

PAULINA DALVA ARTIMONTE ROCCA

A TECNOLOGIA DE FALA APLICADA AO ENSINO DE  
ENTOAÇÃO DA LÍNGUA INGLESA PARA FALANTES NATIVOS DE LÍNGUA  
PORTUGUESA

Doutorado em Lingüística Aplicada e Estudos da Linguagem

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO  
2003

PAULINA DALVA ARTIMONTE ROCCA

A TECNOLOGIA DE FALA APLICADA AO ENSINO DE  
ENTOAÇÃO DA LÍNGUA INGLESA PARA FALANTES NATIVOS DE LÍNGUA  
PORTUGUESA

Tese apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para a obtenção do título de Doutor em Lingüística Aplicada e Estudos da Linguagem, sob orientação da Profa. Dra. Sandra Madureira.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO  
2003

**Comissão Julgadora**

---

---

---

---

---

Autorizo, exclusivamente para fins acadêmicos e científicos, a reprodução total ou parcial desta tese por processos de fotocópias ou eletrônicos.

Assinatura:

São Paulo, 31 de agosto de 2003

À minha mãe

À Profª Drª Sandra Madureira,

inspiradora, incentivadora,

força propulsora em todas as etapas do projeto,

pela amizade e dedicação desmedida,

meu agradecimento;

por proporcionar novas perspectivas de pesquisa para

os Estudos da Linguagem e Ensino de Línguas na PUC/SP com a criação do Grupo de Estudos da Fala e a implantação do Laboratório Integrado de Análise Acústica e Cognição,

minha profunda admiração.

## MEUS AGRADECIMENTOS

Às prof<sup>as</sup> Dr<sup>as</sup> Aglael Gama Rossi e Zaina A. Abdalla Nunes, pelo estímulo e sugestões nas diversas etapas de avaliação deste trabalho.

Aos professores doutores membros da Banca Examinadora, por procederem à análise deste trabalho de pesquisa.

À Comissão de Pesquisa do Conselho de Ensino, Pesquisa e Extensão da Pontifícia Universidade Católica de São Paulo, pelo subsídio financeiro à pesquisa na modalidade de capacitação docente.

Aos alunos do curso de graduação Letras-Inglês da PUC/SP, sem os quais não teria sido possível desenvolver este trabalho, por acolherem minha proposta, participando dos treinamentos com dedicação e entusiasmo.

Aos voluntários, falantes nativos do inglês americano, que, com espírito de cooperação, responderam ao Teste de Percepção, oferecendo horas de trabalho pela Internet.

À Prof<sup>a</sup> Yara de Castro, pela assessoria na análise estatística dos dados.

A Ernesto Luis Foschi, técnico do Laboratório de Rádio da Faculdade de Comunicação e Filosofia pela seriedade e entusiasmo com que se dedicou às horas de gravação e edição dos dados.

Aos funcionários do Laboratório de Linguagem e Informática da Faculdade de Comunicação e Filosofia, Eduardo Vanso Kawanishi, Jefferson dos Santos, Leila Houf Marreiro e Mauricio Alexandre de Souza Campos, pelo apoio nas questões técnicas.

A Valdemir Gonçalves dos Santos, analista de sistemas e Webmaster, por ter atendido a minha solicitação, elaborando o programa Teste de Percepção com muita competência e grande interesse acadêmico.

Ao Prof. Marcello Marcelino, pelo apoio incondicional e pelo empréstimo de seus “olhos e ouvidos” de falante bilíngüe em várias etapas do trabalho.

A Eduardo Vanso Kawanishi, pela dedicação na formatação final do trabalho.

À Prof<sup>a</sup> Karlene Campos, pela revisão de parte do texto.

Às colegas da Pós-graduação e do Laboratório Integrado de Análise Acústica e Cognição, pela cooperação e incentivo constante.

À chefia do Departamento de Lingüística, prof<sup>a</sup> Dr<sup>a</sup> Maria Aparecida C. Borges da Silva, pelo apoio e compreensão.

Ao meu marido, que se fez ausente, estando sempre presente.

Aos meus filhos, por serem como são.

## SÍMBOLOS FONÉTICOS

Os símbolos fonéticos, utilizados para a transcrição das emissões deste trabalho, foram extraídos do IPA, *International Phonetic Alphabet* e encontram-se listados abaixo.

b - **borrow**

p - **please**

d - **could**

dʒ - **judge**

t - **take**

ð - **this**

n - **explain**

ŋ - **going**

j - **you**

l - **please**

g - **sugar**

k - **could**

z - **please**

s - **some**

ʃ - **sugar**

r - **borrow**

v - **vacation**

w - **Word**

aɪ - **I**

oʊ - **borrow**

eɪ - **take**

ɑː - **are**

ɔː - **word**

iː - **please**

ɪ - **this**

ə - **the**

ʊ - **could**

ˈ - sílaba acentuada

| - pausa

- - sílaba “unreleased” – **word**

h - indica aspiração – **take**

ˌ - indica uma consoante silábica – **vacation**

- - indica ensudercimento de um som sonoro – **please**

## LISTA DE FIGURAS

FIGURA 1.	Exemplos de evolução da freqüência fundamental entre uma consoante e a vogal seguinte em (a) apá e (b) aba extraídos de Mira Mateus (1990: 194).....	115
FIGURA 2.	De 1 a 2, medida de duração da unidade fonético acústica əmf da emissão. “ <i>Could I borrow some sugar</i> ” da Grav2 do sujeito Leila. De 1 a 3, medida do segmento fricativo /ʃ/ .....	135
FIGURA 3.	Ana: Grav1: contorno de F0 da oração: <i>Are you going to take a vacation?</i> com 171Hz no início e com 132Hz no final.....	146
FIGURA 4.	Ana, Grav2: contorno de F0 da oração: <i>Are you going to take a vacation?</i> com 214Hz no início e com 268Hz no final.....	146
FIGURA 5.	Célia, Grav1: contorno de pitch ascendente da emissão: <i>Would you please explain this word?</i> com 213Hz no início e com 244Hz no final.....	147
FIGURA 6.	Célia, Grav2: contorno de pitch sem inclinação, com o início e final em torno de 210Hz, da emissão: “ <i>Would you please explain this word?</i> ”.....	148
FIGURA 7.	Espectrograma da emissão modelo de “ <i>Would you please explain this word</i> ”, /wʊdʒu   plɪːz ɪkˈspleɪ ðɪs wɔːrd /. Medidas de duração das Fricativas: (1) - /dʒ/ = 26ms; (2) - /z/ = 15ms; (3) - /s/ = 55ms; - /s/= 78ms. Duração total da emissão = 1.472.	154
FIGURA 8.	Espectrograma da emissão da Grav1 do sujeito Leila de “ <i>Would you please explain this word</i> ”. [ wʊdʒu   plɪːz ɪkˈspleɪ ðɪs wɔːrd ]. Medidas de duração das fricativas: (1) – [dʒ] = 46ms; (2) – [z] = 88ms; (3) – [s] = 90ms; - [s]= 98ms. Duração total da emissão: 1.664ms.....	154
FIGURA 9.	Parte da onda da emissão modelo da oração “ <i>Could I borrow some sugar</i> ” [kʊd aɪ ˈbɔːrou   səm ˈʃʊgə]......	155
FIGURA 10.	Parte da onda da emissão modelo da oração “ <i>Could I borrow some sugar</i> ” [kʊd aɪ ˈbɔːrou   səm ˈʃʊgə]......	155

FIGURA 11.	Espectrograma de banda larga e oscilograma de parte da emissão “Would you please explain this Word” da Grav1 do sujeito Paula.....	160
FIGURA 12.	Espectrograma de banda larga e oscilograma de parte da emissão “Would you please explain this Word” da Grav2 do sujeito Paula.....	160
FIGURA 13.	Grav1 do sujeito Eda da emissão “ <i>Could I borrow some sugar</i> ”	163
FIGURA 14.	Grav2 do sujeito Eda com pausa de 303ms entre a frase entoacional “ <i>Could I borrow</i> ” e “ <i>some sugar</i> ” .....	163
FIGURA 15.	Curva de pitch do sujeito Rita da Grav1 da emissão “ <i>Could I borrow some sugar</i> ” .....	164
FIGURA 16.	Curva de pitch do sujeito Rita da Grav2 da emissão “ <i>Could I borrow some sugar</i> ” .....	164
FIGURA 17.	Curva de pitch da emissão “ <i>Are you going to take a vacation?</i> ” do sujeito Leila, Grav2.....	165
FIGURA 18.	Curva de pitch da emissão “ <i>Could I borrow some sugar?</i> ” do sujeito Paula, Grav1. Subida de 40Hz em 101ms no contorno final de fronteira.....	165
FIGURA 19.	Curva de pitch da emissão “ <i>Could I borrow some sugar?</i> ” do sujeito Paula, Grav2. Subida de 79Hz em 150ms no contorno final de fronteira.....	166

## LISTA DE TABELAS

TABELA 1.	Avaliadores do Teste de Percepção.....	116
TABELA 2.	Cadastro de Falantes.....	122
TABELA 3.	Cadastro de Sentenças.....	122
TABELA 4.	Cadastro de Falantes X Sentenças.....	122
TABELA 5.	Cadastro de Falantes X Avaliador.....	123
TABELA 6.	Tabela de Teste.....	124
TABELA 7.	Resultado da avaliação perceptiva das emissões dos falantes nativos.....	137
TABELA 8.	Resultado da avaliação perceptiva das emissões interrogativas <i>Yes/No</i> dos alunos do grupo1.....	141
TABELA 9.	Resultado geral da avaliação perceptiva da emissão “ <i>Are you going to take a vacation?</i> ” dos alunos do grupo1.....	142
TABELA 10.	Resultado geral da avaliação perceptiva de cada tipo de oração interrogativa <i>Yes/No</i> dos alunos do grupo1.....	142
TABELA 11.	Medidas de F0 das emissões dos sujeitos e do Modelo da oração “ <i>Are you going to take a vacation?</i> ” O fundo cinza sinaliza a emissão com contorno final descendente.....	144

TABELA 12.	Medidas de F0 das emissões dos sujeitos e do Modelo da oração “Would you please explain this word?” O fundo cinza sinaliza as emissões sem contorno final ascendente.....	144
TABELA 13.	Medidas de F0 das emissões dos sujeitos e do Modelo da oração “Could I borrow some sugar?” Todas as emissões apresentam contorno final ascendente.....	145
TABELA 14.	Avaliação da Gravação <sup>2</sup> da emissão “Are you going to take a vacation” de Paula e Leila.....	151
TABELA 15.	Comparação em milisegundos da duração das fricativas com a duração do conjunto de elementos que compõem as umidades fonético-acústicas em <i>oʊs</i> e em <i>əmf</i> nas Grav1 e Grav2 da emissão “Could I borrow some sugar” pelos sujeitos Rita ,Paula e Ana. Os sinais < , > indicam duração menor ou maior da fricativa.....	157
TABELA 16.	Distribuição dos valores de duração relativa dos segmentos [oʊ]e [s] do GIPC <i>oʊs</i> e dos outros Segmentos da emissão “Could I borrow some sugar?” do modelo e da Grav1 e Grav2 dos sujeitos Ana, Paula e Rita.....	158
TABELA 17.	Total Geral por aluno das avaliações da emissão “Could I borrow some sugar” .....	161

## LISTA DE GRÁFICOS

GRÁFICO 1.	Valores de F0 do núcleo da vogal de cada sílaba da emissão modelo “ <i>Could I borrow some sugar?</i> ”.....	133
GRÁFICO 2.	Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do Grupo1 e do Grupo 2.Valores inferiores indicam melhores resultados.....	138
GRÁFICO 3.	Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do grupo1 e em relação aos exercícios1,2,3.Valores inferiores indicam melhores resultados.....	139
GRÁFICO 4.	Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do grupo2 e em relação aos exercícios 1, 2, 3. Valores inferiores indicam melhores resultados.....	139
GRÁFICO 5.	Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 de cada um dos alunos do grupo1.....	140
GRÁFICO 6.	Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 da aluna Rita.....	141
GRÁFICO 7.	Comparação dos dados do contorno de F0 da emissão “ <i>Are you going to take a vacation</i> ” do modelo e do sujeito Ana.....	147
GRÁFICO 8.	Comparação dos dados do modelo e do sujeito Célia do contorno de F0 da emissão “ <i>Would you please explain this word?</i> ” .....	148
GRÁFICO 9.	Linha de tendência de direção dos valores de pitch da emissão modelo “ <i>Would you please explain this word?</i> ”.....	149
GRÁFICO 10.	Linha de tendência de direção dos valores de pitch da grav2 de Célia da emissão “ <i>Would you please explain this word?</i> ” .....	149
GRÁFICO 11.	Comparação de dados de Célia. Grav1, melhor avaliada: curva final ascendente e subida de <i>pitch</i> de 23Hz em 70ms. Grav2: final nivelado com o início da emissão, subida de 24Hz em 57ms.....	169

RESUMO: Esta pesquisa tem por objetivo desenvolver e avaliar práticas pedagógicas para o desenvolvimento da percepção e produção de contornos entoacionais do inglês por falantes do português brasileiro que façam uso da análise, manipulação e visualização do sinal sonoro por meio de instrumentais de análise acústica da fala. O trabalho envolveu 8 alunos universitários brasileiros, divididos em dois grupos: um experimental, cujo treinamento era apoiado em recursos de tecnologia de fala e conceitos teóricos de Fonética Acústica, e outro de controle, que cursou a disciplina com treinamento no laboratório de línguas. O *corpus* constituiu-se de Gravação Estímulo, que serviu de modelo para a repetição, e de Gravação Desempenho feita antes e após o treinamento dos alunos. As emissões-modelo (orações declarativas, interrogativas *Yes/No* e interrogativas com *WH*-) foram extraídas de fitas cassetes originais que acompanham livro didático. O treinamento do grupo experimental empregou o *software WinSal* da Media Enterprise. Os dados da Gravação Desempenho foram submetidos à análise perceptual de 40 falantes nativos do inglês. A análise acústica, realizada com o apoio do CSRE da Avaaz Corporation, concentrou-se nos dados das emissões interrogativas *Yes/No* do grupo experimental, por terem sido as únicas que apresentaram melhora significativa com o treinamento. Os resultados finais apontam para a interação de vários fatores na percepção e produção do sotaque estrangeiro relativo ao contorno entoacional de interrogativas *Yes/No* do inglês, dentre os quais destacam-se a interação do nível segmental e do supra-segmental e o contorno ascendente de fronteira final de grupo entoacional. Sugerem, também, que o treinamento, baseado na análise acústica e com o apoio visual proporcionado pela tela do computador, proporciona vantagens para o ensino da entoação de uma língua estrangeira.

ABSTRACT: This research aims at implementing and evaluating techniques designed to improve Brazilian students' perception and production of English intonation contours by analyzing, manipulating and visualizing the acoustic signal provided by speech analysis devices. Eight undergraduate students of the English language took part in the experiment that compared the achievement of two groups: (1) the experimental group, which exercised the intonation contours with the speech technology support and (2) the control group, which took the same training in the language laboratory. The *corpus* comprises two types of recordings, the Stimulus Recording, employed as a model to be followed by the students and the Production Recordings, produced before and after the training by the subjects. The Stimulus Recording emissions (declarative, Wh- and Yes/No questions) were extracted from tapes designed to teach English intonation. The experimental group training was based on WinSal, a Media Enterprise software. The Production Recording data were submitted to the perceptual analysis of 40 native speakers of English. The Yes/No questions produced by the experimental group were submitted to acoustic analysis, since they displayed significant improvement after the training. The Computerized Speech Research Environment from Avaaz Corporation was employed for the acoustic analysis. The final results point at several factors interfering in the perception and production of foreign accent in relation to the intonation contour of Yes/No questions such as the interaction of segmental and suprasegmental levels and the rising contour at the end of the intonation group. They also suggest that training based on the acoustic analysis and on the visual support provided by the computer screen can improve the teaching of the intonation of a foreign language.

## **INTRODUÇÃO**

Justificativa.....	2
Objetivo.....	7
Experimento piloto.....	8
Organização geral do estudo.....	11

## **CAPÍTULO 1 - TEORIAS DE AQUISIÇÃO, ABORDAGENS SOBRE ENSINO DE LÍNGUAS E PRONÚNCIA**

1.1 Abordagens teóricas de aquisição de primeira e segunda línguas.....	14
1.1.1 Behaviorismo.....	14
1.1.2 Inatismo.....	19
1.1.3 Conexionismo.....	36
1.1.4 Interacionismo / Construtivismo.....	41
1.1.5 Abrangência explicativa das principais abordagens sobre aquisição.....	48
1.1.6 Aquisição de Segunda língua.....	52
1.2 Abordagens sobre ensino de línguas e pronúncia.....	56

## **CAPÍTULO 2 - PERCEPÇÃO E PRODUÇÃO DA FALA**

2.1 Percepção e produção da fala na aquisição de línguas.....	64
2.2 Origem do sotaque estrangeiro.....	67
2.3 Relação entre percepção e produção na aquisição de Segunda língua...	76
2.4 Percepção e aquisição da entoação.....	79
2.5 Entoação e sotaque estrangeiro.....	81
2.6 Identificação de sotaque estrangeiro entoacional.....	83

## **CAPÍTULO 3 - MODELO DE DESCRIÇÃO ENTOACIONAL**

3.1 Modelo de seqüências tonais.....	88
3.2 Introdução ao ToBI.....	91

## **CAPÍTULO 4 - TECNOLOGIA DE FALA E ENSINO DE ENTOAÇÃO**

4.1 Ensino de entoação.....	100
4.2 Tecnologia de fala e ensino de entoação.....	103

## **CAPÍTULO 5 – EXPERIMENTO**

5.1 Objetivo.....	106
-------------------	-----

5.2	Sujeitos do experimento.....	106
5.3	<i>Corpus</i> da pesquisa.....	108
5.4	Gravações.....	111
5.5	Treinamento.....	112
5.6	Procedimentos de análise dos dados.....	116
5.6.1	Análise perceptiva dos dados.....	117
5.6.1.1	Avaliadores do teste de percepção.....	117
5.6.1.2	Procedimentos na elaboração do teste de percepção.....	119
5.6.1.3	Procedimentos na aplicação do teste de percepção.....	124
5.6.1.4	Análise estatística dos dados de percepção.....	129
5.6.2	Análise acústica dos dados.....	132

## **CAPÍTULO 6 - ANÁLISE DOS RESULTADOS**

6.1	Resultados da avaliação perceptiva.....	137
6.2	Resultados da análise acústica.....	143
6.3	Resultados da avaliação do treinamento pelos alunos.....	170

<b>CONSIDERAÇÕES FINAIS.....</b>	<b>172</b>
----------------------------------	------------

## **ANEXOS**

Anexo 1 -	Protocolo de Avaliação.....	181
Anexo 2 -	Questionário 1: dados pessoais dos sujeitos.....	185
Anexo 3 -	Protocolo de avaliação inicial e final dos sujeitos.....	187
	<i>a- Ficha que os alunos receberam na sala de gravação.....</i>	<i>187</i>
	<i>b- Conteúdo da gravação que os sujeitos ouviram.....</i>	<i>189</i>
Anexo 4 -	Organização das aulas laboratoriais.....	193
Anexo 5 -	Avaliação livre do treinamento.....	208
Anexo 6 -	Avaliação dirigida do treinamento.....	209
Anexo 7 -	Questionário 2: Dados pessoais dos falantes nativos.....	211
Anexo 8 -	Emissões incluídas na análise perceptual por falantes nativos do inglês americano.....	212

<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>214</b>
--	------------

## **INTRODUÇÃO**

## Justificativa

O campo de estudos da voz e da fala tem recebido nas últimas décadas contribuições significativas de várias áreas para a ampliação dos conhecimentos acerca da natureza, função e uso da voz e da fala nas situações de comunicação. Parte da evolução desses estudos deve-se ao desenvolvimento tecnológico, que permite investigações mais detalhadas tanto do funcionamento das estruturas do aparelho fonador durante a produção da fala quanto das propriedades físicas do som resultante.

Nos últimos anos, a tecnologia de fala tem sido apontada<sup>1</sup> nos congressos internacionais sobre Fonética como uma das grandes promessas no que diz respeito ao ensino de pronúncia de línguas. Na medida em que novas e mais complexas tecnologias de fala emergem, a aplicabilidade dos recursos disponíveis aumenta, incluindo desde pesquisas de reconhecimento até síntese de fala. Embora muitas aplicações tenham sido feitas na área de treinamento de pronúncia para deficientes auditivos (Vicsi, Roach, Oster, Kacic & Barczikay, 1999; Jonsson, 1998, entre outros), os estudos referentes à sua aplicação no campo da aprendizagem de segunda língua (Bernstein, 1998; Davies & Poesio 1998; Devlieger (1998); Townshend, Bernstein, Todic & Warren, 1998), ainda exploratórios, somente agora passam a despertar o interesse dos pesquisadores.

Muitos dos experimentos aplicados à educação na área de linguagem incluem tecnologia de reconhecimento de fala, síntese, análise, manipulação e visualização do sinal sonoro e *softwares* de avaliação de pronúncia com medidas de

---

<sup>1</sup>. Vejam-se os trabalhos apresentados no *ESCA Workshop sobre Speech Technology in Language Learning*, publicados nos *Proceedings* do encontro, em 1998.

rejeição, de aceitação e de visualização de *feedback*. Apesar de os resultados, na sua maioria, serem considerados preliminares, acredita-se que as tecnologias de fala possam complementar e fortalecer métodos tradicionais de aprendizagem de línguas, ainda que, por si sós, não sejam solução para qualquer problema: “Without people who know how to use them, books do nothing for literacy, computers do nothing for computation, and speech technology does nothing for language learning” (Price, 1998: 81).

O ensino de línguas baseado na tecnologia é uma atividade que requer um trabalho interdisciplinar, muitas vezes dificultado pela necessidade de reunir profissionais de formação bastante diversa, tais como engenheiros, programadores e professores. No entanto, uma abordagem multidisciplinar, que envolva tecnólogos de fala, lingüistas e especialistas em ensino de línguas, pode indicar caminhos promissores para o uso da tecnologia, que culminem com avanços significativos nos métodos atuais de ensino e aprendizagem, levando em conta as novas ferramentas colocadas à disposição dos professores. Coloca-se, como desafio central, a prova incontestável do papel da tecnologia na aprendizagem de línguas.

Nesse universo, o ensino de entoação<sup>2</sup> apresenta uma série de desafios para o professor de línguas estrangeiras, exigindo dele um posicionamento consciente sobre questões que vão desde a escolha de um modelo teórico apropriado até a escolha de uma abordagem de ensino mais eficaz para grupos específicos de línguas.

---

<sup>2</sup> O termo entoação refere-se a alterações na fala que “correspondem a modulações da frequência fundamental (medida em Hertz), da intensidade (medida em decibéis) e da duração (medida em milisegundos)” (Madureira, 1999:55). A frequência fundamental (F0), termo que designa o número de repetições de ciclos de uma onda sonora, é considerada, no entanto, o parâmetro acústico mais importante da entoação.

Quando um aluno começa a aprender um novo idioma, dedica, em geral, um certo tempo para a aprendizagem de sons que não estão presentes em sua língua materna. Contudo, a experiência de professores tem demonstrado que uma pronúncia dos fones de uma língua estrangeira próxima da do falante nativo, mas com ritmo e *pitch*<sup>3</sup> incorretos, pode ser ininteligível<sup>4</sup>. Jenkins (2000) argumenta que a variação segmental interfere menos na inteligibilidade da fala de não-nativos do que os aspectos supra-segmentais, tais como o acento, o ritmo e a entoação. É o padrão entoacional que une a cadeia de fones, formando a mensagem, guiando o ouvinte durante a emissão, fornecendo pistas contextuais, indicando os itens lexicais de conteúdo, solucionando ambigüidades, imprimindo características de sentido e veiculando atitudes e emoções.

Adquirir entoação semelhante à do falante nativo é uma das tarefas mais difíceis na aprendizagem de línguas estrangeiras (Beckman, 1995). O grau de proficiência alcançado pelo aluno é influenciado por muitos fatores que interagem entre si (Mixdorff, 1996), incluindo a natureza da língua envolvida, as práticas pedagógicas adotadas e as características individuais, tais como habilidade, motivação, idade, entre outras.

Embora muitos traços dos padrões entoacionais sejam considerados universais, porque podem ser encontrados em diferentes línguas, existem especificidades que são próprias de cada uma delas. Vaissière (1983: 54) afirma que “It is often assumed that the acoustic properties of speech are partly shaped by

---

<sup>3</sup> O correlato perceptual da frequência fundamental denomina-se *pitch* e o acento produzido pela percepção de variações na frequência fundamental denomina-se *pitch accent*. Madureira (1999) traduz o termo *pitch* por “inflexão tonal” e emprega a expressão “acentos marcados por inflexão tonal” para se referir a *pitch accents*. No entanto, optamos por empregar os termos *pitch* e acento de *pitch* neste trabalho.

<sup>4</sup> Considera-se a fala inteligível, quando o ouvinte apreende a mensagem tal como pretendida pelo falante.

'natural tendencies' that have been conventionalized to various degrees depending on the language”.

Assim, falantes nativos têm competência para variar a frequência fundamental, o parâmetro acústico mais diretamente ligado à produção de contornos entoacionais, do mesmo modo que têm competência para variar a duração e a intensidade. Porém, a importância relativa de cada um desses elementos e o sentido vinculado a tais contornos são específicos de cada língua.

Por esse motivo, de acordo com Daniel Hirst e Albert Di Cristo (1998), a descrição de um sistema entoacional de uma língua específica é uma tarefa muito difícil de ser realizada quando não se é falante nativo da língua em questão. Esses autores concordam que, apesar de a entoação ser um dos traços mais universais da linguagem humana, ela é, paradoxalmente, um dos mais específicos e completam afirmando que, para um falante não-nativo, pode ser impossível dizer se duas emissões são exemplares de um mesmo padrão entoacional ou não.

Também a produção dos padrões entoacionais traz desafios para o aprendiz de uma língua estrangeira. Em estudo anterior (Rocca & Marcelino, 1997), constatou-se que algumas das dificuldades de entoação encontradas por alunos brasileiros que aprendem inglês estão relacionadas à interferência da língua materna quanto ao aspecto de se conferir proeminência. Há que se mencionar em relação a esse fato que os correlatos acústicos do acento no português brasileiro traduzem-se na maior duração da unidade acentuada e na diminuição de intensidade em sílabas pós-tônicas (Barbosa, 1996; Massini, 1991), enquanto o inglês confere proeminência principalmente por meio do *pitch accent*. Na produção

dos alunos, a insuficiência ou ausência de movimento tonal nas sílabas proeminentes é muitas vezes compensada por um alongamento da duração dessas mesmas sílabas.

Esse mesmo estudo demonstrou que a abordagem do ensino de entoação não pode basear-se apenas no uso de gravadores e tutores nativos e não-nativos, porque a análise puramente perceptual provou ser limitada. Desse modo, o aluno brasileiro encontra dificuldade em perceber as diferenças nos padrões entoacionais, da mesma maneira que o falante nativo é incapaz de apontar o que caracteriza o sotaque estrangeiro que ele detecta na fala do brasileiro. Ficou enfatizada a importância da análise acústica para a explicitação objetiva dos problemas e para o desenvolvimento de técnicas de ensino de entoação que utilizem a análise, manipulação e visualização do sinal sonoro, proporcionadas por programas computacionais para o aperfeiçoamento da percepção e produção dos alunos.

Pesquisa posterior (Rocca, 1998) indica vantagens da observação do contorno do *pitch* na tela do computador para o aprimoramento da percepção e produção de padrões entoacionais do inglês por falantes brasileiros, confirmando relatos de experimentos feitos com outras línguas (Simões, 1996; Ziolkowski & Landahi, 1995). Contudo, os dados obtidos nesse estudo são limitados e permitem apenas conclusões preliminares, porque dizem respeito ao desempenho de apenas um sujeito que desenvolveu o treinamento individualmente, em ambiente laboratorial. No entanto, os resultados sugerem que a abordagem proposta apresenta um ponto de partida promissor para uma nova investigação.

## **Objetivo**

Esta pesquisa tem por objetivo desenvolver e avaliar práticas pedagógicas inspiradas nas novas tecnologias de fala para o ensino da entoação do inglês a falantes do português, de modo a garantir o que Jenkins (2000) denominou inteligibilidade fonológica, ou seja, uma comunicação bem sucedida.

Pretende-se propor uma abordagem multidisciplinar na implementação de estratégias inovadoras para o desenvolvimento da percepção e produção de padrões entoacionais. Destaque-se que tais estratégias devem fazer uso da análise, manipulação e visualização do sinal sonoro por meio de instrumentais de análise acústica da fala, que deverão apoiar o ensino da entoação do inglês para alunos brasileiros. Mais especificamente, pretende-se explorar a possibilidade de construção de uma estratégia de ensino que utilize análises espectrográficas em microcomputadores, o que, no caso da entoação, refere-se à extração da curva de F0, a ser utilizada pelo aluno de língua estrangeira como parâmetro orientador no treinamento de sua produção.

Ao se avaliar a prática pedagógica proposta, pretende-se discutir características dos contornos entoacionais que fazem a entoação interferir na comunicação bem sucedida em contextos de inglês como segunda língua.

## **Experimento piloto**

A construção do experimento desta pesquisa foi orientada pelos resultados de um experimento piloto, que não alteraram a composição do *corpus*, mas

provocaram alterações no projeto quanto aos procedimentos relativos à escolha dos sujeitos e treinamento proposto.

Os sujeitos do experimento piloto eram alunos do curso de graduação Letras-Inglês da PUC/SP, que se ofereceram como voluntários para participarem do treinamento em horários extracurriculares. Foram selecionados 24 alunos, divididos em dois grupos: o grupo experimental, que faria o treinamento com o auxílio do computador, e o grupo de controle, que faria o mesmo treinamento no laboratório de línguas. Esses alunos submeteram-se a uma avaliação no laboratório de rádio (anexo1), a fim de que se determinasse seus graus de competência na percepção de sons da língua estrangeira como iguais ou diferentes; na reprodução de sons e palavras da língua estrangeira; na reprodução de sentenças com padrões entoacionais da língua estrangeira, com e sem apoio visual; na reprodução de diferentes tons musicais. A avaliação tinha, ainda, o intuito de medir a capacidade da memória auditiva.

Com os resultados da avaliação e do questionário de dados pessoais (anexo2), pretendia-se que os dois grupos, experimental e de controle, fossem formados por sujeitos com habilidades semelhantes, sendo 4 alunos de nível básico, 4 de nível intermediário e 4 de nível avançado do curso de graduação de Letras:inglês. Entretanto, a interpretação desses resultados não apontou para conclusões que pudessem orientar na escolha dos sujeitos, de modo a formarem grupos homogêneos. Dessa forma, eles acabaram sendo selecionados e classificados tendo em vista o estágio de língua, pois se tinha como premissa que o treinamento atingiria melhor resultado com os alunos de nível mais básico .

Mesmo assim, o objetivo de selecionar sujeitos com habilidades e características semelhantes foi frustrado no projeto piloto. Alguns alunos apresentaram um nível de proficiência em língua mais avançado do que o previsto pelo estágio que cursavam, porque tinham de cumprir carga horária do currículo. Outros apresentaram um nível de proficiência aquém do previsto, porque haviam trancado matrícula durante um certo tempo e estavam voltando aos estudos naquele semestre. Por isso, tanto o teste de habilidades quanto o questionário com dados pessoais não auxiliaram na divisão homogênea dos grupos. Diante desses resultados, a preocupação em selecionar sujeitos que constituíssem um grupo homogêneo e com grau de competência semelhante na língua inglesa foi abandonada.

A proposta do projeto piloto, de que as atividades do treinamento fossem executadas individualmente e sem a presença do professor, também não foi produtiva. Os alunos do grupo experimental consideraram que o *feedback* proporcionado pela tela do computador era insuficiente e ressentiam-se da falta de apoio do professor. Além disso, o comprometimento que os voluntários demonstraram com as atividades foi por demais instável, o que prejudicou qualquer avaliação dos resultados da experiência. Concluiu-se, então, que a presença do professor era imprescindível e que voluntários não eram adequados para uma atividade de longa duração.

Para a seleção das emissões do *corpus* da gravação inicial e final do experimento piloto, foi adotada a abordagem sintática (O'Connor & Arnold, 1973), pela facilidade que oferece para a organização dos contornos em grupos. Porém, a análise acústica dessas emissões comprovou que os contornos de entoação do inglês baseados em padrões sintáticos não se mostraram neutros, talvez porque não existam padrões entoacionais

neutros, uma vez que, mesmo em uma gravação com finalidade didática de exemplificar certos contornos entoacionais específicos, representa-se uma situação e/ou um interlocutor. Assim, a interrogativa com *Wh-*, descrita canonicamente como um exemplo de curva descendente, apresentou variação na fala dos falantes nativos nas orações chamadas modelo. As interrogativas do tipo *Yes/No*, que têm um padrão ascendente, também apresentaram variáveis dependentes da modalidade. Desse modo, uma interrogativa introduzida pelo modal *could* apresenta características um pouco diferentes daquelas presentes em uma interrogativa introduzida pela partícula *do*.

Embora se concorde com Bolinger (1996) que, em se tratando de entoação, não é possível separar o que seja estritamente gramatical, a escolha das emissões para compor o *corpus* com base no padrão sintático foi mantida diante das desvantagens apresentadas pelas abordagens atitudinal e discursiva.

A abordagem atitudinal, que procura descrever os contornos entoacionais relacionados a atitudes vinculadas a atos de fala, tais como surpresa, raiva, indiferença, entusiasmo, entre outras, foi descartada pela dificuldade em se classificar todas as nuances de atitudes e sentimentos expressos em uma língua, que exigiriam minimamente um contexto mais amplo do que o proposto no experimento.

A abordagem discursiva (Brazil, Coulthard & Johns, 1980), que considera a entoação como sendo a forma de os falantes organizarem e relacionarem significados por meio do discurso, já que esse é o contexto maior de interação entre eles, foi também descartada pela dificuldade em restringir os dados a serem analisados em um experimento que teria duração de apenas 10 sessões. Tal

abordagem não permitiria isolar as frases entoacionais de seu contexto discursivo, visto que significados entoacionais e discursivos estão unidos. Desse modo, manteve-se a abordagem sintática na elaboração do Corpus deste trabalho.

### **Organização geral do estudo**

Ainda que este estudo esteja inserido na área de aquisição de segunda língua, a resenha das teorias de aquisição de primeira língua fez-se necessária, uma vez que essas são retomadas e reformuladas como ponto de partida para as posições teóricas em relação à aquisição de segunda língua. Desse modo, o capítulo 1 apresentará os modelos de aquisição de primeira e segunda línguas, tendo em vista a sustentação teórica à metodologia de ensino proposta com base no apoio visual proporcionado pela tecnologia de fala.

O capítulo 2 focalizará teorias de percepção e produção da fala na aquisição de primeira e segunda línguas, as quais enfatizam a relação entre percepção e produção na aquisição de entoação, de modo a explicitar a origem do sotaque em língua estrangeira.

O capítulo 3 apresentará os modelos tradicionais de descrição de entoação, seguidos da explanação do motivo da escolha do modelo de seqüências tonais e do modelo de transcrição do ToBI (1992) para embasar a análise e transcrições das emissões dos sujeitos que participam do experimento.

O capítulo 4 tratará dos problemas envolvidos no ensino de entoação.

O capítulo 5 concentrar-se-á na pesquisa em si, referindo-se à metodologia adotada para a construção do experimento, incluindo procedimentos de escolha de sujeitos e procedimentos de coleta e de análise estatística, acústica e perceptiva dos dados da pesquisa.

O capítulo 6 demonstrará os resultados das análises acústicas e das avaliações perceptivas feitas por falantes nativos do inglês americano das produções dos sujeitos que participaram do experimento; o tratamento estatístico dos resultados e as avaliações que os alunos fizeram da experiência.

Por fim, as considerações finais focalizarão a síntese dos resultados apontados pelo estudo e os anexos apresentarão os protocolos de avaliação dos sujeitos, os questionários de dados pessoais dos sujeitos do experimento e dos informantes americanos e os conteúdos das gravações e das aulas de treinamento. Encerra-se o volume com as referências bibliográficas.

**CAPÍTULO 1 - TEORIAS DE AQUISIÇÃO, ABORDAGENS  
SOBRE ENSINO DE LÍNGUAS E PRONÚNCIA**

## 1.1 Abordagens teóricas de aquisição de primeira e segunda línguas

As várias abordagens teóricas que tentam explicar a aquisição de linguagem<sup>1</sup> podem ser englobadas em quatro grandes correntes: o Behaviorismo, o Inatismo, o Conexionismo e o Interacionismo/Construtivismo, as quais, por sua vez, refletem-se nos modelos desenvolvidos para tratar da questão da aquisição de segunda língua. Desse modo, seus fundamentos teóricos serão apresentados tendo em vista sua abrangência explicativa do processo de aquisição de entoação em língua estrangeira.

### 1.1.1 Behaviorismo

O Behaviorismo (Skinner, 1957), uma teoria psicológica de aprendizagem muito influente nos anos 40 e 50, principalmente nos Estados Unidos, pressupõe que toda aprendizagem, verbal ou não, ocorre por meio dos mesmos processos. Nessa perspectiva, a criança é uma *tábula rasa*, e o meio é detentor de todo o conhecimento. A aprendizagem de língua é resultante de imitação, prática e reforço, os quais formam uma cadeia entre estímulo e resposta, que leva à formação de um hábito e à aquisição de conhecimento do meio. Os behavioristas tradicionais entendem **imitação** como a repetição de palavra por palavra de toda uma emissão ou parte dela e **prática** como a manipulação repetitiva da forma.

Desse modo, as crianças, quando imitam bem os sons e padrões que ouvem, aproximam-se do padrão adulto e recebem reforço positivo na forma de um elogio

---

<sup>1</sup> Empregamos o termo linguagem para designar a faculdade pela qual os humanos se comunicam uns com os outros por meio de símbolos arbitrários, ao passo que o termo língua, enquanto produto social da faculdade de linguagem, refere-se ao conjunto de convenções necessárias adotadas pelo corpo social para permitir o exercício dessa faculdade nos indivíduos (Saussure, 1972: 17).

ou simplesmente de um gesto afetivo; caso contrário, recebem um reforço negativo. Encorajadas pelo ambiente que as cerca, continuam a imitar e a praticar esses sons e padrões, de modo a formarem hábitos corretos de uso de língua. De acordo com essa abordagem, a qualidade e a quantidade de língua à qual são expostas, assim como a consistência do reforço oferecido pelos outros indivíduos do ambiente, influenciam no sucesso que elas atingem no processo de aquisição de um sistema lingüístico.

Mas sabe-se (Bloom, Hood & Lightbown, 1974) que a porcentagem de imitação na emissão das crianças é bastante variável. A fala de algumas delas chega a ter de 30 a 40% de imitações, enquanto a fala de outras chega a ter menos de 10%. É importante também notar que tais imitações não são aleatórias. Pesquisas (Bloom *et al.*, 1974) demonstram que as crianças imitam palavras e estruturas de sentenças novas até que essas se tornem cristalizadas no sistema da língua quando, então, param de imitá-las e começam a imitar novas palavras e estruturas. Ou seja, quando imitam, a escolha do que imitar parece estar baseada em algo que já começaram a *entender* e não no que está disponível no ambiente no momento.

Ressalte-se, ainda, que elas não imitam tudo o que ouvem, como também não imitam o tempo todo, porque fazem outras coisas. Bloom & Lahey (1978) referem-se a crianças que raramente imitam o interlocutor, mas que respondem e colocam questões, que assumem a comunicação, elaborando sobre as perguntas e afirmações do outro falante<sup>2</sup>.

A abordagem behaviorista de como a linguagem é adquirida tem um apelo

---

<sup>2</sup> Veja-se, também, De Lemos (1986) para a discussão sobre outros processos na fala da criança além da imitação.

intuitivo. Segundo Slobin (1980), quando se pergunta ao leigo de que forma ele acha que as crianças aprendem a falar, ele geralmente responde que é por imitação. Não há dúvida de que esse procedimento oferece uma explicação parcial e razoável de como elas aprendem alguns aspectos regulares e rotineiros da linguagem; sabe-se que, por meio do ato de imitar, elas praticam novas estruturas de maneira semelhante à de um aluno de uma classe de língua estrangeira. No entanto, os conceitos de imitação e prática da teoria behaviorista não explicam como elas aprendem todos os aspectos de sua língua nativa, além de não explicarem a aquisição da língua por crianças que raramente imitam ou que sejam surdas.

Uma teoria empírica de aprendizagem baseada no estímulo e resposta, como é o caso do Behaviorismo clássico, atribui ao organismo um mínimo de estrutura inata, responsável pela capacidade de formar associações entre estímulos e entre estímulos e respostas na base de similaridade e contigüidade. Todavia, o problema com a abordagem behaviorista de aquisição da linguagem refere-se ao fato de que aquilo que a criança adquire a respeito de sua língua não pode ser totalmente explicado apenas com base numa porção de conexões entre estímulo e resposta, pois, como mostrou Chomsky (1965), prevê um conjunto de estruturas inatas, sem as quais não é possível explicar o fato de que as crianças produzem coisas que nunca ouviram antes. A aquisição do sistema lingüístico com base em tal evidência requer uma outra explicação.

Contudo, não se pode negar que o fato de tal abordagem explicar a aquisição de linguagem em termos de imitação, prática, reforço e formação de hábito teve um

grande impacto nas abordagens de ensino e aprendizagem<sup>3</sup> de segunda língua, que passaram a prever a necessidade de os aprendizes receberem *input* lingüístico de falantes e formarem cadeias de associações entre palavras, objetos e acontecimentos, as quais se fortalecem conforme a repetição das experiências. Os alunos são encorajados diante das imitações corretas e corrigidos em seus erros.

Já que o desenvolvimento da linguagem é visto como formação de hábito, assume-se que a pessoa que vai aprender uma segunda língua inicia o processo com os hábitos que formou em sua língua materna e que esses interferem na formação dos hábitos novos requeridos pela segunda língua (Lado, 1964). Daí o Behaviorismo ser freqüentemente vinculado à Hipótese da Análise Contrastiva, desenvolvida pelos lingüistas estruturais da Europa e América do Norte, a qual pressupõe que, onde houver similaridade entre as línguas materna (L1) e estrangeira (L2), as estruturas da língua-alvo serão aprendidas com facilidade, enquanto as diferenças implicarão dificuldade de assimilação.

