Unesp UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Faculdade de Ciências e Letras
Campus de Araraquara – SP

MARIANE CARVALHO

A VARIAÇÃO MELÓDICA DA FALA: um estudo da correlação entre uma análise acústica e uma análise auditiva da entoação

ARARAQUARA – S.P.
2013

MARIANE CARVALHO

A VARIAÇÃO MELÓDICA DA FALA: um estudo da correlação entre uma análise acústica e uma análise auditiva da entoação

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – Unesp/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise fonológica, Morfossintática, Semântica e Pragmática

Orientador: Prof. Dr. Luiz Carlos Cagliari

Bolsa: CNPq

ARARAQUARA – S.P.
2013

MARIANE CARVALHO

A VARIAÇÃO MELÓDICA DA FALA: um estudo da correlação entre uma análise acústica e uma análise auditiva da entoação

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – Unesp/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise fonológica, Morfossintática, Semântica e Pragmática

Orientador: Prof. Dr. Luiz Carlos Cagliari

Bolsa: CNPq

Data da defesa: 27/02/2013

MEMBROS COMPONENTES DA BANCA EXAMINADORA:

Presidente e Orientador: Prof. Dr. Luiz Carlos Cagliari
Universidade Estadual Paulista (UNESP - Campus de Araraquara)

Membro Titular: **Profa. Dra. Larissa Cristina Berti**
Universidade Estadual Paulista (UNESP – Campus de São José do Rio Preto)

Membro Titular: **Prof. Dr. Daniel Soares da Costa**
Universidade Estadual Paulista (UNESP- Campus de Araraquara)

Local: Universidade Estadual Paulista
Faculdade de Ciências e Letras
UNESP- Campus de Araraquara

Aos meus pais, à minha irmã e ao Bruno, as três maiores razões da minha vida.

AGRADECIMENTOS

Durante esses dois anos, muitas foram as pessoas que fizeram parte, direta ou indiretamente, de mais essa conquista. Por isso, é com muito carinho e gratidão que venho agradecê-las.

Primeiramente, agradeço a Deus, por estar sempre presente em minha vida, por confortar e iluminar os meus dias e, principalmente, por colocar em meu caminho pessoas tão especiais.

Agradeço ao meu orientador, Luiz Carlos Cagliari, pela confiança depositada em mim, pelas oportunidades concedidas, pelo incentivo e, principalmente, pela ajuda e dedicação durante a elaboração desta dissertação.

Também sou grata à minha família, pelo apoio incondicional e pelo amparo emocional, fundamentais para que eu chegasse até aqui.

Sou grata ainda ao carinho, às dúvidas compartilhadas e, sobretudo aos momentos de distração divididos com minhas amigas, Patrícia, Raisa, Islene, Pricila, Rafaela e Audineia.

Meu muitíssimo obrigada à minha cunhada Cíntia Belli Vischi, com quem pude compartilhar muitos momentos, uns muito alegres e outros nem tanto. Agradeço também à Eloísa, minha sogra querida, pelas palavras de incentivo e pelas infinitas orações.

Agradeço às minhas avós, que em sua simplicidade sabem dizer tudo aquilo que preciso ouvir e entender. Ao meu avô, agradeço pelo carinho e pela preocupação. O senhor faz muita falta.

À minha irmã, pelas conversas e risadas e, a quem muito pude e ainda posso ensinar.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento do meu trabalho durante o período de dois anos.

Também devo agradecer ao Programa de Pós Graduação em Linguística e Língua Portuguesa da Unesp pelo financiamento dos eventos científicos dos quais participei. Meus agradecimentos sinceros também vão aos funcionários do programa, sempre muito pacientes, e à coordenadora do programa Profa. Dra. Rosane de Andrade Berlinck, sempre muito atenciosa e dedicada aos alunos.

Por fim, e nem por isso menos importante, agradeço ao Bruno, pelo companheirismo, pela compreenssão e pelo conforto de suas palavras. Obrigada por acreditar em mim.

Todo lo que puedas imaginar es real
Pablo Picasso

RESUMO

O objetivo desta dissertação de mestrado é estudar alguns aspectos da correlação entre dados acústicos e dados obtidos através de modelos de análise auditiva da entoação, a fim de verificar se a análise acústica se distancia ou não do reconhecimento auditivo. Para isso, investigamos, através de alguns dos principais métodos de descrição da entoação, do ponto de vista perceptivo (fonologia funcional) e do ponto de vista acústico (modelo autossegmental), qual a relação entre a percepção e a realidade física do som na análise do fenômeno da entoação. Nossa intenção é provar que essa correlação é tanto possível quanto viável. Para alcançar esse resultado, fizemos, primeiramente, uma descrição individual das duas abordagens eleitas para essa comparação. Uma das abordagens diz respeito ao estudo auditivo desenvolvido por Halliday (1970) e adaptado por Cagliari (1982/2007) para a análise dos dados do português brasileiro. A outra é de base acústica e segue o modelo de análise autossegmental, desenvolvido por Pierrehumbert (1980). Feito isso, propomos uma comparação entre os dois modelos, trazendo à baila suas semelhanças e diferenças, com o intuito de chegarmos a um possível paralelo entre eles. O resultado foi satisfatório e mostrou que, apesar de formalmente diferentes, é possível constatar semelhanças entre eles. Partindo desse princípio, fizemos uma detalhada análise acústica por meio do programa computacional PRAAT e, auditiva, (seguindo a metodologia de Halliday e Cagliari). A análise descreve a estrutura fonológica da entoação de um pequeno trecho do livro *História sem fim* (sd). A gravação foi feita por um informante adulto da cidade de Araraquara. O resultado mostrou, mais uma vez, uma correspondência entre as formas fonológicas, e provou a aproximação das análises acústica e auditiva. Além disso, também foi nosso objetivo mostrar que a correlação entre as duas análises (acústica e auditiva) permite uma melhor explicação da variação melódica da língua, bem como sua correlação com outros níveis da gramática (sintaxe, semântica e pragmática). Com relação a isso, concluímos que a entoação só é descrita corretamente quando os fatores sintáticos e semânticos são levados em consideração, pois os padrões entoacionais revelam as intuições dos falantes e ouvintes, sendo parte do componente fonológico da gramática da língua.

Palavras-chave: Análise acústica. Análise auditiva. Fonologia entoacional. Halliday, M.A.K. Pierrehumbert, Janet.

ABSTRACT

The aim of this dissertation is the analysis of the correlation between acoustic and auditory data, in order to verify whether the acoustic analysis follows or not the auditory recognition. For this, we investigated, with some of main methods of description of intonation, the perception (functional phonology) and the acoustics (autossegmental model), what is the relation between perception and physical reality in the analysis of intonation. Our intention is to prove that this correlation is as much possible as feasible. We started with the description of the two approaches chosen for our comparison. One of them is concerned with the auditory approach. This theory was developed by Halliday (1970), according to Cagliari (1982/2007), who adapted Halliday's model to describe the intonation of Brazilian Portuguese. The other, is the acoustic approach developed by Pierrehumbert (1980) autossegmental approach. We propose a comparison between the two theories, bringing up their similarities and differences in order to reach a possible comparative parallel between them. The results showed that, despite the formal differences, similarities between the two approaches. Based on this idea, we made a detailed acoustic analysis of our data with the software PRAAT and an auditory analysis of the same data (following the Halliday's and Cagliari's methodology). The analyses describe the phonological structure of intonation of a short passage from the book *História sem fim* (*The Neverending Story*) (sd). The recording was performed by an adult informant from the city of Araraquara. Once more, the results showed a good correspondence between the phonological forms merged from the acoustic and auditory analyses. In addition, our goal was to show that the correlation between the two approachs (acoustic and auditory) allows a better explanation of language melody variation, as well as it relation with other levels of the grammar (syntax, semantics, and pragmatics). In this regard, we conclude that the intonational phenomenon is better described only when the syntactic and semantic factors are considered, because intonational patterns reveal the intuitions of speakers and listeners, being part of the phonological component of the grammar of the language.

Keywords: Acoustic analysis. Auditory analysis. Intonational phonology. Halliday, M.A.K. Pierrehumbert, Janet.

LISTA DE FIGURAS

Figura 1 Exemplo de notação segundo o modelo de Daniel Jones (1972) 31

Figura 2 Representação do tom 1 normal 32

Figura 3 Descrição do tom 2 normal em enunciados interrogativos.....33

Figura 4 Representação: contorno entoacional de nível 3 na frase:
“If Tom’ll do it I will”.....37

Figura 5 Os três exemplos de pausas propostos pelo modelo notacional de Pike 38

Figura 6 Modelo de transcrição no enunciado “If he can, he’ll cooperate.....38

Figura 7 Modelo de transcrição a partir de linhas curvas e tamanhos de bolinhas.....39

Figura 8 Sistema de transcrição desenvolvido para o aprendizado de língua
inglesa 40

Figura 9 Modelo de análise autossegmental para a língua Bantu 40

Figura 10 Exemplo de representação da formação básica de um ennunciado
fonológico (U), sintagma entoacional (I) e regras de proeminência (W/S) 44

Figura 11 Caracterização dos contornos melódicos para o Holandês.....49

Figura 12 Caracterização e especificação da configuração dos Roots e
representação do contorno entoacional da F0 50

Figura 13 Configuração dos prefixos e sufixos, com os respectivos tipos de
contornos melódicos e configuração da F0 51

Figura 14 O mecanismo respiratório.....58

Figura 15 Representação da glote e das cordas vocais, vista através da abertura da
laringe.....59

Figura 16 O aparelho auditivo e suas estruturas 63

Figura 17 Pauta entoacional com os cinco níveis melódicos.....106

Figura 18 Ilustração dos contornos melódicos por meio de linhas verticais e barras
paralelas.....106

Figura 19 Exemplo de análise a partir do modelo de Halliday (1970) e
Cagliari (2007), com caracterização dos símbolos convelcionalmente usados 108

Figura 20 Representação gráfica do padrão entoacional dos tons 109

Figura 21 Representação dos tons primários simples para o inglês britânico.....110

Figura 22 Representação dos tons primários compostos 13 e 53.....110

Figura 23 Tons secundários. Representação pictórica da tônica.....111

Figura 24 Tons secundários. Representação pictórica da pretônica 112

Figura 25 Sistema de tons primários simples e compostos do português brasileiro ...113

Figura 26 Sistema de tons secundários para o português brasileiro.....114

Figura 27 Exemplo de tom primário simples em frases assertivas 115

Figura 28 Exemplo de tom secundário em frases de valor enfático.....115

Figura 29 Exemplo de tom de fronteira inicial 120

Figura 30 Esquema do modelo teórico proposto por Pierrehumbert para o repertório do inglês americano.....121

Figura 31 Representação do *pitch accent*, do acento frasal e do tom de fronteira.....122

Figura 32 Contorno entoacional do tipo *downstep* no enunciado: “There are many intermediate levels 126

Figura 33 Esquema básico para representação do *downstep* 126

Figura 34 Padrão entoacional do Inglês do tipo *terraced downsteping*.....127

Figura 35 Exemplo de um caso de *upstep* no enunciado:
“It’s really tôo good to be true”.....128

Figura 36 Exemplo de como calcular graficamente a linha de base 129

Figura 37 Exemplo de formação de *Plateau*.....130

Figura 38 Evidências de interpretação auditiva no modelo de Pierrehumbert para o enunciado: “It’s indexed by the keywords in the abstract”.....131

Figura 39 Frases entoacionais de Fronteira anotadas com pausas 139

Figura 40 Correspondência entre Pierrehumbert (1980) e o estilo britânico dos tons nucleares.....142

Figura 41 O modelo de Perrehumbert descrito na metodologia de Halliday por Cagliari.....143

Figura 42 O modelo de Halliday descrito na metodologia de Pierrehumbert por Cagliari.....143

Figura 43 Conversão dos tons primários simples e compostos do modelo de Halliday para a forma fonológica de Pierrehumbert.....144

Figura 44 Representação do método de análise a partir do ponto medial das sílabas vozeadas do enunciado 12 – “pequeno estrangeiro” 149

Figura 45 Análise da frequência fundamental no ponto medial das sílabas desvozeadas, a partir da função: *pitch*> *get pitch*, no enunciado 12 – “pequeno estrangeiro” 149

Figura 46 Exemplo de um problema de filtragem gerado pelo PRAAT do enunciado 3 – “Por que está preso aqui? 152

Figura 47 Exemplo de variação entre a análise acústica e a auditiva no segundo grupotonal do enunciado “Sou Gmork, o lobisomen” 168

Figura 48 Representação de um padrão melódico que pode interferir de modo mais significativo na correlação entre uma análise acústica e auditiva 169

Figura 49 Representação de um padrão melódico que não interfere na correlção entre uma análise acústica e auditiva 169

Figura 50 Padrões quedificultam a representação do *target ideal* 170

Figura 51 Exemplo de sentença declarativa segundo os padrões de Halliday (1970)/ Cagliari (1982/2007) e Pierrehumbert no enunciado: “Aqueles que me prenderam com esta corrente 178

Figura 52 Representação acústica do esquema de contorno tônico do padrão entoacional H- + L* L- L% do enunciado “The Uruguyan bulldozer drivers’ union” 179

Figura 53 Exemplo da análise do tom 3 na forma fonológica de Pierrehumbert no enunciado“depois de um longo silêncio,” 181

Figura 54 Representação acústica do esquema de contorno tônico do padrão entoacionalL-+H*H-H% do enunciado “The Uruguayan bulldozer drivers’ union” 182

Figura 55 Representação acústica do esquema de contorno tônico do padrão entoacionalL-+H*L% do enunciado “The Uruguayan bulldozer drivers’ union” 184

Figura 56 Representação do esquema de contorno tônico para o Tom 1-, no enunciado “The Uruguayan bulldozer driver’s union” 186

Figura 57 Exemplo da curva melodic para o enunciado E7 – “E para onde foram?” Com caracterização do *pitch accent*, destacado em vermelho, do acento frasal e do tom de fronteira 189

Figura 58 Contorno da F0 para o enunciado “The Uruguayan bulldozer drivers’ union” 190

Figura 59 Contorno da F0 do enunciado E11 “Você não é daqui” 190

Figura 60 Esquema de representação, em sequência, de todos os tons encontrados no fragmento de texto analisado 191

LISTA DE TABELAS

Tabela 1 Descrição das sílabas que apresentaram maior e menor F0, com sua respectiva soma e média aritmética 156

Tabela 2 Distribuição da variação melódica das sílabas em Hz de acordo com a classificção auditiva 161

Tabela 3 Quantidade e porcentagem de ocorrência dos tons 162

Tabela 4 Intervalo de frequência que caracteriza os tons, dada a tessitura desse informante e a interpretação auditiva dos tons 163

Tabela 5 Número de ocorrência dos tons em função das faixas de frequência 165

Tabela 6 Intervalo da frequência fundamental em função dos cinco níveis tonais e a diferença, em cps, para cada tom 166

Tabela 7 Tons que variaram em relação aos intervalos da F0 estabelecidos e sua respectiva diferença, dada a partir dos valores obtidos na tabela anterior 166

Tabela 8 Especificação dos tipos de tons e número de ocorrência de cada um deles 171

LISTA DE GRÁFICOS

Gráfico 1 Representação gráfica da tessitura, alcançada a partir da média da distribuição fundamental nas sílabas dos grupos tonais.....157

Gráfico 2 Representação gráfica, em porcentagem, dos cinco níveis tonais 162

Gráfico 3 Número de ocorrência de cada tom e sua respectiva variação.....164

LISTA DE QUADROS

Quadro 1 Interpretação dos tons que ocorreram no fragmento do texto *História sem fim* 172

Quadro 2 Conversão do tom 1 primário para a forma fonológica de Pierrehumbert 173

Quadro 3 Conversão do tom 3 primário para a forma fonológica de Pierrehumbert 174

Quadro 4 Os tons secundários de Halliday analisados no modelo fonológico de Pierrehumbert 174

Quadro 5 Correspondência entre as formas fonológicas acústica e auditiva para os modelos teóricos que as ilustram, bem como os resultados obtidos para cada um deles nesta dissertação 192

Quadro 6 Resumo dos significados produzidos pelos diferentes padrões entoacionais do texto analisado 212

LISTA DE ABREVIAÇÕES

A	tom alto segundo o modelo de análise funcional
(A)	variação melódica simples de contorno ascendente
AM	teoria métrica-autossegmental
B	tom baixo
(bb)	contorno melódico brusco
(c)	contorno melódico contínuo
CMN	componente melódico notável
CPT	componente pré-tônico
CT	componente tônico
D	variação melódica simples de contorno descendente
(e)	contorno melódico por etapas
F0	frequência fundamental
GT	grupo tonal
H	tom alto (high) segundo o modelo autossegmental
H*/ L*	sílaba tônica frasal (foco)
H^-/L^-	sílaba tônica
L	tom baixo (low)
L%/H%	tom de fronteira (final ou inicial)
MB	tom meio baixo
M	tom médio
MA	tom meio alto
N	variação melódica simples de contorno nivelado
PB	português brasileiro
P	configuração do tipo prefixo
R	configuração do tipo raiz
RP	Inglês Britânico – Variedade “Received Pronunciation” do Inglês Britânico
(s)	contorno melódico por saltos

(ss)	contorno melódico suave
S	configuração do tipo sufixo
STS	sílaba tônica saliente
ToBI	*Tones and Break Indices* (nome do programa)
ToBiPI	*Transcription* of *Brazilian Intonation*

LISTA DE SÍMBOLOS

U	enunciado fonológico
I	sintagma entoacional
φ	sintagma fonológico
C	grupo clítico
ω	palavra fonológica
∑	pé
σ	sílaba
^	sílabas tônicas silenciosas
[/]	pausa virtual (*tentative pauses*)
[//]	pausas finais
/	divisão das sílabas em pés rítmicos
//	separação dos GTs
═	indicação da STS
‖	limite estabelecido entre a sílaba átona e tônica
[^{o}]	sílaba tônica saliente
[-]	intervalo entre um nível e outro
	foco do enunciado

SUMÁRIO

1 INTRODUÇÃO 21

2 OBJETIVOS 26

2.1 Objetivos gerais 26
2.2 Objetivos específicos 27

3 FUNDAMENTAÇÃO TEÓRICA 28

3.1 Considerações iniciais 29
3.1.1 Contexto histórico 29
3.1.1.1 Os predecessores: modelos teóricos 31
3.1.1.2 O modelo de *Daniel Jones* 31
3.1.2 O modelo de *Kenneth Lee Pike* 34
3.1.2.1 O modelo autossegmental: a visão inicial de *John Goldsmith* 40
3.1.2.2 Fonologia prosódica: o modelo de *Nespor e Vogel (2007)* 41
3.1.3 O modelo de *t'Hart, Colier e Cohen (1990)* 45
3.1.3.1 O modelo de *Robert Ladd* 52
3.1.3.2 A entoação no Português Brasileiro 55

3.2 OS MECANISMOS DE PRODUÇÃO E PERCEPÇÃO DA FALA 57

3.2.1 Considerações iniciais 57
3.2.2 O ar e a produção da fala 57
3.2.3 A produção e a anatomia da audição humana 62
3.2.4 A percepção da fala e as teorias envolvidas 64
3.3 A questão acústica e auditiva para a categorização dos aspectos da fala 70
3.3.1 Considerações iniciais 70
3.3.2 O acústico e o auditivo 70
3.3.3 As discussões 72
3.3.4 Alguns pontos de vista 77

4 OS ELEMENTOS PROSÓDICOS 82

4.1 A tessitura 82
4.2 Propriedades da dinâmica da fala 85
4.3 A qualidade de voz 86
4.4 O ritmo e a continuidade 87
4.5 A entoação 88
4.5.1 As funções da entoação 88
Função semântica e sintática 89
A função fonológica 94
4.6 As expressões de atitude 95

5 EMBASAMENTO TEÓRICO 99

5.1 Descrição dos trabalhos teóricos 100
5.1.2 O modelo descritivo de Halliday 100
5.1.3 O modelo de Cagliari (1982/2007): a adaptação para o PB 112

5.2 O modelo descritivo de Janet Pierrehumbert (1980) 116
5.2.1 A versão original 117
5.2.2 Modificação do modelo original 133
5.2.3 O sistema ToBI (Tones a Break Indices) 133
5.3 O modelo de Halliday e Pierrehumbert: uma comparação 134
5.3.1 Semelhanças e diferenças 134

6 MATERIAIS E MÉTODOS 145

6.1 Procedimentos: considerações iniciais 146
6.1.1 Dados da leitura 146
6.1.2 Dados do informante 147
6.1.3 Dados da análise acústica 147
6.2 Tessitura: procedimentos de análise 150
6.3 O padrão entoacional: procedimentos de análise 152

7 ANÁLISE DOS DADOS E RESULTADOS 154

7.1 Análise da tessitura 155
7.1.2 A tessitura no nível tonal a partir do modelo de Halliday 159
7.1.3 As divergências nos dados 164
7.2 Análise do padrão entoacional dos enunciados 170
7.2.1 Tipos de tons 170
7.2.1.2 Conversão da forma fonológica para os tons primários e secundários 173
7.2.2 Caracterização dos padrões fonológicos para o PB a partir do modelo de Halliday e Pierrehumbert 175
7.2.2.1 Análise: tons primários 176
Tom 1 180
Tom 3 178
7.2.2.2 Análise: tons secundários 183
Tom 1+ 183
Tom +1 185
Tom 1- 186
Tom -1 187
Tom :1 188
Tom 3+ 189
7.3 Análise: o significado dos enunciados de acordo com os diferentes padrões entoacionais 195
7.3.1 Observações gerais 195
7.3.1.1 Apontamentos metodológicos 196
7.3.1.2 Análise dos enunciados 198

8 CONCLUSÃO 214

REFERÊNCIAS 218

ANEXOS 227

ANEXO A – Fragmento do texto História sem fim 227

ANEXO B – Transcrição ortográfica e fonética do texto analisado 228

1 Introdução

Atualmente, há uma grande dificuldade quando se trata dos estudos auditivos, já que muitos foneticistas não admitem a possibilidade de se estudar fonética do ponto de vista auditivo, acreditando apenas em dados estatísticos. Durante a graduação, ao correlacionar dados acústicos e dados auditivos, obtidos através da gravação de um pequeno fragmento do livro *O pequeno papa sonhos* de Michael ENDE e FUCHSHUBER (1998, p.19), demos início a um estudo que tinha por objetivo desmistificar essa ideia. Isto é, mostrar que é possível fazer análises auditivas concomitantemente a análises acústicas. Nesse período, trabalhamos apenas com o modelo descritivo de M.A.K. Halliday (1970) e com a adaptação dele feita por Cagliari (2007) para o Português do Brasil (PB). A pesquisa desenvolvida, mesmo sendo mais limitada, obteve resultados satisfatórios e mostrou a viabilidade das análises auditivas. A partir disso, demos continuidade ao estudo desse tema e estendemos nossa pesquisa para este projeto de mestrado, acrescentando o modelo teórico autossegmental proposto por Pierrehumbert (1980).

De acordo com Cagliari (2012, p. 3), "Os trabalhos mais antigos sobre a entoação foram feitos através de análises auditivas". Assim, nos perguntamos como é possível compreender a acústica da fala sem correlacioná-la com a percepção? A percepção da fala é o que motiva o uso da linguagem entre as pessoas, é "[...] fundamental para o falante e para o ouvinte, e serve como controlador daquilo que se fala e do que se ouve." (CAGLIARI, 2012, p. 3), de modo que o ouvinte é capaz de perceber variações no tom de voz do falante, identificando padrões prosódicos do tipo ofensivo, irônico, de decepção, que expressem alegria, entre outros. Tal discernimento mostra o valor dos estudos auditivos e perceptivos nos estudos linguísticos.

Para Cagliari (2012, p. 3),

> [...] a observação direta e a introspecção do linguista são métodos importantes e essenciais para a descrição das línguas, em todos os seus níveis, incluindo os fenômenos fonéticos. Portanto, a fonética pode prescindir das análises experimentais e instrumentais que se fazem nos laboratórios, mas tais análises físicas não podem nunca prescindir das interpretações perceptuais auditivas, se o resultado desejado é uma interpretação linguística do fenômeno.

Decorre disso o nosso interesse em abordar esse aspecto tão criticado por alguns foneticistas que não admitem a possibilidade de se estudar fonética do ponto de vista auditivo, acreditando apenas em dados estatísticos. Por isso, optamos por fazer um estudo comparativo entre as análises acústica e auditiva, a fim de mostrar que é posível e também

viável desenvolver análises desse tipo e, ainda, mostrar a proximidade entre os dois fenômenos.

Nas transcrições auditivas, as análises são menos detalhadas, baseando-se em padrões interpretativos mais gerais. Em compensação, temos a precisão dos dados acústicos. Além disso, também devemos levar em consideração a natureza física e fisiológica do ouvido que funciona como um filtro. Assim, é da comparação entre ambos que podemos estabelecer as faixas de frequência que são importantes, pois pequenas variações da frequência fundamental (F0) não devem ser consideradas como relevantes para a análise linguística, uma vez que

> [...] Estabelecer esses limites é algo que não se pode fazer apenas estatisticamente, porque há outros fatores que interferem na avaliação da F0 como a tessitura. Tratamentos estatísticos, em geral, evidenciam apenas grandes contrastes ou se aplicam em condições de fala muito controladas. (CAGLIARI, 2010, p. 3).

A presente dissertação também mostra um interesse especial em entender como a entoação tem sido estudada tradicionalmente. As pesquisas mais antigas (PIKE, 1945) descreviam a entoação através de uma análise auditiva, usando um modelo estruturalista em que o fenômeno da variação melódica da fala era descrito por meio de padrões entoacionais em contraste uns com os outros. Em trabalhos antigos, era comum vir, junto com a descrição entoacional, uma descrição do ritmo da fala (PIKE, 1945; HALLIDAY, 1970). Somente na segunda metade do século passado começaram a aparecer estudos acústicos, substituindo os estudos auditivos. Logo, eles se mostraram independentes dos estudos auditivos anteriores pela dificuldade de sistematização dos dados acústicos em padrões bem definidos do tipo que aparecia nas análises auditivas. A solução foi investir em análises estatísticas para a detecção de padrões recorrentes. Com isso, criou-se um abismo entre os modelos de análise acústica e os modelos de análise auditiva. Fato que não deveria ocorrer, uma vez que os dados da variação melódica são os mesmos, podendo ser analisados e interpretados seja através de um modelo descritivo de base perceptual, seja através de equipamentos de análise física dos sons. Tendo em vista essa situação, a melhor solução não está em tratar as duas análises individualmente, mas sim em correlacioná-las, de modo que uma ajudaria a esclarecer fatos que a outra não conseguiria tratar adequadamente.

Como estamos interessados em desenvolver um trabalho que correlacione uma análise auditiva a uma análise acústica, ou vice versa, é essencial apresentarmos uma seleção dos trabalhos, teses, livros e artigos em periódicos dos autores mais representativos dessas

áreas. Nossa intenção não é fazer um apanhado teórico das metodologias empregadas por eles, mas apresentá-los, de modo a separar os estudos acústicos dos estudos auditivos.

A lista apresentada segue as anotações feitas por Cagliari (2011, p. 2). Assim, em relação aos estudos acústicos temos os trabalhos de: LIBERMAN (1975); GOLDSMITH (1980); PIERREHUMBERT (1980); LADD (1980, 1996); CUTLER e LADD (1983); FUJISAKI (1983); t'HART, COLLIER e COHEN (1990); BECKMAN e AYRES (1994); GRABE (1998); HIRST e DI CRISTO (1998); GRICE, LADD e ARVANITI (2000); GUSSENHOVEN (2002), JUN (2003), entre outros. Nesses trabalhos, a variação melódica é tratada como um fenômeno fonológico, a descrição feita é experimental acústica. Para Cagliari (2011, p. 2-3), a noção de fonologia adotada nesses trabalhos nem sempre é muito clara, ficando, na maioria das vezes, somente em considerações de natureza fonética.

Os trabalhos de natureza auditiva, se comparados aos de natureza acústica, são poucos, sendo eles, conforme Cagliari (2011): os estudos de HALLIDAY (1970); BOLINGER (1986); NESPOR e VOGEL (1986); COUPER-KHULEN (1986); CRUTTENDEN (1997).

Dentre os vários modelos apresentados, que foram propostos e usados tanto em análise auditiva quanto em análise acústica, é difícil avaliar todas as teorias e resultados que cada um deles produz. Como sabemos, as visões e as perspectivas frente ao objeto de estudo são distintas, cada uma com suas peculiaridades. O trabalho conjunto, como revela Cagliari (2012), permite uma melhor compreensão e observação do fenômeno, mesmo diante de algumas aparentes contradições, pois um completa as falhas do outro.

Diante disso, este trabalho irá privilegiar, para a análise auditiva, o modelo teórico proposto por M. A. K. Halliday (1970) e adaptado para o Português do Brasil por Cagliari (1982/2007). Para a análise acústica da entoação, o modelo teórico utilizado será o de Pierrehumbert (1980). Em linhas gerais, o primeiro modelo pertence à gramática funcional e faz um estudo sobre o inglês americano. O segundo modelo pertence à gramática gerativa e faz uma análise do inglês americano. A escolha pelos dois tipos de abordagem deve-se ao fato de eles serem, atualmente, os modelos mais desenvolvidos no que diz respeito à descrição da entoação dentro do componente fonológico da gramática funcional e gerativa.

Nesta dissertação também levaremos em consideração a questão da variação melódica da fala dentro de um modelo descritivo linguístico, ou seja, incorporado à gramática da língua, pois todo padrão entoacional ou contorno melódico relaciona-se a algum significado semântico ou pragmático, a alguma estrutura sintática e, especialmente, aos atos de fala e às atitudes do falante.

Portanto, com esta pesquisa, nosso interesse é mostrar que, quando uma abordagem auditiva vem auxiliada pela análise acústica, ou o contrário, a descrição do fenômeno ganha uma maior qualidade. Além disso, a análise acústica acoplada à análise auditiva ajuda a conciliar os detalhes da fala com as unidades fonológicas do sistema da língua.

Sendo assim, esta dissertação está organizada do modo descrito a seguir.

Na primeira parte, apresentamos o objetivo principal da pesquisa, assim como os objetivos específicos, essenciais para a validação dessa teoria.

A segunda parte desta dissertação é destinada à "Revisão de literatura" e subdivide-se em cinco partes. Na primeira parte, para se entender um pouco do que se faz hoje em termos de análise linguística da entoação, traçamos, primeiramente, um breve percurso histórico, a fim de contextualizarmos o fenômeno. Feito isso, fazemos uma revisão cronológica dos trabalhos entoacionais que tiveram uma maior importância nos estudos entoacionais ao longo dos anos. Começamos essa revisão com Daniel Jones (1917), mostrando, por meio de um sistema de pontos, curvas e linhas paralelas, a descrição da variação melódica da língua proposta por ele. Depois, seguimos com o modelo de Pike (1945) e John Goldsmith (1976). O primeiro reconhece a entoação como uma variação melódica e propõe, para sua análise, um sistema de quatro níveis de contorno entoacional. O segundo foneticista é quem dá início ao modelo autossegmental utilizado neste trabalho que, por sua vez, foi empregado para a descrição das línguas tonais.

Em seguida, chegamos a 1986 a partir da teoria de Nespor e Vogel e apresentamos o modelo chamado de Fonologia Prosódica. Nesse modelo, segundo Scarpa e Fernandes-Svartman (2012), o fluxo de fala é organizado hierarquicamente através dos domínios prosódicos. Seguindo essa teoria, as autoras afirmam que a hierarquia prosódica é obtida por meio de regras fonológicas aplicadas no interior e entre as junturas de determinados domínios e que também podem ser bloqueadas no interior e entre junturas de outros domínios prosódicos. A estrutura prosódica, dentro dessa perspectiva é parcialmente determinada pela estrutura sintática, de modo que elas podem coincidir ou divergir, dependendo do caso.

Posteriormente, caracterizamos a abordagem de t'Hart, Collier e Cohen (1990). De uma maneira geral, os autores compreendem a entoação, a partir de variações da frequência fundamental (F0), mostrando, através de testes perceptivos, o que é e o que não é linguisticamente relevante (PERES; CONSONI; NETTO, 2011). Por fim, descrevemos sucintamente um dos trabalhos de Robert Ladd (1996), trazendo alguns de seus apontamentos sobre a concepção e aplicação da Teoria Autossegmental. Por último,

indicamos alguns dos trabalhos que têm sido desenvolvidos para a entoação do Português Brasileiro (PB).

A segunda parte da revisão diz respeito aos mecanismos envolvidos no processo de percepção das variações melódicas da fala. No item "O ar e a produção da fala" procuramos mostrar os vários órgãos envolvidos no processo de produção dos sons da fala e o modo como eles transformam-se em linguagem. Com isso, fazemos uma reflexão sobre a importância da audição no processo de comunicação e, em decorrência, apresentamos, embasados em modelos teóricos específicos, um resumo das principais teorias envolvidas no processo de reconhecimento da fala. Mesmo que apresentado de modo limitado, este momento do texto é necessário, pois nosso trabalho, entre outras coisas, tem a intenção de mostrar o quanto a audição é significativa para os estudos linguísticos e como o processo auditivo vem confirmar a relevância dessa afirmação.

O tópico "A questão acústica e auditiva para a categorização dos aspectos da fala" constitui a terceira parte da revisão de literatura. Nesse item, descrevemos o modo como a análise acústica e auditiva passou a ser reconhecida por diversos teóricos ao longo dos anos, seja de modo negativo ou positivo. A partir disso, nossa pretensão foi estabelecer um paralelo entre as duas metodologias (acústica e auditiva), trazendo justificativas teóricas que comprovassem e/ou dessem indícios de que uma análise linguística tem um melhor desempenho quando une o acústico e o auditivo, de modo que um seja o reflexo do outro.

O item referente à "Tessitura", por sua vez, caracteriza a quarta parte dessa sessão e trata de algumas questões teóricas relacionadas a esse fenômeno da entoação. Por fim, na quinta parte "Análise semântica e sintática dos enunciados", fazemos uma introdução sobre as funções da entoação, bem como uma caracterização sobre as expressões de atitude na fala.

A terceira parte desta dissertação destina-se a responder as questões levantadas durante este estudo. Este momento da dissertação divide-se em duas etapas. A primeira diz respeito ao quadro metodológico usado para embasar este trabalho e subdivide-se em três partes. Primeiro, apresentamos o modelo descritivo desenvolvido por Halliday (1970) e Cagliari (1982/2007). Em seguida, damos continuidade e trazemos para a discussão o modelo autossegmental proposto na tese de doutorado de Pierrehumbert (1980). O primeiro modelo junta a entoação a outros níveis gramaticais. É um modelo de análise funcional que leva em consideração os aspectos fonéticos e fonológicos, fazendo uma análise auditiva da entoação. O modelo autossegmental, por sua vez, favorece a organização hierárquica dos constituintes prosódicos. É um sistema que também conjuga os fatos fonológicos e

fonéticos, mas que se preocupa muito mais com a questão acústica da entoação. Por fim, após trabalhar individualmente com os dois modelos entoacionais, no item "O modelo funcional e autossegmental: uma proposta de correlação" foi possível responder a uma das proposições feitas mais acima, isto é, a possibilidade de ligar em uma análise, duas perspectivas teóricas que, supostamente, seriam tão diferentes entre si. Na segunda parte, seguimos com a apresentação dos materiais e procedimentos utilizados durante a pesquisa, expondo os dados acústicos, de leitura, bem como os dados do *corpus* e do informante. Em relação à análise acústica, destacamos a metodologia desenvolvida e proposta nesta dissertação para a análise da variação melódica da entoação. Também apresentamos nesse item os procedimentos usados para a análise da tessitura e dos padrões entoacionais dos enunciados.

A quarta etapa deste trabalho é destinada à análise dos dados, juntamente com uma discussão dos resultados obtidos. Esse item subdivide-se em três partes, cada uma delas com os respectivos gráficos, tabelas e discussão dos dados. As três etapas tomam como ponto de partida os modelos teóricos acústico e auditivo, buscando mostrar a viabilidade das análises auditivas. Na parte que diz respeito à "Tessitura", fizemos a correlação entre o ponto de vista auditivo e acústico, a fim de comprovar a possibilidade desse tipo de estudo. Em "Os padrões entoacionais", fizemos a conversão entre a forma fonológica dos tons obtida no modelo de Halliday/Cagliari para a de Pierrehumbert, a fim de mostrarmos a correspondência entre os dois métodos descritivos. Na terceira parte, referente à Análise semântica e sintática dos enunciados, foi feita a análise dos enunciados, com o objetivo de definir linguisticamente os padrões entoacionais dentro da gramática da língua, juntamente com os padrões acústicos e auditivos.

Finalmente, a sexta parte contempla, de acordo com os resultados e discussões, as conclusões a que chegamos nesta dissertação.

2 Objetivos

A fim de oferecer uma maior praticidade à leitura, apresentamos a seguir os objetivos gerais e específicos deste trabalho.

2.1 Objetivos gerais

O objetivo principal desta dissertação é investigar o ponto de vista perceptivo com o caráter físico acústico do som, a fim de entender se a percepção corresponde à análise acústica ou se ela apresenta características próprias. Para isso, parte-se da análise auditiva

para compará-la, depois, com os dados acústicos. Para alcançar esse objetivo, foi preciso trabalhar, primeiramente, no nível das sílabas e, depois, integrá-las aos padrões entoacionais, sem os quais não seria possível fazer a correlação proposta.

2.2 Objetivos específicos

Para alcançar o propósito geral apresentado, os objetivos específicos seguidos foram:

1) investigar como as questões acústicas e auditivas foram reconhecidas ao longo dos anos por diversos teóricos;

2) investigar os modelos teóricos selecionados, avaliando-os do ponto de vista metodológico e dos resultados que produzem;

3) avaliar as características de cada uma das abordagens (funcional e autossegmental), observando os pontos divergentes e comuns entre elas;

4) investigar, através da variação de tessitura, a validação ou não da análise auditiva, de acordo com os níveis de altura melódica proposto pelo modelo funcional, bem como estabelecer uma correspondência com a variação acústica da frequência fundamental (F0);

5) investigar a possibilidade de correspondência entre as duas abordagens teóricas, fazendo uma correlação entre suas formas fonológicas;

6) investigar a questão da variação melódica da fala dentro de um modelo descritivo linguístico, incorporado à gramática da língua, bem como sua relação com outros níveis da gramática (sintaxe, semântica e pragmática).

3 Fundamentação Teórica

3.1 Considerações iniciais

Para entendermos o que se faz hoje em termos de análise linguística da entoação, apresentamos a seguir um breve percurso histórico de como o fenômeno passou a ser visto e estudado ao longo dos anos, apontando, de um modo mais geral, alguns dos trabalhos mais ilustrativos de cada época. Em seguida, damos continuidade a essa visão cronológica, apresentando, de modo sucinto, algumas das perspectivas teóricas de maior destaque entre os estudos entoacionais.

3.1.1 Contexto histórico

O termo entoação é novo; a palavra mais antiga usada para descrever a variação melódica da fala é "prosódia". Segundo Nooteboom (1999), o termo "prosódia" teve sua origem na antiga Grécia, fazendo referência a cantos e músicas que eram executadas com instrumentos musicais. Couper-Kuhlen (1986) menciona que esse termo também era usado pelos gregos para se referir a características da fala que não podiam ser representadas graficamente, como, por exemplo, o tom.

Depois, esse termo foi usado, no século XVIII, para a arte da versificação, governando a modulação da voz humana na leitura de poesias em voz alta (NOOTEBOOM, 1999, p. 01). Entretanto, essa perspectiva métrica poética não era suficiente, pois só dava conta da duração e do ritmo, não dizendo nada sobre a variação melódica. Todavia, indo além da metrificação poética, esse fato começou também a interessar, a partir daquela época, aos próprios prosodistas, que queriam mostrar, além do ritmo e da duração, a presença da melodia na língua, fato que também interessou aos estudos dos retóricos, que passaram a ser conhecidos como "prosodistas" (ABERCROMBIE, 1967).

Os estudos sobre a entoação podem ser divididos entre aqueles que foram desenvolvidos na primeira e segunda metade do século passado. No início do século XX, os estudos sobre a variação melódica da fala passaram a ser mais detalhados, a partir da preocupação dos linguistas em descrever as línguas tonais e os padrões entoacionais das demais línguas. O ensino de línguas estrangeiras ajudou a chamar a atenção para a entoação. Na linha americana, segundo Cagliari (2011), os trabalhos vêm desde Bloomfield (1933), Bloch & Trager (1942), Pike (1945) e Trager e Smith (1951). Já na tradição britânica, diz o autor, um dos linguistas mais representativos desse período foi Daniel Jones, seguindo a notação de Paul Passy para marcar a prosódia com curvas, linhas paralelas e bolinhas. Em sua obra *An outline of English Phonetics* (1918), Jones descreveu a entoação do inglês

britânico, apresentando dois tons básicos: os padrões de tons descendentes referentes à asserção e os padrões ascendentes das perguntas. Além da variação melódica dos enunciados, por meio de de uma notação especial (tamanho das bolinhas) registrava também as ocorrências da tonicidade e do ritmo.

Posteriormente, conforme Cagliari (2011), Kenneth Lee Pike publicou um estudo minucioso, intitulado *The intonation of American English* (1945), descrevendo a entoação do inglês americano. Seu objetivo era descrever o fenômeno e ajudar no ensino de inglês como língua estrangeira. Seu modelo descritivo não indicava a variação melódica de cada sílaba, mas somente onde havia um desvio da altura melódica para cima, para baixo, ou estável em um determinado nível. Nessa obra, Pike (1945) discutiu também questões teóricas e descritivas do ritmo das línguas, colocando claramente que as línguas ou são de ritmo acentual ou são de ritmo silábico. As línguas de ritmo acentual, como o inglês, marcam as batidas rítmicas nas sílabas tônicas, como se constata na prática poética de versos metrificados na Língua Inglesa. As línguas de ritmo silábico não fazem isso e o acento tônico dos enunciados não serve para compor versos metrificados.

Ainda de acordo com Cagliari (2011), foi a partir de Palmer (1922) e, posteriormente, com Kingdon (1958), que a notação entoacional passou a ser feita no próprio texto. Palmer também introduziu as noções de *head*, *nucleus* e *tail* e Kingdon, de *head*, *pré-head* e *body*. O' Connor e Arnold (1973) apresentaram estudos sobre a entoação em conversas. Para Cagliari (2011), o problema dos estudos de tradição britânica está no fato de que eles se preocuparam mais com o ensino da língua inglesa para estrangeiros do que em desenvolver uma teoria mais elaborada do fenômeno e, ainda, pelo fato de terem investido muito pouco em análises acústicas. Uma teoria mais bem formulada aparece, segundo o autor, nos trabalhos de Crystal e Quirk (1964), juntamente com a teoria de Halliday (1967, 1970 e 1973).

