tratada com afinco – daí nossa motivação em recorrer aos postulados de Flege (1995) sobre o tema, a serem tratados no parágrafo seguinte. Todavia, o conceito de refonologização apresentado pelos autores é delineado com mais aprofundamento em uma proposta que almeja modelar a percepção fônica a partir da ampliação do léxico, chamada de MODELO DE SINTONIZAÇÃO VOCABULAR DA REFONOLOGIZAÇÃO EM SEGUNDA LÍNGUA [VOCABULARY-TUNING MODEL OF SECOND LANGUAGE REPHONOLOGIZATION (BUNDGAARD-NIELSEN; ET AL., 2012)].

¹¹⁴ No original: “According to the SLM, the full range of L₂ sounds may at first be identified in terms of a positionally defined allophone of the L₁ but, as L₂ learners gain experience in the L₂, they may gradually discern the phonetic difference between certain L₂ sounds and the closest L₁ sound(s). When this happens, a phonetic category representation may be established for the new L₂ sound that is independent of representations established previously for L₁” (FLEGE, 1995, p. 263).

Em nosso ponto de vista, assimilações da L_{NN} para a L₁ sempre deverão ocorrer no evento perceptual fônico quando os elementos acessados forem interpretados como sons da fala [do contrário, pode não haver assimilação, assim como adverte Best (1995)]. Diferentemente, a formação de novas categorias é um fenômeno que pode ou não ocorrer. Dessa maneira, advogamos que a operação conjunta de MECANISMOS DE RACIONALIZAÇÃO E EXPOSIÇÃO AO INPUT ao lado de FATORES TEMPORAIS E EXPANSÃO DO INVENTÁRIO LEXICAL irão responder pela possibilidade de formação das novas categorias. Em outras palavras, OS ESTÁGIOS DESENVOLVIMENTAIS do percebido são pontos determinantes para que categorias novas sejam formadas.

Em termos práticos, as novas categorias fônicas a serem formadas para dar conta das distinções funcionais da L_{NN} seriam destinadas exclusivamente ao idioma não nativo, divergindo daquelas que servem para a manutenção de contrastes na gramática fônica materna. No que compete ao primitivo que dá sustentação às novas categorias, defendemos que se trata de uma unidade gestual, de cunho acústico-articulatório, aos moldes do que preconiza Albano (2001). Assim, para que possa existir como uma categoria fônica voltada ao idioma não nativo, a unidade gestual deverá diferir daquela utilizada na L₁ do aprendiz, configurando outros tipos de constrição, de magnitude, de faseamento e de aspectos temporais, ou seja, a nova categoria deverá se instanciar como outro gesto fônico, diferente dos demais existentes na gramática fônica materna.

A formação de novas categorias fônicas, segundo o que preconizamos neste trabalho, deverá ocorrer na medida em que o percebido adquire experiências cumulativas em relação ao conhecimento da L_{NN}. Em termos essencialmente fônicos, formar novas categorias que estejam a serviço do sistema não nativo implica que o aprendiz realize, constantemente, o exercício de testar a variabilidade dos gestos em termos de suas bordas. Essa ponderação é de extrema importância teórica, já que a projeção simbólica dos gestos diz respeito às suas bordas (ALBANO, 2001). Essa atividade, que permite ao percebido experimentar maior ou menor deslize pelas bordas gestuais reitera a noção de que podemos “aprender a fazer, fazendo”, também sustentada por Albano (2001). Com isso, a fronteira entre o que é físico e o que é abstrato se rompe, levando a gradiência para as representações fônicas dos aprendizes. Nesse quesito, as situações comunicacionais em que o aprendiz se encontra e a racionalização dos insumos ricos e multimodais advindos de suas oportunidades de interação são definitivas para que novas categorias fônicas se instanciem. Trata-se, pois, de um EXERCÍCIO FÔNICO subsidiado por MECANISMOS COGNITIVOS, em que ambos são captados por um ACESSO INDIRETO à realidade.

As novas categorias deverão estar alinhadas, também, ao que concebemos no âmbito cognitivo e filosófico. Em relação à esfera cognitiva, as novas categorias, aos moldes daquelas já existentes, devem ser vislumbradas como um símbolo representacional interno (mas completamente relacionado ao mundo externo), advindo de relações inferenciais e conhecimento prévio que combinam mecanismos de racionalização, referentes ao aparato cognitivo do aprendiz, e informações ricas e essencialmente multimodais, advindas do mundo observável e atestado pelo percebedor/aprendiz. Além disso, as novas categorias coadunam aspectos mentais/simbólicos e aspectos motores/físicos no que se refere à informação fônica disponível em seu entorno e a que ele tem acesso. Frisamos, porém, que a formação de novas categorias fônicas só poderá existir depois de ter havido processos de assimilação, já que o evento perceptual pressupõe que o inventário fônico da L_{NN} seja balizado pela L₁.

A observação que acabamos de ressaltar nos permite conectar o expediente cognitivo com o expediente filosófico, de caráter realista indireto, correspondente à outra esfera que compreende o segundo âmbito mencionado no início deste capítulo. Se as assimilações *PER SE* acontecem de maneira indireta por decorrerem da atuação do sistema fônico materno (um dos elementos intermediários na apropriação do inventário não nativo no momento do evento perceptual), então a formação de novas categorias, que só é possível se houver assimilações previamente, se torna uma instância derivada de um empreendimento já indireto.

Salientamos, portanto, que os mesmos referenciais cognitivo, fônico e filosófico que deram suporte à nossa proposta de percepção fônica não nativa em termos de assimilações podem ser resgatados para que possamos conceber uma tentativa de abordar o fenômeno que diz respeito à formação de novas categorias. Consequentemente, elas teriam as mesmas propriedades encontradas nas categorias da L₁ e dos elementos fônicos da L_{NN}, mas seriam decorrentes das assimilações e da relação entre mecanismos de racionalização mais exposição ao input e o acúmulo de experiência com a L_{NN}.

4.5 Um modelo alternativo para a percepção fônica de línguas não nativas?

Com base nas considerações que tecemos ao longo das seções 4.1, 4.2, 4.3 e 4.4.1, acabamos fornecendo argumentos que levam a uma nova caracterização do PAM-L₂, delineando aspectos teóricos que, coerentemente ajustados, podem explicar e responder, de forma alternativa, à percepção fônica

de línguas não nativas. Não temos, necessariamente, a pretensão de instaurar um novo modelo de percepção fônica não nativa, mas compreendemos que seja fundamental oferecer uma perspectiva alternativa ao *PAM-L2* em termos teórico e metodológicos. Cunhar um novo modelo poderia vir a requerer mais adaptações, e talvez em mais âmbitos, do que aquelas que aqui propomos. No entanto, as modificações que sugerimos podem nos levar, futuramente, a derivar outro modelo perceptual.

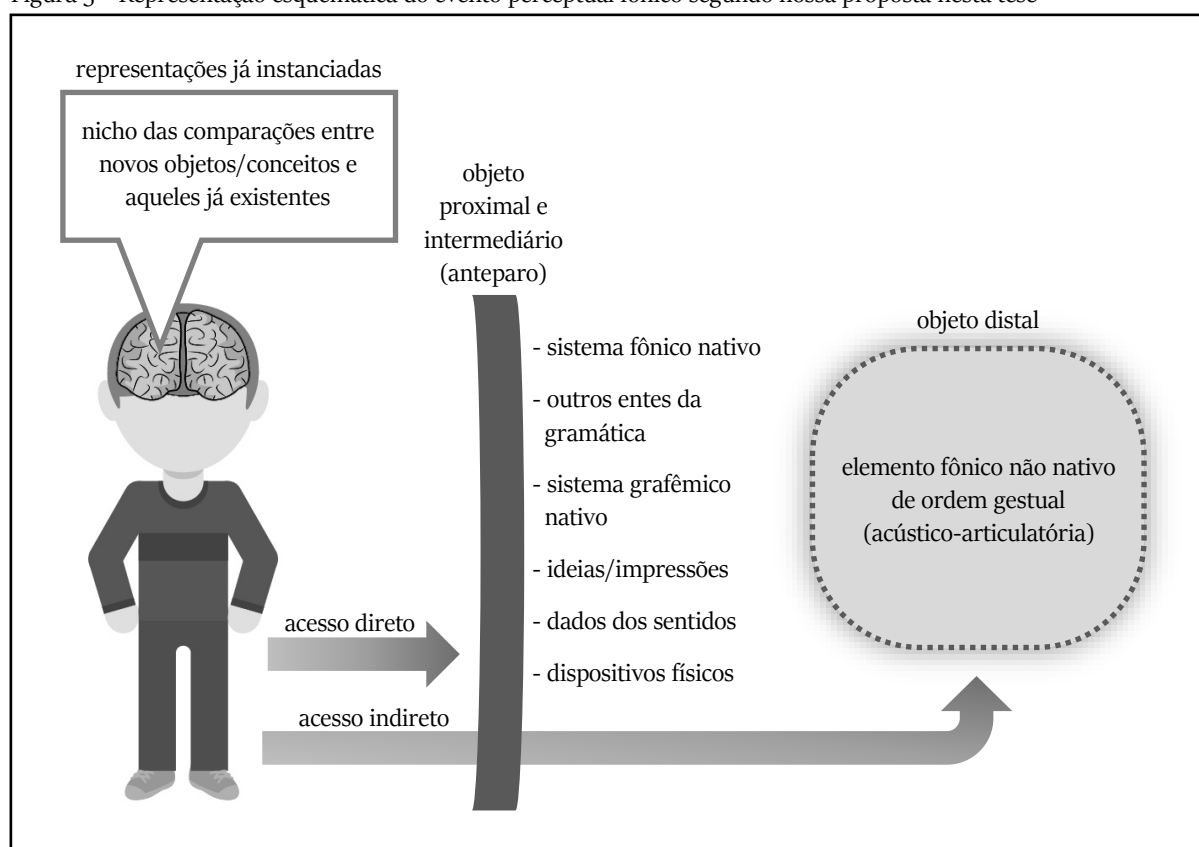
Sáimos de um modelo que dispensa operações cognitivas e trouxemos à tona a necessidade de encarar o evento perceptual sob uma perspectiva da cognição orientada pelas neurociências e pela psicologia. Confrontamos um modelo cuja unidade gestual era tratada por vias articulatórias com uma proposta cuja unidade gestual recebeu um tratamento acústico-articulatório. Migramos de um modelo calcado no realismo direto para uma concepção filosófica (com implicações psicológicas) assentado sobre os pressupostos do realismo indireto. Assim, ao demarcar, como ponto de partida, que a experiência perceptual fônica se configura como um empreendimento cognitivo por excelência, fazemos uso de uma unidade gestual que tem compromisso tanto com o expediente abstrato, simbólico, quanto com o expediente concreto, físico. Isso garante um *ELO* não somente com a cognição, ao tratar de mecanismos abstratos e da internalização de propriedades do mundo físico, mas também com o realismo indireto, que nos faculta conceber a percepção dos objetos do mundo através de um elemento intermediário, como, por exemplo, o sistema fônico materno.

Destacamos, outrossim, a extensão de um modelo que só se voltava para a percepção fônica em ambiente de *L2* para um modelo que engloba percebedores de maneira geral, independentemente de o contexto ser *L2* ou *LE*, uma vez que ilustramos a ineficiência de tal dicotomia. Insinuamos, também, ainda que de maneira incipiente, a necessidade de considerar os aspectos grafêmicos no evento perceptual em que se inserem os elementos fônicos, e procuramos expor a provocação de que, além de assimilações, há que se pensar na formação de novas categorias fônicas. Especificamente, no que compete a esses dois últimos pontos, ainda que timidamente, postulamos que nossas considerações sejam relevantes à área na medida em que explicitam possíveis meios de diagnosticar a influência do código escrito na percepção de elementos fônicos não nativos, e coloca em pauta a necessidade de irmos além das assimilações como sendo o fenômeno resultante da percepção fônica não nativa.

Algo de extrema relevância a ser apontado é o fato de incorporarmos, integralmente, em nossa concepção de percepção fônica, as previsões de assimilação veiculadas em Best e Tyler (2007). Com base no levantamento dos estudos realizado no terceiro capítulo desta tese, resgatamos tais

previsões por entender que elas dão conta de estabelecer com sucesso as relações existentes entre categorias fônicas maternas e gestos fônicos não nativos. Dessa forma, o quadro epistemológico em que nos posicionamos é completamente distinto ao do *PAM-L2*, mas nossa visão sobre as relações fônicas que subjazem às previsões de assimilação encontra-se, de modo geral, em consonância com a dos proponentes desse modelo. Conseqüentemente, em termos procedimentais no que concerne a uma investigação empírica, partiríamos de hipóteses baseadas nas previsões de assimilação de Best e Tyler (2007), porém, teríamos como plano de fundo um arcabouço voltado à percepção como um fenômeno cognitivo, cujas unidades são elementos acústico-articulatórios e cuja realidade é indireta. As implicações teóricas de nosso posicionamento endereçam uma reconcepção do que seja o evento perceptual fônico, e as implicações metodológicas de nossa proposta residem em possíveis diretrizes para a condução de estudos experimentais e abordagens empíricas na área (a serem discutidas no próximo capítulo). A seguir, apresentamos a Figura 5, que corresponde a uma representação esquemática acerca de como concebemos o evento perceptual fônico de uma LNN:

Figura 5 – Representação esquemática do evento perceptual fônico segundo nossa proposta nesta tese



Fonte: elaborada pelo autor.

Para fins de delimitação, portanto, a tese que defendemos é a de que a percepção fônica de línguas não nativas corresponde a um fenômeno *COGNITIVO* por excelência, cujas unidades são representadas por *GESTOS DE ORDEM ACÚSTICO-ARTICULATÓRIA* e são acessadas *INDIRETAMENTE* a partir do filtro do sistema materno e de ideias e hipóteses que levantamos sobre o universo que nos cerca.

4.6 Resumo do capítulo

O objetivo deste capítulo foi apresentar contra-argumentos ao construto do *PAM* e do *PAM-L2*. Iniciamos nossa teorização chamando atenção para o fato de os eventos perceptuais serem um fenômeno cognitivo, fazendo referência à literatura das neurociências e da psicologia cognitiva. Em seguida, argumentamos que a unidade da percepção fônica é de cunho gestual, porém é mais apropriadamente concebida se receber o tratamento teórico de Albano (2001), uma vez que são revistas relações entre gestos e outras unidades gramaticais e se faz igualmente importante a face acústica dos elementos gestuais (o que corresponde a ganhos teóricos e metodológicos, por exemplo). Então, discorremos sobre o realismo indireto quanto à percepção fônica de línguas não nativas, apontando sua preponderância sobre o realismo direto, dadas as circunstâncias que regem o fenômeno perceptual enquanto atividade mental e real.

A congregação desses três elementos – cognição, unidade gestual de ordem acústico-articulatória e realismo indireto – nos permitiu transitar entre o plano abstrato, mental e simbólico e o plano concreto, motor e físico com bastante fluidez, dado o paralelismo que apontamos haver entre eles. Desse modo, temos uma aproximação epistemológica, a partir da qual a cognição opera através de mecanismos de racionalização e inferências aliados às informações disponíveis no mundo físico; o gesto de ordem acústico-articulatória opera através das relações entre o mental e o motor, entre simbolismo e ação concreta; e o realismo indireto opera através da existência de um mundo real, observável, acessado indiretamente através de elementos intermediários (que podem ser de caráter mental ou físico).

A título de tecermos considerações sobre questões que seriam apenas decorrentes de nossa proposta, problematizamos sobre o fim da dicotomia entre *L2 VS. LE*, afirmando que ela não se mostra relevante em função de que, no realismo indireto, a noção ecológica da percepção se perde. Subsequentemente, discutimos sobre a influência do sistema grafêmico na percepção fônica não nativa e como tentar prevê-la e, por fim, tratamos da possibilidade de haver a formação de novas

categorias fônicas relativas a gestos não nativos a partir da experiência do percebedor com a gramática fônica não nativa (aprendizagem fônica), os quais são decorrentes, necessariamente, de um processo prévio de assimilação.

Refletir sobre esses três últimos pontos, ao que sugerimos, parece se mostrar mais consistente epistemologicamente segundo nossa proposta do que segundo a de Best e Tyler (2007). Primeiramente, em função de a dicotomia L_2 -LE não se sustentar segundo o realismo indireto, e também porque as razões trazidas pelos autores se mostram bastante limitadas em diversos aspectos. Em segundo lugar, em virtude de os autores não problematizarem o impacto das relações grafêmicas na percepção de um elemento fônico não nativo, tarefa por nós empreendida, ainda que não substancialmente. Em terceiro lugar, ao avançar na discussão sobre a formação de novas categorias, julgamos ter dado um passo à frente em relação ao modelo de Best e Tyler (2007), tópico essencial a ser abordado quando se trata da percepção fônica não nativa, uma vez que os autores a endereçam periféricamente em sua teorização.

5 PERCEPÇÃO FÔNICA E PROCEDIMENTOS EMPÍRICOS DE TESTAGEM

Naturalmente, ao conceber a língua como uma associação entre propriedades abstratas e físicas, o trabalho do linguista que se dedica à dimensão fônica das estruturas linguísticas deve residir largamente sobre fatos observáveis, os quais resultam de enunciados produzidos pelos falantes. A observação de tais fatos abre caminho para uma série de análises, dedutivas e/ou indutivas, que visam a responder sobre a organização do sistema linguístico em questão. Mais precisamente, para os propósitos desta tese, as “escolhas” que um ouvinte faz acerca de determinado conjunto de estímulos pode revelar aspectos perceptuais relevantes acerca de sua gramática fônica não nativa.

Nos textos que lançam o *PAM* e o *PAM-L2*, nada consta sobre quais seriam os procedimentos empíricos de testagem mais efetivos para que pudéssemos inferir, com relativa segurança, as relações de categorização estabelecidas pelo ouvinte quando exposto a estímulos linguísticos. Entretanto, é bem verdade que, em outras publicações¹¹⁵, Catherine Best e Michael Tyler empregam mecanismos metodológicos para a verificação das categorizações fônicas, mas há diferenças procedimentais a depender da configuração tipológica dos idiomas em questão e dos objetivos de pesquisa que permeiam os estudos. Acreditamos que a ausência de sugestões metodológicas nos dois textos seminais dos modelos decorra de questões referentes à delimitação de espaço nos veículos que difundem os artigos científicos, os quais precisariam se limitar a poucas dezenas de páginas para uma dupla tarefa: apresentar formalmente os modelos e apontar meios empíricos de testagem.

Conforme relatamos na introdução, este capítulo se sustenta a partir de uma discussão sobre os procedimentos empíricos de testagem, sugerindo quais procedimentos metodológicos podem ser mais profícuos e operativos para que possamos aferir adequadamente as categorizações fônicas (sejam elas referentes à assimilação perceptual à L_1 ou referentes à formação longitudinal de categorias no próprio desenvolvimento fônico não nativo) a partir da noção de percepção fônica defendida nesta tese. Desse modo, revisitaremos as pesquisas apresentadas no terceiro capítulo deste trabalho e, com base nelas, avaliaremos o que podem nos oferecer metodologicamente para que

¹¹⁵ Previamente ao lançamento do *PAM*, Catherine Best já vislumbrava empiricamente tentativas de compreender as categorizações fônicas estabelecidas pelos ouvintes que observava (BEST; McROBERTS; SITHOLE, 1988; BEST, 1990, 1993, 1994a, 1994b; BEST; STRANGE, 1992;), mas nunca reuniu seus estudos de forma a propor uma metodologia de verificação. O mesmo ocorreu após Catherine Best e Michael Tyler difundirem o *PAM-L2*: alguns procedimentos metodológicos para testar as categorizações fônicas foram empregados em suas pesquisas (BUNDGAARD-NIELSEN; BEST; TYLER, 2011b), mas nada organizado de maneira mais sistemática até o presente.

conseguamos obter, segundo as premissas de base às quais nos alinhamos, constatações empíricas mais confiáveis no que compete à determinação das categorizações fônicas em línguas não nativas.