Já é bastante evidente que a primeira língua dos aprendizes interfere na aquisição da segunda língua. Assim, Jakobson (1967) afirma que o sistema fonológico da L2 é mapeado na língua materna. Daí o porquê de o falante do português brasileiro, ao falar inglês, não produzir oclusivas surdas aspiradas quando em posição inicial de sílaba e inserir /ɪ/ no final de consoantes mudas ou no início de itens lexicais, criando novas sílabas e alterando a estrutura da palavra, como em [ɪspi:kɪ]. Porém nem todas as dificuldades previstas pela Hipótese da Análise

---

<sup>3</sup> Nesse estudo, o termo aprendizagem de uma segunda língua refere-se a habilidades adquiridas por meio do treinamento condicionado pela cadeia de estímulos e respostas somente no contexto da teoria behaviorista. Em qualquer outro contexto, será empregado como sinônimo de aquisição, não implicando uma distinção teórica entre aquisição de primeira língua e aprendizagem de segunda língua ou, ainda, no contexto de segunda língua, uma distinção entre aquisição e aprendizagem, tal como proposta por Krashen (1982) e exposta no item 1.12.

Contrastiva são apresentadas pelos aprendizes. Pelo contrário, comprovou-se, muitas vezes, ser mais fácil adquirir um fone completamente distinto dos que fazem parte do sistema fonológico da L1 do que um outro semelhante (Flege, 1988, 1991, 1995a).

Alem disso, muitos dos erros cometidos não são previstos pelos estudos contrastivos. Esse é o caso de adultos que empregam, na língua estrangeira, estruturas simples, típicas da linguagem das crianças na fase em que estão adquirindo sua primeira língua, e que em nada lembram traduções de L1 para L2. De fato, muitas das sentenças produzidas por aprendizes de segunda língua nos primeiros estágios de aprendizagem são agramaticais até em sua língua materna e, por vezes, apresentam algumas características comuns entre aprendizes de *backgrounds* lingüísticos variados, cujas respectivas primeiras línguas são diferentes entre si e diferentes da língua-alvo. Por outro lado, alunos, algumas vezes, relutam em transferir traços da L1 para a L2, mesmo quando a tradução equivalente seria apropriada.

Tudo isso sugere que a influência da primeira língua pode não ser simplesmente um processo de transferência, mas um caso mais complexo de identificação de pontos de semelhança, o que leva, embora não necessariamente de maneira consciente, a uma reflexão sobre o fato de certos traços da língua materna pertencerem ou não à estrutura da língua-alvo.

Dessa forma, na aquisição de segunda língua, tanto quanto na aquisição de primeira língua, a teoria behaviorista apresenta-se incompleta na explicação que dá para o processo de aprendizagem de línguas. Um *input* pobre e truncado, mesmo

com um esquema de reforço perfeitamente trabalhado, não explica a capacidade da criança, ou do aprendiz de uma segunda língua, de construir uma gramática baseada em tal *input*. Daí psicólogos e lingüistas proporem novas e mais complexas teorias de aquisição, que vão além da imitação e da prática.

### 1.1.2 Inatismo

A posição teórica conhecida como Inatismo foi proposta em contraposição ao Behaviorismo pelo lingüista Noam Chomsky (1959). Ele argumenta contra a tentativa de se tratar qualquer questão relativa à aquisição de linguagem por meio do condicionamento, pois a complexidade formal das línguas naturais vai além das cadeias associativas do condicionamento operante. Além disso, ele demonstra que essa teoria é inadequada para explicar o mecanismo recursivo de tais línguas, já que esse é infinito e não pode ser explicado por Cadeias de Markov ou seqüências lineares de palavras, onde a última determina a probabilidade de ocorrência da seguinte: ou é aprendido ou já se nasce com ele. As cadeias associativas não conseguiriam explicar, por exemplo, como, na emissão, “Chomsky, aquele grande lingüista, esteve no acampamento dos MST no Brasil”, o verbo **esteve** refere-se à **Chomsky**.

Chomsky (1959) argumenta também que a teoria behaviorista falha ao não reconhecer o problema lógico da aquisição de linguagem, que se refere ao fato de que as crianças aprendem mais sobre a estrutura de sua língua do que se poderia esperar, tendo como base os exemplos degradados de língua que lhes são apresentados devido à pobreza do *input*.

Para ele, as crianças são biologicamente programadas para adquirirem linguagem, que se desenvolve da mesma maneira que outras funções biológicas. Todas elas aprendem a andar, desde que tenham alimentação adequada e liberdade de movimentos, sem que seja necessário ensiná-las. Essa ação é aprendida em torno da mesma idade, e o andar é essencialmente o mesmo em todos os seres humanos normais. Segundo o autor, a aquisição de linguagem é muito semelhante. O ambiente tem uma contribuição básica, que é a de prover amostras da língua que as crianças virão a falar, ou seja, o chamado *input* lingüístico, e então, com sua capacidade gramatical inata, elas agirão sobre esse *input*, e o seu aparato biológico fará o resto.

Assim, uma das evidências nas quais se apóia a posição inatista encontra-se no fato de que todas as crianças aprendem sua língua nativa com uma idade em que não se espera que sejam capazes de dar conta de algo tão complexo. Mesmo aquelas completamente surdas aprendem a linguagem dos sinais se expostas a ela desde cedo e o seu progresso na aquisição da língua é semelhante ao das que ouvem. Crianças com limitações nas suas capacidades cognitivas também chegam a desenvolver um sistema lingüístico se expostas a um ambiente em que haja pessoas falando e envolvendo-as na comunicação. No capítulo 5, *Bases biológicas da linguagem*, Slobin (1980) mostra que, apesar de anões nanocefálicos terem um cérebro menos volumoso do que chimpanzés, eles atingem um grau de linguagem correspondente ao de uma criança de 5 anos, ou seja, existe, mesmo em condições precárias, a manifestação da especificidade do cérebro humano para a linguagem.

Outra evidência recorrente na literatura inatista refere-se às condições bastante variadas em que as crianças adquirem as estruturas básicas de sua língua

nativa, as quais incluem não só condições favoráveis, tais como aquelas em que existe atenção e carinho de pais atentos, como também condições desfavoráveis, proporcionadas por ambiente de pais repressores ou ausentes. Elas adquirem diferentes níveis de vocabulário, criatividade, sociabilidade, mas, virtualmente, todas adquirem a língua falada do seu meio ambiente. Esse fato é visto pelo inatismo como uma confirmação da hipótese de que a língua é, de alguma maneira, independente de outros aspectos do desenvolvimento cognitivo e está alocada em uma parte específica do cérebro.

A língua à qual as crianças estão expostas não contém exemplos - ou muitos exemplos - de todas as regras lingüísticas e padrões que elas virão a demonstrar. Mas tanto as regras quanto os padrões são adquiridos sem que haja alguém consistentemente indicando-lhes erros e acertos. Sendo assim, os inatistas concluem que a aquisição das regras gramaticais é guiada por princípios inatos da gramática universal aplicáveis a todas as línguas e que as crianças adquirem sua língua específica pela exposição a um número limitado de exemplos. Em outras palavras, línguas diferentes têm diferentes regras e as crianças, ouvindo alguns exemplos, aprendem quais são os parâmetros da língua que estão adquirindo.

Chomsky (1965) alerta que o *input*, além de pobre, é repleto de orações truncadas, incompletas, com posições sintáticas invertidas em relação à estrutura profunda. Além disso, as crianças que não são sistematicamente corrigidas ou instruídas a respeito de suas produções, quando o são, freqüentemente ignoram a correção e continuam a se expressar de sua maneira, mostrando-se alheias à língua que as circunda. Elas fazem, por exemplo, generalizações indevidas com as terminações verbais regulares para os verbos irregulares, mesmo depois de já

terem usado corretamente as formas irregulares e resistem a correções prematuras. Nesse sentido, Albano (1990: 10) desafia o leitor: “Tente ensinar FIZ para uma criança que diz FAZI”.

De acordo com Chomsky (1965), a mente das crianças não é uma lousa em branco, que, como pensavam os behavioristas, será preenchida pelo conhecimento advindo do meio ambiente, em que a imitação da língua é o principal recurso de aprendizagem. Ele afirma que elas já nascem com habilidades especiais, às quais se refere, inicialmente, como *language acquisition device (LAD)*, para descobrirem por si próprias as regras subjacentes ao sistema lingüístico.

Esse mecanismo, alocado em algum ponto do cérebro, seria como um extrator de elementos para preencher as categorias gramaticais inatas e extrair as regras da língua a ser falada com base no *input* lingüístico. O *LAD* conteria todos e somente os princípios universais da linguagem humana, que impediriam as crianças de trilharem falsos caminhos na tentativa de descobrirem as regras de sua língua. Para que ele fosse ativado, além da maturação do cérebro, seria necessário somente que elas tivessem acesso a exemplos de língua natural. Então, seriam capazes de descobrir a estrutura da língua a ser adquirida, por meio do confronto entre o conhecimento inato das relações gramaticais básicas e as estruturas da língua específica do ambiente. Apesar de não ter sido localizado no cérebro humano, Pinker (1994) defende a existência do *LAD* e expressa sua esperança de que a genética encontre um gene que funcione tal como o mecanismo proposto por Chomsky.

Em seus trabalhos mais recentes, Chomsky e seus seguidores não usam

mais o termo *LAD*, mas referem-se à capacidade inata das crianças como Gramática Universal (GU), que consiste em um conjunto de princípios comuns a todas as línguas. Assim, pré-equipadas com a GU, as crianças têm de aprender as maneiras como seu sistema lingüístico faz uso desses princípios e como ocorrem as variações desses parâmetros na língua específica que ouvem. (Chomsky, 1981; Cook, 1988; White, 1989).

Ao contrário do pressuposto behaviorista, diferenças ambientais estão associadas apenas à variação na velocidade da aquisição, não chegando a interferir na competência lingüística, pois essa é a mesma entre todos os falantes. Ao adquirirem o sistema intrincado e complexo que forma a língua, as crianças, cujas habilidades cognitivas ainda são bastante limitadas em muitos outros aspectos, dão conta de algo que aprendizes adultos de uma segunda língua invejam. Elas apresentam, em torno do segundo ano de vida ou entre o segundo e terceiro ano (Sacks, 2002), um domínio da estrutura da língua falada ao seu redor, o qual as faz ser vistas pelo interlocutor adulto como um falante em muitos aspectos competente (Gama, 1989; Gonçalves, 1989).

Tal fato é visto como apoio para a hipótese de que a linguagem é, de alguma maneira, uma capacidade separada de outros aspectos do desenvolvimento cognitivo e que deve estar alocada numa parte diferente e específica do cérebro. Daí o termo modular (Fodor, 1983) ser usado para representar a noção de que esse órgão humano tem diferentes módulos para processar diferentes tipos de informação. Para explicar a autonomia da gramática, Fodor opta por incluir a linguagem entre os módulos ou sistemas de entrada inatos, juntamente com os

processos perceptivos. Ele defende a independência desses processos perceptivos com relação a uma cognição geral, apontando a ineficácia de uma percepção dirigida por processos como raciocínio e inferências, os quais poderiam completar as lacunas da nossa percepção, mas poderiam, também, falsear os dados sensoriais em favor de nossas expectativas. Outro fator em prol dessa independência seria o fato de não haver controle das sensações.

As bases biológicas da posição inatista encontram-se no trabalho do biólogo Eric Lenneberg (1967), que também compara o aprender a falar com o aprender a andar (Slobin, 1980; Pinker, 2002). Ele observa que a habilidade para desenvolver conhecimento e comportamentos normais em ambientes variados não continua indefinidamente e que as crianças que nunca aprenderam uma língua por causa de surdez ou isolamento extremo (Sacks, 2002) não são capazes de adquiri-la se essas privações permanecerem por um longo tempo. Além disso, ele argumenta que o mecanismo inato de aquisição de linguagem, como outras funções biológicas, funciona com sucesso somente quando é estimulado no tempo certo, ou seja, dentro do chamado **período crítico**. Essa noção de que existe um período de tempo específico e limitado para aquisição de linguagem é referida como a Hipótese do Período Crítico (Lenneberg, 1967).

Até agora a melhor evidência para tal hipótese são os relatos sobre crianças que cresceram sem serem expostas a uma língua, como os casos de Victor (Itard, 1962), encontrado com cerca de 12 anos vagando nu pelas florestas de Aveyron, na França, e de Genie (Rumer, 1993), de 13 anos, criada isoladamente, presa em um quarto escuro, sem contato com linguagem humana. Apesar do intenso trabalho realizado com essas crianças, elas não foram capazes de adquirir a

língua sem problemas. Victor desenvolveu sociabilidade, memória, julgamento, mas permaneceu não receptivo a qualquer outro som que não fossem aqueles que fizessem sentido para ele, como o som da chuva e sons de animais. Ele aprendeu a falar a palavra *lait* (leite) e a expressão *O Dieu* (Ó Deus), mas nunca as empregou para se comunicar. Genie, após cinco anos de exposição ao inglês, ainda apresentava inconsistência no uso das formas gramaticais, usava exageradamente expressões formulaicas e não utilizava algumas formas gramaticais específicas, sempre presentes no desenvolvimento gramatical normal de uma criança de cinco anos.

A investigação das bases neurofisiológicas da aquisição de linguagem calcada na Hipótese do Período Crítico confere importância significativa à chamada idade de aprendizagem de uma língua, uma vez que faz uma distinção crucial entre a aquisição da língua nativa e a aquisição da língua não-nativa, isto é, da segunda língua. Tal teoria preconiza que a idade jovem é ótima para se adquirir linguagem porque o cérebro nessa fase ainda é capaz de adaptar suas estruturas aos requisitos de uma língua específica. Falantes nativos são definidos como aqueles que adquiriram sua língua nesse período. Mais tarde, o cérebro perde essa plasticidade e a aquisição de linguagem é menos eficiente.

A questão crucial que se coloca para essa hipótese diz respeito exatamente à identificação das mudanças ocorridas durante a maturação do cérebro, que seriam responsáveis por essa perda na capacidade de aquisição de linguagem. A esse respeito, há vários relatos diferentes na literatura. Penfield & Roberts (1959), por exemplo, assumem como causa provável desse fenômeno mudanças na plasticidade do cérebro ocorridas no início da puberdade.

Lenneberg (1967) considerou que a completude da lateralização cerebral seria o fator decisivo, o que levou Krashen (1973) a argumentar, em um debate sobre a idade em que a lateralização hemisférica ocorre, que, com base na lateralização, a aquisição de língua nativa seria impossível já na idade de 5 anos.

A Hipótese do Período Crítico está sujeita a muitas refutações, no entanto tem recebido suporte de pesquisas recentes, como a que foi publicada na revista *Nature Neuroscience* e relatada em artigo da *Gazeta Mercantil* do dia 26 de julho de 2002, intitulado *O Despertar da Linguagem*. Usando imagens obtidas por ressonância magnética, neurocientistas dos Estados Unidos e da Inglaterra constataram a presença de padrões de atividade cerebral diferentes em crianças submetidas a uma segunda língua antes da puberdade.

A posição inatista tem sido bastante persuasiva ao apontar a dificuldade implícita no processo de aquisição do conhecimento complexo manifestado na fala adulta, todavia atribui muita ênfase ao estado final, isto é, à competência do falante nativo adulto, e não dá atenção suficiente aos aspectos do desenvolvimento da aquisição da linguagem.

### *Gramática Universal*

Chomsky não faz afirmações sobre as implicações da Gramática Universal na aquisição de segunda língua, contudo alguns lingüistas argumentam que a GU oferece as melhores perspectivas para a compreensão da L2 (Gass, 1989). Outros, por sua vez, afirmam que, apesar de a GU apresentar uma boa explicação para aquisição de língua materna, ela não dá conta da aquisição de língua estrangeira

em falantes que já ultrapassaram o período crítico da aquisição de linguagem, ainda que a Hipótese do Período Crítico explique, em parte, o sotaque estrangeiro em falantes bilíngües (Rocca, 2003).

Mesmo entre os que acreditam que a Gramática Universal tem um papel na explicação da aquisição de segunda língua, não existe consenso em como a GU opera no desenvolvimento desse processo (White, 1989). Alguns autores asseveram que, embora os alunos comecem a aprender a língua depois do período crítico e apesar de muitos fracassarem ao tentarem dominá-la completamente, ainda há um problema lógico a ser resolvido: aqueles que adquirem uma segunda língua sabem mais a respeito dela do que poderiam ter aprendido se dependessem exclusivamente do *input* recebido. Diante disso, tais autores concluem que a GU deve estar disponível para os aprendizes de L2 da mesma maneira que está disponível para L1, mas Gama-Rossi<sup>4</sup> questiona se não é justamente o fato de já dominarem um sistema lingüístico (e o fato de, geralmente, serem leitores) que os ajuda na nova empreitada. Outros estudiosos defendem que a GU pode estar presente na aquisição da segunda língua, mas que sua natureza foi alterada pela aquisição de pelo menos uma outra língua.

Pesquisadores que trabalham com a GU também diferem em suas hipóteses de como a instrução formal ou a correção de erro afeta o conhecimento do aprendiz de segunda língua. Alguns (Schwartz, 1993; Krashen, 1982) declaram que, como as crianças, os adultos não têm necessidade e nem se beneficiam da correção de erro ou de informação metalingüística. Concluem que esses mecanismos alteram

---

<sup>4</sup> Questionamento colocado por Gama-Rossi em exame de qualificação.

apenas a aparência superficial da performance lingüística e não afetam realmente o conhecimento sistemático subjacente da nova língua. Isso explicaria o caso de Helena, que, ao entender a forma das interrogativas *Yes/No* do inglês, que lhe causava problemas na quinta série do segundo grau, de repente, diz para a mãe: “descobri o segredo” e passa fazer os exercícios escritos corretamente, de maneira consciente. Todavia, ter descoberto o parâmetro das interrogativas da L2 não garante que ele passe automaticamente a fazer parte da competência lingüística de Helena, tal como acontece com a L1.

Por outro lado, aqueles que acreditam ter sido a GU afetada pela aquisição anterior de uma primeira língua sugerem que os alunos de L2 precisam receber informações explícitas a respeito do que não é gramatical na língua estrangeira, sob pena de virem a pressupor que algumas estruturas da L1 têm equivalentes na L2, quando, de fato, não têm. Daí ser difícil pensar na GU operando na aquisição de segunda língua, pois os estudantes tentam usar as mesmas categorias gramaticais e derivacionais para mapeá-la e, se tornar consciente as diferenças não é suficiente para garantir a competência, ignorá-las também não contribui para a aquisição do sistema lingüístico da L2.

Os estudos sobre a aquisição de segunda língua de uma perspectiva da GU estão interessados na competência lingüística de alunos em estágios avançados, mais do que nas estruturas simples dos estágios iniciais. Eles apontam que uma variedade de diferentes teorias explica a performance dos primeiros estágios, mas que somente uma teoria como a GU é capaz de explicar o conhecimento de estruturas sintáticas complexas. A questão que se colocam é a de saber se a competência subjacente à performance de aprendizes de L2 assemelha-se àquela

subjacente à performance de falantes nativos. Por isso, suas investigações sempre envolvem comparações de julgamento de gramaticalidade feito pelos dois grupos e não se baseiam em observações sobre a fala real. Ao procederem assim, esperam descobrir o que os aprendizes realmente sabem a respeito da L2, qual a sua competência diante dela, sem dar atenção à performance.

### *O Modelo Monitor de Krashen*

Uma teoria nativista de aquisição de segunda língua com grande influência na prática de ensino de línguas estrangeiras foi proposta por Stephen Krashen (1982), denominada Modelo Monitor, a qual se fundamenta em cinco hipóteses: hipótese da aquisição/aprendizagem, hipótese do monitor, hipótese da ordem natural, hipótese do *input* e hipótese do filtro afetivo.

#### Hipótese da aquisição/aprendizagem

Essa hipótese pressupõe que há dois modos para aprendizes adultos de língua estrangeira desenvolverem seus conhecimentos: a aquisição e a aprendizagem. A **aquisição** ocorre conforme são expostos a exemplos de L2 e passam a compreendê-los. Isso acontece de uma maneira muito semelhante à de como a criança adquire a primeira língua, ou seja, sem atenção consciente à sua forma. Por outro lado, a **aprendizagem** - ocorre por um processo consciente de estudo e atenção à forma e à regra do sistema lingüístico.

Para Krashen, a aquisição é muito mais importante porque somente a língua adquirida está disponível para a comunicação natural e fluente. Ele afirma, ainda,

que a aprendizagem não se transforma em aquisição, o que é evidenciado pelo fato de que muitos falantes são fluentes sem nunca terem aprendido regras, enquanto outros sabem regras, mas são incapazes de aplicá-las quando focalizam sua atenção no que **querem** dizer e não em **como** vão dizê-lo.

Essa distinção proposta por Krashen parece refletir duas situações diversas de exposição à língua estrangeira, uma formal, em ambiente de ensino-aprendizagem, e a outra informal, em contatos sociais com falantes da L2, como, por exemplo, em viagens ou mudança de país. Ela pode, também, relacionar-se com duas abordagens de ensino, uma com ênfase na conscientização de regras e a outra com enfoque em atividades comunicativas. Entretanto, não fica claro como se dá a almejada aquisição. Se aprendizes adultos adquirem uma língua estrangeira da mesma maneira que a criança adquire a L1, deve-se admitir que a GU funciona depois do período crítico de aquisição de línguas.

#### Hipótese do monitor

De acordo com essa hipótese, o sistema adquirido atua para iniciar as emissões dos falantes e é responsável pela fluência e julgamentos intuitivos sobre correção, ao passo que o sistema aprendido atua somente como um editor ou monitor, operando pequenas mudanças e melhorando o que o sistema adquirido produziu. Os aprendizes usam o monitor somente quando estão mais preocupados com a forma correta daquilo que têm a dizer, do que com o seu conteúdo; quando têm tempo suficiente para procurar na memória as regras relevantes e quando realmente sabem essas regras. Portanto, o ato de escrever faz mais uso do monitor do que o ato de falar, porque naquele existe mais tempo para focalizar a atenção. Saber as regras somente ajuda o falante a complementar o que ele já adquiriu, daí o

fato de que o ensino de línguas deve ter como foco a criação de condições para aquisição e não para a aprendizagem.

Contudo, não há evidências do uso do monitor, por ser impossível diferenciar o que é produzido pelo sistema adquirido e o que é resultado do uso do monitor. Por conseguinte, a afirmação de Krashen de que a emissão produzida rápida e espontaneamente deve ter sido adquirida e não aprendida leva-nos a uma definição circular. Pesquisas desenvolvidas na área de aquisição de pronúncia de língua estrangeira (Rocca & Marcelino, 1997; Rocca, 1998) têm demonstrado a importância de tornar o aluno consciente de sua própria língua e das diferenças entre ela e a língua-alvo, diante das limitações impostas pelo sistema fonológico da L1 à percepção do sistema fonológico da L2.

#### Hipótese da ordem natural

Segundo Krashen, aprendizes de língua estrangeira parecem adquirir os traços da língua-alvo em seqüências previsíveis que, muitas vezes, contrariam a intuição. As regras mais fáceis de serem formuladas e, portanto, mais fáceis de serem aprendidas não são necessariamente as primeiras a serem adquiridas. Estudos sobre morfemas gramaticais demonstram que essa ordem natural é também independente da ordem em que as regras são aprendidas em classes de língua. Assim, a regra morfológica do inglês de que a terceira pessoa do singular do presente do indicativo é formada pelo acréscimo do morfema [-s] à forma infinitiva do verbo é simples, de fácil formulação e uma das primeiras a ser introduzida nas aulas de língua inglesa. No entanto, é uma regra difícil de ser adquirida no sentido proposto por Krashen.

Apesar das críticas que esses estudos receberam, pesquisas posteriores confirmaram que os aprendizes passam por seqüências ou estágios em seu desenvolvimento. A mudança de um estágio para o outro pode levá-los a empregarem de maneira incorreta formas da L2 que, até então, usavam corretamente, mas, mesmo assim, essas generalizações indevidas indicam progresso no processo de aquisição. Então, da mesma maneira que as crianças, aprendizes adultos de L2 aprendem formas irregulares do passado de certos verbos antes de aprenderem a forma dos verbos regulares. No caso do inglês, o aluno que passa a trocar *bought* por *buyed* pode saber mais sobre a gramática do inglês do aquele que não apresenta a troca.

#### Hipótese do *input*

Tal hipótese destaca que a única maneira de se aprender uma língua é pela exposição ao *input* “compreensível”. Conseqüentemente, a compreensão e a aquisição ocorrem somente se o *input* contiver formas e estruturas um pouco além do nível de competência do aprendiz. Porém, não há evidências empíricas que comprovem essa afirmação.

Apesar de reconhecer que algumas pessoas expostas ao *input* compreensível extensivo não alcançam níveis altos de proficiência em L2, Krashen mantém a convicção de que esse, tal como definido por ele, é a origem da aquisição e recorre à hipótese do filtro afetivo para explicar o fracasso de alguns aprendizes.

## Hipótese do filtro afetivo

O filtro afetivo é uma barreira imaginária que impede os aprendizes de adquirirem língua com base no *input* disponível e refere-se a estados emocionais como motivação, necessidade, atitude, entre outros. É ele que limita o que é percebido e o que é adquirido, pois bloqueia a aprendizagem, filtrando o *input* e tornando-o inacessível para a aquisição.

Essa hipótese pode ter implicações imediatas para práticas em salas de aula, pois tenta explicar a razão de alguns alunos terem problemas na **aprendizagem**. A dificuldade, porém, está em provar que fatores afetivos causam as diferenças na **aquisição**, tendo em vista a distinção feita por Krashen, entre aprendizagem e aquisição.

A teoria de Krashen (1982) teve grande influência na abordagem comunicativa do ensino de línguas, principalmente na América do Norte, apesar de receber sérias críticas por ser incapaz de propor hipóteses que pudessem ser comprovadas empiricamente. Mesmo assim, o ensino comunicativo de línguas, que se baseia no uso da língua em interações significativas e no cumprimento de tarefas, em vez de colocar o foco do ensino na aprendizagem de regras, foi largamente aceito por professores e alunos.

Por outro lado, algumas pesquisas (Schmidt,1990) centradas em salas de aula indicam que chamar a atenção para aspectos estruturais da língua pode ser mais importante do que o admitido por Krashen, já que esse tipo de instrução

pode ser incorporado ao ensino, enriquecendo a abordagem comunicativa.

O teórico Richard Schmidt (1990) enfatiza a importância de se focalizar a atenção na aquisição de L2, argumentando que tudo o que o falante vem a conhecer a respeito da língua foi primeiramente notado de maneira consciente. Essa posição, que contrasta com as convicções de Krashen (1982), está de acordo com outros psicólogos cognitivistas, que não acreditam na diferença entre aquisição e aprendizagem.

Assim, os cognitivistas que trabalham com o modelo de processamento de informação tendem a considerar a aquisição de língua estrangeira como um processo de construção de sistemas de conhecimento, acionados automaticamente na fala e na compreensão (McLaughlin 1987). Primeiramente os aprendizes focalizam a atenção no aspecto da língua que estão tentando aprender ou produzir, considerando que existe um limite para a quantidade de informação a que uma pessoa é capaz de dar atenção. Desse modo, aqueles que se encontram nos estágios iniciais de L2 são capazes de se ater às palavras principais da mensagem e podem não ser capazes de notar os morfemas gramaticais ligados a essas palavras. Mas aos poucos a experiência e a prática os tornam aptos para usar o seu conhecimento de maneira tão rápida e automática que nem chegam a ter consciência disso. Tal fato deixa-os livres para focalizar outros aspectos da língua que vão, gradativamente, tornando-se automáticos.

Essa performance automática pode originar-se da aprendizagem intencional, mas nem sempre é o caso. Qualquer coisa que faça uso do espaço

de processamento mental, mesmo que de maneira inconsciente, é uma possível fonte de geração de informação ou de habilidades que podem tornar-se automáticas se houver prática suficiente. Note-se que, nesse contexto, a prática não é considerada uma atividade mecânica, pois envolve esforço por parte do aluno.

Além do desenvolvimento do automatismo pela prática, é possível que ocorram mudanças nas habilidades e no conhecimento devidas a uma reestruturação. Essa noção faz-se necessária diante da observação de que algumas coisas de que se tem conhecimento e as quais são usadas automaticamente não podem ser explicadas em termos de uma construção gradual do automatismo pela prática. Elas parecem estar baseadas na interação do conhecimento que já se possui ou na aquisição de um novo conhecimento que, sem prática extensiva, de alguma maneira se enquadra em um sistema já existente, transformando-o ou reestruturando-o.

Isso parece explicar não só os súbitos progressos dos alunos, mas também as regressões aparentes que ocorrem quando um aspecto sistemático de sua língua incorpora dados em excesso ou dados errados. Portanto, quando aprendizes de inglês finalmente dominam o uso da forma regular do plural de substantivos, o plural de substantivos irregulares, que eles já dominavam, pode ser afetado e dar origem a formas como *mans* e *womans*.

### 1.1.3 Conexionismo

Apesar das várias teorias divergirem sobre a maneira como os humanos processam a linguagem, existe um consenso em torno da convicção de que o cérebro é o órgão da mente. A visão conexionista da aquisição de primeira língua vem enfatizar o poder computacional dinâmico do cérebro de manipular símbolos para explicar o processo cognitivo.

Nos últimos anos, muito se tem feito na tentativa de explicar vários aspectos da fala e da linguagem em termos de sistemas dinâmicos. Elman (1995) tenta mostrar que a abordagem tradicional, que leva em conta a atividade mental com base em representações discretas, estáticas, passivas e livres de contextos, contrasta fortemente com as propriedades computacionais do cérebro. Ele considera que redes neurais com características dinâmicas têm um número de propriedades que capturam importantes aspectos da linguagem, incluindo sua natureza subjacente.

O autor, então, propõe que o processamento de linguagem seja considerado como um sistema dinâmico que prevê representações muito sensíveis ao contexto, probabilísticas e continuamente variáveis, no qual os objetos de representação mental são pensados como trajetórias através do espaço mental e não como objetos construídos.

Nesse contexto, o léxico consiste em regiões de um espaço de estado (*state space*) e a gramática consiste na dinâmica, que, por meio da atração ou repulsão, restringe os movimentos nesse espaço. Embora o modelo de redes difira da visão

tradicional de processadores de linguagem quanto à maneira de representar a informação lexical e gramatical, ele é capaz de induzir a estrutura da categoria lexical baseado em estatísticas de regularidades do uso e, de certo modo, é também capaz de representar a estrutura de constituintes.

Para os conexionistas, o sistema nervoso é formado por um conjunto extremamente complexo de neurônios, no qual a comunicação é realizada por impulsos. Quando um impulso é recebido, o neurônio o processa e, passado um limiar de ativação, dispara um segundo impulso produtor de uma substância neurotransmissora que flui do corpo celular para o axônio (o qual, por sua vez, pode ou não estar conectado a um dendrito de outra célula). O neurônio transmissor do pulso pode controlar a frequência de pulsos, aumentando ou diminuindo a polaridade na membrana pós-sináptica. Conseqüentemente, os neurônios têm um papel essencial na determinação do funcionamento, comportamento e raciocínio dos humanos.

Por meio de simulações computacionais, a teoria conexionista demonstra que um programa de computador relativamente simples, quando comparado a um cérebro humano, pode aprender certas coisas se exposto a elas com frequência suficiente. Essas redes neurais artificiais podem até mesmo fazer generalizações além do que foi exposto e cometer os mesmos tipos de erros criativos das crianças. Rumelhart & McClelland (1986) simularam, numa rede neural, o emprego de regularizações indevidas na aprendizagem das formas de passado dos verbos irregulares do inglês, e a rede, ao aprendê-las, generalizou as formas regulares do passado para os itens que são irregulares. Trabalhos como esse mostram que a aprendizagem de redes neurais pode reproduzir aspectos da aprendizagem

morfêmica das crianças.

O campo de redes neurais artificiais tem se mostrado extremamente vasto e promissor, tornando-se um assunto interdisciplinar bastante atrativo para profissionais de domínios distintos, tais como psicólogos, lingüistas, neurofisiologistas, engenheiros, cientistas cognitivos e cientistas da computação, que buscam, cada um em sua área, novos caminhos por intermédio da computação neural. Essas redes neurais artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento pela experiência. Uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento, já o cérebro de um mamífero pode ter muitos bilhões de neurônios.

Ao contrário das redes neurais artificiais, as naturais não transmitem sinais negativos e sua ativação é medida pela freqüência com que emitem pulsos contínuos e positivos. Elas não são uniformes, como as artificiais, a não ser apenas em alguns pontos do organismo. Além disso, seus pulsos não são sincrônicos ou assincrônicos, por não serem contínuos como nas redes artificiais.

O debate entre os inatistas e os conexionistas tem sido intenso nos últimos anos. Estes, diferentemente daqueles, não vêem necessidade de postular a existência de um módulo neurológico que tenha por finalidade apenas a aquisição de linguagem (Elman, 1995), a qual deve ser explicada em termos de aprendizagem em geral. Como muitos psicólogos cognitivistas, os conexionistas atribuem maior importância ao papel do *input* do que ao conhecimento inato, argumentando que, o que é inato é simplesmente a habilidade de aprender e não uma estrutura lingüística

específica. Eles também argumentam que tudo o que a criança precisa saber está disponível na língua à qual está exposta. Assim, enquanto os inatistas consideram o *input* como o gatilho que ativa o conhecimento inato, os conexionistas o vêem como a principal origem do conhecimento lingüístico.

No caso de aquisição de segunda língua, os alunos constroem gradualmente seu conhecimento por intermédio de sua exposição a milhares de exemplos de amostras de língua. Depois de ouvirem muitas vezes essas amostras em contextos lingüísticos e situações específicas, eles desenvolvem conexões neurológicas ou mentais cada vez mais fortes entre esses elementos. Eventualmente, a presença de um elemento lingüístico ou situacional ativará outros em sua mente e tais conexões podem ser fortes, porque os elementos ocorreram conjuntamente com grande freqüência, ou podem ser relativamente fracas, porque houve poucas oportunidades para experimentá-las juntas.

A proposta conexionista, contudo, ainda está em construção. Dessa forma, algumas questões ficam sem respostas e há dificuldade em capturar as relações semânticas naturais com o mundo, porque as redes neurais desenvolvidas para processamento de língua são desvinculadas da atividade humana. Conseqüentemente, não há espaço para atividade criativa, para fala espontânea ou para linguagem do pensamento. Como colocado por Elman (1995), são redes que não pensam, por isso os lingüistas que trabalham com a Gramática Universal desafiam os conexionistas a provarem que a sua teoria pode dar conta da criatividade da língua.

Apesar das limitações apontadas, que, de alguma maneira, aplicam-se aos

outros modelos de aquisição de língua, a visão conexionista do comportamento lingüístico derivado de um sistema dinâmico parece oferecer mais oportunidades para solucionar suas próprias limitações do que os modelos não dinâmicos.

#### 1.1.4 Interacionismo/ Construtivismo

Os estudos interacionistas, centrados na criança, surgiram como uma crítica à proposta chomskiana de aquisição da linguagem por seu reducionismo ao biológico. Entretanto, os interacionistas representam uma gama enorme de teorias que, embora tenham em comum o pressuposto de que a linguagem desenvolve-se como resultado da interação, propõem interações concebidas de maneiras muito diversas, incluindo interação com o outro, com o meio, com a própria língua. Nesse universo, destacam-se dois autores, Piaget (1923) e Vygotsky (1934), cujas idéias foram de grande impacto. No âmbito nacional, destaca-se o trabalho de Albano (1990), por sua contribuição em relação ao papel fundamental do elemento fonológico na emergência e desenvolvimento de linguagem.

A posição interacionista, articulada no começo do século passado pelo psicólogo suíço Jean Piaget (1923), focaliza o papel da interação entre o ambiente lingüístico e as capacidades inatas da criança no desenvolvimento da linguagem. Observando crianças brincando e interagindo com adultos, ele traça o desenvolvimento cognitivo, concentrando-se na formação de conceitos tais como permanência de objeto na ausência da visão, estabilidade da quantidade independentemente de mudanças na sua aparência e inferências lógicas. Piaget conclui que tal desenvolvimento é construído na interação entre as crianças e as coisas que observam, tocam ou manipulam. Calcado nessa observação, infere que o desenvolvimento cognitivo

determina parcialmente o uso que as crianças fazem da língua, como, por exemplo, o emprego dos termos **maior** ou **mais**, que são dependentes da forma como elas compreendem os conceitos que eles representam.

Piaget, além de admitir estruturas inatas, vê a interação como algo mais amplo do que a díade mãe-criança. A interação, a seu ver, dar-se-ia com os objetos do mundo, entre os quais se encontra o adulto. Todavia, diferentemente dos inatistas, ele não considera que a linguagem seja baseada em módulo específico, independente do desenvolvimento cognitivo. Para ele, linguagem é um dos numerosos sistemas simbólicos desenvolvidos na infância, que pode ser usado para representar o conhecimento que as crianças adquiriram pela interação física com os elementos que as rodeiam. Além disso, ele considera a aquisição de linguagem como sendo semelhante e influenciada pela aquisição de outros tipos de habilidade e conhecimento e não como algo largamente independente da experiência das crianças e do seu desenvolvimento cognitivo.

O autor pressupõe que o sistema simbólico da linguagem somente pode ser adquirido após o final do estágio sensório-motor, ou seja, ele adota a posição de que o desenvolvimento cognitivo antecede em alguns meses o desenvolvimento lingüístico, enfatizando a tão criticada divisão entre um momento pré-lingüístico e um lingüístico que deixa sem explicação as vocalizações da criança, que antecedem ao término do período sensório-motor, o qual ocorre aproximadamente aos 18 meses.

Uma outra perspectiva do papel da interação na aquisição de linguagem abre-se com a teoria sociocultural do processamento mental humano do psicólogo

Lev Vygotsky (1934), a qual pressupõe que todo o desenvolvimento cognitivo, incluindo o da linguagem, é resultado de interações sociais entre indivíduos. Desse modo, enquanto Piaget postula que a linguagem desenvolve-se como sistema simbólico para expressar conhecimento adquirido por meio da interação com o mundo físico, Vygotsky afirma que o pensamento é essencialmente fala internalizada e que a fala emerge na interação social.

Vygotsky, que trabalhou na União Soviética nas décadas de 20 e 30, assevera que, com o suporte de um ambiente interativo, as crianças adquirem conhecimento e têm uma atuação de nível mais alto do que conseguiriam desenvolver de forma independente. Ele denomina zona de desenvolvimento proximal aquilo que elas são capazes de fazer na interação com outras e que seriam incapazes de atingir sozinhas. Ele observou a importância das interações que crianças têm com adultos e com outras crianças, e conclui estar aí, as origens da linguagem e do pensamento.

De fato, pesquisas (Sachs, Bard & Johnson, 1981) demonstram que crianças expostas à língua num ambiente em que não existe interação de um para um não desenvolvem linguagem normalmente, como é o caso de Jim. Filho de pais surdos, Jim, uma criança normal, apresentava desempenho lingüístico abaixo da média porque só tinha contato com a língua oral quando assistia à televisão. Um outro exemplo da importância da interação na aquisição da língua é o atraso geralmente encontrado em crianças institucionalizadas ou mesmo naquelas que passam o dia em creches, onde há um número pequeno de adultos em proporção ao número de crianças.

Na aquisição de segunda língua, conforme Jim Lantolf (1994), a zona de desenvolvimento proximal pode ser observada em estratégias de fala usadas por professores e alunos mais adiantados para criar condições de apoio para o aluno de L2 compreender e produzir a língua. Lantolf afirma que os alunos progredem e atingem níveis mais altos de conhecimento lingüístico quando colaboram e interagem com falantes de L2 que têm mais conhecimento do que eles e que se utilizam de estratégias como repetição, simplificação e modelamento para propiciar a co-construção de experiências de aprendizagem.

A diferença entre a perspectiva de Vygotsky e de outros pesquisadores que vêem na interação um papel importante na aquisição de L2 é que os teóricos socioculturais assumem que a aquisição da linguagem realmente acontece nas interações do aluno com o interlocutor, enquanto outros modelos interacionistas assumem que é a modificação do *input* que fornece o material lingüístico básico a ser processado internamente e de maneira invisível. Estudos (Long,1983) indicam que ajustes conversacionais ocorridos na interação levam a uma melhor compreensão do que a simplificação ou modificação lingüística que é planejada antecipadamente. Essa interação modificada incluiria taxa de elocução de fala mais lenta, gestos, checagem de compreensão, repetição ou paráfrase, pedidos de esclarecimento e também dicas contextuais adicionais.

Albano (1990) propõe quatro condições para o desenvolvimento da linguagem:

- a presença de um interesse subjetivo por ela;
- a existência de um sistema sensório-motor dotado de plasticidade;

- a inserção num meio onde a linguagem faça parte de rotinas significativas;
- a presença de uma língua minimamente auto-referenciada, sinalizando a própria organização.

Ela concorda com a idéia de Piaget de que há graus de abstração vários no conhecimento lingüístico de maneira semelhante ao que ocorre com o conhecimento geral. Desse modo, a abstração progressiva com reestruturações sucessivas no interior da própria linguagem explicaria o porquê de um conhecimento presente num dado momento desaparecer e reaparecer reestruturado depois. No entanto, ela não aceita que tal fato implique a redução da linguagem ao conhecimento geral. Em outras palavras, segundo essa autora, Piaget deu pouca atenção à linguagem na sua teoria geral do desenvolvimento, porque o seu construtivismo buscava um meio de reduzir a linguagem a uma cognição mais geral, ignorando a independência considerável que existe entre a capacidade de aprender uma língua e a de realizar outras tarefas cognitivas.

Outra divergência com Piaget consiste no fato de admitir pré-disposições inatas já bastante específicas e diferenciadas, ao passo que Piaget propõe que o desenvolvimento procede da indiferenciação para a diferenciação através da atividade estruturante de um sujeito. Albano argumenta que, se a criança é responsável por estruturas num contato com o meio, a plasticidade do seu potencial torna a sua tarefa mais fácil se for pressuposto que o recém-nascido pode realizar muitas discriminações ou mesmo categorizações, isto é, que o recém-nascido tem uma sensório motricidade suficientemente pré-estruturada para sustentar processos de abstração e reorganização progressivas.

A criança piagetiana, por outro lado, constrói o saber em interação com o meio por um processo conhecido como auto-regulação, portanto o poder do sujeito é quase absoluto. A autora diverge dessa proposta porque uma subjetividade como essa teria o efeito de uniformizar todos os processos de aquisição de conhecimento e de prever um centro de atividades cognitivas. Para ela, se a sensório-motricidade já possui um subsistema especializado desde o nascimento, o conhecimento se auto-organiza com base nas interações entre esses subsistemas e essa seria a faceta da auto-organização, que diz respeito à estrutura do sujeito.

No que concerne ao inatismo, Albano coloca, como uma evidência contrária aos pressupostos teóricos, as invenções idiossincráticas das crianças que apresentam regularidades inventadas sem nenhuma função comunicativa direta, como as que denominou *o latim do Nico* e que se referem às construções do tipo *midato* (me dá), *tiato* (tira), *sentato* (senta), *tiradas*, provavelmente, de comandos que Nico ouvia da mãe: *deitado*, *sentado*. Além do mais, se certos usos remetem à capacidade de análise e abstração, outros evidenciam aspectos imitativos, como o caso de crianças que repetem “Radio Globoooo” com perfeição, apesar de não serem capazes de emitir o /r/ em outros contextos.