Como os estudos entoacionais britânicos exigiam um treinamento auditivo, sua repercussão não foi possível, uma vez que não havia estudos auditivos fora do Reino Unido (CAGLIARI, 2011). Outro impasse frente aos estudos britânicos foi a grande influência americana. Nesse contexto, Bolinger (1958) é considerado um dos linguistas que mais influenciaram os estudos sobre a entoação inglesa, como afirma Crystal (1969, p. 54), *"Bolinger seems to have written more articles on intonation and related features than any other linguist"* [1]. Foi ele quem, segundo Cagliari (2011), observou que as técnicas usadas

[1]"Bolinger parece ter escrito mais artigos e assuntos relacionados à entoação do que qualquer outro linguista." [tradução nossa].

para descrever os fonemas não eram as mesmas para descrever a entoação, uma vez que os fonemas são descontínuos e a entoação é contínua.

> Diferentemente dos fonemas, a entoação tem a ver mais com as atitudes do falante. Foneticamente, a entoação tem a ver apenas com a variação da frequência fundamental, ao passo que os fonemas dependem de muitos parâmetros fonéticos (ou propriedades distintivas).
> (CAGLIARI, 2011, p. 1).

Entrando na segunda metade do século XX, há a incidência de muitos trabalhos e diferentes pontos de visa sobre os estudos entoacionais. Liberman e Michaels (1962) desenvolvem um estudo mais acústico da entoação, estudando, a partir da síntese da fala, a interpretação do ouvinte em relação às variações de atitudes acarretadas pela curva melódica, como assegura Cagliari (2011). Em seus trabalhos, os autores reconheceram a importância da frequência fundamental (F0), mas viram que ela não é adequada para definir as emoções ocasionadas pela entoação.

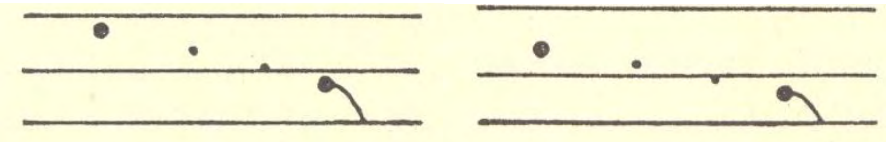
3.1.1.1 Os predecessores: modelos teóricos

3.1.1.2 O modelo de *Daniel Jones*

Daniel Jones (1972) insere-se em um modelo de análise estruturalista, que exige "[...] o reconhecimento de um número limitado de padrões e de regras produtivas de uso desses padrões" (CAGLIARI, 2002, p. 4). Jones escreve o livro *An outline of English Phonetics* (primeira publicação em 1917) com a pretensão de ensinar a Língua Inglesa aos estudantes estrangeiros.

No livro, Jones define a entoação a partir de variações melódicas da voz. A notação entoacional do linguista segue o modelo do *Handbook of English Entonation* de Armstrong e Ward (1926). Em sua metodologia, a entoação é representada pelo sistema de pontos, curvas e linhas paralelas. Os pontos denotam, aproximadamente, os níveis melódicos e as curvas, a elevação e a queda da entoação. Essas marcas entoacionais devem ser colocadas acima da transcrição fonética. Esses símbolos entoacionais aparecem dispostos em três linhas paralelas, seguindo a notação musical (com cinco linhas). Veja o exemplo de transcrição que segue.

Figura 1- Exemplo de notação segundo o modelo de Daniel Jones (1972)



Fonte: (JONES, 1972, p. 277).

As linhas superior e inferior representam, respectivamente, o limite superior e inferior da voz. A linha do meio, por sua vez, representa o nível intermediário da melodia. A proeminência melódica vem indicada por pontos grandes, normalmente, no início da curva. Se não há marcação, entende-se que a sílaba é átona. Os dois tons usados pelo foneticista são o tom 1 e 2, que, segundo ele, são suficientes para uma descrição entoacional.

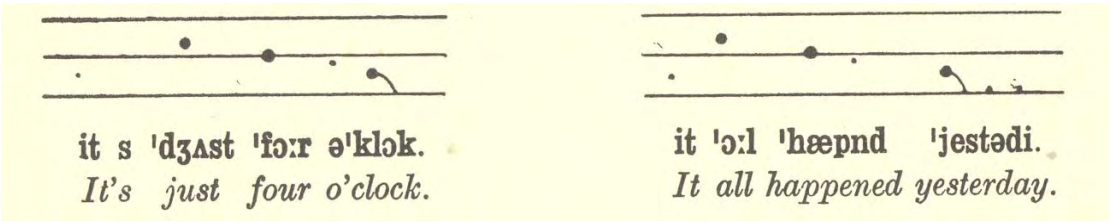
Abaixo, apresentamos, resumidamente, conforme o autor, as principais características do tom 1 e, em seguida, do tom 2. Como será observado nas figuras que serão apresentadas posteriormente, Daniel Jones não coloca, ao lado da descrição entoacional, qual é o tom de cada enunciado. A marca entoacional das sentenças é estabelecida a partir das características de cada tom. Por isso, achamos conveniente anotar algumas das principais características desses tons.

O tom 1

- As sílabas iniciais, átonas, são mais baixas. Quando há duas ou mais sílabas átonas, elas apresentam a mesma variação melódica;
- As sílabas acentuadas formam uma sequência de notas descendentes. A primeira sílaba começa com um tom alto e a última apresenta uma queda de entoação;
- Quando há mais de uma sílaba acentuada, a queda da última sílaba tônica, geralmente, começa com um tom próximo das sílabas iniciais não acentuadas, caindo para uma frequência mais baixa;
- Quando as sílabas átonas estão entre as sílabas tônicas, elas apresentam o mesmo tom da sílaba tônica que a precede, exceto para o caso das sílabas átonas, imediatamente anteriores a última sílaba tônica. Nesse caso, a última sílaba não acentuada é um pouco mais baixa do que a sílaba acentuada anterior.

A fins de exemplificação, veja dois exemplos de sentenças de tom 1 propostas pelo modelo descritivo de Jones (1972).

Figura 2- Representação do tom 1 normal



Fonte: Jones (1972, p. 280).

De acordo com Jones, os dois enunciados mostram o padrão de duas sentenças afirmativas neutras com variação melódica descendente. O modelo notacional segue as mesmas características descritas anteriormente para esse tipo de tom.

Nesses dois exemplos, abaixo da anotação da variação melódica, o autor também faz uma transcrição fonética e ortográfica dos enunciados.

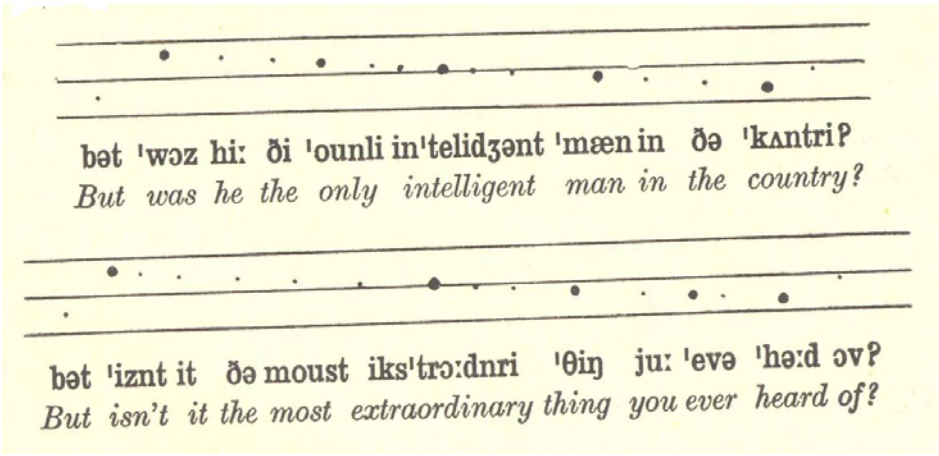
Já em relação ao tom 2, temos as seguintes características, de acordo com o autor:

O tom 2

- As sílabas iniciais são mais baixas, assim como as sílabas do tom 1;
- Quando há mais de uma sílaba tônica em um enunciado, a primeira apresenta uma altura maior do que a última;
- As sílabas átonas, seguidas da última sílaba tônica apresentam-se em uma sequência ascendente. Quando não há ocorrência de sílabas átonas, é a sílaba tônica que recebe o contorno ascendente;
- A altura da última sílaba tônica é, normalmente, mais baixa do que a sílaba átona inicial. No entanto, se ela for a única sílaba tônica do enunciado, terá a mesma altura melódica da sílaba anterior.

Os dois enunciados que seguem mostram o padrão entoacional do tom tipo 2, normal em sentenças interrogativas que exigem respostas do tipo sim e não.

Figura 3- Descrição do tom 2 normal em enunciados interrogativos



Fonte: Jones (1972, p. 283).

Em relação às características e aos tipos de frases, para Jones (1972), o tom 1 é representativo de frases declarativas; interrogativas com palavras interrogativas, assim como *why, when, which* e *how*[2]; imperativas, indicando ordem e , por fim, das frases convidativas. Por outro lado, os enunciados de tom 2 são característicos de frases que apresentam um

[2] porque, quando, que/o qual e como. [tradução nossa].

sentido incompleto; de frases enumerativas e de perguntas alternativas. Esse tom também é típico de perguntas que indicam pedidos e exigem uma resposta do tipo sim ou não[3].

A representação entoacional de Daniel Jones segue uma análise auditiva da entoação, pois, como observado, sua anotação é feita auditivamente, nas linhas paralelas, a partir da altura melódica de cada sílaba, assim como irão fazer Halliday (1970) e Cagliari (1982/2007).

3.1.2 O modelo de *Kenneth Lee Pike*

Seguindo o modelo descritivo de Daniel Jones, Kenneth Lee Pike desenvolveu um estudo minucioso sobre a entoação, no livro *The intonation of American English* em 1945. À semelhança de Jones (1917), a obra de Pike foi escrita como material didático para os estudantes de língua estrangeira, mas o autor estava mais interessado nos falantes da América Latina. Assim como Jones, a abordagem de Pike também se baseia no modelo estruturalista de análise.

Em seu livro, Pike faz, primeiramente, uma revisão dos trabalhos americanos e ingleses que trataram da entoação sobre o ponto de vista da análise perceptual auditiva e daqueles que utilizaram os mecanismos de análises acústicas. Seguido disso, o foneticista relaciona os dois métodos (auditivo e acústico) e afirma que o resultado das duas análises é diferente até mesmo para um observador mais casual. De acordo com o autor, as análises acústicas são mais objetivas devido à quantidade de medições e testes estatísticos, enquanto as análises auditivas trabalham com unidades menos objetivas.

Tendo em vista essas disparidades, Pike salienta que há muitas desavenças entre os adeptos de uma e de outra análise. Em relação às análises auditivas, ele diz que os instrumentalistas acreditam que as suposições feitas pelos adeptos de uma análise auditiva quanto à intensidade, frequência fundamental e duração são muito imprecisas e não têm validade. Por outro lado, segundo ele, os analistas auditivos rejeitam a tendência dos instrumentalistas em medir qualquer coisa, deixando de lado, muitas vezes, a questão linguística do fenômeno.

Na opnião de Pike (1945), cada um dos dois métodos tem seu lado negativo e seu lado positivo. Por exemplo, segundo ele, as análises auditivas não são tão eficientes quanto às análises acústicas no que diz respeito à descrição e à medição das características das ondas sonoras. Por outro lado, nenhum pesquisador pode afirmar, a partir de análises físicas,

[3] Vale lembrar que Daniel Jones também traz outras características entoacionais dos tons 1 e 2. Contudo, elas não serão abordadas nesta dissertação, já que o nosso objetivo é fazer uma revisão geral dos modelos descritivos.

o que determinadas ondas sonoras significam. Já nas análises auditivas, é possível questionar o falante nativo e descobrir, mesmo que vagamente, o significado dos sinais linguísticos. O foneticista segue fazendo alguns apontamentos sobre os dois tipos de análise e chega à conclusão de que o melhor método está na união do auditivo com o acústico, ou vice-versa, já que são análises que se complementam. O autor considera, ainda, que os estudos experimentais contribuem de forma significativa para as análises auditivas. Sobre essas proposições ele diz:

> *[...] the auditory analyst cannot make adequate descriptions of the units which he discovers (although he utilizes temporary rough but highly practical description in terms of producing mechanisms, for vowels and consonants, or vague degrees of intensity, frequency, and duration for pitch and stress and quantity), but must ask the instrumentalist to do it for him. On the other hand, the instrumentalist, if he wishes to describe units significant to communication, must first discover these units by deliberate auditory analysis, or utilize the best analyses available, or fall into the unconscious error of basing his researches on a tacit perceptual analysis of naïve type.*[4] (PIKE, 1945, p. 18)

No que concerne à entoação, Pike nos diz que toda sentença, palavra ou sílaba, apresenta alguma variação melódica durante a pronúncia. De acordo com ele, não há sentença sem variação melódica, de modo que nenhuma língua pode ser considerada monotônica. Segundo ele, essas variações não são casuais, mas altamente organizadas. Alguns enunciados, ele afirma, podem ser desprovidos de qualquer atitude ou sentimento, mas, ainda assim, apresentam uma função mecânica e são importantes para o aprendizado de uma segunda língua. Por outro lado, revela Pike, há aqueles enunciados que apresentam um estado psicológico individual do falante, como alegria, excitação, raiva, tristeza, entre outros. Esses enunciados ajudam na identificação do estado emocional das pessoas, determinando, conforme o autor, os significados atitudinais, que modificam o significado lexical transmitido pelas sentenças.

Em relação à descrição dos contornos entoacionais, Pike considera que não é suficiente dizer que um contorno é do tipo ascendente, descendente ou os dois, já que a mais simples subida apresenta uma estrutura interna complexa e uma série de relações com outros

[4] [...] o analista auditivo não pode fazer descrições adequadas das unidades que ele descobre (mesmo que ele utilize um método de descrição aproximado, mas muito eficiente em termos da produção de mecanismos para a análise das vogais e consoantes, ou vagos graus de intensidade, frequência e duração para a variação melódica, de tonicidade e de quantidade) mas, por outro lado, deve pedir aos instrumentalistas para fazerem isso por eles. Por outro lado, o instrumentalista, se deseja descrever as unidades significativas para comunicação deve, primeiro, descobrir essas unidades por meio de uma análise auditiva deliberada, ou utilizar a melhor análise disponível, ou cair no erro inconsciente de basear suas pesquisas em uma análise perceptual ingênua. [tradução nossa].

contornos. Em linhas gerais, para o autor, há quatro tipos de contornos entoacionais e o significado da sua variação melódica é determinado em detrimento da altura relativa do outro contorno.

Na perspectiva do linguista, a questão da altura relativa, seja ela alta, baixa ou intermediária, é determinada pela proporção da relação entre as sílabas ou frases umas com as outras e não por sua medição física exata, assim como visto em Jones (1972). Os quatro níveis de contornos entoacionais propostos por Pike são: *extra-high*, *high*, *mid* e *low*[5], numerados de um a quatro, indo do mais alto ao mais baixo. Para ele, esses níveis são suficientes para fazer uma descrição e diferenciação de todos os contornos que apresentam distinções de significado encontrados no Inglês. O foneticista também usa símbolos adicionais para descrever a tonicidade, quantidade, pausa, altura e qualidade de voz.

O contorno entoacional de nível 1, sugerido pelo autor, normalmente indica surpresa ou imprevisibilidade; o de nível 2 é o mais frequente para as sílabas tônicas e os de níveis 3 e 4, para as sílabas átonas, sendo que o último ocorre no contorno decrescente. Em qualquer contorno, conforme o autor, seja ele crescente ou decrescente, há a presença de dois níveis entoacionais.

Os níveis tonais são anotados por ele no início e no fim de cada palavra ou enunciado. Também podem ocorrer contornos com três níveis. Esse tipo de contorno ocorre, normalmente, com enunciados que apresentam primeiramente uma subida e, em seguida, uma descida. O terceiro nível ocorre nos lugares em que há uma mudança de direção no contorno entoacional. De acordo com Pike, devemos anotar apenas os locais em que há mudanças na direção da variação melódica do enunciado.

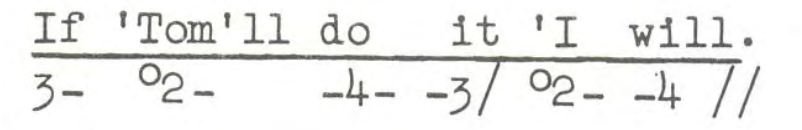
De modo geral, na descrição proposta pelo autor, as barras inclinadas indicam um grupo tonal. A bolinha sobrescrita [^{o}] antes de cada nível melódico representa uma sílaba tônica saliente. Segundo Pike, todo contorno primário é formado por uma sílaba tônica e toda sílaba tônica inicia um novo contorno primário.

Os hífens [-] marcam o intervalo entre um nível e outro. A variação melódica entre esses níveis é irrelevante. Seguidamente às sílabas tônicas do contorno primário ocorre, na maioria das vezes, uma sílaba átona, que é pronunciada com a mesma intensidade da sílaba tônica anterior. Essas sílabas são chamadas de "pré-contorno" e dependem de como as sílabas seguintes foram pronunciadas. Gramaticalmente, elas podem ser consideradas palavras independentes ou somente partes de uma palavra, como acontece com a palavra

[5] muito alto, alto, médio e baixo. [tradução nossa].

re(ceive), (PIKE, 1945, p. 29), em que a sílaba *ceive*, conforme sua pronúncia, pode ser parte da palavra ou uma palavra independente. Diante dessas considerações, veja um exemplo do modelo de descrição proposto pelo autor:

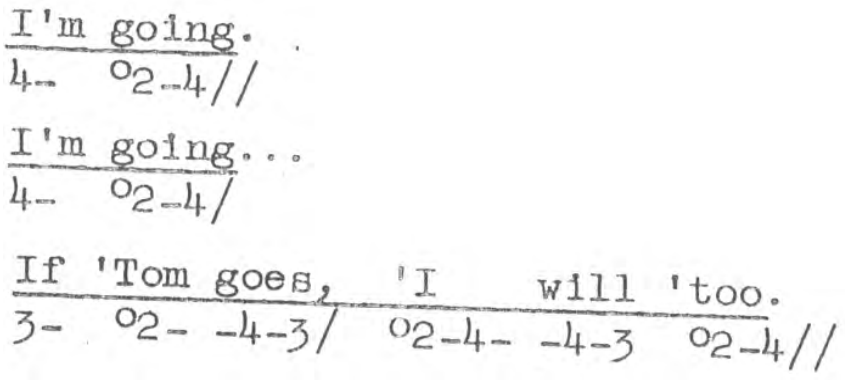
Figura 4- Representação: contorno entoacional de nível 3 na frase: "If Tom'll do it I will"



Fonte: (PIKE, 1945, p. 35).

Em relação às pausas, Pike diz que elas podem ser de dois tipos: 1) *tentative pauses* ou 2) *final pauses*. O primeiro tipo é representado por barras inclinadas simples [/] e o segundo tipo por barras inclinadas duplas [//]. *Tentative pauses* são menores do que as *Final pauses*. As *Final pauses* modificam o fim de um contorno entoacional, indicando o término de um enunciado, a partir do abaixamento da altura final do contorno. Esse tipo de pausa ocorre, normalmente, onde há atitude do falante. As *tentative pauses*, de acordo com o autor, também revelam uma atitude do falante, indicando incerteza e continuidade na fala (isto é, o discurso não acabou), como um enunciado incompleto. A diferença entre esses dois tipos de pausas, segundo Pike, está na intensidade, ou seja, no modo como o falante as enfatiza. Conforme o autor, as *final pauses* são, na maioria das vezes, mais claras, deixando evidente que o enunciado chegou ao fim, diferentemente das *tentative pauses* que acontecem, por exemplo, em momentos de hesitação, em pensamentos suspensos, entre outros, conforme diz Pike. As pausas no meio de um enunciado marcam sequências de grupos tonais e, na maioria das vezes, segundo ele, são classificadas e anotadas assim como as *tentative pauses*. Veja os exemplos dos três tipos de pausas na figura abaixo:

Figura 5- Os três exemplos de pausas propostos pelo modelo notacional de Pike

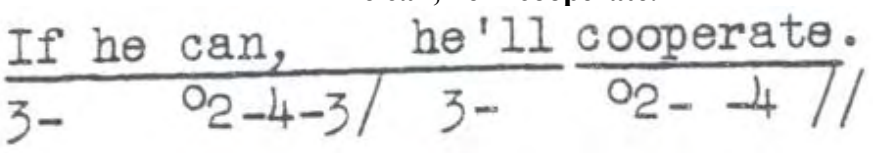


Fonte: Pike (1945, p. 33).

O primeiro exemplo mostra a *final pause*, o segundo, a *tentative pause* e o terceiro, a pausa no meio de enunciado (*midle pauses*). De acordo com Pike, no primeiro exemplo, tem-se uma implicação, uma possibilidade certa. No segundo, por sua vez, a atitude é a mesma, mas demonstra incerteza e possibilidade de mudar de ideia, caso seja convencido do contrário. Por fim, no terceiro exemplo, as pausas, separam as unidades gramaticais, contribuindo para a organização interna do discurso.

É importante lembrar que Pike também tem outras formas de transcrição. Para ele, essas diferentes formas têm uma vantagem para cada propósito de análise. Sendo assim, observe o exemplo de transcrição abaixo:

Figura 6- Modelo de transcrição no enunciado
"If he can, he'll cooperate."



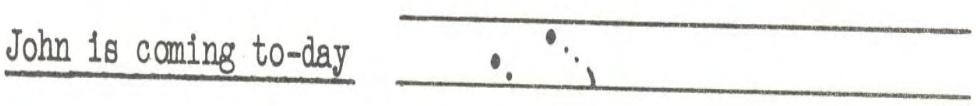
Fonte: Pike (1945, p. 42).

Na figura acima, conforme afirma Pike, temos dois grupos tonais (ou, como quer o autor, duas unidades rítmicas), *If he can* e *he'll cooperate*, separadas por uma *midle pause*. *If* e *he'll* são os precontornos. *Op* (da palavra *cooperate*) e o verbo modal *can* representam o início dos contornos, ao passo que *ate* (também da palavra *cooperate*) e a última parte de *can* representam o final do contorno. No meio da palavra *can* está a mudança de direção no contorno. Entre as palavras *can* e *he'll*, há uma fronteira entoacional.

Segundo Pike (1945, p. 42), a vantagem do sistema descrito anteriormente *"[...] lies in its unambiguous indication of the contrastive levels, the contour points, and the like, with resultant presentation of the structural organization of the material".*[6]

Veja a seguir outro modelo de transcrição utilizado pelo autor:

Figura 7- Modelo de transcrição a partir de linhas curvas e tamanhos de bolinhas



Fonte: Pike (1945, p. 41).

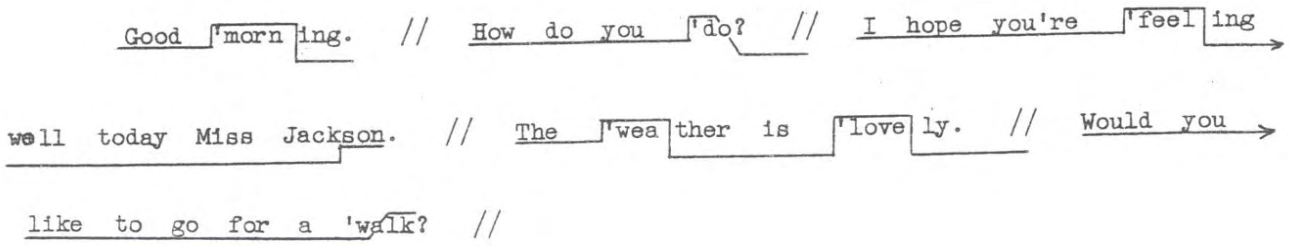
A transcrição acima mostra uma sequência de linhas curvas e vários tamanhos de pontinhos (os pontos maiores representam as sílabas tônicas), dispostos de acordo com a altura melódica no enunciado. Segundo o foneticista, esse modelo é proveitoso para mostrar sequências gerais de variação melódica, antes de uma análise fonêmica. Além disso, ele diz ser também de grande utilidade para a anotação do ritmo, através do contraste entre sílabas tônicas e átonas. Contudo, uma desvantagem desse sistema, de acordo com o autor, está no fato de ele não mostrar a organização interna dos níveis melódicos como um sistema contrastivo e de não indicar a estrutura dos tipos de contorno. (PIKE, 1945).

Outro tipo de notação mostrada por Pike é a desenvolvida pelo Instituto Staff, para o aprendizado de língua inglesa. Nessa descrição, as linhas são anotadas ao longo do enunciado, e indicam o tipo de entoação desejada. A linha que passa acima da palavra mostra o local do foco no enunciado, enquanto o abaixamento da linha sobre a palavra caracteriza a queda da entoação. Por fim, a indicação da pausa na leitura, ou final de enunciado é representada ou por uma quebra na sequência das linhas, ou por uma separação muito grande entre duas palavras ou, ainda, por marcas de pontuação. Com esse método, seguindo o foneticista, os estudantes são capazes de pronunciar e aprender os padrões entoacionais do inglês[7] com maior facilidade e rapidez. Esse exemplo de transcrição está ilustrado a seguir:

[6] "[...] está em sua clara indicação dos níveis contrastivos, dos pontos de contorno, e dos semelhantes, com uma apresentação resultante da organização estrutural do material". [tradução nossa].

[7] Em seu livro, Pike faz uma extensa notação sobre os tipos de contornos entoacionais encontrados para o Inglês Americano. Esses contornos não serão apresentados nesta dissertação, mas, para mais informações, veja em Pike (1945, p. 42-106).

Figura 8- Sistema de transcrição desenvolvido para o aprendizado de língua inglesa



Fonte: Pike (1945, p. 42-43).

3.1.2.1 O modelo autossegmental: a visão inicial de *John Goldsmith*

Mais tarde, em 1976, John Goldsmith desenvolveu um modelo de análise fonológica que ele chamou de "autossegmental". Inicialmente, esse modelo foi utilizado para a descrição dos tons nas línguas tonais e, em seguida, deu origem à Geometria de traços (cf. Clements, 1985). Goldsmith (cf. Pierrehumbert, 1980) propôs também uma abordagem autossegmental para estudar os padrões entoacionais[8]. Neste caso, as sílabas passariam a ter sequências de tons altos (H = de high) ou baixos (L = de low).

Para as línguas tonais, cada tom (H ou L) representa um nível, ligados por uma linha de associação, chamada de "co-registration". A atribuição dos tons nessas línguas era utilizada na distinção de itens lexicais e seguia regras de ajuste fonológico.

Nesse modelo, os tons vêm representados em uma camada chamada "tonal" e as vogais e consoantes em uma camada chamada "segmental". Segundo Gonçalves (2009), os processos que afetam os elementos de uma camada podem não afetar os elementos da outra. O autor mostra, por exemplo, que na língua Bantu, quando duas vogais tornam-se contíguas, a primeira deve ser sempre apagada, mas o tom não é afetado. Na figura abaixo, trazida por Gonçalves, isso pode ser visto na palavra "Kusa" (comprar). A elisão da vogal (a) não leva à eliminação do seu tom alto (H), que se junta à vogal (e). Isso mostra que as camadas são independentes, mas não isoladas, uma vez que há uma organização hierárquica dos elementos.

Figura 9- Modelo de análise autossegmental para a língua Bantu



Fonte: Gonçalves (2009, p. 213)

[8] A análise referente aos padrões entoacionais será abordada no item 7.2 "Análise do padrão entoacional dos enunciados" referente à análise dos resultados.

Atualmente, o modelo de Goldsmith para as línguas entoacionais tem sido muito utilizado, principalmente depois dos trabalhos de Pierrehumbert (1980), t' Hart et alii (1990), Ladd (1996), Hirst e Di Cristo (1998), entre outros. Descrições da entoação do português nesse modelo têm sido feitas por Frota e Vigário (1999); Moraes e Stein (2006); Fernandes (2007), entre outros.

3.1.2.2 Fonologia prosódica: o modelo de *Nespor e Vogel (2007)*

Saindo da fonologia autossegmental, Nespor e Vogel (1986, 2007; SELKIRK, 1984, 1986, 2000, entre outros), inserem-se em um modelo chamado de Fonologia Prosódica[9]. Essa é uma teoria da estrutura fonológica relacionada à sintaxe.

Nespor e Vogel estão mais voltadas para uma abordagem gerativista. De acordo com elas, no início da teoria gerativa, a fonologia era caracterizada por meio da organização linear dos segmentos e por um conjunto de regras fonológicas. Essas regras fonológicas eram definidas de acordo com os limites da estrutura morfo-sintática, sendo limitadas à interação da fonologia com o restante da gramática. No entanto, para as linguistas, em sua visão de fonologia, esse modelo não é o mais adequado, uma vez que o componente fonológico não pode ser considerado como um sistema homogêneo e, sim, como um conjunto de subsistemas em interação, cada um deles com o seu princípio próprio, assim como as fonologias autossegmental, lexical e prosódica. (NESPOR & VOGEL, 2007).

Nespor e Vogel estão, portanto, preocupadas com o subsistema prosódico e com a teoria de domínios. De acordo com elas, na teoria prosódica, a representação mental da fala é dividida em um conjunto de blocos, hierarquicamente organizados. Na fala contínua, os constituintes prosódicos da gramática são assinalados por diferentes tipos de pistas que vão desde a atual modificação segmental até a mais sutil mudança fonética. Dessa maneira, segundo as autoras, cada constituinte prosódico servirá como um domínio para a aplicação de determinadas regras fonológicas e processos fonéticos. (NESPOR & VOGEL, 2007). A hierarquia de constituintes prosódicos definida pelas autoras permite um diferente tipo de informação dos domínios fonológicos.

Em relação à diferença entre as hierarquias morfo-sintáticas e prosódicas, Nespor e Vogel (2007) explicam que elas não diferem somente na maneira como dividem os constituintes em sequências, mas também em relação à profundidade, já que

> *[...] the rules that construct the phonological hierarchy are not recursive in nature, while the rules that construct the syntatic hierarchy are; the depth of*

[9] Para a abordagem da Fonologia Prosódica consideraremos somente a visão proposta por Nespor e Vogel (2007), pois acreditamos que ela é suficiente para caracterizar esse modelo teórico de análise da entoação.

> *phonological structure is finite, while the depth of syntatic structure is, in principle, not finite*[10].
> (NESPOR & VOGEL, 2007, p. 2).

O modelo de Fonologia Prosódica desenvolvido por Nespor e Vogel (2007) é baseado no livro *On prosodic structure and its relation to sytatic structure* (1980) de Selkirk e nos estudos propostos por elas mesmas em 1982. O modelo apresentado consiste, então, na teoria dos domínios fonológicos, isto é, em uma abordagem que organiza, hierarquicamente, um determinado componente da linguagem em uma série de constituintes fonológicos, formando contextos em que as regras fonológicas se aplicam. Para elas, a teoria prosódica forma um subsistema com o componente fonológico da gramática e interage de diferentes maneiras com os outros subsistemas, representados pelas diferentes teorias (autossegmental, métrica, lexical e prosódica). Sendo assim, Nespor e Vogel dizem que a análise mais apropriada está na união entre a teoria autossegmental e a teoria prosódica, uma vez que a primeira representa o modo como as regras de domínio funcionam e, a segunda, representa como deve ser a aplicação dessas regras nos domínios.

O outro tipo de interação é baseado em um conjunto de regras que tem como fenômeno a proeminência relativa e o ritmo, que requerem um tratamento em termos de rede (*grids*) – preocupados com as regras de conteúdo – e, com a teoria prosódica – preocupados com os domínios.

O modelo prosódico do subsistema fonológico difere de dois modos do modelo da fonologia gerativa tradicional. O primeiro diz respeito à representação das estruturas e o segundo, à natureza das regras.

De modo contrário à representação linear da tradicional fonologia gerativa, a representação fonológica consiste em um conjunto de unidades fonológicas organizadas de modo hierárquico (como mencionado). De acordo com Nespor e Vogel (2007), as unidades fonológicas são definidas na base do mapeamento de regras, incorporando informações de vários componentes gramaticais. Assim, para especificar as unidades fonológicas, Nespor e Vogel (2007) propuseram sete unidades prosódicas. De forma decrescente, são elas:

1 o enunciado fonológico (U) – *phonological utterance*;
2 sintagma entoacional (I) – *intonational phrase*;
3 sintagma fonológico (φ) – *phonological phrase*;
4 o grupo clítico (C) – *clitic group*;

[10] "[...] as regras que constroem a hierarquia fonológica não são recursivas por natureza, enquanto as regras que constroem a hierarquia sintática são; a profundidade da estrutura fonológica é finita, enquanto a profundidade sintática é, em princípio, não finita". [tradução nossa].

5 a palavra fonológica (ω) – *phonological word*;
6 o pé (∑) – *foot*;
7 a sílaba (σ) – *syllable*.

Dentro dessa hierarquia prosódica, a sílaba é o menor constituinte. O pé é o próximo constituinte e pode conter uma ou mais sílabas. No modelo proposto, as sílabas não são agrupadas diretamente em palavras, mas, primeiramente, em constituintes de tamanho intermediário: os pés. Segundo as autoras, na fonologia gerativa e, de modo particular, na teoria métrica, a maioria dos argumentos em favor dos pés é baseada na atribuição de tonicidade. Isto é, segundo elas, o pé é o elemento fundamental para a determinação das sílabas tônicas e átonas tanto nas palavras quanto em enunciados maiores. (NESPOR & VOGEL, 2007). Em termos gerais, a estrutura dos pés pode ser caracterizada como uma sequência de sílabas relativamente fortes e um número de sílabas relativamente fracas. Nesse modelo, há somente uma sílaba forte por pé[11].

A palavra fonológica representa a interação entre os componentes fonológicos e morfológicos da gramática. É uma categoria que, imediatamente, domina o pé, sendo que *"[...] all the feet of a given string must be grouped into phonological words, and no other category may be so grouped"*.[12] (NESPOR & VOGEL, 2007, p. 109). Cada pé é incluído em uma palavra fonológica, ou seja, não é possível que sílabas de um único pé pertençam a palavras fonológicas diferentes.

Em relação ao grupo clítico, na fonologia, segundo Nespor e Vogel, o mais comum é considerá-los como pertencente à palavra fonológica (nesse caso, são considerados similares aos afixos) ou como pertencente ao sintagma fonológico (nesse caso, são considerados similares a palavras independentes). Contudo, para elas, um dado elemento será clítico se, junto com a palavra, for afetado por regras de sândi interno. Por outro lado, será uma palavra independente se, juntamente com a palavra, for afetado por regras de sândi externo.

Acima do grupo clítico, está o sintagma fonológico, abrigando um ou mais grupos clíticos. Segundo elas, é o constituinte que melhor define as regras fonológicas em uma série de outras línguas. Além disso, é o constituinte que faz mais uso das noções sintáticas em sua construção e, por isso, é considerado como o componente mais importante da interface entre a sintaxe e a fonologia.

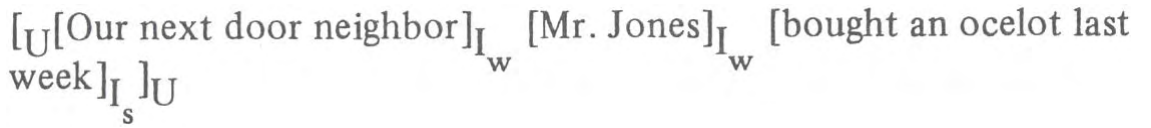
[11] Para saber mais (cf. Liberman e Prince, 1977; Halle e Vergnaud, 1978; Kiparsky, 1979; Selkirk, 1980; Hayes, 1981 – *A metrical theory of stress rules*. Ph.D. diss.: MIT. IULC).
[12] "[...] todos os pés de uma determinada cadeia devem ser agrupados em palavras fonológicas, e nenhuma outra categoria pode ser agrupada de tal modo". [tradução nossa].

Depois do sintagma fonológico, o sintagma entoacional (*I*) é a próxima unidade da hierarquia prosódica. Há fatores semânticos relacionados com a proeminência e desempenho dessa categoria, assim como a velocidade de fala e o estilo, que afetam o número de contornos presentes no enunciado. O contorno entoacional é essencial para a discussão do constituinte da hierarquia fonológica no *I*. A regra básica para a formação do *I* está, exclusivamente, na divisão do sintagma fonológico em *Is*, não permitindo nenhum tipo de variabilidade na estrutura do *I.* A variabilidade é a característica do sintagma entoacional no nível prosódico. A formulação básica para a formação do *I* está baseada na noção de que ele é o domínio do contorno entoacional e que o seu final coincide com a posição em que a pausa deve ser introduzida na sentença. Em relação ao contorno entoacional, devemos observar que há alguns tipos de construção que formam domínios por si próprios, como as expressões parentéticas, sentenças relativas não restritivas, vocativos, *tag questions*, exclamações, entre outros. A sintaxe é relevante para a reestruturação do sintagma entoacional como uma estrutura de argumentação.

O enunciado fonológico (*U*) é o último constituinte da hierarquia prosódica. Ele é constituído por um ou mais sintagmas entoacionais. É a categoria mais alta da hierarquia. Assim como outros constituintes prosódicos, o enunciado fonológico também faz uso da informação sintática em sua definição. É preciso saber que o enunciado fonológico é delimitado em seu início e fim pelos constituintes sintáticos. O enunciado fonológico também está relacionado à tonicidade. Considera-se, então, o nó mais à direita do *U* como sendo o mais forte e os outros nós, como os mais fracos. Na representação de Nespor e Vogel (2007), o nó tônico (mais forte) é representado por (S) – do inglês ***s**trong* (forte) – e o nó átono (mais fraco) é representado por (W) – do inglês ***w**eak (fraco)*.

Diante dessas afirmações, veja a aplicação desse modelo entoacional no exemplo abaixo:

Figura 10- Exemplo de representação da formação básica de um enunciado fonológico (U), sintagma entoacional (I) e regras de proeminência (W/S)



Fonte: Nespor e Vogel (2007, p. 233).

No exemplo acima, está representada a formação básica para *U*, bem como as regras de proeminência. O nó mais a direita, dominado por *U*, é o mais forte; todos os outros nós são fracos.

Ainda de acordo com as autoras, por enquanto, não há uma razão para que a fonologia de uma determinada língua inclua todas as sete unidades apresentadas. Contudo, elas assumem todos esses elementos, pois consideram que a ausência de determinada categoria poderia trazer consequências para a interface entre o componente fonológico e o resto da gramática, já que trabalham com um modelo em que cada hierarquia fonológica é definida em termos de mapeamento de regras.

Portanto, o modelo apresentado é, na fonologia, uma teoria dos domínios, uma teoria em que o fluxo de fala é organizado em um conjunto de unidades fonológicas, interagindo com a fonologia e com os outros componentes da gramática.

> *The interactions, in the form of mapping rules that build phonological structure on the basis of morphological, syntactic, and semantic notions, provide the set of phonological units necessary to characterize the domains of application of a large number of phonological rules*.[13]
> (NESPOR & VOGEL, 2007, p. 299).

3.1.3 O modelo de *t'Hart , Colier e Cohen (1990)*

Seguindo uma linha de análise perceptual, t'Hart, Collier e Cohen (1990) escreveram o livro *A perceptual study of intonation*, a fim de observar o efeito da percepção na fala e entender como a informação presente na variação melódica é processada pelo ouvinte. Segundo Ladd (1996, p. 14), esse tipo de descrição

> [...] *is in many ways the first to make a serious attempt to combine an abstract phonological level of description with a detailed account of the phonetic realization of the phonological elements.*[14]

Como Jones e Pike para o Inglês, o modelo dos autores serviu como base para uma descrição entoacional do Holandês, estendendo-se, posteriormente, para a descrição entoacional de outras línguas[15], como o Inglês, com Pijer 1983; Willems, Collier e Pijer 1988; para o Alemão, com Adriens 1991; para o Russo, com Odé 1989 e, mais recentemente, para o Francês, com Beaugendre 1994 e, para o Indonésio, com Odé e van Heuven 1994.

---

[13] As interações, na formação de regras de mapeamento que constroem a estrutura fonológica com base nas noções morfológicas, sintáticas e semânticas, fornecem um conjunto de unidades fonológicas necessárias para caracterizar os domínios de aplicação de um grande número de regras fonológicas. [tradução nossa].

[14] [...] é, em muitos aspectos, o primeiro a fazer uma série de tentativas para combinar um nível fonológico abstrato de descrição com uma detalhada descrição da realização fonética dos elementos fonológicos. [tradução nossa].

[15] Esses dados foram retirados do trabalho dos próprios autores.

O método de análise dos autores teve início no Instituto de pesquisa para a percepção (IPO) em Eindhoven, na Holanda e, depois, os autores propuseram uma teoria sobre a estrutura geral da entoação. Para o desenvolvimento da pesquisa, eles trabalharam com informantes do Holandês, tendo em vista a disponibilidade de falantes e ouvintes.

Nos capítulos que compõem o livro, os autores discutem, primeiramente, como a entoação está incorporada ao sinal de fala e como a informação mais relevante pode ser extraída de formas mais ou menos automáticas. Em seguida, eles discutem a variação melódica na fala. Para isso, começam com uma visão geral da percepção psico-acústica das impressões auditivas da entoação. Por meio desse método, os autores pretendem observar se essas percepções podem ajudar a reconhecer variações melódicas na fala, a partir de um estímulo maior e mais complexo. Tendo em vista as análises feitas, os resultados alcançados pelos autores foram, em sua maioria, negativos, ou seja, em sinais de fala mais dinâmicos e complexos, a percepção da fala foi menos precisa do que os estímulos aplicados em testes psico-acústicos.

Posteriormente, os autores introduzem uma visão básica das unidades de análise usadas por eles, bem como os procedimentos. Primeiramente, explicam como a manipulação artificial da frequência fundamental (F0) está no centro dos processos que permitem uma descrição das unidades complexas. A manipulação da curva de F0 e a sua consequência para a percepção da entoação é o foco principal do trabalho dos linguistas. Dos resultados obtidos nessa análise, t'Hart, Collier e Cohen apresentam sua teoria da entoação a partir de dez proposições.