O presente capítulo, que contempla o objetivo (f) desta tese, se fundamenta sobre três seções de conteúdo e uma seção de resumo, que encerra o capítulo. A primeira, engendrada a partir de questionamentos de Alves e Silva (2016), discute a aproximação empírica sob a perspectiva realista direta do *PAM* e do *PAM-L2*, já que parece haver um distanciamento entre pressupostos do realismo direto e da abordagem ecológica e pressupostos metodológicos adotados em laboratório pelos proponentes dos modelos. A segunda retoma os estudos apresentados no terceiro capítulo e examina sua organização metodológica e suas abordagens laboratoriais a fim de problematizar a taxionomia, o delineamento empírico, o conhecimento de base testado e o objetivo das tarefas, com base em Gerrits (2001), Schouten, Gerrits e Van Hessen (2003), Gerrits e Schouten (2004) e em Colantoni, Steele e Escudero (2015), bem como apontar a pertinência dos métodos empregados para a verificação da percepção fônica. A terceira toca em questões metodológicas no que se refere à aplicação de tarefas perceptuais para fins de futuras testagens a partir do modelo advindo desta tese, de modo a refletirem os princípios epistemológicos do modelo aqui proposto, no sentido de darem conta acuradamente das previsões de ASSIMILAÇÃO PERCEPTUAL à L₁ e da formação de categorias que respondem pelo DESENVOLVIMENTO FÔNICO não nativo.

5.1 O realismo direto e a arquitetura empírica da pesquisa no âmbito da percepção de sons linguísticos

Quando tratamos de um tema que se institui de maneira interdisciplinar, como é o caso da percepção de sons linguísticos no *PAM* e no *PAM-L2*, que congrega tanto aspectos fônicos como aspectos filosóficos (com implicações psicológicas), e que tem um compromisso empírico, faz-se fundamental haver consistência interna para que o construto a ser delineado se sustente. No momento em que é explicitada a vinculação psicológica, filosófica e/ou linguística de um modelo que se coaduna com investigação experimental, teoria e metodologia devem ser não só complementares, mas, também, consistentes entre si. É neste aspecto que a presente seção propõe contribuir, uma vez que discute o distanciamento existente entre testagem empírica e o eixo realista direto/ecológico do *PAM* e do *PAM-L2*, e advoga a favor da vertente realista indireta, já defendida no capítulo anterior, como sendo a epistemologicamente coerente com testes laboratoriais de percepção fônica não nativa.

A propósito de problematizar teoria de base e metodologia de pesquisa empírica a partir do *PAM/PAM-L2*, Alves e Silva (2016) lançam questionamentos indispensáveis acerca dessa relação, sobre os quais nos deteremos nesta seção. Nesse mérito, os autores iniciam a discussão com base na seguinte ponderação:

Considerando-se a noção ‘realista direta’ da perspectiva que motiva o *PAM-L2*, o primeiro questionamento a ser feito diz respeito a uma questão metodológica, levando-nos a indagar o quão diretos podem ser considerados estudos laboratoriais que se baseiam nesse modelo. Afinal, testes de percepção que se valem de sons gravados, possibilitando o contato com um estímulo não-advindo de uma interação, cujas tarefas levam os participantes a, simplesmente, clicar em botões para identificar ou discriminar sons (sem um fim comunicativo real), poderiam, efetivamente, ser considerados instâncias de uma abordagem ecológica, de modo a retratarem um ambiente real, percebido diretamente (ALVES; SILVA, 2016, p. 116)?

Conforme os próprios pesquisadores apontam, não há qualquer ingenuidade nesse questionamento, e sua resposta irá depender da acepção do termo “direto” que está em jogo. Alves e Silva (2016) salientam que, se nos valermos da acepção¹¹⁶ segundo a qual o evento perceptual não é mediado por mecanismos cognitivos, a percepção seria, de fato, direta para Best e Tyler (2007). Porém, os autores chamam atenção para outra interpretação¹¹⁷ do termo, a qual assume que não deve existir qualquer objeto ou aparato mediador alocado entre o indivíduo e o ambiente. Assim sendo, não podemos esperar que o evento perceptual no âmbito laboratorial seja direto, uma vez que existe um ELEMENTO COMPUTACIONAL que figura entre o percebedor e o dado a ser percebido.

Embora os autores não toquem no quesito da segunda acepção do referido termo, a de que a percepção é direta se não for o produto da percepção de outro indivíduo (percepção em segunda mão), os testes de percepção também encaram dificuldades conceituais no que compete à sua natureza direta, já que, durante a elaboração das tarefas, os estímulos são selecionados e indexados como “certos ou errados” pelo próprio experimentador, antes mesmo de serem apresentados aos sujeitos da pesquisa. Em outras palavras, somente em uma das três acepções contempladas, o evento perceptual em contexto de laboratório poderia ser encarado como uma experiência direta – e, mesmo assim, essa acepção é alvo de severas críticas, como esclarecemos no quarto capítulo deste trabalho.

¹¹⁶ Nesse caso, trata-se da terceira acepção do termo “DIRETO” a que fazemos menção no segundo capítulo desta tese.

¹¹⁷ Referida, nesta tese, como a primeira acepção do termo “DIRETO” veiculado na obra de J. Gibson (1986).

Alves e Silva (2016, p. 117) perguntam sobre o quão mediados seriam os estudos laboratoriais e se “poderiam ser representantes de uma percepção direta”. Ao que nos parece, um estudo laboratorial sempre será mediado porque sempre irá ocorrer através de uma FERRAMENTA que possibilite a apresentação de um estímulo e propicie o fornecimento de uma resposta. Em contexto de experimentação, independentemente de qual seja a ferramenta ou o anteparo que esteja mediando o evento perceptual, independentemente de suas características ou de sua quantidade, o fato de haver mediação implica, necessariamente, que se trata de uma percepção indireta.

Outro questionamento de extrema relevância apresentado por Alves e Silva (2016), muito alinhado ao que acabamos de apresentar, diz respeito a quão direta se faz a percepção de estímulos manipulados acusticamente e/ou sintetizados em laboratório. Inevitavelmente, essa questão esbarra na caracterização ecológica dos estímulos auditivos empregados nos testes de percepção e poderia ser respondida através das observações feitas nos parágrafos anteriores. Mas há que se notar que, em se tratando de estímulos criados ou modificados, eles podem não fazer parte do meio em que se encontra o percebido, pois fogem à naturalidade dos elementos que são encontrados no ambiente. Ademais, as propriedades invariantes de tais estímulos seriam virtuais, e não reais, dado que foram forjadas a partir de uma maquinaria que altera, em maior ou menor grau, a realidade circundante. Logo, a percepção de objetos manipulados ou criados computacionalmente poderia ser de ordem direta [se ligada à concepção (problemática) de que o evento perceptual dispensa processos inferenciais e mecanismos cognitivos], mas os estímulos em si não teriam o compromisso de serem fiéis ao cenário natural e tácito da percepção. Nesse sentido, Alves e Silva (2016) entendem que os estímulos manipulados não podem ser considerados ecológicos.

Pode decorrer do que ponderamos no parágrafo anterior o entendimento de que, se estímulos naturais forem utilizados em testes perceptuais, então, podem ser concebidos como ecológicos e percebidos de maneira direta. Defendemos que esse possível entendimento é errôneo. Por mais que o pesquisador não altere computacionalmente as propriedades acústicas e/ou visuais desses estímulos, isto é, por mais próximos que eles sejam aos estímulos disponíveis no ambiente, tais informações são armazenadas em um dispositivo por meios magnéticos (discos rígidos, *HDS*), óticos (discos compactos, *CDs*) ou eletrônicos (cartões de memória, *SSDs*), resultando em sua completa transfiguração. Nesse sentido, perde-se o que há de ecológico no estímulo. Essas informações tampouco são acessadas diretamente pelos percebidos, pois, além de estarem estocadas em um

equipamento, encontram-se codificadas: é necessário que sejam lidas por um mecanismo específico, como, por exemplo, um feixe de laser de alta precisão projetado sobre a superfície de um *CD*.

Ousamos afirmar, portanto, que de pouco adiantaria a percepção ser direta se os elementos a serem percebidos não correspondem à realidade observada ou vivenciada pelo sujeito que realiza a experiência perceptual. De fato, utilizar estímulos armazenados em qualquer tipo de mídia ou dispositivo, assim como insumos acusticamente alterados e/ou produzidos em laboratório, representa, no mínimo, um contrassenso dentro de um referencial teórico realista direto. Se trouxermos à tona o que discutimos no capítulo anterior, podemos presumir que, diferentemente do que ocorre com o realismo direto, o emprego de estímulos manipulados ou sintetizados não contradiz um aporte realista indireto, proposição que fica evidente se reconhecermos que qualquer informação objetiva proveniente do mundo externo passa por um filtro subjetivo para que seja por nós interpretada. Novamente, segundo os postulados do realismo indireto, não se faz relevante problematizar o *STATUS* ecológico/não ecológico do ambiente circundante e o caráter real/virtual dos objetos a serem percebidos, uma vez que nunca os acessamos diretamente. Assim, podemos igualmente utilizar, em sua totalidade, os recursos de que o ambiente dispõe, mas esse processo é dependente de um objeto intermediário que nos faculta avaliar e apreender o que é pertinente em determinada situação, sejam eles objetos reais ou virtuais que estejam em jogo.

Na mesma direção do que foi problematizado no parágrafo anterior, encontra-se o questionamento feito por Alves e Silva (2016) referente ao *DESIGN* dos estudos de percepção. Os autores ilustram a discussão fazendo referência a uma simples tarefa de identificação perceptual, que geralmente emprega opções de resposta baseadas em convenções escritas, sejam elas da ordem de um alfabeto fonético ou grafêmico. Nesse expediente, é inexorável que o participante relacione as propriedades do que está sendo percebido a uma simbologia específica de representação escrita. Isso quer dizer que a testagem inclui, em qualquer circunstância, endereçar um conhecimento de cunho enciclopédico que faz parte de uma convenção instaurada culturalmente, que assume uma faceta diferente (embora relacionada) do estímulo a ser aferido. Em termos gibsonianos, essa aproximação empírica teria nada ou pouquíssimo a dizer sobre a fidedignidade do evento perceptual.

De todas as questões trazidas por Alves e Silva (2016), as quais não foram tocadas, aqui, em sua exaustividade, podemos resumir a intenção dos autores na saga de indagar sobre os paralelismos entre procedimentos laboratoriais de testagem e uma perspectiva realista direta/ecológica correspondente. No limite, sob os prismas ecológico e realista direto, qualquer tentativa de tarefa

perceptual parece se distanciar, epistemologicamente, das práticas experienciadas pelo sujeito percebedor em seu ambiente. Essa constatação se justifica tanto a partir do instrumento de coleta de dados como a partir da própria natureza do estímulo em questão. No primeiro caso, milita sobre o evento perceptual o TIPO DE TAREFA a ser executado (tarefas de identificação, discriminação, transcrição, repetição, etc.), que sempre irá se configurar através de um dispositivo ou de uma ferramenta, computacional ou física, a qual será sempre mais ou menos mediada/indireta. No segundo caso, é a NATUREZA DO ESTÍMULO que incide sobre o quão ecológico é o evento perceptual: quanto mais alterado computacionalmente, menos ecológico é o estímulo; em última instância, se criado através de sintetizadores, arriscamos afirmar que não se trata, de maneira nenhuma, de um estímulo ecológico.

A situação parece, entretanto, ser outra (e bem mais favorável) se nos posicionarmos em favor de uma postura indireta em relação ao evento perceptual, conforme defendida nesta tese. Nessa esfera, levando em consideração que sempre existe um aparato intermediário alocado entre o percebedor e o objeto a ser percebido, testes laboratoriais que se utilizam de equipamentos específicos para captar as respostas dos participantes são completamente aceitáveis, não repercutindo em violações epistemológicas nesse paradigma. Além disso, como interações ecológicas ou não ecológicas são indiferentes para o rol de discussão do realismo indireto (porque elas não informam sobre o evento perceptual), a natureza do estímulo, segundo argumentamos neste trabalho, não é decisiva para que haja apreensão de conhecimento a partir do fenômeno perceptual. Nesse mérito, a partir da leitura do quarto capítulo desta tese, deve ter ficado claro que, na esfera realista direta de James Gibson, as relações entre percebedor e objeto a ser percebido devem ser ecológicas porque essa é uma das maneiras mais fundamentais pelas quais o evento perceptual ocorre, na medida em que os sentidos do percebedor atuam como sistemas de percepção orientados aos elementos invariantes que compõem os objetos do ambiente circundante. Esse direcionamento dos sentidos às concessões toma forma graças à noção de que sujeito e ambiente fazem parte de um sistema único; um sistema, portanto, ecológico. Consequentemente, alinhar-se a uma perspectiva realista indireta favorece a coadunação de procedimentos metodológicos com teoria de base em um modelo perceptual.

O debate promovido por Alves e Silva (2016) retrata uma reflexão primorosa no que compete à congregação de aspectos filosóficos e empíricos, que, além de servir de ferramenta para que outros pesquisadores não sejam ingênuos nas suas escolhas metodológicas, funciona como argumento adicional para que possamos eleger a versão indireta do realismo como a mais apropriada das duas para tratar da percepção no plano experimental. É nesse sentido que destacamos a grande

potencialidade do realismo indireto, defendido nesta tese, em se coadunar com procedimentos experimentais: se os postulados do realismo indireto acerca da existência de um objeto intermediário entre o sujeito percebedor e o ambiente que o cerca estiverem corretos, e estamos fortemente inclinados a assumir que sim, estão, então a natureza do estímulo percebido (real/virtual) e as condições ambientais em que ocorre o evento perceptual (ecológicas/laboratoriais) são satisfatoriamente captadas por tal posição filosófica, a qual, para os nossos propósitos, suplanta a sua contraparte direta.

Já que estamos endereçando aspectos experimentais, a seção que inicia a partir deste momento tem como objetivo revisitar os estudos apresentados no terceiro capítulo desta tese e examinar sua organização metodológica e suas abordagens empíricas, a fim de discutirmos a taxionomia das tarefas empregadas em testes perceptuais e apontarmos a pertinência dos métodos utilizados para a verificação da percepção fônica.

5.2 Discussão metodológica dos estudos perceptuais e suas implicações para uma nova caracterização do modelo

No terceiro capítulo desta tese, apresentamos diversos trabalhos no âmbito da percepção de sons linguísticos, dentre os quais alguns tiveram enfoque explícito nas previsões assimilatórias do *PAM-L2*, mas outros não. O propósito de tal levantamento era o de verificar a exaustividade das previsões de assimilação elencadas no referido modelo, por isso nos preocupamos em descrever cada estudo em linhas gerais. Nesta seção, todavia, o *LOCUS* da discussão irá compreender os aspectos metodológicos de tais estudos, com o intuito de tecermos considerações sobre a taxionomia, os propósitos e a pertinência das abordagens empíricas utilizadas nos trabalhos. Nesta seção, serão feitas referências não só a estudos que enfocaram as predições do *PAM-L2*, mas também àqueles que não as utilizaram, já que, nos dois conjuntos de trabalhos, o que se quis encontrar foram as possíveis categorizações fônicas feitas pelos sujeitos investigados, e em todas as pesquisas houve o emprego de metodologias experimentais.

Ao examinar os trabalhos apresentados previamente nesta tese, observamos que as tarefas envolvendo a percepção de elementos fônicos não nativos são rotuladas de acordo com diferentes designações, embora todas estejam relacionadas de alguma forma. Essa constatação dá margem para o uso relativamente aleatório de diferentes nomes de tarefas que, no final das contas, sugerem o mesmo tipo (ou um tipo muito similar) de tarefa. Um exemplo claro disso são os trabalhos de Best e

Hallé (2010) e de Bohn e Best (2012), em que o primeiro atribui à tarefa a ser realizada o nome de CATEGORIZAÇÃO e o segundo atribui o nome de IDENTIFICAÇÃO ao mesmo padrão de experimento (a rotulação de elementos fônicos). Assim, há uma pluralidade de termos que, sob a concepção adotada nesta tese, nem sempre caminham por vias diversas, mas que se apresentam terminologicamente como distintos, podendo acarretar confusões ou desentendimentos.

As discussões metodológicas que faremos nesta seção dizem respeito a testes perceptuais que refletem duas caracterizações: (i) quantidade e inter-relação de elementos fônicos que são solicitados aos participantes perceberem, e (ii) tipo de conhecimento testado, gerando constatações para assimilações, para dissimilações (formação de novas categorias) e, em algum nível, para ambas.

Em termos da quantidade e da inter-relação de elementos fônicos em tarefas de percepção, existem aquelas que se refere à apresentação de UM ÚNICO ELEMENTO FÔNICO em um dado estímulo, em que o participante deve indicar, dentre as opções disponíveis e somente em termos de tais opções, qual é o elemento fônico que está sendo reproduzido. Outras tarefas concernem à apresentação de DOIS OU MAIS ELEMENTOS FÔNICOS em um dado estímulo, presumindo que o participante estabeleça comparações entre os itens ouvidos. Realizadas as comparações, pode-se perguntar ao participante, por exemplo, se, dos três itens fônicos ouvidos em sequência no mesmo estímulo, o último item ouvido assemelha-se ao segundo item ou ao primeiro item – o que configura um formato de tarefa conhecido como ABX (X é aquele que deve ser comparado aos dois elementos prévios)¹¹⁸.

No que tange ao tipo de conhecimento averiguado, testes perceptuais podem englobar tarefas que sejam mais eficientes para registrar assimilações e tarefas que sejam mais sugestivas para responder por dissimilações (formação de novas categorias). Com base na literatura em percepção fônica não nativa a que tivemos acesso, podemos dizer que as assimilações tendem a ser medidas coerentemente através de tarefas de assimilação. As dissimilações, por sua vez, tendem a ser mensuradas adequadamente através de tarefas de identificação e de discriminação, mas também podem ser utilizadas para inferir, mesmo que indiretamente, assimilações.

As subseções 5.2.1, 5.2.2 e 5.2.3, apresentadas a seguir, irão tratar das características das tarefas perceptuais mencionadas, levando em consideração o que propõem e fazendo referência a alguns estudos ilustrados no terceiro capítulo desta tese.

¹¹⁸ Outros formatos de tarefas semelhantes a esse serão abordados adiante neste capítulo.

5.2.1 Tarefa de assimilação perceptual

A primeira tarefa perceptual de que iremos tratar encontra-se sob o nome de TAREFA DE ASSIMILAÇÃO PERCEPTUAL. Nesse tipo de tarefa, solicitamos explicitamente aos participantes que eles deverão ouvir um item fônico não nativo, apresentado em um estímulo no momento da testagem laboratorial, e rotulá-lo de acordo com aqueles que compõem seu inventário da L₁. Ao utilizar uma tarefa de assimilação perceptual, estamos interessados em compreender os processos assimilatórios que transitam na via L_{NN}-L₁. Nesse quesito, direcionamos os participantes a se utilizarem unicamente de sua gramática fônica nativa, de modo que ela sirva de base para a possível acomodação dos elementos fônicos não nativos.

A tarefa de assimilação perceptual, segundo Colantoni, Steele e Escudero (2015), geralmente, é acompanhada de uma tarefa de julgamento, conhecida como TAREFA DE ADEQUABILIDADE, a qual, frequentemente, envolve uma escala de Likert (nesse caso, escala de adequabilidade). A partir dessa escala, o participante, após ouvir o item fônico não nativo (primeira parte da tarefa), deve eleger o grau de proximidade ou distanciamento entre as categorias fônicas em jogo (segunda parte da tarefa). Os autores pontuam que as escalas de adequabilidade são utilizadas para se determinar a assimilação perceptual de certo elemento fônico em duas situações: quando os participantes possuem nenhum ou pouco conhecimento da língua-alvo ou quando os pesquisadores encarregados do estudo desejam observar como os elementos fônicos não nativos são alocados nas categorias nativas, em termos de mapeamento.