A proposta inatista dificulta, também, a explicação para o fato de certos conhecimentos serem adquiridos de maneira lenta, enquanto outros aparecem e/ou desaparecem de maneira súbita, uma vez que uma linguagem e uma lógica inatas pressupõem que a regularidade na língua deva ter um caráter instantâneo. Albano afirma, também, que o racionalismo somente garante o espaço para a criatividade

das línguas naturais. Essas, contudo, são sistemas formais muito especiais que exigem múltiplos sistemas formais para captar o seu imenso potencial simbólico.

A explicação proposta por Vygotsky para o salto qualitativo, com base na possibilidade de interiorização do conhecimento socializado, também é rejeitada por Albano, que diz que a utilização da linguagem como instrumento para a interiorização do conhecimento só é possível quando já se a conhece minimamente e que a socialização dos primeiros conhecimentos lingüísticos é bastante precária.

A proposta da autora é a de que a linguagem constrói-se por meio de condutas sensório-motoras, cuja plasticidade neurofisiológica torna-as capazes de interligarem-se a outras condutas sem perderem a própria autonomia. Essas condutas sensório motoras seriam a vocalização/audição e, no caso da surdez, a gesticulação/visão. Albano considera, portanto, que há capacidades inatas em funcionamento já no nascimento, como a capacidade de percepção de fala, mas admite, também, que a criança é influenciada pelo que lhe é oferecido; assim, ela chama seu modelo de “construtivismo”.

De acordo com Albano, a linguagem é um objeto no qual processos de referenciação cruzada da própria estrutura constituíram-se de modo muito eficaz. Essa auto-referência do objeto limitaria o poder da auto-organização do sujeito na medida em que oferece pistas sobre as direções já trilhadas com sucesso pela comunidade. Quando há pistas inscritas no próprio objeto e múltiplas interações possíveis entre aptidões já construídas, torna-se possível construir novos subsistemas cognitivos que tenham os seus próprios modos de operação.

A combinação dessas duas condições garante a formação de um subsistema mais ou menos autônomo do conhecimento, que interage com outros sistemas presentes no indivíduo sem, todavia, subordinar-se a eles. A subjetividade seria responsável pelas decisões diante do incerto, traduzindo-se na ação livre para se dissociar e se recombinar. Esse “ludismo”, assim denominado por Albano, que individualiza a espécie e cada um de seus membros, seria uma forma de lidar com o imprevisível, dando margem a novas estruturações.

A autora conclui, então, que a subjetividade não é o centro do conhecimento, mas é o seu motor. A linguagem é o objeto aprendido, recriado por pistas nele próprio contidas. Os recursos para aprender a falar desenvolvem-se por intermédio do engajamento na “brincadeira” e, posteriormente, são integrados numa atividade auto-organizada, em que a auto-referência do objeto tem um papel fundamental.

#### 1.1.5 Abrangência explicativa das principais abordagens sobre aquisição

No estado da arte da Lingüística, uma multiplicidade de modelos contrapõe-se para elucidar as inúmeras faces, quase sempre entrecruzadas, de um objeto difícil de apreender na sua totalidade. Cada uma das diferentes abordagens teóricas (Behaviorismo, Inatismo, Conexionismo, Interacionismo/Construtivismo) procura explicar o funcionamento da mente humana, fazendo uso de metáforas para representar essa realidade invisível. Tanto lingüistas quanto psicólogos tiram suas evidências da pesquisa neurológica que, no entanto, encontra-se limitada pela ética

e pelo estado atual da tecnologia<sup>5</sup>, obrigando-os a procurar outros tipos de evidência.

O papel de estruturas inatas e do meio ambiente que embasam as diferentes correntes teóricas de aquisição de uma língua explicitam aspectos diferentes do desenvolvimento da linguagem da criança. Explicações behavioristas, inatistas e conexionistas podem dar conta da aquisição de vocabulário e de morfemas gramaticais. A posição inatista, também, parece bastante plausível em sua teorização sobre a aquisição da complexidade gramatical. O processamento de informação e a pesquisa conexionista, por sua vez, envolvem simulações computacionais ou experimentos laboratoriais bastante controlados, que comprovam ser possível aprender um conjunto específico de traços lingüísticos, freqüentemente em uma língua inventada. Já explicações interacionistas podem ser úteis na compreensão de como crianças relacionam forma e significado no uso da língua, de como elas interagem nas conversações e de como aprendem a usar a fala apropriadamente em interações sociais.

Entretanto, essas abordagens apresentam limitações na abrangência de suas explicações para o fenômeno da aquisição. Muitas das afirmações da teoria behaviorista foram baseadas em experimentos com animais que aprendiam uma variedade de respostas a estímulos laboratoriais. Por isso, sua aplicabilidade na aprendizagem natural de línguas por humanos foi fortemente desafiada por psicólogos e lingüistas, principalmente pela inadequação dos modelos para explicar a complexidade envolvida no processo de aprendizagem de língua.

---

<sup>5</sup> Os exames de imagem (*PETSCAN*, *radioisótopo*), no entanto, têm mostrado muito da correlação entre função e área.

Crianças imitam exercitando-se e essa prática pode explicar a aquisição de alguns aspectos da língua, tais como significados de palavras e algumas rotinas de fala. Contudo, a imitação e a prática sozinhas não dão conta da complexidade do conhecimento que todas as crianças eventualmente adquirem. Além disso, a aquisição da linguagem parece depender de algumas capacidades já bastante diferenciadas ao nascimento, como é o caso da percepção de contrastes de sons da fala para as diferentes línguas naturais, independentemente da língua que a criança virá a falar, que lhe permite processar o que ouve e ir além disso, fazendo abstrações e atingindo generalizações complexas (Best, 1999).

Lingüistas que se apóiam nos pressupostos da GU afirmam que os experimentos dos conexionistas não os autorizam a fazer generalizações que incorporem as complexidades implícitas na aprendizagem de linguagem humana normal. Em contraste, os inatistas, que tiram muitas de suas evidências de estudos sobre a complexidade do conhecimento e da performance lingüísticos do falante proficiente, bem como da análise de suas próprias intuições sobre a língua, são também criticados, com base no argumento de que não é suficiente saber o estado final do conhecimento e de que mais atenção deveria ser dada aos passos intermediários do desenvolvimento que levam ao nível final de domínio.

Os interacionistas, que enfatizam o papel da modificação nas interações conversacionais por acreditarem que o suporte de um interlocutor ajuda na compreensão de como os sujeitos acessam novo conhecimento acerca da língua,

são também questionados. Seus críticos argumentam que o *input* é muito limitado para explicar a aquisição de processos complexos, dando maior ênfase aos princípios inatos da linguagem.

Segundo Albano (1990), subjacente ao embate sobre a influência do meio nas ações do ser humano, encontra-se o pressuposto sobre o caráter monolítico do indivíduo. Assim, os modelos behavioristas consideram-no essencialmente passivo e os modelos cognitivistas, essencialmente ativo. Mas a autora afirma que a separação entre o idiossincrático, memorizado, produtivo e regado na fala infantil não é tão clara quanto parece e que, na verdade, essa separação é forçada pelas teorias que tentam explicar o fenômeno.

Ao propor um modelo (Albano, 1990) que considera o indivíduo como dividido, sendo, ao mesmo tempo, passivo e ativo, Albano escapa da polarização e supera a dicotomia entre o racionalismo e o empirismo. A passividade faz-se representar pela repetição e pelo hábito, aos quais diferentes instâncias psíquicas podem ser suscetíveis. A atividade está, por outro lado, garantida pela interação, já que qualquer efeito sobre o todo é necessariamente mediado por um processo interacional que, por sua vez, não é inteiramente previsível entre as partes. “Assim, não há fantoches (...) e nem criadores todo-poderosos. Há apenas indivíduos cuja economia interna dá maior lugar ao hábito ou a criação” (Albano, 1990: 51).

Por meio da crítica aos modelos behavioristas e cognitivistas, Albano propõe seu Construtivismo, que incorpora alguns dos pressupostos das duas correntes na defesa de que o som é utilizado como recurso na construção da linguagem.

### 1.1.6 Aquisição de segunda língua

As teorias desenvolvidas para explicar aquisição de segunda língua estão intimamente relacionadas com as teorias de aquisição de primeira língua. Assim, algumas delas dão ênfase às características inatas do aprendiz, enquanto outras enfatizam o papel essencial do ambiente na aprendizagem e outras, ainda, procuram integrar as características do aluno com os fatores ambientais na explicação de como ocorre a aquisição de segunda língua.

Porém, uma teoria lingüística voltada para a aquisição de L2 apenas se caracteriza como tal se for capaz de explicar esse processo em toda a sua plenitude e isso não acontece. Pesquisadores e educadores almejam por teorias que lhes forneçam subsídios para a sua prática de ensino de línguas, mas frustram-se, com freqüência, diante da inadequação dos vários modelos. Isso ocorre porque as complexidades da aquisição de segunda língua, assim como as da aquisição de língua materna, representam para lingüistas, neurologistas e psicólogos um quebra-cabeça a resolver. Dessa forma, as teorias de aquisição estão ainda em discussão e, mesmo que houvesse concordância em torno de uma abordagem, haveria questões sobre como esse conhecimento poderia afetar o ensino de línguas.

Muitos professores observam o debate teórico com interesse, mas têm de continuar a ensinar. Nessas circunstâncias, é muito comum encontrar transposições diretas de pressupostos teóricos na prática pedagógica, os quais, considerados com exclusividade, fatalmente conduzem a distorções. Daí, então, a grande quantidade de pesquisas aplicadas, desenvolvidas com base nas diferentes formulações

teóricas que se iniciam nas observações da aquisição de segunda língua em contexto natural ou instrucional. Esses estudos fundamentam-se em uma gama muito grande de orientações teóricas, algumas explicitadas, outras apenas sugeridas.

É evidente que a aquisição de uma segunda língua pela criança ou pelo adulto é diferente da aquisição de primeira língua, tanto em termos de características pessoais dos sujeitos envolvidos quanto em termos das condições em que se dá a aprendizagem.

Quanto às características dos sujeitos, é necessário levar em conta que todos os aprendizes de uma segunda língua que não estejam em contexto de bilingüismo, independentemente da idade, já possuem um sistema lingüístico, o que pode ser uma vantagem, uma vez que têm uma noção de como as línguas funcionam. Por outro lado, o conhecimento prévio de um dado sistema lingüístico pode também levá-los a fazer inferências erradas de como a segunda língua opera, podendo gerar erros que não são cometidos pelas crianças ao adquirirem sua língua materna.

O aprendiz da primeira língua não tem a mesma maturidade cognitiva, consciência metalingüística ou conhecimento de mundo que o aprendiz de uma segunda língua que se supõe mais velho. Embora os aprendizes jovens de uma segunda língua já tenham uma certa maturidade cognitiva e consciência metalingüística, eles ainda têm um caminho mais longo a percorrer nessas áreas, assim como na área de conhecimento de mundo, até atingirem os níveis atingidos pelos adultos e adolescentes.

Diferenças de personalidade têm, também, uma importância singular no processo de aquisição de segunda língua. Crianças, geralmente, não se sentem nervosas ou envergonhadas ao tentarem usar língua estrangeira, mesmo que sua proficiência seja bastante limitada, enquanto adultos e adolescentes consideram estressantes situações em que são incapazes de se expressar com clareza e correção. Mas é verdade que até mesmo as crianças podem reagir diferentemente nessas situações. Algumas falam sem inibições, ao passo que outras preferem ouvir e participar silenciosamente da interação social com seus pares.

Quanto às condições em que se dá a aprendizagem, essas também são bastante diversificadas. Aprendizes jovens em um ambiente informal de aquisição de segunda língua têm a liberdade de permanecerem calados, se assim quiserem, até atingirem um estágio em que se sintam em condições de se comunicarem. Aprendizes mais velhos, por sua vez, são freqüentemente forçados a falar para atenderem a requisitos da sala de aula ou às necessidades do dia-a-dia, como fazer compras, ir ao médico ou a entrevistas de emprego.

Na tentativa de reduplicar a situação de aquisição de primeira língua, o estilo de fala ajustado, chamado de *manhês*<sup>6</sup>, é muitas vezes adotado na interação com alunos de língua estrangeira, havendo um ajuste intuitivo que o interlocutor considera necessário para que eles atinjam o entendimento.

No que se refere à correção, quando se trata de primeira língua, tende-se a limitá-la ao significado, incluindo-se escolha errônea de vocabulário. O

---

<sup>6</sup> O *manhês* consiste na fala direcionada à criança, em situação de L1, e é denominado fala de estrangeiro ou de professor em situação de L2.

mesmo ocorre na situação informal de aquisição de segunda língua, em que erros que não interferem na compreensão são ignorados, considerando-se falta de polidez interromper e corrigir quem está tentando se comunicar. Portanto, erros de gramática e de pronúncia são menosprezados, ao passo que a escolha de uma palavra errada que interfere na compreensão pode receber um comentário. A correção de uma falha é comum somente em situação de sala de aula, em situação formal de ensino de línguas. Todavia, essa correção ou a sua intensidade vai depender da abordagem de ensino adotada. Na abordagem comunicativa, na qual se tenta reconstituir a situação informal de aquisição de primeira língua, a correção é considerada “prejudicial”, didaticamente inadequada.

Assim, uma teoria geral de aquisição de segunda língua tem de levar em conta falantes com características variadas, que aprendem em contextos também diversificados. Nesse sentido, o caso específico da língua inglesa sobressai-se, pelo fato de haver no mundo mais falantes não-nativos do que falantes nativos, impondo novas metas para o seu ensino enquanto considerada uma língua internacional (Jenkins, 2000).

## 1.2 Abordagens sobre ensino de línguas e pronúncia

O ensino de línguas mudou muitas vezes seu posicionamento em relação ao ensino de pronúncia, que ora foi priorizado, ora relegado a segundo plano, quando não totalmente ignorado. Ao pesquisar sobre a história do ensino de línguas, Kelly (1969) demonstra que os filólogos dedicaram-se ao estudo da gramática e do vocabulário muito antes de se interessarem pela pronúncia. Essa é uma das razões, segundo ele, por que professores têm melhor compreensão das duas primeiras modalidades do que de pronúncia, a qual só começou a ser sistematicamente estudada um pouco antes do início do século XX.

Conforme Celce-Murcia, Brinton & Goodwill (1996), o ensino moderno de línguas apoiou-se em duas abordagens para o ensino de pronúncia: a abordagem de imitação intuitiva e a abordagem lingüística-analítica.

A **Abordagem de Imitação Intuitiva**, única utilizada até o final do século XIX, era por vezes acompanhada de observações sobre os sons, as quais eram fundamentadas em impressões e feitas por professor ou autores de livros didáticos; entretanto, por serem baseadas na ortografia, eram, freqüentemente, imprecisas do ponto de vista fonético. Essa abordagem pressupunha a capacidade do estudante de ouvir e imitar o ritmo e o som da língua estrangeira, sem a intervenção de qualquer informação explícita, mas tendo acesso a bons modelos, proporcionados pelo uso de gravadores e laboratórios de línguas na metade do século XX e, mais recentemente, pelo uso de videocassetes.

A **Abordagem Lingüística-analítica**, por outro lado, fazia uso de informação

e ferramentas tais como o alfabeto fonético, descrições articulatórias, ilustrações do aparato vocal, informação contrastiva e outros suportes para as tarefas de ouvir, imitar e produzir. Essa abordagem, que explicitamente informava o aluno e focalizava sua atenção nos sons e ritmos da língua estrangeira, foi desenvolvida para complementar e não para substituir a abordagem de imitação intuitiva, que foi mantida na fase de treinamento da informação fonética.

Observe-se que ambas partiam da convicção de que falantes nativos são inconscientes do funcionamento da língua, mas divergiam suas opiniões no que se refere à aplicação dessa idéia para o ensino de línguas estrangeiras. Os métodos de ensino que se desenvolveram, dependendo das suas propostas, podem ser enquadrados em uma dessas duas abordagens. Assim, alguns dos métodos de ensino adotados ao longo do século XX, tais como o ***Grammar Translation*** e a **Abordagem Baseada na Leitura**, não incluíram a pronúncia entre seus objetivos instrucionais. No entanto, outros o fizeram, embora partissem de diferentes pressupostos e práticas pedagógicas.

O **Método Direto** (que vigorou no final de 1800 e começo de 1900) baseava-se em observações intuitivas de como crianças aprendiam sua língua nativa e de como adultos e crianças aprendiam uma língua estrangeira fora de contextos formais de ensino e esperava, em conformidade com a teoria behaviorista, que o aluno aprendesse a pronúncia de uma língua pela imitação intuitiva de um modelo, o qual podia ser o professor ou o gravador.

Os **Métodos Naturalistas**, sucessores do método direto, defendiam a hipótese de que antes de qualquer tentativa de fala, os alunos deveriam ser submetidos a um

período de audição sem qualquer instrução explícita sobre pronúncia, a fim de internalizarem o sistema de som da língua estrangeira.

A primeira contribuição da lingüística para o ensino de pronúncia aconteceu em 1890, como parte do Movimento de Reforma do Ensino de Línguas, amplamente influenciado por foneticistas como Henry Sweet, Wilhelm Viëtor e Paul Passy. Foram eles que criaram a Associação Fonética Internacional em 1886 e desenvolveram o Alfabeto Fonético Internacional (IPA), produto do reconhecimento da Fonética como a ciência que descreve e analisa os sistemas de som das línguas. Com a criação do IPA, tornou-se possível representar os sons de qualquer sistema lingüístico porque, pela primeira vez, havia a possibilidade de se estabelecer uma relação consistente de um para um entre os símbolos da escrita e o som por eles representados.

No ensino de línguas foram instituídas as seguintes práticas, conforme apontado na resenha elaborada por Celce-Murcia *et al.* (1996):

- ensino da língua oral é primordial e deve ser priorizado;
- as descobertas da Fonética devem ser incorporadas ao ensino;
- os professores devem ter treinamento sólido em Fonética;
- o aluno deve se valer do treinamento fonético para estabelecer bons hábitos de fala.

O Movimento de Reforma deu origem, entre 1940 e 1950, ao **Audiolingüismo** nos Estados Unidos e à **Abordagem Oral** na Grã-Bretanha, que enfatizaram o ensino de pronúncia com informação explícita desde o início do processo de aprendizagem. O aluno imitava repetindo um modelo que podia ser o professor ou um gravador, mas recebia também apoio de conceitos da Fonética, de

sistemas de transcrições e de ilustrações da articulação dos sons. Foram também introduzidos os exercícios com pares mínimos que trabalham com as oposições de sons vocálicos e consonantais da língua, como em bad/bed ou bad/pad (Bloomfield, 1933).

Nos anos 60s, a **Abordagem Cognitiva**, influenciada pela Gramática Transformacional Gerativa de Chomsky (1957, 1965) e pela Psicologia Cognitiva (Neisser, 1967), passou a ver a linguagem como um comportamento ditado por regras e não como formação de hábito, voltando a enfatizar o ensino de gramática e vocabulário em detrimento do estudo da pronúncia. Apoiados na Hipótese do Período Crítico de aquisição, teóricos cognitivistas consideravam que uma pronúncia semelhante a do falante nativo era um objetivo impossível de ser alcançado.

Os métodos que apareceram nos anos 70s continuaram a apresentar diferenças significativas quanto ao ensino de pronúncia. O “**Silent Way**” (Gattegno, 1972, 1976), cujo objetivo era o ensino de pronúncia, procurava focalizar a atenção dos alunos em como a combinação de palavras, o acento e a entoação dão forma à produção de uma emissão. O professor evitava falar, e procurava o apoio visual de ilustrações de sons e palavras baseadas em cores. Ainda hoje, muitos educadores acreditam que o princípio utilizado de correspondência som-cor ajuda os alunos a internalizarem a dicção, ritmo e melodia da língua (Blair, 1991).

A abordagem “**Community Language Learning**” (CLL) teve suas origens na abordagem humanística de ensino-aprendizagem centrado no aluno de Carl Rogers (1951). No método CLL, várias técnicas e ferramentas são centrais para o ensino de pronúncia. O gravador, peça central do método por gravar as emissões dos alunos,

proporciona um distanciamento do que foi dito, possibilitando que a atenção recaia no como foi dito e tornando possível a comparação com a emissão do professor. O professor, no papel de computador humano, que podia ser ligado ou desligado conforme desejo do aluno, não oferecia correção de pronúncia, ficando a critério do estudante a seleção do que praticar e o quanto praticar.

A **Abordagem Comunicativa**, que vem dominando o ensino de línguas desde os anos 80, passa a enfatizar a comunicação, que, para ser eficaz, requer uma pronúncia inteligível, de modo a que o ouvinte seja capaz de apreender a mensagem tal como pretendida pelo falante, sendo essa, ao mesmo tempo, aceitável socialmente. No entanto, seus proponentes não desenvolveram as estratégias que levariam ao ensino de pronúncia de uma maneira comunicativa. Portanto, professores continuam a empregar as mesmas técnicas que foram tradicionalmente utilizadas, tais como: ouvir e imitar, treino fonético, pares mínimos, "*tongue twisters*". Assim, a ênfase no ensino de pronúncia continua sendo a produção correta do som no nível da palavra, embora apareçam tentativas de tornar essas estratégias mais comunicativas diante da forte evidência de que a pronúncia de palavras isoladas é distinta de sua pronúncia em fala encadeada.

Influenciado pelas teorias de análise do discurso, o ensino de pronúncia volta-se para os traços supra-segmentais da linguagem, ritmo, acento e entoação, em um contexto discursivo. De acordo com McNermey e Mendelsohn (1992: 186),  
".. a short term pronunciation course should focus first and foremost on suprasegmentals as they have the greatest impact on the comprehensibility of the learner's English."

Atualmente, o debate entre o ensinar traços segmentais versus supra-segmentais encontra-se em bases mais sensatas. Existe consenso em que fracassos na produção de distinções significativas em qualquer dos dois níveis podem ter um impacto negativo no processo comunicativo de falantes não-nativos da língua.

Tem-se procurado, também, enfocar o ensino de pronúncia de forma reflexiva como Lieff (2003) que, em curso para professores de inglês da rede pública, propõe ajudar professores a perceberem o tipo de ensino de pronúncia relevante para aquele contexto específico de sala de aula e a maneira de desenvolvê-lo em sua prática e no trato com os alunos.

Conclui-se, então, que o ensino de pronúncia foi acumulando estratégias adotadas pelos diferentes métodos de ensino de línguas, as quais por sua vez derivam, de certo modo, das teorias de aquisição de primeira e segunda línguas, com pequenos ajustes ao longo do tempo.

Assim, o uso dos *drills* com pares mínimos apresenta-se contextualizado em sentenças, o emprego da técnica do **repita e imite** é hoje apoiado pelo uso do gravador em uma abordagem como a do CLL, proporcionando ao aluno mais autonomia em seu processo de aprendizagem. No entanto, todas as abordagens e estratégias apresentadas pecam por apoiarem o ensino, a produção e avaliação das emissões dos alunos principalmente na percepção auditiva que é falha, imprecisa e exige treino específico, como apontado em pesquisa anterior (Rocca e Marcelino, 1997). Novas estratégias que aprimorem a percepção e/ou sejam independentes do julgamento perceptual devem ser implementadas.

## CAPÍTULO 2 - PERCEPÇÃO E PRODUÇÃO DA FALA

## 2.1 Percepção e produção da fala na aquisição de língua

A percepção da fala constitui a base da sua subsequente produção. Embasadas nessa afirmação, são três as abordagens teóricas mais significativas sobre produção e percepção, tais como apresentadas por Best (1995), que tentam explicar quais componentes da fala são percebidos e processados pelos ouvintes, visando a apresentar conclusões pertinentes sobre a produção da língua adquirida.

A abordagem **Psicoacústica** da percepção da fala (Aslin, Pisoni & Jusczyk, 1983) assume que o elemento básico na percepção da fala é a pista acústica proximal, isto é, a decomposição do sinal da fala em características espectrais e temporais de eventos específicos na forma da onda. Durante a aquisição da primeira língua, as crianças aprendem a associar certas combinações de traços acústicos com unidades lingüísticas de significado e estrutura fonológica. Essas constelações acústicas são, então, estocadas na memória como protótipos auditivos que influenciam fortemente a percepção de fala nova e, mais importante, servem como *templates*/moldes de produção na aquisição de linguagem.

Em contraste com a teoria psico-acústica, **a Teoria Motora** (Liberman & Mattingly, 1985) considera as unidades articulatórias como os primitivos da percepção da fala e não as unidades acústicas. Ela assume que um módulo fonético especializado associa um sinal de fala com uma unidade fonológica por

meio da representação neuromotora dos gestos articulatórios<sup>1</sup>. Também assume que a percepção e a produção são mediadas por um módulo neural inato, que gera os gestos articulatórios necessários. Assim, na aquisição da primeira e segunda línguas, os aprendizes identificam os gestos articulatórios que eles percebem na língua-alvo e, para a produção da fala, as representações estocadas dos gestos são mantidas e realizadas.

A abordagem **Realista Direta** da percepção da fala, por sua vez, deriva do ponto de vista filosófico do realismo direto, também conhecido como teoria da percepção ecológica, que discute as origens do conhecimento perceptual (Gibson & Gibson, 1955). Seu ponto central está na afirmação de que em todas as formas de percepção o sujeito que percebe apreende diretamente o objeto perceptual e não meramente uma representação da qual o objeto deva ser inferido (Best, 1995). Portanto, os primitivos perceptuais da fala são os próprios gestos articulatórios produzidos pelo trato vocal do falante, os quais são diretamente percebidos pelo ouvinte sem qualquer representação intermediária.

É nesse ponto que tal abordagem difere-se da Teoria Motora, que também postula gestos articulatórios como os elementos perceptuais básicos, mas na forma de gestos pretendidos, representados na mente. A abordagem Realista Direta, por outro lado, assume que, de acordo com os princípios de física acústica, a estrutura acústica da fala correlaciona-se exatamente com a forma do trato vocal, tal como determinado pelos articuladores ativos e passivos; portanto, a informação gestual

---

<sup>1</sup> O gesto articulatório é a unidade analítica abstrata do modelo de Fonologia Articulatória (Browman e Goldstein, 1986,1990) e da proposta de Fonologia Acústico-Articulatória (Albano,2001). Segundo Browman e Goldstein (1986:223) os "gestos são padrões de movimentos dos articuladores no trato vocal ou sistemas articulatórios". Constelações de gestos ou pautas gestuais referem-se às estruturas de organização de gestos nas quais representações abstratas dos movimentos articulatórios coordenam-se com representações de outros movimentos (Ball e Kent, 1997). O gesto difere de outras unidades fonológicas básicas postuladas pelos modelos fonológicos, tais como, o fonema e o traço distintivo. O gesto é dotado de tempo intrínseco e pode envolver mais de um articulador. A grande vantagem da adoção dessa unidade analítica é que ela possibilita a consideração do elo entre o elemento lingüístico e o físico na estrutura da fala.

está sempre disponível na fala.

Desse modo, o falante deve inferir da fala percebida quais gestos relevantes estão sendo usados. Uma vez que o falante, em contraste com a **Teoria Motora**, não tem conhecimento inato desses gestos, ele, a princípio, percebe apenas informação não-lingüística e, conforme a percepção muda para um foco lingüístico, começa a descobrir correspondências entre relações de invariantes de ordem maior entre gestos ou constelações gestuais de um lado e elementos lingüísticos, do outro. Gradualmente o sistema perceptual torna-se afinado com os invariantes de ordem superior que especificam as constelações gestuais (que correspondem a eventos articulatórios mais complexos), e que constituem o inventário fonológico da língua. Por conseguinte, o falante-ouvinte torna-se mais eficiente em detectar as principais propriedades distintivas dessas constelações e também em ignorar detalhes irrelevantes.

O resultado final dessa sintonia é a criação do chamado espaço fonológico nativo, região articulatória na qual os gestos cruciais específicos de cada língua ocorrem.

Em suma, a percepção da fala constitui a base da sua subsequente produção e, independentemente das diferentes explicações de como ela se estabelece, essa relação entre percepção e produção é crucial tanto para a aquisição de primeira língua quanto para a aquisição de segunda língua.

## 2.2 Origem do Sotaque estrangeiro

Sotaque estrangeiro, todos nós temos, com exceção dos que falam apenas sua língua materna. No entanto, definir sotaque estrangeiro não é uma tarefa fácil. De uma maneira geral, ele refere-se aos desvios de realização das unidades fonológicas da língua não nativa que têm sua origem em traços específicos do sistema lingüístico do falante (Jilka, 2000). A percepção do sotaque estrangeiro é bastante desenvolvida entre os ouvintes e muitos podem mesmo imitar ou associar certos sotaques com línguas específicas.

Sua ocorrência é constatada quando as estruturas fonológicas estabelecidas, que foram formadas como requisitos para a aquisição da língua materna, são confrontadas com dados de fala de uma segunda língua a ser adquirida. Existe, também, razão suficiente para acreditar que há pistas fonéticas da L2 que, por não terem correspondentes diretos no sistema fonológico nativo, não são percebidas pelos falantes não-nativos. Nesse último caso, o vínculo entre produção e percepção no indivíduo adulto é de tal ordem que impede que ele perceba outras distinções fora de seu sistema fonológico e esse fato contribui significativamente para a impossibilidade de produzir fala semelhante à do nativo.

São várias as teorias que tentam explicar a origem do sotaque estrangeiro, baseadas na relação entre percepção e produção. A seguir, apresentam-se as mais relevantes.

A teoria do **Magnetismo da Língua Nativa** (Kuhl, 1991 a e b; Kuhl e Iverson,

1995) é um modelo de percepção da fala e de aquisição de língua especialmente voltado à estrutura e representação de categorias fonéticas e fonológicas. Seu foco principal está na dependência que a percepção tem de uma dada representação e nas conseqüências para a produção, especificamente na aquisição de segunda língua, quando dados de fala de L2 são confrontados com representações estabelecidas na L1, resultando em representações de categorias de L2, que são, por conseguinte, a origem de sotaque estrangeiro na produção oral.

A origem e a motivação da teoria vem da constatação de que crianças são capazes de ouvir todos os sons das línguas humanas até os 6 meses, enquanto adultos apresentam uma capacidade de percepção de fala mais voltada para sua língua nativa. De acordo com esse modelo, isso ocorre porque a exposição à língua no começo da vida produz uma mudança na percepção de distância no espaço acústico que está subjacente às distinções fonéticas (Kuhl & Iverson, 1995).

É, portanto, a distorção do espaço perceptual que influencia a percepção dos sons da fala de tal modo que alguns sons não podem mais ser identificados independentemente. A distorção é explicada pelo desenvolvimento de protótipos fonéticos, definidos como exemplos de categoria fonológica para uma língua específica, os quais têm o efeito de imãs perceptuais, atraindo outros sons da mesma categoria e reduzindo as distâncias perceptuais.

Resultados de experimentos (Kuhl, Williams, Lacerda, Stevens, e Lindblom, 1992) comprovam o efeito magnético tanto em adultos como em crianças de 6 meses de idade, demonstrando que a exposição à língua ambiental altera o espaço

perceptual já aos 6 meses, quando representações específicas de cada língua estão começando a se formar. Nessa idade, as crianças passam a ouvir apenas os contrastes de sons de sua língua materna, o que propicia o surgimento do balbúcio. A teoria do Magnetismo da Língua Nativa dá, então, conta da percepção da fala no primeiro ano de vida, a qual precede antes da aquisição do significado da palavra e da fonologia contrastiva.

Antes da alteração do espaço perceptual pelo desenvolvimento do magnetismo perceptual, as crianças possuem a habilidade inata de resolver todas as diferenças acústicas entre as unidades fonéticas possíveis na linguagem humana. Alguns dos limites naturais entre essas categorias fonológicas potenciais desaparecem com a aquisição da primeira língua quando essa capacidade perceptual se reduz para atender às necessidades da aquisição de um sistema fonológico específico. Segundo Gama-Rossi<sup>2</sup>, seria um desperdício manter uma capacidade perceptual que desse conta das distinções entre sons das diferentes línguas do mundo, já que a natureza parece não ter previsto a Torre de Babel, a evolução e diversidade das línguas e nem o fato de o ser humano falar mais de uma língua.

Para Jilka (2000: 27), a representação da estrutura e do desenvolvimento perceptual decorrente da teoria de Magnetismo da Língua Nativa pode ser estendida facilmente para explicar a percepção de sons de língua estrangeira e a produção de fala com sotaque estrangeiro.

O problema crucial para aprendizes de segunda língua é que os novos sons de fala com os quais se deparam são confrontados com as representações já estabelecidas dos magnetismos e limites perceptuais da L1. Assim, os sons da L2 que são foneticamente semelhantes às categorias estabelecidas na L1 são atingidos pelo efeito magnético. Por essa razão eles podem, ou não, ser percebidos como sons independentes do magnetismo perceptual ou, ainda, podem ter sua percepção significativamente influenciada pelo espaço perceptual modificado.

Esse fenômeno explicaria exemplos clássicos de sotaque estrangeiro na língua inglesa como o /l/ e o /r/ do japonês ou as oclusivas surdas aspiradas do brasileiro. Se o falante não é capaz de estabelecer representações perceptuais distintas para os dois sons, não é capaz de produzi-los corretamente.

A ligação entre percepção e produção explica-se pelo fato de que as representações perceptuais que são estabelecidas em uma língua servem como alvos para a aquisição dos gestos articulatórios relevantes para a produção. Desse modo, se as representações de alguns sons da L2 não correspondem às características fonéticas reais dos sons produzidos pelos falantes nativos, elas farão surgir o sotaque estrangeiro na fala.

As principais afirmações da teoria do Magnetismo da Língua Nativa sobre a percepção e produção da fala, tal como, a influência das estruturas representacionais estabelecidas na aquisição de L1 sobre a percepção e produção da L2, são compartilhadas pelo ***Speech Learning Model*** (SLM), modelo de

---

<sup>2</sup> Comunicação feita durante o exame de qualificação deste trabalho.

aquisição de segunda língua proposto por Flege (1995a). O SLM tem como meta investigar o papel do fator idade na aquisição da segunda língua, pois, com base na Hipótese do Período Crítico, considera ser a idade o critério que define a distinção entre a aquisição da língua nativa e a aquisição de segunda língua, principalmente no que diz respeito ao nível de proficiência.

Pressupõe-se, assim, que, durante a aquisição de L1, a percepção da fala ajusta-se aos elementos fônicos contrastivos da L1 (Flege, 1995a). Conseqüentemente, os aprendizes de língua estrangeira terão dificuldades em distinguir sons distintos que ocorrem apenas em L2, uma vez que esses não poderão ser assimilados pela categoria nativa. Assim, sem uma representação acurada para servir de alvo perceptual, a produção não soará nativa, mas com sotaque estrangeiro.

Flege, no entanto, ressalta que nem todo sotaque estrangeiro é motivado pela percepção. Muitos casos, como o do /eskul/ do falante do espanhol não tem uma base perceptual, mas pode ser explicado pela aplicação de uma restrição de *output* do espanhol.

A abordagem em foco (SLM) tem o intuito de explicar sistematicamente todos os tipos de desvios segmentais nas produções de L2 (incluindo a causa e a extensão do sotaque estrangeiro, especialmente com relação à idade), a fim de dar conta do processo de aquisição de segunda língua. Para tanto, é constituída de 4 postulados, cujos princípios básicos são expressos nas seguintes hipóteses:

Hipótese 1 (H1): os aprendizes relacionam perceptualmente alofones posicionais concretos da L2 com alofones similares perceptualmente na L1.

Contudo, o *Speech Learning Model* assume um nível abstrato de organização implícito na Hipótese 2 (H2) que afirma: se o aprendiz discernir uma diferença fonética entre um som da L2 e da L1, uma nova categoria fonética pode ser estabelecida para o som da L2.

Hipótese 3 (H3): a probabilidade de que o aprendiz torne-se consciente de tais diferenças aumenta na mesma proporção em que aumenta a diferença fonética entre o som da L2 e o som mais próximo na L1.

Hipótese 4 (H4): conforme aumenta a idade em que se dá a aprendizagem, os aprendizes são menos capazes de detectar essas diferenças fonéticas.

Desse modo, idade de aprendizagem correlaciona-se com a habilidade dos aprendizes em estabelecerem novas representações fonológicas dos sons da L2.

Hipótese 5 (H5): a formação de categoria pode ser impedida pelo mecanismo de classificação equivalente.

Em tais casos, os sons da L1 e da L2 são percebidos como tão semelhantes que são classificados na mesma categoria. Como resultado, o som da L2 não será representado da maneira nativa devido à interferência da L1. A interferência, contudo, é bidirecional, alterando também, de certo modo, a forma fonética da L1 e

fazendo que os dois sons se assemelhem muito, como pode ser observado no caso de bilíngües de infância (Rocca, 2003).

Em resumo, o aprendiz tem duas grandes possibilidades de classificação e processamento do som da L2, já que o som que for idêntico a um som correspondente na L1 será automaticamente representado de maneira correta. Se o som é muito semelhante a um som correspondente na L1, ele será representado pela categoria já existente na L1 com as conseqüências já descritas. Se o som da L2 é claramente distinto de qualquer som da L1, ele estabelecerá uma nova categoria, que reproduz a representação correspondente do falante nativo da L2 e, então, não é realizado com sotaque estrangeiro.

Contudo, em conformidade com a Hipótese 6 (H6), também é possível que a nova categoria estabelecida pelo aprendiz de L2 seja diferente da categoria equivalente do falante nativo monolíngüe. Isso ocorre porque a categoria do falante da L2 pode ser desviada de uma categoria próxima da L1 para manter contraste no espaço fonológico comum da L1 e L2 (Bohn & Flege, 1992). A alteração de uma categoria pode ser causada pelo uso de traços diferentes ou exagerados pelo falante da L2 para discriminá-la de outra. Essas representações categóricas com desvios daquelas equivalentes do falante nativo serão realizadas com sotaque estrangeiro na produção da fala (Rocca, 1999).

Hipótese 7 (H7): estabelece a conexão entre a representação categórica e a produção, afirmando que a produção de um som no curso da aquisição corresponde, eventualmente, às propriedades de sua representação de categoria

fonética.

Desse modo, o SLM propõe-se a descrever o processo de aquisição de L2, dando conta da ocorrência de sotaque estrangeiro no nível segmental, tendo como base a percepção dos contrastes da L1 e L2 e levando em consideração a contribuição de fatores relacionados à idade.

Best (1995), com o seu Modelo de Assimilação Perceptual, também apresenta uma teoria de aquisição de segunda língua e de sotaque estrangeiro. A diferença do modelo anterior é que o **Perceptual Assimilation Model** (PAM) utiliza como referência o realismo direto.

Assim, gestos articulatórios no espaço fonológico nativo são as unidades elementares da percepção e produção da fala. Gestos não nativos, contudo, são percebidos, primeiramente, de um modo direto, de acordo com suas semelhanças e discrepâncias para com as constelações de gestos segmentais nativos, em função da maior ou menor proximidade em relação ao espaço fonológico nativo. Tal espaço fonológico, definido pelo *layout* espacial do trato vocal e pelas características dinâmicas dos gestos articulatórios, determina as dimensões dentro das quais a semelhança é julgada.

A PAM distingue 3 relações possíveis entre os gestos articulatórios, isto é, categorias de um falante de L1 e aquelas requeridas para se falar a L2. A primeira decorre da semelhança entre gestos. Se semelhanças gestuais são detectadas, dependendo do grau de igualdade, as categorias não nativas serão assimiladas a uma categoria nativa ou como bons exemplos, aceitáveis, embora não ideais, ou,

então, serão assimiladas com claros desvios da categoria nativa.

A segunda resulta da impressão de discrepância entre os gestos da L1 e da L2. Nesse caso, ele será assimilado dentro do espaço fonológico nativo como uma constelação gestual que ainda caracteriza ou remete a um som da fala, mas formando uma nova categoria no espaço das categorias nativas já existentes.

Uma terceira possibilidade, por fim, ocorre quando nenhuma constelação gestual de fala é reconhecida. Nesse caso, o gesto não é assimilado pelo espaço fonológico nativo e é ouvido como um som que **não é** de fala. Estariam, por exemplo, incluídos, aqui, os cliques de algumas línguas africanas, que, para nós, seriam classificados como ruídos ou onomatopéias.

Uma vez que, de acordo com o PAM, o que é assimilado não são os gestos articulatórios em si, mas os contrastes auditivos entre os gestos, o modelo descreve padrões de assimilação para os contrastes entre os sons não nativos adquiridos. Esses padrões, no entanto, podem ser inferidos dos tipos de assimilação das categorias de L2 que acabam de ser descritos.

Como é sugerido pelo próprio nome, o PAM focaliza a atenção na percepção da fala não nativa durante a aquisição de L2 e não explicita as conseqüências para a produção e sotaque estrangeiro, mas podemos inferir que seriam semelhantes às previstas pelo SLM. O PAM está de algum modo relacionado ao SLM apesar da diferença fundamental no conceito da natureza da percepção, uma vez que os dois modelos tentam explicar a aquisição de segunda língua pela percepção da relação entre os sons da L1 e L2.

### **2.3 Relação entre percepção e produção na aquisição de segunda língua**

A relação entre a percepção dos sons da língua estrangeira e sua produção por falantes não nativos ainda não é totalmente compreendida: a produção precede a percepção ou o inverso, a percepção antecede a produção no processo de aquisição de segunda língua? A resposta a essa questão tem implicações teóricas no que diz respeito ao processo de aquisição de L2, mas, também, conseqüências práticas no que se refere à metodologia empregada no ensino da pronúncia e entoação.

Relatos de experimentos (Llisterri, 1995) comprovam que as habilidades de produção e de percepção de sons de uma segunda língua estão intimamente ligadas, mas que a precedência de uma habilidade sobre a outra em relação ao treinamento de falantes não nativos não está claramente comprovada, o que difere claramente da aquisição de L1, processo durante o qual há precedência da compreensão sobre a produção. Por tratar-se de uma interação de natureza complexa, a relação percepção/produção em aquisição de L2 requer a consideração de vários fatores na sua descrição. Todavia, algumas considerações apresentam-se como tendências gerais nos relatos de experimentos, tais como:

- **O estágio de aprendizagem na aquisição da L2, a experiência com a língua, o grau de exposição e a idade de aquisição** parecem desempenhar um papel importante na interação entre produção e percepção.

Elsendoorn (1984) comprova em seu estudo que desvios-padrão de percepção em falantes não nativos diminuem na proporção em que aumentam o grau de familiaridade e de conhecimento da língua. Os resultados de Bohn & Flege (1990) confirmam as descobertas de Elsendoorn: uma exposição maior à língua estrangeira tem como resultado uma produção e percepção mais apuradas.