As seis primeiras proposições dizem respeito ao lado fonético da teoria, tratando do modelo fonético de entoação usado pelos linguistas e da relação entre o perceptual e o acústico. As três proposições seguintes referem-se aos aspectos funcionais da variação melódica da fala e sua contribuição com os significados semânticos e pragmáticos do enunciado. A última preposição diz respeito à questão psicolinguística e à quantidade de um planejamento prévio, que é necessário para que o falante possa integrar requisitos melódicos e funcionais no controle de sua entoação.

Por fim, t'Hart, Collier e Cohen dedicam um capítulo especial à declinação e, posteriormente, fazem algumas generalizações linguísticas. Finalmente, no capítulo sete, aplicam os resultados fonéticos obtidos.

No decorrer do livro, t'Hart, Collier e Cohen argumentam que o progresso do estudo da entoação não pode ser alcançado, a menos que se tenha adquirido uma descrição da unidade. Contudo, dizem que isso não pode ser feito somente por intermédio de um critério

linguístico rude de distintividade, de impressões auditivas não confiáveis, nem mesmo por um cuidadoso detalhe acústico, revelado pela medição da frequência fundamental. Isso porque, segundo eles, de um lado, qualquer informação trazida pela fala não pode ser limitada apenas a critérios de distintividade linguística e, de outro, é muito improvável que todas as variações na curva da frequência fundamental sejam percebidas pelo falante.

Tendo em vista essas considerações, os autores acreditam que a solução está em desenvolver um "modelo de audição", pois a fala é *"[...] after all meant to be heard"*.[16] (t' HART, COLLIER e COHEN p. 190). Para isso, é preciso saber quais unidades melódicas são percebidas pelos ouvintes e como eles as associam à percepção do contorno melódico, relacionando a percepção a entidades melódicas mais abstratas, ou seja, os padrões entoacionais. Além disso, também é preciso saber como o ouvinte é capaz de integrar a informação melódica e textual às mensagens linguísticas.

Para eles, o estudo auditivo da entoação não exclui o interesse do estudo da produção e da acústica da fala, já que eles consideram que as impressões auditivas devem ter um respaldo das propriedades acústicas. Os autores ainda afirmam: *"No matter how systematically a phenomenon may be found to occur through a visual inspection of F0 curves, if it cannot be heard, it cannot play a part of in communication".*[17] (t'HART; COLLIER and COHEN, 1990, p. 25).

Portanto, seguindo a visão dos linguistas, entendemos que é no campo intermediário entre o psico-acústico e linguístico que as características entoacionais podem ser melhor empregadas. A solução vem, desse modo, da correlação entre as duas perspectivas, pois a análise da entoação, do ponto de vista perceptivo, deve conter apenas dados que sejam relevantes para a comunicação, uma vez que o ouvido humano não é capaz de distinguir alterações que acontecem no aspecto visual e nos valores da curva de frequência fundamental, tal qual mostrados através de uma análise acústica.

Em relação à descrição entoacional, t'Hart, Collier e Cohen, encontraram dez tipos de padrões entoacionais para o Holandês. Esses dados foram obtidos a partir da gravação de seis fragmentos de quatro peças de teatro diferentes. A leitura foi feita por quatro falantes do sexo masculino em quatro faixas de frequência diferentes. A entoação foi marcada auditivamente por dois autores. A partir da análise dessas gravações, os autores chegaram a

[16] "[...] antes de tudo para ser ouvida". [tradução nossa].
[17] "Não importa o quão sistematicamente um fenômeno possa ser observado através das inspeções visuais das curvas de F0, se isso não pode ser ouvido, não pode desempenhar um papel na comunicação" [tradução nossa].

alguns resultados, obtidos de acordo com as dez proposições elencadas anteriormente. A seguir evidenciamos alguns deles.

Os autores constataram que, a partir dos limiares estabelecidos para a variação da F0, os ouvintes não foram capazes de identificar algumas variações na F0. No entanto, isto não significa que eles não conseguiram identificar mudanças na frequência fundamental, mas mostram, por outro lado, que algumas dessas variações podem ser perceptualmente irrelevantes e, por isso, não foram percebidas.

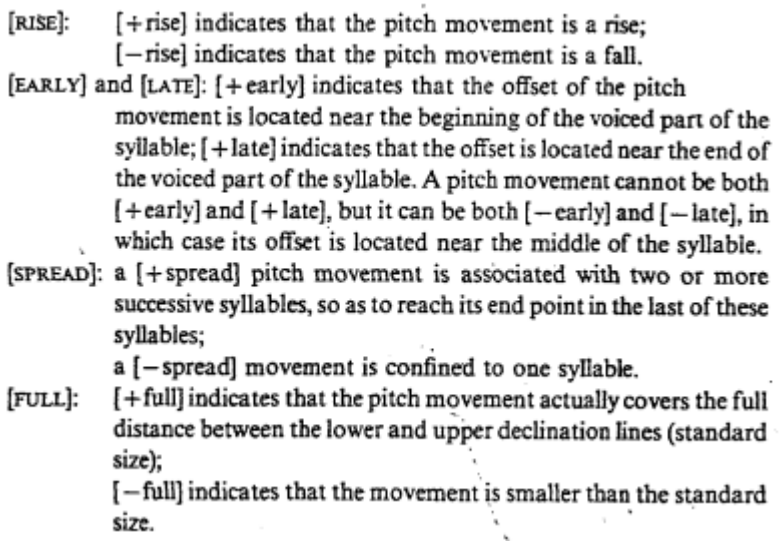
Os autores também verificaram que os movimentos melódicos devem começar e terminar em alturas audíveis distintas. Contudo, seus estudos revelaram que, para o Holandês, os movimentos melódicos podem ser dados por um mesmo valor fixo, sem que isso altere sua identidade. Para eles, é muito simplificado classificar a entoação por níveis melódicos. E, mesmo que isso seja uma tentativa fonológica recomendada, sua aplicação contraria, no nível fonético, a produção e percepção das variações melódicas.

t' Hart, Collier e Cohen (1990) observaram também que pode ocorrer um número ilimitado de contornos melódicos na linguagem, assim como os diferentes padrões entoacionais. Segundo eles, isso ocorre por dois motivos: um deles diz respeito à sequência de movimentos melódicos que pode ser diferente entre os contornos, e o outro se deve ao fato de uma mesma sequência poder ser distribuída diferentemente ao longo do enunciado.

Outra constatação feita pelos autores é a de que a correspondência entre a entoação e a sintaxe não é única e nem obrigatória. É um fato que a entoação afeta a interpretação semântica dos enunciados, sendo determinada, entre outras coisas, pelo significado das palavras e por sua relação com os constituintes prosódicos. No entanto, eles relataram que essas interpretações não apresentam um significado intrínseco, já que não foi possível saber como o falante escolhe entre os diferentes contornos, que são realizações do mesmo padrão entoacional. De acordo com t'Hart, Collier e Cohen, essa é uma questão que ainda está sob investigação.

Quanto à descrição entoacional do Holandês, os autores a caracterizaram, no nível fonético, por uma sequência de subidas e descidas. Apesar disso, eles relataram que a variação melódica não é limitada a um contraste de mudanças de direção, de modo que as subidas e descidas são diferenciadas de acordo com suas variáveis, como sua posição na sílaba, dimensão e *slope*. Para as variações melódicas, eles encontraram cinco tipos de movimentos melódicos, que são caracterizados como, *rise*, *early*, *late*, *spread* e *full*. Abaixo, mostramos a especificação de cada um dos movimentos propostos.

Figura 11- Caracterização dos contornos melódicos para o Holandês



[RISE]: [+rise] indicates that the pitch movement is a rise;
[−rise] indicates that the pitch movement is a fall.

[EARLY] and [LATE]: [+early] indicates that the offset of the pitch movement is located near the beginning of the voiced part of the syllable; [+late] indicates that the offset is located near the end of the voiced part of the syllable. A pitch movement cannot be both [+early] and [+late], but it can be both [−early] and [−late], in which case its offset is located near the middle of the syllable.

[SPREAD]: a [+spread] pitch movement is associated with two or more successive syllables, so as to reach its end point in the last of these syllables;
a [−spread] movement is confined to one syllable.

[FULL]: [+full] indicates that the pitch movement actually covers the full distance between the lower and upper declination lines (standard size);
[−full] indicates that the movement is smaller than the standard size.

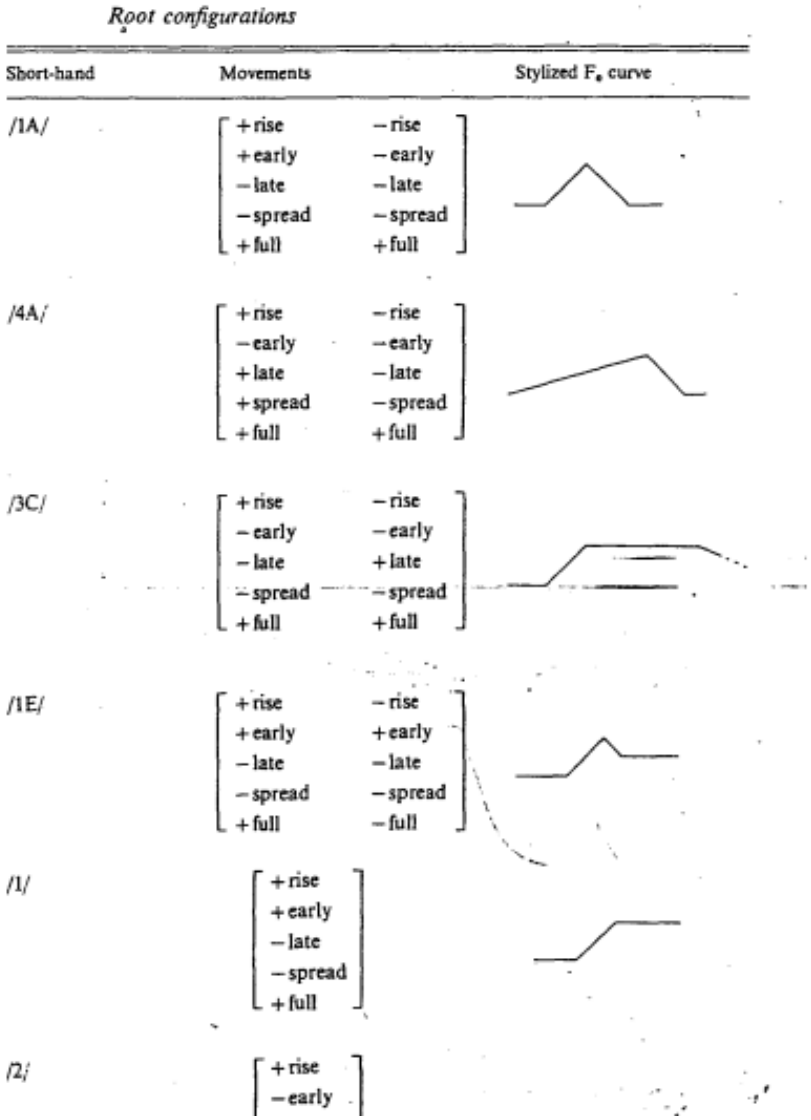
Fonte: t'Hart, Collier e Cohen (1990, p. 153).

De acordo com os autores, por meio das especificações listadas acima, é possível caracterizar todos os movimentos melódicos perceptualmente distintivos em Holandês. Cada um desses movimentos combina-se com outras unidades, chamadas por eles de "configurations" (configurações). As configurações podem ser de três tipos: Prefix (P), Root (R) e Suffix (S). Um contorno tipo (R) pode ocorrer sozinho, mas os dois outros, (P) e (S), devem, necessariamente, estar conectados a um (R).

O contorno é a maior unidade descritiva do sistema entoacional proposto pelos autores. O seu domínio é limitado a uma sentença. Um contorno, por sua vez, pode ser formado por uma sequência de configurações ou pode constituir-se de apenas uma, quando é um Root (R). Um contorno é formado pela seguinte estrutura: Contour = $(\text{Prefix})^{n}$ Root (Suffix). Os parênteses indicam que os contornos são opcionais. O exponencial (n), no prefixo, indica que ele pode ser precedido por outro prefixo. No entanto, um sufixo nunca pode ser seguido por outro sufixo. O Root (R) é o único elemento obrigatório do sistema. Essa unidade determina a configuração melódica de um enunciado e mostra qual padrão entoacional está sendo realizado. Fazendo um paralelo com o modelo de Halliday/Cagliari (1970/2007)[18], podemos dizer que um (R) compara-se aos tipos de tons propostos pelos autores. Sendo assim, as configurações melódicas características de cada unidade são representadas do seguinte modo:

[18] Veja a descrição do modelo entoacional nos itens 4.1.1 "O modelo descritivo de Halliday" e 4.1.2 "O modelo de Cagliari 1982/2007: a adaptação para o PB".

Figura 12- Caracterização e especificação da configuração dos Roots e representação do contorno entoacional da F0



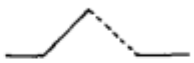
Fonte: t'Hart, Collier e Cohen (1990, p. 155).

Nesse quadro, os autores mostram a configuração dos *Roots*. Na primeira coluna, temos o tipo de contorno, na segunda, os tipos de movimentos e, na terceira, a configuração melódica característica de cada contorno. Assim, um contorno do tipo 1 será *rise*, *early* e *full* com configuração da F0 identificada na terceira coluna.

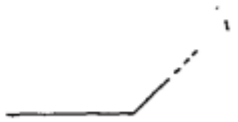
Já para o sufixo e prefixo, as configurações são do seguinte tipo:

Figura 13- Configuração dos prefixos e sufixos, com os respectivos tipos de contornos melódicos e configuração da F0

The Prefix configuration

Short-hand	Movements		Stylized F_0 curve
/1B/	+rise +early −late −spread +full	−rise +early −late −spread +full	

The Suffix configuration

Short-hand	Movement	Stylized F_0 curve
/2/	+rise −early +late −spread +full	

Fonte: t'Hart, Collier e Cohen (1990, p. 156-157).

Algumas das configurações apresentadas constituem-se de dois movimentos melódicos, outras, de apenas um. Cada movimento é associado a uma sílaba diferente. Assim, a configuração é mapeada em uma das sílabas, dependendo do número de movimentos[19] que ela contém. A seleção da sílaba apropriada depende do texto e de sua associação ao tom. A ocorrência de dois ou mais movimentos na mesma sílaba é considerado uma variação da estrutura básica dos enunciados. Os dois apontamentos elucidados são derivados por regras gramaticais.

Resumindo, o modelo proposto pelos linguistas é conhecido como "*Countour-interaction*" e não está baseado na teoria de sequência tonal. Nesse modelo, a percepção dos ouvintes da variação melódica da fala está limitada por uma informação conhecida como "*top-down*", ou seja, os ouvintes interpretam a variação melódica da fala a partir das estruturas dos contornos entoacionais presentes na gramática entoacional, o que permite integrar a melodia à informação textual.

[19] Para saber mais a respeito dessas regras veja t'Hart, Collier e Cohen (1990, p. 157-162).

3.1.3.1 O modelo de *Robert Ladd*

Diferentemente de t'Hart, Collier e Cohen (1990), a descrição entoacional de Robert Ladd (1996) está baseada na teoria de sequência tonal. Como o título do livro sugere - *Intonational phonology* – a ideia central da teoria é a de que a entoação e a variação melódica têm uma organização fonológica própria. Partindo desse pressuposto, os objetivos da sua obra são: primeiro, estabelecer uma hipótese de base teórica e empírica para o modelo métrico-autosegmental; segundo, demonstrar a viabilidade do uso dessa teoria, a partir da discussão das características da entoação, nos moldes da teoria métrica-autosegmental, também conhecida como AM; terceiro, fazer suas próprias considerações sobre a teoria métrica-autossegmental, tentando distinguir seu ponto de vista, principalmente, das ideias de Pierrehumbert (1980).[20]

Ladd insere-se, portanto, em um modelo de análise métrico-autosegmental. De acordo com Gussenhoven (2002), foi Ladd quem primeiramente rotulou o modelo. Dentro dessa abordagem, a interpretação dos tons vem de análises acústicas que podem também ter uma avaliação auditiva.

No modelo proposto por t'Hart, Collier e Cohen (1990), a questão auditiva é considerada como essencial para a análise entoacional. Como vimos, esse modelo assume que certas variações melódicas são interpretadas pelos ouvintes como relevantes e que esses movimentos são caracterizados por comandos discretos nas cordas vocais, que podem ser recuperados como um número de eventos discretos, a partir dos contornos entoacionais. Para Ladd, esse modelo, no entanto, não é suficiente para uma descrição fonológica completa, já que se trata de um método muito abstrato. Para ele, uma descrição fonológica completa consiste em especificar como essas fórmulas são realizadas, ou seja, deve-se haver uma descrição dos elementos fonológicos juntamente com uma descrição dos parâmetros acústicos.

Sobre a teoria AM, Ladd (1996, p. 42) afirma,

> [...] *the AM theory adopts the phonological goal of being able to characterize contours adequately in terms of a string of categorically distinct elements, and the phonetic goal of providing a mapping from phonological elements to continuous acoustic parameters*.[21]

Segundo o autor, a teoria métrica-autossegmental recebe esse nome porque, do lado autossegmental, trata das propriedades fonéticas e fonológicas dos tons e da relação dos

[20] Nesta dissertação, essa questão não será discutida.

[21] A teoria AM adota o objetivo fonológico de ser capaz de caracterizar contornos adequadamente em termos de uma cadeia de elementos categoricamente distintos, e o objetivo fonético de proporcionar um mapeamento, a partir dos elementos fonológicos, para os parâmetros acústicos contínuos. [tradução nossa].

pitch accents[22] com a cadeia segmental. Já do lado métrico, é uma teoria geral das relações sintagmáticas na fonologia. Segundo Prieto (sd, p. 6), é a informação *"[...] about the association of tones with metrical constituents and the relative alignment of tones with the metrically prominent syllable"*[23]. Outra definição é a dada por Gussenhoven (2002, p. 271). Ele diz que

> *The model is autosegmental because it has separate tiers for segments (vowels and consonants) and tones (H,L). It is metrical because it assumes that the elements in these tiers are contained in a hierarquically organized set of phonological constituents […]*[24]

De acordo com Ladd, a teoria AM entende a entoação a partir de quatro princípios gerais. O primeiro diz respeito à linearidade da estrutura tonal, que consiste na caracterização do contorno entoacional em uma sequência de eventos tonais, como os acentos melódicos associados às sílabas proeminentes da cadeia segmental e aos tons adicionais, associados a pontos específicos da cadeia segmental.

O segundo diz respeito à distinção entre o *pitch accent* e a tonicidade (*stress*). O *pitch accent* refere-se às pistas perceptuais concretas para o acento ou proeminência de determinadas sílabas ao longo do enunciado e está ligado à organização prosódica. A tonicidade está relacionada à proeminência melódica das palavras, que são determinadas por outros parâmetros, como duração e intensidade, além da frequência fundamental (F0). Entre outras coisas, a tonicidade, segundo o autor, deve ser entendida como pistas acústicas.

O terceiro, por sua vez, consiste na análise dos eventos entoacionais em níveis tonais, de modo que os *pitch accents* e os tons de fronteira (*edge tones*) são descritos por meio dos tons alto (H *high*) e baixo (L *low*), ou a partir da combinação desses tons. O *pitch accent* é considerado pelo autor como o evento fonológico mais importante em que um contorno melódico pode ser analisado. Por fim, o último princípio está ligado à tendência da F0 em um enunciado, como a declinação. De acordo com o linguista, esses são os princípios básicos para uma boa análise entoacional.

Aprimorando um pouco mais, os *pitch accents*, são definidos na teoria AM por uma regra local e estão relacionados à proeminência melódica da sílaba em um enunciado.

---

[22] O *pitch accent* refere-se a um elemento entoacional saliente como, por exemplo, um pico na curva melódica. Nesse trabalho será mantido o termo inglês, já que se trata de um termo técnico de uma teoria específica.

[23] "[...] sobre a associação dos tons com os constituintes métricos e o alinhamento relativo deles com a sílaba metricamente proeminente". [tradução nossa].

[24] O modelo é autosegmental porque tem as camadas separadas para os segmentos (vogais e consoantes) e tons (H, L). É métrico porque assume que os elementos dessas camadas estão contidos em um conjunto hierarquicamente organizado de constituintes fonológicos [...]. [tradução nossa].

Segundo Ladd, Bolinger (1958) foi o primeiro a propor o uso desse termo sendo, posteriormente, usado por Pierrehumbert (1980). No entanto, conforme o autor, o *pitch accent* não pode ser confundido com a tonicidade, já que o primeiro, segundo ele, é a base para os contornos melódicos e, o segundo, está separado da organização fonológica dos enunciados. Em outras palavras, a teoria AM considera que há em um enunciado um "*stress pattern*", que envolve graus diferentes de percepção da proeminência, manifestado por meio de pistas fonéticas (*stress*). Juntamente com o "*stress pattern*" há um padrão entoacional (*intonational pattern*), formado por uma cadeia de *picth accents* e tons de fronteira. Os *pitch accents* estão alinhados com o texto a partir de relações de proeminência, o que foi denominado primeiramente por Liberman (1975) de "*tune-text association*", para descrever o modo como os elementos fonológicos abstratos estão alinhados com as características segmentais do enunciado.

O alinhamento é uma propriedade fonética relativa ao tempo de eventos no contorno da F0 e dos eventos da cadeia segmental. Segundo Ladd (1996, p. 55), deve ser visto

> *[...] as the manifestation of an overarching structure in which elements of a tune are associated with elements of a text in ways that reflect the prominence relations in the text*.[25]

De um modo geral, segundo Ladd, a teoria AM interpreta a frequência fundamental (F0) como uma sequência de eventos fonológicos discretos. Esses eventos são suficientes para descrever a F0 e são denominados de tonais, já que o tom é a unidade básica. Esses eventos são: 1) os acentos frasais (*phrase accents*), que se relacionam com a sílaba acentuada, e 2) os tons de fronteira (*boundary tones*), ligados às fronteiras dos constituintes prosódicos. Nos *pitch accents*, a sílaba acentuada é formalmente indicada pelas letras H ou L, e o acento frasal é indicado por H* ou L*. Em sua notação, o linguista também faz uso do símbolo de uma seta para cima (↑) do lado esquerdo do acento frasal, para indicar que se trata de uma sequência enfática, o foco do enunciado. Para os tons de fronteira, a indicação é feita pelo símbolo % (por exemplo, H%, L%).

Entre outras coisas, a obra apresentada traz uma discussão sobre a teoria autossegmental e métrica da Fonologia Entoacional, caracterizando e criticando algumas das metodologias que fizeram uso desse modelo, bem como a de Pierrehumbert (1980), bem como uma caracterização da visão universalista da entoação. A partir dessa visão, Ladd enfatiza algumas questões como, a) a tendência à declinação; b) a questão do tom alto com

[25] "[...] como a manifestação de uma estrutura global em que os elementos de um tom estão associados com os elementos de um texto de modo a refletir as relações de proeminência no texto". [tradução nossa].

enunciados interrogativos e não finais; c) a presença dos movimentos locais dos *pitch accents* em palavras novas ou informativas. Outro aspecto levado em consideração por Ladd é a questão melódica da entoação, mostrando que as línguas divergem sistematicamente em seus padrões entoacionais[26].

Tendo em vista o artigo "The structure of intonational meaning: evidence from English", escrito por Ladd em 1980, t'Hart, Collier e Cohen (1990), criticam o modelo descritivo proposto pela teoria AM. Para eles, a redução da variação melódica perceptualmente distintiva em apenas duas unidades abstratas (H e L) prejudica a identificação dos padrões entoacionais, uma vez que contornos entoacionais distintos, derivados da mesma invariável (H ou L), podem ser apenas variantes de um ou de outro "[...] *and so must be the contours that are composed of them"*.[27] (t'HART, COLLIER, COHEN, 1990, p. 168). Por outro lado, a identificação por *Roots*, dizem eles, seria a mais apropriada, pois leva, essencialmente, a diferentes variações melódicas.

3.1.3.2 A Entoação no Português Brasileiro

Em relação à entoação no Português do Brasil, faremos apenas alguns comentários gerais sobre os tipos de trabalhos que vêm sendo desenvolvidos ao longo dos anos, apresentando, quando necessário, algumas de suas ideias teóricas básicas, uma vez que a maioria deles toma como ponto de partida a concepção dos modelos teóricos apresentados acima.

De acordo com Cagliari (2010), foi somente após o surgimento dos cursos de pós-graduação no Brasil, com as teses e dissertações, que começaram a aparecer mais estudos entoacionais. De acordo com o autor, os estudos entoacionais em PB tiveram início com os estudos de linguística aplicada a partir do modelo de Kenneth Pike (1945), como, por exemplo, os trabalhos de Rameh (1962) e Mascherpe (1970).

Conforme Cagliari, as análises mais antigas interpretam os dados tendo em vista a variação do som do fundamental (ou seja, de acordo com o autor, a F0 que corresponde ao primeiro harmônico nos sons periódicos), obtida por meio de tabelas e gráficos. As análises mais recentes, por sua vez, diz ele, além de trabalharem com os dados estatísticos, fazem suas interpretações de acordo com o modelo autossegmental métrico.

---

[26] A apresentação feita nesta subseção consistiu em uma síntese da descrição proposta pela teoria AM de acordo com a visão de Ladd. O seu aprofundamento será feito no item 5.2 "O modelo descritivo de Janet Pierrehumbert (1980)".

[27] "[...] e devem ser contornos que são formados por eles". [tradução nossa].

Segundo Tenani (2002), uma das afirmações típicas dos anos 70 e 80 é justamente sobre a falta de pesquisas nessa área, já que, de acordo com ela, o que chama a atenção é o fato de alguns dos trabalhos dessa época trazerem poucas referências bibliográficas em relação às pesquisas produzidas no Brasil. Segundo a autora,

> [...] o que se observa é uma diversidade de abordagem e de notação em que são feitos esses trabalhos, fato que contribui para a dificuldade em se constituir uma descrição do que possa ser denominado de 'gramática entoacional'.
> (TENANI, 2002, p. 14).

Cagliari (2010) também explica que a maioria dos trabalhos em PB aborda a entoação sob um ponto de vista mais fonético. Assim como o de Rameh, citado, também podemos mencionar o de Lacerda (1950), o de Lemle (1965), entre outros.

O trabalho de Rameh (1962) tem origem em sua dissertação de mestrado. Nela, a autora faz uma comparação entre os sistemas entoacionais do Inglês e do Português, a partir da gravação de entrevistas. A análise da autora é baseada na linha do descritivismo americano. Lemle (1965), por sua vez, além de trabalhar com a descrição fonêmica dos dados, também incorpora em seus estudos as funções linguísticas desempenhadas pela entoação.

Há ainda, os estudos de caráter fonológico, como os de Tenani, 2002; Fernandes, 2007; Serra, 2009; Tenani & Fernandes-Svartman, 2008; Vigário & Fernandes-Svartman, 2010; entre outros. A título de exemplificação desse modelo fonológico, discorremos brevemente sobre o trabalho de Fernandes (2007). A autora desenvolve seu trabalho tendo em vista a interface entre a sintaxe e a fonologia, com a intenção de comparar as construções de focalização em PB e PE (português europeu). Fernandes toma como base o quadro da Fonologia Entoacional, que pressupõe que a entoação apresenta uma organização fonológica (como descrito na parte referente aos "Predecessores", no item 3.1.1.1) e também toma os pressupostos da fonologia prosódica, sob a ótica de Nespor e Vogel (1986).

De acordo com Cagliari (2010), os trabalhos realizados em PB estão mais voltados para as análises acústicas, "[...] sendo mais raros trabalhos que, além da análise acústica, tragam também análises auditivas" (CAGLIARI, 2010, p. 2). Os trabalhos que levam em consideração o modelo auditivo inscrevem-se na linha de pesquisa da tradição britânica, trabalhando sob a ótica do modelo de contornos entoacionais de Halliday (1970). Entre eles, diz o autor, estão a dissertação de Gebara (1976) e Carioni (1978), os trabalhos do mesmo (CAGLIARI, L.C, 1980, 1981, 1982, 1982a 1992, 1992a, 2007) e de Rizzo (1981).

Fora dessa linha, seguindo as anotações de Tenani (2002), temos os estudos desenvolvidos por Moraes (1982, 1984, 1993, 1995, 1998) e orientados por ele, como o de Freitas (1995), Gonçalves (1997) e Cunha (2000). De acordo com Tenani, os trabalhos de Moraes seguem a linha de análise entoacional francesa, a partir da proposta de Fonágy (1981, 1983, 1993). Também nessa perspectiva está a tese de Reis (1995). Segundo a autora, esses estudos têm em comum o fato de realizarem uma análise instrumental do *corpus*.

Cagliari (2010) diz que os cursos de pós-graduação no Brasil foram um dos grandes responsáveis para a incidência de trabalhos sobre a entoação em PB. Entre alguns dos elencados acima, o autor também destaca a dissertação de mestrado de Reis (1984), Antunes (2000), Alves (2002), Queiroz (2004); a tese de doutorado de Hochgreb (1983), Pacheco (2006) e Bodolay (2009). Além desses trabalhos, também podemos citar a tese de Cruz-Ferreira (1983), a dissertação de Souza (2007) e Lucente (2008), o artigo de Abaurre (1981), os trabalho de Madureira (1982, 1991, 1999) e Lucente e Barbosa (2007, 2009, 2010), entre outros. Essa foi apenas uma pequena seleção dos trabalhos feitos em PB, mas procuramos deixá-la o mais completa possível, trazendo as anotações feitas por Cagliari (2010) e Tenani (2002), além de acrescentarmos mais algumas.

3.2 Os mecanismos de produção e percepção da fala

3.2.1 Considerações iniciais

Para entendermos as variações melódicas da fala, é importante esclarecermos como e onde ocorrem os sons, para que essas variações sejam possíveis. E, ainda, entender como o ouvinte consegue perceber os sons que chegam até ele.

3.2.2 O ar e a produção da fala

> *"[A linguagem] É um método puramente humano e não-instintivo de comunicação de ideias, emoções e desejos por meio de um sistema de símbolos voluntariamente produzidos. Entre eles avultam primacialmente os símbolos auditivos, emitidos pelos chamados 'órgãos da fala'."*
> (SAPIR, 1954, p. 22).

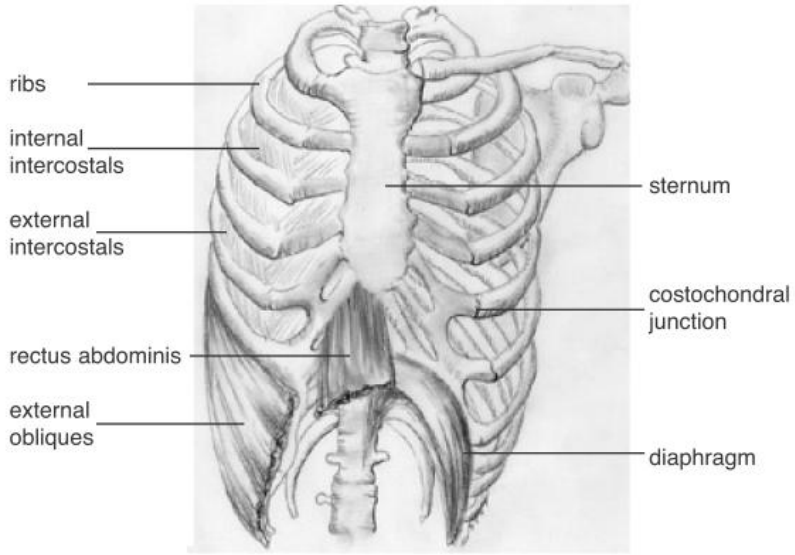
O homem utiliza mais da metade do corpo para falar. Conforme afirma Cagliari (2007), não há um órgão específico envolvido nesse processo: "A fala aparece como uma modificação do funcionamento de certas partes do corpo, de tal modo que resulte na produção de sons como forma de expressão de uma linguagem" (CAGLIARI, 2007, p. 17).

Segundo Massini-Cagliari e Cagliari (2001), a primeira etapa de produção de fala é um processo neurolinguístico. É o momento em que as ideias se juntam aos sons correspondentes do que se quer dizer em uma determinada ordem, a partir das regras da língua. Em seguida, ocorre o processo neuromuscular: "[...] o cérebro começa a enviar para os músculos mensagens para diferentes partes do corpo preparando-se para dizer o que foi planejado." (MASSINI-CAGLIARI; CAGLIARI, 2001, p. 108).

De acordo com Ladefoged (2002), a fala envolve, primeiramente, a entrada de ar pelos pulmões e, em seguida, sua saída, fornecendo energia suficiente para a produção dos vários sons até chegar ao ouvinte. Assim, as primeiras mensagens enviadas pelo cérebro agem sobre o processo da respiração que ativa os diferentes mecanismos aerodinâmicos. (MASSINI-CAGLIARI; CAGLIARI, 2001).

A aspiração e a expiração do ar pelos pulmões são realizadas por meio dos músculos intercostais. Como pode ser observado na figura 14 a seguir:

Figura 14- O mecanismo respiratório



Fonte: Ladefoged (2002, p.1).

Segundo Santos e Souza (2005), o ar pode tomar duas direções para a produção da fala: uma egressiva e a outra ingressiva. A primeira direção caracteriza a saída do ar do corpo e a segunda a entrada do ar no corpo. No português, ocorrem somente sons egressivos, podendo ser pulmonares, glotais ou velares.

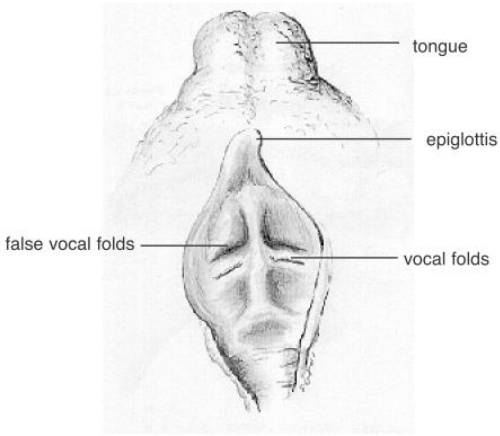
Os sons da fala são originados e modulados em função da presença do ar. De acordo com Cagliari (2007, p. 22), "O conjunto de partes envolvidas na produção dos sons da fala chama-se aparelho fonador". O aparelho fonador está organizado em três estruturas. A

primeira é a respiratória e constitui as estruturas envolvidas no processo da respiração, formando as cavidades subglotais. A segunda é a fonatória e "[...] refere-se ao modo como o ar é excitado acusticamente ao passar pela glote" (CAGLIARI, 2007, p. 26). Por fim, temos a estrutura articulatória, que se constitui pelas cavidades supraglotais.

Em geral, os sons da fala se originam na passagem do ar que ocorre entre as cordas vocais, denominada *glote*. A glote divide o trato vocal em sistema subglotal e supraglotal. Para a maioria dos sons da fala, o sistema subglotal fornece a energia necessária para o fluxo de ar. O sistema supraglotal, por outro lado, assim como a laringe, é responsável pela modulação do fluxo de ar para a produção dos sons (STEVENS, 2000). Desse modo, ao passar pelas cavidades supraglotais (faringe, nasofaringe, nasal, oral e labial), o ar é modulado e recebe o timbre característico, de acordo com a configuração articulatória do aparelho fonador.

Na laringe, as cordas vocais, conforme Stevens (2000), desempenham o papel principal na produção da fala e são, segundo Edward Sapir (1954, p. 56), "[...] para os órgãos da fala humana o que são para uma clarineta as suas duas hastes vibratórias ou para um violino as suas cordas", ou seja, essenciais. A seguir, vemos uma ilustração da representação da glote e das cordas vocais:

Figura 15- Representação da glote e das cordas vocais, vista através da abertura da laringe



Fonte: Ladefoged (2002, p. 41).

Ao passar pela glote, o ar é acusticamente excitado pela vibração das cordas vocais que "[...] libertam a cada intervalo uma pequena porção de ar, resultando a excitação acústica periódica do ar numa frequência igual ao número de batidas das cordas vocais por segundo." (CAGLIARI, 2007, p. 27). Essa frequência, mencionada por Cagliari, também é

denominada “fundamental do som”. Ela é responsável pelas características melódicas da fala. A frequência das batidas varia em decorrência da necessidade de se variar a melodia do enunciado ou entoação.

A corrente de ar vinda dos pulmões, ao ser modificada, pode apresentar características de um som periódico, quando ocorre o vozeamento ou, de um som aperiódico, “[...] isto é, fricção, no caso do efeito de turbulência causado pela obstrução no tubo laríngeo e, sobretudo, pelos estreitamentos das cordas vocais.” (MASSINI-CAGLIARI; CAGLIARI, 2001, p. 110).

As cordas vocais podem produzir três diferentes tipos de movimento, todos, conforme Sapir, importantíssimos. “Elas podem aproximar-se ou afastar-se uma da outra, podem vibrar como as hastes ou as cordas de um instrumento de música, e podem tornar-se tensas ou frouxas na direção do comprimento”. (SAPIR, 1954, p. 56). O último movimento faz com que as cordas vocais vibrem com distintos graus de tensão, levando à ocorrência das diferentes variações de altura presentes na fala, contribuindo para os diferentes modos de fonação (LADEFOGED, 2002).

Para a produção de qualquer som, ocorre uma produção de energia. De acordo com Stevens (2002), essa geração de energia envolve um processo de modulação da corrente de ar, nas cordas vocais, que pode estar próximo ou na constrição do trato vocal.

> *A source of sound can also be generated by creating turbulence or random fluctuations in the airflow, usually in the vicinity of a constriction or an obstacle. It is also possible to generate a source of sound by causing an abrupt stepwise change in the airflow in the vocal tract, through rapid release of air pressure within an enclosed volume in the vocal tract*[28]. (STEVENS, 2002, p. 55).

Portanto, o som produzido é resultado das diferentes modulações do fluxo de ar. Assim, levando em conta essas considerações, descrevemos, sucintamente, baseados em Massini-Cagliari e Cagliari (2001, p. 111-112), os principais tipos de fonação e segmentos que podem ser produzidos através das possibilidades articulatórias das cordas vocais:

Oclusiva glotal – é produzida pelo fechamento da glote. Em Português, são as expressões de surpresa que costumam apresentar a oclusiva glotal, como: [aʔ], [Eʔ] – (Ahʔ, Éhʔ);

Fricativa glotal – pode ser um segmento surdo ou sonoro que é articulado com a duração equivalente de uma consoante. É semelhante ao que acontece em inglês com a consoante inicial da palavra *horse*;

[28] Um tipo de som também pode ser gerado por meio da turbulência ou flutuações aleatórias da corrente de ar, normalmente pela proximidade de uma constrição ou de um obstáculo. Também é possível que um tipo de som seja originado por uma mudança abrupta gradual da corrente de ar no trato vocal, através da rápida liberação da pressão do ar, presente em um volume interno do trato vocal. [tradução nossa].

Vozeamento – nesse processo, os sons são produzidos por meio de vibrações das cordas vocais, o caso típico das vogais.

Ensurdecimento – é o processo por meio do qual os sons são produzidos sem que ocorra vibração das cordas vocais. Nesse caso, a corrente de ar passa pela laringe sem sofrer alterações acústicas. Como exemplo, podemos citar os sons surdos, como em massa;

Aspiração – ocorre quando a corrente de ar, ao passar pela glote, produz uma fricção local, produzindo um sussurro;

Murmúrio – é o processo que acontece quando, juntamente com as vibrações das cordas vocais, ocorre um escape de ar fricativo, produzindo, assim, um som murmurado;

***Creaky voice* –** é o processo em que a corrente de ar é modificada por meio de lentas vibrações das cordas vocais. Esse som é característico de pessoas que possuem voz grossa;

Falseto – esse tipo de fonação ocorre quando as cordas vocais estão muito esticadas, imprimindo à corrente de ar um som fundamental agudo.

Ainda tendo em vista o processo fonatório, os sons também são classificados por meio da oposição entre sons vozeados (sonoros) e sons desvozeados (surdos). Quando não há vibração nas cordas vocais, os sons são considerados desvozeados. Nesse caso, como as cordas vocais estão separadas *"[...] the air from the lungs will have a relatively free passage into the pharynx and the mouth."*[29] (LADEFOGED; JOHNSON, 2011, p. 4). Por outro lado, se as cordas vocais *"[...] are adjusted so that there is only a narrow passage between them, the airstream from the lungs will set them vibrating*."[30] (LADEFOGED; JOHNSON, 2011, p. 4), teremos, então, um som vozeado (ou sonoro).

Como vimos, a articulação das cordas vocais permite a produção de diferentes tipos de fonação e segmentos. O movimento dos lábios e da língua também resulta na produção de muitos sons. De acordo com Ladefoged e Johnson (2011), se pensarmos esses movimentos como um processo gestual, veremos que eles podem formar sons particulares. Também sabemos que muitas informações são transmitidas por meio de processos visuais, como, por exemplo, o movimento das mãos. Mesmo assim, segundo Ladefoged e Johnson (2011, p. 2), o ser humano *"[...] have found a marvelously efficient way to impart information."* [31], isto é, através do desenvolvimento dos órgãos responsáveis pela produção da fala. Mas, para que esses sons possam ser identificados, é preciso, primeiramente, que eles sejam ouvidos.

[29] "[...] o ar vindo dos pulmões terá uma passagem relativamente livre na faringe e na boca." [tradução nossa].
[30] "[...] são modificadas de forma que exista apenas uma passagem estreita entre elas, a corrente de ar vinda dos pulmões irá encontrá-las vibrando." [tradução nossa].
[31] "[...] o ser humano encontrou um modo admiravelmente eficiente de transmitir informações". [tradução nossa].

3.2.3 A produção e a anatomia da audição humana

Muitos eventos que ocorrem ao nosso redor produzem a sensação do som, como o vento, o som de algum instrumento musical, o barulho de janelas e portas batendo, a voz humana, entre outros. De acordo com Johnson (2003, p. 3), as sensações de som listadas acarretam um tipo de movimento que causam flutuações na pressão ao redor do ar. E, quando essa pressão atinge o tímpano, geram um movimento que é traduzido como uma experiência de som, que é percebida pelo ouvinte através da audição. Assim, juntamente com o olfato, o paladar, a visão e o tato, a audição ajuda a compor os cinco sentidos do ser humano (PACHECO, 2006).