Deve ficar claro, portanto, que o propósito da tarefa reside em contemplar o quão semelhante ou distinto é o estímulo não nativo ouvido em relação à sua possível categoria fônica correspondente na L₁. Logo, uma tarefa de assimilação perceptual se concretiza a partir de dois “exercícios” distintos, mas complementares: a rotulação de um elemento fônico não nativo com base no inventário materno SEGUIDA da atribuição de adequabilidade do elemento fônico não nativo à(s) categoria(s) da L₁.

Defendemos o emprego de uma escala de adequabilidade nessa tarefa de múltipla escolha para a verificação da assimilação perceptual de elementos fônicos a categorias nativas justamente porque se faz como um mecanismo de o pesquisador, ante os resultados encontrados, vislumbrar as assimilações ensejadas pelos participantes no momento em que se ocuparam da tarefa. Como o grande propósito da assimilação perceptual é estabelecer quais são as categorias maternas utilizadas pelos participantes para acomodar os elementos fônicos não nativos, resulta imprescindível que tais

elementos sejam classificados em uma escala. Essa classificação pode vir a refletir diferenças de adequação à categoria, designada por Best (1995), conforme vimos no segundo capítulo desta tese, como *CATEGORY-GOODNESS DIFFERENCE*.

Argumentamos, pois, que a escala de adequabilidade não só responde pelo estabelecimento das categorias fônicas nativas que acomodam as não nativas, como também faculta ao pesquisador compreender os limites perceptuais resultantes das diferenças gestuais dos itens fônicos em questão. Se, de acordo com os preceitos de Albano (2001), com os quais concordamos, o local de projeção simbólica dos gestos são as suas bordas, então, segundo o que expusemos no capítulo anterior, contemplar a proximidade ou o distanciamento entre as fronteiras gestuais é de suma importância para verificarmos as relações de compatibilidade entre elemento não nativo e categoria fônica materna. Em outras palavras, diferenças de adequação à categoria encontram-se em estreita relação com a elucubração de Albano (2001) acerca das bordas dos gestos fônicos.

Em uma tarefa de assimilação, ainda que haja questões de múltipla escolha e que elas sejam categóricas, o fato de o participante julgar (através da escala de Likert) o quão próximo ou distante se encontra determinado item fônico da L_{NN} em relação à respectiva categoria materna pode indicar os limites entre os gestos de ambos os inventários. Assim, no momento em que se vale de suas categorias maternas para acomodar o item fônico não nativo, o participante responde essencialmente por questões *ASSIMILATÓRIAS*, ressaltando a simbiose entre L_{NN} - L_1 .

O número de pontos em uma escala de adequabilidade é variável. Nos estudos de Gilichinskaya e Strange (2010) e de Strange *ET AL.* (2011), os pesquisadores se utilizaram de uma escala de 9 pontos, em que o ponto 1 indicava um elemento fônico muito diferente daquele da L_1 , enquanto que o ponto 9 indicava um elemento fônico muito similar ao da L_1 . No estudo de Escudero, Sisinni e Grimaldi (2014), a escala utilizada era composta por 5 pontos, em que o ponto 1 indicava um elemento fônico muito diferente daquele da L_1 , enquanto que o ponto 5 indicava um elemento fônico muito similar ao da L_1 . Por outro lado, o trabalho de Holliday (2016), em vez de se utilizar de uma escala de Likert, dispôs de três opções com referência à L_1 dos participantes, a serem assinaladas após a audição do estímulo: “exatamente o mesmo”, “similar, mas um pouco diferente” e “muito diferente”. Nesse último caso, julgamos que as opções de adequabilidade sejam mais limitadas do que nos outros estudos, fato que pode obscurecer, em algum nível, o detalhamento das relações L_{NN} - L_1 , tão importante para o mapeamento dos sistemas fônicos em jogo.

Foram adotadas escalas de adequabilidade nos instrumentos de CATEGORIZAÇÃO de Antoniou, Best e Tyler (2012), Escudero, Simon e Mitterer (2012), e Antoniou, Best e Tyler (2013), possibilitando ao participante eleger o quão similar foi o estímulo ouvido àquele correspondente ao de sua L₁. Tais escalas também foram incorporadas às tarefas de IDENTIFICAÇÃO nos estudos de Bundgaard-Nielsen, Best e Tyler (2011a, 2011b) e Tyler *ET AL.* (2014), e à tarefa de ASSIMILAÇÃO proposta na pesquisa de Gilichinskaya e Strange (2010), por exemplo. Embora se apresentem sob rótulos distintos, as três tarefas propõem a mesma verificação: os participantes devem apontar o elemento fônico ouvido com base apenas em sua L₁, sugerindo, então, o quão adequado ele soa frente às categorias nativas. Reiteramos, nesse sentido, que o nome das tarefas designadas “categorização” e “identificação” incorre em deslizos taxionômicos se nos respaldarmos em Colantoni, Steele e Escudero (2015), já que a tarefa que prevê rotulação com base na L₁ mais escala de adequabilidade chama-se TAREFA DE ASSIMILAÇÃO PERCEPTUAL, e, não, tarefa de categorização ou de identificação, conforme os autores citados mencionaram em suas pesquisas.

5.2.2 Tarefa de identificação

A segunda tarefa perceptual a que nos reportamos é denominada TAREFA DE IDENTIFICAÇÃO. Conforme explanam Colantoni, Steele e Escudero (2015), em uma tarefa dessa natureza, o participante deverá ouvir um item fônico não nativo e relacioná-lo a opções de respostas que correspondam ao seu inventário NATIVO OU NÃO NATIVO. Todavia, como deve ter ficado claro a partir da caracterização da tarefa anterior, quando for o caso de as respostas se basearem nas categorias fônicas nativas, então a tarefa é comumente referida como TAREFA DE ASSIMILAÇÃO PERCEPTUAL. Desse modo, por questões operacionais, quando nos referirmos à tarefa de identificação, estaremos assumindo que se trata de uma tarefa em que o participante deva ouvir determinado elemento fônico não nativo e registrar sua resposta com base em seu CONHECIMENTO FÔNICO TAMBÉM NÃO NATIVO.

Entendemos que, ao solicitar que os participantes atribuam rótulos aos itens fônicos pertencentes ao inventário não nativo com base na própria L_{NN}, estamos empenhados em testar seus conhecimentos em termos da língua-alvo (certamente, o sistema de L₁ estará operando também nas tomadas de decisão, mas o foco da tarefa é a L_{NN} e não a L₁). Assim, neste tipo de tarefa, objetivamos detectar as categorias formadas pelos aprendizes ao longo de sua experiência com a L_{NN}, ou seja, entram em cena aquelas categorias que não são pertencentes à L₁, mas, sim, à L_{NN}. Nesse sentido,

prevê-se a recuperação de uma categoria que reside somente no âmbito não nativo – uma CATEGORIA NOVA, formada a serviço da LNN, a qual difere daquelas da L1, mas que tampouco precisa ser idêntica às dessa. Quando se vale de seu inventário não nativo em uma testagem laboratorial, o participante responde, sobretudo, por seu conhecimento das unidades fônicas em termos DESENVOLVIMENTAIS no que compete ao idioma-alvo. Baseados nessa definição, podemos ter uma TAREFA DE IDENTIFICAÇÃO.

Quanto às tarefas que se constituíram a partir da reprodução de um estímulo auditivo isolado e que solicitavam apenas que o participante informasse qual elemento fônico ele percebeu ou reconheceu¹¹⁹, podemos enumerar dois instrumentos (de acordo com a taxionomia dos próprios autores): um de CATEGORIZAÇÃO (BEST; HALLÉ, 2010; ANTONIOU; TYLER; BEST, 2012; ESCUDERO; SIMON; MITTERER, 2012; ANTONIOU; BEST; TYLER, 2013) e outro de IDENTIFICAÇÃO (ESCUDERO; CHLÁDKOVÁ, 2010; BUNDGAARD-NIELSEN; BEST; TYLER, 2011a, 2011b; ESCUDERO; VASILIEV, 2011; ESCUDERO; WILLIAMS, 2011; ALBUQUERQUE, 2012; BOHN; BEST 2012; ESCUDERO; WILLIAMS, 2012; SHAFIRO *ET AL.* 2012; PEROZZO, 2013; DUFOUR *ET AL.* 2014; STÖLTEN; ABRAHAMSSON; HYLSTENSTAM, 2014; TYLER *ET AL.* 2014; WILLIAMS; ESCUDERO, 2014; ALVES; ZIMMER, 2015; HOLLIDAY, 2016).

Podemos observar que os termos “categorização” e “identificação” foram utilizados intercambiavelmente nos estudos acima relacionados. Desse modo, sugerimos que o termo “TAREFA DE CATEGORIZAÇÃO” seja suplantado pelo termo “TAREFA DE IDENTIFICAÇÃO” por duas razões essenciais. A primeira é a de que, se já existe uma tarefa (a de identificação) que propõe o mesmo procedimento e que está devidamente referenciada na literatura da área, não há necessidade de se criar outra tarefa que assuma os mesmos pressupostos, porém divirja apenas no nome. A segunda, e talvez mais basilar, é a de que pode ocorrer de o participante da pesquisa não necessariamente ser capaz de encontrar determinada categoria para um dado elemento fônico (como ao associar, por exemplo, um som clique a um barulho qualquer, e não a um som da fala), possibilidade mencionada por Best (1995), o que faria com que o nome da tarefa, em nosso entendimento, perdesse seu sentido. Por

¹¹⁹ Uma tarefa de assimilação perceptual também prevê a reprodução de um estímulo auditivo isolado e também requer que o participante informe de qual elemento fônico se trata; porém, com base nas opções de múltipla escolha referentes à sua L1. No entanto, argumentamos que a tarefa de assimilação perceptual deva ser sempre subsidiada por uma escala de adequabilidade. Os estudos que contaram com tarefas de assimilação, retratados no terceiro capítulo desta tese, são os de Gilichinskaya e Strange (2010), Strange *ET AL.*, (2011) Escudero, Sisinni e Grimaldi (2014), Williams e Escudero (2014) e Holliday (2016). Embora a pesquisa de Williams e Escudero (2014) seja a única que não contemple uma escala de adequabilidade, os autores são suficientemente cuidadosos ao chamar seu instrumento apropriadamente de tarefa de assimilação perceptual, já que se tratava do reconhecimento de itens fônicos segundo a L1 dos participantes.

consequente, recomendamos a diluição da distinção puramente terminológica entre IDENTIFICAÇÃO e CATEGORIZAÇÃO, a partir da qual a tarefa seja chamada unicamente de TAREFA DE IDENTIFICAÇÃO.

5.2.3 Tarefa de discriminação

A última tarefa perceptual sobre a qual iremos discorrer corresponde à TAREFA DE DISCRIMINAÇÃO, uma opção de tarefa rica em detalhes e vasta em peculiaridades (teóricas e metodológicas). Esse tipo de tarefa se configura de maneira que o participante ouça dois ou mais itens fônicos não nativos em sequência, em um único estímulo auditivo, e os compare entre si (COLANTONI; STEELE; ESCUDERO, 2015).

As tarefas de discriminação diferem em formato e em objetivo, e também em relação às tarefas de identificação e de assimilação perceptual. Quanto aos formatos distintos, são utilizados os paradigmas AXB, ABX, XAB, AX e, ainda, os três primeiros juntos na tarefa conhecida como *ODDITY DISCRIMINATION*¹²⁰. Em relação ao primeiro grupo, os participantes devem responder se o elemento x é igual a outro, e, em relação ao da *ODDITY DISCRIMINATION*, os participantes devem responder qual é o elemento diferente ou destoante.

Para Colantoni, Steele e Escudero (2015), o formato AX é o mais simples de todos, pois prevê uma metodologia pouco elaborada e, em termos de demanda cognitiva, requer pouca escansão do conteúdo veiculado, o que é bastante positivo no caso de testagens envolvendo aprendizes de baixo nível de proficiência no idioma-alvo. Nesse sentido, Gerrits e Schouten (2004) pontuam que esse formato de tarefa de discriminação conta com baixa exigência da memória auditiva, o que facilita a sua execução. Por outro lado, tal formato tem recebido críticas em função de estar mais suscetível a captar diferenças puramente acústicas¹²¹ entre os elementos fônicos apresentados do que diferenças linguísticas propriamente. Ademais, segundo Gerrits e Schouten (2004), outra desvantagem desse paradigma diz respeito ao fato de os sujeitos poderem responder que os estímulos são diferentes somente se eles tiverem muita certeza de sua decisão, desencadeando respostas enviesadas na medida em que elas podem ser totalmente dependentes de critérios mais subjetivos.

¹²⁰ Referenciada nas tarefas que correspondem a *ODDITY VOWEL CATEGORIZATION* e *ODDITY VOWEL DISCRIMINATION*, as quais serão endereçadas mais adiante nesta seção.

¹²¹ Esse termo se distancia de nossa concepção de gramática fônica, mas optamos por mantê-lo de acordo com o que os próprios autores mencionam.

Diferenças “fonéticas ou fonológicas”, de fato, seriam bem acomodadas através da utilização dos paradigmas AXB e ABX , segundo esclarecem Colantoni, Steele e Escudero (2015, p. 97). Talvez essa seja uma das razões pelas quais esse formato foi largamente empregado nas pesquisas apresentadas nesta tese, já que, conforme elencam Gerrits e Schouten (2004), o paradigma ABX corresponde à tarefa de discriminação prototípica. A ideia por detrás desses formatos é a de que, com o emprego de três elementos fônicos em um dado estímulo, os participantes devem ignorar diferenças acústicas do sinal e processá-las enquanto unidades significativas, fônicas. Todavia, constatações importantes são mencionadas por Gerrits e Schouten (2004), que afirmam que, ao se considerar o tempo relativamente curto da memória auditiva para registrar eventos sonoros (de 200 a 300 ms), os índices de percepção categórica encontrados na tarefa do tipo ABX podem refletir o uso exclusivo do que os autores chamam de uma “memória fonética”. Desse modo, interpretamos que, para os referidos pesquisadores, tal memória estaria ligada ao fato de os participantes poderem tentar recuperar tanto a informação da memória auditiva quanto os rótulos atribuídos aos elementos fônicos que correspondem a A e B e, quando X for apresentado, tais informações auditivas podem ter desaparecido. Além disso, ao que observam Schouten, Gerrits e Van Hessen (2003), há grandes chances de, no paradigma ABX , os sujeitos reportarem que o elemento B é igual ao elemento X, também por questões referentes à recência em termos de memória.

No que concerne ao paradigma AXB , variante da opção ABX , há nebulosidade quanto a seus resultados. Conforme relatam Gerrits e Schouten (2004), há chances de que os sujeitos ignorem o terceiro estímulo, representando a anulação da esperada vantagem em relação ao paradigma ABX ; porém, o enviesamento das respostas é significativamente menor do que na versão ABX . Todavia, Gerrits (2001) explana que, por vezes, no formato AXB , os sujeitos podem fornecer suas respostas antes mesmo de o elemento B ser apresentado e, nesse caso, a tarefa AXB se resume simplesmente a uma tarefa AX . Um recurso metodológico possível para conter esse efeito, apresentado na tarefa de discriminação utilizado em Perozzo (2013)¹²², foi o de habilitar as opções de resposta somente após o término da reprodução do estímulo auditivo. Assim, os participantes registraram sua impressão inevitavelmente depois de terem ouvido toda a sequência de itens fônicos.

¹²² Cabe destacar que o formato da tarefa de discriminação utilizado em Perozzo (2013) foi o ABX . No entanto, o recurso que sugerimos também se mostra aplicável ao formato AXB .

Em relação ao formato X_{AB} , encontra-se pouca informação na literatura. Além de ser encarado como uma variação do paradigma ABX , seu conteúdo é brevemente sinalizado em Escudero e Williams (2012) e Elvin, Escudero e Vasiliev (2014). Curiosamente, esse formato foi utilizado pelos autores em tarefas de discriminação que possuíam tanto estímulos naturais quanto estímulos sintéticos. Especulamos, no entanto, que a mesma desvantagem em termos de memória encontrada no formato ABX possa ser comum ao paradigma X_{AB} , uma vez que espelha a mesma configuração. Contudo, pesquisas experimentais que tratem do tema podem dar maior sustentação a tal possibilidade [esse padrão não é testado em Gerrits (2001), a qual discute sobre formatos de tarefas perceptuais em sua tese].

Em termos de processos cognitivos, sugerimos que a tarefa de discriminação esteja mais proximamente relacionada com aqueles referentes à L_{NN} do que com a assimilação à L_1 em virtude de duas razões. A primeira é a de que esse tipo de tarefa incorre na comparação de um contraste genuinamente não nativo (que pode ou não ser funcional), a qual prevê a distinção de propriedades gestuais que acontecem essencialmente na L_{NN} e se manifestam para as finalidades próprias da L_{NN} . Advertimos, no entanto que, a depender do nível de proficiência do participante, poderá haver ou não o reconhecimento de um contraste funcional (caso esse seja configurado na tarefa). A segunda é a de que a tarefa envolve a percepção de propriedades não nativas que não têm a pretensão de se revelar ou conectar com a L_1 , isto é, NÃO HÁ UM COMPROMISSO COM O SISTEMA MATERNO DO PARTICIPANTE. Possivelmente, solicitar que, nessa tarefa, os participantes comparem os elementos fônicos do estímulo e também apontem a qual eles se relacionam na L_1 fugiria aos propósitos da tarefa de discriminação, bem como implicaria um trabalho perceptual bastante oneroso ao ouvinte¹²³.

Tarefas de discriminação com base na L_{NN} podem ser empregadas com o intuito de cobrir diferenças fônicas que não necessariamente veiculam distinções funcionais, atreladas ao que Best (1995) chama de variações de “ordem mais baixa” (*lower-order variations*), responsáveis por alofonias e contrastes no âmbito da articulação. O estudo de Nichols (2014) ilustra, por exemplo, a discriminação de um contraste alofônico no francês do Quebec por ouvintes norte-americanos. Assim, uma das atribuições de uma tarefa de discriminação seria a de estabelecer a percepção de propriedades fonéticas gradientes, conforme apontam Best e Tyler (2007). Além disso, Tyler *ET AL.*

¹²³ Por outro lado, *a posteriori*, tal tarefa pode sugerir, também, que os sons da L_2 ainda estão sendo assimilados à L_1 . Por exemplo, em resultados cujos graus de discriminação sejam mínimos, podemos ser levados a sugerir que ainda não houve dissimilações. Nesse quesito, é possível que os participantes ainda estejam ouvindo os estímulos a partir dos padrões assimilatórios do sistema materno.

(2014, p. 16) mencionam que “as previsões de discriminação do *PAM* dependem da distância fonética entre os fones assimilados¹²⁴”. Diferentemente, tarefas de identificação, que têm por base o conhecimento fônico não nativo, são associadas, geralmente, a distinções funcionais [a exemplo da pesquisa de Holliday (2016)], que remonta ao que Best (1995) define como sendo invariantes de “ordem mais alta” (*higher-order invariants*).

Contudo, Colantoni, Steele e Escudero (2015) mencionam que uma tarefa de discriminação no formato *AXB* pode dar conta de abarcar diferenças tanto fonéticas quanto fonológicas (em termos tradicionais), em vez de distinções que residiriam somente no âmbito da fonética (gradiência em e entre os gestos). Por outro lado, ao assumirmos uma perspectiva gestual, tanto aos moldes de Browman e Goldstein (1989, 1992) como aos moldes de Albano (2001), estamos rompendo as fronteiras entre fonética e fonologia, as quais parecem ainda permear os trabalhos de Best (1995) e de Best e Tyler (2007). Assim, nossa teoria de base gestual, de cunho acústico articulatório (ALBANO, 2001), tem vantagens diretas para um impasse metodológico que assola as tarefas de discriminação: se não há fronteiras entre aspectos gradientes da fonética e aspectos categóricos da fonologia, então, o problema de não se saber se estamos testando uma habilidade tida como fonética ou tida como fonológica desaparece¹²⁵.