- **Não parece ser possível inferir habilidades de produção das habilidades de percepção e vice-versa.** Apesar da hipótese de filtro fonológico (Trubetzkoy, 1949) e do princípio de classificação por equivalência (Flege, 1991) indicarem que a percepção de um novo contraste fonético deve preceder sua produção, outros experimentos, como os de Borrell (1990), Neufeld (1988) e Brière (1966), concluem que nem todo som que é percebido corretamente será produzido de maneira adequada. De acordo com Flege (1988), existe uma dicotomia entre a produção e a percepção da fala, e os falantes não nativos têm uma habilidade maior em detectar produções com sotaque do que em reproduzir os sons da língua estrangeira. Em certos casos, no entanto, a produção de sons da L2 pode preceder sua percepção. Estudos sobre bilíngües como o de Sheldon (1985) indicam que os falantes bilíngües adaptam melhor sua produção do que sua percepção em sua segunda língua.

- **Fatores sociais, como a necessidade de se fazer compreender, podem explicar casos em que a produção precede a percepção.** Tanto Sheldon (1985) quanto Mack (1989) concluem em suas pesquisas que as conseqüências sociais da produção não-nativa são mais importantes do que

as conseqüências sociais da percepção não-nativa. Assim, o desenvolvimento da produção pode preceder ao da percepção em falantes não nativos, desde que seja mantido um grau de compreensão mínimo que garanta o diálogo.

- **A relação entre produção e percepção pode diferir, dependendo da classe de sons, dos correlatos acústicos e perceptuais dessas classes e de efeitos de contextos.** Bohn e Flege (1990) argumentam que as estratégias usadas pelos falantes da segunda língua podem não ser as mesmas utilizadas pelos falantes nativos. Assim, o falante não nativo pode se basear em dados de duração para perceber e produzir vogais de diferentes timbres, enquanto o falante nativo emprega diferenças espectrais com a mesma finalidade<sup>3</sup>. Em seu estudo, os autores comprovam que relações diferentes entre percepção e produção podem ser estabelecidas para diferentes correlatos acústicos das categorias fonéticas. De acordo com seus dados, falantes não nativos podem vir a distinguir a duração da vogal baseados no contexto, mas a mesma situação contextual não os auxilia na produção das mesmas vogais.

## 2.4 Percepção e aquisição da entoação

Entende-se por entoação “*the occurrence of recurring pitch patterns, each of which is used with a set of consistent meanings, either on single words or on groups of words of varying length*” (Cruttenden, 1986: 9). Portanto, a entoação

---

<sup>3</sup> É o caso de alunos brasileiros que tendem a fazer a distinção entre certas vogais do inglês com base apenas na duração, como por exemplo, as vogais [ɪ] e [i:].

*envolve “all aspects of the perceived pitch pattern that the speaker intends for the hearer to use in understanding the utterance, or that the hearer does use whether intentionally controlled by the speaker or not” (Beckman, 1995: 119).*

Todos os conceitos e teorias de aquisição de segunda língua descritos até agora são orientados para os segmentos. A pergunta que se coloca é se eles são aplicáveis na aquisição da entoação e na explicação do sotaque estrangeiro motivado por ela. A entoação é representada de maneira idêntica, semelhante ou de uma maneira completamente diferente da dos segmentos?

Há pesquisas na área da neurolingüística que defendem que a prosódia é processada em regiões especializadas do cérebro que diferem das regiões que processam os segmentos (Ross, 1980; Mayer, Dogil, Wildgruber, Riecker, Ackermann & Grodd, 1999; Baum & Pell, 1999). Entretanto, outros estudos apontam para a especificidade das tarefas. Se essas envolverem análise, independentemente de serem segmentos ou prosódia, o processamento envolveria o hemisfério esquerdo (Locke, 1997).

De acordo com Kuhl & Iverson (1995), as propriedades prosódicas de uma língua já são aprendidas no útero. Ramus e Mehler (1999) demonstram, também, que recém-nascidos mostram preferência pela voz de suas mães, por certos padrões de acento e podem discriminar línguas com base em sinais de fala filtrados por uma banda baixa.

Essas características da prosódia poderiam explicar a percepção menos

consciente que se tem dela, em comparação com a percepção dos segmentos, principalmente em línguas como o inglês e o português, que a empregam para expressar funções de discurso e/ou funções pragmáticas e não fazem uso dela, ou somente raramente, para expressar significado lexical. Daí ser difícil relacionar estruturas entoacionais particulares com um conteúdo semântico específico, uma vez que o significado semântico ou pragmático geralmente associado a um certo padrão entoacional pode variar segundo o contexto.

Contudo, estudos demonstram que os aspectos psicoacústicos da percepção dos traços supra-segmentais constituem parte importante da capacidade humana de percepção da fala. Nesse sentido, Clark e Yallop (1995: 306) apontam que o ser humano é capaz de discernir mudanças na frequência de apenas 0,5%. Tais alterações, em frequências abaixo de 1000 Hz, são de suma importância para perceber movimentos importantes de F0, ou seja, de entoação, ao passo que mudanças em regiões acima de 1000 Hz são cruciais para a distinção de frequência de formantes que diferenciam as categorias segmentais.

A resolução no tempo, a duração, também contribui para a percepção de traços segmentais e os humanos são capazes de distinguir diferenças temporais de apenas 20 ms em frequências entre 500 e 1500 Hz (Pisoni, 1977). Os estudos também demonstraram que a relação entre o volume do som e sua captação pelo ouvido humano está intrinsecamente relacionada com a frequência. O limite de intensidade para um som ser detectado foi estabelecido em cerca de 70dB entre frequências de 20 Hz e 1500 Hz (Clark & Yallop, 1995).

Além dessas considerações gerais sobre a percepção dos traços supra-segmentais, vários estudos tratam especificamente da percepção de características específicas do contorno de F0, sendo que os mais pertinentes a esse respeito tratam da investigação de dois aspectos muito importantes no contorno de F0: a incursão do *pitch* e o alinhamento temporal. Gussenhoven, Repp, Rietveld, Rump & J. Terken (1997), por exemplo, examinam a relação entre a proeminência dos picos de F0 em sílabas acentuadas do holandês e sua posição na emissão. Concluem que picos mais baixos são percebidos como tendo o mesmo grau de proeminência que picos mais altos, anteriores na emissão, porque a proeminência do pico é controlada por uma linha abstrata referencial que reflete a extensão global do *pitch*.

## **2.5 Entoação e sotaque estrangeiro**

Embora se reconheça que a prosódia e, principalmente, a entoação contribuem para a impressão de sotaque estrangeiro, os estudos e as análises que tratam especificamente do problema ou da descrição do sotaque estrangeiro decorrente da prosódia/entoação são raros, principalmente no contexto de modelos de aquisição de segunda língua. Aspectos segmentais de sotaque estrangeiro, no entanto, têm sido descritos com maiores detalhes. Conseqüentemente, os modelos de aquisição de segunda língua têm se desenvolvido com base em aspectos segmentais da aquisição e de sotaque estrangeiro.

Por outro lado, estudos contrastivos de entoação, na medida em que comparam padrões entoacionais de duas línguas, podem contribuir para a pesquisa do sotaque estrangeiro. Eles implicitamente sugerem que diferenças

fonológicas e fonéticas detectadas serão percebidas como sotaque estrangeiro quando um falante expressa-se na língua não nativa e apontam para possíveis causas e manifestações desse fenômeno (Grabe, 1997). Alguns tentam identificar e confirmar a relevância de desvios prosódicos descritos, mas muitos, entretanto, não se referem exclusivamente à prosódia e muito menos à entoação, além de não estabelecerem uma relação com conceitos de aquisição de segunda língua.

Algumas pesquisas que contrastam aspectos pontuais da prosódia do português e do inglês (Barbosa, 1996; Massini, 1991; Rocca & Marcelino, 1997) apontam diferenças que podem ser resumidas nos seguintes itens:

- direção do movimento do *pitch* em interrogativas do tipo *Yes/No*. No inglês, o movimento do *pitch* ascendente é diferente daquele que ocorre no português;
- magnitude do movimento do *pitch* como, por exemplo, a magnitude da ascensão final em *Yes/No questions* que, no português, tem uma largura de banda mais estreita do que no inglês;
- realização de acentos de *pitch*;
- diferenças no contorno de F0 associadas com contextos tonais frasais específicos e situações discursivas;
- nível de F0 no começo da emissão: o português tem um *onset* mais alto no começo da frase entoacional e, portanto, efeitos de declinação são mais graduais do que no inglês;

No entanto, os resultados dos estudos comparativos de entoação não podem

ser diretamente equacionados com sotaque estrangeiro em relação aos aspectos entoacionais.

## **2.6 Identificação de sotaque estrangeiro entoacional**

A identificação do que constitui um caso de sotaque estrangeiro entoacional parece clara: a entoação na fala de um não nativo deve apresentar um desvio tal, de modo a ser considerada inapropriada pelo nativo.

No entanto, a decisão do que é inapropriado em termos de entoação depende amplamente do contexto, mais do que no caso de julgamentos de desvios na articulação segmental. Constituem, então, um pré-requisito para a análise do sotaque estrangeiro entoacional, a clareza e a delimitação do contexto, de modo a permitir o julgamento de uma realização particular.

O fato de a entoação permitir um alto grau de variação na escolha e distribuição de categorias tonais é uma agravante no processo de identificação do sotaque estrangeiro, pois as variações que são notadas podem ser interpretadas como iguais ou um pouco diferentes, mas não serem consideradas inapropriadas, ou seja, com sotaque estrangeiro.

A fim de garantir uma análise homogênea, algumas hipóteses de trabalho possíveis serão adotadas. Seguindo a metodologia proposta por Jilka (2000), parte-se do princípio de que, se um desvio entoacional é detectado em uma das emissões dos sujeitos que participaram deste experimento, isso é, de fato, devido ao sotaque

estrangeiro, já que as emissões são todas produzidas por falantes não nativos.

Na verdade, vários aspectos adicionais de variação entoacional devem ser levados em consideração, visto que a variação não afeta somente a escolha e distribuição de categorias tonais, mas também a realização fonética da categoria<sup>4</sup>. Variações de uma realização tida como protótipo que são comprovadas por medidas podem não ser percebidas e, por essa razão, ser irrelevantes; ou podem ser percebidas como diferentes, mas não interpretadas como tal; ou, ainda, podem ser interpretadas como diferentes, mas não como estrangeiras.

Tal fato explica-se porque categorias entoacionais podem ter mais realizações fonéticas diferentes dependentes de contexto (alotons) do que as segmentais, aumentando a dificuldade de identificar o sotaque estrangeiro quanto aos aspectos referentes à entoação, embora muitas dessas realizações fonéticas adicionais não contribuam para que ele ocorra. Como variação desse tipo instaura-se tanto com falantes nativos quanto com não nativos, considera-se que o sotaque estrangeiro pode estar presente nas microvariações quando o falante não nativo apresenta um maior desvio-padrão das formas prototípicas.

Enquanto as diferenças na variação podem ser observadas facilmente em termos de medidas acústicas, o efeito perceptual das variações individuais é difícil de ser confirmado. As impressões de sotaque estrangeiro nos ouvintes provavelmente serão desencadeadas pelo acúmulo de pequenos desvios individuais, incluindo os segmentais. Em outras palavras, um único desvio tonal pode ser considerado normal, mas no contexto com outros fenômenos da mesma

---

<sup>4</sup> Para discussão sobre a variação na realização fonética dos segmentos sob diferentes condições prosódicas, veja-se Albano, Barbosa, Gama Rossi, Madureira e Silva, (1998).

natureza, um efeito acumulativo de sotaque estrangeiro pode ser criado. Essa pode ser uma das razões porque o sotaque estrangeiro, em relação aos aspectos entoacionais, seja percebido, como erro, de maneira menos evidente do que ocorre com o sotaque estrangeiro segmental.

Além de afetar a realização de categorias tonais, o fenômeno acumulativo pode ter um efeito no nível da colocação do acento<sup>5</sup>. Do mesmo modo, um ou vários desvios de colocação de acento e/ou escolha inapropriada de acentos de *pitch* podem não ser percebidos como sotaque estrangeiro devido à possibilidade de interpretações diferentes, mas a combinação de associações incompatíveis com certos significados discursivos criará também essa impressão.

---

<sup>5</sup> No exame de qualificação deste trabalho, Gama-Rossi relatou sua experiência com a palavra *temperature* do inglês. Por não saber colocar o acento lexical adequadamente, tinha dificuldade em se fazer compreender em uma situação em que o termo era largamente empregado pelo fato da temperatura estar a – de 30º Cº

## **CAPÍTULO 3 - MODELO DE DESCRIÇÃO ENTOACIONAL**

Na história da pesquisa entoacional há vários modelos de descrição e representação entoacional que competem entre si e que influem na nossa interpretação de quais sejam as unidades entoacionais relevantes do ponto de vista da percepção.

Assim, na abordagem do ***Tune Approach*** (Jones, 1918; Armstrong & Ward, 1926), que considera apenas dois tons básicos para dar conta de todas as formas possíveis de contornos entoacionais, o contorno do movimento do *pitch* é visto como um bloco, sem nenhuma estrutura interna ou hierarquia de elementos discretos.

Mais tarde a escola britânica da análise da entoação desenvolveu uma abordagem mais complexa segundo a qual o contorno é decomposto em grupos tonais cujo inventário inclui 10 tipos diferentes (O'Connor & Arnold, 1973). No grupo tonal, a sílaba acentuada da última palavra acentuada carrega o movimento de *pitch*, o chamado núcleo, sendo identificados 7 tipos de tons nucleares. Opcionalmente, o núcleo pode ser precedido de um pré-núcleo (*pre head*) ou seguido de coda (*tail*) de movimento tonal.

Em contraste, a escola americana de análise da entoação (Bloomfield, 1933) adotou uma abordagem fonêmica ao fenômeno da prosódia, a qual resultou na postulação de fonemas de *pitch* e de acento (*stress*). Desenvolvimentos posteriores (Trager & Smith, 1951) acrescentaram níveis hierárquicos.

Elementos importantes da escola americana, tal como a adoção de níveis de *pitch* distintos assim como unidades discretas na ordem linear, ainda estão na base

de abordagens mais recentes como a **Teoria Autosegmental** (Goldsmith, 1976) e também no modelo de **Seqüências Tonais** (Pierrehumbert, 1979, 1980).

A escolha da teoria de representação entoacional, segundo a qual os contornos entoacionais são examinados, é importante na medida em que modela nossa percepção de suas unidades relevantes. Dentre os modelos existentes, optou-se pelo modelo de seqüência tonal cujo sistema se aplica para a descrição do inglês e do português.

A abordagem de seqüências tonais foi preferida para apoiar as análises por várias razões. Apresenta maior compatibilidade com teorias de aquisição de segunda língua que postulam categorias. É um princípio comum nas ciências cognitivas considerar que as representações perceptuais em geral, não apenas a percepção da fala ou auditiva, sejam categoriais. Desse modo, a representação da entoação proposta compara-se à representação de outros elementos da fala, de modo a permitir que a entoação seja interpretada dentro de um nível mais alto comum de representação da percepção da fala. Além disso, as unidades entoacionais definidas em seu sistema de representação entoacional refletem relevância perceptual em termos de associação com conteúdos semânticos e pragmáticos específicos e refletem relevância perceptual em aspectos significativos dos contrastes do português e inglês.

### **3.1 Modelo de Seqüências Tonais**

O modelo de Seqüências Tonais para a descrição e geração de entoação

(Pierrehumbert, 1979 e 1980) foi desenvolvido com o objetivo primeiro de ser um modelo de geração de entoação. No entanto, ele apresenta, também, os princípios básicos correspondentes para uma descrição entoacional.

A versão do modelo de seqüência tonal de Pierrehumbert segue uma linha de representação entoacional que começou com Pike (1945). Baseia-se na idéia de que, do mesmo modo que se consideram as variações contínuas de ressonâncias do trato vocal como a implementação de uma cadeia de segmentos discretos da fala, considera-se a variação contínua no contorno de F0 como a implementação de elementos discretos no nível supra-segmental (Pierrehumbert, 1981). De acordo com essa visão, os contornos de F0 são analisados como uma série de valores alvos conectados por funções transicionais.

Na análise de Pierrehumbert da entoação do inglês (1980), os contornos tonais frasais são decompostos em uma série de alvos que estão associados com as sílabas tônicas (correspondendo aos *pitch accents*, ou acento de *pitch*) ou com as margens da frase entoacional (correspondendo aos tons de fronteira prosódica).

Os alvos são definidos como *High* ou *Low*<sup>2</sup> em uma área delimitada por linha abstrata alta ou superior e uma linha de base, inferior, que limitam a extensão do *pitch* do falante. Assim, um contorno entoacional é descrito como uma seqüência de alvos *High (H)* e *Low (L)* e a extensão do *pitch (pitch range)* é o espaço compreendido entre o valor de F0 do alvo mais baixo (L) e o valor de F0 do alvo mais alto (H) em uma frase entoacional.

---

<sup>2</sup> O termo *High* refere-se a um tom alto e o termo *Low* a um tom baixo.

A incursão de um padrão específico de F0 pode assim ser controlada dependendo dos valores concretos respectivos para a linha superior e de base, dando suporte para a visão de que os mesmos padrões de *pitch*, isto é, o mesmo evento tonal, pode vir a ter diferentes realizações fonéticas em diferentes extensões de *pitch* ou contextos variados (que dependem, por exemplo, de contextos segmentais e taxa de elocução<sup>3</sup>).

A realização fonética das categorias tonais não é predeterminada dentro do modelo Seqüencial de Tons, deixando espaço para diferentes possíveis abordagens que reflitam a influência de diferentes bases teóricas na interpretação do contorno entoacional, embora em um nível fonético mais baixo.

Em resumo, pode-se afirmar que a descrição e análise da entoação baseadas na seqüência tonal é recomendável, uma vez que possui uma estrutura clara e é facilmente acessível para posterior análise e processamento lingüístico.

Portanto, a investigação dos dados que servirão de base a este trabalho vai ser conduzida segundo esse modelo. Os movimentos de F0 serão analisados através de uma abordagem fonológica orientada para categoria, i. e., um contorno entoacional será visto como uma sucessão de eventos tonais relevantes em ordem linear que correspondem a categorias tonais individuais. A diferenciação dos tipos específicos de categorias tonais será dada pelo padrão descritivo de entoação do **ToBI** (Índices de Tons e Pausas; Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert & Hirschberg, 1992) que dará rótulos tonais específicos para as categorias tonais. A grande vantagem do esquema do ToBI é

---

<sup>3</sup> A taxa de elocução refere-se ao tempo despendido na articulação de um enunciado. Diferentemente da taxa de articulação, inclui as pausas no cômputo do tempo do enunciado (Laver, 1994).

sua definição de eventos entoacionais relacionados com significados discursivos específicos, que proporcionam uma orientação valiosa na procura de padrões entoacionais que são relevantes para a percepção do sotaque estrangeiro. Em termos mais gerais, o principal princípio do modelo de seqüência tonal, que é o da análise do contorno entoacional como uma sucessão linear de eventos tonais discretos, estabelece uma estrutura para o exame do contorno geral.

Mais importante ainda, o conceito de categorias entoacionais com diferentes realizações fonéticas dependendo de contexto tonal, frasal e segmental corresponde à representação de categorias fonológicas segmentais. Portanto, tanto a representação quanto a aquisição da entoação por não nativos podem ser cobertas por um modelo baseado em categorias que seja equivalente às teorias de aquisição de segunda língua orientadas aos segmentos.

### **3.2 Introdução ao TOBI**

ToBI, Índices Tonais e de Pausa, é um sistema de transcrição dos fenômenos prosódicos (Silverman et. al. 1992) concebido originalmente para ser um padrão de descrição da prosódia do inglês americano. Seu desenvolvimento foi incentivado por cientistas de fala de várias disciplinas, que desejavam um método comum e largamente aceito para representar a prosódia e, ainda, que fosse compatível com a pesquisa em áreas relacionadas, principalmente com o processamento de linguagem (síntese de fala e reconhecimento; *parsing* sintático) e com as representações formais na semântica e pragmática.

Em sintonia com o Modelo de Seqüências Tonais, os eventos de entoação

discretos são interpretados como categorias tonais. Os diferentes tipos de categorias, i. e., os movimentos tonais que elas representam, consistem de seqüências de valores alvos *High (H)* e *Low (L)* e o ToBI designa essas categorias com rótulos tonais correspondentes que também são compostos por elementos básicos *H* e *L*. Duas grandes classes de configurações tonais são distinguidas: acentos de *pitch* que descrevem o movimento de F0 específico associado com sílabas acentuadas dentro de uma frase entoacional e tons frasais que representam as configurações de F0 nas margens das frases entoacionais. Os vários contornos tonais individuais formam um inventário de acentos de *pitch* e de categorias de tom frasal.

A visão de entoação expressa pelo sistema do ToBI, orientada para categoria, tem conseqüências importantes no tratamento do fenômeno de declinação de *pitch* ao longo da frase entoacional, o qual tem sido geralmente interpretado como um processo que afeta o contorno como um todo. ToBI assume que a declinação está sob o controle consciente do falante como parte da escolha de um tipo específico de acento de *pitch* e define-o como *downstep*. Uma compressão da extensão de *pitch*, acionada fonologicamente, abaixa as metas de F0 de qualquer tom H subsequente ao acionamento do *downstep*<sup>4</sup> (Beckman & Ayers, 1994).

Assim, após o tom rebaixado, a linha superior com a qual o modelo de seqüência tonal define a margem superior da extensão de *pitch* do falante também é abaixada de modo a que todo e qualquer tom alto que se segue é colocado no novo topo de linha estabelecido. A declinação gradual no *pitch* não é interpretada como

---

<sup>4</sup> O termo *downstep*, equivalente ao termo “abaixamento” no português, será mantido neste trabalho.

um abaixamento ou estreitamento da extensão de *pitch* ao longo do curso da frase entoacional, mas como sendo causada por um evento controlado fonologicamente, associado com um evento tonal particular.

Enquanto o esquema do ToBI coloca a ênfase principal na representação da entoação e do acento em nível da camada tonal, ele também inclui uma camada de indicadores de pausa para prover uma descrição da estrutura frasal e outras camadas para comentários adicionais sobre outros aspectos prosódicos. Informação sobre aspectos duracionais, assim como a posição temporal relativa das categorias tonais rotuladas dentro dos segmentos fonéticos, é apresentada através de camadas de transcrição fonética e ortográfica.

### **Inventário de Tons do Inglês Americano**

A forma dos padrões entoacionais relevantes e, conseqüentemente, a identidade de categorias tonais é específica das línguas. Com base nessa afirmação, as categorias tonais de vários idiomas puderam ser descritas e o ToBI foi fundamental nesse processo. Tal sistema, como se sabe, foi criado para elaborar o inventário do inglês americano (Beckman & Hirschberg, 1994), o qual será descrito a seguir, tendo em vista sua relevância para este trabalho.

O inventário do inglês americano contém 5 tipos de acento de *pitch*, dois dos quais são monotons, H\* e L\*, sendo que o símbolo (\*) indica que o alvo, assim marcado, está associado com a sílaba acentuada. Os outros 3 são bitonais e implicam em um movimento inerente de F0 (ascendente ou descendente) entre os dois alvos:

L\*+ H

L + H\*

H + H\*

Os tons L\*+ H e L + H\* são diferentes no que se refere ao tempo do movimento ascendente. Assim, em L\*+ H, o movimento ascendente começa na sílaba acentuada e, em L + H\*, termina na sílaba acentuada.

H+!H\* representa uma queda pequena de um High para um alvo High com *downstep* representado pelo símbolo (!). Esse símbolo substituiu o símbolo (L\*) original de Pierrehumbert (1980), porque o tom com (\*) está na verdade alocado na metade superior da extensão do *pitch* e não na metade inferior como (L\*) sugeria. Do mesmo modo, o (H\* + L) foi substituído pela combinação de acento de *pitch* (H\*) regular seguido de um acento de *pitch* mono tom e com *downstep* (H\* ...! H\*).

A análise do contorno entoacional baseia-se em uma hierarquia de frases entoacionais (IP's) que contém uma ou mais frases intermediárias. Três tipos de eventos tonais são marcados:

- 1- Um tom de fronteira H% ou L% marca o final de cada frase entoacional (IP).
- 2- Um tom frasal H- ou L- é acrescentado após o último acento de *pitch* da frase intermediária (ip). Este tom preenche o espaço até o final da frase intermediária ou até o tom de fronteira.
- 3- Acentos de *pitch* estão associados com as sílabas acentuadas das palavras proeminentes. Deve haver pelo menos um acento de *pitch* na palavra mais proeminente em cada frase intermediária.

Desse modo, o fraseamento no inglês americano é determinado por duas unidades de nível mais alto, as frases intermediárias (ip's) e as frases entoacionais (IP's). Os tons frasais alto (H) ou baixo (L), na extensão do *pitch* do falante, marcam o final dessas frases. Para as frases intermediárias eles são chamados *phrase accents* (H-, L-) e, para frases entoacionais, o termo tom de fronteira é empregado (H%, L%).

Como sugerido pela terminologia, as ips e as IP's são ordenadas hierarquicamente. Uma IP consiste de uma ou mais ip's, e, uma ou mais IP's faz uma emissão. Por essa razão, o final de uma IP é por definição o final de uma ip e um tom de fronteira é sempre acompanhado por um acento frasal permitindo 4 possíveis combinações : L-L%, L-H%, H-L% e H-H%.

Os começos de ip's e IP's não costumam ser marcados. Há, contudo, um tom de fronteira inicial alto (%H) no inglês americano, usado em casos raros onde uma IP começa alta na extensão de pitch de um falante. Esse fato não pode ser atribuído a um acento de *pitch* alto na primeira sílaba da emissão. A realização fonética de começo de frase não é discutida mais no ToBI.

Categorias tonais individuais ou combinações, daqui para frente denominadas de tons (normalmente uma configuração de fronteira precedida por um acento de *pitch*), podem expressar significados discursivos específicos, uma vez que a descrição tonal proporcionada pelo esquema do ToBI relaciona diretamente padrões entoacionais com estrutura semântica. A

complexidade da configuração tonal é refletida no significado discursivo expresso conforme descrito pelo modelo de Pierrehumbert e Hirschberg (1990) de uma abordagem composicional ao significado do *tune*.

Resumidamente, essa abordagem associa os alvos tonais subjacentes, H e L do modelo de sequência tonal com conceitos básicos de estrutura de informação, e relaciona *tunes*, que são composições complexas desses elementos básicos, com significados discursivos complexos correspondentes.

Vários *tunes* particulares, que envolvem as 4 configurações básicas de tom de fronteira precedidas de acento de *pitch* nuclear, estão associados com certas situações discursivas bem definidas, facilitando, assim, consideravelmente o exame das diferenças entre falantes nativos e não nativos na realização entoacional dessas situações discursivas. Daí as transcrições utilizadas neste estudo concentrarem-se nos acentos de *pitch* e nos tons de fronteira. As pausas só serão mencionadas quando relevantes para a análise.

No inglês americano, por exemplo, o *tune* (H\*) L-L% é característico de declarativas e de *Wh\_questions*, enquanto o (H\*) H-H% e (L\*) H-H% ocorrem em *Yes/No questions*. A configuração H-L% cria um plateau (usado, por exemplo, como um contorno de chamamento), e o (H\*) L-H% assinala uma ascendente de continuidade.

No português, essas configurações de fronteira parecem ser mais simples, mas não essencialmente contornos diferentes. Assim, para as

emissões declarativas, teremos movimento descendente com (H\*) L-L% e, para *Yes/No questions*, o movimento será ascendente (H\*) H-H% como no inglês americano, com a diferença de que, no português, a extensão do *pitch* ascendente e descendente se apresenta menor do que no inglês.

A descrição particular de uma língua com base numa certa visão teórica de sua estrutura já constitui uma avaliação de suas propriedades relevantes. Desse modo, o método de descrição entoacional que será usado nesse estudo implica em que certas diferenças já estão pressupostas pela natureza das categorias tonais específicas da língua conforme são dadas pelo inventário de tons respectivos de cada uma das línguas.

No entanto, não temos à disposição uma descrição completa do ToBI para o português, da mesma maneira que temos para várias línguas como alemão, espanhol, holandês, entre outras, que possibilite a comparação do inventário de tons do inglês e português e de regras de realização detalhadas das categorias tonais em ambas as línguas de modo a identificar diferenças fonéticas específicas e prever causas possíveis de sotaque estrangeiro entoacional.

A descrição do inglês será, então, a base para a análise das emissões dos não nativos.

## **CAPÍTULO 4 - TECNOLOGIA DE FALA E ENSINO DE ENTOÇÃO**

#### 4.1 Ensino de entoação

Segundo Beckman (1995), o ensino de entoação é um dos aspectos de fala mais difícil de ser ensinado por ser difícil a modulação de sua estrutura. Uma das dificuldades diz respeito aos vários tipos de significados expressos por muitas das categorias entoacionais, que vão desde a sinalização de diferenças lexicais e categoriais até a sinalização de relacionamentos sutis entre uma emissão e sua inserção em um contexto maior do discurso. Assim, por exemplo, o sistema entoacional do inglês americano apresenta um contraste categorial entre acentos de *pitch* de tom alto (H\*) e de tom baixo (L\*) que, em diferentes contextos discursivos, pode ser interpretado como o elemento diferenciador entre uma afirmação e uma pergunta do tipo falsa ou verdadeira, entre uma informação nova e uma retomada de informação já dada, entre uma repetição para confirmar ou expressar incredulidade sobre o que o falante acabou de emitir, ou entre o emprego literal de uma palavra como *well* e seu emprego como marcador discursivo.

Pierrehumbert & Hirschberg (1990) sugerem que o tom alto (H\*) expressa comprometimento, isto é, o falante espera que o ouvinte acrescente uma proposição ou uma entidade à informação compartilhada até aquele ponto do discurso, ao passo que o tom baixo (L\*) expressa descomprometimento, ou porque o falante espera que o ouvinte acrescente a informação correta, ou porque sabe que o ouvinte tem a informação (lembrete), ou porque questiona a intenção do ouvinte de acrescentar a informação (repetição incrédula), ou, ainda, porque pretende que a palavra seja entendida como um marcador de discurso.

Diferenças de significados como essas, sinalizadas por padrões entoacionais

distintos, são inerentemente mais difíceis de serem caracterizadas do que muitas das diferenças lexicais usadas para apoiar análises de consoantes e vogais de uma língua. Dessa forma, é mais fácil explicar a diferença entre um **viu/fio**, com base na distinção de fonemas, do que aquela existente entre **agora** (adjunto adverbial de tempo) e **agora** (conectivo de discurso), com base na morfologia entoacional<sup>5</sup>. Conseqüentemente, tais diferenças são também mais difíceis de serem ensinadas, tanto do ponto de vista da percepção quanto do da produção, mesmo para crianças adquirindo sua língua nativa (Cruttenden, 1985).

O ensino da entoação é também dificultado quando embasado em teorias fonológicas inadequadas. Para o ensino de vogais e consoantes, assume-se a análise fonológica segmental linear do alfabeto fonético internacional e concentra-se em outros aspectos, tais como em desenvolver a habilidade de ouvir e reproduzir contraste de consoantes e vogais ou em introduzir os princípios básicos da física necessários para que os alunos entendam como os contrastes são produzidos e percebidos.

Esse método, bastante difundido, apesar de ter implementado o ensino da fonética, leva o aluno a ter a ilusão de que consoantes e vogais existem na natureza já segmentadas daquela maneira, incapacitando-o de entender traços segmentais ou qualquer outro traço da estrutura do som da língua que desafie esse modelo fonológico. A gramática tradicional, quando aborda a questão da entoação, ensina apenas os contornos entoacionais associados aos sinais de ponto final, interrogação, exclamação e reticências. Por isso, quando se procura focar a

---

<sup>5</sup> Veja-se os exemplos: “Você **viu** aquilo? O menino ficou pendurado por um **fio**”. “Não quero escrever **agora**. Falar é fácil. **Agora**, escrever é que são elas”.

questão de um outro prisma, percebe-se que é fácil apresentar aos alunos a fisiologia da laringe e a relação física entre a tensão das cordas vocais e a frequência fundamental, mas que é extremamente difícil fazê-los entender as dimensões da variação entoacional que caracterizam até sua própria língua nativa, como, por exemplo, o fato de que um mesmo contorno de entoação ter funções distintas em diferentes partes do enunciado ou em enunciados de outras línguas, uma vez que essas noções têm a ver com a estrutura do som da língua e não são contempladas pela maioria dos modelos teóricos.

Uma outra dificuldade que se refere ao ensino de entoação mencionada por Beckman (1995) diz respeito à própria natureza da prosódia. Muitos dos aspectos entoacionais estão vinculados a construtos prosódicos, como a sílaba, o acento e a frase entoacional, que são, por si sós, difíceis de serem caracterizados, pois não se adequam ao modelo alfabético. A exploração de construtos entoacionais tais como acento de *pitch* e tom de fronteira, por exemplo, requer a compreensão de unidades prosódicas como a sílaba e frase entoacional. Do mesmo modo, não dá para explorar acento de *pitch* sem recorrer ao acento.

A prosódia diz respeito à estrutura que organiza a fala, por isso apresenta maior dificuldade de ser apreendida do que traços contrastivos. Pode-se focalizar traços como timbre de vogais, ponto e maneira de articulação da consoante, tipo de fonação e altura de *pitch* sem pensar como esses elementos funcionam na estrutura organizacional de uma dada língua. O problema é fazer o aluno entender o papel que eles desempenham nessa organização, bem como os aspectos que se referem à sua variação.

Contudo, não há maneira de se proceder a análise entoacional em qualquer língua sem resolver o problema da prosódia e os aprendizes devem desenvolver a habilidade de identificar os construtos prosódicos usados pelos falantes para estruturar as categorias entoacionais e outros traços contrastivos. Cabe ao professor desenvolver estratégias que os auxiliem nessa tarefa.

#### **4.2 Tecnologia de fala e ensino de entoação**

Computadores parecem oferecer uma pista a mais na captação de aspectos prosódicos para o ensino de Fonética e Fonologia. Os programas, que permitem visualizar na tela frequência fundamental, duração e intensidade do som de uma emissão, levam professores e alunos a realizarem análises de padrões de entoação, ritmo, acentuação e características fonéticas, facilitando a compreensão e percepção dos fenômenos envolvidos. Além de ser um excelente recurso didático para a exploração de conceitos complexos, tais programas podem ser utilizados como suporte para a prática de pronúncia. Os alunos têm oportunidade de gravar suas emissões, analisá-las e compará-las na tela com as emissões do modelo, recebendo *feedback* imediato e objetivo, que independe de avaliações perceptuais.

Resultados de pesquisa anterior (Rocca, 1998) apontam para as vantagens da combinação do apoio oral e visual proporcionado por um sistema computacional para o ensino da entoação. Tal abordagem chama a atenção do aluno para os padrões acústicos relevantes para o seu caso específico e vai ao encontro das características que Jamieson (1995: 101) considera essenciais para um treinamento proveitoso: “training is more likely to be effective when the training task is designed to provide prompt, unambiguous feedback concerning the L2 category appropriate to

each training token”.

Os alunos consideram o processo altamente estimulante porque o *feedback* visual é imediato e permite auto-avaliação do seu desempenho, além de orientá-los para as novas tentativas de autocorreção dos erros detectados. Computadores têm uma paciência infinita e permitem que o estudante faça tentativas intermináveis, sem demonstrar irritação com os erros. Eles também julgam de uma maneira menos ameaçadora e mais particular.

Esknazi (1996) afirma que o emprego da máquina para melhorar a pronúncia é atraente por duas razões : “it affords the user more practice time than a human teacher since it is usually more available, and second, the user is not faced with the **sometimes** overwhelming problem of human judgement of his production of ‘foreign’ soundings”. Combinada com métodos tradicionais de ensino de pronúncia, a tecnologia de fala abre novas perspectivas para o ensino da entoação: o ensino mediado por computador.

## **CAPÍTULO 5 - EXPERIMENTO**

O experimento que fundamenta este trabalho foi realizado dentro de uma proposta de curso de Fonética e Fonologia para alunos universitários brasileiros. Nesse curso, privilegiou-se o ensino da entoação da língua inglesa mediado por *software* de análise acústica em laboratório de informática.

### **5.1 Objetivo**

O experimento tinha como objetivo verificar se o ensino mediado pelo computador, que, entre outras vantagens, apresenta a possibilidade do acréscimo do apoio visual às abordagens tradicionais de ensino da entoação, pode ser de alguma ajuda no desenvolvimento de um curso sobre entoação de uma língua estrangeira.

### **5.2 Sujeitos do experimento**

Os sujeitos eram alunos que estavam inscritos na disciplina **Fonética e Fonologia: Entoação do Inglês**, no segundo semestre de 2001. Com o consentimento do professor da disciplina, seis alunos foram convidados para fazer parte de um grupo experimental que cursaria a disciplina com um conteúdo programático que introduziria conceitos teóricos de Fonética Acústica apoiados em recursos de tecnologia de fala. O curso semestral, com duas horas semanais, fazia parte da grade horária, e a avaliação final foi feita com base na presença e participação dos alunos nas atividades propostas. Eles não eram, portanto, considerados voluntários.

Fizeram, também, parte do experimento dois alunos que não compunham o grupo experimental, mas que cursaram a disciplina com o restante da classe do professor. Esses alunos, chamados de grupo de controle, dispuseram-se a fazer a gravação dos mesmos dados gravados pelo grupo experimental no início e no final do curso.

Todos os alunos do experimento responderam a um questionário, que se encontra no anexo 2. A aplicação do questionário não teve o objetivo de selecionar sujeitos com habilidades e características semelhantes, como no experimento piloto, mas sim de coletar dados pessoais que pudessem vir a ser úteis na interpretação dos resultados.

Nesse questionário, procura-se obter informação sobre a idade atual e idade em que o sujeito começou a estudar a língua estrangeira, seu estágio de aprendizagem e o grau e qualidade de exposição à língua estrangeira, pois relatos de experimentos (Llisterri, 1995; Eisendoorn, 1984; Bohn & Fledge, 1990) indicam que o estágio na aquisição da L2, a experiência com a língua, o grau de exposição e a idade de aquisição desempenham um papel importante na interação entre produção e percepção em L2.

As informações coletadas pelo questionário indicam que os oitos alunos que participaram do experimento eram brasileiros, na faixa etária de 19 a 24 anos, filhos de pais também brasileiros. Todos tinham o inglês como sua primeira língua estrangeira e já haviam cursado os estágios de língua inglesa previstos no curso de Letras - Inglês da PUC/SP. Apenas um deles era do sexo masculino e apenas uma aluna havia morado durante um mês em país de língua inglesa. Três dos

participantes tinham noções bastante elementares de uma outra língua estrangeira (francês, espanhol e alemão). Somente quatro deles haviam feito outro curso de inglês além da graduação na PUC. O dado mais variável entre os participantes diz respeito à idade em que começaram a estudar a língua inglesa:

- 11 anos: 1 aluno
- 10 anos: 2 alunos
- 12 anos: 1 aluno
- 13 anos: 2 alunos
- 17 anos: 2 alunos

Os participantes permitiram, por escrito, que seus nomes próprios fossem divulgados. No entanto, optou-se pela identificação através de nomes fictícios, quando, assim, se fizesse necessária.

### **5.3 *Corpus* da pesquisa**

A escolha das emissões para compor o *corpus* do experimento baseou-se em padrões sintáticos que, nos livros de ensino de entoação para falantes não-nativos, são descritos como tendo um padrão entoacional canônico não marcado, se excluídos os fatores pragmáticos e emocionais que possam vir a interferir.

Embora a abordagem sintática (O'Connor & Arnold, 1973) tenha a vantagem de facilitar a organização dos padrões, ela não consegue dar conta da criatividade do falante que emprega a mesma estrutura com diferentes entoações para

comunicar atos de fala opostos ou emoções variadas. Mesmo assim, elegeu-se o critério sintático na composição do *corpus*, por ser ele, ainda, aquele que permite um maior controle dos dados experimentais.

O *Corpus* foi composto de emissões com as seguintes características sintáticas:

- orações declarativas afirmativas;
- orações interrogativas com pronome interrogativo do tipo *Wh*;
- orações imperativas;
- orações interrogativas do tipo *Yes/No*, isto é, interrogativas que questionam se *X* é falso ou verdadeiro.

O experimento concentrou-se em dois tons de fronteira, L% (Low) e H% (High), sendo o movimento final de *pitch* descendente representado pelas orações declarativas, interrogativas com *Wh* e as imperativas enquanto que o movimento final de *pitch* ascendente é representado pelas interrogativas do tipo *Yes/No*.

As emissões que constituíram o *Corpus* desta pesquisa faziam parte de dois tipos de gravações:

- a Gravação Estímulo, que era apresentada para o aluno como estímulo à sua produção,
- e
- a Gravação Desempenho, que apresenta os dados de produção desses alunos.

A Gravação Estímulo (anexo 3) propunha, aos alunos do experimento, 3 tipos de atividades:

1 - repetição do que é ouvido com apoio na escrita (19 emissões ao todo; 3 imperativas e 8 em situação de diálogo, das quais 4 são interrogativas com *Wh* e 4 interrogativas do tipo *Yes/No*, sendo que todas recebem como resposta 8 afirmações);

2 - repetição do que é ouvido sem apoio da escrita (9 sentenças isoladas sendo 2 imperativas, 2 afirmativas, 4 interrogativas do tipo *Yes/No*, e 1 interrogativa com pronome interrogativo do tipo *Wh*);

3 - produção espontânea induzida<sup>1</sup> (8 sentenças isoladas, sendo 2 declarativas afirmativas, 2 imperativas, 2 interrogativas do tipo *Yes/No*, e 2 interrogativas com pronome interrogativo do tipo *Wh*).

A inclusão do apoio da escrita na atividade (1) teve como objetivo excluir interferências relativas à capacidade de memória e de compreensão oral da produção dos alunos. Por outro lado, a atividade (2), de repetição apoiada só na emissão oral do falante nativo, procurou excluir interferência da escrita.

---

<sup>1</sup> Situação em que, mediante um estímulo gravado, o aluno é induzido a produzir uma emissão específica. Assim, ao ouvir a gravação "Ask me if it is going to rain", o aluno emite "Is it going to rain?"

As duas primeiras atividades foram constituídas de emissões gravadas por falantes nativos do inglês americano, extraídas de fitas cassetes originais<sup>2</sup> que acompanham o livro didático. A terceira atividade da Gravação Estímulo foi elaborada com emissões com os padrões sintáticos focalizados pela pesquisa, e foi gravada por falante bilíngüe português/ inglês em laboratório de rádio com isolamento acústico.

A Gravação Desempenho dos alunos é formada por duas versões de gravações das atividades propostas pela Gravação Estímulo, uma gravação inicial antes do treinamento e uma gravação final, após o treinamento.

As emissões que compuseram o *corpus* da Gravação Estímulo e da Gravação Desempenho dos alunos foram registradas em fita DAT com o uso do gravador digital Sony DCT A6 em modalidade de gravação estéreo. Posteriormente, foram registradas em CD ROM.

#### **5.4 Gravações**

Cada um dos sujeitos gravou o *corpus* individualmente em ambiente com isolamento acústico no laboratório de rádio da Faculdade de Comunicação e Filosofia da PUC/SP, de forma a minimizar o efeito de interferências técnicas de gravação nas análises a serem realizadas. Receberam instruções orais e escritas sobre a atividade de gravação.