De acordo com Stevens (2000), o processamento da fala pelo ouvinte exige que o sinal de fala seja interpretado em termos de uma representação linguística discreta, consistindo em uma sequência de palavras. Nesse processo, segundo ele, a forma de onda da fala passa pelo ouvido, onde é transformada, primeiramente, em movimentos mecânicos de estruturas presentes no próprio ouvido. Em seguida, transforma-se em atividade elétrica no nervo auditivo e nas estruturas que estão acima do sistema nervoso central.

Pacheco (2006) explica que a audição é muito importante, pois além de participar ativamente do processo de comunicação, permite que as pessoas ouçam os sons da fala sem, necessariamente, serem vistas pelo seu interlocutor.

O aparelho auditivo, segundo Cagliari (1982) é altamente sofisticado, pois é "[...] capaz de fazer um trabalho de análise dos sons que máquina nenhuma até hoje é capaz: discernir a fala de uma pessoa da fala de outra, por exemplo, em meio ao barulho de uma rua movimentada" (CAGLIARI, 1982, p. 18). Por isso, para o autor, ouvir é uma técnica, uma arte dentro da fonética, uma vez que o ouvido humano é capaz de detectar muito mais detalhes do que somos capazes de identificar.

O som é produzido por uma onda vibratória que, deslocando as partículas de ar, gera uma onda sonora. A onda sonora chega ao ouvido por meio de um processo contínuo de compressão e rarefação das partículas de ar. (PACHECO, 2006)

Os sons podem ser periódicos ou aperiódicos (JOHNSON, 2003). Os sons periódicos possuem um padrão que se repete regularmente, podendo ser simples ou complexos. As ondas sonoras simples, também conhecidas como ondas senoidais resultam do movimento de um único harmônico, como o balanço de um pêndulo. Embora, na fala humana, esse tipo de onda sonora ocorra raramente, ela é, ainda assim, muito importante, uma vez que os sons mais complexos são descritos a partir da combinação de ondas senoidais. No que diz respeito à repetição do padrão da forma de onda e a presença dos ciclos, as ondas sonoras

complexas são como as ondas sonoras simples. No entanto, elas são compostas por, pelo menos, duas ondas. (JOHNSON, 2003).

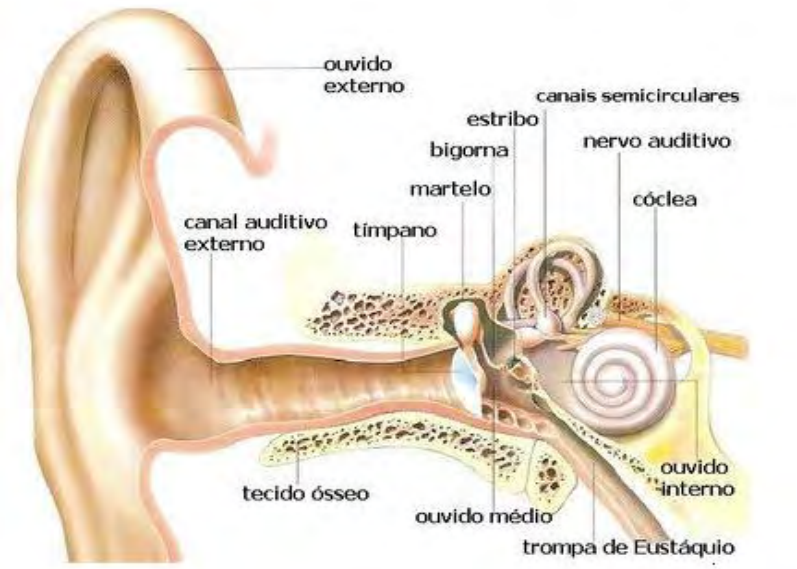
Diferentemente das ondas simples e complexas, as ondas aperiódicas não apresentam um padrão regular de repetição, formando, por sua vez, uma onda aleatória, ou seja, um padrão que não se repete. Segundo Johnson (2003, p. 12), esse tipo de onda é chamado de *"white noise"* e pode ser comparada com o som de um vento passando através das árvores.

Devido à sensibilidade do ouvido humano, o homem pode perceber e interpretar ondas sonoras "[...] em uma gama muito ampla de frequências (16 a 20.000 Hz para sons em geral , sendo o limite de percepção dos sons da fala de 8000 Hz)". (LENNERBERG, 1967; MARTINS, 1988 *apud* PACHECO, 2006, p. 16).

A captação do som, segundo Baguunyà e Peña-Casanova, 1997 (*apud* PACHECO, 2006), até a sua interpretação, acontece através de uma sequência de transformações de energias, que se iniciam com a onda sonora, passam, em seguida, à energia mecânica e terminam com a energia elétrica dos impulsos nervosos que chegam ao cérebro. Essa sequência de transformações ocorre, segundo Pacheco (2006), devido à anatomia do ouvido humano.

O ouvido humano é formado por três partes: o ouvido externo, médio e interno. Essas três estruturas estão representadas abaixo:

Figura 16- O aparelho auditivo e suas estruturas



Fonte:(http://osilenciopermanente.blogspot.com/2010/01/estrutura-do ouvido-humano.html).

É pelo ouvido externo que o som entra no canal auditivo, recolhendo os sons e os enviando ao ouvido médio. Conforme afirma Pacheco (2006), o ouvido externo é

responsável por localizar e selecionar as ondas sonoras, funcionando como um filtro acústico. “É no ouvido externo que ocorre a transformação das vibrações sonoras em vibrações mecânicas que são comunicadas aos ossículos (martelo, bigorna e estribo).” (PACHECO, 2006, p. 17).

O ouvido médio, por sua vez, transforma a energia acústica em energia mecânica. Ele é formado pela membrana timpânica, pela cadeia de ossículos e pelos músculos tensor timpânico e estapédico. (PACHECO, 2006). O ouvido médio é responsável por conduzir as ondas sonoras do tímpano para a janela oval, sem que haja uma perda significativa de energia. Uma outra função dessa estrutura é proteger o sistema auditivo dos efeitos dos sons muito altos.

Por fim, no ouvido interno, situam-se a cóclea, os canais semicirculares e o nervo auditivo. A cóclea, segundo Pacheco (2006), está imersa em um líquido e insere-se em uma cavidade óssea – o labirinto. É a parte do ouvido responsável por transformar os sons em estímulos elétricos. O labirinto, por sua vez, caracteriza-se por manter o corpo em equilíbrio e nada tem a ver com a audição. De acordo com Pacheco, os fluídos da cóclea sofrem uma compressão, deslocando a membrana basilar (localização das células ciliadas) para cima e para baixo, gerando diferentes frequências. As células ciliadas formam os primeiros receptores neurológicos da audição, que iniciam a excitação nervosa nas áreas cerebrais, específicas para a percepção auditiva. Após receber os impulsos elétricos, de acordo com a autora, há uma análise das excitações nervosas recebidas, que são, em seguida, analisadas pelo cérebro e decodificadas para a compreensão da mensagem verbal, ocorrendo assim, a percepção da fala.

3.2.4 A percepção da fala e as teorias envolvidas

Segundo Mateeuscu (2003 apud PACHECO 2006), a audição e a percepção estão intimamente ligadas, uma vez que a percepção organiza os sons de acordo com os padrões já presentes em nossa mente, distribuindo-os em imagem acústica. Nessa etapa da audição, juntam-se os processos psicológicos que decodificam, interpretam e arranjam os sons conforme os padrões linguísticos da língua. Assim, segundo Pacheco, quando “[...] ouvimos alguém falar, duas operações distintas estão ocorrendo: a audição propriamente dita, ou seja, a captação do som pelo nosso aparelho auditivo, e a transformação dessa informação elétrica em mensagem”. (PACHECO, 2006, p. 18).

Do ponto de vista acústico, Keith Johnson (2003) considera que a representação perceptual da fala é muito diferente da representação auditiva. Para ele (JOHNSON, 2003, p. 59)

> *Acoustically, speech sounds are sound waves that are produced by vocal tracts in motion. Auditorily, they are shaped by the nonlinearities of the auditory system. Perceptually, speech sounds are no longer even limited to audition – the visual image of the talker plays a role in perception – and perception takes place in the context of a lifetime of experience with language*[32].

Como se vê, o autor parte da descrição física do sistema auditivo. Nesse sentido, ele considera a distintividade entre a representação perceptual e auditiva. Contudo, no decorrer do seu trabalho, ele mostra que a caracterização da fala é construída a partir de um *input* auditivo e mais algumas outras informações.

Para Pacheco (2006), o maior problema entre a questão perceptiva e auditiva é saber identificar como o ouvido humano é capaz de interpretar, classificar e arranjar os sons a partir dos padrões fonológicos da língua, pois, segundo ela, o ouvido pode captar invariâncias na fala,

> [...] já que a fala possui um caráter contínuo e tem uma realização que se dá em tempo real, o que faz com que sons adjacentes sejam co-articulados, ou seja, sofram influências acústicas e articulatórias mútuas, não havendo limite preciso entre um e outro.
> (PACHECO, 2006, p. 18).

A fim de explicar questões como essas, revela a autora, muitas hipóteses foram propostas. Trataremos aqui de duas delas. Uma delas diz respeito à suposição de que a percepção parte de gestos articulatórios e, partindo da produção, possuem uma base articulatória. A outra hipótese propõe que a percepção da fala tenha início com o sinal acústico, sendo, portanto, de base essencialmente auditiva.

A primeira hipótese é chamada de Teoria Motora da percepção. Essa teoria foi proposta por Liberman; Shankweiler e Studdert-Kennedy em 1967 e teve como incentivo inicial o desconforto com a variabilidade do sinal acústico. Liberman foi um dos primeiros teóricos a assumir que havia uma relação direta entre os aspectos articulatório e fonológico.

[32] Acusticamente, os sons da fala são ondas sonoras produzidas pelas cordas vocais em movimento. Auditivamente, elas são moldadas pela não linearidade do sistema auditivo. Perceptualmente, os sons da fala não são tão limitados à audição – a imagem visual do falante desempenha um papel na percepção – e a percepção tem um lugar no contexto de experiência de vida com a linguagem. [tradução nossa].

Segundo ele, os sinais acústicos recuperam os *"gestures"* articulatórios que os produzem. Ele diz:

> *These gestures are the primitives that the mechanisms of speech production translate into actual articulator movements, and they are also the primitives that the specialized mechanisms of speech perception recover from the signal*[33].
> (LIBERMAN; MATTINGLY, 1989, p. 491).

Os gestos articulatórios são caracterizados como invariantes e ativadores da percepção. Eles "[...] compreendem o comando neuromotor e a intenção de produção de sons por parte do falante, sendo, portanto, de natureza abstrata." (DIEHL; LOTTO; HOLT, 2004 apud PACHECO, 2006, p. 5).

Para Fowler e Galantucci (2006), o conceito da teoria motora resume-se em três pontos básicos: aquele que os ouvintes percebem como gestos pretendidos, aquele em que a percepção é alcançada por um módulo do sistema nervoso dedicado à percepção e produção da fala e, por fim, aquele em que a percepção da fala recruta o sistema motor da fala.

O primeiro conceito diz respeito aos *gestures* (gestos) e parte do princípio de que eles são elementos da competência fonológica, primeiramente reconhecidos no trato vocal e não na mente, ou seja, segundo os autores, ações linguisticamente significativas do trato vocal. (FOWLER; GALANTUCCI, 2006).

Os *gestures* não são movimentos fonológicos de articuladores individuais, mas, sim, ações coordenadas de um ou mais articuladores, responsáveis por alterar a forma do trato vocal. Essas ações, conforme Fowler e Galantucci, são responsáveis por criar e liberar constrições, como, por exemplo, o fechamento bilabial, que ocorre tanto para o inglês quanto para o português com os fonemas /b/, /p/ e /m/. Já as duas últimas asserções estão intimamente relacionadas e dizem respeito ao modo como a percepção e a produção fazem uso do sistema motor da fala. Para eles,

> *[...] the motor theoretical interpretation of duplex perception would be buttressed by evidence favoring the third claim of the theory, that there is a motor system or motor competence involvement in perceiving speech"* [34].
> (FOWLER; GALANTUCCI, 2006, p. 636).

[33] Esses gestos são os primeiros que os mecanismos de produção da fala traduzem em efetivos movimentos articulatórios, eles também são os primeiros que os mecanismos especializados na percepção da fala recuperam a partir do sinal. [tradução nossa].

[34] [...] a interpretação teórica motora da dupla percepção poderia estar sustentada pela evidência que favorece a terceira afirmação da teoria, a de que há um sistema motor ou um envolvimento da competência motora na percepção da fala. [tradução nossa].

Segundo os autores, há uma ligação entre os mecanismos de produção e de percepção da fala na teoria motora, que é realizada em uma parte do circuito neural, responsável pela percepção e produção da fala. Em outras palavras, a percepção dos sons da fala não é atribuída ao mecanismo de audição e percepção, mas sim, a um módulo específico da fala, inato e único para a espécie humana. Em 1967, Liberman *et alli* já haviam reconhecido a participação de um centro de controle motor na representação acústica dos fonemas. *"Some control center must therefore 'know' what the syllable is to be in order that its component subphonemic features appropriately combined and overlapped."*[35] (LIBERMAN *et alli*, p. 449).

Sendo assim, de modo simples, a teoria motora evidencia a relação entre a percepção e a produção, de modo que os ouvintes lançam mão do conhecimento da articulação e produção dos sons para a sua percepção, o que é efetivamente corroborado por Liberman *et alli* (1967, p. 452) *"[...]speech is perceived by process that are also involved in its production."*[36] Esse ponto de vista é motivado pelo fato de que quem ouve a fala é também um falante da língua e, por isso, possui todos os mecanismos necessários para permitir, através de sucessivas operações de codificação, que a linguagem resulte em um sinal acústico. (LIBERMAN *et alli*, 1967). E isso, conforme Libermam *et alli* e Fowler e Galantucci (2006), só é possível porque *"[...] a linkange exists in speech mechanisms supporting speech production and perception receives additional support when it is considered in a large context of research."*[37] (FOWLER; GALANTUCCI, 2006, p. 644).

De acordo com Liberman (1996), essa ligação entre os mecanismos de percepção e produção é responsável pela implementação da fala. Segundo ele, o comando neuro-motor é uma entidade abstrata e invariante que constitui o verdadeiro subconjunto dos sinais neurais em que o modelo geral de percepção e produção estão interessados. Assim, em resumo, a teoria motora propõe que os ouvintes interpretam a fala a partir de uma atividade motora articulatória que já lhe é interiorizada.

Sem apresentar nenhuma relação com o modo como o som é produzido, a segunda hipótese, idealizada por Stevens (1972, 1989), diz respeito à Teoria Quântica da percepção. Essa teoria teve início em 1972, quando Stevens observou que a relação entre as características distintivas da acústica e da articulação pareciam ter características quânticas.

---

[35] "Algum centro de controle deve, portanto, 'saber', qual sílaba está em ordem para que as características dos seus componentes subfonêmicos estejam apropriadamente combinadas e sobrepostas. [tradução nossa].
[36] "[...] a fala é percebida por um processo em que sua produção também está envolvida." [tradução nossa].
[37] "[...] porque existe uma ligação nos mecanismos da fala, apoiando sua produção e a percepção recebe um apoio adicional quando é considerada em um amplo contexto da pesquisa." [tradução nossa].

Isto é, "*[...] when a particular articulatory dimension is manipulated through a range of values, there is a nonlinear relation between this dimension and its acoustic consequence."*[38] (STEVENS, 2010, p. 10).

Stevens propôs, então, a teoria quântica, a fim de explicar que as línguas não usam todas as possibilidades disponíveis de articulação da fala, o que ocorre, segundo ele, devido ao mapeamento entre a articulação e a acústica, que é não linear. A teoria quântica tem como referência principal o processo auditivo, assumindo que o sinal físico é suficiente para a decisão perceptual.

Quanto à ideia da não linearidade, mencionada na citação anterior, Johnson (2003, p. 83) faz uma analogia ao botão que controla a sintonização de um rádio. Ele diz:

> *Nonlinear control is like the tuning knob of the radio. As you turn this knob, the radio receives different stations. So a small turn of the knob may result in a large change in the sound produced by the radio, and a large turn may cause very little change (all static)* [39].
> (JOHNSON, 2003, p. 83)

Para Pacheco (2006), a não linearidade proposta pela teoria quântica também pode ser explicada pela configuração de platôs ou saltos nos parâmetros acústicos, devido à variação do parâmetro articulatório.

De acordo com Stevens,

> *[...] there appear to be ranges of the articulatory parameter for which there is a very little change in the acoustic parameter and other ranges where the acoustic parameter is more sensitive to changes in articulation.*[40]
> (STEVENS, 1989, p.3).

Assim, segundo o autor, a relação entre a articulação e a acústica é quântica no sentido em que o padrão acústico mostra uma mudança de um estado para outro, conforme a variação na taxa de valores dos parâmetros articulatórios.

Para essa teoria, as regiões de platôs mencionadas por Pacheco são correlatos de traços distintivos, de modo que os contrastes linguísticos levem a diferenças nas regiões quânticas. "Todas as regiões quânticas definem sons contrastivos e todos os sons contrastivos diferem quanticamente" (PACHECO, 2006, p. 23). Além disso, a teoria

[38] "[...] quando uma dimensão articulatória particular é manipulada através de uma série de valores, há uma relação não linear entre essa dimensão e a sua consequência acústica." [tradução nossa].
[39] O controle não linear é como o botão de sintonização do rádio. Conforme você ativa esse botão, o rádio sintoniza diferentes estações. Assim, uma pequena volta no botão pode resultar em uma grande mudança no som produzido pelo rádio, por outro lado, uma volta grande pode provocar uma mudança muito pequena (tudo estática). [tradução nossa].
[40] [...] parece haver valores dos parâmetros articulatórios para os quais há uma mudança muito pequena no parâmetro acústico e outros valores onde o parâmetro acústico é mais sensível a alterações na articulação. [tradução nossa].

quântica da percepção também considera que traços como ponto e modo de articulação são percebidos de modo mais rápido pelos ouvintes do que os traços secundários como palatalização, labialização, entre outros. (PACHECO, 2006).

Portanto, por ser uma teoria de base auditiva, ela se fundamenta na percepção a partir da audição, de modo que a percepção da fala não tem nenhuma relação com a produção. De acordo com Stevens (2010), o parâmetro acústico é relativamente insensível a alterações em apenas uma parte dos parâmetros articulatórios, mostrando, conforme ele, *"[...] a relatively rapid change with articulation over another part of its range"* [41] (STEVENS, 2010, p. 10).

Segundo o autor, foi proposto que as regiões de insensibilidade dos atributos acústicos poderiam proporcionar uma base quantitativa para a definição das características distintivas. A tendência à relação quântica entre os parâmetros acústico, auditivo e articulatório, de acordo com Stevens (1989), é o principal fator para a formação do inventário dos atributos acústicos e articulatórios que são utilizados para indicar as diferenciações na língua.

Assim, por estar embasada em um aspecto sensorial, sensitivo dos padrões acústicos, a produção dos sons da fala ganha um papel secundário. A partir dessa sensitividade, a teoria quântica propõe um padrão abstrato, que, estocado na memória, torna-se a base para o reconhecimento dos sons da fala.

Dessa maneira, o principal objetivo da teoria quântica é saber o porquê de os sistemas fonológicos das línguas possuírem um número restrito de sons, já que as possibilidades articulatórias são consideravelmente grandes. Para Pacheco (2006), essa desproporção entre as propriedades articulatórias e o número de sons do sistema fonológico das línguas ocorre, justamente, devido à não linearidade entre os parâmetros articulatórios e o *output* acústico, "[...] podendo ter casos em que os parâmetros acústicos sejam relativamente insensíveis a uma mudança grande no parâmetro articulatório". (PACHECO, 2006, p. 22). Sendo assim, as mudanças articulatórias podem não produzir efeito no *output* da fala, uma vez que movimentos no trato vocal produzem uma situação acústica estacionária e, por consequência, a articulação não afeta a perceptibilidade do som.

Como se viu, as duas teorias apresentam conceitos diferentes em relação ao entendimento da percepção da fala. A primeira, Teoria Motora da percepção, compreende que o sinal acústico não é suficiente para o entendimento da percepção, sendo necessário o conhecimento sobre a produção da fala, de modo que a articulação é o princípio fundamental

[41] "[...] uma mudança relativamente rápida com articulação sobre a outra parte da sua escala." [tradução nossa].

para a percepção. A segunda, Teoria Quântica, por sua vez, compreende que a decisão perceptual pode ser extraída a partir dos padrões acústicos, descartando a ação da produção, devido à não linearidade entre os parâmetros articulatórios e o *output* acústico, pois, para essa teoria, os parâmetros acústicos são indiferentes a mudanças no parâmetro articulatório.

Assim, à sua maneira, as duas teorias assumem que a percepção da fala é de natureza mono modal, isto é, ocorre por meio da ação exclusiva da audição (ponto de vista da Teoria Quântica) ou através da ação da visão (ponto de vista da Teoria Motora). Pacheco (2006), no entanto, discorda dessa visão mono modal. Para ela, a percepção da fala ocorre por meio da correlação entre a audição e a visão, o que, segundo a autora, já havia sido demonstrado em 1976 em um trabalho de McGurk e McDonald.

3.3 A questão acústica e auditiva para a categorização dos aspectos da fala

3.3.1 Considerações iniciais

Na parte referente aos predecessores, abordamos brevemente a questão do trabalho acústico e auditivo, a partir da concepção de Pike (1945). Nesta parte, damos continuidade, de forma mais abrangente, ao estudo desses dois aspectos (acústico e auditivo), mostrando os diferentes interesses de cada um deles. Nossa intenção não é criar um abismo entre eles, mas, pelo contrário, justificar a importância da sua correlação, levando em consideração a integração entre pesquisas instrumentais e pesquisas auditivas da fala[42].

3.3.2 O acústico e o auditivo

O estudo da acústica da fala, segundo Cagliari (2007b), teve início no final do século XIX e início do século XX, a partir do surgimento de vários laboratórios de fonética. Nesses laboratórios, o interesse principal era

> [...] desenvolver sistemas mais aperfeiçoados de telefonia e trabalhar no desenvolvimento de sistemas de fala sintética. Vários deles construíram aparelhos ou modelos de fala sintética. Por outro lado, desenvolveram teorias acústicas da fala como, por exemplo, a noção de formante e de transição.
> (CAGLIARI, 2007b, p. 17).

[42] É importante salientar que muitas das citações e dados referentes a essa etapa da dissertação, principalmente no que diz respeito à cooperação entre as análises auditivas e acústicas, fazem parte do trabalho escrito por Luiz Carlos Cagliari, Mariane Carvalho e Mariana Moretto Gementi para a apresentação do GSCP 2012 – *VII International GSCP Conference: Speech and Corpora*, intitulado "Revisitando a questão da fonética acústica e auditiva dentro e fora dos laboratórios". Será publicada somente uma versão em inglês do trabalho com o título: *Revisiting the acoustic and auditory approach to speech analysis*.

Até o surgimento dos primeiros computadores pessoais (década de 60), os fenômenos fonéticos eram, em geral, descritos por meio de uma interpretação auditiva, a partir de uma descrição fonológica. Posteriormente, essa tradição foi perdendo espaço para as análises acústicas, com o surgimento de um novo grupo de foneticistas: os formados em laboratórios. Com esses foneticistas, as análises acústicas eram obtidas por meio de espectrógrafos e outros aparelhos para a análise da forma de onda dos sons. (CAGLIARI, 2007b). Em consequência disso, segundo Cagliari *et alli* (2012), criou-se uma divisão, por alguns estudiosos, entre fonética e fonologia, sendo atribuído à primeira uma investigação física experimental dos sons da fala e à fonologia um aspecto "[...] psicoacústico ou auditivo da representação mental do aspecto sonoro dos sistemas das línguas." (CAGLIARI *et alli*, 2012, p. 1).

Com o surgimento de novos equipamentos, entre eles os computadores, Cagliari diz que

> [...] muitos foneticistas se formaram fora da tradição lingüística, tornando-se especialistas em análises acústicas da fala. Os objetivos mais ambiciosos não estavam voltados para a descrição lingüística das línguas, mas no uso da fala para a comunicação (telefonia e produção de software).
> (CAGLIARI, 2007b, p. 18-19).

O aparecimento dos computadores também permitiu uma maior sofisticação das pesquisas, além de gerar um novo banco de dados para a fonética, devido ao grande número de pesquisas realizadas nos laboratórios de fonética.

> O volume passou a ser tão grande que a tradição de treinamento em reconhecimento e produção de sons da fala, feito auditivamente, teve uma drástica redução, a ponto de muitos foneticistas trabalharem apenas com os equipamentos, sem contar com o devido treinamento auditivo.
> (CAGLIARI *et alli*, 2012, p. 3)

De acordo com os autores, com essas pesquisas mais experimentais, surgiram também novos linguistas, que, sem apresentar um conhecimento linguístico mais sofisticado, passaram a considerar a interpretação estatística como o mais correto e válido instrumento de análise. Seguido disso, as interpretações auditivas passaram a ter menos valor e as variações físicas dos elementos linguísticos passaram a ser mais importantes, deixando a análise línguística para um segundo plano. Em decorrência, também houve uma explosão de trabalhos acústicos de fala, mostrando,

> [...] em geral, que a fala é um caos enorme, que somente tratamentos estatísticos muito sofisticados conseguiriam fazer ver algum resultado relevante e constante no meio de um turbilhão de variações sem sentido.
> (CAGLIARI, 2007b, p. 20).

Esse fato, portanto, conforme o autor, reduziu significativamente os trabalhos auditivos, de modo que muitos foneticistas esqueceram o treinamento auditivo e passaram a trabalhar somente com equipamentos. Segundo Cagliari (2007b), esse acontecimento também foi responsável por transformar a fonética em um "caos linguístico", de modo que as interpretações linguísticas mais antigas foram substituídas pelas interpretações estatísticas e o falante vivo foi substituído por um informante de laboratório. De acordo com o autor,

> Os falantes ouvem e falam, entendem e dizem o que querem, expressam suas idéias e sentimentos através da fala, usando os recursos segmentais, supra-segmentais e prosódicos com a maior facilidade. Então, o fenômeno deve se apresentar à mente dos falantes como algo bem sistematizado e com acesso fácil às informações requeridas. Em outras palavras, assim como as palavras são formadas de morfemas, e estes de fonemas, assim também o conteúdo semântico da prosódia deve se estruturar em unidades de fácil identificação para falantes e ouvintes, não podendo ser o resultado de fórmulas estatísticas complicadas.
> (CAGLIARI, 2007b, p. 23-24).

Conforme Cagliari (2007b), a estatística não pode definir o que é invariante e o que é variante, porque se trata de uma questão de qualidade do sistema e não de quantidade. Os resultados estatísticos não valem como resultados linguísticos.

> Para se saber a que fonema um som está ligado, é preciso determinar o sistema fonológico, de nada adianta chegar a uma conclusão estatística qualquer, partindo de análises acústicas. Para determinar o sistema fonológico, o lingüista precisa de sua intuição como sujeito falante, não de taxas estatísticas.
> (CAGLIARI, 2007b, p. 24).

O foneticista complementa afirmando que as investigações acústicas revelam detalhes da natureza dos sons da fala que ou são descartadas como irrelevantes pelos ouvintes ou nem sequer são percebidas, dada a natureza biológica da audição. O que é irrelevante é irrelevante, diz o autor. O que não se percebe auditivamente, ele continua, não pode constituir parâmetro linguístico, pois estaria fora do alcance dos falantes.

Nesse sentido, Cagliari (2007b) prossegue e acrescenta que os trabalhos de investigação acústica da fala destinam-se a finalidades de interpretação linguística. Suas limitações, exageros ou incapacidades precisam ser bem declaradas, uma vez que a fala é uma realidade linguística e esta é uma realidade mental e não física.

3.3.3 As discussões

A partir do crescente interesse pelas análises acústicas, em que os dados estatísticos são vistos como os mais "exatos" e mais "fiéis" aos estudos linguísticos, a questão física (acústica) e psicológica (auditiva) passa a ser alvo de muitas discussões e polêmicas entre

aqueles que são a favor dos estudos auditivos e aqueles que não os consideram como uma forma de análise possível e segura (CAGLIARI *et alli*, 2012). Como afirma Crystal (1969, p.3),

> *The realization that there are two sides to the understanding of the fact of speech – the physical and the psychological – and two main techniques of recognition which complement rather than alternate with each other, has led many linguists and phoneticians to criticize any approach which would be based on only one or the other*[43].

Diante dessa polêmica, surge um conflito entre os pesquisadores, criando uma divisão histórica entre os foneticistas tradicionais (que têm como base uma interpretação auditiva) e os foneticistas de laboratórios (que têm como base uma interpretação acústica). O problema tem início com as pesquisas instrumentais que passam a ter problemas específicos, ligados à teoria ou ao ponto de vista do novo linguista. A esse respeito, Kohler (1981, p. 175) faz um comentário sobre Pancocelli-Calzia:

> *In Hamburg, Giulio Panconcelli-Calzia became head of the phonetics department. A pupil of Rousselot's, he built up experimental phonetics as a natural science (Phonetik als Naturwissenschaft 1948), regarding phonetics as a part of physiology, a sub-section of the study on motion like walking, running, jumping, at the same time ridiculing the philologus auricularius furibundus. Panconcelli-Calzia never came to grips with the linguistic point of view and he had no criteria for separating normal and deviant phenomena. His theoretical outlook and his experimental and instrumental methods caused all phonetic categories (sound, syllable etc.) to disintegrate. He turned them into fictions of investigation without relevance to the communication process (Das Als Ob in der Phonetik – 1947). He thus helped to prepare the ground for the division of the discipline into phonetics and phonology*[44].

Dessa maneira, alguns foneticistas passaram a considerar que somente os dados avaliados instrumentalmente seriam suficientes para descrever corretamente os dados da fala. Outros, por sua vez, passaram a considerar que a fala só poderia ser explicada

---

[43] A compreensão de que há dois lados para o entendimento da fala – o físico e o psicológico – e duas técnicas principais de reconhecimento que se complementam, ao invés de alternarem entre si, levou muitos linguistas e foneticistas a criticarem qualquer outra abordagem que estivesse baseada apenas em uma das duas [tradução nossa].

[44] Em Hamburgo, Giulio Pancocelli-Calzia tornou-se chefe do departamento de fonética. Um discípulo de Rousselot que fez da fonética experimental uma ciência natural (*Phonetik als Naturwissenschaft 1948*), quanto à fonética como parte da fisiologia, uma subseção sobre o estudo do movimento como caminhar, correr e saltar, ao mesmo tempo ridicularizando o *philologus auricularius furibundus*. Panconcelli-Calzia nunca chegou a lidar com o ponto de vista linguístico e ele nunca teve critério para a separação de fenômenos normais e desviantes.

Sua perspectiva teórica e seus métodos experimentais e instrumentais causaram a desintegração de todas as categorias fonéticas (som, sílaba etc).

Ele as transformou em investigações fictícias sem relevância para o processo comunicativo. Com isso, ele ajudou a preparar o terreno para a divisão da disciplina em fonética e fonologia. [tradução nossa].

linguisticamente, com relação ao sistema da língua, dispensando qualquer dado instrumental.

Straka (1965) considera que a fonética acústica faz uso apenas de instrumentos laboratoriais para suas análises, enquanto a fonética auditiva estaria baseada somente no ouvido do pesquisador para a análise dos sons em questão.

De acordo com Catford (1977), Rousselot é um dos grandes pioneiros em análise instrumental e seu trabalho apresenta uma riqueza de informações impressionante. Contudo, no que diz respeito ao estudo auditivo, Rousselot (1892), assim como Straka, não o considera como um instrumento suficiente de análise. Para ele, o ouvido não é capaz de distinguir diferenças sonoras ao ouvir a linguagem humana.

De fato, segundo Cagliari *et alli* (2012), só um foneticista devidamente treinado consegue categorizar os sons da fala em unidades linguisticamente distintivas. Uma pessoa sem um treinamento especial simplesmente entende a fala como um contínuo sonoro no qual reconhece uma realização de sua língua. Conforme os autores, caso aconteça de o falante ouvir sons que não batem com o sistema de sua língua, poderá achar que está diante de uma língua estrangeira ou de sons sem sentido. Se lhe é difícil descrever os sons da fala em unidades, por outro lado, lhe é muito fácil reconhecer que está diante de uma realização sonora que não pertence à língua que ele conhece como usuário falante. Como exemplo, Cagliari *et alli* (2012) dizem que alguém pode dizer a palavra *cama* sem nasalizar as vogais, mas centralizando a primeira vogal e não a segunda. Um usuário não linguista pode achar que a vogal foi nasalizada, simplesmente porque uma foi falada de modo diferente da outra. Isso acontece, segundo eles, porque a nasalização é esperada nesse contexto e uma pronúncia sem nasalização e sem centralização seria estranha na língua.

Os autores continuam e dizem que, como os falantes e ouvintes se guiam pela semântica e pela sintaxe, a observação detalhada da fonética é muito difícil. Os falantes reagem mais em função da semântica do que de qualquer outro aspecto da língua. Esse fato, segundo eles, mostra a grande diferença entre a interpretação que um linguista faz da língua e a de um falante comum. De acordo com Cagliari *et alli*, o primeiro a interpreta como um especialista que cria modelos teóricos para explicar a realidade. Já o falante comum, não sabe o que é uma oclusiva palatoalveolar surda, porque é um termo técnico descritivo, previsto pela teoria. Na verdade, eles dizem, o falante nem sabe como mexe os órgãos fonadores quando fala. Ele simplesmente vai saber que se pode dizer [pɔʧɪ] (*pote*), mas não pode dizer [paʧʊ] (*pato*).

Portanto, conforme os autores, a consciência interpretativa pertence aos cientistas e somente deles se deve exigir uma descrição adequada e completa, teoricamente consistente. Porém, quando há choques de teorias irreconciliáveis, o problema sempre estará nos argumentos de cada uma delas. Diante de um mesmo dado, de uma mesma realidade, Cagliari *et alli* consideram que as teorias podem se complementar, mas não podem se contradizer. Uma realidade não pode ser e não ser ao mesmo tempo sob um mesmo ponto de vista. Como exemplo, eles mostram que uma vogal, como unidade fonética, não pode pertencer a duas categorias diferentes ao mesmo tempo, quer do ponto de vista auditivo, quer do ponto de vista acústico, porque representa uma única realidade.

De acordo com os autores, um programa de computador, por sua vez, pode reconhecer e descrever as diferenças, mas só pode fazer isso com referência a uma análise linguística que foi, primeiramente, descrita em termos auditivos, para que a comunicação entre falantes e ouvintes seja transmitida corretamente. Além disso, eles afirmam, a análise acústica não acrescenta nada de novo à análise auditiva. Mas, a análise auditiva, por outro lado, segundo os autores, pode reconhecer, na análise acústica, detalhes e informações físicas sofisticadas (e.g. duração) que permitem uma análise linguística mais aperfeiçoada. Para Cagliari *et alli* (2012), a perfeição está nos detalhes e não na destruição de parte do sistema da língua.

Cagliari *et alli* (2012) também falam que não é da natureza acústica da fala re-interpretar o sistema de uma língua. Somente a análise auditiva pode fazer isso. O máximo que a análise acústica faz, eles dizem, é sinalizar para algo, a respeito do qual o falante ainda não tomou consciência. A partir daí, pode até reconstruir parte do sistema.

> Por exemplo, uma análise sofisticada da qualidade vocálica leva a acústica a apresentar variações muito sutis, permitindo o estabelecimento de muitas categorias acústicas para realidades fonéticas que, tradicionalmente, são agrupadas em uma ou outra apenas.
> (CAGLIARI *et alli*, 2012, p. 17).

Essas variantes, na visão dos autores, podem ser relevantes para modelos acústicos da fala, como, por exemplo, na produção de fala sintética ou no reconhecimento automático. Mas, linguisticamente, elas são irrelevantes.

> A análise auditiva, porque orientada pela fonologia, interpreta a realidade em um certo nível de abstração (a percepção de muitos casos é agrupada em uma mesma categoria funcional). Dessa maneira, há filtros acústicos, assim como há filtros auditivos, para se chegar às unidades significativas da fala e, por consequência, da língua como um todo.
> (CAGLIARI *et alli*, 2012, p. 17).

Por isso, não podemos dizer que a análise acústica dispensa a análise auditiva, pois,

> *[...] even with simpler, speedier, cheaper, more sensitive and reliable instruments to analyse a corpus of speech than exist at present, the results would still be of limited value in trying to reach any understanding of the meaning of such vocal effects as intonation and other prosodic features when perceived by the listener, for the obvious reason that the instrumental analyses produce pictures of speech which are too sensitive to detail to provide any clear pattern. The physical correlates of those features, or accumulations of features, which are of linguistic significance for the native speaker, are obscured by the presence of a large amount of accompanying but less relevant (or irrelevant) phenomena*[45]
> (CRYSTAL, 1969, p. 3).

Por outro lado, Cagliari *et all* (2012) acreditam que é possível dizer que a análise auditiva dispensa a análise acústica. Afinal, de acordo com eles, os falantes não têm embutido na mente um espectrógrafo, nem um programa como o PRAAT. A audição é processada biologicamente, por meio da faculdade da linguagem. O aspecto acústico funciona apenas como mecanismo de transmissão de uma fonte (falante) para um receptor (ouvinte).

De acordo com os autores, é certo que a transmissão precisa ser exata para quem a produz e para quem a recebe, caso contrário, quem fala não diz o que devia e quem ouve fica privado de informações necessárias. Portanto, não é o aspecto acústico da fala que define a fala linguisticamente, mas os mecanismos de produção e de recepção da fala, que são, pela sua natureza, biológicos e não acústicos.

Aqueles que acreditam que a análise acústica é autossuficiente, dizem os autores, incorrem no erro de analisar apenas o aspecto físico dessa realidade, porque toda atribuição de segmentos acústicos a uma unidade fonética (vogal, consoante, entoação, entre outros) só pode ser feita se o sistema da língua estiver avaliando, justificando ou negando as análises feitas. Segundo eles, mesmo no caso de se operar com um reconhecimento automático da fala, o confronto com o sistema linguístico é imprescindível e, na realidade, é o regulador e avaliador das análises acústicas automáticas.

---

[45] "[...] mesmo com os mais simples, rápidos, baratos, mais sensíveis e confiáveis instrumentos do que os que existem até o momento para a análise de um *corpus* de fala, os resultados para se chegar a qualquer compreensão dos efeitos vocálicos, como a entoação e outras características prosódicas quando percebidas pelo ouvinte, ainda teriam valores limitados, pela simples razão de que as análises instrumentais produzem imagens da fala que são muito sensíveis ao detalhe para proporcionar qualquer padrão claro. Os correlatos físicos desse recurso ou acumulação dessas funções, que têm importância linguística para o falante nativo, são obscurecidos pela presença de uma grande quantidade acrescentada, mas menos relevante (ou irrelevante) de um fenômeno." [tradução nossa].

Sendo assim, de acordo com CAGLIARI *et alli* (2012), mesmo tendo, nos dias atuais, um grande número de trabalhos acústicos da fala, a decisão linguística da análise acaba caindo sempre na decisão auditiva do foneticista que atribui um símbolo (categoria) a um som. "Mesmo quando ele apenas analisa um espectrograma, a atribuição de um símbolo significa que ele acha que o que vê corresponde a algo que ele ouve. Todo foneticista sabe que não ouvimos espectrogramas, mas sons da fala." (CAGLIARI *et alli*, 2012, p. 11).

Eles afirmam que, se fosse possível ler espectrogramas diretamente, sem uma análise auditiva da gravação e atribuir a eles os segmentos fonéticos da fala, teríamos um mecanismo de reconhecimento automático de fala. Mas isso, segundo os autores, ainda não foi encontrado. Eles afirmam que, no máximo, diante de um espectrograma, podemos identificar alguns tipos de segmentos, pela simples observação visual. Todavia, essas propriedades acústicas não garantem o reconhecimento automático da gravação que o espectrograma reproduz.

Tendo em vista todas essas considerações, entendemos que a melhor saída ainda está em se fazer uma união entre os dados acústicos e os dados auditivos. Não faz sentido desenvolver um estudo que englobe somente uma análise acústica ou somente uma análise auditiva, já que essas duas abordagens são complementares.

3.3.4 Alguns pontos de vista

O reconhecimento da análise auditiva como condição para uma interpretação acústica, segundo Cagliari *et alli* (2012), tem sido objeto de interesse de muitos foneticistas dedicados a análises acústicas. Para eles, a cooperação entre análises instrumentais e auditivas é imprescindível, mesmo que as teorias e os instrumentos acústicos sejam os mais sofisticados. Como diz F. Winckel,

> *Speech as an acoustic process can be defined as a continually changing structure of waves, comprising all frequencies and amplitudes over a certain frequency spectrum. Physically, such phenomena are called transients, and from an analysis of these made by his ear the hearer is able to identify linguistic signals*[46].
> (WINCKEL, 1968, p. 17)

Para Cruz-Ferreira (1983), é lamentável que alguns pesquisadores, devido ao uso das mais sofisticadas técnicas instrumentais, tenham deixando de lado o caráter perceptual da entoação, pois, segundo ela, os fenômenos linguísticos, além de serem físicos, são também

[46] A fala como um processo acústico pode ser definida como uma estrutura de ondas que muda constantemente, compreendendo todas as frequências e amplitudes de um certo espectro de frequência. Fisicamente, tal fenômeno é chamado de transiente, e a partir de uma análise dessas, feita com seu ouvido, o ouvinte é capaz de identificar os sinais linguísticos. [tradução feita a partir do artigo apresentado no GSCP 2012].

cognitivos. A autora diz que, em seu ponto de vista, a análise perceptual é um dado essencial para a análise fonética da entoação, pois a relevância linguística não é mais decidida somente no caráter físico da fala, mas também no perceptual.

Currie (1979) e Martin (1975) também concordam que a caracterização da fala não é somente uma propriedade acústica, mas também uma propriedade relacionada ao ouvinte, ou seja, estão de acordo que existe uma correlação entre o acústico e o perceptual. Sobre isso, Currie (1979) diz que nem tudo o que é medido é relevante, por isso, para diferenciar o que é irrelevante do que não é, é preciso encontrar algum tipo de equivalência entre o sinal físico e o símbolo perceptual.

Daniel Jones (1960) considera que a fonética experimental (acústica) é corroborada pelo o que o ouvido interpreta. O mesmo ponto de vista é compartilhado por Boix, ao afirmar que *"[...] cualquier buen profesional escucha atentamente a la vez que utiliza sus instrumentos habituales de análisis"*[47]. (BOIX, 1991, p. 46).