O questionamento decorrente dessa posição é: se um dos papéis da tarefa de discriminação é apontar o contraste de propriedades fonéticas de “ordem mais baixa”, gradientes, em um conjunto de itens sonoros, então, qual é o papel dessa tarefa em uma perspectiva que dilui os limites entre aspectos gradientes e categóricos da gramática fônica? Para os propósitos desta tese, uma tarefa de discriminação é essencial para que possamos, de alguma maneira, mensurar O ALCANCE DAS BORDAS DOS GESTOS NÃO NATIVOS. Em outras palavras, a variabilidade contida em um estímulo que apresenta diferentes gestos fônicos se revela no julgamento que o participante faz acerca dos itens ouvidos, e as fronteiras perceptuais responsáveis pela associação de uma opção a outra (por exemplo, em um formato *ABX*, *X* ser associado a *A* em detrimento de *B*) advém da maior ou menor similaridade entre as variáveis que desencadeiam um gesto ou outro.

¹²⁴ No original: “[...] *PAM* discrimination predictions are dependent on phonetic distance between the assimilated phones [...]” (TYLER, *ET AL.*, 2014, p. 16). Cabe ressaltar que essa própria citação prevê a noção de dois níveis, o do fone (explícito) e o do fonema (implícito).

¹²⁵ Observação realizada verbalmente pela professora Adelaide Silva durante a etapa de qualificação desta tese.

Em estágios iniciais de desenvolvimento da LNN, o participante poderá estar menos ciente das propriedades fônicas que respondem pela gradiência de um gesto e de outro, e apresentar baixos índices de acuidade no que compete à discriminação de tais unidades. Todavia, em estágios mais otimizados de desenvolvimento da LNN, em que a sua experiência com o idioma-alvo já se tornou mais robusta, espera-se que o participante consiga interpretar com maior segurança as gradiências entre os gestos. Dessa forma, propomos que UMA TAREFA DE DISCRIMINAÇÃO TENHA IMPLICAÇÕES TANTO PARA A DISTINTIVIDADE COMO PARA A NÃO DISTINTIVIDADE DOS GESTOS, a depender da experiência do participante com a LNN. Nesse quesito, fazemos uma ressalva: não estamos afirmando que, aprioristicamente, uma tarefa de discriminação esteja em íntima relação com fatores apenas gradientes (fonéticos), conforme consta em trabalhos da literatura – defendemos que a tarefa de discriminação tem IMPLICAÇÕES para o que se concebe tradicionalmente como fonético e fonológico, e pode servir para ambos os aspectos, para nós, aspectos gestuais.

Nesse sentido, uma tarefa de discriminação não se comprometeria em informar se a habilidade discriminatória de contrastes fônicos tem a ver com propriedades de baixa ordem ou de alta ordem (*lower order* ou *higher order*, segundo a distinção do PAM e do PAM-L2), ou seja, não é pertinente determinar se levam em consideração características que restariam no nível da produção da fala ou no nível das oposições funcionais. Em nosso ponto de vista, a polarização que Best (1995) apresenta em termos de baixa ou alta ordem fragiliza a adoção a uma tarefa de discriminação justamente por não se saber se o que está em jogo para os ouvintes são diferenciações que estariam para a fonética ou para a fonologia. Logo, um primitivo de análise que conte com a diluição dessa fronteira tradicional se mostra plenamente inserível em uma ferramenta empírica como uma tarefa de discriminação, já que seus resultados respondem pela organização da gramática fônica como um todo, sem pender mais para um lado ou para outro.

Ao considerar os trabalhos retratados no terceiro capítulo desta tese, quando, em um mesmo estímulo, apresentaram-se dois ou mais elementos fônicos a fim de comparação de determinado contraste funcional, pudemos também observar uma variada gama de designações para tal tarefa. Os nomes encontrados para as tarefas que compreenderam tal procedimento foram DISCRIMINAÇÃO, conforme Best e Hallé (2010), Bundgaard-Nielsen, Best e Tyler (2011b), Strange *ET AL.* (2011), Albuquerque (2012), Antoniou, Best e Tyler (2012), Bohn e Best (2012), Antoniou, Best e Tyler (2013), Patihis, Oh e Mogilner (2013), Elvin, Escudero e Vasiliev (2014), McKelvie-Sebileau e Davis (2014), Tyler *ET AL.* (2014), Alves e Zimmer (2015) e Holliday (2016), DISCRIMINAÇÃO CATEGÓRICA, segundo Fabra e

Romero (2012), Perozzo (2013) e Escudero, Sisinni e Grimaldi (2014), CATEGORIZAÇÃO¹²⁶, conforme Escudero e Williams (2012), CATEGORIZAÇÃO DO ELEMENTO VOCÁLICO DIFERENTE¹²⁷, segundo Darcy e Krüger (2012), e DISCRIMINAÇÃO DO ELEMENTO VOCÁLICO DIFERENTE¹²⁸, por Escudero, Sisinni e Grimaldi (2014).

Nesse perfil de tarefa, foram três os formatos encontrados no que tange às sequências dos elementos fônicos, a saber, AXB , XAB e AX . No formato AXB (BEST; HALLÉ, 2010; BUNDGAARD-NIELSEN; BEST; TYLER, 2011b; STRANGE *ET AL.* 2011, ANTONIOU; TYLER; BEST, 2012; BOHN; BEST, 2012; FABRA; ROMERO, 2012; ANTONIOU; BEST; TYLER, 2013; PATIHIS; OH; MOGILNER, 2013; TYLER *ET AL.* 2014; ALVES; ZIMMER, 2015), A e B são elementos fônicos diferentes, e X corresponde a um terceiro elemento, o qual pode ser igual a A ou a B . Nesse padrão, solicita-se ao participante que ele informe se o item em segunda posição, X , corresponde ao item A , que o precede, ou se corresponde ao item B , que o segue. O formato XAB (ESCUADERO; WILLIAMS, 2012; EVLIN; ESCUDERO; VASILIEV, 2014) é bastante similar em procedimento ao AXB , mas o elemento-alvo (X) é o primeiro a ser apresentado na tríade, e não o segundo. No formato AX (MCKELVIE-SEBILLEAU; DAVIS, 2014), entretanto, em vez de ouvir três elementos fônicos em um mesmo estímulo (como ocorre no AXB e no XAB), o participante ouve apenas dois, e deve julgar se o segundo (X) é igual ao primeiro (A) – ou seja, deve informar se os dois elementos são iguais ou diferentes.

Especialmente no que tange ao formato AX , Yang e Fox (2014) configuraram sua tarefa de percepção, que, por sua vez, correspondeu a uma tarefa de discriminação, de modo que operasse em conjunto com uma ESCALA DE SIMILARIDADE. Essa escala, contendo 9 pontos que variavam de “muito similar” a “muito diferente”, permitiu que os participantes avaliassem o quão similares eram os elementos fônicos ouvidos em um mesmo estímulo. Os pesquisadores afirmam que os participantes monolíngues tenderam a acomodar os itens fônicos da LNN em grandes grupos de categorias nativas, enquanto que os participantes bilíngues reconheceram com mais precisão as fronteiras entre os elementos fônicos da LNN.

Desdobramentos dos formatos mencionados até o momento são encontrados nos estudos de Darcy e Krüger (2012) e de Escudero, Sisinni e Grimaldi (2014), que contaram com as tarefas ODDITY

¹²⁶ Embora os autores designem o nome de “*TWO-ALTERNATIVE FORCED CHOICE CATEGORIZATION TASK IN AN XAB FORMAT*” [tarefa de categorização com duas alternativas em formato XAB (ESCUADERO; WILLIAMS, 2012, p. 409)] à tarefa utilizada, trata-se de uma tarefa de discriminação. Dessa forma, consideramos que o termo “CATEGORIZAÇÃO” se apresenta de maneira inapropriada na pesquisa em questão.

¹²⁷ O nome da tarefa em inglês é *ODDITY VOWEL CATEGORIZATION TASK*. Nesse caso, julgamos ocorrer a mesma inadequação terminológica observada em Escudero e Williams (2012).

¹²⁸ O nome da tarefa em inglês é *ODDITY VOWEL DISCRIMINATION TASK*.

VOWEL CATEGORIZATION TASK (porém, sem referência à L₁ dos participantes) e ODDITY VOWEL DISCRIMINATION TASK, respectivamente. Tanto em uma como em outra tarefa, são apresentados aos participantes três elementos fônicos em um mesmo estímulo e, ao responder, eles devem apontar qual é o elemento destoante entre os três itens – o qual pode ocupar a primeira, a segunda ou a terceira posição no estímulo auditivo. Conforme mencionamos previamente, julgamos que a utilização do termo “CATEGORIZAÇÃO”, na pesquisa de Darcy e Krüger (2012), seja imprecisa, uma vez que os participantes não rotularam qual elemento era destoante (informaram somente que um dos elementos era diferente, mas não o caracterizaram), o que poderia vir a justificar o uso do referido termo; além disso, a tarefa não contou com apenas um elemento fônico veiculado em cada estímulo, mas, sim, com três, o que implica, necessariamente, uma tarefa de DISCRIMINAÇÃO.

Após termos discorrido sobre as tarefas de assimilação, identificação e discriminação em suas respectivas subseções (5.2.1, 5.2.2 e 5.2.3), vislumbramos a necessidade de retomar, de modo resumido, as características de cada uma das tarefas, endereçando, particularmente, para fins de taxionomia, o delineamento empírico, o tipo de conhecimento de base a ser testado e o objetivo de cada instrumento perceptual. A propósito dessas informações, apresentamos o Quadro 3 a seguir:

Quadro 3 – Resumo das tarefas perceptuais e suas especificidades (continua)

Taxionomia	Delineamento empírico	Conhecimento de base testado	Objetivo
Assimilação	Rotulação de um item fônico não nativo em um único estímulo auditivo com base em opções de múltipla escolha, mais escala de adequabilidade.	L ₁	Compreender os processos assimilatórios instaurados na via L _{NN-L₁} , a propósito de semelhanças ou distinções entre as categorias fônicas nativas e não nativas.
Identificação	Rotulação de um item fônico não nativo em um único estímulo auditivo com base em opções de múltipla escolha.	L _{NN}	Verificar as categorias fônicas formadas pelos aprendizes ao longo de sua experiência com o idioma-alvo, resgatando sua gramática fônica não nativa.

Quadro 3 – Resumo das tarefas perceptuais e suas especificidades (continuação)

Taxionomia	Delineamento empírico	Conhecimento de base testado	Objetivo
Discriminação	Comparação de dois ou mais itens fônicos não nativos em um único estímulo auditivo com base em opções de múltipla escolha.	LNN	Mensurar a variabilidade dos gestos fônicos que permeiam as categorias não nativas, de modo a estabelecer os limites perceptuais (funcionais ou não) entre tais categorias.

Fonte: elaborado pelo autor.

Conforme podemos observar a partir do Quadro 3, cada uma das tarefas é dotada de especificidades e respondem por finalidades distintas. Faz-se válido mencionar, no entanto, que, embora a assimilação perceptual seja aferida diretamente através da tarefa homônima, ela também pode ser indiretamente inferida por meio das tarefas de identificação e de discriminação, quando os participantes exibirem baixos níveis de acuidade, por exemplo. Assim, inferências interpretativas acerca de assimilações podem ser indiretamente vislumbradas em tarefas de identificação e de discriminação, mas reservamos a tais instrumentos a mensuração das dissimilações, que evidenciam as categorias a serviço do sistema não nativo.

Na próxima subseção, trataremos de algumas ponderações adicionais no que se refere aos estudos apresentados neste capítulo e no capítulo 3. Tais considerações poderiam ter sido acopladas às discussões já desenvolvidas nesta e nas duas subseções anteriores, mas julgamos mais pertinente que elas sejam endereçadas em particular porque tocam em questões cruciais à maneira como os gestos fônicos são concebidos para fins empíricos.

5.2.4 Ponderações teórico-metodológicas em relação a estudos retratados no capítulo

Uma tarefa de discriminação, como pudemos observar, pode fomentar discussões teóricas cujos argumentos não são sempre consensuais. Incitamos, pois, uma breve digressão cabível neste momento, dada a cisão entre relações de baixa e alta ordem estabelecidas em Best (1995).

Um caso clássico de alofonia em português brasileiro é representado pelo par [t] e [tʃ]. Em termos “fonéticos”, de baixa ordem, [t] e [tʃ] correspondem a categorias distintas e de fácil

discriminação. Contudo, em termos “fonológicos”, de alta ordem, tais elementos fazem parte de uma única categoria, a de /t/. Consequentemente, essa exemplificação retoma a nossa indagação sobre uma tarefa de discriminação (considerando-se fonética e fonologia como expedientes distintos) refletir uma habilidade discriminatória baseada em pressupostos fonéticos ou na manutenção da contrastividade. Assim sendo, como saberíamos, hipoteticamente, que [t] e [tʃ] pertencem a uma mesma categoria (a de /t/), mas [p], diferentemente, pertence a outra categoria (a de /p/)?

Uma resposta intuitiva, mas circular, seria afirmar que não temos registrados casos de alofonias envolvendo /p/, por isso esses elementos fariam parte de categorias fonológicas distintas. No entanto, uma resposta mais convincente, que encontra sustentação a partir dos postulados de Repp (1984), seria a de que existe maior variabilidade articulatória no contínuo [t]-[tʃ] do que na relação [t]-[p]. Em outras palavras, as mudanças nas configurações gestuais de [t] e [tʃ] seriam bem menos abruptas do que aquelas de [t] e [p].

Essa hipótese pode ser evidenciada articulatoriamente se observarmos os articuladores que estão ativos a propósito dessas respectivas constrictões: no caso de [t] e [tʃ], o articulador ativo envolve a região coronal da língua (e os articuladores passivos seriam a região alveolar e a região alveopalatal), enquanto que no caso de [t] e [p], os articuladores ativos estão para domínios completamente diferentes; região coronal da língua vs. lábios. Dessa forma, mesmo considerando que a percepção do contraste entre gestos consonantais tenda a ser categórica e a percepção do contraste entre gestos vocálicos tenda a ser contínua, ainda podemos verificar gradiência em gestos consonantais cujas fronteiras são tidas como categóricas.

Essa continuidade em classes assumidas como mais rígidas provavelmente advenha do compartilhamento de propriedades gestuais entre regiões acústico-articulatórias adjacentes. Logo, a tarefa de determinar as fronteiras entre o que é “fonético” e o que é “fonológico” depende, em larga escala, de como o pesquisador vislumbra as relações entre os gestos fônicos.

Em nossa concepção teórica, a empreendida por Albano (2001), não assumimos uma fronteira entre aspectos fonéticos e aspectos fonológicos, pois acreditamos que as representações fonológicas incluem a gradiência fonética em sua totalidade, o que nos leva entender que as relações sonoras nas línguas se instanciam em uma gramática fônica, cuja unidade é gestual e de caráter acústico-articulatório. Transpondo essa noção para um contexto empírico de testagem, a exemplo de uma tarefa de discriminação, é preciso que nosso referencial teórico de base gestual esteja em consonância com o nosso objetivo metodológico de aferir as habilidades discriminatórias dos

participantes. É por isso que, na subseção 5.2.3, propusemos que o julgamento que o participante faz acerca dos itens fônicos ouvidos no momento em que os compara está associado à variabilidade dos gestos fônicos, em que a percepção das propriedades (ou variáveis) limítrofes de um gesto é dependente da experiência do participante com a L_{NN}. Uma potencialidade da tarefa de discriminação, nesse sentido, é a de apresentar variações gestuais (que podem ou não ser funcionais) que respondam por fronteiras perceptuais sinalizadas na medida em que a gramática fônica não nativa se desenvolve.

Mudemos, agora, o foco de nossa problematização. Se no expediente teórico existe toda uma complexidade que envolve o modo como nos remetemos às unidades fônicas em termos empíricos, no expediente laboratorial, o cenário conta com discussões tão ou mais incessantes. Chamamos atenção para uma questão polêmica que se faz presente no estudo de Escudero e Chládková (2010).

Ao desempenhar a tarefa perceptual do experimento conduzido, que envolvia a identificação de vogais inglesas pertencentes aos dialetos americano e britânico, os participantes (falantes nativos de espanhol peruano) foram advertidos de que estariam ouvindo vogais espanholas, extraídas de gravações de fala rápida, que poderiam, eventualmente, soar como instâncias deficitárias dessas vogais. O mesmo ocorreu em Escudero e Williams (2011): falantes nativos de espanhol peruano foram advertidos de que ouviriam vogais espanholas articuladas em fala rápida, que poderiam soar deficientes, mas, na verdade, estavam ouvindo estímulos pertencentes ao holandês. O que podemos observar é que há uma distorção das informações, resultando em uma inverdade sobre a natureza dos estímulos. Segundo Colantoni, Steele e Escudero (2015), a experiência com a língua-alvo não afeta o mapeamento L_{NN-L₁} quando os participantes acreditam que os estímulos advêm de seu próprio sistema materno.

Em nossa visão, esse tipo de procedimento pode incidir em dois problemas, um de ordem ética e outro de ordem teórico-metodológica. No âmbito ético, os pesquisadores informam os participantes acerca de uma condição que não corresponde verdadeiramente aos objetivos do estudo, já que sinalizam que os elementos fônicos a serem ouvidos fazem parte do inventário materno, porém em situações específicas (uma suposta fala rápida). No âmbito teórico-metodológico (e este é um ponto que, de certo modo, impõe questionamentos a tarefas de assimilação como um todo), por sua vez, induzir os participantes a responder ao estímulo apresentado com base no inventário fônico materno acaba restringindo uma opção que poderia ser útil ao pesquisador: a de o item fônico ouvido não corresponder a um som da fala [possibilidade levantada em Best (1995)].

Teoricamente, a utilidade dessa informação reside no fato de tornar o pesquisador ciente de que tal elemento fônico, eleito como não pertencente a um som da fala, soa como não linguístico, e isso tem consequências em termos de como a gramática fônica não nativa do aprendiz está organizada no momento em que se depara com itens muito distintos (gestual e tipologicamente) daqueles da L1. Metodologicamente, tarefas que apresentassem a opção “o elemento ouvido não corresponde a um som da fala”, poderiam ser configuradas de modo a englobar, em princípio, perturbações acústicas de qualquer natureza, estendendo-se a todas as possibilidades de sons, linguísticos ou não linguísticos.

Tratadas essas duas questões mais específicas, daremos seguimento ao capítulo. Esta penúltima seção irá abordar aspectos referentes à implementação da pesquisa no âmbito da percepção fônica não nativa, levando em consideração discussões realizadas para os propósitos da presente tese.

5.3 Implementação das tarefas perceptuais fônicas não nativas à luz das considerações desta tese

Aceitamos e acreditamos na proposição de que a investigação empírica deve estar a serviço de pressupostos teóricos. Desse modo, é nosso dever destacar, nesta seção, como podemos dar conta de proceder metodologicamente de modo a satisfazer os aspectos teóricos que baseiam a percepção fônica de línguas não nativas, defendidos nesta tese. De antemão, advertimos que não estaremos fornecendo uma “receita” acerca de como deve ocorrer o trabalho laboratorial do pesquisador no expediente em que se situa a nossa proposta, mas apontaremos algumas questões metodológicas que fazem sentido frente à implementação de tarefas perceptuais no quesito fônico. Especificamente, elencaremos algumas premissas metodológicas alinhadas à nossa concepção teórica, discorreremos sobre a operacionalização da percepção fônica não nativa em termos metodológicos e, então, caracterizaremos as tarefas laboratoriais indispensáveis à nossa proposta de percepção fônica não nativa.