---

<sup>2</sup> As fitas cassetes, das quais foram extraídas as emissões modelo, acompanham o livro didático *Sound Advantage - A Pronunciation Book* (Hagen & Grogan, 1992).

Foram instruídos de que haveria 3 tipos de atividades diferentes. Na primeira e na segunda propostas, deveriam repetir as sentenças produzidas pelos falantes nativos do inglês, procurando imitar o melhor possível a entoação do modelo. Na primeira proposta, em forma de diálogo, teriam o apoio da escrita juntamente com o estímulo auditivo para auxiliar a produção, enquanto que, na segunda atividade, em forma de sentenças isoladas, haveria apenas o estímulo auditivo orientando a repetição. A terceira atividade não era de repetição, mas de produção induzida. Deveriam emitir perguntas, comandos ou afirmações, conforme requerido pela gravação.

Foi-lhes permitido tirar dúvidas sobre o material escrito. As emissões com problemas de compreensão ou de gravação foram descartadas e re-gravadas na mesma sessão.

Após as 10 sessões de prática com material variado e diverso, os sujeitos refizeram a gravação do mesmo *corpus* proposto pela Gravação Estímulo.

## **5.5 Treinamento**

Todos os sujeitos do experimento, grupo de controle e o grupo experimental, tiveram cinco aulas introdutórias ao ensino de entoação, juntamente com os alunos do curso regular. Os dois sujeitos do grupo de controle continuaram com o curso regular e receberam o tipo de instrução de um curso de pronúncia tradicional, durante 15 semanas, com o apoio de técnicas de ensino tais como: alfabeto fonético, prática de transcrição fonética, descrição detalhada do sistema articulatório, descrição do sistema prosódico, tarefas de

reconhecimento e discriminação, exercícios de repetição de orações e de pares mínimos, prática com sentenças contextualizadas, leitura de textos curtos e diálogos, “*tongue twisters*” entre outros. Além das 15 semanas em sala de aula, treinaram a produção dos padrões entoacionais em atividade extracurricular, que consistia em 10 sessões no laboratório de línguas. Essa prática, direcionada para os padrões ascendentes e descendentes dos padrões sintáticos incluídos no experimento, baseou-se na imitação e repetição do sinal sonoro com apoio do sinal auditivo apenas.

O treinamento dos alunos do grupo experimental teve a duração de um semestre com uma aula semanal de duas horas, totalizando 15 semanas. Além das cinco aulas de introdução ao estudo da entoação que foram assistidas com os alunos do curso regular, foram previstas dez sessões no laboratório de informática e duas sessões de gravação de desempenho, uma inicial e uma final.

As aulas, que incluíam o treinamento, foram ministradas pela professora pesquisadora no Laboratório de Linguagem e Informática da Faculdade de Comunicação e Filosofia com um computador por aluno.

O treinamento foi realizado com apoio da tecnologia de fala. Os alunos empregaram o *software WinSal* da Media Enterprise que permite a análise, manipulação e visualização do sinal sonoro através de instrumentais de análise acústica da fala, para orientar sua percepção e produção durante o treinamento.

Os alunos podiam ouvir a emissão do modelo, ou parte dela, quantas vezes

quisessem, gravar a repetição, auto-avaliar suas emissões e comparar a curva da frequência fundamental das suas produções com a do modelo na tela do computador ou em gráficos impressos. Desse modo, “*feedback*” impresso, oral e visual estava disponível no treinamento.

As atividades planejadas (anexo 4) visaram desenvolver a atenção a fenômenos de coarticulação<sup>3</sup> e de interação prosódia/segmento<sup>4</sup> e procuraram dar ênfase ao ensino das habilidades necessárias para que os sujeitos fossem capazes de fazer uso do contorno de F0 como representação fonética. Para poder ler os gráficos adequadamente, precisavam aprender sobre as distorções de microprosódia, que estão relacionadas às transições articulatórias e dinâmicas de um segmento para o outro e são impostas pelo padrão vibratório das pregas vocais (Laver, 1994). No *onset* de vogais seguidas de oclusivas surdas, por exemplo, observam-se picos de F0.

Assim, segundo Mira Mateus (1990: 193), “durante a fase de vozeamento das obstruintes, é necessário relaxar as cordas vocais para que o vozeamento se mantenha o que tem como efeito baixar também a frequência fundamental. Pelo contrário, para evitar o vozeamento durante a produção das consoantes não vozeadas, a tensão das cordas vocais é aumentada e a frequência fundamental eleva-se”.

---

<sup>3</sup> A coarticulação refere-se ao fato de que um segmento fonológico não se realiza da mesma maneira em todos os contextos, mas varia suas propriedades em relação aos segmentos adjacentes ou próximos a ele. A coarticulação pode ser antecipatória (influência de um segmento posterior em um anterior no enunciado) ou perseverativa (influência de um segmento anterior em um posterior no enunciado) (Hardcastle e Hewlett, (1999); Laver, 1994).

<sup>4</sup> Para exemplos de interação prosódia/segmento, veja-se, também, Albano, Barbosa, Gama-Rossi, Madureira e Silva, (1998).

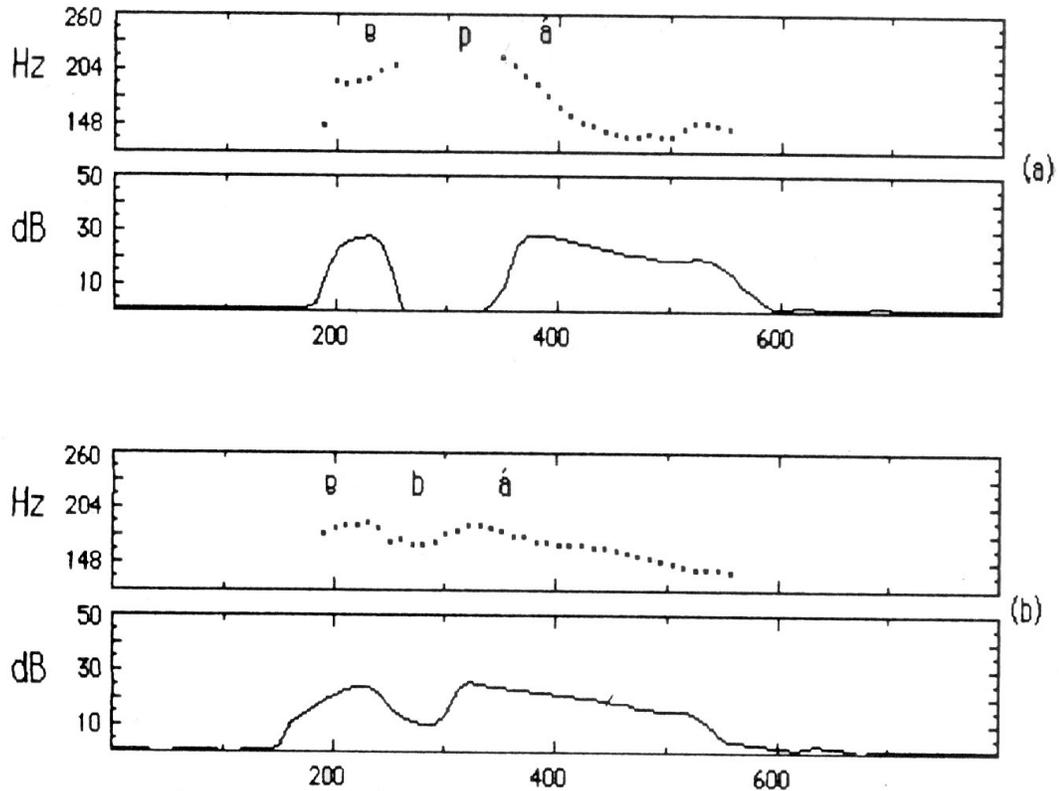


Figura1: Exemplos de evolução da freqüência fundamental entre uma consoante e a vogal seguinte em (a) apá e (b) aba extraídos de Mira Mateus (1990: 194)

Esse efeito, exemplificado na figura 1, deve-se à microprosódia e os alunos precisam aprender a não considerá-lo quando analisam os aspectos macroprosódicos como os padrões de *pitch*.

Visando atingir essas metas, foram propostos exercícios de desenhos estilizados dos contornos de *pitch* das emissões que ouviam. Esses contornos eram depois comparados com os contornos de F0 da tela do computador. Receberam, também, orientação de como medir freqüência e duração e de como usar a linha da freqüência fundamental enquanto apoio visual para suas emissões e autocorreções.

Foi preciso ensinar sobre a natureza da prosódia e dar-lhes os instrumentos para descobrir os construtos prosódicos que o falante nativo usa para estruturar as categorias entoacionais e outros traços contrastivos, procurando desenvolver a intuição sobre construtos prosódicos relevantes, tais como quantas sílabas há na palavra e quantas frases entoacionais há em uma emissão longa.

Focalizando-se os fenômenos de coarticulação e interação prosódia/segmento, foram propostas atividades de análise de eventos fonéticos em fronteiras prosódicas, reduções de segmentos vocálicos, ressilabificação, identificação do acento frasal e as implicações semântico-pragmáticas decorrentes da localização desse acento.

## **5.6 Procedimentos de análise dos dados**

As Gravações de Desempenho iniciais e finais dos sujeitos foram submetidas a dois tipos de análises, análise perceptual e análise acústica dos dados.

A análise perceptual foi feita através de um teste de percepção por parte de falantes nativos residentes nos Estados Unidos. Tal teste fez-se necessário para se verificar se, após o treinamento, houve uma melhora na produção dos alunos que fosse perceptível aos ouvidos dos falantes nativos.

A análise acústica das orações modelo e das gravações desempenho dos alunos foi introduzida para se investigar características de F0 e duração e para subsidiar a interpretação dos resultados da avaliação perceptiva dos falantes nativos.

#### 5.6.1 Análise perceptiva dos dados

A avaliação perceptiva tinha como objetivo responder às seguintes questões:

1 - Os falantes nativos do inglês avaliariam da mesma maneira as emissões referentes à gravação inicial e final dos alunos?

2 - Haveria diferença entre os dois treinamentos no que se refere aos resultados obtidos pelos alunos?

##### 5.6.1.1 Avaliadores do teste de percepção

Participaram do teste de percepção 40 falantes nativos do inglês americano, leigos no assunto, de ambos os sexos, na faixa etária de 15 a 64 anos, de regiões variadas dos Estados Unidos, a maioria residentes nos estados de Nova York, Califórnia, Carolina do Norte, Massachusetts, Washington, Florida e Ohio. Os informantes foram convidados a participar da pesquisa por e-mail. Não foram aceitos falantes bilíngües do português e inglês, filhos de brasileiros morando nos Estados Unidos e professores de línguas em geral. Os voluntários avaliadores do teste de percepção preencheram ficha com dados pessoais, cujo modelo encontra-se no anexo 7. Os dados coletados nessas fichas encontram-se na tabela abaixo:

Tabela 1: Avaliadores do Teste de Percepção

Falante Nativos	Inglês Americano	40	
Faixa de Idade	de 15 a 20 anos	1	40
	de 21 a 30 anos	6	
	de 31 a 40 anos	11	
	de 41 a 50 anos	11	
	de 51 a 60 anos	8	
	de 61 a 64 anos	3	
Sexo	Masculino	25	40
	Feminino	15	
Falantes de Outras Línguas	Não	24	40
	Sim	16	
Falantes de Inglês e de	Espanhol	7	16
	Chinês	3	
	Francês	1	
	Italiano	1	
	Alemão	1	
	Chinês e Francês	1	
	Chinês e Espanhol	1	
	Urdu e Punjabi	1	
Língua Nativa dos Pais	Inglês	34	40
	Chinês	2	
	Italiano	1	
	Alemão	1	
	Inglês/ Espanhol	1	
	Urdu / Punjabi	1	

### 5.6.1.2 Procedimentos na elaboração do teste de percepção

#### *O Corpus da Avaliação*

As frases escolhidas para o teste de percepção (anexo 8) foram as mesmas para todos os sujeitos e para as duas gravações, sendo 4 emissões resultantes da repetição com apoio da escrita, 4 repetições sem apoio da escrita e 4 produções induzidas. A seleção tentou incluir um exemplo de cada padrão gramatical que fez parte do experimento para cada tipo de proposta de atividade, isto é, uma sentença *Wh*, uma pergunta *Yes/No*, uma declarativa e uma imperativa. No entanto, para a atividade de repetição sem apoio visual, foi necessária a substituição da emissão com *Wh* por uma segunda pergunta com *Yes/No*, porque procurou-se construir o *corpus* com produções que não apresentassem problemas graves na pronúncia dos segmentos<sup>6</sup> e/ou que não fossem truncados por titubeios, de modo a minimizar a interferência dessas ocorrências no julgamento dos nativos. Todas as emissões selecionadas para o teste de percepção dos falantes nativos foram equalizadas no laboratório de rádio a 0 decibel.

Foram montados 4 blocos de dados, sendo cada um dos blocos formado por 50 arquivos de som, distribuídos da seguinte maneira: 12 arquivos da gravação inicial e 12 arquivos da gravação final de dois dos alunos, totalizando 48 emissões, às quais foram acrescentadas 2 emissões do modelo, selecionadas das gravações dos falantes nativos e incluídas no teste na tentativa de medir o grau de rigor do julgamento dos nativos. As emissões foram apresentadas aleatoriamente para serem avaliadas pelos voluntários.

---

<sup>6</sup> Problemas graves na pronúncia de segmentos, tais como, por exemplo, “th” sendo pronunciado como [f] e não [ð].

Cada um dos blocos de dados foi acessado, via Internet, por 10 falantes nativos. Desse modo, a produção de cada sujeito do experimento foi apreciada por 10 avaliadores.

### *O Sistema do Teste de Percepção*

Para o desenvolvimento do teste de percepção foi utilizado o banco de dados MySQL (Sequence Query Language) e a linguagem PHP na sua construção. O sistema, sintetizado no quadro abaixo, composto de banco de dados e aplicações, é formado de dois módulos: Módulo de Manutenção e Módulo de Teste que alimentam os Dados Estatísticos.

***Sistema de Teste de Percepção***

Manutenção das Tabelas

Áreas de Testes

Estatísticas Geral

Estatísticas por Gravação

Please rate the degree of foreign accent by clicking on a number of the scale (0=no foreign accent at all/ 6=very strong foreign accent)

Estatísticas por Gravação

Please rate the degree of foreign accent by clicking on a number of the scale (0=no foreign accent at all/ 6=very strong foreign accent)

O módulo de Manutenção destina-se à alimentação dos diversos cadastros e tabelas relacionais do banco de dados, que estão exemplificados nas tabelas abaixo.

<p style="text-align: center;"><b>Sistema de Teste de Percepção</b></p> <p style="text-align: center;"><b>Manutenção das Tabelas</b></p> <p><a href="#"><u>Cadastro de Falantes</u></a></p> <p><a href="#"><u>Cadastro de Sentenças</u></a></p> <p><a href="#"><u>Cadastro de Falantes x Sentenças</u></a></p> <p><a href="#"><u>Cadastro de Falantes x Avaliador</u></a></p> <p><a href="#"><u>Tabela de Testes</u></a></p> <p><a href="#"><u>Gerenciamento de Arquivos</u></a></p>
--

Na tabela 2 abaixo, relativa ao cadastro de falantes, foram enumerados todos os participantes do teste de percepção. Os alunos e modelo, que seriam julgados pelos falantes nativos, foram identificados como tipo A(luno) e os avaliadores como tipo N(ativo). Uma vez que os avaliadores eram desconhecidos no momento da formação do banco de dados, foi-lhes conferido um nome de usuário aleatório (inform1) para fazer o cruzamento com a senha de acesso (1inform). Assim, **inform1** recebia a senha **1inform** e o sistema gerava automaticamente um número de código, exemplificado pelo número 56 na tabela abaixo. No momento em que o voluntário, devidamente autorizado e reconhecido pela senha de acesso, entrava na área de teste e preenchia a ficha de identificação, seu nome real era registrado na tabela de cadastro de falantes, como o de Linda Rottenberg na tabela 2. Selecionando-se **editar**, tinha-se acesso aos dados dos informantes, ao passo que **apagar** excluía o voluntário do processo.

**TABELA 2. CADASTRO DE FALANTES**

Tipo	Código	Acesso	Nome	[INCLUIR]	
A	5	Ana	Ana	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
A	13	Modelo	Modelo	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
N	56	Inform1	Linda Rottenberg	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>

Na tabela 3 foram cadastradas todas as emissões selecionadas para o teste de percepção, sendo que o código as identificava. Assim, **1f** refere-se à sexta emissão do exercício 1, no qual a repetição era baseada na escrita.

**TABELA 3 .CADASTRO DE SENTENÇAS**

Código	Sentença	[INCLUIR]	
1f	How often does your mother call?	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
1h	Close your eyes, please.	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>

Na tabela 4, cadastrou-se os arquivos de som da Gravação Desempenho de cada um dos sujeitos, identificando-os como referentes à gravação inicial (1) ou à gravação final (2).

**TABELA 4. CADASTRO DE FALANTES x SENTENÇAS**

Nº	Falante	Gravação	Sentenças	[INCLUIR]	
1	1-Leila	1	Leila11f.wav	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
2	1- Leila	1	Leila11h.wav	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
3	1- Leila	1	Leila11q.wav	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
4	1- Leila	1	Leila11r.wav	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
5	1- Leila	1	Leila21a.wav	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>

Na tabela 5, fez-se o cruzamento dos dados, estabelecendo-se, por exemplo, que o avaliador que havia recebido o número de código 56 avaliaria os arquivos de som da Gravação 1 e 2 dos sujeitos Leila e Paula.

**TABELA 5. CADASTRO DE FALANTES x AVALIADOR**

Nº	Falante	Avaliador	Gravação	[INCLUIR]	
9	1- Leila	56-Linda Rottenberg	1	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
10	1- Leila	56-Linda Rottenberg	2	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
14	1- Leila	12-Greg Davenport	1	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
15	1- Leila	12-Greg Davenport	2	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
11	2-Paula	56-Linda Rottenberg	1	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>
12	2-Paula	56-Linda Rottenberg	2	<a href="#">EDITAR</a>	<a href="#">APAGAR</a>

A Tabela de Teste registrava os dados das respostas dos informantes, em tempo real, conforme essas iam sendo enviadas, de modo que era possível verificar o andamento do processo e identificar qualquer problema, tal como o registro duplicado de uma avaliação. Essa falha, que ocorreu algumas vezes quando o falante nativo teve problemas de conexão, pode ser facilmente localizada e o registro duplicado apagado. Então, na tabela 6, temos a informação que a emissão 2a (primeira emissão do exercício 2) da segunda gravação do sujeito Leila foi avaliada pelo informante identificado pelo número de código 18, no dia 16 de fevereiro de 2003, às 16 horas, 10 minutos e 51 segundos; o avaliador conferiu grau 4 para a primeira pergunta do teste, e grau 3 para a segunda.

**TABELA 6. TABELA DE TESTE**

Nº	Falante	Sentença	Gravação	Avaliador	AV-1	AV-2	Data-Hora	Incluir	
888	1-Leila	2a	2	18	4	3	16-02-03 16:10:51	Editar	Apagar

Somente o programador tinha acesso ao Gerenciamento de Arquivos, área destinada à construção do programa em si.

O Módulo de Teste, área de realização da avaliação, era o único módulo que podia ser acessado pelos participantes da pesquisa.

#### 5.6.1.3 Procedimentos na aplicação do teste de percepção

Os falantes nativos foram convidados através de e-mail a participar da pesquisa. Uma vez aceito o convite, receberam o endereço do “*site*” onde o programa de avaliação estava hospedado, juntamente com um nome de usuário e uma senha, específicos para cada um dos participantes. Receberam, também, informações gerais sobre o programa de avaliação, que os alertava das dificuldades que podiam encontrar com conexão de baixa velocidade. Além disso, foram avisados de que as emissões foram gravadas mais de uma vez pelos falantes, de modo que não estranhassem quando percebessem que as orações repetiam se. No entanto, não tinham conhecimento de que as emissões repetidas eram relativas a dois estágios diversos no treinamento dos alunos.

Ao conectarem-se ao “*site*”, abria-se a tela de

**ÁREA DE TESTE**

**Perception Test**

User Name:

Password:

Digitando-se o nome do usuário e a senha, eram levados a uma tela que deveria ser preenchida com os dados pessoais (anexo 7) que, ao serem enviados ao banco de dados, remetia os avaliadores à tela abaixo:

PERCEPTION TEST

Hello, Welcome!

You are going to listen to 50 English sentences produced by Brazilian learners of English.  
I need your help to evaluate them.

You can replay the sentences as many times you wish, but, after giving them a score, you won't be able to return to them any more.

The task will take you about one hour.  
You can interrupt it at any moment.

When you return to finish it, the program will start at the sentence you have stopped at.  
If the sentences are reproduced with interruptions, play them again till the problem is solved.  
I'd like to thank you for your valuable help.

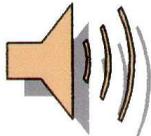
Paulina Rocca  
Pontifícia Universidade Católica de São Paulo – Brasil  
[paulina@pucsp.br](mailto:paulina@pucsp.br)

Start the test

O teste de percepção propriamente dito tinha a seguinte forma na tela do computador:

## PERCEPTION TEST

Informer: 9 –Paulina



Please, rate the degree of foreign accent by clicking on a number of the scale  
(0=no foreign accent at all/ 6=very strong foreign accent)

0 1 2 3 4 5 6

No Foreign Accent Very Strong Foreign Accent

Does the speaker sound natural as he asks a question or makes a statement?  
(0 = natural/ 6 = not natural)

0 1 2 3 4 5 6

Natural Not Natural

**Next Sentence**

**N. 0/50 ... 2003**

O avaliador só era capaz de mudar o arquivo de som e prosseguir na tarefa se houvesse completado a avaliação da emissão anterior. Caso não registrasse sua

avaliação para as duas questões, era avisado de que deveria fazê-lo e remetido à questão incompleta. A notificação do término da avaliação era acompanhada de agradecimento pela cooperação.

Foram contatados mais de 70 voluntários, mas, conforme planejado, apenas 40 deles foram incluídos na pesquisa. A coleta demorou cerca de três meses e foi dada como encerrada quando foram obtidas as 10 primeiras avaliações para cada um dos alunos. As demais avaliações recebidas foram desconsideradas. Os dados foram armazenados no banco de dados e as informações disponibilizadas nas seguintes telas:

1 –

<b>Total geral</b> Todas as Avaliações.	0	1	2	3	4	5	6	Total
--	---	---	---	---	---	---	---	-------

2 –

Total geral por Avaliador Todas as Avaliações	0	1	2	3	4	5	6	Total
--	---	---	---	---	---	---	---	-------

3 –

Total geral <b>por Sentença</b> Todas as Avaliações.	0	1	2	3	4	5	6	Total
---	---	---	---	---	---	---	---	-------

4 –

Total geral <b>por Falante</b> Todas as Avaliações.	0	1	2	3	4	5	6	Total
--	---	---	---	---	---	---	---	-------

As telas de 5 a 10 geram resultados em separado para as duas questões do experimento além de fazerem a separação entre as gravações 1 e 2.



#### 5.6.1.4 Análise estatística dos dados de percepção

Os dados coletados foram submetidos a tratamento estatístico, empregando-se o programa *SPSS - Statistical Package for Social Sciences for Windows - Standard Version*.

Foram aplicados os seguintes testes:

- teste de comparação de duas médias para amostras independentes, sendo consideradas variáveis independentes o grupo, tipo de exercício, tipo de sentença e sujeito;

- teste de comparação de duas médias para amostras dependentes, sendo a variável dependente o *Score* da avaliação anterior e da posterior;

- comparação múltipla de médias - teste de Scheffé.

No tratamento estatístico dos dados, foram empregadas as seguintes abreviações:

Grav1 = avaliações obtidas sobre as emissões da Gravação Desempenho, gravadas antes do treinamento;

Grav2 = avaliações obtidas sobre as emissões da Gravação Desempenho, gravadas depois do treinamento;

Grupol = grupo de alunos que fez o treinamento com o auxílio do computador;

Grupo2 = grupo de alunos que fez o treinamento no laboratório de línguas;

Scorel = respostas relativas à questão I do teste de percepção (***Rate the degree of foreign accent by clicking on a number of the scale***);

Score2 = respostas relativas à questão 2 do teste de percepção (***Does the speaker sound natural as he asks a question or makes a statement?***).

Score Total = soma das respostas obtidas no score1 e no score2;

Exercício1 = repetição das emissões do nativo com apoio da escrita;

Exercício2 = repetição das emissões do nativo sem apoio da escrita;

Exercício3= produção induzida sem apoio da escrita.

Apesar de terem sido submetidos à análise estatística, os dados identificados como Score2, relativos à segunda questão, foram desprezados, devido a vários depoimentos de falantes nativos dizendo da dificuldade em fazer distinção entre a questão 1, relativa ao sotaque, ***Please, rate the degree of foreign accent by clicking on a number of the scale*** e a questão 2, ***Does the speaker sound natural as he asks a question or makes a statement?***, direcionada ao contorno entoacional. Como consequência, o Score Total também foi desconsiderado nas análises dos dados estatísticos.

Optou-se por manter as respostas relativas a questão 1 por ser a avaliação perceptiva do falante nativo mais espontânea, uma vez que aparecia em primeiro lugar na tela do computador. A segunda questão levava-o a ponderar sobre o que pudesse ser a diferença entre **sotaque** e **soar natural**. Os comentários que foram enviados espontaneamente pelos voluntários<sup>7</sup> indicam ser praticamente impossível para o leigo separar o sotaque de origem segmental do sotaque entoacional com base na percepção auditiva apenas. Para se ter um resultado mais confiável, a influência do sotaque segmental teria de ser anulada através de um experimento baseado em síntese de fala que focalizasse a entoação, como o apresentado por

---

<sup>7</sup> Tais comentários podem ser sintetizados em duas afirmações: “Não consigo perceber a diferença entre a primeira e a segunda questão, portanto, vou dar a mesma avaliação para as duas questões”. “É muito difícil distinguir a primeira pergunta da segunda. Não tenho certeza de estar avaliando corretamente”.

Jilka (2000).

Desse modo, os resultados que serão apresentados dizem respeito apenas à comparação dos resultados do Scorel da Grav1 com os resultados da Grav2:

1 - do conjunto de alunos do grupo1,

e

- do conjunto de alunos do grupo2;

2 - do conjunto de alunos do grupo1 e em relação a cada tipo de exercício, separadamente: exercício1, exercício2 e exercício3,

e

- do conjunto de alunos do grupo2 e em relação a cada tipo de exercício, separadamente: exercício1, exercício2 e exercício3;

3 - do conjunto de alunos do grupo1 e em relação a cada tipo de sentença: *wh*, *yes/no*, imperativa e declarativa,

e

- do conjunto de alunos do grupo2 e em relação a cada tipo de sentença: *wh*, *yes/no*, imperativa e declarativa.

Fez-se, também, uma análise estatística com as variáveis apresentadas acima, de 1 a 3, considerando-se os resultados individuais de cada sujeito do experimento.

## 5.6.2 Análise acústica dos dados

A análise acústica teve como objetivo inspecionar as emissões dos alunos que, segundo a percepção dos falantes nativos, melhoraram seu desempenho após o treinamento, na tentativa de se estabelecer uma relação entre a percepção e as características acústicas da produção. Assim, somente as emissões cujo resultado da análise estatística apresentou diferença significativa na comparação da avaliação perceptiva feita pelos falantes nativos entre gravação 1 e gravação 2 foram submetidas à análise acústica. Os dados foram digitalizados com frequência de amostragem de 22kHz em um *PC Intel, Pentium*. A análise empregou o *CSRE 45 (Canadian Speech Research Environment)* da *Avaaz Corporation*.

Padrões de frequência fundamental dos dados das emissões dos falantes nativos da Gravação Estímulo e das repetições iniciais e finais dos dois grupos de sujeitos consideradas relevantes foram extraídos e comparados. Para extração do contorno de frequência fundamental das emissões, foi usado o módulo de análise do *CSRE* com o algoritmo *Cepstrum* com os seguintes parâmetros:

*Window Size* – 512

*Overlap %* - 50

*Silence Threshold (db)* - 5

*Zero Threshold (HZ)* - 1500

*Upper Calc. Freq (Hz)* – 400

*Sample Frequency* – 22.05 KHz

O algoritmo *Comb*, empregado somente para confirmação de valores não claramente identificados pelo sistema *Cepstrum*, apresentava os seguintes parâmetros:

*Low Pass Cut off (Hz) – 1000*

*Wind – 128 pts*

Foram extraídos os valores de F0 do núcleo da vogal de cada sílaba das emissões, sendo que, na sílaba em que se dá o movimento de *pitch* que sinaliza a fronteira final do enunciado, foram tiradas duas medidas, uma no início e outra no final do movimento, tal como exemplificado no gráfico 1 abaixo.

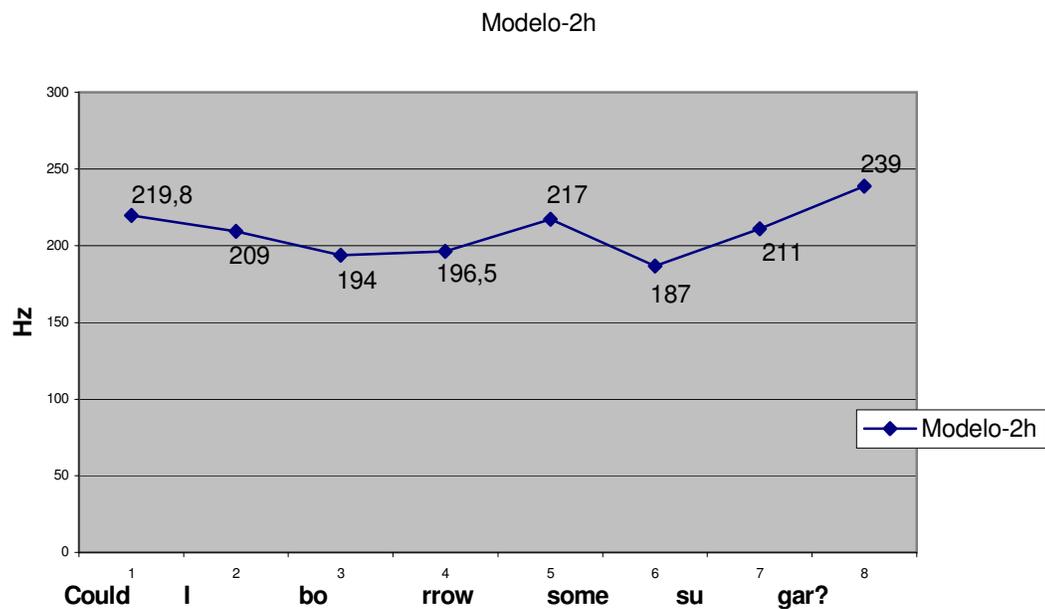


GRÁFICO 1: Valores de F0 do núcleo da vogal de cada sílaba da emissão modelo "Could I borrow some sugar?"

As medidas de duração dos dados da emissão modelo e das Grav1 e Grav2 das emissões dos alunos foram feitas na forma da onda. Utilizou-se o espectrograma de banda larga, sincronizado com a forma da onda, na tela do computador, como referência para o estabelecimento dos limites das unidades fonéticas. As características espectrais foram geradas a partir das seguintes especificações:

*Process – AC/LPC*

*Wind ( tamanho da janela) – 512 pts*

*Bands* ( número de bandas ) - 256

*Overlapping* – 50%

*Function* – *Hanning*

*Sample Frequency* – 22.05 KHz

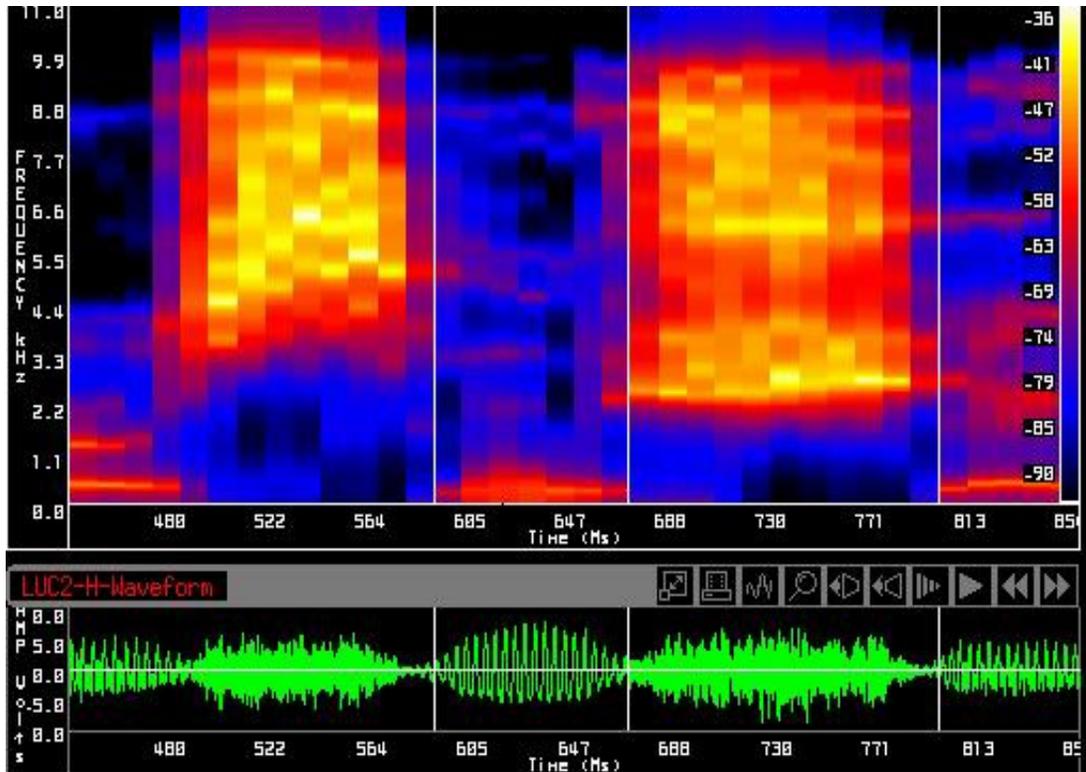
*Pre emphasis*- 98%

*Order* –20

Foram medidas, em milisegundos, a duração total dos enunciados, a duração de suas unidades fonéticas – os *GIPC* ( *Group Inter-Perceptual Center*)<sup>8</sup> (Barbosa & Bailly, 1994) e a duração relativa de segmentos vocálicos e consonantais que se mostraram diferenciados na produção do modelo e dos alunos, quando se inspecionou acusticamente os dados. Veja-se, na figura 2 abaixo, a medida de duração da unidade fonético-acústica, *GIPC*, **amf** da emissão “*Could I borrow some sugar*”, a qual compreende o espaço de tempo que vai do *onset* da vogal de **some**, sinalizado no espectrograma pelo aparecimento de dois formantes da vogal e, na forma da onda, por ciclos periódicos, até o ponto onde se inicia o *onset* da vogal da sílaba seguinte. Na mesma figura, exemplifica-se uma medida de duração do segmento fricativo /ʃ/.

---

<sup>8</sup>O *GIPC* é uma unidade fonético-acústica que compreende a rima de uma sílaba e o ataque da sílaba seguinte, quando esse se faz presente.



1 \_\_\_\_\_ 2 \_\_\_\_\_ 126,80ms \_\_\_\_\_ 3  
 \_\_\_\_\_ 209,14ms \_\_\_\_\_ 3

FIGURA2: De 1 a 3, medida de duração da unidade fonético acústica **əmʃ** da emissão "Could I borrow some sugar", da Grav2 do sujeito Leila.  
 De 2 a 3, medida do segmento fricativo /ʃ/

## **CAPÍTULO 6 - ANÁLISE DOS RESULTADOS**

## 6.1 Resultados da avaliação perceptiva

A avaliação das emissões dos alunos por parte dos falantes nativos deu-se em conformidade com uma escala de **0** a **6**, sendo que as emissões consideradas próximas da produção nativa deveriam receber avaliação **0** ou avaliações mais próximas de **0**. Portanto, na leitura dos dados, os melhores resultados são os que apresentam as médias mais baixas.

As sentenças modelos, que eram exemplos de emissões de falantes nativos, receberam avaliação de **0** a **2** pelos voluntários avaliadores, distribuída da seguinte maneira:

TABELA 7: Resultado da avaliação perceptiva das emissões dos falantes nativos

Total Geral por <b>Sentença</b>	0	1	2	3	4	5	6	Total
1f-How often does your mother call?	21	8	11	0	0	0	0	40
1r- Are you going to take a vacation?	31	8	3	0	0	0	0	40

Ao conferirem grau **0** para a maioria das emissões de falantes nativos ou grau próximo de zero, os voluntários dão confiabilidade aos dados coletados na pesquisa perceptiva, pois demonstram ter percepção do sotaque estrangeiro ao se posicionarem de maneira diversa diante das emissões dos alunos participantes do experimento, além de proporcionarem um outro parâmetro de comparação para as avaliações dos alunos quando avaliadas dentro da escala de **0** a **2**.

Comparando-se os resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do Grupo1, conclui-se que os falantes nativos do

inglês acusam perceber uma melhora na produção desses alunos após o treinamento com o apoio visual (gráfico 2), sendo essa melhora considerada significativa ( $p < 0,05$ ).

Por outro lado, a comparação entre os resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do Grupo2 demonstrou que os alunos desse grupo foram mais mal avaliados na Grav2, como pode ser visto no mesmo gráfico 2 abaixo. Entretanto, não foi encontrada diferença significativa ( $p > 0,05$ ), o que nos leva a concluir que os alunos que fizeram o treinamento no laboratório de línguas não tiveram alteração em sua produção: o grupo não melhorou nem piorou.

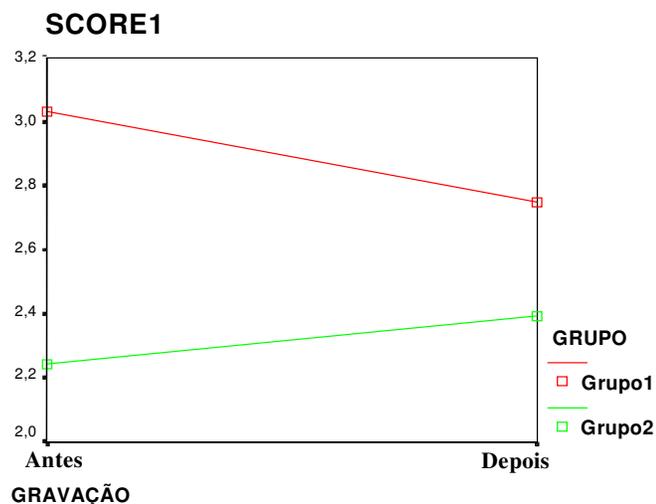


Gráfico 2 - Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do Grupo1 e do Grupo 2.

Valores inferiores indicam melhores resultados

A comparação dos resultados referentes aos tipos de exercícios demonstrou que a variável “exercício” não exerceu influência significativa no resultado para o conjunto de alunos dos dois grupos. Apesar da avaliação do Grupo1 apresentar uma tendência de melhores resultados, principalmente no exercício 3 (gráfico 3), e a avaliação do Grupo 2 apresentar uma tendência de piores resultados nos exercícios

1 e 3 (gráfico 4), essas alterações não são significativas.

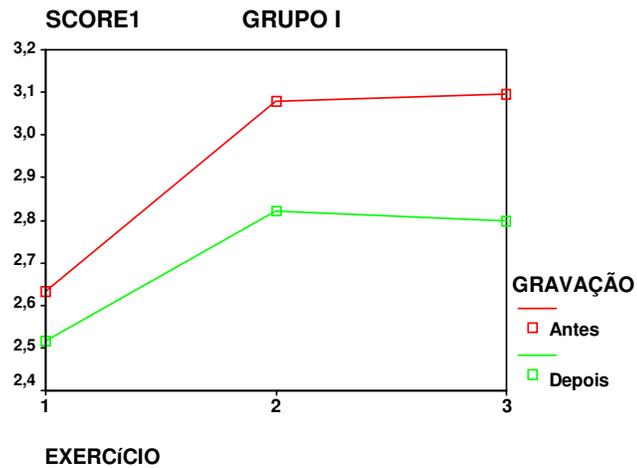


Gráfico3 - Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do grupo1 e em relação aos exercícios 1, 2, 3.  
Valores inferiores indicam melhores resultados

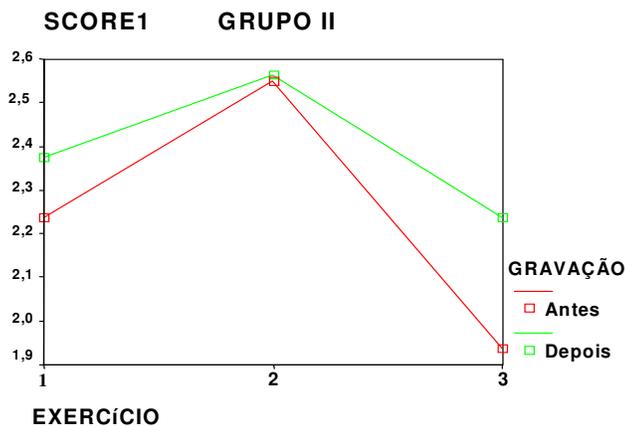


Gráfico4 - Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 do conjunto de alunos do grupo2 e em relação aos exercícios 1, 2, 3.  
Valores inferiores indicam melhores resultados

A variável “sentença”, entretanto, mostrou-se influente, visto que se verificou diferença significativa ( $p < 0,05$ ) entre as produções das orações interrogativas Yes/No dos alunos do Grupo1.

Os resultados relativos ao conjunto dos alunos do Grupo2 foram desconsiderados na análise que se segue, uma vez que o grupo não se alterou com o treinamento.

Considerando-se os resultados das comparações das emissões de cada um dos sujeitos do Grupo1, chega-se à conclusão que a aluna Rita foi a que apresentou maior progresso de acordo com a avaliação dos falantes nativos do inglês (gráfico 5), sendo que a melhora com significância estatística ocorreu na emissão das interrogativas *Yes/No* (gráfico 6).