Bertil Malmberg também está de acordo com Jones e Boix. Para ele, a melhor análise linguística é aquela que parte, primeiramente, de uma interpretação auditiva, para, em seguida, juntar-se a uma análise experimental. Ele afirma:

> *It seems reasonable […] to presume […] that a combination of a strictly structural approach on the form level with an auditory based description on the substance level will be the best basis for a scientific analysis of the expression when manifested as sound. This description has to start by the functional analysis, then it must establish in auditory terms the distinctions used for separating phonemic units, and finally, by means of appropriate instruments, find out which acoustic and physiological events correspond to these different units. The interplay between the different sets of phenomena will probably for a long time remain a basic problem in phonetic research*[48] (MALMBERG, 1968, p. 15).

Ladefoged (1973) também não opta por uma análise individual, ou seja, somente por uma interpretação acústica ou auditiva. Mas, por outro lado, revela a necessidade das duas em uma análise. Pois, segundo ele, uma descrição fonética só é considerada adequada se tiver o mesmo significado para todos os que a usam. Sendo assim, quanto à importância de uma análise auditiva para a categorização dos sons, ele diz:

---

[47] "[...] qualquer bom profissional ouve atentamente, enquanto utiliza seus instrumentos habituais de análise." [tradução nossa].

[48] Parece razoável [...] supor [...] que a combinação entre uma abordagem estritamente estrutural no nível físico com uma descrição auditiva no nível da substância será a melhor base para a análise científica da expressão quando ela manifestar-se como um som. Essa descrição tem que se iniciar por uma análise funcional e só então, se deve estabelecer, em termos auditivos, as distinções usadas para a separação das unidades fonêmicas e, finalmente, por meio dos instrumentos adequados, saber quais eventos acústicos correspondem a essas diferentes unidades. A interação entre os diferentes conjuntos de fenômenos continuará, provavelmente, por muito tempo, a ser um problema básico na pesquisa fonética. [tradução nossa].

> *Understanding speech is, in essence, a process of obtaining information from an auditory stimulus. This process involves discriminating between some sounds and considering other sounds to be similar*[49].
> (LADEFOGED, 1967, p.143).

Já para a análise acústica nos estudos linguísticos, o autor justifica:

> *Furthermore, although we could (with difficulty) characterize all possible systematic phonetic contrasts entirely in physiological terms, it would be ridiculous to overlook the fact that some phonological rules obviously work in terms of acoustic properties of sounds*[50]
> (LADEFOGED, 1971, p. 4).

Cagliari *et alli* (2012) comentam que Ladefoged (1973) conduziu um importante experimento que, além de contribuir para revelar a importância de uma interpretação auditiva para a fonética, também mostrou a sua indispensabilidade para as análises acústicas. Esse método consistiu na avaliação e descrição das vogais e é nomeado de "método das vogais cardeais" (JONES, 1917; ABERCROMBIE, 1967; LADEFOGED, 1975; LAVER, 1994; CAGLIARI, 2007). Os resultados obtidos nesse trabalho confirmaram o que já havia sido previsto pelo procedimento, ou seja, que a precisão e a categorização fonética da qualidade vocálica depende, necessariamente, de um referencial auditivo.

Ball e Rahilly também declaram a importância da correlação entre as duas formas de análise. Eles consideram que a análise auditiva, assim como a análise acústica, também precisa ser avaliada, estando sujeita a provas e a avaliações independentes. Nesse sentido, eles asseveram:

> *A short cautionary note is needed, however, before we proceed (with instrumental phonetics). Students of phonetics may well develop an idea that speech description using impressionistic transcription is wholly subjective and imprecise. On the one hand, transcription if learned properly can be an objective and testable approach to describing speech sounds; on the other, the result from instrumental analyses may themselves be difficult to interpret, as it may be unclear which part of a printout refers to which section of the speech. All instruments have in-built limitations (perhaps to do with speed of data acquisition, or ability to cover all relevant parameters), so no approach to phonetic description can give all the answers, but using*

[49] Entender a fala é, essencialmente, um processo de obtenção de informação do estímulo auditivo. Esse processo envolve discriminar alguns sons e considerar outros como sendo semelhantes. [tradução feita a partir do artigo apresentado no GSCP 2012].

[50] Além disso, embora possamos (com dificuldade) caracterizar todos os contrastes fonéticos sistemáticos inteiramente em termos fisiológicos, seria ridículo desprezar o fato de que algumas regras fonológicas obviamente funcionam em termos de propriedades acústicas dos sons. [tradução feita a partir do artigo apresentado no GSCP 2012].

> *transcription and instrumentation together can get us quite a way along the road*[51].
> (BALL; RAHILLY, 1999, p. 2001).

Para Gunnar Fant (1967, p.95), conforme Cagliari *et alli* (2012), não é possível fazer uma análise acústica se não tivermos uma referência aos fatos articulatórios dos sons,

> *The rules relating speech waves to speech production are in general complex since one articulatory parameter, e.g., tongue height, affects several of the parameters of the spectrogram. Conversely, each of the parameters of the spectrogram is generally influenced by several articulatory variables. However, to establish and learn these analytical ties is by no means a hopeless undertaking.*
> *Some elementary knowledge in acoustics is valuable, but the main requirement is a sound knowledge of articulatory phonetics*.[52]

D. Fry também não dispensa a análise acústica ou a auditiva. Para ele, deve haver um compromisso entre os dois métodos, senão, os estudos da fala serão insatisfatórios. No entanto, para ele, tudo está relacionado com o funcionamento da mente:

> *Almost all linguistic behavior is organized at a brain level where the subject, speaker or listener, is unaware of what happens. We may say, indeed, that successful experiments are possible because the subject is unaware of the processes. If we devise the tasks and frame the questions carefully enough, the subject will supply information which is indispensable to an understanding of the way in which language works.*
> *It must be made quite clear that such experimental work depends on the observations that the linguist and the phonetician can make directly; without these, there would be no basis for experiment. On the other hand, linguistic theory, if it is to have any validity as a summary of linguistic behavior,*

[51] Porém, é necessária uma observação de precaução, antes de continuarmos (com a fonética instrumental). Alunos de fonética podem muito bem achar que a descrição da fala, que usa transcrição impressionística, é totalmente subjetiva e imprecisa. Por um lado, a transcrição se for aprendida adequadamente, pode ser uma abordagem objetiva e testável na descrição dos sons da fala; por outro lado, o resultado de análises instrumentais pode ser difícil de ser interpretado, porque pode não estar claro que parte do registro se refere a que parte da fala. Todos os instrumentos têm limitações em sua construção interna (talvez poderia ser feito pela velocidade de aquisição de dados ou pela capacidade de cobrir somente os parâmetros relevantes), de tal modo que nenhuma abordagem da fonética descritiva pode dar todas as respostas, mas usando a transcrição e a instrumentalização juntas, podemos nos sair bem nessa jornada. [tradução feita a partir do artigo apresentado no GSCP 2012].

[52] As regras que relacionam ondas da fala com a produção da fala são, em geral, complexas, uma vez que um parâmetro articulatório, por exemplo, altura da língua, afeta vários parâmetros do espectrograma. Por outro lado, geralmente, cada parâmetro do espectrograma é influenciado por algumas variáveis articulatórias. Contudo, estabelecer e entender esses laços analíticos é um empreendimento, sem dúvida, inalcançável. Alguns conhecimentos elementares de acústica são valiosos, mas o requisito principal é o conhecimento dos sons da fonética articulatória. [tradução feita a partir do artigo apresentado no GSCP 2012].

> *cannot afford to ignore a whole area of the subject which is revealed only by experimental methods*[53]
> (FRY, 1973, p. 86).

J. Catford (1988) também reconhece a necessidade de cooperação entre essas análises. De acordo com ele, os conhecimentos fonéticos são adquiridos por meio de uma introspecção de como os sons são produzidos. Ainda segundo Catford, não é possível aprender fonética somente fazendo uso da teoria. É preciso algo mais, isto é, uma habilidade prática que permite ao linguista distinguir e interpretar os sons.

Keith Johnson, no livro *Acoustic and auditory phonetics* (2003), considera que a análise acústica padrão só mostra verdades parciais. Para ele, o sistema auditivo é de extrema relevância, pois, além de permitir a compreensão do sinal acústico da fala, mostra características interessantes do discurso *"[...] The auditory system warps the speech signal in some very interesting ways, and if we want to understand the linguistic significance (or lack of it) of speech acoustic, we must pay attention to the auditory system."*[54] (JOHNSON, 2003, p. 1-2). Mesmo assim, para o autor, o que verdadeiramente importa é a relação entre um e outro, já que, assim como a acústica apresenta análises parciais, também não podemos atribuir a validação de uma análise ao ouvido humano, uma vez que *"The human auditory system is not a high-fidelity system. Amplitude is compressed; frequency is warped and smeared; and adjacent sounds may be smeared together."* [55] (JOHNSON, 2003, p. 46). Nesse sentido, ele diz:

> *Because listeners experience auditory objects, not acoustic records like waveforms and spectrograms, it is useful to consider the basic proprieties of auditory perception as they relate to speech acoustics*.[56]
> (JOHNSON, 2003, p. 46).

---

[53] Quase todo comportamento linguístico é organizado em nível cerebral, em que o sujeito, falante ou ouvinte não está a par do que acontece. Podemos dizer, de fato, que as experiências bem sucedidas são possíveis porque o sujeito não tem conhecimento desses processos. Se elaborarmos as tarefas e enquadrarmos as questões com cuidado suficiente, o sujeito irá fornecer as informações que são indispensáveis para o entendimento de como a linguagem funciona. Deve ficar bem claro, que o trabalho experimental depende das observações que o linguista e o foneticista podem fazer diretamente; sem isto, não haveria base para o experimento. Por outro lado, a teoria linguística, se é para ter alguma validade como representação do comportamento linguístico, não pode se permitir ignorar toda uma área do assunto, que é revelada apenas por métodos experimentais. [tradução nossa].

[54] "O sistema auditivo deforma o sinal de fala em alguns aspectos muito interessantes, e se nós quisermos entender o significado linguístico (ou a falta dele) da acústica da fala, nós devemos prestar atenção no sistema auditivo". [tradução nossa].

[55] "O sistema auditivo humano não é um sistema de alta fidelidade. A amplitude é comprimida; a frequência é distorcida e borrada; e os sons adjacentes podem ser confundidos". [tradução nossa];

[56] Devido à experiência dos ouvintes com os objetos auditivos, e não com registros acústicos como formas de onda e espectrogramas, é útil considerar a relação das propriedades básicas da percepção auditiva com a acústica da fala. [tradução nossa].

Yi Xu, no trabalho intitulado *In defense of lab speech* (2009), a fim de desmistificar o trabalho experimental, mostra que os laboratórios acústicos são indispensáveis para a compreensão dos mecanismos da linguagem humana. No entanto, segundo ele, o problema é que eles têm sido considerados como algo não natural e monótono. E, ainda, por serem mais tecnológicos, são vistos com falta de qualidade prosódica e de funções comunicativas, interativas e emotivas, o que, para ele, não se justifica. Mesmo contrário a essas opniões, o autor reconhece que o trabalho acústico deve ir além dos laboratórios, pois, ele diz, que mesmo a amostra da fala sendo a mais natural possível, o objetivo principal é sempre entender o tipo de fala que ocorre fora dos laboratórios.

Assim como Xu, foi preciso que outros foneticistas reconhecessem a necessidade das análises auditivas para resolver não só as descrições linguísticas como também suas análises acústicas. Segundo Cagliari *et alli* (2012), as pessoas que fazem análise acústica sem uma finalidade linguística exata, como os engenheiros em telecomunicação, apresentam grandes dificuldades em trabalhar, ao mesmo tempo, com dados físicos e interpretações linguísticas. Por outro lado, segundo os autores, do ponto de vista da percepção, a audição humana é altamente sofisticada. Através de testes específicos, sabe-se que o ouvido humano é capaz de distinguir diferenças de 1 hertz, com intensidade normal. (JOHNSON, 2003). Linguisticamente, a variação melódica da prosódia das línguas precisa de intervalos bem maiores.

Assim, é necessário saber lidar com realidades diferentes, como, por exemplo, reconhecer a importância da colaboração entre as análises acústicas e auditivas, mesmo que elas não apresentem uma relação amigável entre si. O cérebro humano não é capaz de distinguir certas variações físicas da frequência fundamental e os dados acústicos vêm apoiar o que o ouvido processou, mostrando detalhes que ajudam o linguista a correlacionar fatos da fala com fatos do sistema da língua, pois

> Os olhos que veem os espectrogramas são ajudados pelos ouvidos que interpretam os sons como unidades psicoacústicas, interiorizadas na mente dos falantes, como resultado de um processo de aquisição da linguagem. (CAGLIARI *et alli*, 2012, p. 11).

4 Os elementos prosódicos

4.1 A tessitura

Tradicionalmente, a tessitura é o elemento da língua que menos tem sido pesquisado. No *Dicionário de fonética e fonologia* (2000), esse elemento prosódico é definido como: "Termo da terminologia musical usado por alguns foneticistas para indicar um compasso

característico, ou variação de PITCH, da voz de uma pessoa, quando fala normalmente". (CRYSTAL, 2000, p. 254). Em outras palavras, segundo Cagliari (2002) é o intervalo entre o som mais grave e o mais agudo que uma pessoa usa quando fala, e ocorre de acordo com os ajustes da laringe que podem aumentar ou abaixar sua F0.

O autor (CAGLIARI, 1992) explica que esse elemento prosódico é o que permite reconhecermos o tom de voz de uma pessoa mesmo sem nunca tê-la visto.

> A razão disto é que as pessoas têm expectativas com relação ao tom melódico e seus valores. Se ouvirmos alguém entrando numa sala ao lado esbravejando, deduzimos isto pelo tom de sua voz (tessitura aguda, volume alto, fala mais silábica etc).
> (CAGLIARI, 1992, p. 57).

Na maior parte das vezes, os falantes usam uma tessitura constante. Entretanto, podem haver variações que, por sua vez, podem estar relacionadas com as intenções discursivas do indivíduo, "[...] ocorrendo em contextos que precisam ser destacados como mais relevantes (tessitura alta) ou como irrelevantes (tessitura baixa) ou como encaixados ou inseridos fora do lugar sintático mais típico, (tessitura baixa) [...]" (CAGLIARI, 2002, p. 28).

Conforme Cagliari e Massini-Cagliari (2003, p. 67), essas variações também estão relacionadas com os objetivos expressivos do falante, como expressões de raiva, fúria, alegria, tristeza, desespero, entre outras. A tessitura pode ser encontrada, ainda, em início de parágrafo, estabelecendo relações com o tempo da narrativa (CAGLIARI, 2002). Esse fato, também mostra que a tessitura tem um papel fundamental na determinação dos elementos prosódicos da fala, pois, além de organizar a cadeia sonora, revela a estrutura do discurso. E, por isso, a tessitura não pode ser estudada isoladamente, devendo ser levado em conta o ritmo, a entoação, a qualidade de voz e a continuidade durante a sua análise.

De acordo com Bollela (2006), a tessitura pode desempenhar quatro tipos de funções linguísticas no decorrer do discurso. São elas: função sintática, dialógica, discursiva e pragmática.

Em relação à primeira, a autora diz que a sua função mais recorrente é a de destacar os elementos que estão deslocados no texto, como por exemplo, situações em que estão presentes as conjunções adversativas, como em "A professora, porém, ignorou o acontecido." (BOLLELA, 2006, p. 119). De acordo com ela, nesses casos, a tessitura é mais grave, típica de um *creaky voice*.

A segunda função, por sua vez, diz respeito aos turnos de fala. Conforme Bollela, a tessitura mais aguda nos turnos conversacionais está relacionada àqueles momentos em que o interlocutor quer se pronunciar e pede o turno durante a fala do outro.

A terceira, discursiva, está relacionada aos elementos coesivos. Nesses casos, a tessitura mais grave é típica das digressões e a mais aguda caracteriza a retomada do assunto principal.

Por fim, a função pragmática corresponde às atitudes do falante. Para ela, nesses casos, os tons mais graves denotam mais razão e autoridade, os agudos, contestação e exaltação. Além disso, segundo Bollela, uma tessitura muito grave ou uma muito aguda seria uma forma de estratégia do interlocutor para não ser interrompido durante o discurso.

Para Cagliari e Massini-Cagliari (2003), a tessitura e a entoação são os elementos prosódicos da língua que mais se aproximam. Isso ocorre porque, segundo eles, na sua constituição, ambas se utilizam de variações na frequência fundamental do falante. Contudo, são dois elementos independentes um do outro e não podem ser confundidos, pois, dependendo da intenção do falante, um contorno entoacional pode ter uma tessitura mais alta ou mais baixa.

A tessitura "[...] focaliza as variações nos intervalos entre a frequência mais baixa (mais grave) e a mais alta (mais aguda) do indivíduo." (CAGLIARI; MASSINI-CAGLIARI, 2002, p. 2). A entoação, entretanto, se constroi a partir das variações da F0 nos limites internos da tessitura, mostrando que a mesma não pode ser confundida com os padrões entoacionais dos enunciados.

> Assim, enquanto as variações de freqüência fundamental constituem os padrões entoacionais dos enunciados, as variações de registro (tessitura) podem deslocar esses padrões para níveis mais graves ou mais agudos, mantendo intactos o "desenho" dos padrões.
> (CAGLIARI; MASSINI-CAGLIARI, 2002, p. 2).

Esse elemento prosódico, também tem a característica de não alterar a forma típica dos padrões entoacionais. Segundo os autores, ela (a tessitura) somente os desloca para cima ou para baixo de acordo com a faixa de frequência fundamental, sendo, portanto, variações diferentes das dos padrões entoacionais. Alguns padrões entoacionais, segundo Cagliari (2007) também podem evidenciar quedas ou elevações de tessitura, como é o caso do tom 6, em que seu padrão entoacional mostra uma espécie de abaixamento na tessitura.

Na maioria das vezes, quando se analisa os grupos tonais, observa-se que as variações de tessitura ocorrem em ambientes muito específicos. Segundo Cagliari (1992),

não há discurso sem tessitura, de modo que sua variação é responsável pelas diferentes significações. Assim, de acordo Cagliari e Massini-Cagliari (2003, p. 70), as quedas na tessitura ocorrem, normalmente,

> [...] em uma parte do texto em que o falante está fazendo uma digressão semântica com relação ao tópico principal, concatenando vários grupos tonais em tessitura baixa, ou querendo colocar 'entre parênteses' uma idéia, julgada menos importante.

As tessituras mais altas, segundo os autores, por sua vez, têm função de destaque, no sentido de chamar a atenção para o que se diz, dando maior ênfase para a ideia central, assim como ocorre com os vocativos ou como resultado da perda de controle (Não levante a voz para mim!).

Nos diálogos, as variações de tessitura são bem visíveis e comuns, principalmente em situações agitadas de debate ou de discussão, afirmam Cagliari e Massini-Cagliari (2003). Um exemplo claro pode ocorrer, de acordo com os autores, quando um falante eleva a tessitura em um momento de alta discussão, fazendo com que seu interlocutor, querendo rebater seus argumentos, aumente ainda mais o nível da tessitura. Uma questão pragmática, portanto. Nesses casos, a qualidade de voz característica é denominada de falseto e representa a tessitura mais alta a que um falante pode chegar. (CAGLIARI; MASSINI-CAGLIARI, 2002).

Uma questão que deve ser levada em consideração para a análise da tessitura é que a fala obtida por meio de uma leitura (caso típico dessa dissertação) é bem diferente de uma fala espontânea, seja ela acalorada ou calma, pois cada qual tem suas características próprias. Segundo Cagliari e Massini-Cagliari (2003), há leituras, como as escolares, que apresentam somente tons entoacionais primários, afirmativo, interrogativo e suspensivo, todos com variações normais, seja de tessitura, ritmo e tempo. Entretanto, as leituras interpretativas são mais rebuscadas e enfeitadas, usando recursos prosódicos mais ornamentados, a fim de se aproximar da fala espontânea.

4.2 Propriedades da dinâmica da fala

Como pudemos ver, a tessitura é um elemento prosódico da língua que não pode ser estudado sozinho, devendo ser trabalhado juntamente com a qualidade de voz, com o ritmo, com a continuidade e com a entoação. É importante salientar que Abercrombie (1967) classifica esses elementos prosódicos como propriedades da *dinâmica da fala*. Para o autor, esses elementos, além de serem os responsáveis pela particularização dos indivíduos,

também os caracterizam como grupos sociais, podendo, ainda, desempenhar funções linguísticas. Diante disso, tratamos, brevemente, desses quatro elementos da língua.

4.3 A qualidade de voz

A qualidade de voz é uma propriedade peculiar da fala, responsável pela sua individualização. De acordo com Abercrombie (1967),

> *The term 'voice quality' refers to those characteristics which are present more or less all the time that a person is talking: it is a quasi-permanent quality running through all the sound that issues from his mouth.*
> (ABERCROMBIE, 1967, p. 91)[57]

Segundo Cagliari (2007), a qualidade de voz tem origem nos aspectos anatômicos e no modo como as pessoas articulam os sons, bem como os segmentos e tipos especiais de fonação (brevemente caracterizados no item que diz respeito a 'A audição humana').

Sobre os aspectos anatômicos, Cagliari (2007, p.127) diz:

> A voz de um homem se distingue da voz de uma mulher e ambas da voz das crianças, porque eles têm aparelhos fonadores com dimensões absolutas diferentes. Diferenças anatômicas peculiares são responsáveis também pela diferenciação da voz de pessoas que têm aproximadamente o mesmo tamanho de aparelho fonador.

Essas características são responsáveis, justamente, por individualizar e distinguir a fala das pessoas. De acordo com Cristófaro Silva (2011), a qualidade de voz também é responsável por marcar variações dialetais.

Para Massini-Cagliari e Cagliari (2001), a maneira mais fácil de identificar a qualidade de voz é por meio da produção individual. Para exemplificar, os autores argumentam que algumas pessoas podem falar sons com qualidades secundárias ao deslocarem a língua em direção a um dos lugares de articulação, podendo produzir uma qualidade de voz dentalizada, palatalizada, retroflexa, nasalisada, velarizada, entre outras. De acordo com Cagliari (2007), as maiores mudanças nas qualidades de voz acontecem no nível da glote.

> Nesse caso, os falantes usam certos tipos de fonação constantemente, por exemplo, juntando continuamente à fala uma fricção glotal, uma voz tremulada (*creaky*) ou murmurada (*breathy*).
> (CAGLIARI, 2007, p. 127).

[57] O termo 'qualidade de voz' refere-se àquelas características que estão presentes, mais ou menos, durante todo o tempo em que a pessoa está falando: é uma qualidade quase permanente que passa por todos os sons que saem da sua boca. [tradução nossa].

4.4 O ritmo e a continuidade

O ritmo e a continuidade dizem respeito à variação da duração (bem como, o acento, a pausa e a velocidade de fala). Quanto ao ritmo, entende-se, de maneira geral, o modo como as línguas organizam no tempo os elementos salientes da língua, assim como os acentos e as durações silábicas. Cagliari (2007, p. 132) define o ritmo como "[...] um tipo de simetria, uma harmonia resultante de certas combinações e proporções regulares." Para Abercrombie (1967, p. 96), *"Rhythm, in speech as in other human activities, arises out of the periodic recurrence of some sort of movement, producing an expectation that the regularity of succession will continue"*[58].

Para Cagliari (2007), o ritmo está intimamente ligado ao tempo e à duração. O autor também diz que não existe um parâmetro específico para a geração do ritmo na fala. "Na realidade, o ritmo da fala é manifestado por todos os elementos que, na dinâmica da fala, apresentam momentos de saliência e momentos de redução." (CAGLIARI, 2007, p. 133). As línguas, portanto, irão variar conforme o modo como organizam o ritmo nos grupos tonais. Sendo assim, segundo o autor, elas podem ser de ritmo silábico, ou acentual. No primeiro caso, a língua "[...] pode ter todas as sílabas com uma duração aproximadamente igual, como por exemplo o francês, o japonês, etc". (CAGLIARI, 2007, p. 134). No segundo caso, as línguas se caracterizam "[...] por terem as sílabas acentuadas ocorrendo em intervalos de duração aproximadamente iguais ou isocrônicos". (CAGLIARI, 2007, p. 134), o que ocorre, por exemplo, com o português e inglês.

Na definição de Abercrombie, a *continuidade*[59] *"[...] refers to the incidence of pauses in the stream of speech – where they come, how frequent they are, and how long they are"*[60]. (ABERCROMBIE, 1964, p. 96).

Segundo Bolella (2011), esse elemento prosódico, por meio da junção das palavras, define o modo como as pausas ocorrem em um enunciado, além de ser o responsável pela reestruturação da produção de fala, concatenando-a "[...] em sintagmas de um jeito e não de outro." (BOLLELA, 2006, p. 122). Por isso, por estar relacionado à incidência de pausas em um enunciado, a continuidade está intimamente ligada ao tempo do discurso.

De acordo com Cagliari, "[...] as pausas têm lugares certos para ocorrer e quando o falante não as usa adequadamente, a fala se torna truncada, geralmente com consequências

---

[58] "Ritmo, na fala, bem como em outras atividades humanas, surge da recorrência periódica de algum tipo de movimento, produzindo uma expectativa de que a regularidade da sucessão continuará". [tradução nossa].
[59] É válido salientar que esse elemento também é designado por Abercrombie por *fluência*.
[60] "[...] refere-se à incidência de pausas no fluxo de fala – de onde elas vêm, quão frequentes elas são, e quão longas elas são [tradução nossa].

de uma desorganização na produção da estrutura fonética do enunciado." (CAGLIARI, 2007, p. 128).

O ritmo e a continuidade, por serem propriedades da dinâmica da voz além de particularizarem os grupos sociais, permitem que o falante tenha controle sobre eles, podendo, até mesmo, serem adquiridos e copiados por outras pessoas (ABERCROMBIE, 1964; CAGLIARI, 2007).

4.5 A entoação

A entoação, assim como a tessitura, está relacionada à variação de altura melódica. Ela também diz respeito a variações na frequência fundamental, nas pausas, no tempo, no volume, entre outros. Para Cristófaro Silva (2011), esse elemento prosódico ou suprassegmental caracteriza a fronteira entre os constituintes sintáticos, podendo indicar o tipo de enunciado (se é declarativo, afirmativo, interrogativo etc) e, ainda, expressar as atitudes do falante (alegria, raiva, tristeza, ironia etc). Para Abercrombie (1967), são essas as características que levam a entoação a ser, provavelmente, um dos elementos mais importantes da fala, já que o fluxo conversacional depende tanto das flutuações tonais quanto dos indicadores afetivos que, consequentemente, acarretam mudanças na tessitura. Além disso, segundo o autor, a entoação caracteriza-se por estar presente na fala de todas as pessoas. *"The pitch of the voice continually fluctuates while we are talking [...] found in the speech of all communities."*[61] (ABERCROMBIE, 1967, p. 102).

4.5.1 As funções da entoação

> *"If a man's tone of voice belies his words, we immediately assume that the intonation more faithfully reflects his true linguistic intentions".*[62]
> (PIKE, 1945, p. 105)

A entoação pode desempenhar diversas funções, que são resultado da quantidade de vibrações por segundo das cordas vocais, cujo efeito altera a altura melódica dos sons. São justamente essas alterações que levam às discussões em torno do significado dos enunciados.

De acordo com Souza (2007), essa variação na entoação leva à ocorrência de diferentes padrões prosódicos, permitindo que falante e ouvinte expressem e decodifiquem o que se quer dizer com determinada entoação. Cagliari (1992, p. 47) complementa: "A

[61] "Enquanto falamos, a entoação da voz está em constante variação [...] sendo encontrada na fala de todas as comunidades". [tradução nossa].
[62] "Se o tom de voz de um homem contradiz suas palavras, nós assumimos imediatamente que a entoação reflete mais fielmente suas verdadeiras intenções linguísticas". [tradução nossa].

prosódia comanda o valor que se deve atribuir a diferentes elementos do discurso, dentro do contexto desse próprio discurso, dando, por assim dizer, uma chave de interpretação".

De acordo com o autor, a entoação pode estar presente até mesmo em um texto escrito, pois,

> Mesmo quando uma pessoa vê o texto escrito, para entendê-lo necessita, obrigatoriamente, de decodificá-lo e processá-lo como se fosse dizê-lo espontaneamente. E somente assim consegue dispor de todos os elementos necessários para ler algo e entender.
> (CAGLIARI, 1992, p. 47).

Segundo Córdula (2011, p. 2733), a entoação também é o elemento que "[...] exerce funções semânticas (conotações, subentendidos) e pragmáticas não exercidas pelos outros elementos, tom e tessitura no Português Brasileiro".

Nesse sentido, classificamos a entoação a partir de três funções: 1) Semântica, 2) Sintática e 3) Fonológica.

Funções semântica e sintática

As funções semântica e sintática podem ser analisadas concomitante, já que, segundo Cagliari (2007), as atitudes do falante expressas pela entoação devem ser vinculadas à função sintática e deixá-las "[...] de lado na descrição linguística [...] é negar à linguagem oral parte integrante da semântica da linguagem". (CAGLIARI, 1992, p. 48).

Cagliari (1992a) complementa e diz que uma das funções dos supra-segmentos[63], nesse caso, a entoação,

> [...] depende do significado a que eles estão servindo. Assim de um modo geral pode-se dizer que há um significado estrutural ('sintático') e um significado interpretativo ('semântico'). A distinção entre ambos fica relativamente fácil, quando se define significado interpretativo como sendo tudo aquilo que traz consigo uma referência ao falante, ou seja, representa a atitude do falante.
> (CAGLIARI, 1992a, p. 138).

Segundo o autor, o significado estrutural mostra a sistematização da língua, como pode ser visto pelas pausas e pelos padrões entoacionais. Quando acontece isso, fica "[...]

[63] Cagliari (1992a) caracteriza os elementos supra-segmentais em dois tipos: 1) os que são responsáveis por modificar os segmentos, entre eles, o autor cita: a labialização, a palatalização e a nasalização. São os "[...] elementos tidos como portadores de uma articulação secundária". (CAGLIARI, 1992a, p. 137), e 2) os elementos que caracterizam unidades maiores do que os segmentos. De acordo com o autor, esse segundo tipo é dividido entre os a) elementos supra-segmentais e os b) elementos prosódicos. Os elementos supra-segmentais são agrupados por ele em "Elementos da melodia da fala: tom, entoação, tessitura; elementos da dinâmica da fala: duração, mora, pausa, tempo, acento, ritmo, ársis/tesis; elementos da qualidade da voz: volume, registro, qualidade da voz". (CAGLIARI, 1992a, p. 137). Os elementos prosódicos não são mencionados pelo autor. Nesta dissertação, não caracterizamos todos os elementos supra-segmentais, mas somente os elementos da melodia da fala e a qualidade da voz. Mas, para saber mais sobre eles, veja o artigo completo de Cagliari: "Prosódia: algumas funções dos supra-segmentos" de 1992.

fácil reconhecer a função dos elementos supra-segmentais prosódicos, como também é fácil decidir como realizar foneticamente através desses elementos uma estrutura desejada". (CAGLIARI, 1992a, p. 149). Por outro lado, segundo ele, isso não ocorre com a função interpretativa (semântica e pragmática), uma vez que

> Dados os contextos linguísticos e de uso da linguagem, é relativamente fácil prever o valor linguístico dos elementos supra-segmentais prosódicos envolvidos na fala real. Porém, ao organizar o que vai falar, o falante tem diante de si, em geral, várias opções para realizar através destes elementos um determinado efeito semântico.
> (CAGLIARI, 1992a, p. 149).

Sendo assim, a função semântica é responsável por marcar as atitudes ou emoções do falante. Conforme Cagliari (1992a), essa função, como também a pragmática, é responsável por caracterizar o sujeito e permitir suas interpretações próprias. Isso, na visão dele, é o que ocorre justamente com as atitudes do falante e com os turnos conversacionais. A função sintática, por outro lado, além de diminuir a ambiguidade entre as sentenças, permite a diferenciação entre frases declarativas, interrogativas, suspensivas, entre outras.

Segundo Gebara (1976), a teoria dos atos de fala foi introduzida na filosofia da linguagem pelos filósofos analíticos de Oxford e tem como maiores expoentes Austin (1970), Searle (1972) e Lyons (1973).

Segundo ela, um dos pressupostos dessa teoria é reconhecer uma dimensão social e interpessoal ao fato linguístico, já que, nesse, a transmissão das informações não tem um fim em si mesmas. Isto é,

> Mesmo quando o falante 'comunica' algo linguisticamente, esta comunicação não é neutra, pois não constitui uma simples transmissão de informações, mas é um modo pelo qual o falante tenta influenciar o ouvinte, quer chamando a atenção deste para um determinado universo de discurso – que é seu, do falante -, quer fazendo com que o ouvinte tome uma atitude frente a algo novo que se lhe apresenta, ou ainda marcando linguisticamente o assunto e os locutores participantes do discurso.
> (GEBARA, 1976, p. 69).

Portanto, de acordo com a autora, produzir um enunciado é estabelecer uma interação tanto linguística quanto social entre falante e ouvinte. Assim, ela diz que a compreensão das funções linguísticas só é possível quando se leva em consideração a interação entre "[...] falante-ouvinte-contexto, a atividade linguística, a criatividade individual e as intenções do falante no uso da linguagem". (GEBARA, 1976, p. 70).

Dentro da teoria dos atos de fala, insere-se o que Austin (apud GEBARA, 1976) nomeia de "enunciados performativos", responsáveis por realizar um ato ilocucional, isto é, "[...] a prática de um ato através do uso da linguagem". (GEBARA, 1976, p. 70).

Os atos de fala ou atos ilocucionários (entoação, altura de voz, etc) são entendidos por Searle (1984) como uma forma de comportamento que adotamos quando falamos, constituindo a unidade básica da entoação, sendo a força ilocutória responsável por fornecer ao enunciado o seu valor de ato de fala. De acordo com o autor, eles estão relacionados ao domínio da frase, pois

> [...] toda a frase dotada de significação pode, em virtude da sua própria significação, ser utilizada para executar um ou uma série de actos de fala particulares, e desde que todo o acto de fala realizável pode, em princípio, receber uma formulação exacta no interior de uma ou de várias frases (admitindo que a situação permita), segue-se, então, que o estudo da significação das frases e o estudo dos actos de fala não constituem dois domínios independentes, mas um único apenas visto sob dois aspectos diferentes.
> (SEARLE, 1984, p. 29).

Cook (1968, p. xvii) também diz que os atos de fala dependem do contexto em que estão inseridos, pois *"A remark that sounds friendly in one situation may sound rude in another; the way of being polite to one's wife is different from the way of being polite to one's employer"*[64]. Em outras palavras, devemos considerar, em uma situação de diálogo, o tipo de falante e a relação que os envolve, uma vez que esses fatos interferem na interpretação do ouvinte.

Os atos ilocucionais básicos são, segundo Gebara (1976): afirmar, perguntar e ordenar (ou pedir). Eles se enquadram, segundo ela, nas categorias de enunciado, e podem ser declarativos, interrogativos e imperativos, manifestando-se no plano da expressão a partir de características entoacionais específicas de cada um deles. Além das expressões mencionadas pela autora para designar os atos de fala, Searle (1984) também acrescenta outras, tais como: *descrever, advertir, observar, comentar, comandar, pedir, criticar, desculpar-se, censurar, prometer, saudar, exigir,* entre outras. Segundo ele, essas expressões, além de serem de suma importância para o entendimento de uma determinada frase, podem trazer ao ouvinte diferentes consequências como: *assustá-lo, persuadi-lo, convencê-lo, alarmá-lo, esclarecê-lo,* etc.

[64] "Uma observação que soa amigável em um contexto pode soar de modo rude em outro; o modo de ser educado com a esposa de alguém é diferente do modo de ser educado com o empregado de alguém". [tradução nossa].

De acordo com Gebara, a entoação típica de uma frase declarativa neutra apresenta as seguintes características:

"- tonicidade neutra: o elemento tônico incide sobre o último item lexical do grupo tonal;

- tom descendente; a direção da curva de altura tem âmbito médio-baixo;" (GEBARA, 1976, p. 75).

Contudo, segundo ela, algumas sentenças declarativas podem ter esse padrão, mas não expressar um enunciado desse tipo, como, por exemplo, 1) os enunciados interrogativos indiretos, com forma gramatical declarativa, mas com sentido interrogativo, como "Eu quero saber se a terra é um planeta". (GEBARA, 1976, p. 76), e as 2) sentenças declarativas interpretadas como uma ordem: "Você vai comigo. (e isto é uma ordem)" (GEBARA, 1976, p. 76).

Gebara classifica os enunciados imperativos entre ordem e pedido. Conforme ela, "[...] são enunciados que impõem ou propõem um certo curso de ação ou padrão de comportamento e indicam que ele deve ser seguido". (GEBARA, 1976, p. 77). Esses atos ilocucionais, de acordo com a autora, estão vinculados aos mandos e caracterizam-se, de uma maneira geral, por terem um modo e flexão verbal específicos e um sujeito na segunda pessoa.

Por fim, os enunciados interrogativos são definidos por Gebara no plano do conteúdo como uma codificação do ato ilocucional de perguntar. No plano da expressão, ela os divide entre polares e não-polares[65]. Os enunciados polares referem-se à resposta do tipo sim/não, e são caracterizados por uma sequência segmental em PB igual ao dos enunciados declarativos, divergindo deles apenas no nível supra-segmental. Entoacionalmente, ela o caracteriza por ter um tom ascendente e uma tonicidade neutra. Os não-polares, por sua vez, aparecem com palavras interrogativas em posição temática ou em posição de final de enunciado. Entoacionalmente, Gebara diz que esses enunciados podem manifestar-se de dois modos. No primeiro modo, a tonicidade é marcada, a proeminência (foco) está na palavra interrogativa, o tom é descendente e a altura inicial é alta com tendência terminal baixa, como, por exemplo, "Quem vai comprar carne" (GEBARA, 1976, p. 95). Já para o segundo modo, há dois elementos proeminentes no GT, sendo o primeiro em posição temática e o segundo na última sílaba tônica do GT. O tom caracteriza-se por ser descendente-ascendente, de modo que o aspecto descendente é determinado pelo primeiro elemento

[65] Para mais informações sobre essas duas divisões dos enunciados interrogativos, veja Gebara (1976, p.88-98).

tônico e o ascendente pelo segundo, como em "Por onde você entrou" (GEBARA, 1976, p. 95).

As caracterizações entoacionais apresentadas por Gebara para a descrição dos enunciados (declarativos, imperativos e interrogativos) não são conclusivas, mas, por outro lado, têm "[...] o caráter de uma primeira abordagem aproximativa do complexo problema da intonação numa língua como o Português, e o de uma reflexão sobre este problema". (GEBARA, 1976, p. 105).

Cagliari (2007), por sua vez, diz que, para o Português e para muitas línguas, é válido afirmar que os contornos descendentes indicam algo com a conotação de certeza e os contornos ascendentes, algo com a conotação de incerteza. Como exemplo, ele cita o tom 4 (descendente-ascendente) que indica "[...] uma mudança de ideia do falante no meio do enunciado e é usado para orações declarativas com reserva, para exprimir restrições do tipo 'mas', condições, etc" (CAGLIARI, 2007, p. 173). Por outro lado, o tom 5, oposto ao tom 4, indica uma dúvida e o tom 6, por ser nivelado, é utilizado para indicar suspense.

Para mostrar essa relação sintática da entoação, Cagliari (2007, p. 174) traz dois exemplos desses tons, lembrando a todo momento que devemos sempre levar em conta a atitude do falante. Veja os exemplos:

(91)[66] //4 Ele/ vem a/qui// (acho que ele não vem aqui; é de se admirar que ele venha).

(92) //5 Ele /vem a/qui// (não duvide que ele venha, porque ele vem mesmo aqui).

Outro aspecto que também deve ser levado em consideração, de acordo com Mattoso Camara (1980), para o entendimento dos enunciados é a ordem vocabular e determinadas partículas, como *que*, *quão*, *como*, *quanto*, entre outras. Segundo ele, essas partículas também podem ser consideradas marcadores entoacionais de um enunciado, pois exercem funções diferentes quando inseridas em uma frase. Nesse sentido, mesmo que em um contexto individual os atos de fala sejam suficientes para a sua compreensão, os componentes morfológicos, sintáticos, lexicais e prosódicos também podem ser considerados marcadores entoacionais.

Portanto, função sintática está correlacionada também com o papel *enfático* da entoação pois, ao oferecer reforço ao enunciado, proporciona um destaque ao que se

[66] A numeração dos dois exemplos é a utilizada pelo autor.

pretende enfatizar. Por exemplo, um tom baixo, em geral, expressa uma significação conotativa mais intensificada do que um tom alto.

A função sintática e semântica, bem como a pragmática, segundo Crystal (1969), em suma, estruturam o enunciado, construindo o seu sentido referencial e manifestando as relações entre os signos e seus intérpretes.

A função fonológica

Outro papel desempenhado pela entoação é o fonológico**,** compreendendo o foco e o acento frasal. O foco indica qual é o elemento mais importante da frase. Com ele divide-se o GT em componente pretônico e componente tônico. Antes do foco, vem o elemento semântico dado ou tema e, a partir do foco, vem o elemento semântico novo ou rema (HALLIDAY, 1970, 1974), como já destacado. De acordo com Pereira (2009, p. 29), baseada em Couper-Kuhlen (1986), a estrutura informacional da frase e seu grupo tonal podem-se definir por meio do elemento dado e do elemento novo.

Em relação ao acento frasal, Gimson (1962, p. 243) nos diz que as mudanças entoacionais estão relacionadas a ele e dependem do contexto em que estão inseridos. O acento frasal é considerado por ele como um dos aspectos mais eficientes na interpretação de um enunciado para *"[...] a listener those parts of a utterance on which the speaker wishes to concentrate attention*[67]*"*.

As funções entoacionais, portanto, permitem relacionar a linguagem oral com o mundo e com a vida de diferentes maneiras, possibilitando ao ouvinte perceber as intenções do falante e o sentido lexical dos enunciados. Além disso, analisar a intenção do falante é um dos fatores essenciais para o estudo da entoação, uma vez que, ao dizer algo ou, até mesmo, fazer uma leitura, ele irá construir uma estrutura de frase, acrescentando aos valores sintáticos e semânticos uma forma de expressão entoacional. O mesmo acontece com os atos de fala. Isso ocorre porque a fonética e a fonologia estão diretamente ligadas à sintaxe e à semântica, de modo que uma não existe isolada da outra. A identificação das atitudes do falante é feita claramente através da análise auditiva. A análise acústica, geralmente, vem confirmar o que já foi percebido antes pelo ouvido.