5.3.1 Premissas metodológicas essenciais quanto à percepção fônica não nativa com base em nossa proposta

Nossa caracterização da percepção fônica de línguas não nativas tem como base três esferas diferentes que se relacionam e se complementam: um aporte cognitivista para o evento perceptual,

um primitivo acústico-articulatório para a informação gestual das línguas e uma postura realista indireta quanto às experiências de percepção. Assim, nossas ponderações serão voltadas a atender essas três premissas, ao considerar que nossa proposta pode ser aplicada tanto a PADRÕES DE ASSIMILAÇÃO ENTRE CATEGORIAS FÔNICAS NA RELAÇÃO L_{NN}-L₁ COMO AO DESENVOLVIMENTO DE CATEGORIAS FÔNICAS NÃO NATIVAS.

Defendemos a visão de que a experiência perceptual corresponde a um mecanismo COGNITIVO por excelência e pode ser medida de diversas maneiras. Por se tratar de um evento cognitivo que coaduna mecanismos de racionalização e informação (rica e multimodal) proveniente do ambiente, acreditamos que testes psicológicos de percepção, contemplando diferentes tarefas, sejam ferramentas cabíveis à sua mensuração. Para fins de compreendermos como ocorrem as assimilações perceptuais, no sentido de compreendermos quais são as categorias fônicas da L₁ a acomodarem aquelas da L_{NN}, sugerimos a utilização de tarefas de assimilação perceptual subsidiadas por uma escala de adequabilidade, conforme explanado na segunda seção deste capítulo.

Com o intuito de verificar as categorias fônicas formadas pelos aprendizes ao longo de sua experiência com o idioma-alvo, o que implica resgatar sua gramática fônica não nativa, recomendamos tarefas de identificação. Além disso, de modo que mensuremos a variabilidade dos gestos fônicos que permeiam as categorias não nativas, e que repercutem no estabelecimento dos limites perceptuais entre tais categorias, sejam tais gestos contrastivos ou não, vemos a possibilidade da aplicação de uma tarefa de discriminação.

Na esfera FÔNICA, a caracterização que propomos se sustenta de modo a conceber uma unidade gestual de cunho acústico-articulatório. Assim, por uma questão de coerência epistemológica, não esperamos que parâmetros de ordem estritamente acústica ou de ordem puramente abstrata operem no modelo. A especificação de um primitivo gestual de caráter acústico-articulatório certamente restringe o leque de fenômenos fônicos descritos na literatura capazes de serem vislumbrados segundo essa ótica, mas, por outro lado, confere consistência interna e externa ao modelo. Frisamos que não é o caso de análises envolvendo parâmetros puramente acústicos ou puramente fonológicos não poderem ser endereçadas segundo a nossa proposta, apenas deixamos claro que tais unidades podem se distanciar das assunções que dão forma ao modelo que desenvolvemos.

Em nossa visão, frente a contrastes funcionais, o plano fônico é tocado de maneira distinta por percebedores sem experiência e com experiência em relação à L_{NN}. No primeiro caso, esperamos que, dado o sistema fônico da L₁, contrastes funcionais não nativos possam ser interpretados como variações no mérito da produção da fala ou como alofonias. Para passar desse cenário em que a

distintividade é pouca ou nenhuma para um cenário em que a forma fônica se coaduna com a significação, é necessário que o percebedor experiencie o sistema fônico da LNN, que se insira em contextos em que possa testar, gradativamente, os limites “daquilo que pode e daquilo que não pode” em termos de distinções lexicais. São as suas tentativas e as suas produções, de fato, que lhe darão possibilidade de conceber os limites das bordas gestuais que terão implicações para as fronteiras perceptuais. Deve ficar evidente, portanto, que cognição e gramática fônica operam em consonância para que novos contrastes na LNN sejam observados.

Levando-se em conta nossos pressupostos filosóficos (com implicações psicolinguísticas), no que concerne ao REALISMO INDIRETO, o cenário metodológico para a testagem da percepção fônica parece bastante claro e positivo. Isso se dá tanto em termos dos estímulos que podemos utilizar nas tarefas de percepção, como em termos da viabilidade de se testar, em ambiente laboratorial, a experiência perceptual fônica. No primeiro caso, quer utilizemos estímulos naturais, quer utilizemos estímulos sintéticos, essa decisão não impacta de forma negativa o modelo aqui proposto porque entendemos que os dados/estímulos não “falam por si”, mas são racionalizados e passam pela comparação com nossas experiências prévias, são filtrados pelas nossas sensações e pelas nossas impressões – o dado, embora seja rico e multimodal, não é puro e não pretende revelar uma única verdade, absoluta e inegável. No segundo caso, por abandonarmos a noção de ambiente ecológico, perfeito às condições de um sujeito percebedor que se vale puramente das concessões oferecidas, e também por seguirmos um referencial realista indireto, julgamos que as condições laboratoriais podem ser tão satisfatórias quanto aquelas não laboratoriais, e o fato de as tarefas serem administradas via algum *SOFTWARE* ou maquinaria computacional, além de fones de ouvido, se encontra em consonância com o que se espera de uma percepção indireta (especialmente, se considerarmos a primeira acepção do termo indireto segundo as considerações gibsonianas).

5.3.2 Operacionalizando aspectos particulares da percepção fônica não nativa

Esta subseção foca em como podemos operacionalizar aspectos da percepção fônica não nativa e faz referência, especificamente, a duas situações importantes. Primeiramente, preocupamo-nos, aqui, em discorrer sobre como estabelecer o ponto de corte entre o que foi percebido categoricamente ou não, no sentido de vislumbrar até onde se pode afirmar que determinado item fônico não nativo foi assimilado a uma categoria nativa. Em segundo lugar, trazemos um exemplo

de estudo que contou com três tarefas perceptuais, de modo a evidenciar a possibilidade concreta de uma pesquisa abarcar tarefas de assimilação perceptual, identificação e discriminação.

Uma dúvida que pode surgir no desenvolvimento de uma pesquisa no âmbito da percepção, essencialmente quando consideramos tarefas que proponham verificar as nuances do processo de assimilação perceptual, é relativa ao fato de, muitas vezes, desconhecermos quais seriam os pontos a partir dos quais pudéssemos afirmar que houve ou não a alocação de determinado item fônico a uma dada categoria. Desse modo, no que tange a uma possível maneira de se poder evidenciar os pontos de corte entre aquilo que foi ou não percebido categoricamente, Repp (1984) menciona que um parâmetro virtuoso seria o ESTATÍSTICO.

Nesse quesito, Bundgaard-Nielsen, Best e Tyler (2011b) prestam uma importante contribuição analítica: ao conduzir sua análise, os pesquisadores estabelecem que a assimilação perceptual deva ocorrer a partir de dois intervalos percentuais; primeiramente, em 50% e, em seguida, em 70% das categorizações. Se determinado item fônico de uma L₂ é julgado como relacionável a certo item fônico da L₁ do participante em 50% e em 70% dos casos em que aquele estímulo aparece, então o elemento não nativo é percebido como uma instância apropriada de sua contraparte na L₁. Nas palavras dos autores, “uma vogal da L₂ foi definida como categorizada se houver sido assimilada como uma instância de uma categoria específica da L₁ em mais de 50% das ocorrências” (BUNDGAARD-NIELSEN; BEST; TYLER, 2011b). Na mesma pesquisa, os estudiosos afirmam que elevaram o ponto de corte acerca dos índices de assimilação para 70% em função de que tal prática era adotada também em outros trabalhos e poderiam gerar resultados mais tangíveis. Para todos os efeitos, a recomendação dos autores é a de que o ponto de corte para as categorizações deva ser essencialmente de 70%.

São três as razões pelas quais Bundgaard-Nielsen, Best e Tyler (2011b) justificam o baixo ponto de corte das assimilações em 50% das ocorrências. A primeira diz respeito à constatação de que, por se tratar de um experimento envolvendo vogais, há de se pressupor que os elementos fônicos sejam muito mais contínuos entre si, o que tornaria a classificação em categorias mais permissiva dada a natureza dos itens em questão. A segunda se sustenta sobre o fato de que mesmo os falantes nativos do idioma testado podem encontrar dificuldades na categorização apurada dos elementos fônicos. A terceira razão se relaciona à complexidade da tarefa aplicada aos aprendizes do estudo, que envolvia dois tipos de registros gráficos (o alfabético/fonético, da L₂, e o silabário katakana, da L₁). Para os efeitos da proposta que defendemos, contudo, presumimos que o ponto mínimo de corte

em relação às assimilações seja de 70%, mas esse valor pode ser alterado a critério do pesquisador, caso o julgue necessário.

Em se tratando do segundo aspecto a ser levantado nesta subseção, chamamos atenção para um estudo primoroso no que compete à percepção fônica, resenhado no terceiro capítulo desta tese, conduzido por Holliday (2016). Bastante satisfatória do ponto de vista de seus objetivos e mecanismos de análise, a pesquisa de Holliday (2016) propõe analisar a identificação e a discriminação de elementos fônicos (ambas em relação à L_{NN} dos participantes), bem como avaliar a assimilação de tais elementos (em relação à L₁ dos participantes) – o que julgamos configurar uma abordagem bastante completa em termos da testagem da percepção fônica.

Basicamente, em sua tarefa de assimilação perceptual, solicitou-se aos participantes que eles ouvissem certo estímulo da L_{NN} e, após escutá-lo e rotulá-lo com base em seu inventário materno, marcassem em uma escala de adequabilidade o quão similar foi o estímulo ouvido ao de seu correspondente na L₁. As demais tarefas, de identificação e de discriminação, levaram em conta a percepção dos elementos fônicos com base na L_{NN} apenas. Salientamos, pois, a pertinência do método amplo de Holliday (2016) nas testagens acerca da percepção fônica, contando com tarefas de identificação e discriminação com base na L_{NN}, mais tarefa de assimilação com base na L₁. A metodologia empregada pelo referido autor se encontra em consonância com o que, de fato, compreendemos como sendo um fator fundamental da assimilação: A APROXIMAÇÃO OU O DISTANCIAMENTO ENTRE UNIDADES FÔNICAS NATIVAS E NÃO NATIVAS, o que responde pela incorporação de um som qualquer da fala a determinada categoria, segundo Best (1995). A tarefa de identificação (a partir das categorias da L_{NN}) se faz, sim, essencial, mas a tarefa de assimilação (a partir das categorias da L₁, contendo diferentes opções de escolha e, também, uma escala de Likert), além de revelar qual elemento fônico o sujeito seleciona como aquele que corresponde ao de sua L₁, indica a adequabilidade do elemento em si, o qual pode funcionar como um “bom” ou “mau” (e suas diversas gradações) representante da categoria fônica em jogo.

As considerações metodológicas apontadas nesta subseção têm a finalidade de sugerir recursos aos pesquisadores no que se refere a um maior controle no gerenciamento de seus estudos, incidindo tanto no mérito de um critério analítico para as assimilações, como no que compete a um exemplo promissor de pesquisa, o qual envolve tarefas com objetivos bastante específicos e procedimentos que se complementam em prol de um resultado confiável. Sublinhamos, assim, a pertinência do trabalho de Holliday (2016) para a condução de estudos que tratem da percepção

fônica não nativa. Para uma compreensão mais ampla desse processo, outro instrumento empírico profícuo corresponde a um questionário, o qual pode se voltar para aspectos linguísticos ou procedimentais da tarefa, conforme será abordado na próxima seção.

5.3.3 Tarefas laboratoriais indispensáveis à nossa proposta de percepção fônica não nativa

Em uma perspectiva que objetive contemplar tanto a ASSIMILAÇÃO PERCEPTUAL COMO a FORMAÇÃO DE CATEGORIAS na LNN, defendemos a inclusão de quatro tarefas, as quais podem ser administradas aos participantes a depender de alguns critérios. Os referidos instrumentos se referem a tarefas de assimilação perceptual, identificação e discriminação, bem como a um questionário sobre as tarefas citadas.

Ao que podemos observar a partir dos postulados de Colantoni, Steele e Escudero (2015), tarefas envolvendo a rotulação de determinado elemento fônico são chamadas de TAREFAS DE IDENTIFICAÇÃO, enquanto que as tarefas envolvendo elementos fônicos intervalados em um contínuo (como em casos de comparação de elementos em um mesmo estímulo) são chamadas de TAREFAS DE DISCRIMINAÇÃO. Argumentamos que essas duas tarefas devam se voltar para a percepção fônica não nativa com base no sistema da própria LNN, segundo mencionamos nas subseções 5.2.2 e 5.2.3. Além disso, tarefas envolvendo a rotulação de determinado elemento fônico seguidas de uma escala de adequabilidade são chamadas de TAREFAS DE ASSIMILAÇÃO. Paralelamente, defendemos que essa tarefa deva se voltar para a percepção fônica não nativa com base no sistema materno, conforme apontamos na subseção 5.2.1. Como deve ter ficado nítido a partir de considerações prévias, sugerimos a manutenção dessa taxionomia para a pesquisa em percepção fônica não nativa, pois contempla satisfatoriamente as especificidades de cada tarefa.

Dois questionamentos legítimos seriam referentes a quando utilizar cada uma das tarefas e para qual público-alvo. Responderemos em conjunto a ambos os questionamentos. Deve restar claro, sobretudo, que cada tarefa toca em instâncias específicas e contempla aspectos distintos da percepção fônica. Uma TAREFA DE ASSIMILAÇÃO PERCEPTUAL pretende responder pelas associações ou mapeamentos que o participante realiza ao conectar determinado item fônico não nativo ao seu conhecimento fônico materno. Necessariamente, sistema materno e sistema-alvo estão em jogo nesse tipo de tarefa, e cada um possui sua parcela de contribuição. Segundo defendemos, uma tarefa de assimilação perceptual deve sempre ser seguida por uma tarefa de adequabilidade que empregue uma escala de Likert. Desse

modo, essa tarefa deve ser administrada somente se estivermos interessados na simbiose ente L_1 e L_{NN} , o que implica mapear as correspondências entre categorias fônicas de ambos os inventários, materno e não nativo, tais quais evidenciam Best e Tyler (2007).

Esse mapeamento é útil tanto para INDIVÍDUOS QUE TÊM POUCA OU NENHUMA EXPERIÊNCIA com o idioma-alvo, como para aprendizes do idioma-alvo que possuem DIFERENTES NÍVEIS DE PROFICIÊNCIA. No primeiro caso, a interpretação dos resultados da tarefa pode demonstrar quais são as categorias maternas que um indivíduo utiliza ao ouvir/reconhecer itens fônicos não nativos (achado que pode ser utilizado para fins pedagógicos, por exemplo, no sentido de o professor fornecer instrução explícita sobre o contraste em questão). No segundo caso, a interpretação dos resultados da tarefa pode permitir o estabelecimento de comparações entre aprendizes de nível básico, intermediário e avançado em termos de quais são as categorias nativas que tais aprendizes utilizam para acomodar os itens não nativos: uma dada categoria da L_1 pode estar sendo utilizada por aprendizes básicos para assimilar determinado item fônico não nativo, ao passo que outra categoria da L_1 esteja sendo usada por aprendizes avançados para assimilar o mesmo item fônico não nativo. Esses possíveis resultados têm implicações para que possamos comparar indivíduos com e sem conhecimento do idioma-alvo.

Uma TAREFA DE IDENTIFICAÇÃO tem o intuito de verificar a percepção de um item fônico não nativo com base no conhecimento da gramática fônica não nativa do participante. O que está em jogo nessa tarefa é a constatação de categorias fônicas formadas pelos aprendizes ao longo de sua experiência com o idioma-alvo. O conhecimento acerca das categorias maternas utilizadas pelos participantes na assimilação de um item fônico não nativo não é objeto de análise aqui, uma vez que, com essa tarefa, observamos o reflexo dos estágios desenvolvimentais dos aprendizes¹²⁹. Nesse sentido, a tarefa deve ser proposta a INDIVÍDUOS QUE TENHAM EXPERIÊNCIA (em diferentes níveis) com o idioma-alvo. Por não conter uma escala de adequabilidade, a resposta que o participante fornece é mais contundente, prestando-se a indicar um domínio do idioma-alvo que o participante monolíngue não possui.

No que se refere à TAREFA DE DISCRIMINAÇÃO, argumentamos que ela seja particularmente interessante porque, em vez de apresentar apenas um item fônico em cada estímulo auditivo, exhibe uma sequência de dois ou mais itens fônicos no mesmo estímulo auditivo. Isso ocorre para que o participante possa comparar os itens fônicos e responder, por exemplo, se o item intermediário se assemelha ao primeiro ou ao último. Com base nas respostas dos participantes, temos subsídios para

¹²⁹ Todavia, frisamos que inferências sobre assimilações podem ser realizadas com base na análise de uma tarefa de identificação e de discriminação, paralelamente ao que expusemos na subseção 5.2.3.

examinar a variabilidade dos gestos fônicos que permeiam as categorias não nativas, o que nos facultaria apontar, com algum nível de precisão, onde residem os limites perceptuais (funcionais ou não) entre tais categorias. Como se trata de uma tarefa que prevê conhecimento do sistema fônico não nativo, esse instrumento se direciona a APRENDIZES da língua-alvo (assim como sugerido para a tarefa de identificação). Faz-se pertinente que grupos pertencentes a diferentes níveis de proficiência respondam a essa tarefa, no sentido de que há a possibilidade de aprendizes mais avançados reconhecerem com maior facilidade (determinada no que se refere a índices de acuidade) os contrastes envolvidos no estímulo se comparados a aprendizes intermediários ou básicos do idioma-alvo. Isso revela, também, os estágios desenvolvimentais dos aprendizes em termos do sistema fônico não nativo. No plano teórico, segundo o que mencionamos na seção anterior, os contrastes funcionais poderão ser reconhecidos por aprendizes de nível mais avançado de proficiência, enquanto que aprendizes de nível mais básico poderão não os reconhecer, e o momento aproximado em que se passa a reconhecer a distintividade poderá ser constatado através da tarefa de discriminação.

Como pudemos observar, tarefas de assimilação perceptual são adequadas a participantes que possuem ou não conhecimento do idioma-alvo, enquanto que tarefas de identificação e discriminação devem ser direcionadas apenas a aprendizes do idioma-alvo, que já possuem contato com o sistema fônico não nativo. No que compete à ordem de aplicação das tarefas, considerando-se um público-alvo que irá desempenhar as três, a literatura se mostra divergente, pois, em alguns estudos, tarefas de assimilação perceptual e/ou identificação são administradas previamente à tarefa de discriminação (BUNDGAARD-NIELSEN; BEST; TYLER, 2011b; ESCUDERO; SISINNI; GRIMALDI, 2014), enquanto que, em outros, a tarefa de discriminação é a primeira a ser empregada (BEST; HALLÉ, 2010; STRANGE; *ET AL.* 2011; ANTONIOU; BEST; TYLER, 2012; ESCUDERO; WILLIAMS, 2012; ANTONIOU; BEST; TYLER, 2013; TYLER; *ET AL.*, 2014). No entanto, Best e Hallé (2010) sugerem que a tarefa de discriminação seja a primeira a ser aplicada, de maneira a minimizar possíveis efeitos de categorização dos estímulos. Para todos os efeitos, seguimos a recomendação de Best e Hallé (2010) quanto à ordem de administração das tarefas e sugerimos que cada uma seja aplicada em um dia diferente, para que possíveis efeitos instrucionais, a partir da tarefa perceptual, também sejam minimizados. Destacamos que, com a inclusão de uma tarefa de assimilação à testagem perceptual, uma ordem coerente de aplicação das tarefas seria, primeiramente, a de assimilação, seguida pela de discriminação, e finalizada com a de identificação.