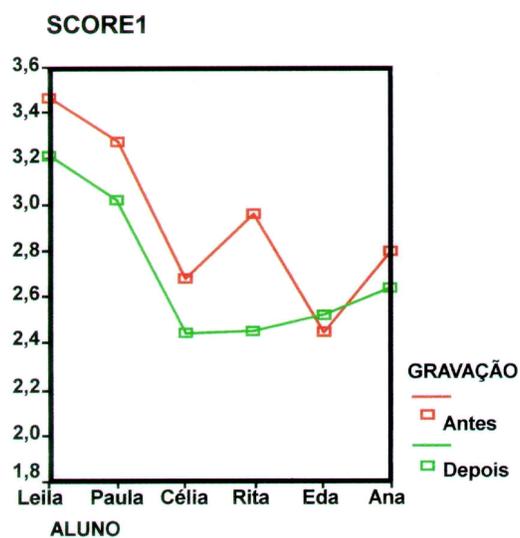


Gráfico 5 - Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 de cada um dos alunos do grupo1

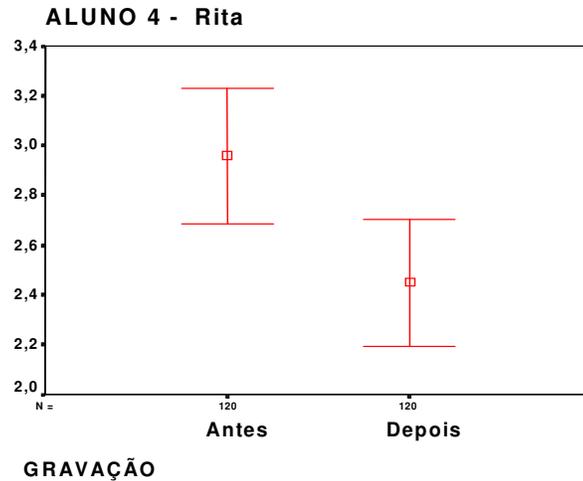


Gráfico 6 - Comparação dos resultados do Score1 da Grav1 com os resultados da Grav2 da aluna Rita

As emissões das orações *Yes/No* que constaram do experimento, tais como produzidas pelos alunos do Grupo1, receberam as seguintes avaliações:

TABELA 8: Resultado da avaliação perceptiva das emissões interrogativas *Yes/No* dos alunos do grupo1

Total geral por sentença. Todas a avaliações	0		1		2		3		4		5		6	
	G1	G2												
1r-Are you going to take a vacation?	2	4	5	12	14	16	18	15	16	10	5	3	0	0
2h-Can I borrow some sugar?	3	3	1	11	13	15	16	19	20	9	6	3	1	0
2d-Would you please explain this word?	1	1	2	3	12	12	16	17	16	22	12	5	1	0

Considerando-se, como exemplo, que a oração 1r recebeu, na gravação1, duas avaliações de valor 0 (ausência de sotaque), cinco avaliações de valor 1, catorze avaliações de valor 2, 18 avaliações de valor 3, dezesseis avaliações de valor 4, cinco avaliações de valor 5 e nenhuma avaliação de valor 6, totaliza-se 176 pontos para essa emissão na gravação1, conforme demonstrado na Tabela 9 abaixo:

TABELA 9: Resultado geral da avaliação perceptiva da emissão “Are you going to take a vacation?” dos alunos do grupo1

Total geral por sentença. Todas as avaliações	0		1		2		3		4		5		6		Total	
	G1	G2	G1	G2												
1r-Are you going to take a vacation?	2	4	5	12	14	16	18	15	16	10	5	3	0	0		
Total de pontos da avaliação	0	0	5	12	28	32	54	45	64	40	25	15	0	0	176	144

Empregando-se o processo exemplificado na Tabela 9, obtém-se os resultados que estão disponibilizados na Tabela 10, relativos às três orações Yes/No do experimento.

TABELA 10: Resultado geral da avaliação perceptiva de cada tipo de oração interrogativa Yes/No dos alunos do grupo1

SENTENÇAS	GRAVAÇÃO1	GRAVAÇÃO2	DIFERENÇA
<i>Are you going to take a vacation?</i>	176	144	32
<i>Could I borrow some sugar?</i>	191	149	42
<i>Would you please explain this word?</i>	204	191	13

Esses resultados indicam uma tendência de melhora na produção dos alunos do grupo 1, mas apontam, também, para diferenças importantes que serão inspecionadas e comentadas com base nos achados da análise acústica. A oração “Are you going to take a vacation?” recebe a melhor avaliação tanto na primeira quanto na segunda gravação, ao passo que a avaliação de “Would you please explain this word?” apresenta os piores resultados, na gravação inicial, e na final, e

a menor diferença entre os resultados da gravação1 e gravação2. Por outro lado, os resultados referentes a “*Could I borrow some sugar?*” atestam maior progresso dos sujeitos nessa emissão.

Concluimos, portanto, que o teste de percepção aponta para uma melhora na qualidade de produção do conjunto de alunos do grupo1, que fizeram o treinamento no computador, mas apenas em relação à produção de emissões interrogativas *Yes/No*. Sugere, também, que a melhora no desempenho dos alunos depende da interação de outros fatores além da estrutura sintática da emissão, os quais serão apontados na análise acústica dos dados.

Por serem as emissões *Yes/No* do Grupo1 as únicas que indicaram uma melhora significativa na qualidade de desempenho dos alunos, elegemos esse tipo de oração como objeto da análise acústica.

## **6.2 Resultados da análise acústica**

Procurando-se respaldo na análise acústica para os resultados da avaliação perceptiva dos falantes nativos, analisou-se os valores obtidos da medida de F0 do núcleo da vogal de cada sílaba das emissões interrogativas *Yes/No* dos alunos do Grupo1. Essas medidas em Hz estão expostas nas tabelas 11, 12, 13 abaixo.

TABELA 11: Medidas de F0 das emissões do Modelo e dos sujeitos da oração

*“Are you going to take a vacation?”.*

O fundo cinza sinaliza a emissão com contorno final descendente

Modelo-1r	214	223	229	230	220	217	187	185	268	
Leila:Grav1	203	205	198	196	196	198	188	255	282	
Leila:Grav2	205	212	206	210	210	210	198	262	319	
Paula:Grav1	110	122	127	114	108	98	97	149	178	
Paula:Grav2	110	113	119	114	110	104	105	165	200	
Rita:Grav1	195	205	231	249	236	231	198	306	333	
Rita:Grav2	193	195	212	220	200	204	200	272	302	
Eda:Grav1	225	220	225	220	220	209	198	279	319	
Eda:Grav2	211	213	226	215	215	207	202	268	301	
Ana:Grav1	171	215	226	234	231	117	106	124	132	
Ana:Grav2	220	227	232	214	214	202	195	220	265	
Célia:Grav1	196	184	227	198	191	185	191	231	289	
Célia:Grav2	188	193	220	188	191	180	174	259	334	
	<i>Are</i>	<i>you</i>	<i>going</i>	<i>to</i>	<i>ta</i>	<i>ke</i>	<i>a</i>	<i>va</i>	<i>ca</i>	<i>tion</i>

TABELA 12: Medidas de F0 das emissões do modelo e dos sujeitos da oração

*“Would you please explain this word?”*

O fundo cinza sinaliza as emissões sem contorno final ascendente

Modelo2d	211	217	250	235	219	211	183	266
Rita:Grav1	229	222	234	222	214	198	195	324
Rita:Grav2	223	220	232	216	212	195	188	290
Eda:Grav1	225	220	240	234	229	197	202	253
Eda:Grav2	231	236	250	241	227	206	200	232
LeilaGrav1	220	214	220	229	218	211	211	262
LeilaGrav2	216	218	223	250	240	242	222	245
Paula:Grav1	119	133	143	136	115	107	113	122
Paula:Grav2	120	125	138	141	117	106	101	112
Ana:Grav1	220	224	232	253	224	214	188	196
Ana:Grav2	226	227	234	248	218	210	197	200
Celia:Grav1	220	250	250	209	195	211	234	234
Celia:Grav2	212	272	259	188	208	185	184	210
	<i>Would</i>	<i>you</i>	<i>please</i>	<i>ex</i>	<i>plain</i>	<i>this</i>	<i>word</i>	

TABELA 13: Medidas de F0 das emissões do modelo e dos sujeitos da oração  
 “*Could I borrow some sugar?*”

Todas as emissões apresentam contorno final ascendente

Modelo-2h	220	209	194	196	217	187	211	239
Paula:Grav1	117	116	113	117	112	98	118	158
Paula:Grav2	112	112	114	117	105	102	116	195
Rita:Grav1	211	204	207	224	211	202	229	305
Rita:Grav2	206	197	195	242	206	178	210	298
Celia:Grav1	195	196	200	218	184	180	206	250
Célia:Grav2	210	204	187	200	199	182	204	257
Ana:Grav1	214	206	196	216	207	206	212	272
Ana:Grav2	206	208	193	206	202	197	198	218
Eda:Grav1	259	237	206	216	208	204	208	282
Eda:Grav2	253	245	218	237	237	113	115	258
Leila:Grav1	237	224	213	214	214	216	224	323
Leila:Grav2	235	227	202	227	242	251	245	306
	<i>Could</i>	<i>I</i>	<i>bo</i>	<i>rrow</i>	<i>some</i>	<i>su</i>	<i>gar?</i>	

Comparando-se os resultados das Tabelas 11, 12 e 13, nota-se a incidência maior de contornos entoacionais com curva final descendente na oração “*Would you please explain this word?*”. Verificou-se, também, que toda emissão de interrogativa *Yes/No* que termina com valores de F0 inferiores ao valor apresentado pela sílaba inicial é mal avaliada pelos sujeitos do experimento. Assim, na gravação2 da sentença 1r (*Are you going to take a vacation?*), a emissão do sujeito Ana recebe melhor avaliação do que a mesma emissão da gravação1 sinalizada pela cor cinza na tabela 11. Na gravação1, a emissão recebe 37 pontos, enquanto que a emissão da mesma oração recebe apenas 30 pontos na gravação2. Comparem-se os contornos de F0 nas figuras 3 e 4 abaixo:

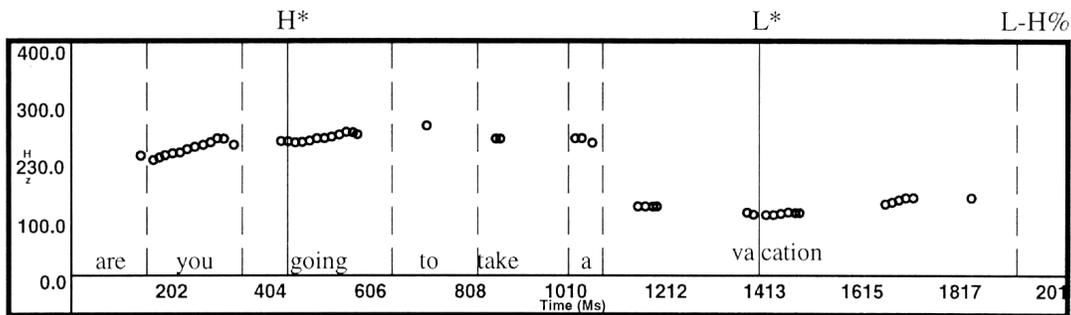


FIGURA3: Ana: Grav1: contorno de F0 da oração: *Are you going to take a vacation?*  
com 171Hz no início e com 132Hz no final

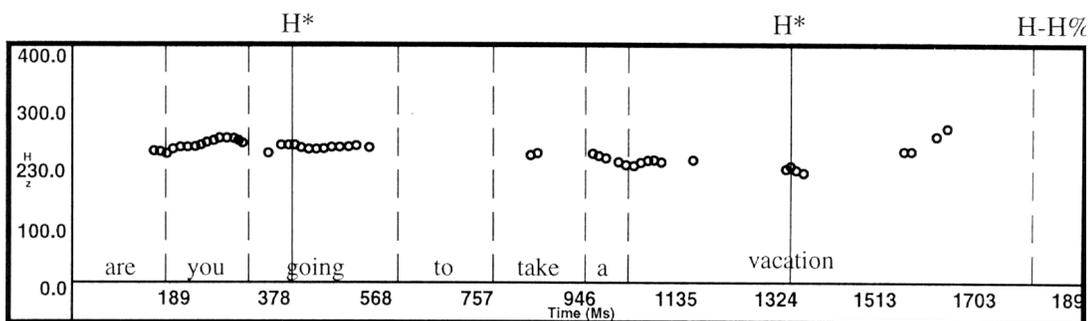


FIGURA4: Ana, Grav2: contorno de F0 da oração: *Are you going to take a vacation?*  
com 214Hz no início e com 268Hz no final

Como se pode observar na Figura 3, referente à gravação 1, o valor de F0 da sílaba inicial é superior ao da sílaba final do grupo entoacional, apresentando o contorno global H\* L\* L-H% que contraria a expectativa de uma curva global ascendente para esse tipo de estrutura do inglês, que se caracteriza pelos contornos H\* H\* H-H% e H\* L\* H-H%. Conseqüentemente, a emissão da gravação 2, com curva global ascendente, apresentada na Figura 3, é melhor avaliada, pois o contorno de F0 se aproxima mais da emissão do nativo H\* H\* H-H%. No Gráfico 7, a seguir, comparamos os contornos gerados a partir de valores de F0 obtidos na porção medial de cada núcleo silábico dos enunciados gravados pelo sujeito Ana (gravação 1 e 2) e pelo nativo (modelo).

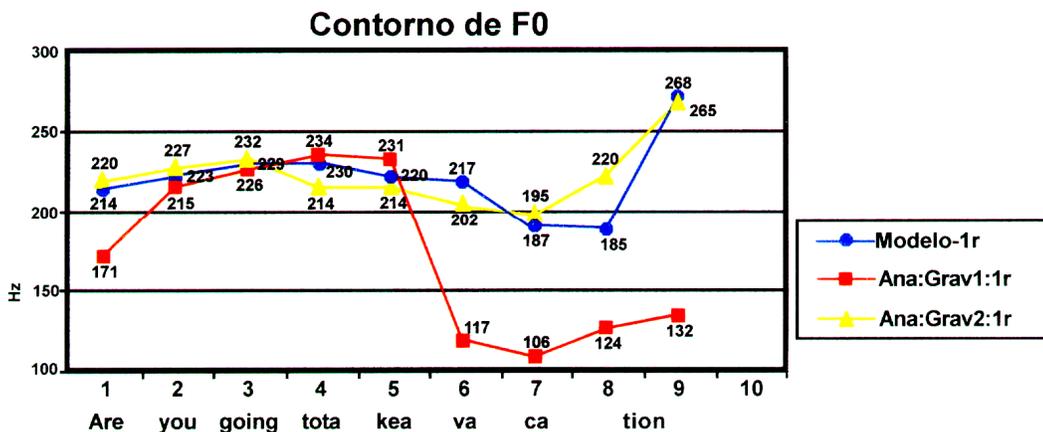


GRÁFICO 7: Comparação dos dados do contorno de F0 da emissão “Are you going to take a vacation” do modelo e do sujeito Ana.

Esse vínculo, que o falante nativo estabelece entre a avaliação do desempenho de interrogativas *Yes/No* e o contorno entoacional ascendente, explicitados no caso do sujeito Ana acima, é reforçado com os dados de Célia. A emissão da oração 2d do sujeito Célia (*Would please explain this word?*) que se apresenta sem inclinação na gravação2 (Figura 6), obteve 41 pontos na avaliação perceptiva, ao passo que a emissão da mesma oração na gravação1 (Figura 5), com curva ascendente, recebeu 37 pontos, sendo, portanto, melhor avaliada.

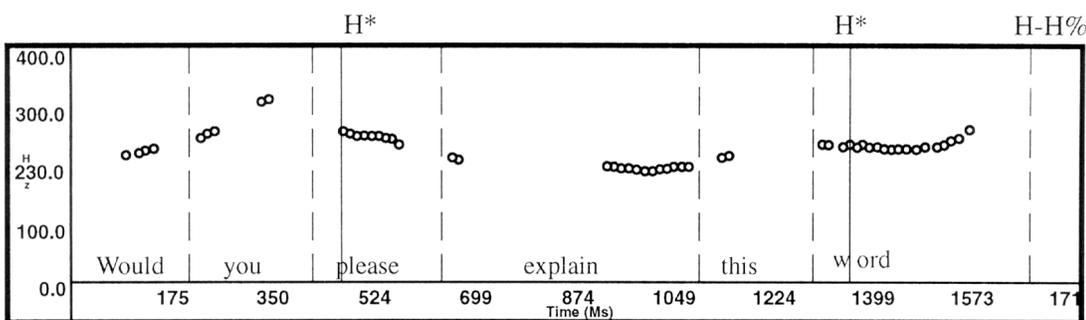


FIGURA 5: Célia, Grav1: contorno de *pitch* ascendente da emissão: *Would you please explain this word?* com 213Hz no início e com 244Hz no final

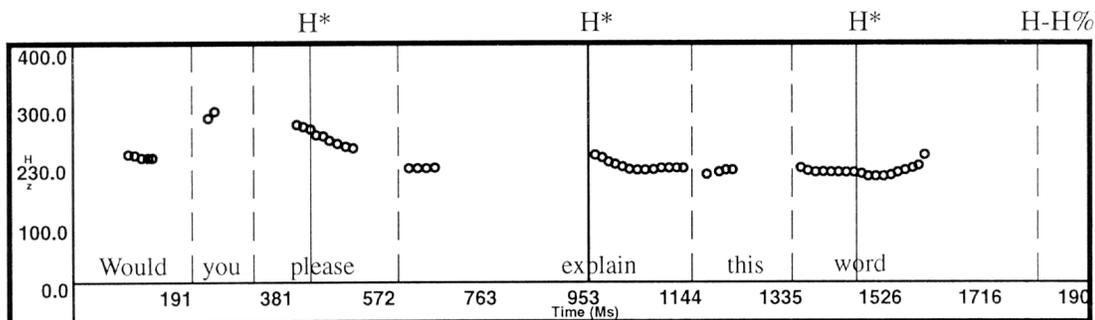


FIGURA6: Célia, Grav2: contorno de *pitch* sem inclinação, com o início e final em torno de 210Hz, da emissão: “*Would you please explain this word?*”

Célia apresenta movimento final ascendente na gravação 2, mas o valor de F0 da sílaba final de fronteira não é superior ao valor da sílaba inicial. Desse modo, o contorno entoacional da emissão como um todo não é ascendente, contrastando com o contorno de F0 ascendente do falante nativo. Veja-se a comparação dos valores de F0 das sílabas iniciais e finais das emissões de Célia e do modelo no Gráfico 8, que se segue.

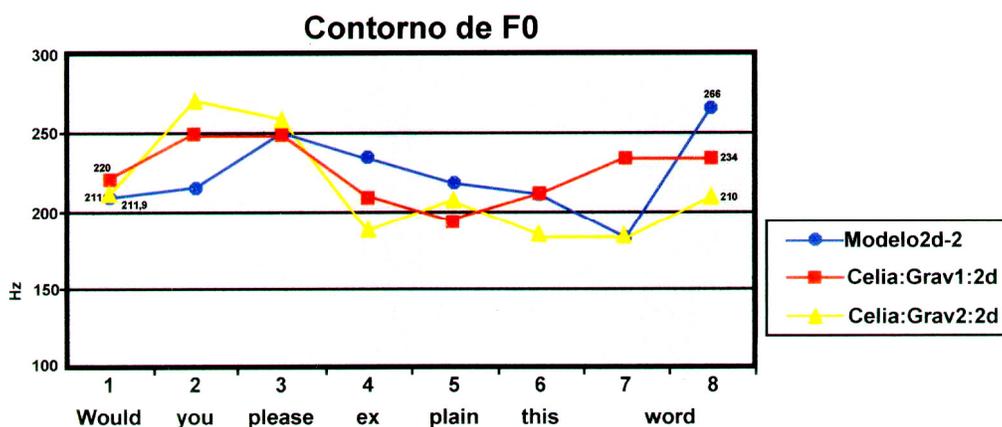


GRÁFICO 8: Comparação dos dados do modelo e do sujeito Célia do contorno de F0 da emissão “*Would you please explain this word?*”

Conclui-se, assim, que a curva ascendente a partir da sílaba que carrega o acento de *pitch* na fronteira final das interrogativas *Yes/No* do inglês é um traço que identifica fortemente o sotaque estrangeiro na fala de falantes não nativos do inglês. Conclui-se, também, que o valor de F0 final de emissão está relacionado não só

com o valor do núcleo vocálico que o precede, mas com os valores de F0 da emissão como um todo. Compare-se nos gráficos 9 e 10 as tendências de direção das emissões do modelo e da gravação 2 de Célia relativas ao gráfico 8.

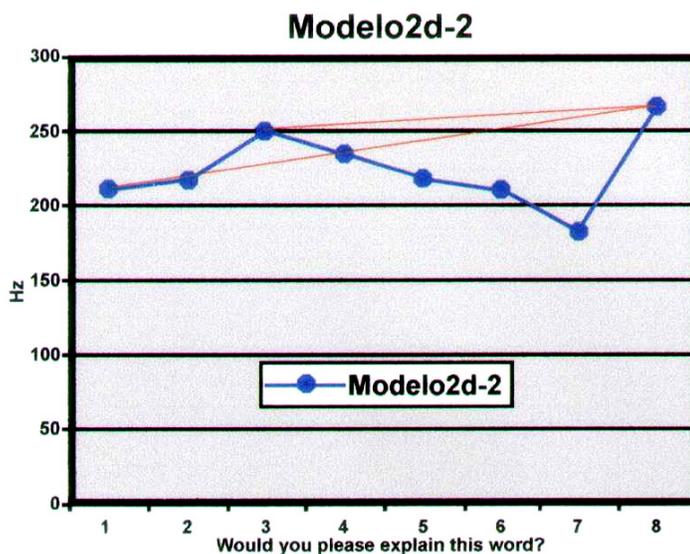


GRÁFICO 9: Linha de tendência de direção dos valores de *pitch* da emissão modelo "Would you please explain this word?"

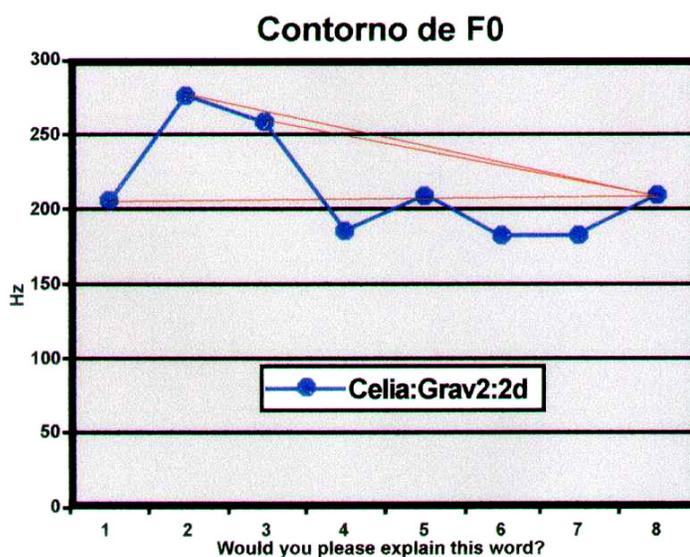


GRÁFICO 10: Linha de tendência de direção dos valores de *pitch* da Grav2 de Célia da emissão "Would you please explain this word?"

No entanto, interrogativas *Yes/No*, produzidas pelos sujeitos do experimento com contornos finais e globais ascendentes, receberam avaliações diferenciadas

por parte dos falantes nativos, sugerindo que outros fatores interagem com o contorno de F0 na percepção dessas emissões.

Considerando-se os resultados da Tabela 10, que repetimos abaixo para facilitar a comparação, nota-se uma gradação na avaliação dos sujeitos, tanto na gravação1 quanto na 2, que sugere maior dificuldade na emissão da oração “*Would you please explain this word*” do que na emissão de “*Are you going to take a vacation*” e “*Could I borrow some sugar*” que já havia sido sinalizada pelos valores de F0 das tabelas 11, 12 e 13.

TABELA 10: Resultado geral da avaliação perceptiva de cada tipo de oração interrogativa Yes/No dos alunos do grupo1

SENTENÇAS	GRAVAÇÃO1	GRAVAÇÃO2	DIFERENÇA
<i>Are you going to take a vacation?</i>	176	144	32
<i>Could I borrow some sugar?</i>	191	149	42
<i>Would you please explain this word?</i>	204	191	13

Analisando-se a curva do *pitch* no final dessas emissões, comprova-se que o aluno tem mais facilidade em fazer o movimento ascendente quando a fronteira final termina com a paroxítone [veɪ'keɪʃn̩] (*vacation*). O movimento inicial da curva localiza-se no ditongo [eɪ] e completa-se na sílaba seguinte [ʃən]. Como esse exemplo de emissão foi o melhor avaliado na produção inicial, a diferença do resultado do desempenho dos alunos obtido na gravação 2, após o treinamento, é pequena, pois, espera-se que, quanto mais próxima a gravação inicial está do modelo, menor a margem de melhora. Considerando-se que as emissões dos falantes nativos das orações modelo receberam, dos americanos, avaliações de 0 a

2, os resultados obtidos pelos sujeitos Paula e Leila na avaliação da gravação2 dessa emissão indicam, por exemplo, uma produção muito próxima de um nativo, sendo que, de 10 avaliações, Paula recebeu duas e Leila três avaliações acima de 2 pontos.

TABELA 14: Avaliação da Grav2 da emissão “*Are you going to take a vacation*” de Paula e Leila

AVALIAÇÃO	0	1	2	3	4	5	6
PAULA	1	5	2	1	1	0	0
LEILA	2	1	4	3	0	0	0

A emissão de “*Can I borrow some sugar*” apresenta, como a anterior, fronteira final em paroxítona com movimento ascendente na última sílaba, apesar de [ˈʃʊgəɾ] (“*sugar*”), diferentemente de [vɛɪˈkeɪʃn̩] (“*vacation*”), ser dissílaba. Além disso, a sílaba final da palavra (“*sugar*”) apresenta coda, a qual oferece mais espaço de manobra para a realização do movimento ascendente por parte dos falantes nativos do português que, normalmente, têm dificuldade de realizar movimentos ascendentes dentro dos limites de apenas um segmento fonético. Uma vez conscientizado da questão, o aluno tem mais facilidade em reproduzir [ˈʃʊgəɾ] (“*sugar*”) em contexto final de fronteira ascendente do que [wɜːd̩] (“*word*”), que ocorre na fronteira final da terceira oração. Esta, por ser monossílaba, exige movimento de ascensão dentro de uma única sílaba, sendo esse fato um dos motivos de “*Would you please explain this word*” ter sido a emissão de pior avaliação nas duas gravações e a que apresentou menor diferença no desempenho dos sujeitos depois do treinamento.

Comparando-se, ainda, as interrogativas *Yes/No* do experimento, fica evidenciado que as orações “*Are you going to take a vacation?*” e “*Would you please explain this word?*” caracterizam-se por contextos fonéticos mais problemáticos para o aluno brasileiro do que “*Could I borrow some sugar?*”.

A interrogativa “*Are you going to take a vacation?*”, [ɑːjə'gouɪŋ|tʰeɪkəveɪ'kheɪʃn], apresenta junturas, tais como, “*Are you*” [ɑːjə] e “*to take*” [tʰeɪk] que exigem redução de segmentos mais difíceis de serem reproduzidas pelo aluno do que “*Could I*” [kʊdaɪ] em “*Could I borrow some sugar?*”, [kʊdaɪ'ba:rou|səm'ʃugəʊ].

Com relação a essa última, alguns dos sujeitos usaram, ainda, o recurso de substituir “*Could I*” por “*Can I*”, uma vez que o modal “*Can*”, por ser apresentado antes do que o modal “*Could*” nas lições de inglês, tem, em geral, uso mais produtivo na fala dos alunos. Além disso, é provável que o falante nativo tenha avaliado melhor “*Can I*” do que “*Could I*” que envolve o julgamento de variantes [kʊd aɪ], [kʊa aɪ].

A oração “*Would you please explain this word?*”,

[wʊdʒu 'plɪz | ɪks'pleɪn ðɪs 'wɜːd̩] ,

além de terminar em monossílaba, apresenta junturas (*Would you*, [wʊdʒu], *this word*, [ðɪs 'wɜːd̩], grupo consonantal (*explain*, [ɪks'pleɪn]), o segmento /ð/ na palavra *this*<sup>9</sup> [ðɪs], o *unreleased* /d/ no final da emissão de [wɜːd̩], contextos fonéticos considerados difíceis para o falante do português brasileiro.

---

<sup>9</sup> O inglês possui dois fonemas fricativos interdental /θ/ surdo e /ð/ sonoro. Esses sons não fazem parte do inventário fonêmico do português e geralmente são realizados como [t], [s] ou [f] por brasileiros aprendizes da língua inglesa.

Na inspeção espectrográfica que realizamos, notamos que os intervalos correspondentes às oclusões total e parcial nas produções da junção entre “*would*” e “*you*” [wʊdʒu] mostraram-se maiores nas produções dos sujeitos do experimento do que na do falante nativo do inglês. Além disso, ambas as palavras são acentuadas e verifica-se um aumento expressivo do F0 nas realizações de “*you*”, o que não acontece na junção de “*Could I*” da emissão “*Could I borrow some sugar?*”. Na junção que ocorre entre “*this*” e “*word*”, o falante do português tende a produzir uma fricativa sonora [z] e no final da palavra “*word*” utilizar uma africada [dʒ] antes de um suposto [ɪ] final.

Comparando-se os espectrogramas das emissões dos sujeitos do experimento com os espectrogramas das orações modelo, evidencia-se a maior duração das sibilantes em *please, this, explain* do que as realizadas nas emissões dos falantes nativos<sup>10</sup> do inglês.

Compare-se a duração das fricativas nas produções de *please, this e explain* no espectrograma de banda larga e na forma da onda da emissão do modelo na Figura 7 com o espectrograma da emissão do sujeito Leila na Figura 8 abaixo, sinalizadas pelas barras brancas numeradas de 1 a 4.

---

<sup>10</sup> Segundo depoimentos dos bilíngües, Marcello Marcelino e Ellen Osborn, o alongamento da sibilante do inglês falado por brasileiros é uma das características de sotaque que mais incomoda.

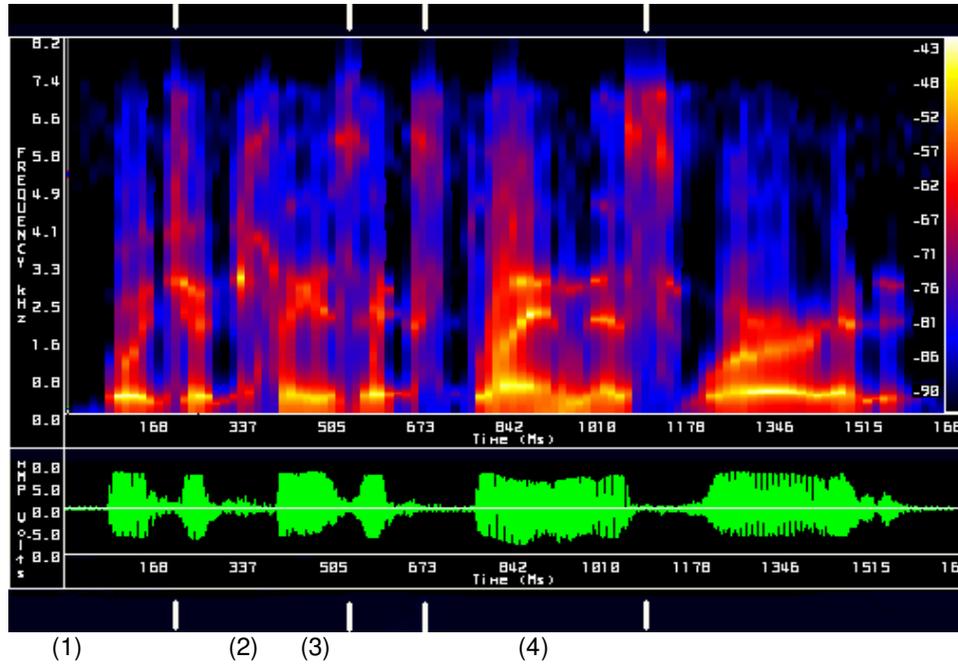


Figura 7: Espectrograma da emissão modelo de "Would you please explain this word",

[ wʊdʒu | plɪːz ɪk'splɛɪ ðɪs wɔːrd ]. Medidas de duração das Fricativas:

(1) - /dʒ/ = 26ms; (2) - /z/ = 15ms; (3) - /s/ = 55ms; - /s/ = 78ms

Duração total da emissão = 1.472ms

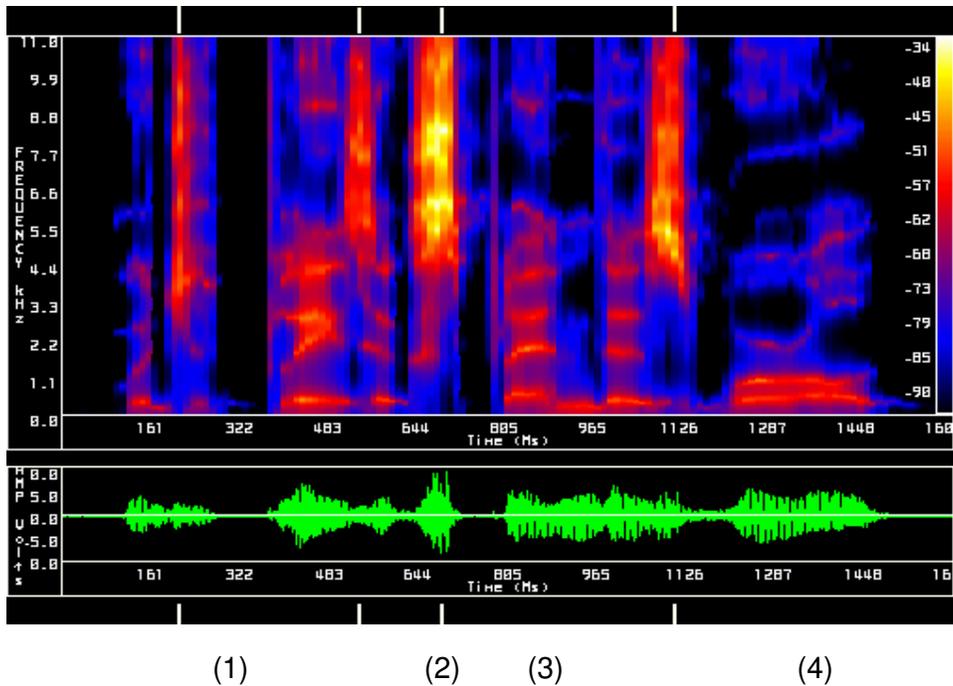


Figura 8: Espectrograma da emissão da Grav1 do sujeito Leila de "Would you please explain this word".

[ wʊdʒu | plɪːz ɪk'splɛɪ ðɪs wɔːrd ]. Medidas de duração das fricativas:

(1) - [dʒ] = 46ms; (2) - [z] = 88ms; (3) - [s] = 90ms; - [s] = 88ms

Duração total da emissão: 1.664ms

Comparando-se, ainda, as emissões do modelo e dos sujeitos, nota-se, também, diferenças quanto à amplitude relativa da onda aperiódica referente à consoante fricativa e da onda periódica referente à vogal, exemplificadas nos oscilogramas abaixo.

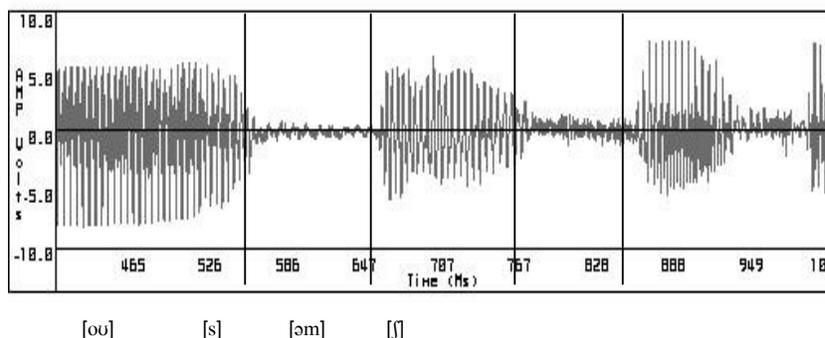


Figura 9: Parte da onda sonora da emissão modelo da oração  
*"Could I borrow some sugar"* [kʊd ai 'bɔ:rou | səm 'ʃʊgəə]

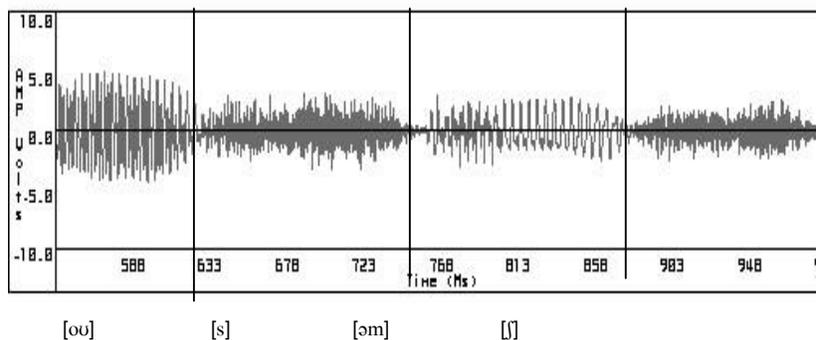


Figura 10: Parte da onda sonora da emissão modelo da oração  
*"Could I borrow some sugar"* [kʊd ai 'bɔ:rou | səm 'ʃʊgəə]

Esses resultados, como um todo, vão de encontro aos apresentados por Magen (1998) que, ao pesquisar a percepção de sotaque estrangeiro na fala de dois falantes do espanhol, conclui que os falantes do inglês são sensíveis a fatores relacionados à estrutura silábica, à maneira de articulação, ao acento lexical e frasal.

Gostaríamos de assinalar que as observações que fazemos aqui, quanto à duração das consoantes e à amplitude do ruído fricativo, são meramente especulativas, pois, para considerarmos os parâmetros acústicos de intensidade e duração são necessários controles rigorosos em relação às condições de gravação, no caso das medidas de intensidade, e em relação à taxa de elocução e uso de normalização, no caso de medidas de duração.

Considerando-se a duração das unidades fonético-acústicas que incluem fricativas, observa-se que, na fala do nativo, a relação entre a duração das fricativas e a duração dos outros segmentos da unidade da qual ela faz parte (GIPC) é diferente da relação que está presente na fala dos sujeitos. Assim, na emissão do falante nativo da oração (“*Could I borrow some sugar?*”), a duração da fricativa [s] é sempre menor do que a duração dos outros segmentos da unidade GIPC da qual faz parte *ous*. Ao contrário, na emissão dos sujeitos do experimento, a duração da fricativa é sempre maior do que a duração do conjunto dos outros segmentos.

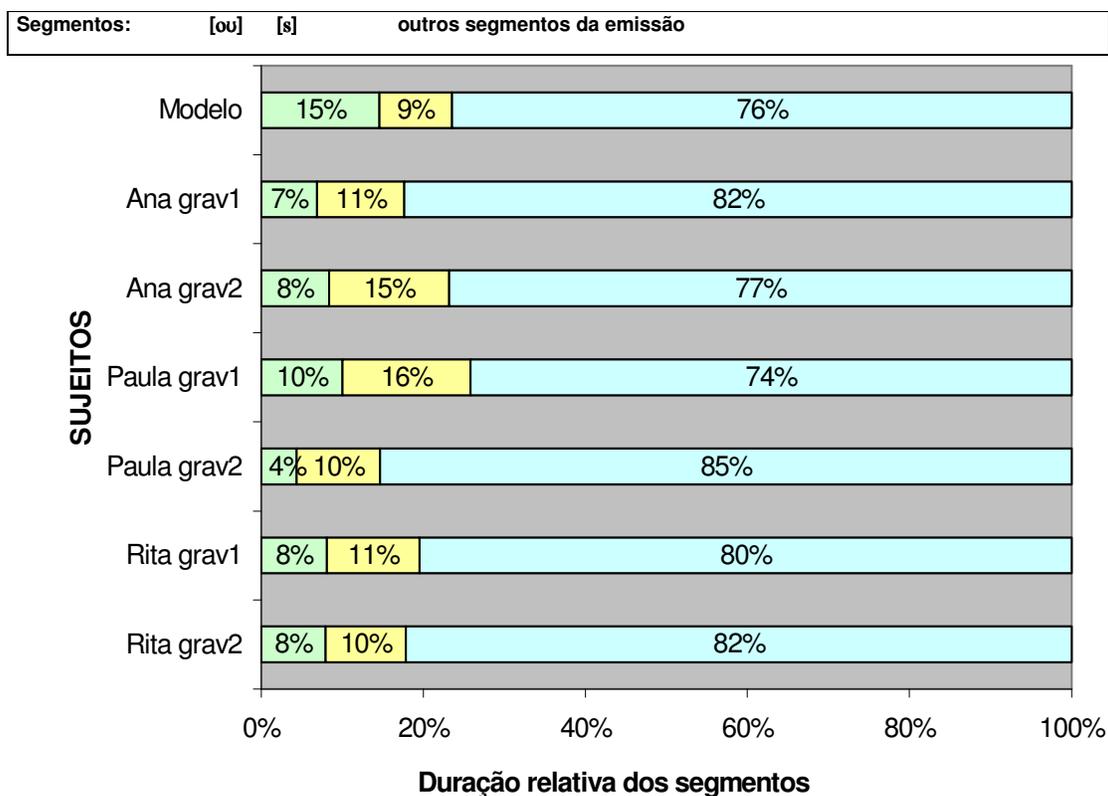
Podemos interpretar tal fato em relação à interação entre prosódia e segmento, visto que [ou] ocorre em posição pós-tônica na palavra “*borrow*”. Como pode se observar, pelas durações em milisegundos indicadas abaixo, os falantes do português reduzem a vogal nessa posição de acento.

TABELA 15: Comparação em milissegundos da duração da fricativa com a duração do conjunto de elementos que compõe a unidade fonético-acústica em *ous* nas Grav1 e Grav2 da emissão “*Could I borrow some sugar*” pelos sujeitos Rita, Paula e Ana. Os sinais < , > indicam duração menor ou maior da fricativa

Sujeito	Gravação	[ou]	[s]	Duração total da emissão
modelo		158	97 <	1.083
Rita	Grav1	92	130 >	1.137
	Grav2	89	111 >	1.120
Paula	Grav1	86	135 >	856
	Grav2	43	101 >	982
Ana	Grav1	86	135 >	1.250
	Grav2	105	185 >	1.249

Analisando-se a duração relativa dos segmentos [ou] e [s] em relação aos outros elementos da emissão desses mesmos dados dos sujeitos Ana, Paula, Rita e da emissão modelo, chega-se à mesma conclusão.

TABELA16: Distribuição dos valores de duração relativa dos segmentos [ou] e [s] do GIPC *ous* e dos outros segmentos da emissão “*Could I borrow some sugar?*” do modelo e da Grav1 e Grav2 dos sujeitos Ana, Paula, e Rita



Os dados expostos na Tabela 16 indicam diferenças em termos de ajustes temporais. Demonstram, também, a interferência da produção dos segmentos fonéticos da L1 na produção da L2. A soma desses fatores ocasiona, no caso da oração em questão, problemas relativos à interação prosódia segmento.