Antes de darmos início a análise dos significados entoacionais, é importante fazermos algumas considerações sobre as expressões de atitude na fala. Vale ressaltar que a nossa intenção não é desenvolver uma conceituação detalhada desse elemento, mas, sim,

[67] "[...] o ouvinte que participa das partes de um enunciado em que pretende concentrar atenção. [tradução nossa].

apontar algumas de suas características principais, bem como sua importância para a prosódia.

4.6 As expressões de atitude

As atitudes da fala estão diretamente relacionadas à prosódia. Sobre isso, Antunes afirma:

> [...] a prosódia é um vetor responsável pela transmissão de informações a respeito das emoções e atitudes do falante dentro do discurso. Em outras palavras, através dos traços prosódicos presentes na fala, é possível perceber se o locutor está triste ou feliz, interessado ou desinteressado, se ele apóia ou critica seu interlocutor, etc.
> (ANTUNES, 2007, p.71).

De acordo com Antunes (2007), os primeiros trabalhos sobre o papel das emoções e atitudes da fala relacionados à prosódia datam da década de sessenta do século passado, mas são ainda pouco expressivos. Depois disso, segundo ela, essa questão ganha uma maior notabilidade e, a partir do século passado e início deste, os trabalhos ficam mais numerosos. A autora faz uma revisão cronológica desses estudos. Entre eles, ela destaca os trabalhos de Lieberman; Michael (1962), Udall (1964), Willians; Stevens (1972), Wallbott e Scherer (1986), Tench (1990), Laukkanen et alii (1995), Mozziconacci (1995), Aubergé; Grépillat; Rilliard (1997), Moraes (1998), Morlec; Bailly; Aubergé (2001), Gussenhoven (2002), Campbell (2006), entre outros.

Antunes faz uma seleção completa desses trabalhos, mas, seguindo sua classificação, nós apresentamos somente aqueles que tiveram seus resultados a partir de mudanças nos padrões da F0. Não podemos deixar de mencionar também a própria tese de Antunes (2007), que, analisando os valores da F0 e da duração, encontrou resultados positivos em relação ao papel da prosódia nas atitudes do falante. A autora trabalhou apenas com questões interrogativas e, a partir desses enunciados, fez as mudanças necessárias nos valores e movimentos da F0, a fim de diferenciar as atitudes. As atitudes analisadas em seu trabalho consistiam em saber se o enunciado interrogativo era neutro, se indicava crítica, incredulidade, dúvida, interesse, condutividade ou provocação. Posteriormente, a autora pode estabelecer as principais características prosódicas das atitudes mencionadas. Vale ressaltar que, em nosso trabalho, não faremos uma análise detalhada dos padrões acústicos (como medidas da F0, duração, intensidade, entre outros) para a caracterização das atitudes

da fala, mas, sim, uma análise mais geral, relacionando os tons, os GTs, a estrutura sintática dos enunciados e os aspectos acústicos e auditivos às atitudes do falante.

Um fato muito comum nas análises é confundir as atitudes com as emoções. Para Fónagy (1993 apud SILVA, 1999), as atitudes designam "[...] um comportamento determinado, consciente, controlado, tendo um componente moral, intelectual [...]", enquanto "[...] as emoções são descargas espontâneas de uma tensão psíquica". (FÓNAGY, 1993 p. 27 apud SILVA, 1999, p. 252). Assim, segundo ele, a cólera, a alegria, a tristeza e a angústia, por exemplo, são consideradas emoções e a ironia, a reprovação e a reserva, atitudes.

Entre os autores que distinguem as atitudes das emoções, também podemos citar, de acordo com Antunes (2007), Couper-kuhlen (1986), Wichmann (2000), Reis (2001), Alves (2001), entre outros.

Antunes (2007) parte de uma extensa revisão de literatura sobre o modo como as atitudes e as emoções têm sido conceituadas ao longo do tempo para depois chegar a sua caracterização. Dessa maneira, após ter agrupado todas as características principais no que diz respeito à emoção e à atitude, a autora chega à seguinte definição:

> [...] **atitudes** são expressões controladas pelo falante (voluntárias, cognitivas, intencionais, motivadas, mais corticais), convencionadas (dependentes do sistema lingüístico, e por isso aprendidas) e que não teriam conseqüências tão evidentes na prosódia (falar com uma atitude qualquer não mudaria de forma geral a prosódia da frase; é necessário observar nuances prosódicas que dariam pistas sobre a atitude utilizada em um enunciado), através das quais o falante informa seu ponto de vista dentro da interação verbal, dando ao ouvinte pistas para que seu comportamento seja percebido ou inferido.
> (ANTUNES, 2007, p. 90, grifos do autor).

Assim, conforme a citação acima e as ideias de Vasconcelos *et alli* (2011), podemos resumir essa questão dizendo que as atitudes são mais controladas, ao passo que as emoções são menos controladas, mais involuntárias.

Bodolay (2009) segue o mesmo caminho de descrição dos autores acima, mas parte do esquema de um contínuo das funções da prosódia feito por Aubergé (2002) e considera que as atitudes estão inseridas em um nível intermediário, variando entre o controle involuntário (quando estão próximas das emoções) e voluntário (quando estão próximas da função linguística). Para a autora, assim como as expressões de atitude estão correlacionadas à prosódia, também estão à expressão emotiva e às expressões de função linguística (asserção, pergunta).

Além disso, a partir de alguns modelos teóricos, Bodolay classifica as atitudes dentro dos domínios prosódicos de duas maneiras. A primeira é chamada por ela de "atitude proposicional" e está ligada ao modo como o falante relaciona-se com o discurso que ele proferiu. Como exemplo dessa atitude, ela cita uma situação comunicativa em que o falante encontra-se em dúvida sobre determinado fato: "eu tranquei a porta?" (BODOLAY, 2009, p. 112). Além da dúvida, a autora acrescenta à atitude proposicional expressões de opinião e crença. A segunda classificação dada por ela é a de "interpessoal ou ilocucional" e relaciona-se ao posicionamento do falante em relação ao seu ouvinte. Como exemplo, Bodolay cita a polidez e a autoridade. A autora também lembra que é muito importante levar em consideração o contexto para a interpretação dos enunciados. No entanto, ela não considera essa divisão como um procedimento metodológico viável, uma vez que o mesmo enunciado, segundo ela, pode apresentar as duas atitudes ao mesmo tempo. De acordo com ela, algumas pistas prosódicas, como o prolongamento, alteração brusca do contorno melódico e aumento da intensidade, também devem ser consideradas indícios de atitudes.

Segundo Cagliari (1992a), há muito mais coisas nas atitudes do que podemos imaginar e que, posteriormente, vão determinar tanto a escolha dos elementos supra-segmentais quanto o uso das palavras no discurso. Em outras palavras, seria como dizer que a mesma atitude pode gerar vários significados. Assim, ele diz que "Uma ironia pode revelar uma coisa ridícula, um desrespeito, uma inverdade, etc". (CAGLIARI, 1992a, p. 150). Para ilustrar a afirmação anterior, o autor apresenta dois exemplos:

(22)[68] Não quero rir na sua cara, por isto digo que você se enganou na escolha do livro.

(23) Você pensa que 'O Banquete de Platão' é livro de culinária?
(CAGLIARI, 1992a, p. 150)

Conforme ele, os enunciados 22 e 23 permitem ao falante usar uma ironia descortês, mas, para isso, o falante tem que optar se vai deslexicalizar ou não a ideia de ironia. Se o falante opta pela primeira ação, diz ele, o enunciado típico é o 23 e, para isso, é preciso fazer uso de outros elementos supra-segmentais.

> Desse modo, o enunciado (23) será dito com um sorriso no canto da boca (contexto pragmático gerando uma certa qualidade de voz). Além disto, o falante usará outros elementos supra-segmentais prosódicos adequados para transformar uma simples pergunta em uma pergunta irônica.
> (CAGLIARI, 1992a, p. 150).

[68] A numeração dos exemplos é a utilizada pelo autor.

Segundo Cagliari, os elementos supra-segmentais mencionados para modificar a pergunta do enunciado 23 também podem ser gerados por uma modificação na tessitura, pelo destaque do nome do livro ou até mesmo por um aumento da velocidade de fala[69]. Por outro lado, se o falante não quiser realizar esse tipo de ação, ele pode optar pelo enunciado 22 e simplesmente dizê-lo como uma afirmação.

[69] Para Cagliari (1992a), a velocidade de fala também pode ser descrita como tempo. A velocidade de fala (tempo) diz respeito ao modo como aceleramos ou desaceleremos nossa fala no decorrer do discurso. Segundo ele, "Ocorrências de aceleração ou desaceleração no discurso servem para se dar maior valor a algo que se diz (desacelerando) ou preparar um argumento mais importante logo adiante com uma aceleração da fala" (CAGLIARI, 1992a, p. 143).

5 Embasamento teórico

5.1 Descrição dos trabalhos teóricos

Essa etapa da dissertação apresenta as duas principais abordagens metodológicas escolhidas para embasar o trabalho, dividindo-se em duas partes. Na primeira, fazemos uma revisão bibliográfica das duas perspectivas teóricas e, em seguida, uma comparação entre elas. Na segunda, apresentamos os métodos desenvolvidos para a análise dos dados.

5.1.2 O modelo descritivo de Halliday

M.A.K. Halliday (1970/1973) descreve a entoação a partir de um modelo de gramática funcional. Para Cagliari (2011), ler o texto de Halliday não é uma tarefa fácil, pois "[...] requer um conhecimento específico da teoria gramatical do autor e do modelo chamado de *prosodic analysis*, aplicado aos estudos fonológicos, seguindo uma tradição que vem de Firth (1957)" (CAGLIARI, 2011, p. 1). Além disso, de acordo com o autor, Halliday utiliza-se de termos que têm significados particulares em sua teoria, como, por exemplo, "níveis gramaticais", "grau de delicadeza", "expoente fonológico", "aspecto formal", entre outros. Mesmo assim, apesar dessa precisão e riqueza de detalhes, aplica-se apenas a línguas de ritmo acentual.

De acordo com Cagliari (2011), Halliday trabalhou com um *corpus* de mais de dois mil tons, tirando daí sua interpretação para os padrões fonológicos do inglês britânico. As análises acústicas desenvolvidas por Katherine Patch (cf. HALLIDAY, 1973, p. 107) também o ajudaram a evidenciar, "[...] as variações de ritmo dos pés, dependendo do número de sílabas, numa proporção de 5: 6: 7 para pés rítmicos com uma, duas e três sílabas" (CAGLIARI, 2011, p. 1). A partir disso, Halliday pode desenvolver seu sistema prosódico, descrevendo a tonicidade, os pés, a tonalidade e os tons.

Trabalhando com análise do inglês britânico, mais especificamente com a variedade do inglês RP (*received pronunciation*) do sudoeste da Inglaterra, Halliday propôs um sistema de cinco tons primários simples e dois tons compostos (também nomeados de secundários) e suas variantes. Sobre isso, ele diz:

> *Natural conversation in British Standard English may be represented as involving continuous selection from a set of five tones. These five tones constitute a phonological system, at a primary degree of delicacy, that is both chain-exhausting and choice-exhausting. That is to say, connected speech can, it is suggested here, be analyzed into an unbroken succession of tone groups each of which selects one or other of the five tones.*[70]
> (HALLIDAY, 1970, p.103)

Na citação anterior, o autor explica o porquê de sua representação fonológica ser caracterizada por um conjunto de cinco tons e menciona, ainda, a utilização de sequências contínuas de tons. Conforme Cagliari (2011), as sequências contínuas não destroem os grupos tonais, como acontece, por exemplo, quando se trabalha com um *corpus* constituído de elementos de conversas. Nas sequências contínuas, isso não ocorre porque o final de um grupo tonal é o início do outro, o autor explica. Além disso, nessas sequências, "[...] as pausas são incorporadas na descrição do ritmo, ocupando lugares silábicos nos pés ou mesmo ocupando a duração dos pés inteiros" (CAGLIARI, 2011, p. 2).

Para cumprir seus propósitos durante a análise, Halliday também diz ser importante levar em consideração os níveis sintagmáticos e paradigmáticos. Sobre isso, ele diz:

> *For purposes of analysis, the selection can be regarded as discrete on both axes, both syntagmatically and paradigmatically: we can make a good description, that is, if we postulate that each tone group begins where the previous one ends, with no overlap and no hiatus, and that each tone group can be unambiguously assigned to one tone, this assignment thereby excluding all the other tones*[71].
> (HALLIDAY, 1973, p. 103).

Segundo Cagliari (2011), a referência a esses dois eixos significa que o componente fonológico apresenta somente cinco padrões entoacionais que devem ser devidamente selecionados para fazer parte da descrição fonológica entoacional no eixo sintagmático. Essa relação de exclusão e oposição, conforme o autor, está relacionada à abordagem estruturalista funcional de Halliday.

Cada um dos cinco tons propostos por Halliday apresenta subdivisões, que são reconhecidas quando a análise vai além do primeiro grau de delicadeza. A respeito disso, Halliday (1973) diz que cada um dos tons apresentam subdivisões que devem ser

[70] A conversação natural em Inglês Britânico padrão deve ser representada como se envolvesse a seleção contínua de um conjunto de cinco tons. Esses cinco tons formam um sistema fonológico do primeiro grau de delicadeza, o que é ao mesmo tempo uma escolha cansativa e desgastante. Isto quer dizer que a fala pode, como é sugerido aqui, ser analisada por meio de uma sucessão ininterrupta de grupos tonais, cada um deles selecionando um ou outro tom entre os cinco tons. [tradução nossa].

[71] Para fins de análise, a seleção pode ser discretamente considerada em ambos os eixos, tanto sintagmaticamente quanto paradigmaticamente; nós podemos fazer uma boa descrição, isto é, se nós postularmos que cada grupo começa onde o anterior termina, sem sobreposição nem hiato e que cada grupo tonal pode ser anotado sem nenhuma ambiguidade tonal, essa marcação deve excluir, portanto, todos os outros tons. [tradução nossa].

consideradas quando a análise é feita além do primeiro grau de delicadeza. Segundo Cagliari, essas subdivisões são completamente possíveis, já que "[...] os cinco tons são níveis de altura melódica distintivos e opositivos no sistema fonológico da língua. Como qualquer fenômeno linguístico, são suscetíveis de variações" (CAGLIARI, 2011, p. 2).

O grau de delicadeza, de acordo com Halliday (1973, p. 104), *"[...] is a variable to which no theoretical limit can be set; nor is there any valid an objective means of measuring it*".[72] Mas, por outro lado, de acordo com o autor, as declarações fonológicas permitem que nos guiemos na escolha adequada de determinado grau de delicadeza, através de algumas considerações formais. Em sua descrição da entoação do inglês, essas questões formais são decididas pela gramática, como ele mesmo confirma quando diz *"[...] we let the grammar decide how delicate we should be"*[73]. (HALLIDAY, 1973, p. 104).

Halliday acrescenta que a decisão pela gramática só é possível por meio da relação entre a fonologia e a forma linguística. De acordo com ele, reconhecer a fonologia como um nível independente de descrição linguística é

> *[...] the only way in which it can provide a bridge between form and substance. In phonology we make a separate abstraction from phonic substance, and represent this in statements which show how the given language organizes its phonic resources in such a way to carry (or expound) its grammatical and lexical patterns. If we tie phonology directly to form, the result is likely to be either excessive complication or distortion.*[74]
> (HALLIDAY, 1973, p. 104).

O grau de delicadeza caracteriza-se, portanto, por ser a estrutura mais abstrata da língua em que se pode definir alguns níveis. Essa questão está mais evidente na teoria funcionalista, pois, uma vez que a fala é investigada diretamente, é imprescindível definir o grau de delicadeza.

> Se alguém grava conversas aleatórias e faz uma análise fonética detalhada pode encontrar ocorrências de regras contraditórias fonologicamente. Como o sistema não pode suportar contradições, deve-se organizar os dados em sistemas, ou seja, em graus de delicadeza.
> (CAGLIARI, 2011, p. 2).

No livro de Halliday (1973), essa discussão é mais longa do que a que foi apresentada. O autor parte do pressuposto de que a questão do grau de delicadeza foi um dos

---

[72] "[...] é uma variável para a qual nenhum limite teórico pode ser definido; nem existe qualquer meio objetivo que possa medi-la". [tradução nossa].
[73] "[...] nós deixamos que a gramática decida o quão delicados nós devemos ser". [tradução nossa].
[74] [...] é o único modo de se fornecer uma ponte entre forma e substância. Na fonologia, nós fazemos uma abstração separada da substância fônica, representando-a de forma que ela mostre o quanto uma dada língua organiza seus recursos fônicos, de modo a realizar (ou explicar) seus padrões gramaticais e lexicais. Se ligarmos a fonologia diretamente com a forma é provável que o resultado seja uma distorção ou uma complicação excessiva. [tradução nossa].

motivos para que Chomsky optasse por uma teoria dedutiva (*top-down)* e não por uma teoria indutiva (*bottom-up*). Conforme Cagliari (2011), essa discussão ajudou a esclarecer o que Chomsky chama de intuição do falante, "[...] ou seja, há um grau de delicadeza que define algumas estruturas aceitas pelo falante como 'normais' e outras como 'não pertencentes à sua língua' ou a algum outro grau de delicadeza definido sociolinguisticamente" (CAGLIARI, 2011, p. 2).

Tomando como base as considerações teóricas mostradas anteriormente, Halliday apresenta a sua abordagem para a análise do inglês britânico. Antes de darmos início à caracterização do seu modelo, citamos a formalização geral do sistema proposta por ele e que, posteriormente, será descrita nesta dissertação.

> *It can be seen, therefore, that any utterance in English has three distinct meaningful choices, or set of choices, are made which can be, and usually are, subsumed under the single heading of 'intonation'. These are: first, the distribution into tone group – the number and location of the tone group boundaries: second, the placing of the tonic syllable (in double tonic tone groups, the two tonic syllables) – the location, in each tone group, of the pretonic and tonic sections; third, the choice of primary and secondary tone I propose to call these three systems tonality, tonicity and tone*.[75]
> (HALLIDAY, 1973, p. 112).

O autor propõe, então, quatro unidades em ordem decrescente de valor: o grupo tonal, os pés, as sílabas e o fonema.

Os pés (P) são descritos de acordo com o ponto de vista de Abercrombie (1964), ou seja, como uma "[...] *unit of rhythm in English, and has a structure of two elements, 'ictus' and 'remiss'*"[76] (HALLIDAY, 1963, p. 106). O fato de o pé ser a unidade de ritmo e não a sílaba refere-se, segundo Abercrombie, ao "*stress-timing*". Um pé é marcado pelo espaço entre dois *ictus*. A unidade abaixo dos pés é formada pelas sílabas, que podem ser fracas ou fortes. A sílaba forte opera como um *ictus* e a sílaba fraca como uma *remiss*. A sílaba forte representa a tônica e a sílaba fraca a átona. Fazendo uma analogia com a música, os pés vêm representados pelo compasso e o *ictus* e a *remiss* seriam, respectivamente, a batida e o descanso.

---

[75] Pode ser visto, portanto, que qualquer enunciado em inglês tem três possibilidades distintas de escolha, ou conjunto de escolhas que podem ser feitas, e normalmente são, a partir dos seguintes pontos: distribuição em grupo tonal – o número e localização dos grupos tonais de fronteira: segundo, o lugar da sílaba tônica (dois grupos tonais, duas sílabas tônicas) – a localização, em cada grupo do componente pretônico e tônico; terceiro, a escolha do tom primário e secundário. Eu propus chamar esses três sistemas de: tonalidade, tonicidade e tom. [tradução nossa].

[76] "[...] como uma unidade de ritmo em Inglês, formada por uma estrutura de dois elementos 'ictus' e 'remis' [tradução nossa].

Todo pé apresenta, pelos menos, uma sílaba acentuada e um número relativo de sílabas fracas. Se um pé for seguido de uma pausa ou iniciar um grupo tonal, ele será considerado silencioso e caracterizado como incompleto. Por outro lado, os pés completos são aqueles que não apresentam uma sílaba átona no início de um grupo tonal.

A estrutura do grupo tonal (GT) é operada pelos pés. Assim como o pé, cada grupo tonal apresenta dois elementos em sua estrutura. O elemento tônico e pretônico. O elemento tônico é obrigatório e está presente em cada grupo tonal. Já o elemento pretônico é opcional. Segundo Halliday, a razão para o estabelecimento do componente pretônico deriva do seguinte fato:

> *[...] all primary tone contrasts are carried by the tonic, but some secondary contrasts are carried independently by an element preceding the tonic – these operate only, and always, if there is at least one strong syllable in this position*[77].
> (HALLIDAY, 1973, p. 7).

A variação melódica do componente pretônico em um grupo tonal se realiza devido às características do tom da sílaba tônica saliente (STS). Essa sílaba se individualiza pela maior variação do contorno melódico, ou seja, por apresentar a marca entoacional mais proeminente do enunciado, compreendendo a parte do enunciado que o falante julga como sendo a mais importante (o foco do GT).

Os pés são marcados por uma isocronicidade fonológica, em outras palavras, por uma regularidade nas batidas tonais, já que *"[...] there is a tendency for strong syllables to occur at roughly regular intervals of time whatever the number of weak syllables, including zero, in between"*[78]

Os pés rítmicos estão relacionados com o componente gramatical. De acordo com Cagliari (2011), Halliday caracteriza essa relação por meio de uma organização sintática entre os enunciados, a partir de dois componentes: o tema e o rema; o dado e o novo. Segundo o autor, essas noções prosódicas têm objetivos informativos diferentes e apresentam características fonológicas próprias.

As sílabas constituem outra categoria de importância para a descrição desse modelo teórico, já que o ritmo é obtido através das saliências silábicas. Cada sílaba apresenta uma determinada altura melódica (*pitch*). A variação da altura melódica nos permite destacar

[77] Todos os contrastes nos tons primários são marcados pela tônica, mas alguns contrastes secundários são marcados independentemente pelo elemento que precede a tônica – ele funciona apenas, e sempre, se há pelo menos uma sílaba forte nessa posição. [tradução nossa].
[78] "[...] há uma tendência para que as sílabas fortes ocorram em intervalos de tempo mais ou menos regulares, independentemente do número de sílabas fracas, incluindo zero, entre elas." [tradução nossa].

partes do enunciado, seja com movimentos crescentes ou decrescentes, o que fica mais evidente nas sílabas tônicas. (CAGLIARI, 2002, p. 15). Além disso, esse processo contribui para a mudança de direção na última sílaba tônica do enunciado, no contorno melódico e em frases afirmativas e interrogativas. O contorno melódico, por sua vez, define os padrões entoacionais no interior de uma frase.

A altura melódica difere em relação às línguas tonais e entoacionais. Nas línguas tonais, como é o caso do chinês, cada palavra apresenta uma sílaba com altura melódica fixa, enquanto nas línguas entoacionais, como o português, "[...] diferentes tipos de enunciados carreiam padrões melódicos predeterminados pelo sistema" (MASSINI-CAGLIARI; CAGLIARI, 2001, p. 117).

Juntamente com a atribuição de valores aos pés rítmicos, Halliday descreve os padrões de tonicidade, de tonalidade e dos tons. Esses três sistemas são independentes um do outro, mas não o são do sistema rítmico.

A tonicidade está relacionada com a acentuação e é caracterizada por ser uma propriedade das sílabas e não das vogais. É responsável por definir o ritmo e marcar o ponto focal de cada unidade de atividade, relacionando-se, dessa forma, com o tema e o rema em um enunciado. A tonalidade, por sua vez, descreve os grupos tonais e suas relações sintáticas e semânticas, marcando um tipo de unidade em atividade na língua. Por fim, os tons marcam o tipo de atividade envolvida e descrevem os padrões entoacionais (padrões fonológicos e pragmáticos), caracterizando a variação melódica de um enunciado. É um padrão complexo, feito de uma posição simples entre polaridade e não polaridade, ou seja, pelo que é oposto, como, por exemplo, quando uma pergunta exige respostas do tipo sim ou não.

Em Ladd (1996), o que Halliday chama de tonicidade é parte do que o autor considera como proeminência relativa. A tonalidade, por sua vez, tem a ver com a divisão do enunciado em grupos tonais e os tons são os *tunes*.

Halliday utilizou-se de alguns símbolos para anotar a diferenciação desses três aspectos, bem como a marcação dos pés rítmicos, durante a análise de um enunciado. Cagliari (2011) nomeia essas anotações de "*devices gráficos"* e afirma que, por Halliday trabalhar somente com uma análise auditiva, é perfeitamente possível que o autor faça o uso direto da ortografia.

Diferentemente dos outros modelos que vínhamos apresentando, Halliday propõe o uso de cinco níveis de altura melódica: alto, meio-alto, médio, meio-baixo e baixo. Não há outros tons fora dessa pauta. A pauta entoacional pode ser usada tanto em análises auditivas

quanto em análises acústicas. Neste último caso, os padrões entoacionais se adaptam à tessitura. Sua transcrição simbólica pictográfica vem descrita sobre uma pauta musical. Cagliari (2007), por sua vez, em sua anotação, usa uma pauta entoacional de três níveis. Veja abaixo a pauta entoacional adaptada para o modelo de Halliday (1970) de acordo com Cagliari (1982/2007):

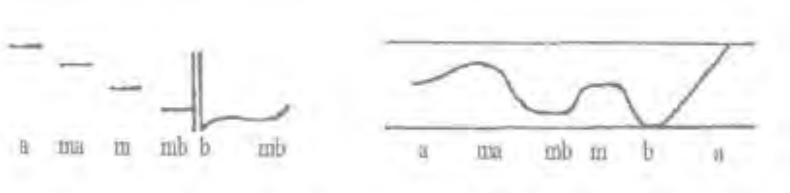
Figura 17- Pauta entoacional com os cinco níveis melódicos

_____ alto
..... meio-alto
_____ médio
..... meio-baixo
_____ baixo

Fonte: Halliday (1970) e Cagliari (1982/2007).

De acordo com Cagliari (2007), esses níveis tonais também podem ser representados pictoriamente, com referência a barras verticais ou entre linhas paralelas, como se pode ver na figura 18 abaixo.

Figura 18- Ilustração dos contornos melódicos por meio de linhas verticais e barras paralelas



Fonte: Cagliari (2007, p. 168).

Nessa figura, os níveis altos e baixos são característicos dos limites de variação melódica do falante e marcam a sua tessitura. O nível alto significa o tom mais elevado e o baixo, o tom mais inferior que o indivíduo usa quando fala. Em modo proporcional, um tom alto apresenta um maior número de ciclos por segundo do que um tom meio alto, no momento em que o falante usa uma mesma tessitura entoacional. "Em outras palavras, um tom será, por exemplo, baixo, porque, se o falante tiver que usar um tom meio-baixo no mesmo contexto, o tom meio-baixo seria mais alto do que o tom considerado baixo" (CAGLIARI, 2007 p. 168).

As anotações das variações melódicas das sílabas são assinaladas com um traço no esquema de linhas mostrado na figura 18 acima. Portanto, há apenas cinco alturas para

serem registradas. Todos os padrões entoacionais podem ser descritos com variações desse esquema de tons. Os padrões são chamados tons entoacionais ou simplesmente tons e são identificados através de números: 1, 2, 3, etc.

Diferentemente de outros modelos, o de Halliday (1970) não faz referência a tons mais baixos do que o tom baixo, nem a tons mais altos do que o tom alto. Às vezes, a variação da altura melódica nos tons alto e baixo pode apresentar um intervalo grande, mesmo sem uma influência específica da variação da tessitura. Nestes casos, o ouvido ajusta essas alturas a um valor linguístico determinado, igual ao dos tons alto ou baixo. Esse ajuste mostra a ação da interpretação fonológica sobre dados fonéticos de natureza física, detectados por meios técnicos de análise acústica.

A escolha de um tom pode relacionar-se com os tipos de oração (afirmativas, interrogativas, entre outras), com as noções de modalidade (possibilidade, probabilidade, entre outras), com os atos de fala (como ordem, pedido, etc) ou ainda, de acordo com o comportamento linguístico do falante, bem como suas atitudes (indiferença, polidez, surpresa, entre outras).

Durante a análise entoacional, não é preciso anotar todas as variações melódicas que ouvimos na fala. Segundo Cagliari (2007), o mais importante, é evidenciar se essas variações relacionam-se com as funções gramaticais ou com algumas manifestações semânticas da língua. Sendo assim, somente os pontos essenciais serão identificados. Segundo Cagliari (2007), as variações melódicas podem ser: 1) simples, de contorno descendente (D); 2) ascendentes (A); 3) niveladas (N); ou 4) complexas, quando combinam os três movimentos de contorno caracterizados. Um contorno, por sua vez, conforme o autor, pode ser: contínuo (c); por etapas (e); por saltos (s); suave (ss) ou até mesmo, brusco (bb)[79].

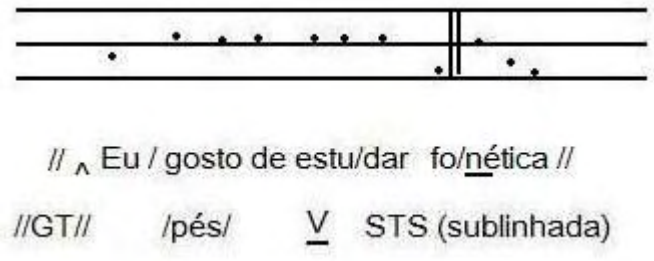
O espaço compreendido entre o som mais grave e o mais agudo, na fala de uma pessoa, é chamado de tessitura. Como acontece na música, uma melodia da fala pode continuar com a mesma curva, localizando-se, porém, em uma escala superior ou inferior. A fala costuma abranger o intervalo (tessitura) de uma oitava e meia. O ato de mudar os valores de frequência dessa escala para cima (fala aguda) ou para baixo (voz grave) em um indivíduo acarreta acréscimo de significação ao discurso. (CAGLIARI; MASSINI-

[79] Os símbolos que aparecem entre parêntesis são usados na descrição de sistemas entoacionais. As letras maiúsculas (A, D, N) descrevem os contornos entoacionais. Já as letras minúsculas (c, e, s, ss) descrevem o tipo de movimento dos contornos.

CAGLIARI, 2001, p. 120). Portanto, a mudança de tessitura não está ligada à entoação, mas ao discurso a longo prazo.

Resumindo, na teoria de Halliday (1970) e de Cagliari (1982/2007), os enunciados são divididos em grupos tonais (GT). Um GT tem sempre uma sílaba tônica saliente (STS) que o divide em componente pretônico (CPT) e componente tônico (CT). O GT inicia-se na STS. Cada componente é dividido em pés (P), começando na sílaba tônica interna dos componentes. Veja o exemplo a seguir:

Figura 19- Exemplo de análise a partir do modelo de Halliday (1970) e Cagliari (2007), com caracterização dos símbolos convencionalmente usados



Fonte: própria.

No exemplo acima, a pauta melódica segue o modelo de três linhas proposto por Cagliari (1982/2007). Os pés rítmicos estão separados por barras inclinadas e são anotados a partir da sílaba tônica. A sílaba tônica saliente (STS) que marca o foco e separa o enunciado em componente pretônico (CPT) e tônico (CT) vem simbolizada por uma barra que a sublinha.

A sílaba tônica pode recair, de acordo com a pronúncia do falante, em qualquer palavra do grupo tonal (GT). A partir disso, pode-se saber se o enunciado apresenta características de asserção, de exclamação, de interrogação, de declaração, entre outras. A STS, além de definir o GT, também contribui para a definição dos tons secundários, variantes dos tons básicos ou primários (CAGLIARI, 2007, p.174), de acordo com o modelo descritivo entoacional utilizado.

A separação entre CPT e CT é simbolizada por duas barras verticais duplas e vem anotada na pauta entoacional. Há a opção de simbolizá-las, também, antes da STS, mas, para maior clareza, escolhemos a primeira opção. A pausa está marcada por um (v) de cabeça para baixo (^) [80].

[80] Esse exemplo de notação é o que será utilizado em nossas análises.

Os grupos tonais podem ser simples ou compostos. Um GT é considerado como simples quando apresenta somente uma sílaba tônica saliente. Nesse caso, ocorre apenas uma mudança de contorno melódico notável (CMN). Um contorno melódico notável, por sua vez, é definido como um padrão do tipo que caracteriza um tom no sistema. Veja um exemplo de GT simples, a partir do enunciado abaixo (CAGLIARI, 2007, p. 165):

(a) //Eu não /<u>a</u>cho que/ seja as/sim// (falando com certeza) [81].

Por outro lado, um GT composto acontece quando há duas sílabas tônicas salientes, apresentando, portanto, duas mudanças notáveis de contorno melódico, como na frase a seguir, também retirada de Cagliari (2007, p. 165).

(b) //Eu não/<u>a</u>cho que/seja as/<u>sim</u>// (falando com hesitação, dúvida).

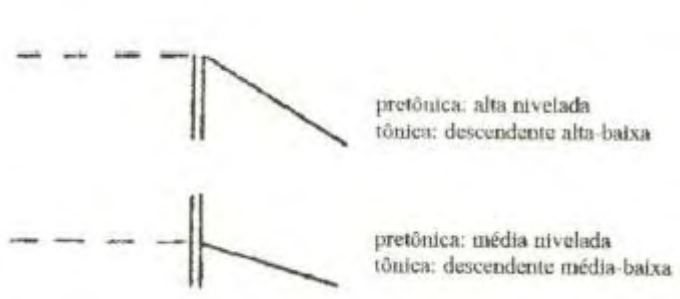
Quando ocorre um GT com tom composto, no caso do Português (e também do inglês), a segunda STS tem sempre um CMN do tipo do tom 3.

Fora da pauta melódica mostrada como esquema tonal, (local da STS), o tom é representado com duas barras verticais || indicando a

> [...] configuração do contorno melódico do componente pretônico, assinalada com pequenos traços, e seguidas do desenho do contorno melódico da sílaba tônica saliente e da parte restante do componente tônico. (CAGLIARI, 2007, p. 169).

É o que podemos observar na figura abaixo, que mostra o tom 1 normal que vai de médio a baixo e sua variante alta, que vai de alto a baixo, com sílaba tônica saliente alta:

Figura 20- Representação gráfica do padrão entoacional dos tons



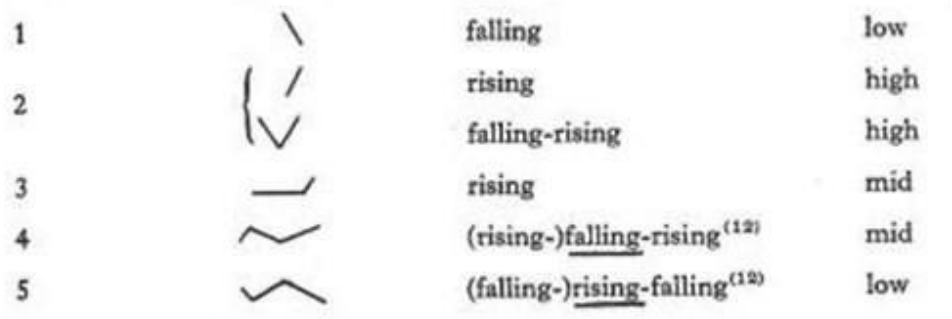
Fonte: Cagliari (2007, p. 169).

[81] As sílabas tônicas que representam a mudança no CMN aparecem sublinhadas nos exemplos (a) e (b).

Além de simples e compostos, os tons também podem ser primários ou secundários. Os tons primários são próprios de uma enunciação neutra, diferenciando-se entre si por meio do contorno melódico (CMN) que o GT apresenta no componente tônico.

No sistema entoacional proposto para o inglês britânico, os tons primários são caracterizados por Halliday da seguinte forma:

Figura 21- Representação dos tons primários simples para o inglês britânico

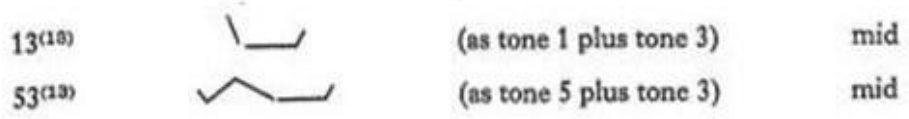


Fonte: Halliday (1973, p. 110).

Da esquerda para a direita, a primeira coluna, em números, indica os tipos de tons. A segunda coluna mostra a configuração visual de cada um dos tons, com seu respectivo contorno melódico, descrito na terceira coluna. Por fim, a última coluna caracteriza a tendência final de cada tom.

Já para os tons primários compostos, o autor encontrou apenas dois:

Figura 22- Representação dos tons primários compostos 13 e 53



Fonte: Halliday (1973, p. 110)

Essa figura tem a mesma disposição da anterior, porém, a terceira coluna deve ser lida como "tom 1 mais tom 3", uma vez que um GT desse tipo apresenta duas sílabas tônicas salientes, sendo a segunda subordinada à primeira, uma do tipo 1 e outra do tipo 3, no caso de um tom composto 13. Nas palavras de Halliday (1970, p. 43), esses dois tipos de tons são caracterizados da seguinte forma:

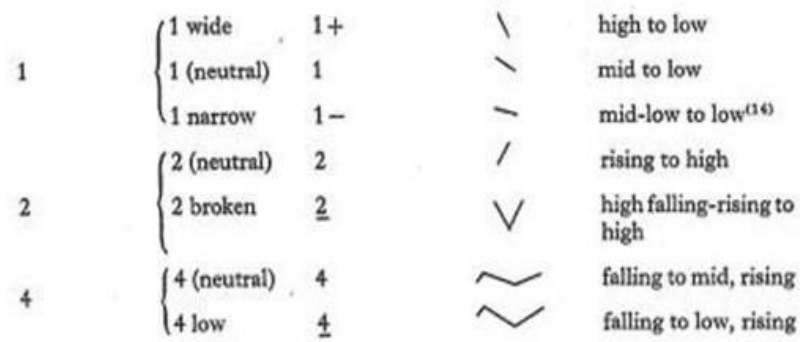
> *Tone groups with [...] tone 13 and 53 have the same general pattern of information as other tone groups, with the tonic marking the new*

> *information. The difference is merely that there are two places where the speaker has decided to focus the information in the information unit, instead of one.*
> *The two tonics are unequal in value. The first tonic, with tone 1 or tone 5, is the 'major'one; the second, with tone 3, is the 'minor' one. The major tonic carries the principal new information in the tone group. The second tonic express information which is in some way secondary or subsidiary to it*[82]

Os tons secundários, em oposição, indicam o uso marcado de um tom, ou seja, trazem consigo uma conotação semântica mais forte do que o tom primário. Os tons secundários são simbolizados pelo número do tom primário correspondente, mais um diacrítico do tipo –, +, entre outros. Em geral, eles acrescentam uma ideia a mais em relação ao significado.

O componente pretônico só é significativo nos tons secundários, enquanto nos primários apresentam-se como fixos e previsíveis. Isso ocorre porque há variações melódicas no tom secundário, tanto no CPT quanto no CT. De acordo com o modelo de Halliday, os tons secundários são reconhecidos da seguinte forma:

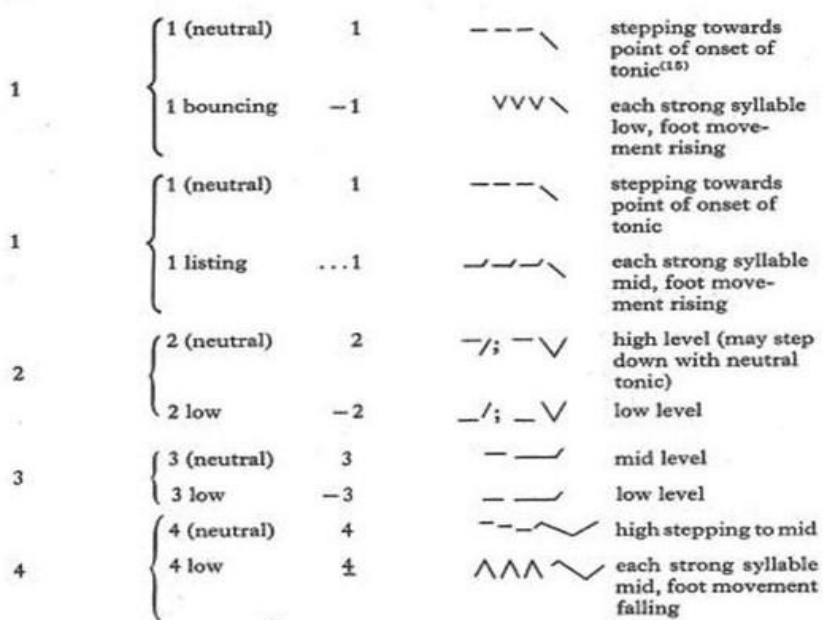
Figura 23- Tons secundários. Representação pictórica da tônica



Fonte: Halliday (1973, p. 110).

[82] Grupos tonais [...] de tom 13 e 53 têm o mesmo padrão geral de informação dos outros grupos tonais, com a tônica marcando a informação. A diferença está no fato de que há dois lugares para o falante decidir onde focalizar a informação na unidade informativa, ao invés de um. As duas tônicas não tem valores iguais. A primeira tônica, com tom 1 ou 5 é a maior; a segunda, de tom 3 é a menor. A tônica maior leva a informação principal no grupo tonal. A segunda tônica expressa a informação que é de alguma maneira secundária ou subsidiária a ele. [tradução nossa].

Figura 24- Tons secundários. Representação pictórica da pretônica



Fonte: Halliday (1973, p. 111).