Uma questão relevante que se apresenta em termos das testagens perceptuais se relaciona ao TIPO DE ESTÍMULO veiculado e ao IDIOMA que empregado nas tarefas e em possíveis instruções orais. Encontramos, nos estudos perceptuais, relatos da utilização de estímulos puramente auditivos, assim como estímulos em vídeo, que abarcam tanto características articulatórias como auditivas do item-alvo. Independentemente de qual tipo for utilizado, defendemos que os estímulos apenas auditivos podem ser bastante informativos, uma vez que parâmetros articulatórios podem ser recuperados via propriedades acústicas, conforme estabelece Albano (2001). Talvez um dos fatores que restrinja a possibilidade de se empregarem estímulos em vídeo (em detrimento de estímulos auditivos unicamente) resida no âmbito metodológico, pois nem todas as plataformas disponíveis atualmente suportam arquivos de vídeo. Nesse caso, chamamos atenção para as ponderações de Albano (2001), que argumenta que muitas propriedades acústicas podem indicar correlatos articulatórios.

Em relação ao idioma empregado na testagem, se estivermos em paralelismo com o tipo de conhecimento de base testado em cada tarefa, podemos recomendar que, no caso das tarefas de assimilação perceptual, passemos a conduzir o experimento em nossa L₁, já que o foco da tarefa de assimilação perceptual é a categorização fônica não nativa com base em categorias maternas. Igualmente, no caso das tarefas de identificação e discriminação, sugerimos que a língua de instrução e condução da tarefa seja a língua-alvo. Esse tipo de postura é atestado na pesquisa de Antoniou, Tyler e Best (2012) e Antoniou, Best e Tyler (2013), em que o idioma da tarefa ou dos participantes é consistente com o idioma falado pelo pesquisador no momento da testagem.

Talvez um dos maiores impasses no que tange aos testes perceptuais seja a FORMA DE APRESENTAÇÃO das opções de múltipla escolha em cada tarefa. Alinhada com o nosso referencial teórico seria a apresentação de PAUTAS GESTUAIS nas opções de resposta, uma vez que estaríamos sendo fiéis ao primitivo que adotamos e que julgamos ser o mais apropriado para tratar dos eventos sonoros nas línguas. No entanto, as desvantagens de tal medida possivelmente superariam suas vantagens. Se a elaboração de pautas gestuais configura um árduo trabalho de análise acústico-articulatória, a interpretação ou a leitura de tais pautas por parte dos sujeitos das pesquisas é completamente impraticável. Se precisamos descartar a apresentação de pautas gestuais nas opções de resposta, o que colocar em seu lugar?

Possivelmente, a primeira opção seria recorrer aos SÍMBOLOS FONÉTICOS veiculados pelo IPA. A vantagem seria a de que estaríamos (ainda que com todas as ressalvas contextuais) nos utilizando de um símbolo que representa um som da fala, estando próximo daquele reproduzido no estímulo. As

desvantagens, para citar apenas três, seriam as de que (a) os participantes deveriam, previamente à execução da tarefa, ser familiarizados com os símbolos fonéticos em questão; (b) não se trata da mesma unidade com a qual compactuamos em um aporte gestual; e (c) por serem unidades discretas, o próprio instrumento de testagem imporá categorias prontas aos sujeitos. Assim, por mais que, idealmente, os participantes dominassem o conjunto de símbolos fonéticos utilizados em prol da tarefa, estaríamos empregando outra unidade de análise, e não a gestual.

Outra provável alternativa seria apresentar apenas as LETRAS de um alfabeto, por exemplo, do alfabeto latino. A vantagem pode ser clara em línguas cuja relação grafo-fônica seja mais transparente, como é o caso do espanhol e do finlandês. No entanto, em línguas cuja relação grafo-fônica seja mais opaca, como é o caso do francês e do inglês, temos uma grande desvantagem. Além disso, a grafia, de modo geral, pode influenciar a percepção fônica, conforme argumentamos no quarto capítulo desta tese, gerando vieses indesejáveis para a mensuração da tarefa. Ademais, mesmo em línguas em que tal relação seja mais transparente, qualquer sistema gráfico torna-se incompleto frente à pluralidade de sons existentes em uma língua.

Embora nosso leque de possibilidades esteja se esgotando, apresentamos outro meio de resposta: a apresentação de PALAVRAS em que esteja presente o elemento fônico em jogo. A vantagem é que, dada a consciência fonológica desenvolvida em adultos, o item fônico em questão seria, em princípio, facilmente recuperado. A desvantagem de tal proposta reside na configuração estrutural das adjacências do gesto (ou da constelação de gestos) que se quer testar, o que teria consequências para a coarticulação. Uma saída para essa desvantagem talvez fosse a de empregar palavras cujas adjacências do gesto a ser testado sejam bastante semelhantes entre si. Porém, essa estratégia está condicionada ao inventário lexical da língua em que figuram os estímulos. Além disso, variáveis como acento, extensão do vocábulo, posição da sílaba na palavra e classe morfológica, por exemplo, estão em constante interação, e poderiam impactar o julgamento do participante acerca da resposta esperada – a menos que fossem devidamente controladas, mas esses fatores poderiam reduzir o número de itens lexicais presentes no experimento.

Das alternativas enumeradas, julgamos que as formas mais adequadas de apresentação das respostas aos estímulos sejam aquelas em que os participantes possam recorrer à palavra e/ou ao segmento para registrar suas escolhas. Embora, obviamente, apresentem desvantagens, ainda parecem ser as mais apropriadas dadas as circunstâncias. O que não podemos perder de vista, contudo, é que, inevitavelmente, incidiremos na situação em que deve haver relações próximas entre

os gestos fônicos e os grafemas, e tais relações são substanciais para que possamos analisar os resultados das categorizações empreendidas pelos aprendizes/percebedores.

Em termos especificamente cognitivos, ressaltamos que as tarefas perceptuais de cunho discriminatório, especialmente aquelas que envolvem a comparação de três elementos fônicos, exigem uma carga de memória bastante alta, conforme apontam Gerrits e Schouten (2004) e Colantoni, Steele e Escudero (2015). Uma vez elaboradas de modo a apresentar um grande número de estímulos, tais tarefas podem vir a comprometer o instrumento não pela dificuldade dos participantes em discriminar entre itens fônicos, mas por demandar excessivamente de seus mecanismos de atenção e memória. Incita-se, portanto, uma tensão entre a quantidade de estímulos adequada à capacidade de processamento perceptual do participante e a quantidade de estímulos necessária para que haja dados suficientes em termos de uma análise estatística robusta e/ou confiável. É preciso, pois, que o pesquisador elabore seu experimento de maneira cautelosa na medida em que previna, dentro do possível, uma sobrecarga cognitiva daqueles que irão responder à tarefa e, ao mesmo tempo, consiga gerar de um número suficiente de respostas (dados) para conduzir suas análises.

Complementarmente, deve restar nítido que o tipo de atividade cognitiva para a realização das tarefas de percepção é diferente em cada uma delas. Isso fica evidente se compararmos, por exemplo, a demanda da memória auditiva (e da memória de trabalho) em uma tarefa de identificação e em uma tarefa de discriminação do tipo *ABX*. Muitas vezes, é possível que baixos índices de acuidade perceptual em dada tarefa não sejam resultado de uma minimizada capacidade de percepção de contrastes, mas que advenham do alto empenho cognitivo utilizado para a resolução da tarefa.

Embora componham um arsenal robusto para a mensuração da percepção fônica de elementos não nativos, nossas possibilidades de experimentos não se resumem somente às tarefas de assimilação, identificação e discriminação. Nesse sentido, o estudo de Bohn e Best (2012) traz uma contribuição importante para a pesquisa no âmbito perceptual: a inclusão de um questionário sobre os elementos fônicos que os participantes ouviram. Embora não discorram largamente sobre a aplicação e sobre as respostas referentes a tal instrumento, os autores afirmam que o questionário poderia configurar um espaço para que os sujeitos pudessem fornecer descrições impressionísticas e informais sobre os estímulos a que tiveram acesso. Esse procedimento pode ser bastante eficaz na interpretação dos dados decorrentes das testagens, uma vez que exprime os relatos dos participantes

sobre a tarefa, contribuindo para um maior entendimento das razões pelas quais se optou por uma ou outra resposta¹³⁰.

Os questionários referentes às tarefas empregadas nas testagens podem apresentar questões que exijam respostas amplas (abertas) ou restritas (fechadas). Qualquer que seja a configuração adotada, esse instrumento nos parece fundamental para poder captar as nuances que uma tarefa mais rígida (como as três apontadas anteriormente) não consegue. As respostas expressas nesse tipo de mensuração podem informar o pesquisador sobre aspectos interessantes referentes às tarefas, seja no que compete a propriedades linguísticas envolvidas na testagem (como a clareza dos contrastes funcionais e as características acústico-articulatórias dos estímulos), seja no que concerne às impressões do respondente sobre as tarefas (número de questões e nitidez do estímulo).

Outra potencialidade do questionário é informar o pesquisador sobre o nível de dificuldade das tarefas. Em uma testagem envolvendo uma bateria de tarefas, é indispensável que o pesquisador esteja ciente do nível de dificuldade que as tarefas impuseram aos participantes, além das impressões desses sobre a extensão das tarefas. Essas peculiaridades podem ser úteis à medida que o pesquisador analisa seus dados, podendo estar diante de uma situação em que baixos índices de acuidade em determinada tarefa não sejam o reflexo de uma tarefa que empregue contrastes fônicos sutis, por exemplo, mas decorram de uma tarefa muito extensiva ou cognitivamente onerosa.

5.4 Resumo do capítulo

O presente capítulo, que se ocupou do objetivo (f) desta tese, pretendeu estabelecer algumas ponderações acerca de quais experimentos perceptuais podem ser mais apropriados para verificar a assimilação perceptual e a formação de novas categorias no que diz respeito à percepção fônica não nativa.

Iniciamos a problematização em termos da discussão realizada por Alves e Silva (2016), a qual focou na relação entre teoria de base e metodologia de pesquisa empírica a partir do *PAM/PAM-L2*, acerca da qual os autores elencaram questionamentos de extrema relevância. Uma estratégia

¹³⁰ Alves e Silva (2016) comentam que, no estudo de Best, McRoberts e Sithole (1988), que envolvia a percepção de sons cliques da língua zulu, os participantes percebiam tais unidades como sons que não pertenciam à fala. Essa constatação, de extrema relevância para que o *PAM* fosse cunhado anos adiante, foi acessada através de um questionário aplicado aos sujeitos da pesquisa, após a conclusão das tarefas perceptuais.

interessante de Alves e Silva (2016) foi a de tentar prover esclarecimentos sobre como o realismo direto poderia responder por diferentes questões metodológicas que envolvem a testagem da percepção fônica. Após uma série de reflexões, presumimos que o realismo direto não se sustenta no que concerne ao ambiente laboratorial em que conduzimos nossas pesquisas perceptuais, abrindo espaço para que seu antagonista, o realismo indireto, figure como um candidato virtuoso para que posição filosófica e procedimentos empíricos estejam em consonância.

Neste capítulo, também, discorreremos sobre pontos metodológicos fundamentais quando propomos testar a percepção fônica não nativa. De posse dos trabalhos levantados no terceiro capítulo, traçamos um percurso taxionômico em relação às tarefas utilizadas, o qual se estendeu sobre o delineamento empírico, o tipo de conhecimento de base testado e o objetivo de cada uma delas, a saber, tarefa de assimilação, tarefa de identificação e tarefa de discriminação. Tais constatações nos levaram a problematizar com maior afinco as relações existentes no âmbito da discriminação de elementos fônicos, especialmente no que se refere à demarcação daquilo que é funcional ou não em um sistema fônico, que está altamente relacionada à experiência linguística do aprendiz com a gramática fônica não nativa.

Por fim, essencialmente no que diz respeito à implementação das tarefas perceptuais segundo o enfoque que defendemos nesta tese, tocamos no fator “ponto de corte” entre aquilo que podemos conceber como categorizado ou não categorizado, e também demos destaque a um estudo perceptual que se mostrou bastante completo do ponto de vista metodológico. Ademais, recomendamos a administração de tarefas de assimilação perceptual, identificação e discriminação, assim como a aplicação de questionários de informações sobre as tarefas, no momento em que desejarmos verificar os mapeamentos existentes entre categorias fônicas nativas e não nativas, bem como a instanciação de novas categorias fônicas, formadas a propósito da LNN. Outro aspecto relevante do capítulo foi a discussão sobre a forma de apresentação das opções de respostas em uma dada tarefa perceptual, e, com base em critérios levantados, julgamos que as opções de respostas deveriam contemplar a apresentação de uma palavra em que se insere o item-alvo ou de um segmento com base no IPA.

Retratar questões ligadas a aspectos metodológicos no que compete à percepção fônica não nativa se faz extremamente fundamental para que estejamos alinhados a uma concepção teórica bem delineada, além de possibilitar comparações mais efetivas entre estudos perceptuais que propõem aferir o conhecimento fônico de aprendizes com diferentes históricos linguísticos e sociais. As ponderações elencadas neste capítulo, portanto, não devem ser vistas como instruções normativas,

completas e imutáveis, mas devem ser compreendidas como sugestões metodológicas embasadas no referencial teórico circunscrito nesta tese.

6 CONSIDERAÇÕES FINAIS

O empreendimento teórico que delineamos nesta tese tomou forma a partir do *PAM-L2*. Após nossa aproximação ao modelo e tentativas de compreender seus objetivos e seu alcance, passamos a refletir sobre questões trazidas na proposta, as quais não se faziam coerentes entre si. De modo a oferecer maior consistência aos eixos que sustentam o *PAM-L2*, portanto, percorremos uma linha de argumentação que se voltou a problematizar seus expedientes cognitivo, fônico e filosófico, dada a descontinuidade que notamos haver entre essas três esferas.

Iniciamos o percurso através do segundo capítulo desta tese, em que procedemos com a apresentação do *PAM*, modelo que, mais tarde, daria origem ao *PAM-L2*. Abordamos como o modelo se sustenta, referimos suas características e explanamos a abordagem ecológica de James Gibson, um psicólogo experimental norte-americano, bem como endereçamos as constatações de Carol Fowler sobre a percepção da fala. Tornamos explícita, também, a noção de gesto articulatório como primitivo de análise que se instancia no modelo, e salientamos o evento perceptual, no que tange às unidades da fala, como sendo resultado de condições de assimilação. Esse ensaio nos permitiu avançar na discussão sobre o modelo que derivou do *PAM*, o *PAM-L2*, que tem por objetivo dar conta dos aspectos perceptuais comuns e complementares entre ouvintes sem experiência com o idioma-alvo e aprendizes de uma *L2*. Até esse estágio, não propusemos objeções ao *PAM-L2*, uma vez que nosso intuito era tão somente o de familiarizar o leitor com esse modelo perceptual.

Como um princípio geral, julgamos que toda a problematização deva ser fundamentada. De posse dos esclarecimentos acerca das propriedades gerais inseridas no *PAM-L2*, passamos a examinar as suas potencialidades. Então, no terceiro capítulo, cujo objetivo estava relacionado a verificar a exaustividade das previsões de assimilação do *PAM-L2*, exploramos pesquisas (artigos e dissertações) que tinham enfoque explícito em tais previsões, assim como nos ativemos a estudos que não focalizaram, explicitamente, os referidos padrões. Com base nos trabalhos selecionados, constatamos que o *PAM-L2* se revela um modelo exitoso em termos de suas previsões de assimilação e dá conta de abarcar, no mérito dos resultados, os achados apontados até mesmo nos estudos que não fizeram menção explícita às suas previsões.

Do ponto de vista preditivo, o *PAM-L2* se mostrou bastante eficaz, contemplando as diversas possibilidades de assimilação, tendo em vista os resultados com os quais o confrontamos. Porém, do ponto de vista teórico, o modelo apresentava fragilidades, as quais se instanciavam no âmago de sua

proposta. A primeira delas se referia ao fato de o modelo dar pouca importância para os aspectos cognitivos envolvidos no evento perceptual. Recorrer a representações mentais ou a mecanismos inferenciais, cognitivos por excelência, se mostrou, para os proponentes do modelo, como algo que não só se fazia irrelevante, mas também secundário à sua noção de percepção. A segunda fragilidade, em nossa opinião, era seguir um tratamento puramente articulatório quanto à unidade gestual em questão, o qual parecia ter compromisso majoritário com fatos físicos e observáveis, e minoritário com uma gramática fônica, que tem sua grande parcela de abstração e que se relaciona com outras unidades da língua. O terceiro ponto que entendemos ser passível de críticas era o de o modelo conceber a percepção da fala não nativa como sendo direta, aos moldes do realismo direto de James Gibson. Conforme o próprio psicólogo apontou, suas considerações estavam para fora do âmbito da linguagem, e não davam conta de abarcá-la.

De modo a propor soluções a tais incongruências, oferecemos concepções alternativas a esses três pontos essenciais do *PAM-L2*, as quais se mostram em consonância com os aspectos cognitivo, fônico e filosófico que circunscrevem a percepção fônica não nativa. O primeiro passo foi o de oferecer um aporte cognitivo (BRAISBY; GELLATLY, 2005; MILLER; VANDOME; MCBREWSTER, 2009; QUINLAN; DYSON, 2010; GAZZANIGA *ET AL.*, 2012; KANDEL *ET AL.*, 2014; KEMMERER, 2014; EYSENCK, 2015; STERNBERG, R.; STERNBERG, K., 2016) para endereçar os fenômenos perceptuais humanos, a partir do qual estabelecemos que a percepção constitui um processo fundamentalmente cognitivo, construído e gerenciado pelo encéfalo.

O segundo passo se instaurou com base no oferecimento de um tratamento acústico-articulatório (ALBANO, 2001) à unidade gestual defendida em Best e Tyler (2007), de cunho puramente articulatório. Basicamente, argumentamos que vislumbrar o gesto através do prisma acústico-articulatório traz, no mínimo, três grandes vantagens a um modelo de percepção: (a) a proporcionalidade com que Albano (2001) trata as faces física e abstrata da unidade gestual evidencia sua preocupação tanto com a abordagem precisa das rotinas articulatórias motoras, como com a instanciação simbólica da fala; (b) a relação que Albano (2001) traça entre uma unidade gestual de caráter acústico-articulatório e uma unidade outra presente na língua, como, por exemplo, o morfema, permite que possamos pensar, de fato, em uma gramática fônica, que está amplamente conectada ao léxico; e (c) a saliência dos aspectos acústicos do gesto, defendida por Albano (2001), propicia dar conta de diferenças gestuais que podem não ser vislumbradas se tomarmos como apoio somente aspectos articulatórios, pois existem discrepâncias perceptuais que derivam de

características acústicas específicas, cujo reflexo articulatório se mostra quase nulo. Além disso, tratar a unidade gestual segundo uma noção acústico-articulatória confere multimodalidade a tal primitivo, essencial para que possamos nos valer com afinco das potencialidades que residem nas informações linguísticas.

Repensar essas três esferas nos facultou ponderar sobre outras questões, decorrentes dessas e advindas de nossa postura alternativa aos construtos de base do *PAM-L2*. Nesse sentido, pudemos tocar em pontos resultantes dessa reconcepção do modelo, a saber, a falsa dicotomia entre L_2 e LE , a influência do conhecimento grafêmico na percepção fônica não nativa e a formação de novas categorias fônicas. Embora possamos evidenciar alguns indícios para que nossa proposta venha a se configurar como um modelo de percepção fônica não nativa, alternativo ao *PAM-L2*, tratamos a questão com bastante parcimônia, uma vez que nosso empreendimento teórico se voltou apenas a elencar aspectos que viessem a tornar o próprio *PAM-L2* consistente em relação às suas premissas, e não a que lançássemos um novo modelo perceptual.