Tais considerações sugerem a importância da contribuição de estudos da análise acústica para avaliar e orientar a composição de material pedagógico para o ensino de entoação de uma língua estrangeira. Esse material deveria privilegiar contextos fonéticos importantes do ponto de vista perceptivo para os falantes nativos e que possam oferecer desafios para falantes não nativos das diferentes

línguas. No caso específico do ensino da entoação do inglês para alunos brasileiros, além de contextos de coarticulação particularmente problemáticos<sup>11</sup>, o movimento de *pitch* dentro de uma única sílaba apresenta-se como um elemento gerador de sotaque entoacional importante, reforçando resultados apresentados em Rocca e Marcelino (1997), que indicavam a dificuldade do aluno brasileiro em perceber e reproduzir o movimento de *pitch* dentro de uma única sílaba.

No entanto, apesar da diferença no grau de dificuldade apresentado pelos contextos fonéticos das emissões, o treinamento parece ter capacitado os sujeitos a produzirem mais eficazmente junturas entre unidades lingüísticas não acentuadas. Tome-se, como exemplo, a gravação1 e 2 do sujeito Paula. Comparando-se a trajetória dos formantes nos espectrogramas das figuras 11 e 12 abaixo, evidenciam-se as diferenças na coarticulação dos segmentos em posição de juntura que apontam para um melhor desempenho na gravação2. Na gravação1, *would you* tem a duração de 151ms e, na gravação2, 135ms. O mesmo acontece com a juntura entre “*could*” e “*I*” na emissão “*Could I borrow some sugar*”, que apresenta 130ms na gravação1, e 97ms na gravação2.

---

<sup>11</sup> Contextos de coarticulação problemáticos, tais como, grupos consonantais em final de palavra (“*glimpsed*” [glimpst]) e processos de sândi (Jack’s Czech friend [dʒæks’tʃek’frend]).

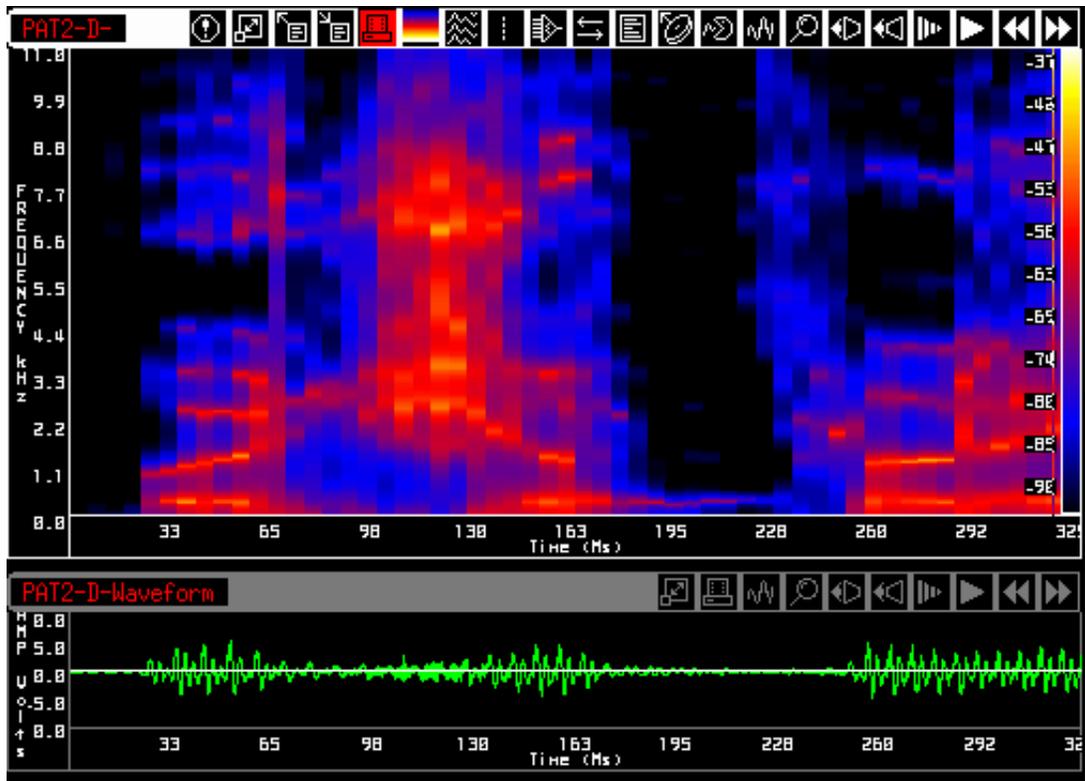


Figura 11: Espectrograma de banda larga e oscilograma de parte da emissão  
 “Would you please explain this word” da Grav1 do sujeito Paula

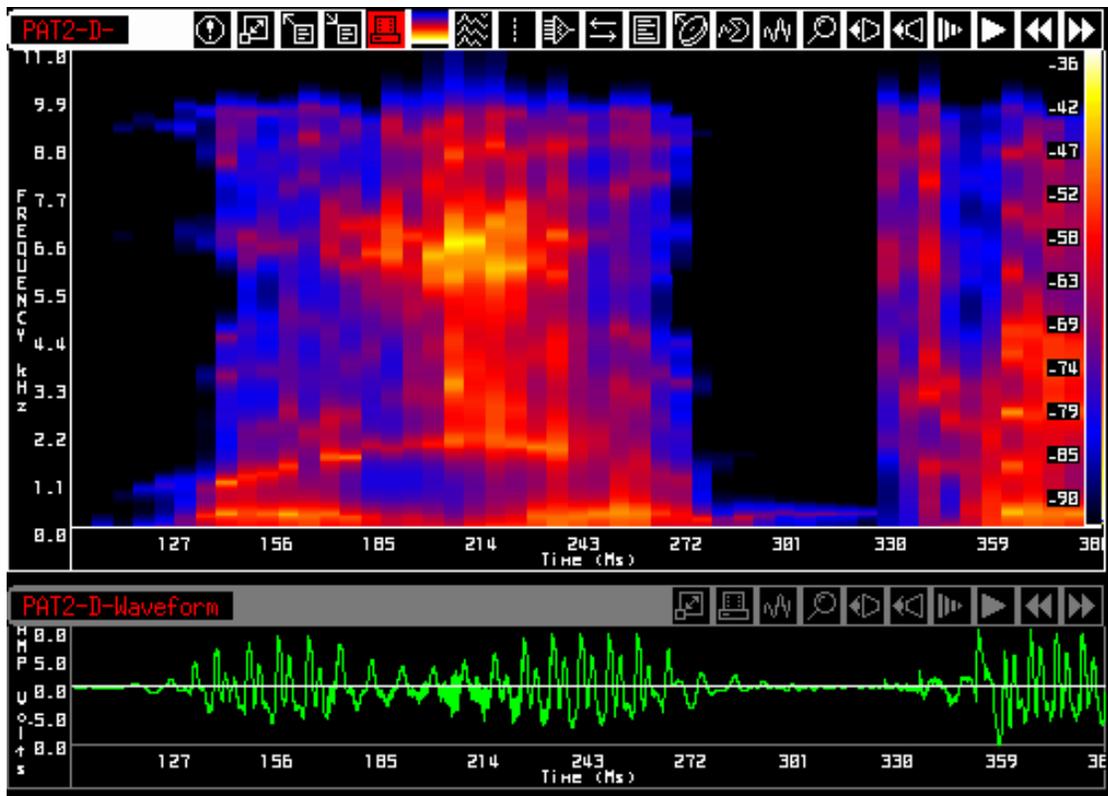


Figura12: Espectrograma de banda larga e oscilograma de parte da emissão  
 “Would you please explain this word” da Grav2 do sujeito Paula

A constatação de que o contexto fonético influi no resultado do aluno explica a gradação dos resultados globais relativos as 3 emissões *Yes/No* do experimento, mas não explica o motivo de a mesma oração, emitida pelo mesmo sujeito, com curva ascendente nas gravações 1 e 2, receber avaliação bastante diferenciada.

As avaliações relativas à oração *“Could I borrow some sugar?”* tornam bastante evidente a questão.

TABELA 17: Total Geral por aluno das avaliações da emissão *“Could I borrow some sugar”* nas Grav1 e 2

CAN I BORROW SOME SUGAR			
ALUNOS	GRAVAÇÃO1	GRAVAÇÃO2	DIFERENÇA
RITA	23	31	+ 8
PAULA	29	18	- 11
LEILA	28	21	- 7
EDA	44	24	- 20
ANA	33	30	- 3
CÉLIA	35	31	- 4

Os dados de Rita e Eda, expostos na Tabela 17 acima, surpreendem porque contrariam as expectativas. Rita foi o sujeito que apresentou melhor aproveitamento no experimento. No entanto, na oração considerada fácil pelos resultados da análise perceptiva, na qual todos os outros participantes foram melhor avaliados, apresentando maior diferença na pontuação entre as gravações 1 e 2, o sujeito Rita

recebe 8 pontos a mais na gravação<sup>2</sup>, indicando uma piora após o treinamento, pois os graus mais altos indicam mais sotaque.

O sujeito Eda, pelo contrário, surpreende pela melhora em 20 pontos, a qual não foi registrada em nenhuma outra emissão do sujeito em questão ou dos outros participantes do experimento. Essa melhora é surpreendente, também, porque a avaliação do todo da produção do sujeito não apresentou diferença significativa com o treinamento.

Procurando-se respaldo para a avaliação perceptiva dos falantes nativos na análise acústica dessas emissões dos dois sujeitos, analisou-se o fenômeno da pausa. Conclui-se que a localização das pausas não explica tal diferença nas avaliações, pois todos os alunos do experimento colocaram-nas em posição idêntica à dos falantes nativos das sentenças modelo.

No entanto, a duração da pausa parece interferir, porque a segunda emissão do sujeito Eda, que foi avaliada 20 pontos melhor do que a primeira, introduz uma pausa silenciosa de 303 milissegundos no final da frase entoacional intermediária "*Could I borrow*" muito mais longa do que a da emissão modelo que, de acordo com nossa avaliação auditiva, evoca uma situação de fala espontânea. Tal pausa deve ter sido interpretada como intencional pelos avaliadores, os quais lhe conferiram sentido, uma vez que a emissão foi muito melhor avaliada do que a emissão da gravação<sup>1</sup>. Compare-se a forma da onda da emissão com pausa na Figura 14 com a emissão da gravação<sup>1</sup>, sem pausa, na Figura 13, abaixo.

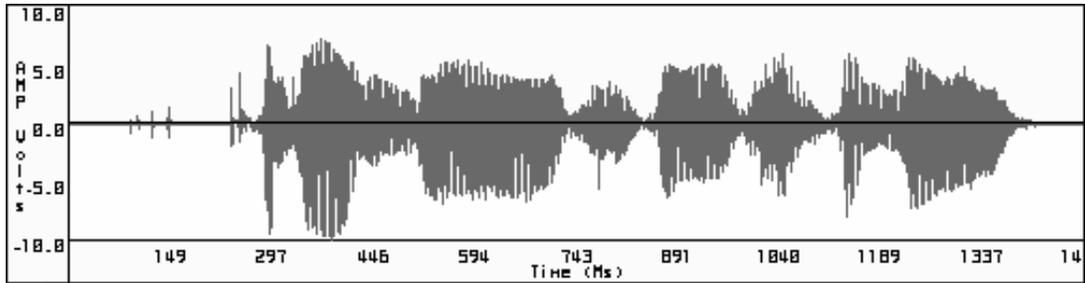


FIGURA 13: Grav1 do sujeito Eda da emissão "Could I borrow some sugar"

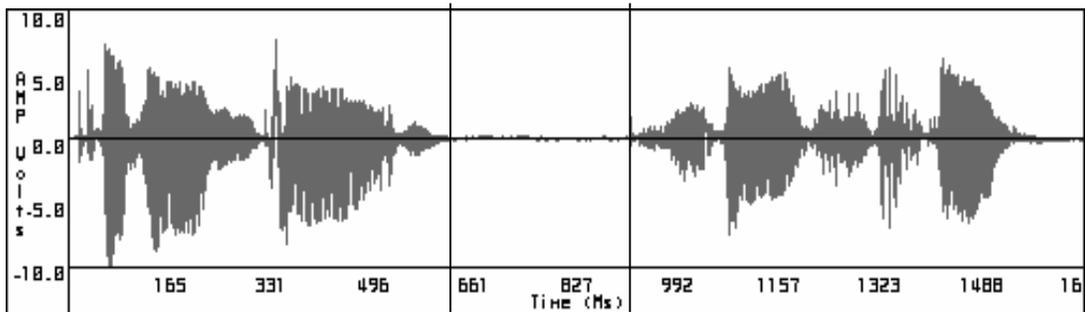


FIGURA 14: Grav2 do sujeito Eda com pausa silenciosa de 303ms entre a frase entoacional "Could I borrow" e "some sugar"

O contorno de F0 na fronteira da pausa também se mostrou importante na percepção do falante nativo. Comparando-se os dados da extração dos valores de F0 das gravações 1 e 2 da emissão do sujeito Rita, observa-se que, na gravação 2 (Figura 16), ocorre uma pausa perceptível, sinalizada por elevação de *pitch* e tom final de fronteira H-H%. No contorno de F0, verificou-se um aumento de 52Hz entre os núcleos vocálicos de "borrow", o que pode ter levado a atribuição de sentidos tais como expectativa, não-continuidade, lista de opções etc., sentidos não confirmados pelo final da emissão. Tal aumento acentuado de F0 não se verificou na gravação 1 (Figura 15), onde constatou-se um aumento de apenas 18Hz entre os referidos grupos vocálicos. Portanto, o contorno entoacional H\* H\* H-H% , típico da interrogativa Yes/No do inglês, é reproduzido.

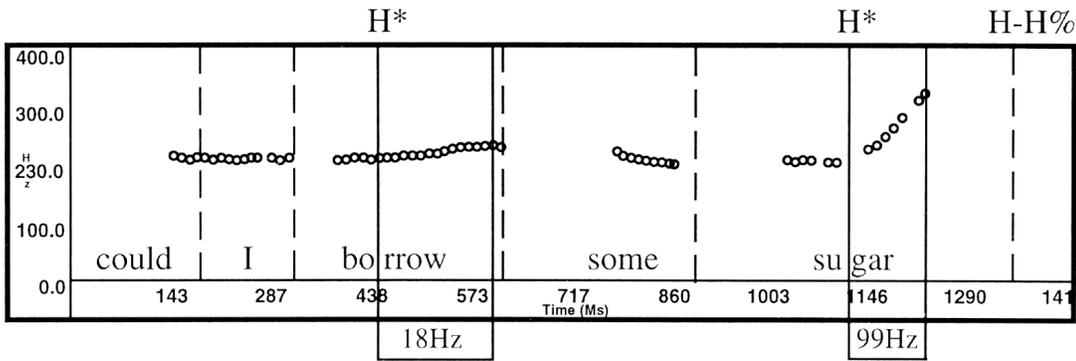


FIGURA 15: Curva de pitch do sujeito Rita da Grav1 da emissão "Could I borrow some sugar"



FIGURA 16: Curva de pitch do sujeito Rita da Grav2 da emissão "Could I borrow some sugar"

Comparando-se os valores de F0 de segmentos anteriores e posteriores à pausa perceptível no interior dos enunciados, nota-se que as emissões dos alunos do experimento que apresentam valores aproximados nesses dois contextos, ou seja, que reiniciam ("resetting") a nova frase entoacional com valores aproximados, são mais bem avaliadas porque mantêm o contorno global H\* H\* H-H%. Este é o caso de Rita, exemplificado na figura 15 acima, e o caso de Leila na gravação 2 de "Are you going to take a vacation?" ilustrado na figura 17 abaixo.

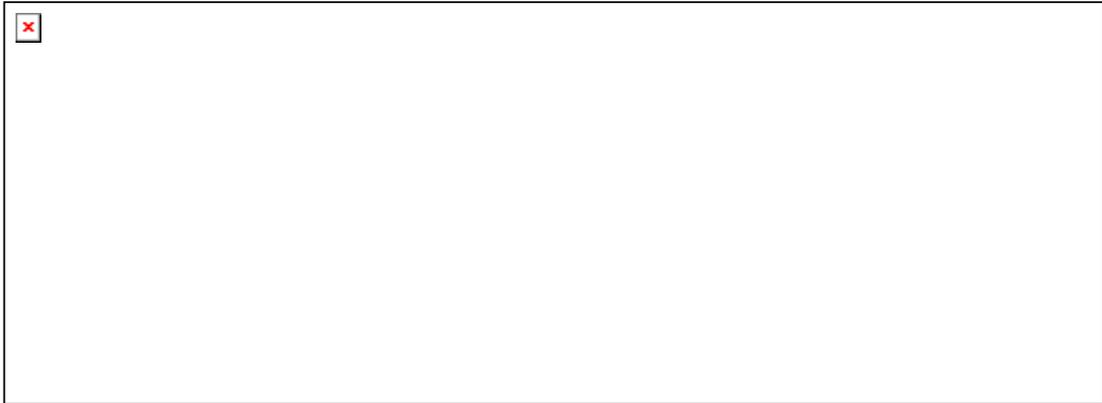


FIGURA 17: Curva de pitch da emissão “Are you going to take a vacation?” do sujeito Leila, Grav2

Inspecionando-se as gravações 1 e 2 do experimento, no que se refere ao intervalo entre o início da subida de F0 e o ponto da emissão que atinge o valor máximo de F0 na fronteira final das interrogativas *Yes/No*, nota-se que as emissões que apresentam aumento maior no valor de F0, em espaço de tempo menor, recebem melhor avaliação, tal como exemplificado nas figuras 18 e 19 abaixo.

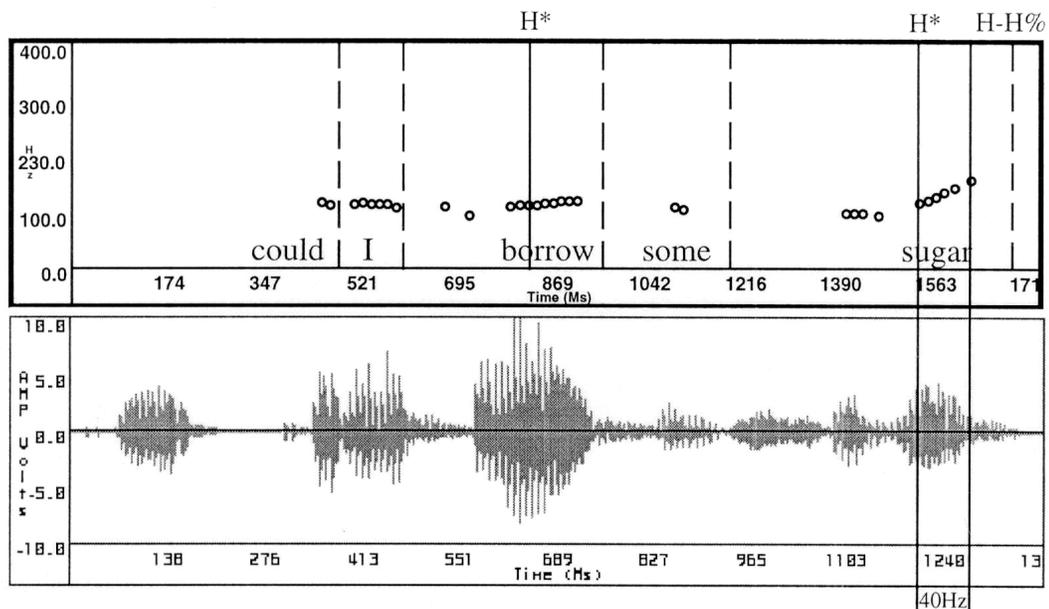


FIGURA 18: Curva de *pitch* da emissão “Could I borrow some sugar?” do sujeito Paula, Grav1.

Subida de 40Hz em 101ms no contorno final de fronteira

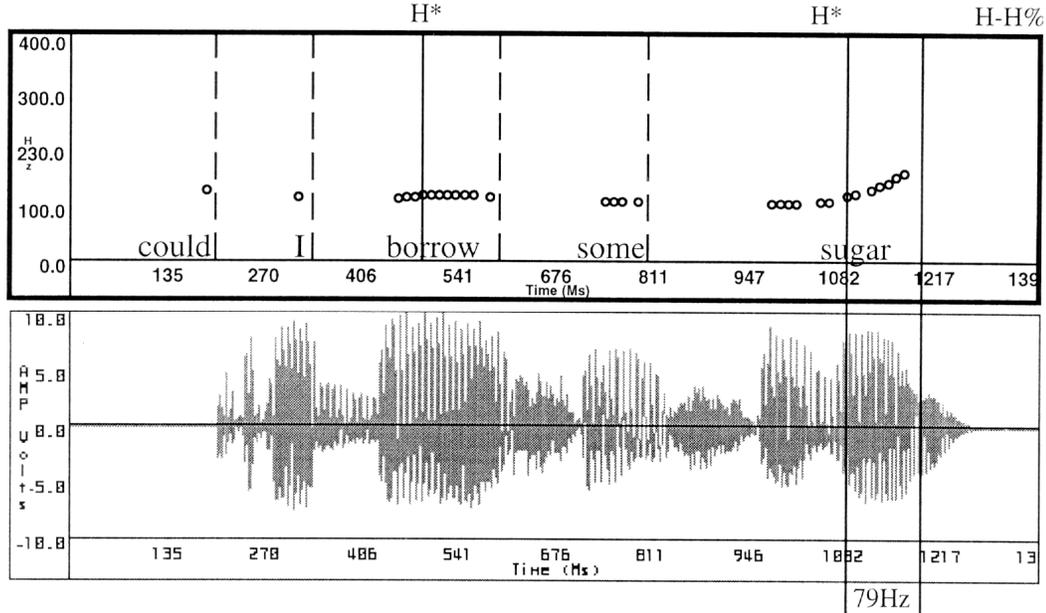


FIGURA 19: Curva de *pitch* da emissão “Could I borrow some sugar?” do sujeito Paula, Grav2. Subida de 79Hz em 150ms no contorno final de fronteira

Enquanto, na gravação 1, o sujeito eleva a frequência fundamental em 40Hz em 101ms, na gravação 2, melhor avaliada, a elevação é de 79Hz em 150ms. Também os dados de Rita, expostos nas Figuras 15 e 16 da página 164, apresentam essas mesmas características. Na gravação1 (Figura 15) melhor avaliada, a aluna eleva a frequência fundamental em 99Hz em cerca de 84ms, enquanto na gravação 2 (Figura 16) essa elevação é mais lenta (89Hz em cerca de 100ms).

Os dados do experimento parecem apontar para a conclusão de que, quanto mais fatores conjugados a emissão apresentar, tais como contorno entoacional global ascendente, movimento do *pitch* final amplo e em espaço de tempo curto, pouca alteração na altura do *pitch* após a pausa de fronteira medial na frase, menos interferência da língua nativa nos efeitos de coarticulação possibilitando a redução

de segmentos não acentuados, melhor avaliadas são as emissões dos alunos.

Dessa maneira é possível explicar a avaliação surpreendente de Eda, referente à gravação 2 da oração “*Could I borrow some sugar*” que apresenta muitas dessas características:

- curva final ascendente,
- demora de 59 ms para subir 143Hz;
- a sílaba anterior à pausa de final de frase entoacional intermediária tem o mesmo valor de F0 da sílaba que inicia a nova frase entoacional (236Hz) .

Além disso, a pausa silenciosa inesperada recebe interpretação adequada, não interferindo na avaliação dos falantes nativos do inglês.

Por outro lado, a gravação 2 de Rita, pior avaliada, apresenta as seguintes características:

- contorno global ascendente de final de fronteira (da mesma maneira que a gravação1);
- aumento de 89Hz em cerca de 100ms, isto é, a subida do valor de F0 é menor e leva mais tempo do que na gravação1;
- introdução de tom final de fronteira H-H%, onde o falante nativo apresenta apenas um acento de pitch H\* em frase entoacional intermediária, o que dificultou a comunicação.

Apesar de uma série de fatores interferirem na avaliação das emissões dos alunos, o contorno entoacional global ascendente continua sendo o fator mais marcante que não pode estar ausente, no caso das interrogativas *Yes/No*, pois, a

presença dos outros fatores não exerce efeito compensatório. Assim, os dados de Célia, exemplificados nas figuras 5 e 6 , repetidas abaixo, indicam que, na gravação1, melhor avaliada e com curva final ascendente, o aumento dos valores de F0 é menor e em tempo maior (subida de 23Hz em 70ms) do que o que se apresenta na gravação 2 (subida de 25Hz em 57ms). Entretanto, a gravação 2 é pior avaliada porque, mesmo apresentando um aumento de F0 maior no final da emissão do que o da gravação 1, no seu todo, ela não se caracteriza por uma configuração global ascendente. Veja-se Gráfico 11.

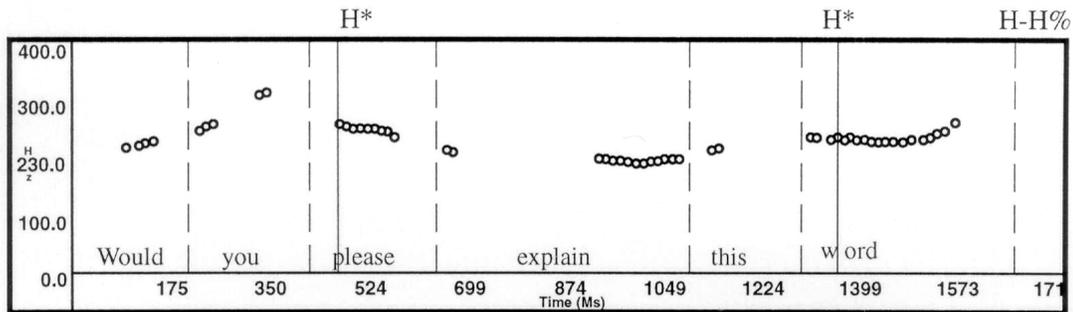


FIGURA 5: Célia, Grav1: contorno de *pitch* ascendente da emissão: *Would you please explain this word?* com subida de 23Hz em 70ms ( 213Hz no início e 244Hz no final)

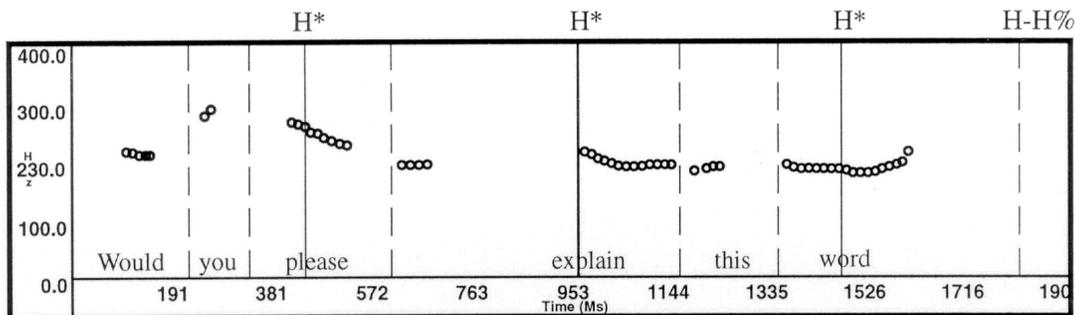


FIGURA 6: Célia, Grav2: contorno de *pitch* sem inclinação, com o início e final em torno de 210Hz, da emissão : "Would you please explain this word?" e com subida final de 25Hz em 57ms

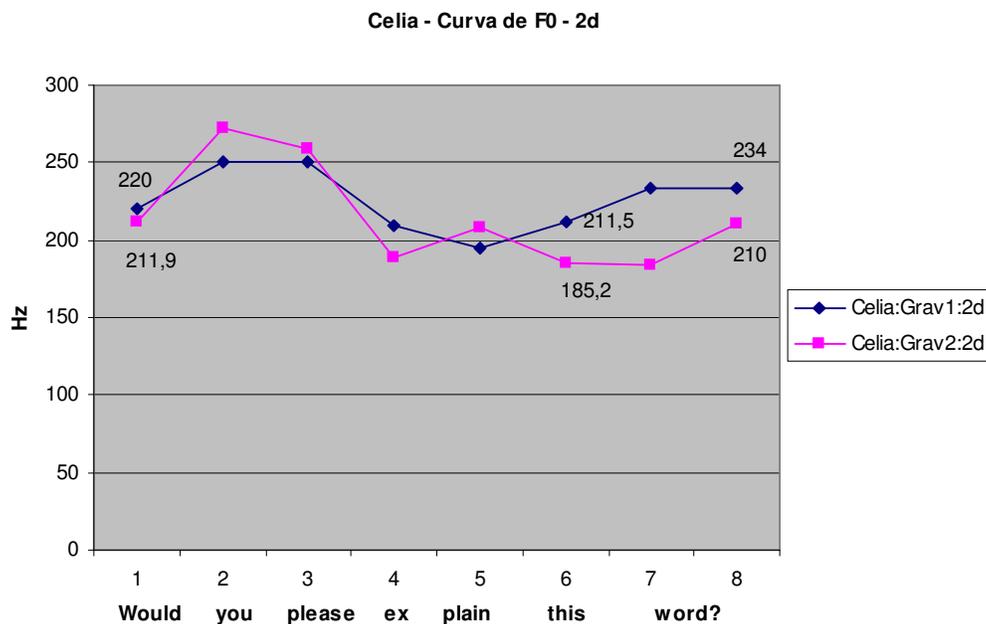


GRÁFICO 11: Comparação de dados de Célia. Grav1, melhor avaliada: curva final ascendente e subida de *pitch* de 23Hz em 70ms. Grav2: final nivelado com o início da emissão, subida de 25Hz em 57ms

Em suma, os dados apontaram para a interação de vários fatores que intervêm na percepção do sotaque estrangeiro na produção de interrogativas Yes/No: o tipo de padrão entoacional global (ascendente/descendente); o tom de fronteira final; a altura de *pitch* anterior e posterior à fronteira intermediária de enunciado; a interação entre prosódia e segmento; a variação dentro de único núcleo vocálico; o padrão acentual das palavras; o acento da frase; os fenômenos de ressilabação; os aspectos fonéticos em junturas de palavras; a articulação de segmentos. Dentre esses fatores, o contorno entoacional global ascendente mostrou-se o mais importante, porque sua ausência não é compensada pela presença de todos os outros fatores. A análise acústica dos dados das emissões interrogativas Yes/No sugere que o treinamento, baseado na análise acústica e com o apoio visual proporcionado pela tela do computador, proporciona vantagens para o ensino da entoação de uma língua estrangeira porque os sujeitos do experimento

apresentaram desenvolvimento não só em relação à implementação do padrão entoacional ascendente nas *Yes/No questions*, mas também em relação à implementação de junturas entre palavras, reduções de segmentos e colocação acentual.

### **6.3 Resultados da avaliação do treinamento pelos alunos**

O comentário dos alunos sobre o treinamento fez-se importante para a coleta de informação sobre a reação que tiveram frente a uma experiência pedagógica que propunha uma postura inovadora por parte de todos os sujeitos envolvidos, alunos e professor.

Apreciando-se as avaliações, tanto as dirigidas quanto as livres, feitas pelos sujeitos participantes no final do experimento, constata-se que todos os alunos julgaram a experiência interessante, proveitosa e estimulante. Não houve comentário negativo a não ser um único relativo à duração do treinamento que foi considerado insuficiente para a assimilação de todos os aspectos entoacionais abordados.

A experiência foi considerada positiva porque:

- permitia uma análise da própria produção:  
“Posso me corrigir e treinar sabendo onde tenho de melhorar”.
- o feedback visual proporcionou maior segurança:  
“Ali, você vê o que fala”.

- havia “feedback” de colegas e do professor, sendo possível discutir com o grupo sobre descobertas, dificuldades e resultados;
- grupo pequeno permite o atendimento personalizado do professor;
- a atividade é mais livre do que a atividade em grupo ou atividade dirigida: “Fico mais a vontade”.
- liberdade de dedicar maior atenção àquilo que o aluno considerava importante na sua emissão para atingir melhor desempenho;
- possibilidade de focar a atenção em partes pequenas da emissão para ouvir várias vezes o trecho considerado de maior dificuldade e poder repeti-lo em partes: “Posso seleccionar o trecho que apresenta dificuldade”.

Foi considerada estimulante, principalmente em relação às emissões de difícil compreensão devido às reduções, normalmente ausentes na produção dos alunos:

“Ajuda a perceber sons que nem ouvíamos”.

“Permite ouvir só as palavras que tivéssemos dificuldades de entender”.

Todos consideraram que progrediram principalmente no que se refere à atenção às pistas importantes para desenvolverem uma entoação mais adequada na língua estrangeira:

“Não consigo avaliar o quanto melhorei, mas percebo melhor os detalhes da fala”.

“O que não é comum aos nossos ouvidos não é fácil de ser assimilado”.

## CONSIDERAÇÕES FINAIS

Os resultados do experimento levam à conclusão de que a tecnologia de fala pode ser um instrumento importante no ensino de línguas, pois os alunos submetidos ao treinamento experimental apresentaram um desenvolvimento significativo em suas produções.

A introdução do apoio visual na tela do computador, proporcionado por *software* de análise acústica, auxiliou os sujeitos a transporem a dificuldade de percepção/produção de traços entoacionais da L2. Sabe-se que até os 6 meses de idade a criança faz uma seleção perceptual dos contrastes de sons pertinentes à língua ou línguas a serem aprendidas e que, a partir daí, dá-se a parametrização<sup>12</sup> (Laver, 1994) do uso que se faz do trato vocal para se atingir acusticamente os contrastes das línguas a serem faladas. Sabe-se, ainda, que é no final do período crítico de aquisição de uma língua que ocorre a fixação dessa parametrização com uma espécie de enrijecimento das deformações que a língua nativa provoca no tubo do trato vocal. Esses efeitos do período crítico têm sido atribuídos a uma perda de plasticidade neurológica relacionada à idade, a alguma espécie de reorganização neurofuncional que ocorre durante o desenvolvimento (Lenneberg, 1967).

No entanto, embora todos os sujeitos do experimento tenham estabelecido contato com a língua depois da idade de 9 anos, quando as possibilidades de deformação do tubo vocal já se enrijeceram em função dos espaços de percepção estabelecidos pela língua nativa, o treinamento proposto, neste experimento, apresentou resultado positivo, que contrasta com as conclusões de muitos estudos,

---

<sup>12</sup> Por parametrização, referimo-nos às tendências recorrentes em termos de combinações particulares de valores concernentes a parâmetros diversos que representam os padrões da fala, como, por exemplo, a estrutura formântica (Laver, 1994:101-103).

os quais, ao examinarem a influência da instrução formal no grau de sotaque estrangeiro, não apresentam resultados animadores para professores de línguas (Thompson, 1991; Flege, 1995a).

Nos últimos trinta anos, começando com o artigo de Asher & Garcia (1969) o fenômeno de sotaque estrangeiro perceptível na produção oral de falantes aprendizes de L2 tem sido investigado por um grande número de estudos experimentais detalhados (Piske, Mackay & Flege, 2001), e há boas razões para isso. Uma delas é que ele pode ajudar a resolver algumas questões teóricas referentes à aquisição de segunda língua, tais como se deve ou não haver limitações quanto à idade nessa aprendizagem. Além disso, a identificação de fatores que influenciam o grau de sotaque estrangeiro na L2 pode ser também importante para o ensino de segundas línguas.

Contudo, os estudos publicados sobre sotaque estrangeiro diferem muito em termos da natureza dos sujeitos e das línguas examinadas, assim como nos procedimentos empregados para elucidar e avaliar a fala não-nativa. De certo modo, essas diferenças metodológicas parecem ser responsáveis por pelo menos parte dos resultados freqüentemente divergentes. Então alguns apontam a variável instrucional e o número de anos de instrução como muito pouco significativa para o fenômeno do sotaque estrangeiro (Thompson 1991 Flege *et al.* 1995). Já Suter (1976) identifica duas variáveis instrucionais interagindo no fator pronúncia, sendo que uma delas, o conjunto total de treino formal na sala de aula em inglês, está inversamente relacionada com a pronúncia acurada.

---

No caso deste experimento, a variável instrucional que conjuga o treino de contornos entoacionais com o apoio visual de gráficos de F0 no ensino de entoação mostrou-se significativa, enfatizando a importância da abordagem lingüística-analítica no ensino de língua estrangeira. Aprender uma língua estrangeira depois dos primeiros anos de vida não é o mesmo que aprender a língua nativa e requer, portanto, uma prática pedagógica que repouse em um arcabouço teórico próprio.

A aprendizagem de uma segunda língua, após os cinco anos de idade, quando a gramática e o sistema fonológico da língua nativa estão adquiridos, havendo, portanto, espaço para o aumento lexical, ajustes prosódicos etc., é algo que não pode ser comparado ao processo de aquisição de língua materna. Assim, não é possível deixar de levar em conta o sistema lingüístico no qual se é nativo, fluente e que funciona de modo inconsciente. No que concerne à aquisição de uma segunda língua, essa tem de iniciar de forma consciente para, mais tarde, atingir um certo nível de proficiência, no qual ela torna-se inconsciente. O aprendiz de L2 precisa, então, ser conscientizado daquilo que não pode ser mapeado de sua língua para a L2 (acento, silabificação [spi:k], etc) e que será o elemento constituinte do seu sotaque ao comunicar-se em L2.

A conscientização “daquilo que o ouvido não ouve” e “daquilo que o aparato vocal não produz” é essencial para o desenvolvimento da aquisição de línguas. A manipulação dos dados tal como permitida pelo computador e o apoio visual proporcionado pelo *software* de análise acústica permitiram que os sujeitos deste experimento se apoderassem dos dados e os manipulassem de uma maneira consciente, e que essa conscientização orientasse seu treinamento. Esse objetivo

foi de tal maneira alcançado que, em seus relatos, os alunos expressam que se tornaram mais conscientes das pistas a que devem dar atenção em relação à entoação do inglês. Eles também consideraram a estratégia empregada estimuladora e essencial para sua aprendizagem. A manipulação consciente do sinal sonoro foi uma prática constante, entendida como necessária para o desenvolvimento da percepção e conseqüente produção de traços ausentes na língua materna.

A experiência os levou a reformularem seus objetivos enquanto aprendizes e futuros professores de língua estrangeira. Tomam consciência de que não serão capazes de reproduzir exatamente o modelo nativo e se colocam uma nova meta, a de atingir uma entoação adequada à comunicação efetiva em língua estrangeira, na qual a intenção pretendida pelo falante seja captada pelo ouvinte. Nas palavras de Paula: o objetivo não é a pronúncia “perfeita”. Aprendem as pistas relevantes para orientar o seu desempenho, mas também estão cientes de que o treinamento não garante a incorporação inconsciente à suas produções.

A proficiência, ou seja, a incorporação do que foi treinado conscientemente, não poderá ser alcançada em sala de aula. O emprego proficiente da língua pelo falante não nativo vai ser determinado pelo conhecimento de mundo que subjaz àquela língua e pela especificidade de seu uso nas situações em que seus falantes nativos a empregam, em contatos com seus produtos culturais (trocas sociais, música, cinema, literatura, teatro).

Embora haja pouca evidência para se afirmar categoricamente que a instrução formal, como tal, afeta o nível de sotaque estrangeiro, este experimento

indica que o ensino embasado em treino que associa aspectos segmentais e supra-segmentais na percepção e produção dos sons de L2 pode ter um efeito maior na acuidade de produção de segunda língua. Também Moyer (1999), que examinou o alemão de falantes nativos de inglês, concluiu que os sujeitos que receberam treinamento supra-segmental e segmental obtiveram avaliações melhores. A relação íntima entre parâmetros segmentais e prosódicos está também indicada pelas descobertas de Missaglia (1999) que, estudando um grupo de italianos, aprendizes adultos do alemão, chegou às mesmas conclusões de Moyer (1999), além de concluir que o treinamento fonético centrado na prosódia tinha um efeito de melhora tanto nos aspectos segmentais quanto prosódicos da pronúncia dos não nativos. Até o momento, no entanto, não há como quantificar a contribuição relativa dos parâmetros segmentais e parâmetros prosódicos e fluência e grau de sotaque. Os aspectos segmentais e supra-segmentais da fala estão intimamente relacionados, de modo que é muito difícil estabelecer-se uma distinção clara entre os dois.

Este trabalho evidenciou a impossibilidade de se separar o nível segmental do nível supra-segmental na análise perceptiva dos falantes nativos do inglês e na análise acústica das emissões dos sujeitos. Quando perguntados, falantes nativos relataram que percebem sotaque estrangeiro por causa de ambos, erros segmentais e supra-segmentais. Além disso, o grau de sotaque percebido na fala não nativa será, provavelmente, também determinado pela fluência, pausa de hesitação, tais como pausas de silêncio ou pausas preenchidas, repetições, falsos começos, e taxa de elocução da fala (Hieke, 1980).

Os falantes nativos do inglês que participaram da avaliação perceptiva dos sujeitos não só expressaram com palavras, a dificuldade que tinham em fazer a

distinção entre os dois níveis, segmental e supra-segmental, como documentaram essa dificuldade ao avaliarem as emissões dos sujeitos da mesma maneira, tanto para a primeira pergunta do experimento, mais voltada para o segmento, quanto para a segunda pergunta, que procurava focalizar o nível supra-segmental.

O trabalho também, depreendeu evidência a favor da interação entre prosódia e segmento da análise acústica dos acertos ou desvios do modelo, presentes nas emissões dos sujeitos. Assim, as emissões melhor avaliadas foram as que apresentaram reduções adequadas à estrutura silábica e rítmica do inglês. Os alunos também se conscientizaram desse fato ao perceberem que, sem a redução de alguns segmentos, tornava-se impossível a reprodução da curva entoacional do falante nativo. Por vezes, essa redução não programada no universo de percepção do aluno, impedia a compreensão. Nesse caso, o uso do computador foi de extrema valia porque permitiu a segmentação de pequenos trechos que podiam ser selecionados e repetidos até que fosse depreendida a compreensão de seu significado e a maneira de sua produção pelo nativo. Assim, uma vez instrumentalizado com as pistas apropriadas, os sujeitos podiam dedicar-se ao treinamento da dificuldade específica.

O resultado de melhora significativa no desempenho das orações *Yes/No* pode ser explicado pela confluência de uma série de fatores. O conhecimento que o professor tinha da importância da curva ascendente na língua inglesa nesse tipo de estrutura, explicitado por pesquisa anterior (Rocca & Marcelino, 1997), permitiu que a atenção dos alunos fosse orientada para as pistas apropriadas que, por sua vez, são fáceis de serem visualizadas na tela do computador. O treinamento, apoiado na extração do *pitch*, por ser de fácil aplicação, conseguiu o resultado pretendido,

apesar do tempo exíguo dedicado ao trabalho. Os alunos aprenderam mais sobre as interrogativas *Yes/No*, porque o professor deu as pistas mais eficazes, o que enfatiza a importância da abordagem adequada para se conseguir atingir as metas pretendidas. A variável ensino pode ser altamente significativa no desenvolvimento dos aprendizes, no que diz respeito ao ensino de pronúncia/entoação, desde que fundamentada em descrição adequada aos métodos empregados.

Nesse tocante, a análise acústica mostra-se essencial não só como instrumento de apoio ao treinamento do aluno, mas também como instrumento de identificação dos traços da L2 que não são percebidos/produzidos pelo não-nativo, sendo, portanto, uma base importante para a construção de material adequado ao ensino da entoação.