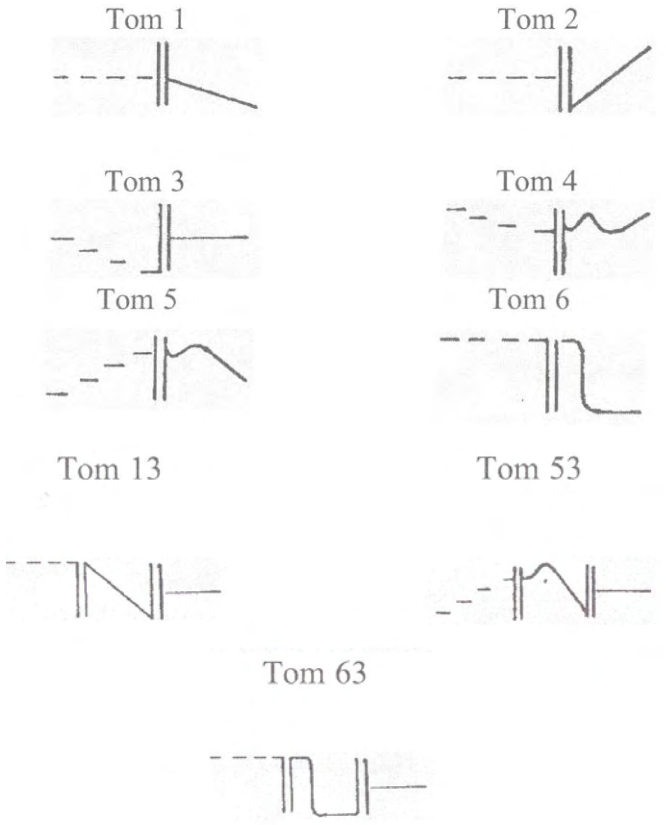
Nas figuras 23 e 24 vemos, respectivamente, os tons secundários com variações no CT e CPT. Nas duas figuras, a primeira coluna evidencia como os tons são anotados no sistema primário de tons. Em seguida, temos sua identificação no sistema de tons secundários e, posteriormente, segue a transcrição simbólica, com a anotação visual na quarta coluna. A última coluna mostra o tipo de movimento, primeiro no CT (figura 23) e depois no CPT (figura 24). Os diacríticos podem vir antes ou depois do número.

5.1.3 O modelo de Cagliari (1982/2007): a adaptação para o PB

A adaptação do modelo feita por Cagliari (1982/2007) para os padrões entoacionais do português brasileiro (dialeto paulista) não diverge muito da análise feita por Halliday para o inglês britânico. No que concerne aos tons primários, Halliday (1970; 1973) trabalhou com cinco tons primários simples e dois compostos, como mostrado. Cagliari (1982, 2007), por sua vez, propôs seis tons primários simples e três tons compostos.

Veja a adaptação da configuração entoacional proposta na figura que segue.

Figura 25- Sistema de tons primários simples e compostos do português brasileiro

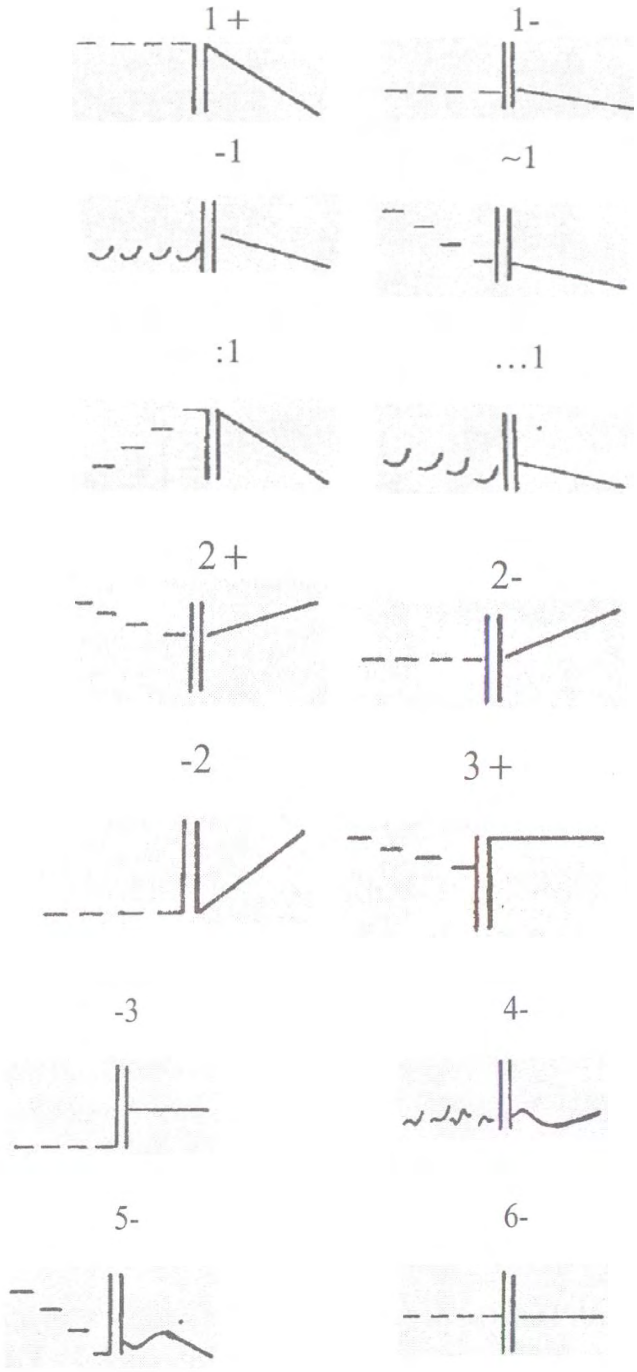


Fonte: CAGLIARI (2007, p. 189-190).

Os tons secundários sugeridos por Cagliari apresentam uma configuração bem diferente do inglês britânico. Essa distinção é esperada porque as línguas têm padrões entoacionais diferentes, bem como os falantes, que fazem interpretações de leitura distintas, e, além disso, uma pessoa pode mudar mais os tons do que a outra.

Os tons secundários são representados do seguinte modo:

Figura 26- Sistema de tons secundários para o português brasileiro

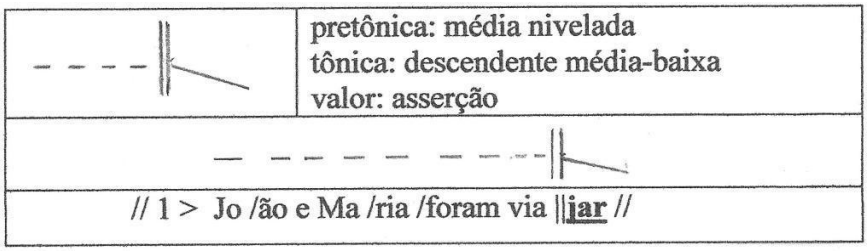


Fonte: CAGLIARI (2007, p. 190-191).

A descrição dos tons feita por Cagliari (2007) é mais completa do que a apresentada. Nos exemplos acima, mostramos somente o padrão de cada tom, sem a descrição do contorno melódico na parte pretônica e tônica. No entanto, para fins representativos, segue abaixo um exemplo de representação das principais variações melódicas da sílaba tônica

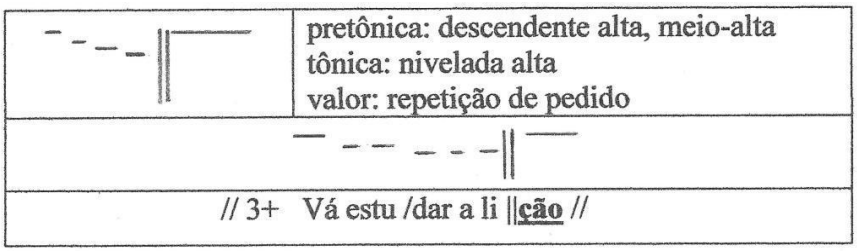
saliente e do componente tônico proposto pelo autor. O primeiro (figura 27) identifica o tom 1 primário simples e a segunda (figura 28), o tom 3+ secundário.

Figura 27- Exemplo de tom primário simples em frases assertivas



Fonte: Cagliari (2007, p. 171).

Figura 28- Exemplo de tom secundário em frases de valor enfático



Fonte: Cagliari (2007, p. 177).

No que diz respeito ao significado dos padrões entoacionais, Cagliari (2007) diz ser difícil fazer generalizações a esse respeito.

> Entretanto, parece ser válido dizer que os contornos descendentes significam algo certo e os contornos ascendentes, algo incerto. Essa noção geral do significado dos tons, em princípio, ajuda de maneira especial a interpretação de tons que mudam de direção.
> (CAGLIARI, 2007, p. 173).

Nesse sentido, o tom 4, conforme o autor, de padrão descendente-ascendente, revela uma mudança de ideia do falante em um enunciado. Esse tom é usado em orações declarativas com reservas, ou seja, que exprimem algumas restrições. Em outras palavras, não há envolvimento do interlocutor com o que está sendo dito, como, por exemplo, na frase de tom secundário 4- (variante do tom 4 simples) e de STS (**lan**): "Mas ele não estava fa**lan**do sério". (CAGLIARI, 2007, p. 177). Para o português brasileiro, o tom 4 aparece, na maioria da vezes, em frases interrogativas. Em oposição, o tom 5 indica um tipo de dúvida, reserva. Mas, diferentemente do tom 4, tudo está claro. O tom 6, que é nivelado, é usado em frases em que algum termo é colocado em suspense. (CAGLIARI, 2007).

Em casos de enunciados afirmativos, o tom 1 simples, apresenta, na maioria das vezes, um valor neutro. Em enunciados imperativos, por sua vez, indica uma ordem. O tom

1 simples também pode aparecer em enunciados declarativos e interrogativos, com palavras interrogativas. O tom 2, assim como no inglês, também é classificado como neutro. É menos frequente nos enunciados imperativos, ocorrendo, especialmente, nas sentenças negativas de valor contraditório, indicando agressividade ou indignação sobre fato. O tom 3 é típico de frases suspensivas, ou seja, que não apresentam sentido completo. Além disso, também pode ser usado para respostas não comprometedoras.

Halliday (1970) embasa o esclarecimento de Cagliari (2007), dizendo que não se pode fazer uma afirmação precisa sobre o significado geral de um tom, pois "*[...] the meaning of a choice of tone is bound up with other grammatical choices in the utterance"*.[83] Para o autor, a explicação dos tons está na gramática, uma vez que a entoação não é caracterizada apenas pela emoção, mas também pela descrição dos fatos.

Trabalhar com a transcrição dos tons a partir do modelo de análise proposto por Halliday não é uma tarefa fácil, uma vez que essa abordagem, além de exigir um treinamento auditivo muito específico, também requer uma habilidade fonética. Segundo Cagliari (2011), o treinamento exigido é semelhante ao que é feito para transcrição dos tons nas línguas tonais ou para a transcrição das qualidades vocálicas, o qual segue o método proposto por Abercrombie (1967) para as vogais cardeais. De acordo com Cagliari (2011, p.6),

> Esse treinamento acostuma o ouvido com as variações de altura melódica, à moda das notas musicais, o que permite ao linguista transcrever os tons entoacionais com uma simples análise auditiva. Sem isto, obviamente, como acontece com as notas musicais, a pessoa se sentirá confusa e pode achar que tal empreitada é impossível, dada a variação que ouve.

5.2 O modelo descritivo de Janet Pierrehumbert (1980)

Pierrehumbert, assim como Ladd (1996), também propôs uma abordagem métrica autossegmental para estudar os padrões entoacionais. Sem sombras de dúvidas, nos dias atuais, a maior influência para os trabalhos em fonologia entoacional tem sido a tese de doutorado da autora "The phonology and phonetics of English Intonation", apesar de muitos estudiosos tomarem como base apenas as ideias gerais do modelo, deixando de lado alguns pontos importantes para a análise. A abordagem de Pierrehumbert (1980) foi construída a partir da teoria métrica de Liberman (LIBERMAN, 1975; LIBERMAN e PRINCE, 1977) e sobre a análise de Bruce (1977) para o acento de palavra do Sueco.

[83] "[...] o significado da escolha de um tom está relacionada a outras escolhas gramaticais no enunciado". [tradução nossa].

Pierrehumbert desenvolveu uma teoria simples e eficiente. Segundo Liberman (1996, p. 3), seu modelo sobre a fonologia entoacional *"[...] effectively resolved the long running 'levels vs. configurations' debate and focused the attention of investigators on certain well-defined questions that have occupied much of their time"*.[84] Em sua proposta, Pierrehumbert explica, através de uma descrição fonológica, vários detalhes sobre a medição instrumental da frequência fundamental (F0) e estabelece novos padrões fonéticos.

Atualmente, há três versões diferentes para a análise entoacional de Pierrehumbert. A versão original aparece de modo completo em sua tese de doutorado de 1980. Posteriormente, em colaboração com Mary Beckman (BECKMAN; PIERREHUMBERT, 1986; PIERREHUMBERT; BECKMAN, 1988) aparece uma versão modificada da original. Por fim, a terceira, segundo Ladd (1996), é uma versão modificada e simplificada da revisão que faz parte do sistema ToBI – *Tones and Break Indices* – (SILVERMAN et al. 1992).

No decorrer da análise desta descrição metodológica, vamos apresentar sucintamente cada uma dessas três versões. No entanto, focaremos mais a primeira, já que foi o modelo escolhido como embasamento teórico desta dissertação. Vale lembrar também que, além dessas sugestões propostas por ela, foram feitas outras, por outros pesquisadores, como Ladd (1996) e Gussenhoven (2002).

5.2.1 A versão original

A versão original deu origem à teoria de Pierrehumbert, conhecida como teoria métrica autossegmental (AM) da entoação. A fonologia autossegmental teve sua evolução através de contribuições da fonologia gerativa (versão padrão: CHOMSKY, 1965), procurando investigar os aspectos fonológicos das línguas a partir do seu componente fonológico, como um conjunto de subsistemas em interação, cada um deles apresentando aspectos particulares e específicos. Segundo Cagliari (2012), é muito importante levarmos em consideração esse aspecto, porque a análise entoacional de um enunciado nesse modelo exige também uma análise da grade métrica. Pela grade métrica definem-se as sílabas tônicas, descrevendo a relação de força entre elas, e os pés métricos, em um nível superior.

O sistema proposto por Pierrehumbert é um modelo de representação abstrato, desenvolvido para dar conta da entoação do inglês americano. A intenção da autora é evidenciar os diferentes padrões entoacionais de um texto e mostrar como um mesmo padrão pode ser implementado em textos com padrões acentuais distintos. Sobre isso, ela diz que a

[84] "[...] efetivamente resolveu a longa discussão entre 'níveis vs configurações', debatendo e focando a atenção dos investigadores em questões bem definidas que ocuparam muito do tempo deles"[tradução nossa].

mesma melodia alinha-se a uma mesma sentença de diferentes modos. Segundo a autora, esse fato corresponde aos diferentes modos de se atribuir os acentos frasais.

Gussenhoven (2002) acrescenta ainda que a diferença no modo de representação de um enunciado está na intenção do falante, no contexto em que o enunciado foi dito. Para outro falante, a notação pode ser completamente diferente. Tudo irá depender do propósito comunicativo. Por isso, há um número infinito de pronunciações de um mesmo enunciado. Assim, Pierrehumbert propõe uma série de regras que transformam essas representações em realizações fonéticas, pois, segundo ela,

> *The rules will be shown to explain both quantitative features of the F0 contour, and also characteristics which one might at first take to be qualitative, such as the overall shape.*[85]
> (PIERREHUMBERT, 1980, p. 11).

De acordo com Pierrehumbert, para caracterizar e descrever esse padrão fonológico da entoação, é preciso levar em consideração três componentes. O primeiro é o componente gramatical - responsável pelos tons da frase entoacional. Essas unidades fonológicas são geradas a partir de uma estrutura gramatical de estado finito, que define os acentos melódicos e, além disso, estão alinhadas à sílaba tônica, tendo como base o padrão métrico do texto, mais dois tons adicionais que caracterizam a entoação no final do enunciado (os acentos frasais e os tons de fronteira). O segundo componente da representação fonológica é a grade métrica do texto. A grade métrica irá dizer quais sílabas são acentuadas e quais não são, além de descrever a relação de força entre as sílabas tônicas. Finalmente, o terceiro componente diz respeito às características da frequência fundamental (F0) entre os tons, explicando a relação do controle do padrão entoacional com a variação melódica. Em outras palavras, é um conjunto de regras que alinha os tons com o texto, de acordo com a estrutura do tom e a representação métrica do texto.

Partindo desse pressuposto, a autora propõe um sistema de notação que permite fazer uma associação fonológica entre os tons e o texto. Ela parte do princípio de que o contorno entoacional é formado por uma sequência de *pitch accents* e *edge tones*. Os *pitch accents* são caracterizados por um tom alto (H - *high*) ou por um tom baixo (L - *low*), ou pela combinação dos dois tons (chamados de *bitonals*). O tom principal do *pitch accent* é simbolizado por um asterisco (*) – *starred tone* – e representa a sílaba tônica frasal, o foco entoacional do enunciado como, por exemplo, H* e L*. O *pitch accent* também pode ser

[85] "As regras irão explicar as características quantitativas do contorno da F0 e também as características que poderiam ser consideradas qualitativas, assim como a configuração geral" [tradução nossa].

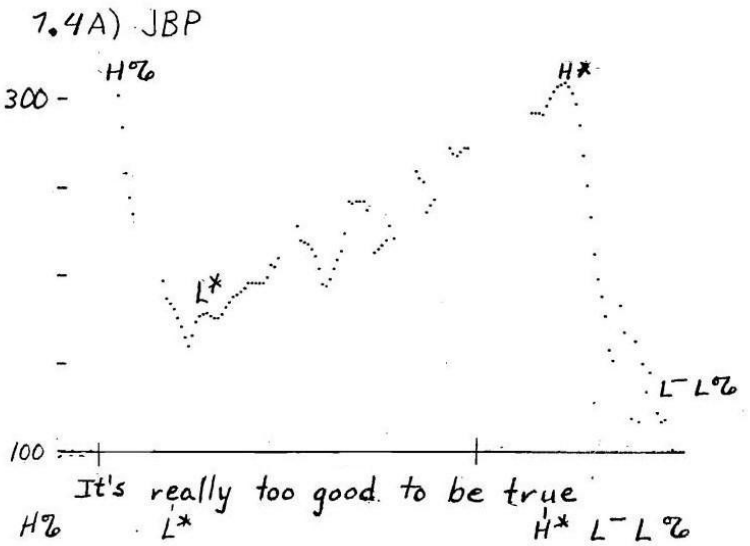
formado por um tom secundário, é *"The tone that leads or trails the starred tone"*.[86] (PIEHUMBERT, 1980, p. 23). Esse tom é indicado por um hífen sobrescrito, como em: H⁻ L⁻. *"A single tone accent behaves like a starred tone in that it lines up with the accented syllable rather than leading or trailing".*[87] (PIERREHUMBERT, 1980, p. 26).

Os *edge tones*, por sua vez, estão divididos em acentos frasais e tons de fronteira. Os acentos frasais, como mencionado, são as unidades fonológicas básicas, podendo ser formados por um tom (H ou L) ou por um par de tons (H+L)[88]. As sílabas do acento frasal nunca serão marcadas com um asterisco, pois elas são menos salientes que os *pitch accents*. Esse tom é atribuído entre o último acento frasal e o tom de fronteira, onde ocorre a variação mais notável da F0.

Os tons de fronteira, por sua vez, estão associados ao final da frase entoacional e referem-se ao valor fonológico de um tom. Na notação da autora, eles são indicados por um símbolo de porcentagem (%) ao lado de cada unidade tonal (H% ou L%). Eles serão sempre formados por um único tom. Eles estão associados às fronteiras dos domínios prosódicos e não estão relacionados à marcação de proeminência. De acordo com Pierrehumbert, o tom de fronteira ocorre no acento frasal, independentemente do padrão entoacional anterior. E, o acento frasal, por sua vez, localiza-se logo após o acento nuclear. Pode acontecer de um tom de fronteira ocorrer no início de um enunciado. Nesses casos, ele será atribuído quando o valor da F0 for alto, de tom H inicial e contorno descendente. Na maioria das vezes, o tom de fronteira inicial ocorre quando o enunciado é muito longo e o foco está no fim da frase. Veja o exemplo:

[86] "O tom que precede ou que segue o tom marcado com asterisco (*starred)*". [tradução nossa].
[87] "Um tom simples acentuado comporta-se como um tom com asterisco (*starred*) no sentido em que se alinha à sílaba (tônica), ao invés das sílabas precedentes ou subsequentes à sílaba tônica". [tradução nossa].
[88] No sistema proposto por Pierrehumbert (1980), dois tons de um *bitonal pitch accent* são unidos por um sinal de adição (+).

Figura 29- Exemplo de tom de fronteira inicial



Fonte: Pierrehumbert (1980, p. 260).

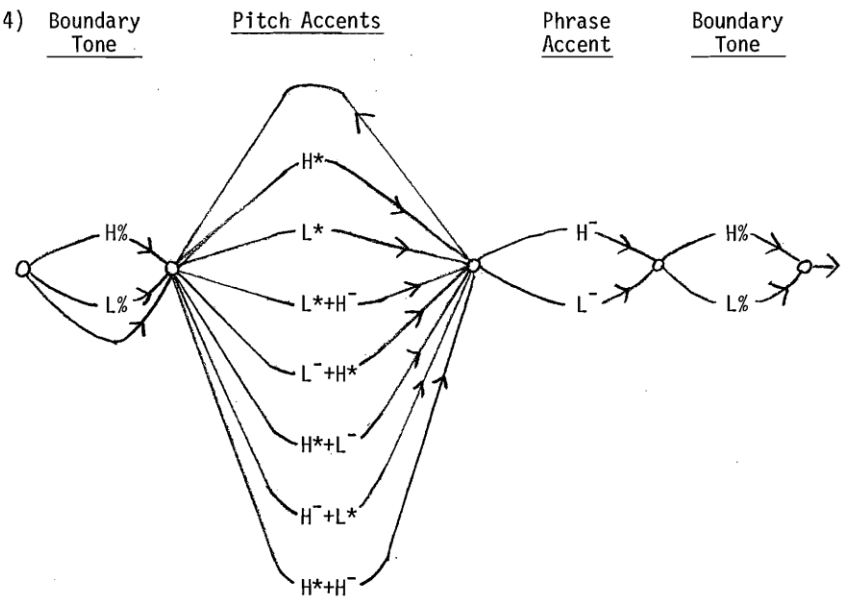
O símbolo de porcentagem (%) também é utilizado para indicar uma marcação de pausa no enunciado, como no exemplo: "*This is the cat % that ate the rat % that stole the cheese ...*"[89] (PIERREHUMBERT, 1980, p. 21).

Em seu modelo, os diacríticos (*, %, e ⁻) propostos não estão relacionados ao valor tonal, ou seja, os tons do tipo H*, H⁻ H% são igualmente tons altos, não há uma diferença de altura entre eles, mas, sim, na maneira como eles estão associados ao texto.

Com essas informações, a autora propõe uma gramática de estado finito da entoação, que gera, no enunciado, um conjunto de sequências tonais bem definidas. O padrão mostrado abaixo representa o esquema de notação simplificado da autora, pois não há a identificação do acento frasal extra usado com as *tag questions*.

[89] "Esse é o gato % que comeu o rato % que roubou o queijo..." [tradução nossa].

Figura 30- Esquema do modelo teórico proposto por Pierrehumbert para o repertório do inglês americano



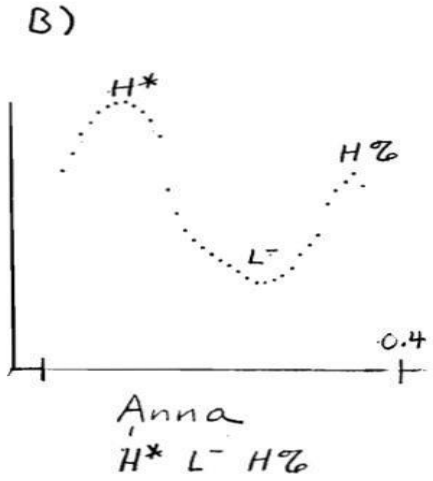
F**onte:** Pierrehumbert (1980, p. 29).

De acordo com essa gramática, os tons são formados por um ou mais *pitch accents*, que vêm seguidos de um acento frasal e um tom de fronteira obrigatórios. Segundo Pierrehumbert, os tons de fronteira ocorrem independentemente do padrão entoacional que os precede e do acento frasal, que é localizado logo depois do acento nuclear. Os tons de fronteira iniciais também estão representados. Normalmente, os contornos entoacionais do tipo H% final serão ascendentes. Por outro lado, se o tom de fronteira final L% for antecedido por acento frasal do tipo L, teremos um contorno entoacional do tipo descendente. Já se o tom de fronteira final L% for precedido por um acento frasal do tipo H, teremos um contorno entoacional nivelado.

Nas palavras de Pierrehumbert (1980, p. 22), os *pitch accents*, os acentos frasais e os tons de fronteira são *"[...] the well-formed tunes for an intonational phrase [...]"*[90]. Na figura 31, podemos ver claramente o modo como esses três elementos entoacionais estão distribuídos em um enunciado.

[90] "[...] os tons bem formados para a frase entoacional [...]" [tradução nossa].

Figura 31- Representação do *pitch accent*, do acento frasal e do tom de fronteira



Fonte: Pierrehumbert (1980, p. 255).

Segundo Ladd (1996), essa gramática também sugere que todas as combinações entre *pitch accents* e tons de fronteira são possíveis, além de mostrar que não há uma estrutura constituinte de contorno. Para Pierrehumbert, esses dois fatos revelam que não existe uma diferença entre os acentos nucleares e prenucleares, mas que eles diferenciam-se apenas em sua posição na cadeia melódica. Contudo, para Ladd (1996), esse não seria o caminho mais adequado a seguir. Para ele, o acento nuclear deveria ser abordado de um modo mais especial, uma vez que está situado no núcleo do contorno melódico, que é a parte obrigatória para a formação do enunciado e, por isso, exerce uma influência maior sobre os significados semânticos e pragmáticos dos enunciados.

Em seu levantamento, Pierrehumbert destaca sete tipos de *pitch accents*: H*, L*, L*+H^{-}, L^{-}+ H*, H*+L^{-}, H^{-}+L*, H*+H^{-}. Entre eles estão representados os tons simples (*single tones*) e os tons compostos (chamados por ela de *bitones*). Os tons compostos, de acordo com a autora, assemelham-se aos pés bissilábicos, no sentido em que eles são classificados por uma relação de força, Por outro lado, os tons simples são equivalentes aos pés monossilábicos.

Como vimos, essas unidades fonológicas são descritas pelos rótulos de H (*high*) e L (*low*) e pelas possíveis combinações entre eles. Não há outro tom além desses. O objetivo de Pierrehumbert (1980) ao desenvolver seu modelo entoacional centrada em dois tons foi fazer um apanhado entre os modelos anteriores de quatro níveis tonais, como os de Trager e Smith (1951), Pike (1945) e Liberman (1975) e teorias que procuraram explicar a entoação em

termos de mudança de F0 ao invés de níveis de F0, como a de Bolinger (1951, 1958), Ladd (1980) e Clark (1978).

Segundo Pierrehumbert, *"[...] a theory framed in terms of target levels is attractive because it affords good facilities for describing how the same intonation pattern lives up with different texts*".[91] (PIERREHUMBERT, 1980, p. 54). Ela explica, ainda, que uma teoria de dois tons e de regras de implementação contextos-sensitiva ajuda a resolver muito dos problemas com as teorias de quatro tons, além de ter a vantagem de descrever a entoação em sequências de níveis tonais[92]. Para a autora, e assim como já havia sido notado por Bolinger (1951), a anotação com quatro níveis tonais gera uma ambiguidade porque, segundo ela, confunde diferenças tonais com diferenças na escolha do *pitch range* (tessitura, escala melódica). Por exemplo, ela explica que um contorno do tipo H% L* H* L- L% seria interpretado, segundo uma teoria de quatro tons, como LM L LM L L ou HM L HM L L em uma alta taxa de elocução ou, ainda, como H L H L L em uma taxa de elocução mais moderada. Por outro lado, de acordo com ela, isso não ocorreria em uma descrição de dois tons, porque o contorno tem apenas a possibilidade de transcrição dada.

Gussenhoven (2002) também justifica a representação por apenas dois tons. Para ele, as regras propostas por Pierrehumbert para a geração de um número infinito de valores da F0 entre o mais alto e o mais baixo *pitch* permite que a distinção entre a representação fonológica e o contorno fonético torne-se mais clara, pois, além de descrever somente o que é linguisticamente possível, caracteriza os contrastes da língua. E, para isso, bastam apenas dois tons, tanto para o inglês quanto para outras línguas, de modo que a implementação fonética cria os alvos contextualmente apropriados para cada tom.

Por sua vez, as regras de reajuste entoacional (geradas a partir dos eventos tonais) não estão baseadas em um sistema de derivação (o que seria esperado, já que seu modelo teórico é de base gerativa, mas em regras locais contexto-sensitivas que mapeiam a sequência dos tons de acordo com os valores quantitativos que determinam o contorno da frequência fundamental (F0). Esse sistema representacional é chamado de *mapping* e caracteriza, segundo Cagliari (2011), um padrão fonológico entoacional que "[...] é atribuído a um enunciado (*text*) através de uma representação (*mapping*) que configura o enunciado entoacionalmente". (CAGLIARI, 2011, p.12). Essas regras se aplicam da direita para a esquerda e incluem movimentos de *downstep* e *upstep*, regras de espraiamento (*spreading)*,

[91] "[...] uma teoria baseada em termos de alvos tonais é atrativa porque proporciona boas facilidades para descrever como o mesmo padrão entoacional alinha-se com diferentes textos". [tradução nossa].

[92] Uma discussão mais detalhada sobre o que motivou Pierrehumbert a trabalhar com apenas dois níveis tonais (H e L) pode ser melhor aprofundada em sua tese de doutoramento (Pierrehumbert, 1980, p. 54-59).

de formação de *plateau* e de implementação das linhas de base (*baseline units*). Elas são específicas da língua inglesa e estão mais ligadas aos tons do que aos padrões entoacionais.

Antes de darmos início à caracterização dessas regras, é necessário fazer algumas considerações sobre a interpretação fonética dos eventos tonais apresentados mais acima. Em relação aos tons H* e L*, o primeiro é caracterizado pelo pico local e vem alinhado à sílaba tônica; o segundo, caracteriza o vale local. Paradigmaticamente, L* é mais baixo que H*, no mesmo contexto. O valor fonético de L* diminui se a proeminência aumenta. Nas mesmas circunstâncias, ocorre o aumento de H*. A diminuição de L* e o aumento de H* não são simétricos, uma vez que L e H estão em diferentes relações com o falante.

O contorno do tipo L*+H^{-} é o exemplo mais comum de representação de um *pitch accent* envolvendo dois tons. É um contorno entoacional baixo para a maior parte da sílaba tônica, que aumenta bruscamente em direção à sílaba átona, se ela está presente. Caracteriza uma atitude de incredulidade e também pode indicar, de modo menos enfático, uma resposta incompleta.

Outros dois *pitch accent* são: L^{-}+H* e H^{-}+L*. Segundo Ladd (1996), esses contornos são caracterizados por movimentos da sílaba anterior e não por movimentos que a seguem, como ocorre em L*+H^{-}. Nesses dois tons compostos, o tom com asterisco *(starred)* (*) é o nuclear, alinhado com a sílaba tônica, enquanto o tom sem asterisco *(unstarred)* determina, na maioria das vezes, a variação melódica da sílaba anterior. L^{-}+H* é, segundo o autor, basicamente igual ao tom H*, com um pico local, porém, em um nível muito mais alto. A diferença entre L^{-}+H* e H* está na presença de uma sílaba anterior que leva o tom L^{-}. O *pitch accent* H^{-}+L* pode ser superficialmente comparado ao *pitch accent* L^{-}+H*, no sentido em que esse envolve um pico local na sílaba anterior e aquele, uma queda local. O *pitch accent* H*+L^{-} representa um tom alto que desencadeia uma regra de *downstep* a partir do tom H que o segue.

O *pitch accent* H*+H^{-}, por sua vez, merece um destaque especial. Na análise original, esse padrão H*+H^{-} é conhecido como um tom de contorno e representa *"[...] a high accent that is followed by a sustained high level transition to a following high accent [...]*." [93] (LADD, 1996, p. 273-247). Contudo, na versão revisada do modelo (que será vista mais adiante) apresentada por Beckman e Pierrehumbert (1986), esse padrão tonal foi eliminado, restando, então, somente seis combinações possíveis (H*, L*, L*+H^{-}, L^{-}+ H*, H*+L^{-}, H^{-}+L*).

[93] "[...] um acento (melódico, H) alto que é seguido por uma transição de nível alto sustentado em sequência a um acento alto [...]" [tradução nossa].

Segundo Ladd, a omissão desse *pitch accent* deve-se ao fato de ele não se relacionar bem com os outros padrões entoacionais. Um dos maiores problemas, de acordo com o autor, é o fato de ele violar o Princípio do Contorno Obrigatório (PCO)[94], uma vez que os outros tons compostos envolvem uma combinação de tons H e L, enquanto esse H*+H^{-} contém apenas dois tons do mesmo tipo (H). Portanto, há razões teóricas para a eliminação desse *pitch accent*. No entanto, as autoras não apresentam justificativas suficientes para essa reanálise, diz Ladd. Segundo ele, elas somente explicam que esse ajuste foi uma consequência natural do novo tratamento dado por Liberman e Pierrehumbert (1984) para a escala entoacional, mas não explicam a razão dessa eliminação.

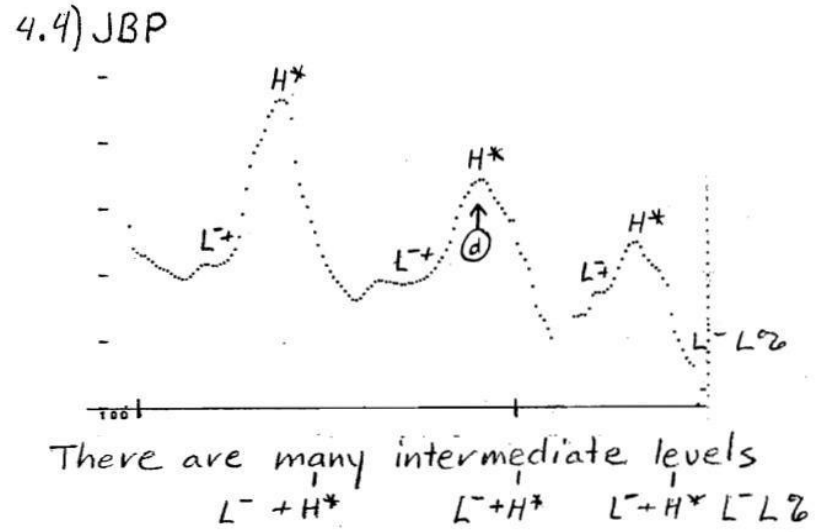
Em suma, a propriedade comum dos padrões entoacionais dos tons compostos descritos, com exceção de H*+H^{-}, é a configuração de uma queda abrupta. Isso ocorre porque há um abaixamento mais rápido do que a declinação poderia considerar. O último pico tem sempre um valor menor do que o seu anterior, mesmo que ele ocorra no acento nuclear mais proeminente. Os contornos diferenciam-se na característica do contorno da F0, que estão localizadas em cada sílaba acentuada.

Dando sequência às regras de reajuste entoacional, as duas mais discutidas por Pierrehumbert são as de *downstep* e de *upstep*[95]. Em relação às regras de *downstep*, os quatro tons compostos que podem desencadeá-las são: H^{-}+L*, H*+L^{-}, L*+H^{-}, L^{-}+H*. Todos eles, exceto o H*+L^{-}, ocorrem transparentemente nos contornos da F0. Observe um exemplo na figura 32 abaixo:

[94] Em nota, Ladd (1996) explica que o PCO foi desenvolvido por Goldsmith (1976). De acordo com esse princípio, as representações fonológicas serão mal formadas se elas apresentarem o mesmo tom adjacente ou, características adjacentes idênticas.

[95] *Downstep* é um termo técnico que se refere à queda melódica que ocorre do começo ao fim do enunciado de modo contínuo. *Upstep* é o contrário do downstep, isto é, ocorre com movimento crescente da curva melódica de modo contínuo no enunciado.

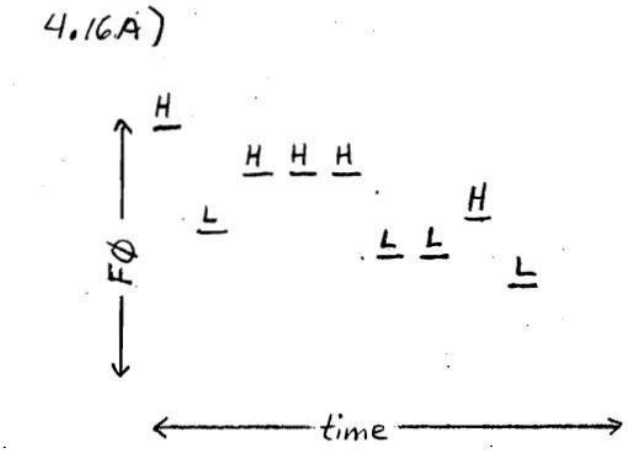
Figura 32- Contorno entoacional do tipo *downstep* no enunciado: "There are many intermediate levels"



Fonte: Pierrehumbert (1980, p. 331).

Segundo Pierrehumbert, o esquema básico para a notação do *downstep* é o seguinte:

Figura 33- Esquema básico para representação do *downstep*



Fonte: Pierrehumbert (1980, p. 339).

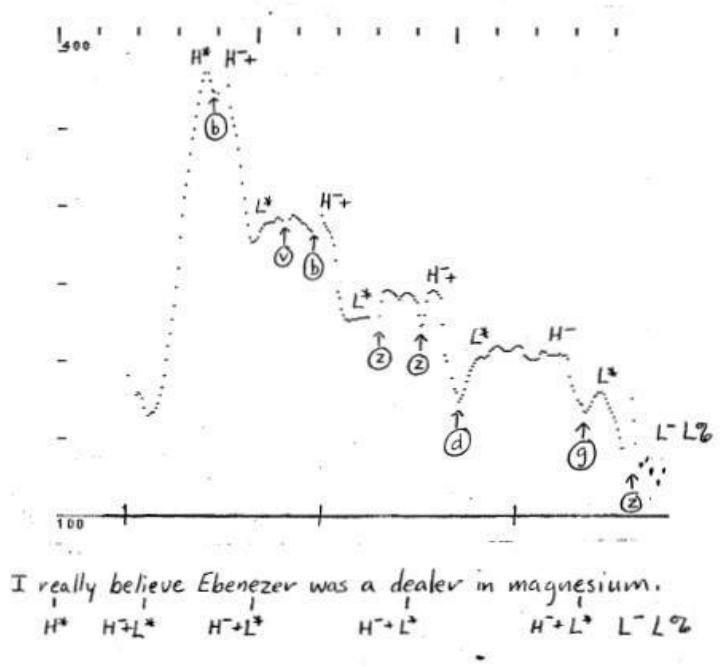
No *pitch accent* H-+L* também há uma regra de *downstep* porque o acento nuclear não está no vale local ou na F0 mínima, mas

> [...] *at the level of a 'downstepped' or locally lowered high tone: that is the pitch drop indicated by the* H+L* *notation is from one high pitch to another*

> *high pitch that is slightly lower, not from a high pitch to a level that would independently be describe as a low.*[96]
> (LADD, 1996, p. 85).

Esse tipo de contorno é chamado por Pierrehumbert de "*terraced downstepping*" e está representado abaixo:

Figura 34- Padrão entoacional do Inglês do tipo *terraced downsteping*



Fonte: Pierrehumbert (1980, p. 331).

Nessa sequência tonal, há um abaixamento nos alvos, sendo que o segundo H é sempre mais baixo do que o primeiro e o terceiro sempre mais baixo do que o segundo e assim sucessivamente. O valor descendente da curva melódica é o que irá produzir o efeito auditivo de "terraços" (*terrace)* ou, de acordo com Cagliari (2011), um efeito de andares que descem. Para Gussenhoven (2002, p. 274), o *"Downstep may be seen as a grammaticalization of declination, just as many tone languages have grammaticalizations of final lowering"*[97].

Segundo Cagliari (2011), a partir dessa regra, Pierrehumbert chegou às linhas de base com fórmulas de cálculo específicas que estão relacionadas às variações de tessitura.

[96] [...] no nível do 'downstep' ou do abaixamento local do tom alto que é a queda do *pitch* indicado por H+L*, a notação vem de um tom alto para outro que é ligeiramente mais baixo, não a partir de um tom alto para outro que poderia ser independentemente descrito como baixo" [tradução nossa].

[97] "Downstep deve ser visto como uma gramaticalização da declinação, assim como muitas línguas tonais têm gramaticalizações no abaixamento final" [tradução nossa].

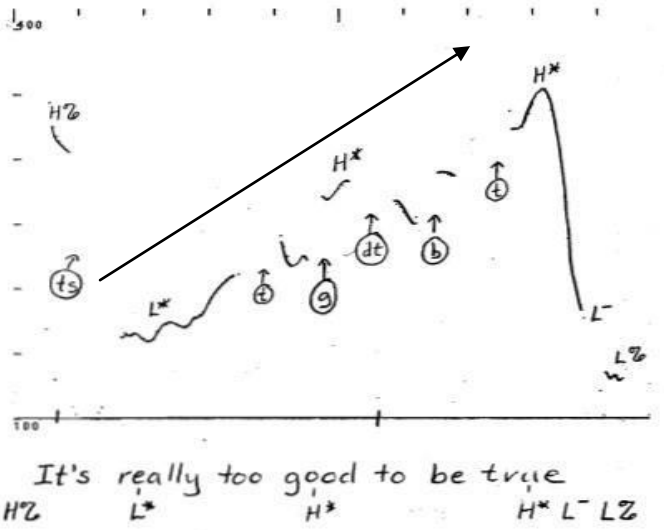
A regra de *upstep*, por sua vez, é local e está relacionada ao final do enunciado e não ao que vem antes, como é o caso do *downstep*. Nessa regra, há um aumento do L ou H que vier depois de um acento prasal do tipo H^-

> [...] *the final rise is what one would expect phonetically from a sequence of L and H, but from a sequence of H and H one might expect only a sustained level, and the upstep rule ensures that the level of the H% boundary tone is higher than that of the immediately preceding H phrase tone.*[98]
> (LADD, 1996, p. 87).

Segundo Pierrehumbert, pode acontecer no máximo dois *upsteps* em um enunciado, em contraposição aos *downstep*, em que não há um número mínimo ou máximo definido. Essas duas regras estão relacionadas à variação de tessitura e são responsáveis por reajustar o valor fonético de um tom em um determinado contexto.

A figura 35 abaixo mostra um caso típico de *upstep*, com contorno melódico ascendente, como indicado pela seta crescente na figura. O início do enunciado é marcado com um tom H%, já que temos uma altura melódica alta. O foco vem anotado na última sílaba tônica do enunciado (*true*).

Figura 35- Exemplo de um caso de *upstep* no enunciado: "It's really too good to be true"



Fonte: Pierrehumbert (1980, p.364).