O quinto capítulo se constituiu através de ponderações sobre a percepção fônica não nativa no plano experimental. Problematizamos a postura filosófica original do *PAM-L2* em termos do ambiente laboratorial, discutimos questões metodológicas envolvendo a percepção fônica não nativa, principalmente, as tarefas que compõem um teste perceptual, e discorremos sobre a implementação de tais tarefas sob a perspectiva que defendemos nesta tese. Esse capítulo, em especial, dá margem a infundáveis discussões que residem no contínuo teoria-experimentação, pois visa a conceber as premissas a favor das quais advogamos dentro do ambiente laboratorial. Nesse sentido, para mensurar a percepção fônica não nativa, tecemos considerações sobre o delineamento, o tipo de conhecimento de base testado e o objetivo de cada tarefa perceptual que julgamos ser adequada.

Considerando-se toda a exposição que fizemos ao longo deste trabalho, defendemos a tese de que a percepção fônica de línguas não nativas advém necessariamente de um mecanismo cognitivo, cujas unidades fônicas são instanciadas por meio de gestos de caráter acústico-articulatório e acessadas de maneira indireta, na medida em que nossa gramática fônica materna, bem como nosso aparato cognitivo, filtram o conhecimento que temos do ambiente ao nosso redor.

Evidentemente, nossa proposta apresenta limitações em relação a questões teóricas e empíricas. Embora tenhamos nos dedicado a prover consistência interna e externa ao *PAM-L2*, ainda precisamos de mais fundamentação no que diz respeito, principalmente, às questões decorrentes de nossa abordagem, elencadas, sobretudo, na seção 4.4. Destacamos, especificamente, a influência do

sistema grafêmico na percepção fônica e a formação de novas categorias fônicas. No que diz respeito à primeira, sabemos que não há modelos de percepção que contemplem aspectos da grafia, e, por essa razão, poderíamos nos eximir de tal compromisso. No entanto, diversos estudos [(ALVES, 2004; SILVEIRA; 2004) por exemplo] evidenciam o papel da escrita na percepção fônica, o que nos leva a desejar problematizar tal aspecto. Ainda que tenhamos ponderado sobre tal fenômeno, no sentido de chamar atenção para o fato de que algumas variáveis que respondem pelas discrepâncias grafêmicas na percepção fônica possam ser elencadas (como tipologia e nível de opacidade) e vislumbradas frente aos resultados, não fomos categóricos em afirmar como um modelo poderia abarcar tal influência dentro de seu núcleo. No que compete à formação de novas categorias fônicas, procuramos explicar de que maneira elas são previstas e como podem ser arquitetadas, mas não somos capazes de informar precisamente os fatores, além dos cognitivos e gestuais, que poderiam acelerar ou retardar a formação de tais categorias.

No que se refere às limitações empíricas, julgamos apropriado mencionar que não contamos com nenhum experimento criado para os propósitos desta tese. Muito embora possamos ir além da proposta referente ao *PAM-L2*, a qual não apontava, em seu texto seminal, como poderíamos testar as previsões de assimilação, nosso trabalho não comportou a aplicação de tarefas no plano empírico. Recorrer a instrumentos de coleta de dados e, de fato, aplicar as tarefas que mencionamos ser apropriadas constituem passos que deverão ficar para pesquisas futuras, portanto.

Finalmente, esperamos ter podido contribuir para com o rol de discussões que permeiam a área, de maneira que nossas considerações teóricas e metodológicas possam ser úteis para os demais pesquisadores. É nosso intuito, também, que esta tese suscite discussões sobre o fenômeno perceptual acerca das unidades fônicas das línguas e que seja problematizada por nossos pares, quer no plano teórico, quer no plano empírico.

REFERÊNCIAS

- ADAMI, C. What is complexity? *Bioessays*, v. 24, n. 12, p. 1085-1094, 2002.
- ADOLPH, K.; KRETCH, K. Gibson's theory of perceptual learning. In: WRIGHT, J. (Ed.). **International Encyclopedia of the Social & Behavioral Sciences**. 2. ed. Nova Iorque: Elsevier, 2015. 127-134.
- ALBANO, E. **O gesto e suas bordas**: Esboço de fonologia acústico-articulatória do português brasileiro. Campinas: Mercado de Letras, 2001.
- ALBANO, E. Representações dinâmicas e distribuídas: Índícios do português brasileiro adulto e infantil. *Letras de Hoje*, v. 42, n. 1, p. 131-150, 2007.
- ALBUQUERQUE, J. **Aspectos da percepção da dessonorização terminal do inglês por falantes nativos de português brasileiro**. 2012. 213 f. Dissertação (Mestrado em Letras)—Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 2012.
- ALVES, U. Estudios recientes sobre la adquisición fonético-fonológica de lenguas extranjeras desarrollados en Brasil. In: LUCHINI, P.; JURADO, A.; ALVES, U. (Eds.). **Fonética y fonología: articulación entre enseñanza e investigación**. Mar del Plata: Editorial Universidad Nacional de Mar del Plata, 2015. p. 98-109.
- ALVES, U. **O papel da instrução explícita na aquisição fonológica do inglês como L2: Evidências fornecidas pela Teoria da Otimidade**. 2004. 335 f. Dissertação (Mestrado em Letras)—Escola de Educação, Universidade Católica de Pelotas, Pelotas, 2004.
- ALVES, U.; SILVA, A. Implicações de uma perspectiva realista direta para o PAM-L2: Desafios teórico-metodológicos. *Revista do GEL*, v. 13, n. 1, p. 107-131, 2016.
- 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*, v. 59, n. 1, p. 157-180, 2015.
- AMARAL, D. A organização funcional da percepção e do movimento. In: KANDEL, E.; SCHWARTZ, J.; JESSEL, T.; SIEGELBAUM, S.; HUDSPETHET, A. (Eds.). **Princípios de neurociências**. 5. ed. Porto Alegre: Artmed, 2014. p. 315-325.

- ANTONIOU, M.; BEST, C.; TYLER, M. Focusing the lens of language experience: Perception of Ma'di stops by Greek and English bilinguals and monolinguals. **Journal of the Acoustical Society of America**, v. 133, n. 4, p. 406-412, 2013.
- 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, n. 4, p. 582-594, 2012.
- BAK, P.; CHEN, K. Self-organized criticality. **Scientific American**, v. 264, n. 1, p. 26-33, 1991.
- BASSETTI, B. Orthographic input and phonological representations in learners of Chinese as a foreign language. **Written Language and Literacy**, v. 9, n. 1, p. 95-114, 2006.
- BASSETTI, B. Orthographic input and second language phonology. In: PISKE, T; YOUNG-SCHOLTEN, M. (Eds.). **Input matters in SLA**. Clevedon: Multilingual Matters, 2008. p. 191-206.
- 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, n. 1, p. 1-26, 2009.
- BECKNER, C.; BYBEE, J. A usage-based account of constituency and reanalysis. **Language Learning**, v. 59, n. 1, p. 27-46, 2009.
- BENOIT, C.; MASSARO, D.; COHEN, M. Modality integration: Facial movement & speech synthesis. In: COLE, R.; MARIANI, J.; USZKOREIT, H.; VARILLE, G.; ZAENEN, A.; ZAMPOLLI, A.; ZUE, V. (Eds.). **Survey of the state of the art in human language technology**: Cambridge: Cambridge University Press, 1997. p. 311-312.
- 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. Adult perception of nonnative contrasts differing in assimilation to native phonological categories. **Journal of the Acoustical Society of America**, v. 88, n. s1, p. s177, 1990.
- BEST, C. Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In: BOYSSON-BARDIES, B.; SCHONEN, S.; JUSCZYK, P.; MACNEILAGE, P.; MORTON, J. **Developmental neurocognition**: Speech and face processing in the first year of life. Dordrecht: Kluwer Academic Publishers, 1993. p. 289-304.

BEST, C. Learning to perceive the sound pattern of English. In: ROVEE-COLLIER, C.; LIPSITT, L. (Eds.). **Advances in infancy research**, Hillsdale: Ablex Publishers, 1994b. p. 217-304.

BEST, C. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In: GOODMAN, J.; NUSBAUM, H. **The development of speech perception: The transition from speech sounds to spoken words**. Cambridge: MIT Press, 1994a. p. 167-224.

BEST, C.; HALLÉ, P. Perception of initial obstruent voicing is influenced by gestural organization. **Journal of Phonetics**, v. 38, n. 1, p. 109-126, 2010.

BEST, C.; McROBERTS, G.; SITHOLE, N. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. **Journal of Experimental Psychology: Human Perception and Performance**, v. 14, n. 3, p. 345-360, 1998.

BEST, C.; STRANGE, W. Effects of phonological and phonetic factors on cross-language perception of approximants. **Journal of Phonetics**, v. 20, p. 305-30, 1992.

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**. Philadelphia: John Benjamins Publishing Company, 2007, p. 13-34.

BOHN, O.; BEST, C. Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. **Journal of Phonetics**, v. 40, n. 1, p. 109-128, 2012.

BORGHI, A. Object concepts and action. In: PERCHER, D.; ZWAAN, R. (Eds.). **Grounding cognition: The role of perception and action in memory, language, and thinking**. Nova Iorque: Cambridge University Press, 2005. p. 8-34.

BRAISBY, N.; GELLATLY, A. **Cognitive psychology**. Nova Iorque: Oxford University Press, 2005.

BROWMAN, C.; GOLDSTEIN, L. Articulatory gestures as phonological units. **Haskins Laboratories Status Report on Speech Research**, v. 100, p. 69-101, 1989.

BROWMAN, C.; GOLDSTEIN, L. Articulatory phonology: An overview. **Haskins Laboratories Status Report on Speech Research**, v. 112, p. 23-42, 1992.

BROWMAN, C.; GOLDSTEIN, L. Dynamics and articulatory phonology. **Haskins Laboratories Status Report on Speech Research**, v. 113, p. 51-62, 1993.

BROWMAN, C.; GOLDSTEIN, L. Gestural syllable position effects in American English. In: BELL-BERTI, F.; RAPHAEL, L. (Eds.). **Producing speech: Contemporary issues**. Woodbury: AIP Press, 1995. p. 19-33.

BROWMAN, C.; GOLDSTEIN, L. Towards an articulatory phonology. **Phonology Yearbook**, v. 3, p. 219-252, 1986.

BROWN, C. The interrelation between speech perception and phonological acquisition from infant to adult. In: ARCHIBLAD, J. (Ed.) **Second language acquisition and linguistic theory**. Oxford: Blackwell, 2000. p. 4-63.

BROWN, D. Indirect perceptual realism and demonstratives. **Philosophical Studies**, v. 145, n. 3, p. 377-394, 2009.

BUNDGAARD-NIELSEN, R.; BEST, C.; KROOS, C.; TYLER, M. Second language learners' vocabulary expansion is associated with improved second language vowel intelligibility. **Applied Psycholinguistics**, v. 33, n. 3, p. 643-664, 2012.

BUNDGAARD-NIELSEN, R.; BEST, C.; TYLER, M. Vocabulary size is associated with second language vowel perception performance in adult learners. **Studies in Second Language Acquisition**, v. 33, n. 3, p. 433-461, 2011b.

BUNDGAARD-NIELSEN, R.; BEST, C.; TYLER, M. Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. **Applied Psycholinguistics**, v. 32, n. 1, p. 51-67, 2011a.

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. **Clearinghouse on Languages and Linguistics**. Washington: Georgetown University, 1993.

CHOMSKY, N. **Linguística cartesiana**. São Paulo: Editora da Universidade de São Paulo, 1972.

CHOMSKY, N.; HALLE, M. **The sound pattern of English**. Nova Iorque: Harper & Row, 1968.

CILLIERS, P. **Complexity and postmodernism**. Londres: Routledge, 1998.

COLANTONI, L.; STEELE, J.; ESCUDERO, P. **Second language speech: Theory and practice**. Cambridge: Cambridge University Press, 2015.

COLE, R. A.; SCOTT, B. Toward a theory of speech perception. **Psychological Review**, v. 81, n. 4, p. 348-374, 1974.

CORNISH, H.; TAMARIZ, M.; KIRBY, S. Complex adaptive systems and the origins of adaptive structure: What experiments can tell us. **Language Learning**, v. 59, n. 1, p. 187-205, 2009.

CRISTÓFARO-SILVA, T. **Dicionário de fonética e fonologia**. São Paulo: Contexto, 2011.

CUNHA, A. **A hipo e a hipersegmentação nos dados de aquisição da escrita: um estudo sobre a influência da prosódia**. 2004. 130 f. Dissertação (Mestrado em Educação)—Programa de Pós-Graduação em Educação, Universidade Federal de Pelotas, Pelotas, 2004.

CUNHA, A; MIRANDA, A. A hipersegmentação nos dados de aquisição da escrita. In: SEMINÁRIO NACIONAL SOBRE LINGUAGEM E ENSINO (SENALE), 4., 2006, Pelotas. **Anais...** Pelotas: Universidade Católica de Pelotas, 2006. p. 1-10.

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

DARCY, I.; KRÜGER, F. Vowel perception and production in Turkish children acquiring L2 German. **Journal of Phonetics**, v. 40, n. 4, p. 568-581, 2012.

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.

DE LACY, P. **The Cambridge handbook of phonology**. Nova Iorque: Cambridge University Press, 2007.

DELATTRE, P. C.; LIBERMAN, A. M.; COOPER, F. S. Acoustical loci and transitional cues for consonants. **Journal of the Acoustical Society of America**, v. 27, n. 4, p. 769-773, 1955.

DETEY, S.; NESPOULOUS, J. Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. **Lingua**, v. 118, n. 1, 2008. P. 66-81.

DUFOUR, S.; KRIEGEL, S.; ALEESAIB, M.; NGUYEN, N. The perception of the French /s/-/[ʃ] contrast in early Creole-French bilinguals. **Frontiers in Psychology**, v. 22, n. 5, p. 1-8, 2014.

ELLIS, N.; LARSEN-FREEMAN, D. Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. **Language Learning**, v. 59, n. 1, p. 90-125, 2009.

ELPO, M. **A construção da fonologia das vogais frontais arredondadas do francês**. 2015. 86 f. Dissertação (Mestrado em Letras)- Programa de Pós-Graduação em Letras, Universidade Católica de Pelotas, Pelotas, 2015.

ELVIN, J.; ESCUDERO, P.; VASILIEV, P. Spanish is better than English for discriminating Portuguese vowels: acoustic similarity versus vowel inventory size. **Frontiers in Psychology**, v. 5, p. 1-8, 2004.

ESCUADERO, P.; BOERSMA, P. The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. In: ANNUAL BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 26., 2002, Boston. **Anais...** Boston: Boston University, 2002. p. 208-219.

ESCUADERO, P.; CHLÁDKOVÁ, K. Spanish listeners' perception of American and Southern British English vowels. **Journal of the Acoustical Society of America**, v. 128, n. 5, p. 254-260, 2010.

ESCUADERO, P.; SIMON, E.; MITTERER, H. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarities predict and explain cross-linguistic and L2 perception. **Journal of Phonetics**, v. 40, n. 2, p. 280-288, 2012.

ESCUADERO, P.; SISINNI, B.; GRIMALDI, M. The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. **Journal of the Acoustical Society of America**, v. 135, n. 3, p. 1577-1584, 2014.

ESCUADERO, P.; VASILIEV, P. Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. **Journal of the Acoustical Society of America**, v. 130, n. 5, p. 277-283, 2011.

ESCUADERO, P.; WILLIAMS, D. Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. **Journal of the Acoustical Society of America**, v. 131, n. 5, p. 406-412, 2012.

ESCUADERO, P.; WILLIAMS, D. Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. **Journal of the Acoustical Society of America**, v. 129, n. 1, p. 1-7, 2011.

EYSENCK, M. **Cognitive psychology**. Londres: Sage Publications, 2015.

FABRA, L.; ROMERO, J. Native Catalan learners' perception and production of English vowels. **Journal of Phonetics**, v. 40, p. 491-508, 2012.

FERNÁNDEZ, H. **Lecciones de psicología cognitiva**. Buenos Aires: Editorial UAI, 2013.

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. 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. Coarticulation and theories of extrinsic timing control. **Journal of Phonetics**, v. 8, p.113-133, 1980.

FOWLER, C. Listeners do hear sounds, not tongues. **Journal of the Acoustical Society of America**, v. 99, n. 3, p. 1730-1741, 1996.

FOWLER, C. Production and perception of coarticulation among stressed and unstressed vowels. **Journal of Speech and Hearing Research**, v. 24, n. 1, p. 127-139, 1981.

FOWLER, C. Speech production. In GASKELL, M. (Ed.) **The Oxford handbook of psycholinguistics**. Oxford: Oxford University Press, 2007. p. 489-501.

FOWLER, C.; SALTZMAN, E. Coordination and coarticulation in speech production. **Language and Speech**, v. 36, p.171-195, 1993.

GARDNER, E.; JOHNSON, K. A codificação sensorial. In: KANDEL, E.; SCHWARTZ, J.; JESSEL, T.; SIEGELBAUM, S.; HUDSPETHET, A. (Eds.). **Princípios de neurociências**. 5. ed. Porto Alegre: Artmed, 2014. p. 392-413.

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

GERRITS, E. **The categorisation of speech sounds by adults and children: A study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues**. Tese (Doutorado em Linguística)—Netherlands Graduate School of Linguistics, Universidade de Utrecht, 2001.

GERRITS, E.; SCHOUTEN, M. Categorical perception depends on the discrimination task.

Perception & Psychophysics, v. 66, n. 3, p. 363-376, 2004.

GIBSON, E. Perceptual development from the ecological approach. In: LAMB, M.; BROWN, AL;

ROGOFF, B. (Eds.). **Advances in developmental psychology**, v. 3, 1984b, p. 243-286.

GIBSON, E. Perceptual learning. **Annual Review of Psychology**, v. 14, p. 29-56, 1963.

GIBSON, E. **Principles of perceptual learning and development**. Nova Iorque: Appleton-Century Crofts, 1969.

GIBSON, E. The concept of affordances in development: The renaissance of functionalism. In:

COLLINS, W. (Ed.). **The concept of development: The Minnesota symposia on child psychology**, v. 15, 1982. p. 55-81.

GIBSON, E. The development of perception as an adaptive process: Perception of events in space develops early, while perception of objects shows evolution and greater dependence on learning.

American Scientist, v. 58, n. 1, p. 98-107, 1970.

GIBSON, E.; GIBSON, J. The senses as information-seeking systems. In: GIBSON, E. (Org.). **An**

odyssey in learning and perception. Cambridge: MIT Press, 1972, p. 502-510.

GIBSON, E.; PICK, A. **An ecological approach to perceptual learning and development**. Nova

Iorque: Oxford University Press, 2000.

GIBSON, E.; RADER, N. Attention: The perceiver as performer. In: HALE, G.; LEWIS, M. (Eds.).

Attention and cognitive development. Nova Iorque: Plenum Publishing Corporation, 1979. p. 1-21.

GIBSON, J. **The ecological approach to visual perception**. Nova Iorque: Psychology Press, 1986.

GIBSON, J. **The senses considered as perceptual systems**. Boston: Houghton Mifflin, 1966.

GIBSON, J.; GIBSON, E. Perceptual learning: Differentiation or enrichment? **Psychological**

Review, v. 62, p. 32-41, 1955.

GILBERT, C. A natureza construtiva do processamento visual. In: KANDEL, E.; SCHWARTZ, J.;

JESSEL, T.; SIEGELBAUM, S.; HUDSPETH, A.(Eds.). **Princípios de neurociências**. 5. ed. Porto Alegre: Artmed, 2014. p. 483-501.