## ANEXOS

## Anexo 1

### PROTOCOLO DE AVALIAÇÃO

Nome:

Data:

I- Ouça duplas de palavras e marque se os pares são iguais ou diferentes até o número 10. Do número 10 ao 20, apenas repita o que ouve.

	same	different	
1			1 - sick/think
2			2 - jam/dam
3			3 - beard/veard
4			4 - rich/ridge
5			5 - rich/ritz
6			6 - lion/Lion
7			7 - hair/air
8			8 - when/gwen
9			9 - catch/cats
10			10 - top/top

- 11 - cake/cake
- 12 - then/then
- 13 - Sue/zoo
- 14 - figure/finger
- 15 - den/den
- 16 - chase/chase
- 17 - ring/rink
- 18 - thank/sank
- 19 - safe/safe
- 20 - set/send

II - Ouça o modelo e repita as sentenças, procurando imitar o melhor possível o falante nativo. Algumas sentenças estão escritas, outras não. Você deve apoiar-se na leitura para imitar o modelo, quando a escrita estiver disponível.

1 - She is studying Japanese.

2 - The train is late.

3 - Tell me the truth.

4 - *(Now, write a brief summary for me). (gravação apenas)*

5 - *(When can we get together for lunch?) (gravação apenas)*

6 - *(How can I help you?) (gravação apenas)*

7 - *(Do I have to memorize all this ?) (gravação apenas)*

8 - *(Would you shut the window?) (gravação apenas)*

9 - Could you lend me a dollar?

10 - Has she finished it?

III - Você ouvirá 5 seqüências de números. Após cada uma das seqüências, você deverá repetir os números na mesma seqüência:

1- 548-32-19                    *(gravação apenas)*

2- 38-92-145                    *(gravação apenas)*

3- 638-97-54                    *(gravação apenas)*

4- 98-41-53-68    *(gravação apenas)*

5- 1-3-5-4-3-8-2-1 *(gravação apenas)*

IV- Repita os tons como se fossem a melodia de uma música:

1- tom simples baixo \_\_\_\_\_

2- Tom simples alto -----

4- tom complexo ascendente \_\_\_\_\_-----

5- tom complexo descendente -----\_\_\_\_\_

## **Anexo 2**

### QUESTIONÁRIO 1 : Dados Pessoais

1- NOME:

2- Nacionalidade: \_\_\_\_\_ Idade: \_\_\_\_\_

3- Qual a nacionalidade de seus pais?

Pai: \_\_\_\_\_ Mãe: \_\_\_\_\_

4- Que língua é falada em sua casa, com seus familiares?

5- O Inglês é a sua Primeira Língua Estrangeira?

6- Fala alguma outra língua estrangeira além do inglês? Não ( )

Sim ( ) Qual?

7- Com que idade começou a estudar Inglês?

8- Há quantos anos estuda Inglês?

9- Já morou em país de Língua Inglesa? Não ( )

Sim ( )

Por quanto tempo? \_\_\_\_\_

10- Em que situações usa o inglês oral ?

Em casa – ( )

Socialmente- ( )

Trabalho- ( )

Viagem- ( )

11- Nível de Língua no Curso da PUC/SP:

Básico I - ( )

Básico II - ( )

Intermediário I – ( )

Intermediário II – ( )

Avançado – ( )

Estudos Individuais – ( )

12- Faz algum outro curso de Língua Inglesa ?

- não ( )

- Sim ( )

Qual ? \_\_\_\_\_

Há quanto tempo? \_\_\_\_\_

Em que nível está? \_\_\_\_\_

13- Você já estudou fonética e fonologia do Inglês?

Não ( )

Sim ( ) Onde?\_\_\_\_\_.

Quanto tempo?\_\_\_\_\_.

14- Você já estudou entoação do Inglês?

Não ( )

Sim ( ) Onde?\_\_\_\_\_.

Quanto tempo?\_\_\_\_\_.

### Anexo 3

#### PROTOCOLO DE AVALIAÇÃO INICIAL E FINAL DOS SUJEITOS

a) Ficha que os alunos receberam na sala de gravação

Exercício 1- Repita as perguntas e respostas, procurando imitar o melhor possível a entoação do modelo:

*Question*

*Answer*

1- Who's that?

That's my brother.

2- Face the facts, Smith.

3- Are you feeling any better today?

No, I'm afraid not.

4- How often does your mother call?

She calls every other day.

5- Close your eyes, please.

6- You're leaving without me?

You can't go. You've got the flu.

7- How's the weather outside?

It's still snowing.

8- Please, speak slowly.

9- Have you got enough health insurance? Well, I bought a policy.

10- Why didn't she sign the lease? She says the house is just a mess.

11- Are you going to take a vacation? No, my parents are coming.

Exercício 2- Repita as sentenças, procurando imitar o melhor possível a entoação do modelo:

1

2

3

4

5

6

7

8

9

Exercício 3- Faça perguntas, afirmações, dê ordens, conforme o que lhe for pedido:

1

2

3

4

5

6

7

8

Protocolo de Avaliação Inicial e final dos sujeitos:

b) Conteúdo da Gravação que os Sujeitos Ouviram

Exercício 1- Repita as perguntas e respostas, procurando imitar o melhor possível a entoação do modelo:

*Question*

*Answer*

1- Who's that?

That's my brother.

2- Face the facts, Smith.

3- Are you feeling any better today?

No, I'm afraid not.

4- How often does your mother call?

She calls every other day.

5-Close your eyes, please.

6- You're leaving without me?

You can't go. You've got the flu.

7- How's the weather outside?

It's still snowing.

8- Please, speak slowly.

9- Have you got enough health insurance?                      Well, I bought a policy.

10- Why didn't she sign the lease?                      She says the house is just a mess.

11- Are you going to take a vacation?                      No, my parents are coming.

Exercício 2- Repita as sentenças, procurando imitar o melhor possível a entoação do modelo:

1- She's studying Japanese.

2-Tell me the truth.

3-Can you come?

4-Would you please explain this word?

5-The train is late.

6- How much do those tires cost ?

7-And this is your brother?

8-Can I borrow some sugar?

9- Be careful.

Exercício 3- Faça perguntas, afirmações, dê ordens, conforme o que lhe for pedido:

1- Tell me that you can't drink coffee. (afirmação)

2- Tell me to come at 8:00 o'clock tonight. ( ordem ).

3- Ask me why I didn't call. ( Wh-question )

4- Tell me his mother is an architect. ( afirmação)

5- Ask me if I think I can do it. ( yes-no question )

6- Tell me to stop it. ( command )

7- Ask me what kind of car I'm looking for. ( Wh- question )

8- Ask me if this is going to be on the test. ( yes-no question )

## Anexo 4

### ORGANIZAÇÃO DAS AULAS LABORATORIAIS:

1ª aula:

-Preenchimento de ficha (anexo2) com os dados pessoais dos sujeitos.

-Apresentação do problema: Aquisição de entoação de língua estrangeira.

Aula expositiva comparando-se os padrões entoacionais do inglês e português, focalizando as diferenças nos processos utilizados para produzir proeminência nas duas línguas : alteração no *pitch* em inglês e alongamento da sílaba no português.

-Apresentação da Proposta do trabalho: Treinar a percepção da entoação em língua inglesa, usando o computador como apoio, concentrando-se em dois padrões entoacionais:

- 1) Padrão (LOW/ HIGH) ascendente-descendente do acento de pitch e do contorno entoacional de orações declarativas, imperativas e questões com Wh- do inglês.
- 2) Padrão ascendente do contorno entoacional das sentenças interrogativas do tipo yes/no.

2ª aula:

Gravação inicial (anexo 3) dos sujeitos no Laboratório de Rádio.

Discussão de texto sobre Análise Acústica no qual se explica o que é um oscilograma, espectrograma, frequência fundamental.

Treinamento no manuseio do software WINSAL, o qual foi utilizado no experimento.

3ª aula:

Objetivo:

- identificar as sílabas proeminentes e a sílaba que carrega o *pitch accent*.
- perceber que tipo de movimento ocorre na sílaba com *pitch accent*.
- treinar a produção.

Treino: Wh-questions seguidas de respostas declarativas.

Sentenças extraídas da lição 1/ part III-intonation 1, do livro didático:  
Pronunciation Exercises for English as a Second Language- 2nd Edition- Gary  
Esarey- Ann Arbor- The University of Michigan Press-1996

- |                                       |                          |
|---------------------------------------|--------------------------|
| 1- What did you think of the test?    | I thought it was easy.   |
| 2- When did you thank him?            | Before I left.           |
| 3- What's wrong with the cloth?       | I think it's too thick.  |
| 4- How much is it worth?              | It's not worth anything. |
| 5- When did your tooth start to hurt? | Two or three days ago.   |

Diretrizes do treinamento:

Prezado aluno, no treinamento de hoje, deverá seguir os seguintes passos:

- 1- Ouça cada uma das sentenças, quantas vezes necessárias, e assinale as sílabas proeminentes:

- 2- De todas as sílabas proeminentes, qual é a mais proeminente:
- 3- Clique em cima dos picos do gráfico da frequência fundamental. Os picos coincidem com as sílabas que apontou como proeminentes?
- 4- Examine o gráfico da frequência fundamental e anote o que acontece com o movimento do pitch na sílaba que você considerou mais proeminente:
  - existe movimento?
  - ascendente?
  - descendente?
  - em uma única sílaba ?
- 5- Ouça a emissão do modelo novamente e repita-a. Quando considerar que a repetição estiver adequada, grave-a.
- 6- Compare o gráfico da frequência fundamental do modelo com a sua gravação:
  - os picos de proeminência estão nas mesmas sílabas?
  - os movimentos de subida e descida coincidem nas mesmas sílabas?
- 7- Se o resultado for adequado, passe para a sentença seguinte. Caso contrário, repita o treino e grave novamente.

4ª aula:

Objetivo:

- identificar as sílabas proeminentes e a sílaba que carrega o pitch accent.
- perceber que tipo de movimento ocorre na sílaba com pitch accent.
- treinar a produção.

Treino: entoação de perguntas com wh e de respostas declarativas.

As emissões examinadas foram extraídas do livro : Pronunciation Exercises for English as a Second Language- 2nd Edition- Gary Esarey- Ann Arbor- The University of Michigan Press-1996

Lição 3/ part III- intonation 1.

- |                                     |   |
|-------------------------------------|---|
| 1- Who's living there now?          | Nobody. The place is vacant.              |
| 2- How fast can you drive it?       | Only about 40 miles an hour.              |
| 3- Why did she refuse the offer?    | She said they didn't offer good benefits. |
| 4- How often do you visit New York? | I've only been there once.                |

Diretrizes do treinamento:

Os sujeitos deveriam ouvir as sentenças uma a uma, indicar quais as sílabas proeminentes, checar as respostas examinando a extração da frequência fundamental da emissão, responder se o padrão era ascendente ou descendente. Em seguida, repetir a emissão, gravá-la e comparar a extração

do *pitch* de sua emissão com a do modelo. A comparação deveria focalizar o alinhamento da sílaba proeminente e sua duração. Detectadas diferenças significantes, repetir até atingir uma produção próxima da sentença modelo.

5ª aula:

Objetivo: perceber e reproduzir o contorno ascendente da *Yes/No question*

Treino:

As emissões examinadas foram extraídas do livro : Pronunciation Exercises for English as a Second Language- 2nd Edition- Gary Esarey- Ann Arbor- The University of Michigan Press-1996

Lição 13/ Part III. Complex Sentences and Intonation –

1-Do you want to drive if Kim will lend you her car?

2- Can you rent a car if you don't have a credit card?

3- Is she going to buy it even if they raise the price?

Diretrizes do treinamento:

ouvir as emissões,

repetir, gravar

e comparar as curvas de F0, focalizando a altura do pitch no início e no final das emissões.

6ª aula:

Objetivo: perceber que a proeminência em pontos diferentes de uma mesma emissão produz contornos de F0 diferentes e significados discursivos também diferentes. Identificar em que diferentes contextos as emissões seriam empregadas. Tentar reproduzir, mas sem extrair F0 na comparação.

Treino: Escute os enunciados quantas vezes forem necessárias e em seguida:

- 1) Escreva o enunciado procurando reproduzir as reduções.
- 2) Escute novamente, e marque as sílabas proeminentes.
- 3) Desenhe o contorno global da linha entoacional, baseando-se nas sílabas proeminentes
- 4) Compare sua curva com a curva de F0.
- 5) Tente repetir as sentenças reproduzindo as contrações. Segure a seta e vá repetindo por partes.
- 6) Compare sua produção com a do original

As emissões examinadas foram extraídas do livro : Pronunciation Exercises for English as a Second Language- 2nd Edition- Gary Esarey- Ann Arbor- The University of Michigan Press-1996

Lição 9 Part 1/ 1- repetition

- 1- What are you doing this weekend?
- 2- Are you going to wash your car Saturday?
- 3- Are you going to wash your car Saturday?
- 4- Are you going to wash your car Saturday?
- 5- Do you already have tickets for Friday night's concert?

6- Do you already have tickets for Friday night's concert?

7- Do you already have tickets for Friday night's concert?

7ª aula:

Objetivo: localizar pausas em enunciados com mais de um grupo entoacional.

Treino: Sentenças declarativas com mais de uma frase entoacional.

Material extraído do Intonation Workbook do Practice Material for the ToBI (Tones and Break Indices) System . J Harrington & J. Fletcher, Speech Hearing and Language Research Centre, Macquarie University, Sydney, Australia December 1994;

10 sentenças com proeminência dada pelo *pitch low high*.

- 1) They asked if I wanted to come along on the barge trip.
- 2) Amongst her friends, she was considered beautiful.
- 3) The huge castle was encircled by a deep moat
- 4) Jane adored maths and French but hated the rest of school.
- 5) The museum hires musicians every evening.
- 6) Her sylph-like form was entrancing the young man.
- 7) I've told you a million times.
- 8) The Vikings looked for new lands with fleets of small ships.
- 9) Would a tomboy often play outdoors?
- 10) Planned parenthood organizations promote birth control.

Diretrizes do treinamento: Ouça as 10 emissões e marque onde ocorrem pausas. Discuta suas respostas com o colega do lado.

8ª aula:

Objetivo: identificar as pausas, os *pitch accents*, o tipo de *pitch accent*, e reproduzir.

Treino: sentenças de estruturas sintáticas variadas e de diferentes variantes do inglês.

Material extraído do Intonation Workbook do Practice Material for the ToBI (Tones and Break Indices) System . J Harrington & J. Fletcher, Speech Hearing and Language Research Centre, Macquarie University, Sydney, Australia December 1994;

- 1) They asked if I wanted to come along on the barge trip.
- 2) Amongst her friends, she was considered beautiful.
- 3) The huge castle was encircled by a deep moat
- 4) Jane adored maths and French but hated the rest of school.
- 5) The museum hires musicians every evening.
- 6) Her sylph-like form was entrancing the young man.
- 7) I've told you a million times.
- 8) The Vikings looked for new lands with fleets of small ships.
- 9) Would a tomboy often play outdoors?
- 10) Planned parenthood organizations promote birth control.

Diretrizes do treinamento:

Ouçã as 10 primeiras sentenças e escreva-as:

Marque as sílabas Tônicas.

Marque as pausas.

Desenhe o contorno da entoação unindo as Tônicas

Marque os *pitch accents*.

Repita as sentenças.

Grave-as e compare-as com o modelo.

9ª aula

Objetivo: identificar as pausas, os *pitch accents*, o tipo de *pitch accent*.

Treino: Ouça as sentenças e escreva-as:

Marque as sílabas tônicas.

Marque as pausas.

Desenhe o contorno da entoação unindo as tônicas.

Marque os acentos de *pitch*.

Repita as sentenças extraídas do Intonation Workbook do Practice Material for the ToBI

(Tones and Break Indices) System . J Harrington & J. Fletcher, Speech Hearing and Language Research Centre, Macquarie University, Sydney, Australia, December 1994;

1-There's a spoon in here.

2- Do atypical farmers grow oats?

3-Are you looking for employment in April?

4-You know the school, other people.

5-Oh, I don't know, it's got oregano 'n marjoram 'n some fresh basil.

6-Legumes are a good source of vitamins.

7-Do you hear the sleigh bells ringing?

8-You know flour 'n butter and sugar.

9-I thought it was good.

10-I'm not gonna drive to school today.

11-They were guided by nothing but their knowledge of the stars.

12-Bananas aren't poisonous.

13-Elephantiasis isn't incurable.

14-Do you really think it's that one?

15-Stein is not a bad man.

10ª aula :

Avaliação Livre do Treinamento (anexo 5)

Gravação final no laboratório de radio em câmara com isolamento acústico.

Questionário de Avaliação Dirigida das atividades (anexo 6).

## **Anexo 5**

AVALIAÇÃO LIVRE DO TREINAMENTO:

Como você avalia a atividade que lhe foi proposta nesse semestre?

Sinta-se à vontade para dizer o que quiser:

## Anexo 6

### AVALIAÇÃO DIRIGIDA:

1- Quantas sessões você fez?

2- Você acha que aprendeu alguma coisa? Sim ( ) não( )

3-Se sua resposta foi afirmativa, diga-me o que aprendeu.

4-Se sua resposta foi negativa, diga-me o porquê de você achar que não aprendeu nada.

5-Em sua opinião, qual, das atividades propostas, foi a :

a) mais chata- por que ?

b) mais interessante- por que?

c) mais inútil- por que?

d) mais proveitosa- por que?

6-Comparando o treino desenvolvido no laboratório de informática com o treino desenvolvido no laboratório de línguas, qual deles :

e) É mais instigante ? Por que?

f) Deixa você mais à vontade? Por que?

g) Fornece um feedback mais seguro? Por que?

7-Você considera que o treino que fizemos desenvolve a

h) percepção dos padrões entoacionais? Sim ( ) Não( ) Justifique;

i) produção dos padrões entoacionais? Sim ( ) Não( ). Justifique

8-Você acha que o trabalho que fizemos neste semestre provocou alguma alteração em você?

Obrigada pela ajuda durante o semestre.

Se permitirem que o nome de vocês seja mencionado na tese de doutoramento, escreva-o de maneira completa na folha de resposta.

Boas Férias!

## Anexo 7

### QUESTIONÁRIO 2 : DADOS PESSOAIS FALANTES NATIVOS

Name :

Age:

Nationality:

Native Language:

Speaker of Foreign Language: Yes ( )      Which one(s)

No ( )

Parents'

Native

Language:

## **Anexo 8**

EMISSÕES INCLUÍDAS NA ANÁLISE PERCEPTUAL POR FALANTES NATIVOS DO INGLÊS AMERICANO:

A- Imitação do modelo com apoio oral e escrito:

1- How often does your mother call?

2- Close your eyes, please.

3- She says the house is just a mess.

4- Are you going to take a vacation?

B- Imitação do modelo sem apoio da escrita:

1- She's studying Japanese.

2- Tell me the truth.

3- Would you please explain this word?

4- Can I borrow some sugar?

C- Produção induzida:

1- I can't drink coffee

2- Come at eight o'clock tonight.

3- Do you think you can do it?

4-What kind of car are you looking for?

D- Produção do Modelo falante nativo do Inglês:

1-How often does your mother call?

2-Are you going to take a vacation?

## **REFERÊNCIAS BIBLIOGRÁFICAS**

- AKHANE-YAMADA, R.; TOHKURA, Y.; BRADLOW, A. & PISONI. (1996). Does training in speech perception modify speech production? *Proceedings of ICSLP 96*. Philadelphia.
- ALBANO, E. C. (1990). *Da fala à linguagem - tocando de ouvido*. São Paulo: Martins Fontes.
- ALBANO, E. C; BARBOSA P; GAMA-ROSSI A; MADUREIRA, S. & A. SILVA (1998). *A Interface Fonética-Fonologia e a Interação Prosódia-Segmentos*. In: Estudos Lingüísticos XXVII: Anais do XLV Seminário do Grupo de Estudos Lingüísticos do Estado de São Paulo, GEL, São José do Rio Preto IBILCE,135-143.
- ALBANO, E. C. (2001). *O gesto e suas bordas*. Editora Mercado de Letras Edições e Livraria Ltda
- ARMSTRONG, L. E. & WARD, I. C. (1926). *A handbook of English intonation*. Cambridge: Heffer.
- ASHER, J. J. & GARCIA R. (1969). The optimal age to learn a second language. *The Modern Language Journal*, 53, p. 334-341.
- ASLIN, R. N; PISONI, D. & JUSCZYK, P. (1983). Auditory development and speech perception in infancy. In: Haith, M. M. & Campos, J. J. (eds.). *Infancy and the biology of development*. New York: Wiley.
- BALL, M. J. & KENT R. D. (1997). *The new Phonologies. Developments in Clinical Linguistics*. Singular Publishing Group. Inc. San Diego –London.
- BARBOSA, P. A. (1996). At least two macrorhymic units are necessary for modeling brazilian portuguese duration. *Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech Production Modeling - 4th Speech Production Seminar*.
- BARBOSA, P. A. & BAILLY, G. (1994). Characterization of rhythmic patterns for text to – speech synthesis. *Speech Communication* 15, p.127-132.
- BAUM & PELL (1999). The neural bases of prosody. Insights from lesion studies and neuroimaging. In: Lynch, M. (ed.). *The cognitive science of prosody: Interdisciplinary perspectives*. Amsterdam: Elsevier North Holland Science Publisher.
- BECKMAN, M. E. (1995). Problems of intonation. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Sweeden, I, p. 450-457.
- BECKMAN, M. & AYERS, G. (1994). *Guidelines to ToBI labelling. Version 2.0*. Ohio:

- Ohio State University.
- BECKMAN, M. & HIRSCHBERG, (1994). *The ToBI annotation conventions*. Ohio: Ohio State University.
- BERNSTEIN, J. (1998). New uses for speech technology in language education. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweden: Marholmen Conference Center.
- BEST, C. T. (1999). Development of language-specific influences on speech perception and production in pre-verbal infancy. *Proceedings of the XIVth International Congress of Phonetic Sciences (ICPhS)*, 2, p.1261-1264. San Francisco.
- BEST, C. T. (1995). A direct realist view of cross-language speech perception. In: W. Strange, (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. York Press, Timonium, MD, p 171 – 204.
- BLAIR, R. W. (1991). Innovative approaches. In: Celce Murcia, (ed.). *Teaching English as a second or foreign language*. 2nd. ed. New York: Newbury, p. 23-45.
- BLOOM, L. & LAHEY, M. (1978). *Language development and language disorders*. New York: John Wiley.
- BLOOM, L.; HOOD, L.; & LIGHTBOWN. (1974). Imitation in child language: if, when, and why. *Cognitive Psychology*, p. 380-420.
- BLOOMFIELD, L. (1933). *Language*. New York: Holt.
- BOHN, O. S. & FLEGE, J. E. (1990). Perception and production of a new vowel category by adult second language learners. In: Leather, J. & James (eds.) *New Sounds 90. Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second –Language Speech*. Amsterdam: University of Amsterdam Press, p. 37-56.
- BOLINGER, D. L. (1996). *Intonation and its parts*. Edward Arnold Publishers.
- BOLINGER, D. L. (1975). *Aspects of language*. 2nd. ed.. New York: Harcourt Brace.
- BORREL, A. (1990). Perception et (re)production dans l'apprentissage des langues étrangères. Quelques réflexions sur les aspects phonético-phonologiques. *Revue de Phonétique Appliquée* 95-96-97, p. 107-114.
- BRAZIL, D.; COULTHARD, M & JOHNS, C. (1980). *Discourse intonation and language teaching*. London: Longman.

- BRIÈRE, E. (1966). An investigation of phonological interference. *Language*, 42 (4), p.769-796.
- BROWMAN, C. P. & GOLDSTEIN L. M.. (1986). Towards an articulatory phonology. *Phonology Year Book*, 3:219-252.
- BROWMAN, C. P. & GOLDSTEIN L. M. (1990) Tiers in articulatory phonology, with some implications for casual speech. In: J Kingston and M. Beckman (eds.) *Papers in Laboratory Phonology: Between The Grammar and the Physics of Speech*, Cambridge: Cambridge University Press. p. 341-376.
- CELCE-MURCIA, M.; BRINTON, D. M. & GOODWIL, J. M. (1996). *Teaching pronunciation: a reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- CHOMSKY, N. (1986). *Knowledge of language: its nature, origin and use*. New York: Praeger.
- CHOMSKY, N. (1981). *Lectures on governing and binding*. Dordrecht: Foris.
- CHOMSKY, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- CHOMSKY, N. (1959). Review of verbal behaviour. *Language* 35, p. 26-58.
- CHOMSKY, N. (1957). *Syntactic structures*. Mouton: The Hague.
- CLARK, J. & YALLOP, C. (1995). *An introduction to phonetics and phonology*. Oxford: Blackwell.
- COOK, V. (1991). *Second language learning and language teaching*. London: Edward Arnold .
- CRUTTENDEN, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- CRUTTENDEN, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language*, 12, p. 645-661.
- DAVIES, S. & POESIO, M. (1998). A CSLUrP-based spoken dialogue system for teaching English as a foreign language. In: *Proceedings of the European Speech Community Association on Speech Technology in Language Learning*. Sweeden: Marholmen Conference Center.
- DE LEMOS, C. (1986). Interacionismo e aquisição de linguagem. *D.E.L.T.A.* São Paulo: EDUC, 2 (2), p. 231-248.

- DEVLIEGER, M. (1998). The applicability of speech recognition in the context of task-based language learning for young children. *Proceedings of the European Speech Community Association on Speech Technology in Language Learning*. Suécia: Marholmen Conference Center.
- ELMAN, J. L. (1995). Language as a dynamical system. In: Port, R. F. & Van Gelder, T. (eds.). *Mind as motion: explorations in the dynamics of cognition*. Cambridge: MA: MIT Press, p. 195-223.
- ELSEENDOORN, B. A. G. (1984). Production and perception of English vowel duration by dutch speakers of english. In: Van Den Broeck, M. P. R & Cohen, A. (eds.). *Proceedings of The Tenth International Congress of Phonetic Sciences*. Dordrecht: Foris, p. 673-676.
- ESCA WORKSHOP (1998). Proceedings of the European Speech Community Association on Speech Technology in Language Learning. Sweeden: Marholmen Conference Center.
- ESKNAZI, M. (1996). Detection of foreign speakers pronunciation errors for second language training- preliminary results. *Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech-Production Modeling - 4th Speech Production Seminar*.
- ESKNAZI, M. & HANSMA, S. (1998). The fluency pronunciation trainer. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweeden: Marholmen Conference Center
- FLEGE, J. E.; YENI-KOMSHIAN, G. & LIU, H. (1999). Age constraints on second language acquisition. *Journal of Memory and language*. 41, p.78-104.
- FLEGE, J. E. (1995a). Second language speech learning: theory, findings and problems. In: Strange, W. (ed.). *Speech perception and linguistic experience: theoretical and methodological Issues*. Timonium, MD: York Press, p. 233-277.
- FLEGE, J. E.; MUNRO, J. R. & MACKAY, I. R. A. (1995b). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, p. 3125-3134.
- FLEGE, J. E. (1991). Perception and production: the relevance of phonetic input to L2 phonological learning. In: Hueber, T. & Ferguson, T. (eds.). *Crosscurrents in second language acquisition and linguistic theories*. Amsterdam: John Benjamins. p. 249-289.
- FLEGE, J. E. (1988). Factors affecting degree of perceived foreign accent in english

- sentences. *Journal of the Acoustical Society of America*. 54 (2), p.70-79.
- FLEGE, J. E. (1987). A critical period for learning to pronounce foreign languages?. *Applied Linguistics*, 8, p.162-177.
- FODOR, J. (1983). *The modularity of mind*. Cambridge: MIT Press.
- FRY, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27, p. 765-768.
- GAMA, A. (1989). *Fala e ação no cuidado materno ao bebê*. Dissertação de mestrado. UNICAMP, Campinas.
- GASS, S. M. & SCHACHTER, J. (eds.). (1989). *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press.
- GATTEGNO, C. (1972). *Teaching foreign languages in schools: the silent way*. New York: Educational Solutions.
- GATTEGNO, C. (1976). *The common sense of teaching foreign languages*. New York: Educational Solutions.
- GIBSON, J. J. & GIBSON, E. J. (1955). Perceptual learning: differentiation or enrichment? *Psychological Review*, 62, p. 32-42.
- GOLDSMITH, J. (1976). An overview of autosegmental phonology. *Linguistic Analysis*, 2, p. 23-68.
- GONÇALVES, M. J. (1989). A construção da fala por uma criança. Dissertação de mestrado. UNICAMP, Campinas.
- GRABE, E. (1997). Comparative intonational phonology: English and German. *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Applications*. Athens, p. 157-160.
- GRABE, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics*, 26, p. 129 – 143.
- GUSSENHOVEN, C.; REPP, B. H.; RIETVELD, T. RUMP, H. H. & TERKEN, J. (1997). The perceptual prominence of fundamental frequency peaks. *Journal of the Acoustic Society of America*, 102 (5), p. 3009-3022.
- HAGEN, S. & GROGAN, P. (1992). *Sound advantage - a pronunciation book*. Prentice Hall Regents.

- HARDCASTLE, W. J. & HEWLETT N (1999). *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press.
- HATCH, E. (1992). *Discourse and language education*. Cambridge: Cambridge University Press.
- HIEKE, A. E. (1980). Aspects of native and non-native fluency skills. PhD Dissertation, University of Kansas
- HIRST, D. & DI CRISTO, A. (1998). A survey of intonation systems. In: Hirst, D & Di Cristo, A. (eds.). *Intonation systems: a survey of twenty Languages*, p.1-44. Cambridge. Cambridge University Press.
- ITARD, J. M. G. (1962). *The wild boy of aveyron*. New York: Meredith House.
- JAKOBSON, R. (1967) In: Mattoso, C. J. J. (ed.) *Fonema e Fonologia*. Rio de Janeiro: Acadêmica
- JAMIESON, D. G. (1995). Techniques for training difficult non-native speech contrasts. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Stockholm, Sweden, 4, p. 100 - 107.
- JENKINS, J. (2000). *The phonology of English as an international language: new models, new norms, new goals*. Oxford: Oxford University Press.
- JILKA, M. & MÖHLER, G. (1998). Intonational foreign accent: speech technology and foreign language teaching. *Proceedings of ESCA Workshop on Speech Technology in Language Learning - Sweden: Marholmen Conference*.
- JILKA, M. (2000). *The contribution of intonation to the perception of foreign accent*. Tese de Doutorado. Universidade de Stuttgart. Alemanha.
- JONES, D. (1918). *An outline of English phonetics*. Cambridge: Heffer
- JONSSON, I. (1998). Multi-sensory stimulation of voice, speech and sounds from surroundings in hard-of-hearing preschool children. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweden: Marholmen Conference Center.
- KELLY, L. G. (1969). *25 centuries of language teaching*. Rowley, Ma: Newbury House.
- KRASHEN, S. D. (1985). *The input hypothesis*. London: Longman.

- KRASHEN, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- KRASHEN, S. D. (1973). Lateralization, language learning, and the critical period: some new evidence. *Language Learning*, 23, p. 63-74.
- KUHL, P. K. e IVERSON, P. (1995). Linguistic experience and the perceptual magnet Effect. In: Strange, W. (ed.). *Speech perception and linguistic experience: theoretical and methodological issues*. Timonium, MD: York Press, p. 121-154.
- KUHL, P. K. (1991a). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, p. 93-107.
- KUHL, P. K. (1991b). Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech. In: Brauth, S. E.; Hall, W. S. & Dooling, R. J. (eds.). *Plasticity of development*. Cambridge: MA, MIT Press.
- KUHL, P. K. (1987). The special-mechanisms debate in speech research: categorization tests on animals and infants. In: Harnad, S. (ed.). *Categorical perception: the groundwork of cognition*. New York: Cambridge University Press.
- LADD, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- LADEFOGED, P. (1993). *A course in Phonetics*. 3rd. ed. London: Harcourt Brace Jovanovich.
- LADO, R. (1964). *Language teaching: a scientific approach*. New York: McGraw-Hill.
- LANTOLF, J. P. & APPEL, G. (1994). *Vygotskian approaches to second language research*. Norwood, N. J: Ablex.
- LAVIER, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.
- LAVIER, J.; HILLER, S. M; MACKENZIE, J. & ROONEY, E. (1986). An acoustic system for the detection of laryngeal pathology. *Journal of Phonetics*, 14, p. 517-524.
- LEMOS, C. (1986). Interacionismo e aquisição de linguagem. *D.E.L.T.A.*. São Paulo: EDUC, 2, p. 231-248.
- LENNEBERG, E. (1967). *The biological foundations of language*. New York: John Wiley.

- LIBERMAN, A. & MATTINGLY, I. (1985). The motor theory of speech perception revised. *Cognition*, 21, p. 1-36.
- LIEFF, C. D. (2003). O ensino da pronúncia do inglês numa abordagem reflexiva. In: Celani, M. A. A. (org.). *Professores e Formadores em Mudança: Relato de um processo de reflexão e transformação da prática docente*. Campinas: Mercados de Letras.
- LIGHTBOWN; PATSY, M. & SPADA, N. (1999). *How languages are learned. handbooks for language teachers*. Oxford: Oxford University Press.
- LLISTERRI, J. (1995). Relationships between speech production and speech perception in a second language. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Sweden, (4), p. 92-99.
- LOCKE, J. (1997) *Brain and Language*, 5:.p.265-326.
- LONG, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, p. 126-41.
- MACK, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception and Psychophysics*, 46 (2), p.187-200.
- MADUREIRA, S. (1999). Entoação e síntese de fala: modelos e parâmetros. In: Scarpa, E. M. (org.). *Estudos de prosódia*.Campinas: UNICAMP, Campinas.
- MAGEN, S. H. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26, p. 381-400.
- MARKHAM, D. J. & MADSEN, Y. N. (1996). Input modality effects in foreign accent. *Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech-Production Modeling- 4th Speech Production Seminar*.
- MASSARO, D. (1994). Psychological aspects of speech perception- implications for research and theory. In: Gernsbacher (ed.). *Handbook of psycholinguistics*. NewYork Academic Press, p. 219-263.
- MASSINI, G. (1991). *A Duração no Estudo do Acento e do Ritmo em Português*. Tese de Mestrado. UNICAMP, Campinas.
- MAYER, J.; DOGIL, G; WILDGRUBER, D; RIECKER, A; ACKERMANN, H. & GRODD, W. (1999). Prosody in speech production: a fMRI study. *Proceedings*

- of the XIVth International Congress of Phonetic Sciences. San Francisco: vol.1, p. 635-638.
- MCALLISTER, R. (1998). Second language perception and the concept of foreign accent. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweden: Marholmen Conference Center.
- MCLAUGHLIN, B. (1987). *Theories of second language learning*. London: Edward Arnold.
- MCNERMEY, M. & MENDELSON, D. (1992). Suprasegmentals in the pronunciation class: setting priorities. In: Avery, P. & Ehrlich, S. (eds.). *Teaching american English pronunciation*. Oxford: Oxford University Press. p. 185-196.
- MIRA MATEUS, M. H. (1990). *Fonética, Fonologia e Morfologia do Português*. Lisboa: Universidade Aberta.
- MISSAGLIA, F. (1999). Contrastive prosody in SLA- an empirical study with adult Italian learners of German. *Proceedings of the XIVth International Congress of Phonetic Sciences*. San Francisco, vol. 1, p. 551-554.
- MIXDORFF, H. (1996). Foreign accent in intonation patterns - a contrastive study applying a quantitative model of the F0 contour. *Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech-Production Modeling - 4th Speech Production Seminar*.
- MOYER, A. (1999) Ultimate attainment in L2 phonology . *Studies in Second Language Acquisition*, 21, p. 81-108.
- NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- NEUFELD, G. G. (1988). Phonological asymmetry in second language learning and performance. *Language Learning*, 38 (4), p. 531-559.
- OCKE-SCHWEN BOHN (1995). What determines the perceptual difficulty encountered in the acquisition of non-native contrasts? *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Sweden, (4), p. 85-91.
- O'CONNOR, J. D & ARNOLD, G. F. (1973). *Intonation of colloquial English*. London: Longmans.
- PENFIELD, W. & ROBERTS, L. (1959). *Speech and brain mechanisms*. NJ, Princeton: Princeton University Press.
- PIAGET, J.(1923). *Le langage et la pensée chez l'enfant*. Neuchâtel: Delachaux et Niestlé, 1956.

- PICA, T. (1994). Research on negotiation: what does it reveal about second language acquisition/ conditions, processes, and outcomes. *Language Learning* ,44, p. 493-527.
- PIERREHUMBERT, J. (1981). Synthesizing intonation. *Journal of the Acoustic Society of America* 70, p. 985-995.
- PIERREHUMBERT, J. (1980). *The phonology and phonetics of English intonation*. Ph. D. Dissertation, Cambridge: MA- M.I.T Press.
- PIERREHUMBERT, J. (1979). The perception of fundamental frequency delineation. *Journal of the Acoustic Society of America* 66, p. 363-369.
- PIERREHUMBERT, J. & HIRSCHBERG, J. (1990). The meaning of intonation contours in the interpretation of discourse. In: Cohen, P. R.; Morgan, J. & Pollack, M. E. (eds.). *Intentions in communication*. Cambridge: MA- MIT Press. p. 271-331.
- PIKE, K. L. (1945). *The intonation of American English*. Michigan: University of Michigan, Ann Arbor.
- PINKER, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*. Viking Press.
- PINKER, S. (1994). *The Language Instinct*. New York: Penguin Books.
- PISONI, D. (1977). Identification and discrimination of the relative onset times of two component tones: implications for voicing perception in stops. *Journal of the Acoustic Society of America*, 61, p. 1352-1361.
- PISKE, T.; MACKAY, I. R. A. & FLEGE, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, p.191-215.
- PRICE, P. (1998). How can speech technology replicate and complement good language teachers to help people learn language?. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweeden: Marholmen Conference Center, p. 81.
- RAMUS, F. & MEHLER, J. (1999). Language identification with suprasegmental cues: a study based on speech resynthesis. *Journal of the Acoustic Society of America*, 105 (1), p. 512-521.
- READ, C.; BUDER, E. H.; KENT, E. R. D. (1990). Speech analysis systems: a survey. *Journal of Speech and Hearing Research*, 33, p. 363 -374, June.

- ROCCA, P. D. A. (2003). *O desenvolvimento de falantes bilíngües: evidências advindas da investigação do VOT de oclusivas surdas do inglês e português*. No prelo.
- ROCCA, P. D. A. (1998). The efficacy of computer-driven visual feedback in the teaching of intonation to Brazilian learners of English. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweeden: Marholmen Conference Center.
- ROCCA, P. D. A. & MARCELINO, M. (1997). Teaching intonation to Brazilian learners of English. *Proceedings of ESCA Tutorial and Research Workshop on Intonation: Theory, Models and Applications*. Athens, Greece, p. 223-226.
- ROSS, E. (1980). The aprosodias: functional-anatomical organization of the affective components of language in the right hemisphere. *Archives of Neurology*, 38, p. 561-569.
- RUMELHART D. E. & MCCLELLAND, J. L. (1986). On learning the past tenses of English verbs. In: McClelland, J. L. & Rumelhart, D. E. (eds.) *Parallel distributed processing, explorations in the microstructure of cognition, psychological and biological models*. Cambridge/ Boston/ London.: MIT Press.
- RUMER, R. (1993). *Genie: a scientific tragedy*. New York: Harper Collins.
- SACHS, J.; BARD, B. & JOHNSON, M. (1981). Language learning with restricted input: case studies of two hearing children of deaf parents. *Applied Psycholinguistics*, 2 (1), p. 33-54.
- SACKS, O. (2002). *Vendo vozes*. São Paulo: Companhia das Letras.
- SAUSSURE, F. (1972). *Curso de Lingüística Geral*. São Paulo: Editora Cultrix.
- SCHMIDT, R. (1990). The role of conciousness in second language learning. *Applied Linguistics*, 11, p. 17-46.
- SCHWARTZ, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15, p. 147-63.
- SHELDON, A. (1985). The relationship between production and perception of the /r/-/l/ contrast in Korean Adults learning English. A reply to Borden, Gerber and Milsark. *Language Learning*, 35 (1), p. 107-113.
- SHLRC. (1994). *ToBI intonation analysis samples*. Sydney, Australia: Macquarie University.

- SILVERMAN, K.; BECKMAN, M.; PITRELLI, J.; OSTENDORF, M.; WIGHTMAN, C.; PRICE, P.; PIERREHUMBERT, J. & HIRSCHBERG, J. (1992). ToBI: a standard for labelling English prosody. *Proceedings of the 1992 International Conference on Spoken Language Processing*, p. 867 – 870.
- SIMÕES, A. R. M. (1996). Assessing the contribution of instructional technology in the teaching of pronunciation. *Proceedings of the 1st ESCA Tutorial Research Workshop on Speech-Production Modeling- 4th Speech Production Seminar*.
- SKINNER, B. (1957). *Verbal behavior*. New York.: Appleton, Century, Crofts.
- SLOBIN, D. J. (1980). O desenvolvimento da linguagem na criança. In: *Psicolingüística*. São Paulo: Companhia da Editora Nacional.
- SUTER, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26, p. 233-253.
- THOMPSON, I. (1991). Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning*, 41, p. 177-204.
- TOWNSHEND, B.; BERNESTEIN, J.; TODIC, O. & WARREN, E. (1998). Estimation of spoken language proficiency. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Sweeden: Marholmen Conference Center.
- TRAGER, G. L. & SMITH, H. L. (1951): An outline of English structure. *Studies in Linguistics, Occasional papers 3*.
- TRUBETZKOY, N. S. (1949). Grundzüge der Phonologie, 1939. Tradução de J. Cantineau. Principes de Phonologie. *Travaux du Cercle Linguistique de Prague*, 7; Paris: Klincksieck.
- VAISSIÈRE, J. (1983). Language - independent prosodic features. In: Cutler, A. & Ladd, R. (eds). *Prosody: models and measurement*. Berlin: Springer, p. 53-66.
- VICSI, K.; ROACH, P.; ÖSTER, A; KACIC, Z. & BARCZIKAY, P. (1999). SPECO- a multimedia multilingual teaching and training system for speech handicapped children. *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 2, p.1361-1364.
- VYGOTSKY, L. (1934). *Thought and language*. Cambridge: MIT Press.
- WERKER, J. F & LOGAN, J. S. (1985). Cross-language evidence for three factors in speech perception. *Child Development*, 52, p. 349-355.

WHITE.L.(1989). *Universal grammar and second language acquisition*. Amsterdam, Philadelphia, Pa.: John Benjamins.

WILLEMS, N. (1983). *English intonation from a dutch point of view*. Doctoral Dissertation, University of Utrecht.

ZIOLKOWSKI, M. & LANDAHI, K. (1995). How successfully does visual feedback train listeners to produce and perceive non-native phonological contrasts?. *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Sweeden, (3), p. 330-334.