- GILICHINSKAYA, Y.; STRANGE, W. Perceptual assimilation of American English vowels by inexperienced Russian listeners. **Journal of the Acoustical Society of America**, v. 128, n. 2, p. 80-85, 2010.
- GOLDSHEN, A. Modality integration: Facial movement & speech recognition. In: COLE, R.; MARIANI, J.; USZKOREIT, H.; VARILLE, G.; ZAENEN, A.; ZAMPOLLI, A.; ZUE, V. (Eds.). **Survey of the state of the art in human language technology**: Cambridge: Cambridge University Press, 1997, p. 309-311.
- GOLDSTEIN, L.; BYRD, D; SALTZMAN, E. The role of vocal tract gestural action units in understanding the evolution of phonology. In: ARBIB, M. (Org.). **Action to language via the mirror neuron system**. Cambridge: Cambridge University Press, 2006. p. 215-249.
- GOLDSTEIN, L.; FOWLER, C. Articulatory Phonology: a phonology for public language use. In: MEYER, A.; SCHILLER, N. (Eds.). **Phonetics and phonology in language comprehension and production: Differences and similarities**. Berlin: Mouton de Gruyter, 2003. p. 159-207.
- GUSSENHOVEN, C.; JACOBS, H. **Understanding phonology**. Londres: Hodder Education, 2011.
- HANCIN-BHATT, B. **Phonological transfer in second language perception and production**. 1994. 236 f. Tese (Doutorado em Linguística)—Escola de Linguística, Universidade de Illinois, Urbana-Champaign, 1994.
- HAYES, B. Compensatory lengthening in moraic phonology. **Linguistic Inquiry**, v. 20, n. 2, p. 253-306, 1989.
- HAYES, B. **Introductory phonology**. Oxford: Blackwell Publishing, 2009.
- HOLLAND, J. **Hidden order**: How adaption builds complexity. Nova Iorque: Perseus Books, 1995.
- HOLLAND, J. Studying complex adaptive systems. **Journal of Systems Science and Complexity**, v. 19, n. 1, p. 1-8, 2006.
- HOLLIDAY, J. Second language experience can hinder the discrimination of nonnative phonological contrasts. **Phonetica**, v. 73, n. 1, p. 33-51, 2016.
- IMAI, M.; KANERO, J.; MASUDA, T. The relation between language, culture, and thought. **Current Opinion in Psychology**, v. 8, p. 70-77, 2016.

ITO, K.; MASUDA, T.; HIOKI, K. Affective information in context and judgment of facial expression: Cultural similarities and variations in context effects between North Americans and East Asians. **Journal of Cross-Cultural Psychology**, v. 43, n. 3, p. 429-445, 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.

KANDEL, E. Das células nervosas à cognição: As representações internas de espaço e ação. In: KANDEL, E.; SCHWARTZ, J.; JESSEL, T.; SIEGELBAUM, S.; HUDSPETHET, A. (Eds.). **Princípios de neurociências**. 5. ed. Porto Alegre: Artmed, 2014. p. 327-344.

KANDEL, E.; SCHWARTZ, J.; JESSEL, T.; SIEGELBAUM, S.; HUDSPETHET, A. **Princípios de neurociências**. 5. ed. Porto Alegre: Artmed, 2014.

KAUFFMAN, S. **At home in the universe**: The search for laws of complexity. Londres: Viking Press, 1995.

KEMMERER, D. **Cognitive neuroscience of language**. Hove: Psychology Press, 2014.

KIPARSKY, P. From cyclic phonology to lexical phonology. In: VAN DER HULST, H; SMITH, N. (Orgs.). **The structure of phonological representation**. Dordrecht: Foris, 1982, p. 131-175.

KOSSLYN, S.; KOENIG, O. **Wet mind**: The new cognitive neuroscience. Nova Iorque: Free Press, 1995.

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

KUHL, P.; MILLER, J. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. **Science**, v. 190, n. 4209, p. 69-72, 1975.

KUHN, P. **A percepção de vogais do inglês por falantes monolíngues (português) e bilíngues (pomerano/português) e o papel do bilinguismo no processamento grafo-fônico-fonológico da L3**. 2011. 91 f. Dissertação (Mestrado em Letras)- Programa de Pós-Graduação em Letras, Universidade Católica de Pelotas, Pelotas, 2011.

LADEFOGED, P. **Vowels and consonants**: An introduction to the sounds of languages. Padstow: Blackwell Publishing, 2005.

LADEFOGED, P.; MADDIESON, I. **The sounds of the world's languages**. Oxford. Blackwell Publishers, 1996.

LAMPRECHT, R. (Ed.). **Aquisição da linguagem**: estudos recentes no Brasil. Porto Alegre: EDIPUCRS, 2009. p. 263-276.

LARSEN-FREEMAN, D. Chaos/complexity science and second language acquisition. **Applied Linguistics**, v. 18, n. 2, p. 141-165, 1997.

LARSEN-FREEMAN, D.; CAMERON, L. **Complex systems and applied linguistics**. Oxford: Oxford University Press, 2008.

LE GOFF, B.; BENOIT, C. A text-to-audiovisual speech synthesizer for French. In: INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 4., 1996, Filadélfia. **Anais...** Filadélfia: Wyndham Franklin Plaza Hotel, 1996. p. 3-6.

LE MORVAN, P. Arguments against direct realism and how to counter them. **American Philosophical Quarterly**, v. 41, n. 3. p. 221-234, 2004.

LECLERCQ, M.; ZIMMERMANN, P. **Applied neuropsychology of attention**: Theory, diagnosis and rehabilitation. Londres: Psychology Press, 2005.

LEWIN, R. **Complexity**: Life on the edge of chaos. Londres: Phoenix, 1993.

LIBERMAN, A.; MATTINGLY, I. The motor theory of speech perception revised. **Cognition**, v. 21, n. 1, p.1-36, 1985.

LOWE, J. Indirect perception and sense data. **The Philosophical Quarterly**, v. 31, n. 125, p. 330-342, 1981.

LUND, N. **Attention and pattern recognition**. Filadélfia: Routledge, 2001.

MACHRY DA SILVA, S. **Aprendizagem fonológica e alofônica em L2: percepção e produção das vogais médias do português por falantes nativos do espanhol**. 2014. 258 f. Tese (Doutorado em Letras)—Instituto de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2004.

MARIANI, J. **Language and speech processing**. Hoboken: John Wiley and Sons, 2009.

MASUDA, T.; GONZALEZ, R.; KWAN, L.; NISBETT, R. Culture and aesthetic preference: Comparing the attention to context of East Asians and European Americans. **Personality and Social Psychology Bulletin**. v. 34, n. 9, p. 1260-1275, 2008.

MASUDA, T.; MESQUITA, B.; TANIDA, S; ELLSWORTH, P.; LEU, J.; VEERDONK; E. Placing the face in context: Cultural differences in the perception of facial emotion. **Journal of Personality and Social Psychology**. v. 94, n. 3, p. 365-381, 2008.

MCGURK, H.; MACDONALD, J. Hearing lips and seeing voices. **Nature**, v. 264, p. 746-74, 1976.

McKELVIE-SEBILEAU, P.; DAVIS, C. Discrimination of foreign language speech contrasts by English monolinguals and French/English bilinguals. **Journal of the Acoustical Society of America**. v. 135, n. 5, p. 3025-3035, 2014.

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.

MILLER, F.; VANDOME, A.; MCBREWSTER, J. **Cognitive neuroscience**. Saarbrücken: Alphascript Publishing, 2009.

MILLER, G. The cognitive revolution: a historical perspective. **Trends in Cognitive Science**. v. 7, n. 3, p. 141-144, 2003.

MIRANDA, A. Aspectos da escrita espontânea e da sua relação com o conhecimento fonológico. In:

MIRANDA, A. Um estudo sobre o erro ortográfico. In: HEINING, O.; FRONZA, C. (Orgs.). **Diálogos entre linguística e educação**. Blumenau: EDIFURB, 2010. p. 141-162.

MOREIRA, L. **A percepção da fricativa interdental surda do inglês (L2) por falantes nativos do português**. 2015. 92 f. Dissertação (Mestrado em Letras)–Programa de Pós-Graduação em Letras, Universidade Católica de Pelotas, Pelotas, 2015.

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

NICHOLS, J. Perception of allophonic vowel variation among French L2 learners. **Concordia Working Papers in Applied Linguistics**, v. 5, p. 451-462, 2014.

NICOLIS, G.; PRIGOGINE, I. **Exploring complexity**. Nova Iorque: Freeman and Co. Publishing, 1989.

NISHIDA, G. **Sobre Teorias de percepção da fala**. 2013. 174 f. Tese (Doutorado em Letras)— Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 2013.

OHALA, J. Speech perception is hearing sounds, not tongues. **Journal of the Acoustical Society of America**, v. 99, n. 3, p. 1718-1725, 1996.

OLIVEIRA, R. **A influência da instrução fonética explícita na produção e na percepção dos fricativos sonoros por hispanofalantes aprendizes de português como língua estrangeira – PLE**. 2016, 161 f. Dissertação (Mestrado em Letras)— Centro de Artes e Comunicação, Universidade Federal de Pernambuco, Recife, 2016.

ORMROD, J. **Essentials of educational psychology: Big ideas to guide effective teaching**. Nova Iorque: Pearson, 2015.

PATIHIS, L.; OH, J.; MOGILNER, T. Phoneme discrimination of an unrelated language: Evidence for a narrow transfer but not a broad-based bilingual advantage. **The International Journal of Bilingualism**, v. 19, n. 1, p. 3-16, 2015.

PEEREMAN, R.; DUFOUR, S.; BURT, J. Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. **Psychonomic Bulletin & Review**, v. 16, n. 2, p. 363-368, 2009.

PEROZZO, R. **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 (Mestrado em Letras)—Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.

PEROZZO, R.; ALVES, U. Uma discussão acerca da aplicação do Perceptual Assimilation Model-L2 à percepção fônica de língua estrangeira: Questões de pesquisa e desafios teóricos. **Domínios de Linguagem**, v. 10, n. 1, p. 733-764, 2016.

PERRE, L.; ZIEGLER, J. On-line activation of orthography in spoken word recognition. **Brain Research**, v. 1188, p. 132-138, 2007.

PERRE, L.; BERTRAND, D.; ZIEGLER, J. Literacy affects spoken language in a non-linguistic task: an ERP study. **Frontiers in Psychology**, v. 2, p. 1-8, 2011.

PETAJAN, E.; BISCHOFF, B.; BODOFF, D.; BROOKE, N. An improved automatic lipreading system to enhance speech recognition. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, CHI88., 1988, Washington. **Anais...** Washington, 1988. p. 19-25.

PETROVA, A.; GASKELL, M.; FERRAND, L. Orthographic consistency and word-frequency effects in auditory word recognition: new evidence from lexical decision and rime detection. **Frontiers in Psychology**, v. 2, p. 1-11, 2011.

PRIMI, R.; NASCIMENTO, S.; SOUZA, A. Critérios para avaliação dos testes psicológicos. In: MACEDO, G.; MORETZSOHN, R.; VASCONCELOS, R. (Orgs.). **Avaliação dos testes psicológicos**. São Paulo: Conselho Federal de Psicologia, 2004. p. 21-30.

QUINLAN, P.; DYSON, B. **Cognitive psychology**. Londres: Prentice Hall, 2010.

RASTLE, K.; MCCORMICK, S.; BAYLISS, L.; DAVIS, C. Orthography influences the perception and production of speech. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, v. 37, n. 6, p. 1588-1594, 2011.

REPP, B. Categorical perception: Issues, methods, findings. **Speech and Language: Advances in Basic Research and Practice**, v. 10, p. 243-335, 1984.

SALTZMAN, E. Task dynamic coordination of the speech articulators: A preliminary model. **Haskins Laboratories Status Report on Speech Research**, v. 84, p. 1-18, 1985.

SALTZMAN, E.; KELSO, J. Skilled actions: A task dynamic approach. **Haskins Laboratories Status Report on Speech Research**, v. 76, p. 3-50, 1983.

SALTZMAN, E.; MUNHALL, K. A dynamical approach to gestural patterning in speech production. **Ecological Psychology**, v. 1, n. 4, p. 333-382, 1989.

SANTOS, G. **Percepção e produção das vogais médias do espanhol por falantes do português brasileiro**. 2014. 170 f. Tese (Doutorado em Letras)—Centro de Educação e Comunicação, Universidade Católica de Pelotas, Pelotas, 2014.

SANTOS, G.; RAUBER, A. Percepção e produção das vogais médias do espanhol/LE. In: ALVES, U. (Org.). **Aquisição fonético-fonológica de língua estrangeira: Investigações rio-grandenses e argentinas em discussão**. Campinas: Pontes, 2016, p. 151-178.

SCHACTER, D.; WAGNER, A. Aprendizado e memória. In: KANDEL, E.; SCHWARTZ, J.; JESSEL, T.; SIEGELBAUM, S.; HUDSPETHET, A. (Eds.). **Princípios de neurociências**. 5. ed. Porto Alegre: Artmed, 2014. p. 1256-1273.

SCHOENEMANN, P. Evolution of brain and language. **Language Learning**, v. 59, n. s1, p. 162-186, 2009.

SCHOUTEN, B.; GERRITS, E.; VAN HESSEN, A. The end of categorical perception as we know it. **Speech Communication**, v. 41, n. 1, p. 71-80, 2003.

SCHWARTZHAUPT, B. **Testing intelligibility in English: the effects of positive VOT and contextual information in a sentence-transcription task**. 2015. 86 f. Dissertação (Mestrado em Letras)—Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.

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, n. 2, p. 311-334, 2015.

SEIDENBERG, M.; TANNENHAUS, M. Orthographic effects on rhyme monitoring. **Journal of Experimental Psychology: Human Learning and Memory**, v. 5, n. 6, p. 546-554, 1979.

SHAFIRO, V.; LEVY, E.; KHAMIS-DAKWAR, R.; KHARKHURIN, A. Perceptual confusions of American-English vowels and consonants by native Arabic bilinguals. **Language and Speech**, v. 56, n. 2, p. 145-161, 2012.

SHOWALTER, C.; HAYES-HARB, R. Unfamiliar orthographic information and second language word learning: A novel lexical study. **Second Language Research**, v. 29, n. 2, p. 185-200, 2013.

SILVA, A. A variável tempo nos estudos em aquisição. In: BROWERMANN-ALBINI, A.; GOMES, M. (Orgs.). **O jeitinho brasileiro de falar inglês**: Pesquisas sobre a pronúncia do inglês por falantes brasileiros. Campinas: Pontes, 2014. p. 33-47.

SILVA, A. **As fronteiras entre fonética e fonologia e a alofonia dos róticos iniciais em PB**: dados de dois informantes do sul do país. 2002. 213 f. Tese (Doutorado em Linguística)—Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2002.

SILVEIRA, R. **The influence of pronunciation instruction on the perception and production of English word-final consonants**. 2004. 271 f. Tese (Doutorado em Letras)—Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2004.

SOUZA, M. **Produção e percepção das vogais e das fricativas /θ/ e /ð/ da língua inglesa por alunos de um curso de Letras**. 2012. 170 f. Tese (Doutorado em Letras)—Faculdade de Ciências e Letras, Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara, 2012.

STERNBERG, R.; STERNBERG, K. **Cognitive psychology**. Belmont: Wadsworth Publishing, 2016.

STÖLTEN, K.; ABRAHAMSSON, N.; HYLSTENSTAM, K. Effects of age of learning on voice onset time: Categorical perception of Swedish stops by near-native L2 speakers. **Language and Speech**, v. 57, n. 4, p. 425-450, 2014.

STRANGE, W. Automatic selective perception (ASP) of first and second language speech: A working model. **Journal of Phonetics**. v. 39, p. 456-466, 2010.

STRANGE, W.; HISAGI, M.; AKAHANE-YAMADA, R.; KUBO, R. Cross-language perceptual similarity predicts categorical discrimination of American vowels by naïve Japanese listeners. **Journal of the Acoustical Society of America**. v. 130, n. 4, p. 226-231, 2011.

STYLES, E. **Attention, perception and memory: An integrated introduction**. Nova Iorque: Psychology Press, 2005.

TRUBETZKOY, N. **Principes de Phonologie**. Paris: Encyclopaedia Universalis, 2016.

TUINMAN, A.; CUTLER, A. Casual speech processes: L1 knowledge and L2 speech perception. In: INTERNATIONAL SYMPOSIUM ON THE ACQUISITION OF SECOND LANGUAGE SPEECH (NEW SOUNDS), 6., 2010, Poznań. **Anais...** Poznań: Adama Mickiewicz University, 2010. p. 512-517.

TYLER, M.; BEST, C.; FABER, A.; LEVITT, A. Perceptual assimilation and discrimination of non-native vowel contrasts. **Phonetica**, v. 71, n. 1, p. 4-21, 2014.

WILLIAMS, D.; ESCUDERO, P. Influences of listeners' native and other dialects on cross-language vowel perception. **Frontiers in Psychology**, v. 5, p. 1-10, 2014.

WOOD, S. X-ray and model studies of vowel articulation. **Working Papers in Linguistics** (Lund University), v. 23, p. 1-191, 1982.

YANG, J.; FOX, R. Perception of English vowels by bilingual Chinese-English and corresponding monolingual listeners. **Language and Speech**, v. 57, n. 2, p. 215-237, 2014.

YOLTON, J. Gibson's realism. **Synthese**, v. 19, p. 400-407, 1968.

ZIEGLER, J.; FERRAND, L. Orthography shapes the perception of speech: The consistency effect in auditory word recognition. **Psychonomic Bulletin & Review**, v. 5, n. 4, p. 683-689, 1998.

ZIMMER, M. **A transferência do conhecimento fonético-fonológico do português brasileiro (L1) para o inglês (L2) na recodificação leitora**: Uma abordagem conexionista. 2004. 187 f. Tese (Doutorado em Letras)—Instituto de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2004.

ZIMMER, M.; ALVES, U. A produção de aspectos fonéticos/fonológicos da L2: instrução explícita e conexionismo. **Revista Linguagem e Ensino**, v. 9, n. 2, p. 101-143, 2006.

ZIMMER, M.; ALVES, U. Uma visão dinâmica da produção da fala em L2: O caso da dessonorização terminal. **Revista da ABRALIN**, v. 11, n.1, p. 221-272, 2012.

ZIMMER, M.; SILVEIRA, R.; ALVES, U. **Pronunciation instruction for Brazilians**: Bringing theory and practice together. New Castle: Cambridge Scholars Publishing, 2009.

APÊNDICES

APÊNDICE A – Lista de periódicos consultados para a revisão da literatura

Quadro 4 – Periódicos consultados para a presente tese (continua)

AILA Review
Alfa: Revista de Linguística (UNESP. Online)
Applied Linguistics
Applied Psycholinguistics
Attention, Perception, & Psychophysics
Australian Journal of Psychology
Avances en Psicología Latinoamericana
Biolinguistics
Brain and Cognition
Brain and Language
Cadernos de Estudos Linguísticos (UNICAMP)
Calidoscopio (Online)
Clinical Linguistics & Phonetics
DELTA Documentação de Estudos em Linguística Teórica e Aplicada (Online)
Developmental Psychobiology
Frontiers in Psychology
Functions of Language
Ilha do Desterro
Infancy
Infant behavior and development
International Journal of American Linguistics
International Journal of Bilingualism
Journal of Clinical and Experimental Neuropsychology
Journal of Experimental Child Psychology
Journal of Experimental Psychology
Journal of Language and Linguistic Studies
Journal of Phonetics

Quadro 4 – Periódicos consultados para a presente tese (continuação)

Journal of the Acoustical Society of America
Journal of the International Phonetic Association
Laboratory Phonology
Language
Language Acquisition
Language and Cognitive Processes
Language and Speech
Language Learning & Technology
Letras de Hoje (Online)
Linguistic Inquiry
Linguistics in the Netherlands (Print)
Lodz Studies in Language
Logopedics, Phoniatics and Vocology
Phonetica
Poznań Studies in Contemporary Linguistics
Psychological Science
Revista Brasileira de Linguística Aplicada
Revista da ABRALIN
Revista da ANPOLL (Online)
Revista de Estudos da Linguagem
Revista do GEL
Revista Linguagem & Ensino (Online)
RLA. Revista de Linguística Teórica y Aplicada (Impresa)
Speech Communication (Print)
Studi e Saggi Linguistici
Studies in Second Language Acquisition
Toronto Working Papers in Linguistics (Online)

Fonte: elaborado pelo autor.