Em relação às outras regras de reajuste entoacional, a discussão de Pierrehumbert sobre as linhas de base começou a partir do estudo da declinação (*declination*) da F0, a partir

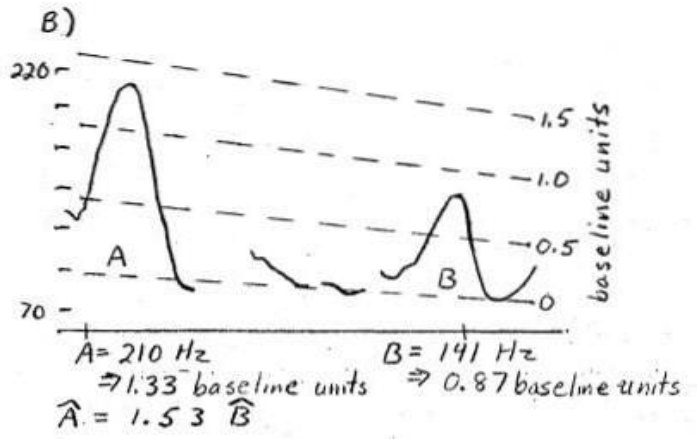
[98] [...] o aumento final é o que se poderia esperar foneticamente de uma sequência de L e H, mas de uma sequência de H e H pode-se esperar somente a manutenção no nível, e a regra de *upstep* garante que o nível do tom alto de fronteira seja mais alto do que o tom frasal imediatamente anterior. [tradução nossa].

da determinação de como as especificações tonais são mapeadas em valores da F0. De acordo com Cagliari (2011, p. 12), "Essa preocupação é decorrente do fato de Pierrehumbert se apoiar nos valores da F0 e não na percepção da variação melódica da fala".

Para Pierrehumbert, é importante entender a declinação por duas razões. A primeira, para modelar o valor da F0 com precisão e dar conta de como o ouvinte recupera as relações entre os valores tonais e a proeminência, a partir da F0 que ele ouve. E a segunda, para o entendimento das distinções categoriais possíveis no sistema entoacional. Contudo, de acordo com Cagliari (2011), essa regra é um dos menores problemas para a definição fonológica dos padrões entoacionais atribuídos pelos enunciados.

A linha de base também ajuda a resolver os problemas das regras de *downstep* e *upstep*, marcando a F0 mais baixa da tessitura do falante, que acontece, normalmente, no final do enunciado. Para isso, Pierrehumbert calcula um ângulo com regras da direita para a esquerda, implementando a linha de base em uma sequência crescente até o início do enunciado. Dessa forma, ela afirma que todos os valores da F0 terão uma proporção de altura melódica compensada pela tessitura, do início ao fim do enunciado. O falante apresenta, em geral, uma linha de base que lhe é própria. Veja um exemplo de implementação da linha de base:

Figura 36- Exemplo de como calcular graficamente a linha de base



Fonte: Pierrehumbert (1980, p.322).

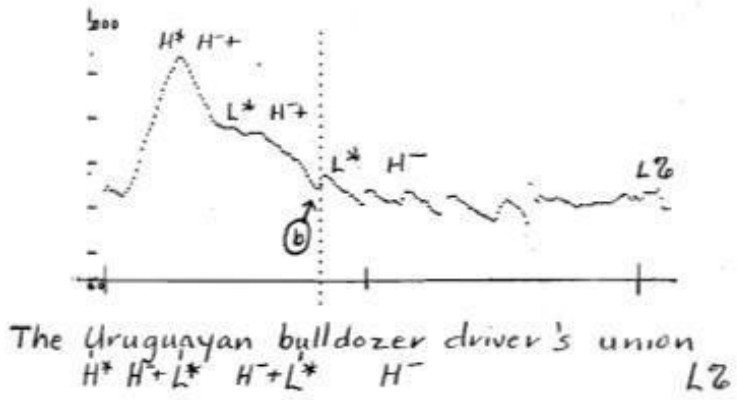
Segundo Cagliari (2011), quando há mais de uma frase entoacional em sequência, o valor da linha de base será calculado para cada uma delas, a partir da F0 mais baixa, que

costuma situar-se na frase entoacional que está mais à direita. Ainda nas palavras do autor, uma pausa pode reestruturar o sistema e criar uma nova linha de base.

A regra de *tone spreading* (espraiamento de tom), por sua vez, acontece em variações melódicas mais longas, ou seja, quando a distância entre T*___ T T% é muito grande. É um caso de propagação do tom. Isso ocorre, normalmente, com as sílabas tônicas, que costumam ser as mais longas em um enunciado. Conforme Cagliari (2011), na maior parte das vezes, o espraiamento forma um tom composto, como, por exemplo, sequências de H+L ou L+H. Nos casos observados por Pierrehumbert, o tom que se propaga é o T^{-}, devido ao seu valor fonético e do tom que o segue.

A formação do *plateau* (formação melódica nivelada), por sua vez, é gerada, na maioria das vezes, por um tom espraiado. O tom espraiado caracteriza-se por abranger toda a palavra, enquanto um tom composto compreende apenas a sílaba. Em contextos de formação de *plateau*, depois de um tom T^{-}, há uma neutralização da variação melódica, que se mantém constante. Pode acontecer de termos um *plateau* no meio do enunciado e, logo em seguida, uma subida repentina. Esse é um padrão comum para enunciados afirmativos que iniciam com uma altura melódica mais alta, mantendo, em seguida, um padrão entoacional mais constante. No enunciado *"The Uruguayan bulldozer driver's union"* abaixo, vemos um caso característico dessa regra.

Figura 37- Exemplo de formação de *Plateau*



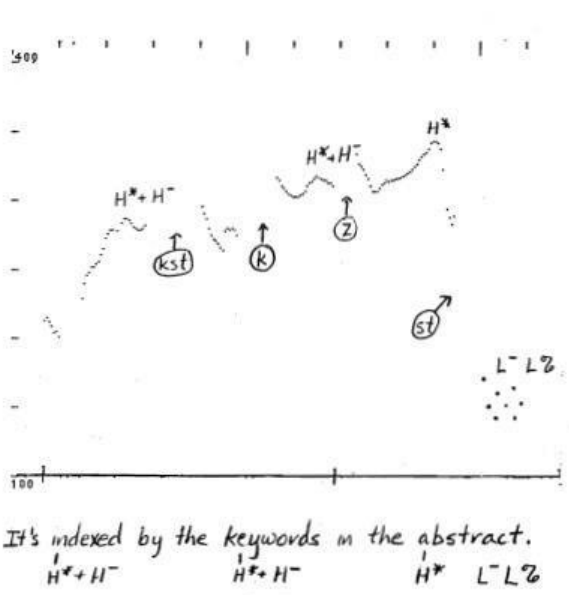
Fonte: Pierrehumbert (1980, p.354).

No exemplo, o enunciado começa com uma sílaba átona. Em seguida, ocorre um pico entoacional na sílaba *bull*, de *bulldozer*. Na sílaba nuclear, temos um tom composto do tipo H^{-}+L*. Posteriormente, temos um espraiamento no tom H^{-}, formando um *plateau* em

um nível médio, de aproximadamente 100 cps, em um enunciado em que a variação melódica da frequência fundamental (F0) vai de 100 a 200 cps.

Em relação à questão auditiva, percebemos que, mesmo não sendo um aspecto considerado pela metodologia apresentada, é uma questão que está implícita no modelo. Sendo assim, seria muita pretensão tomar esse modelo somente do ponto de vista acústico, pois algumas evidências nos levam a acreditar que Pierrehumbert não está fazendo sua análise somente pela frequência fundamental, mas também está ouvindo, assim como pode ser visto na figura 38 abaixo. Pela figura, observamos que a autora não trabalhou somente com uma análise acústica da F0, mas também precisou de um respaldo auditivo, já que o final do enunciado, sendo surdo, só poderia ter sido decidido por meio de uma interpretação auditiva.

Figura 38- Evidências de interpretação auditiva no modelo de Pierrehumbert para o enunciado: "It's indexed by the keywords in the abstract"



Fonte: Pierrehumbert (1980, p. 363).

Outro fator que comprova a análise auditiva é a questão da segmentação. Pois como a autora poderia saber em que local está situada determinada palavra ou sílaba, como por exemplo, *"keywords"*, sem ouvir o enunciado? Em nossas análises, fazemos uma segmentação pelo PRAAT, mas não deixamos de lado a interpretação auditiva.

À parte dessas considerações metodológicas, Gussenhoven (2002) acredita que a abordagem apresentada foi responsável por inaugurar um novo período na pesquisa entoacional, trazendo um modelo que separa a representação fonológica da sua

implementação fonética. Além disso, de acordo com o autor, o outro apecto inovador no trabalho de Pierrehumbert diz respeito ao conceito de alvo (*target)* como sendo distinto de um tom fonológico.

Contudo, de acordo com o autor, as inovações propostas foram poucas, de modo que muitos dos elementos utilizados já haviam sido mencionados em outros trabalhos. Dentre eles, está, primeiramente, a ideia de que há camadas (*tiers*) separadas para os tons e fonemas segmentais. Conforme o autor, essa questão já estava implícita em muitas descrições da entoação, uma vez que as características melódicas não eram consideradas como sendo parte da composição dos segmentos na tradição da descrição Britânica da entoação (cf. O'Connor & Arnold, 1973), ou mesmo na descrição da entoação para o Alemão (cf. 't Hart, Collier & Cohen, 1990), ou, ainda, como o modelo autossegmental de Goldsmith (1976), que já o havia aplicado para a entoação do inglês.

Outro fato previamente estudado foi a ideia de que a fala reflete uma representação fonológica que consiste na organização hierárquica dos constituintes, como, por exemplo, a representação da frase entoacional, que já tinha sido entoacionalmente definida na pesquisa prosódica de Selkirk (1980).

O terceiro fato, segundo Gussenhoven, é a separação em *pitch accents* e tons de fronteira, que já haviam sido descrito por Trager & Smith em 1951 como fonemas de juntura que existiam ao lado dos *pitch phonemes.* A divisão entre *pitch accents* e tons de fronteira também já havia sido reconhecida, conforme o autor, nos trabalhos de 't Hart & Collier (1970) e Cohen (1980).

Outra questão trazida por Gussenhoven diz respeito à interpolação entre os níveis, ou seja, a noção de que o início e o fim dos pontos de um determinado contorno resultem em um alvo e em uma especificação tonal, o que também já havia sido tratado por Pike (1945) e Trager & Smith (1951). A noção das sequências tonais que contém itens lexicais e entoacionais também já havia sido abordada por Bruce (1977). Por fim, o autor refere-se à notação por meio dos dois tons H (high) e L (low), que, segundo ele, já havia sido feito por Liberman (1975).

Sendo assim, podemos dizer que, mesmo Pierrehumbert tendo feito uma análise entoacional muito detalhada, não se pode dizer que ela tenha sido a precursora de todos os métodos e conceitos utilizados em sua tese.

5.2.2 Modificação do modelo original

A grande modificação da tese de doutoramento de Pierrehumbert foi proposta seis anos depois por Beckman e Pierrehumbert em 1986. As autoras propuseram uma revisão geral da estrutura tonal do inglês, mas a maior mudança centrou-se na questão da análise das regras de *downstep*. Essa modificação ocorreu com uma descrição da entoação do japonês, uma língua de ritmo silábico, mas interpretada como uma língua de ritmo "moraico". Para Cagliari (2011, p. 10), "A interpretação do japonês como língua moraica é muito problemática e baseada em uma teoria do ritmo que não se aplica à tradicional divisão entre língua de ritmo acentual e língua de ritmo silábico".

O ponto central dessa questão está entre a frase entoacional *(intonational phrase)* e a frase intermediária *(intermediate phrase)* como níveis de estrutura prosódica no inglês. De acordo com Ladd (1996), a motivação de Beckman e Pierrehumbert para essa distinção está no fato de que o inglês não apresenta uma frase acentual como as encontradas para o japonês, isto é, uma unidade prosódica baseada em um único *pitch accent*.

5.2.3 O sistema ToBI (Tones and Break Indices)

A mais recente versão do modelo modificado de Pierrehumbert foi a incorporada pelo sistema ToBI, uma proposta padrão desenvolvida por um grupo de pesquisadores para a rotulagem das características prosódicas da fala digital, a partir da base de dados do inglês. Entre outras coisas, uma ferramenta desenvolvida para auxiliar na síntese e no reconhecimento da fala. Na transcrição proposta por esse sistema, a marcação é feita a partir de diferentes camadas (*tiers*) que estão associadas a um programa que possa reproduzir e apresentar a curva de F0.

> As diferentes camadas servem para a marcação de tons, transcrição ortográfica, agrupamento prosódico e para quaisquer observações sobre a transcrição consideradas importantes para a interpretação de elementos nas outras camadas e são nomeadas respectivamente por *tone tier*, *ortographic tier*, *break índex tier* e *miscelaneous tier*.
> (LUCENTE, 2008, p. 8).

Nesse sistema, diferentemente do modelo proposto em 1980, não são possíveis combinações de tons compostos que se iniciem por um tom H, como, por exemplo, H*+H, mas somente marcações iniciadas pelo tom L, como L+H*. Por isso, de acordo com Lucente (2008), as combinações do tipo H+L* e H*+H foram substituídas, respectivamente, por H* e H+!H*. A representação de *downstep* foi abandonada e passou a ser indicada pelo diacrítico (!), igual ao utilizado pelas línguas africanas.

Segundo Lucente (2008), o sistema ToBI foi desenvolvido para a transcrição entoacional do inglês americano e, por esse motivo, sua simbologia e combinações de rótulos pressupõem um conhecimento dos aspectos entoacionais e fonológicos específicos dessa língua. Então, foi preciso que as línguas que o utilizaram desenvolvessem uma adaptação dele, conforme a sua transcrição, criando uma simbologia útil para a descrição dos seus próprios fenômenos.

5.3 O modelo de Halliday e Pierrehumbert: uma comparação

5.3.1 Semelhanças e diferenças

Nos dois modelos, a entoação é vista como um fenômeno fonológico. A primeira diferença está em como os autores enquadram a entoação. Pierrehumbert (1980) desenvolve seu modelo teórico a partir das ideias propostas pela Gramática Gerativista. Uma gramática que parte dos princípios gerativistas tem como objetivo geral definir e elaborar regras para a construção de enunciados. O arcabouço de Halliday (1970), por outro lado, é o da gramática sistêmica funcionalista, que entende as funções linguísticas como uma relação entre forma e estrutura. Os funcionalistas acreditam, de acordo com Lyons (1987), que as funções da linguagem determinam as estruturas fonológicas gramaticais e semânticas. Dessa forma, o uso da linguagem é modelado pelo sistema linguístico. Partindo dessa ideia, para fazer um estudo linguístico, é preciso selecionar as unidades e relacioná-las à forma e ao produto, criando, assim, uma unidade fonológica.

Para Halliday, não é possível trabalhar só com a fonologia ou só com a fonética. É preciso juntar as duas, pois o trabalho individual leva a distorções, tanto em uma quanto em outra. Por isso, diferentemente de Pierrehumbert, o autor acredita não ser possível descrever a entoação exclusivamente com a análise (auditiva ou acústica) fonética, bem como não é possível associar os padrões dos tons às categorias gramaticais diretamente. Isso porque essas descrições são independentes umas das outras. “Ocasionalmente, há uma relação entre categorias gramaticais de outro nível com o sistema entoacional, mas não há nenhuma categoria que exija um padrão entoacional único e exclusivo ou vice-versa” (CAGLIARI, 2011, p. 3). Contudo, conforme Cagliari, a característica abstrata e geral dos padrões entoacionais também é considerada pela autora.

Nesse sentido, Halliday acredita que a entoação, como um processo gramatical, depende de outros itens. É um processo fonológico que está inserido em um sistema ligado à sintaxe, à pragmática e à semântica. Essa visão vai contra o que Pierrehumbert faz. Para ela, a entoação não é algo sistemático e fluído, enquanto, para ele, tudo na língua é sistemático, com limites controlados de variação. Ambos são sistemas fonológicos, e não um arranjo

ocasional de dados. O próprio autor justifica seu ponto de vista, considerando que os contrastes fonológicos são tratados por um sistema de termos discretos, aplicando-se tanto para os sistemas fonológicos quanto para os contrastes gramaticais, nos quais os termos fonológicos identificados atuam como *exponents*[99]. Portanto, como Halliday (1973, p. 124) afirma,

> *It is possible to describe English intonation in this way, and justifiable on the grounds that in present-day linguistic theory we can handle discreteness more effectively than non-discreteness, at least at the level of grammar*[100].

Além disso, segundo o autor, sempre que descrevemos, a entoação estamos preocupados com o significado e todos os contrastes significativos podem ser justificados, seja na gramática ou no léxico. Assim,

> *If we regard intonation in English as meaningful – if, for example, the choice between two possible utterances which differ only in that one has tone 1 and the other has tone 4 is a true choice between different utterances – then we should seek to state the place which such choices occupy relative to the total set of formal patterns in the language; and there only two kinds of formal pattern: grammatical and lexical*[101].
> (HALLIDAY, 1973, p. 104).

Porém, isso não significa que o sistema entoacional deva ser tratado por um conjunto de nuances emocionais sobrepostas a um conjunto de itens lexicais e gramaticais.

Halliday e Pierrehumbert começam a descrever seu modelo teórico a partir da caracterização dos dialetos com que vão trabalhar. Halliday com a entoação do inglês britânico (variedade do "Received Pronunciation" - RP) e Pierrehumbert com a entoação do inglês americano (*standard)*. Com essa descrição, Cagliari (2011) diz que os autores querem mostrar que diferentes dialetos podem ter padrões entoacionais distintos.

> A entoação funciona como qualquer outro contraste gramatical (tempo, modo, etc.). Por outro lado, o sistema entoacional pode ser *polisystemic*, ou seja, pode haver mais de um sistema entoacional para uma língua, dependendo de quão *comprehensive* é o nível de delicadeza com que se define a gramática.
> (CAGLIARI, 2011, p. 2).

---

[99] Entenda como *exponed* (expoente, termo técnico da teoria de Halliday) a propriedade (prosódica) distintiva fonológica, atribuída aos vários níveis da entoação.

[100] É possível descrever a entoação do Inglês dessa maneira e justificá-la com base no fato de que na teoria linguística atual, nós podemos lidar com o discreto mais apropriadamente do que com o não discreto, ao menos no nível da gramática [tradução nossa].

[101] Se considerarmos a entoação do inglês como algo significativo – se, por exemplo, a escolha entre dois possíveis enunciados que diferem somente no sentido de que um tem o tom 1 e o outro o tom 4 é uma verdadeira opção entre enunciados diferentes – então nós devemos procurar o lugar que essas escolhas ocupam em relação ao conjunto total de padrões formais da língua; e há apenas dois tipos de padrão formal: o gramatical e o lexical. [tradução nossa].

Cagliari também diz que, assim como Halliday, Pierrehumbert quer mostrar que não há uma relação do contraste de variação melódica com o significado lexical, como ocorre, por exemplo, com o tailandês ou com o vietnamita, ambas línguas tonais. Na visão teórica dos autores, é possível descrever as línguas tonais e entoacionais seguindo o mesmo modelo.

Em relação aos padrões melódicos, os dois autores partem, inicialmente, da atribuição de valores aos pés rítmicos (para ele) e aos pés métricos (para ela). Para o primeiro, os pés estão organizados a partir de sua estrutura sintática. Em seu modelo, o enunciado é dividido em dois componentes, tema e rema, ambos com objetivos informativos. Esses componentes são chamados "componente tônico" (CT) e "componente pré-tônico" (CPT). Este caracteriza as variações entoacionais e as atitudes do falante e aquele, os elementos sintáticos. A divisão entre eles é dada pela sílaba tônica. A autora, por sua vez, não fala em níveis gramaticais, mas também separa o enunciado em dois componentes, nomeados por ela de "componente nuclear" e "pré-nuclear".

No modelo funcional, a divisão entre componente tônico e componente pretônico é chamada de grupo tonal (GT). São os pés rítmicos que operam nessa separação. No autossegmental, por outro lado, não há menção a essa nomenclatura. Mesmo assim, podemos comparar o GT de Halliday ao que Pierrehumbert chama de *tunes*, ou seja, entidades linguísticas que estão alinhadas com o texto através de regras linguísticas com um padrão determinado.

Segundo Cagliari (2011), o componente pretônico para Halliday, e o componente pré-nuclear para Pierrehumbert, em um primeiro grau de delicadeza são, fonologicamente, secundários pois, "Somente secundariamente, o que ocorre nesse contexto passa a ser relevante, definindo atitudes do falante, padrões sintáticos (enumerações...), semânticos (atitudes do falante...) ou pragmáticos (atos de fala...). (CAGLIARI, 2011, p. 3-4).

No sistema entoacional de Halliday, os cinco níveis entoacionais são definidos pela configuração do contorno melódico do CT. Já em Pierrehumbert, essa anotação é feita a partir da sílaba tônica frasal e do tom de fronteira final.

Quanto à atribuição de tonicidade, os autores classificam as sílabas entre tônicas e átonas. Ambos desconsideram as sílabas átonas na análise entoacional. Para saber onde está localizada a sílaba de maior proeminência, chamada por ela de *pitch accent*, ela parte da estrutura métrica do enunciado. Ele, por sua vez, a nomeia de sílaba tônica saliente (STS), definindo-a a partir do foco entoacional do enunciado, ou onde ocorre a mudança mais notável da direção do contorno entoacional. Nos dois modelos, a consideração de que a

sílaba tônica mais proeminente é a que ocorre na sílaba tônica frasal e é a que determina uma maior saliência na curva melódica é a mesma (CAGLIARI, 2011).

Para a marcação dos tons dos padrões entoacionais, Halliday propõe o uso de cinco tons melódicos. Pierrehumbert trabalha somente com os tons alto (H) e baixo (L), os quais são gerados por um sistema fonológico, como parte da gramática da língua. O não reconhecimento de um tom intermediário, nesse caso, o médio, traz alguns problemas para o modelo teórico da autora, pois, algumas vezes, ela acaba traduzindo um tom médio como um tom alto ou baixo, o que não ocorre na abordagem de Halliday.

A partir dos níveis melódicos, o autor propõe então o uso de cinco tons, numerados de 1 a 5 e obtidos por meio de uma abstração fonológica. O que diferencia o uso de um tom e não o de outro é o sentido que eles têm na língua, de modo que o sentido fica determinado pelo uso. Para Halliday, é possível desenvolver um sistema fonológico único, independentemente dos papeis desempenhados pela gramática. O tom 1, por exemplo, é o *exponed* de muitos termos diferentes no sistema gramatical do inglês. Mas, ao mesmo tempo, é foneticamente idêntico em todos os seus usos.

Pierrehumbert também parte de uma abstração fonológica, como ela mesma diz:

> *One main aim of this thesis is to develop an abstract representation for English intonation which makes possible to characterize what different patterns a given text can have, and how the same pattern is implemented on texts with different stress pattern.*[102]
> (PIERREHUMBERT, 1980, p. 10).

Contudo, para tratar dessa abstração, ela precisa investigar as regras que mapeiam essas representações fonológicas em representações fonéticas, o que não precisa ser feito no modelo do autor.

Halliday também propôs dois tons compostos, o 13 e o 53, o que não foi estabelecido no modelo da autora, o que mostra, segundo Cagliari (2011), que a autora não "prevê" uma variação no sistema. Em Pierrehumbert os tons compostos representam um espraiamento dentro de uma sílaba.

Quanto ao significado entoacional, Pierrehumbert faz a sua descrição da entoação a partir da observação das características da frequência fundamental (F0). Feita essa observação, o próximo passo seria comparar os usos do padrão de F0 que são fonologicamente distintivos. Porém, a autora não segue adiante, parando no primeiro

[102] Um dos objetivos principais desta tese é desenvolver uma representação abstrata para a entoação do inglês, que torna possível caracterizar quantos padrões diferentes um texto pode ter, e como o mesmo padrão é implementado em textos com diferentes padrões entoacionais. [tradução nossa].

método. Ela acredita que o sistema de investigação fonológica sobre o significado entoacional seja útil, mas não desenvolve essa questão em sua tese. Em alguns casos, a fim de ajudar o leitor a entender o contorno que está em discussão, ela faz uma descrição mais superficial do significado ou do uso de algum contorno específico. Para ela, não interessa se a frase é afirmativa ou interrogativa, mas, sim, o que o falante fez e não o que o falante quis dizer com o enunciado. Por isso, é que ela observa a curva entoacional, fazendo uma relação entre os fatos acústicos e fonológicos. Tal atitude teórica é decorrente do fato de Pierrehumbert trabalhar no interior da gramática gerativa (CHOMSKY, 1965), que separa o componente fonológico do sintático e do semântico, cada um operando com regras próprias.

Em contrapartida, Halliday faz uma relação entre os padrões entoacionais e as atitudes do falante. Diferentemente de Pierrehumbert, seu objetivo é descrever como o ouvinte percebe a entoação. Para ele, o significado de um tom está diretamente vinculado a outras escolhas gramaticais em um enunciado, questão essencial porque, de acordo com ele, toda vez que descrevemos uma determinada língua há uma preocupação com o significado, e todo contraste no significado pode ser justificado ou na gramática ou no léxico.

Para Pierrehumbert, atribuir padrões sintáticos aos tons não é possível, pois eles são extremamente dependentes do contexto. De acordo com ela, a impressão de que a entoação varia dependendo do contexto semântico de um enunciado ou dos objetivos do falante também não é significativa, pois, segundo ela, um mesmo padrão pode ser usado, por exemplo, tanto em perguntas interrogativas quanto em enunciados afirmativos. Halliday também concorda com essa afirmação, mas, para ele, os padrões entoacionais podem definir, sim, os aspectos semânticos, sintáticos e pragmáticos associados aos enunciados. Pierrehumbert não poderia levar esse tipo de questão em consideração, uma vez que ela está interessada no lado acústico da entoação, optando por estudar os valores quantitativos dos tons. Para Halliday, isso não se aplica, pois, assim como foi colocado por ele na citação anterior, se a língua é um conjunto de ideias e sons, quando a descrevemos, também estamos preocupados com o seu significado.

Dessa maneira, como diz Cagliari (2011), na descrição da gramática funcional, "[...] a interpretação semântica dos padrões entoacionais é uma consequência natural e não pode ser deixada para o componente semântico, como faz a gramática gerativa". (CAGLIARI, 2011, p. 5). É por isso, ainda segundo o autor, que todos os tons são descritos por seus padrões fonológicos.

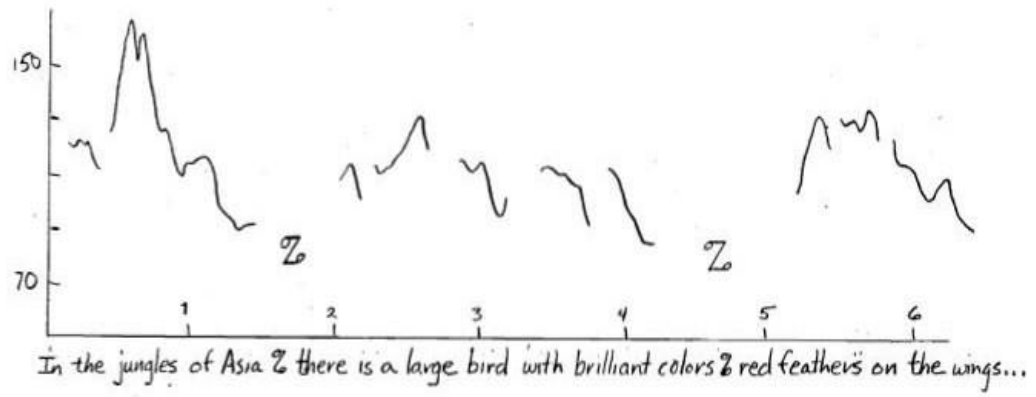
O modelo de Pierrehumbert exige uma marcação para o final do enunciado. O autor também considera o contexto terminal do enunciado como sendo muito importante para a

descrição entoacional. Para isso, através de símbolos visuais e linhas horizontais, ele marca o movimento da altura melódica para cada pé no final do enunciado. Para eles, a importância mencionada pode ser descrita pelas palavras de Pierrehumbert (1980, p. 26): *"[...] The end of the intonational phrase has distinctive tonal characteristics, apart from those attributable to the picth accents"*[103].

Assim como Halliday, Pierrehumbert também admite pausas entre as *intonationals phrase*. De acordo com a autora, seria uma espécie de pausa interna, ocorrendo onde há uma pausa não hesitacional ou onde se pode introduzir uma pausa sem que o contorno entoacional seja prejudicado.

Veja um exemplo, retirado por Pierrehumbert do *corpus* de Maeda (1976), de um contorno entoacional em que a frase entoacional de fronteira é anotada por pausas:

Figura 39- Frases entoacionais de fronteira anotadas com pausas



Fonte: Pierrehumbert (1980, p. 264).

Em Halliday, conforme Cagliari (2011), as pausas são incorporadas na descrição do ritmo e podem ocupar lugares silábicos nos pés ou, até mesmo, na duração de pés inteiros. As pausas ocorrem por uma questão sintática e semântica. Pierrehumbert, por sua vez, não leva esse ponto em consideração, já que, como ela mesma menciona, é um assunto que está fora dos objetivos da sua tese.

Em relação aos vocativos, denominados pelos dois autores de *tags*, Halliday os interpreta como sendo característicos do tom anterior ou como um tom que se constitui por si próprio. Já para o modelo de Pierrehumbert, eles são um problema. De acordo com ela, a dificuldade com as *tags expressions*, está no fato de elas não terem uma marcação de proeminência, ou seja, não apresentam *pitch accents*. Em outras palavras, segundo Cagliari

[103] "O final da frase entoacional tem características tonais distintivas, além daquelas atribuídas aos *pitch accents*" [tradução nossa].

(2012), expressões desse tipo apresentam um número considerável de pés métricos que não permitem a variação melódica nos *pitch accents*. [104]

Em relação à tessitura, Halliday mostra que esse efeito iguala realizações fonéticas diferentes no sistema, "[...] mas que apresentam um contraste estabelecido como sendo o mesmo no nível fonológico." (CAGLIARI, 2011, p.5). Nesse sentido, conforme Cagliari, uma sílaba pode apresentar uma variação da altura melódica de acordo com o valor da frequência fundamental, mas permanecer sendo alta ou baixa, por exemplo. Segundo o autor, isso ocorre porque, mesmo com a variação acústica, o que caracteriza um tom não é a realização fonética de um enunciado, mas o valor fonológico atribuído a ele.

O efeito da tessitura em Pierrehumbert pode ser comparado à sua discussão sobre as regras de *downstep*. De acordo com Cagliari, as variações no valor de H e L são um fato fonético e não fonológico e a tessitura é um fato fonético e fonológico. Entretanto, "[...] sua função não é re-ordenar os valores da F0, mas indicar marcas discursivas, como enunciados parentéticos ou enfatizados (por exemplo, em altura melódica de falseto)." (CAGLIARI, 2011, p. 6).

Outro aspecto que coloca as duas metodologias em confronto é a questão de uma (Pierrehumbert) trabalhar com a questão acústica da entoação e a outra (Halliday) trabalhar com a questão auditiva da entoação. Para a autora, sua opção em representar e mapear a entoação a partir da realização do contorno da frequência fundamental (F0) tem uma importância particular, devido às falhas apresentadas pelos outros métodos ao fazerem uma representação mais abstrata da entoação (embora a descrição auditiva não seja exatamente uma interpretação mais abstrata).

Entre as metodologias consideradas por ela como problemáticas para uma análise entoacional, ela traz, como um dos seus exemplos, justamente, o método de transcrição auditiva. Sobre isso, ela afirma que uma

> *[...] segmental transcription by ear has been a useful tool even though the resulting representation may not be a well-defined level in the theory. Transcription by ear of intonation is the source of so many gross errors in the literature that we feel it cannot be relied on.*[105]
> (PIERREHUMBERT, 1980, p. 13)

[104] Para solucionar casos desse tipo, Pierrehumbert acrescentou em sua gramática um acento frasal composto, isto é, um tom atribuído normalmente e outro atribuído às *tags*.

[105] [...] transcrição segmental feita auditivamente tem sido uma ferramenta útil, embora a representação resultante não seja um nível bem definido na teoria. A transcrição auditiva da entoação é fonte de tantos erros graves na literatura que nós achamos que esse não pode ser um método confiável. [tradução nossa].

Como se vê, ela apenas critica o modelo de transcrição auditiva e não caracteriza, significativamente, quais seriam esses "*gross errors*". Ela apenas justifica sua opinião afirmando que primeiro, ela precisa obter os seus dados para, depois, chegar a uma conclusão sobre os efeitos auditivos em uma análise entoacional. Porém, no decorrer da tese isso não é feito. Halliday, por outro lado, apesar de dispensar a análise acústica, não faz nenhuma crítica quanto ao modelo e nem sequer invalida o seu uso.

Portanto, a comparação apresentada mostrou que os dois modelos em discussão não estão tão distantes assim, pois, apesar das diferenças, as semelhanças são muitas. Como afirma Cagliari (2011, p. 17),

> Embora os modelos sejam diferentes formalmente, quando comparados, eles são surpreendentemente semelhantes ou, pelos menos, apresentam uma ponte que permite passar uma análise realizada em um modelo para uma análise no outro modelo.

Ladd (1996) também comenta que, apesar das aparências, há muita coisa no modelo de Pierrehumbert que pode ser comparada a outras análises, principalmente, segundo ele, com as análises de tradição britânica. Esse fato é justificado na teoria da autora no apêndice referente às figuras, em que ela mostra a possibilidade de combinação de vinte e dois tons.

Diante disso, Ladd faz uma correlação entre a análise de Pierrehumbert e a tradição de tons nucleares britânicos. Veja a seguir:

Figura 40- Correspondência entre Pierrehumbert (1980) e o estilo britânico dos tons nucleares

Pierrehumbert			British-style
H*	L	L%	fall
H*	L	H%	fall–rise
H*	H	L%	stylised high rise
H*	H	H%	high rise
L*	L	L%	low fall
L*	L	H%	low rise (narrow pitch range)
L*	H	L%	stylised low rise
L*	H	H%	low rise
L+H*	L	L%	rise–fall
L+H*	L	H%	rise–fall–rise
L+H*	H	L%	stylised high rise (with low head)
L+H*	H	H%	high rise (with low head)
L*+H	L	L%	rise–fall (emphatic)
L*+H	L	H%	rise–fall–rise (emphatic)
L*+H	H	L%	stylised low rise
L*+H	H	H%	low rise
H+L*	L	L%	low fall (with high head)
H+L*	L	H%	low fall–rise (with high head)
H+L*	H	L%	stylised high rise (low rise?) with high head
H+L*	H	H%	low rise (high range)
H*+L	H	L%	stylised fall–rise ('calling contour')
H*+L	H	H%	fall–rise (high range)

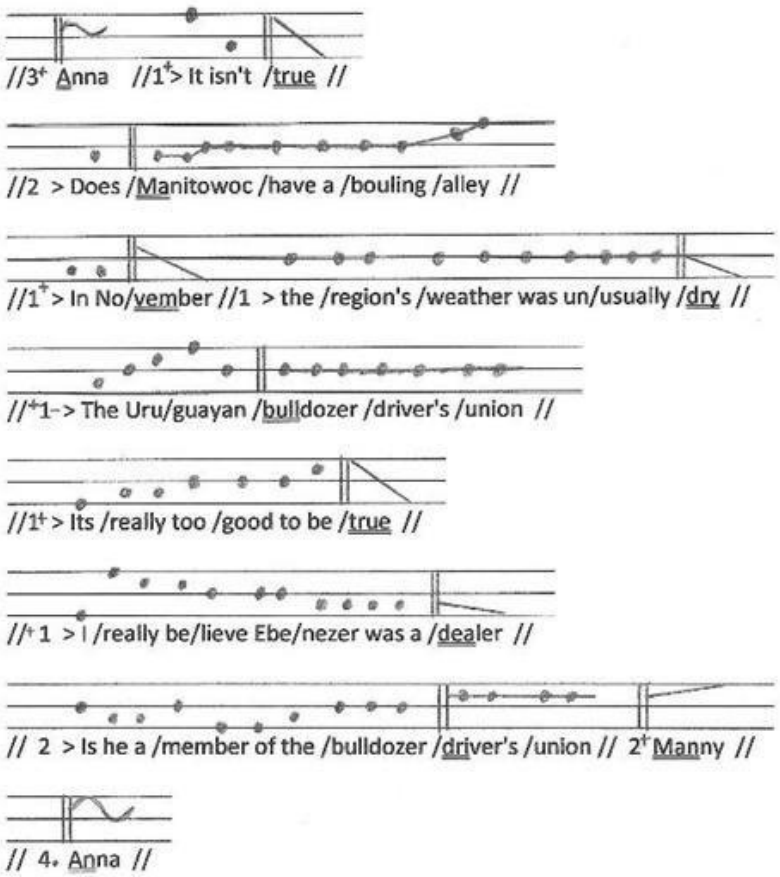
Fonte: Ladd (1996, p. 82).

Nessa figura[106], a primeira coluna mostra as vinte e duas possibilidades de tons no modelo de Pierrehumbert e a segunda coluna apresenta a correspondência de acordo com o modelo britânico.

Cagliari (2011) também fez uma interpretação parecida com a proposta por Ladd. Em sua análise, foram feitas duas correlações. Na primeira, identificada pela figura 41, que segue, os exemplos de Pierrehumbert foram analisados de acordo com o modelo de Halliday (1970). E, na segunda, exemplificada na figura 42, foi feito o contrário.

[106]Entre outros comentários, Ladd menciona que o modelo de Pierrehumbert gerou 28 possibilidades de combinações entre os tons, mas, por razão de neutralização das regras de implementação, restaram apenas 22 possibilidades.

Figura 41- O modelo de Pierrehumbert descrito na metodologia de Halliday por Cagliari



Fonte: Cagliari (2011, p. 16).

Figura 42- O modelo de Halliday descrito na metodologia de Pierrehumbert por Cagliari

Arthur and Jane left for Italy this morning H* H* H* H* L L%
Do they take the car when they go abroad H% H* L* H⁻ H%
Arthur likes to have it while he's there H* H* L*+ H⁻ L⁻ H%
They didn't take the car last time they went H% H* H* H⁻ H⁻+L* L⁻ H%
I didn't know they'd ever been to Italy L* H* H* L⁻+H* L⁻ L%
Arthur's been there twice in the last year or so H* H* H*+L⁻ L* L*+ H⁻ L⁻ H%
He's never taken Jane on any of this visits though L* L* L⁻+H* H⁻ L*+ H⁻ L⁻ H%

Fonte: Cagliari (2011, p. 17).

A análise acima corresponde aos tons primários e secundários. Segundo Cagliari, os tons compostos foram os mais difíceis de localizar, já que não se tem os valores da F0 para

os exemplos de Halliday. Nessa interpretação, eles são correspondentes aos tons 3, 4 e 5 de Halliday, sendo neutralizados para os demais. As pretônicas, por serem previsíveis, não foram representadas.

Sendo assim, fazendo a conversão da descrição de Halliday para a de Pierrehumbert, temos, segundo o autor, a seguinte correspondência:

Figura 43- Conversão dos tons primários simples e compostos do modelo de Halliday para a forma fonológica de Pierrehumbert

Tom 1:	H* L L%	
Tom 2:	L* H⁻ H%	
Tom 3:	L*+ H⁻ L⁻ H%	
Tom 4:	H⁻+ L* L⁻ H%	
Tom 5:	L⁻ + H* H⁻ H%	
Tom 13:	(H*+ L⁻) *prenuclear*	(L*+ H⁻ L⁻ H%) *nuclear accent*
Tom 53:	(L⁻ + H*) *prenuclear*	L*+ H⁻ L⁻ H%) *nuclear accent*

Fonte: Cagliari (2011, p. 17).

Assim, diante da comparação feita entre os dois modelos teóricos, podemos dizer que, mesmo com suas diferenças e singularidades, é perfeitamente possível correlacionar uma abordagem e outra, além de mostrar a viabilidade deste trabalho.

6 Materiais e Métodos

6.1 Procedimentos: considerações iniciais

A seguir, apresentamos os procedimentos realizados para o desenvolvimento desta pesquisa. Primeiro, descrevemos como foi feita a seleção do *corpus* e do informante, bem como suas características principais. Posteriormente, seguimos com as informações sobre a análise acústica e a metodologia utilizada para a anotação da frequência fundamental (F0) para as sílabas vozeadas e desvozeadas. Por fim, descrevemos o procedimento utilizado durante a análise da tessitura e do padrão entoacional dos enunciados.

6.1.1 Dados da leitura

Os dados a serem analisados na presente dissertação fazem parte da leitura de um pequeno trecho do livro *História sem fim* de Michael Ende (sd) [107]. Esse fragmento foi selecionado porque a história apresenta um conteúdo semântico que facilita a presença de variações prosódicas na fala, além de apresentar vários tipos de frases, tais como: assertivas, interrogativas, exclamativas, suspensivas, entre outras. Optamos por não utilizar frases isoladas, mesmo que atualmente seja um tipo de *corpus* muito utilizado nos laboratórios de análise acústica. Essa opção deveu-se ao fato de que, às vezes, elas dão margem a uma intuição do falante diferente da esperada em experimentos programados com a investigação de fatos específicos, dificultando a interpretação semântica dos enunciados. Além disso, segundo Cagliari (2010, p. 3), um problema da utilização da fala "espontânea", não planejada de antemão, é que nem sempre ocorrem frases completas. Já em relação à fala "espontânea de laboratório", planejada, é que ela não se constitui de frases isoladas. Também optamos por não estudar a fala comum devido às suas dificuldades, como, por exemplo, a falta de funções comunicativas, interativas e emotivas, causadas por gravações de pesquisa.

A leitura de pequenos fragmentos, por outro lado, tem a vantagem de facilitar a investigação do ritmo, da tessitura e das pausas, além de possibilitar uma melhor interpretação semântica do texto. Nesse tipo de gravação, também há a vantagem de o falante sentir-se mais confortável e, em consequência, utilizar uma linguagem mais natural. Contudo, temos consciência de que a leitura de pequenos textos ou mesmo a fala espontânea mais longa nem sempre traz à tona muitos dos fatos relevantes do sistema entoacional de uma língua, mas apenas alguns deles.

[107] O fragmento do texto utilizado pode ser visto nos anexos com sua respectiva transcrição ortográfica e fonética.

6.1.2 Dados do informante

O material sonoro selecionado foi lido por um homem de 65 anos, falante do dialeto paulista da cidade de Araraquara e com instrução superior completa. Atualmente, o falante atua como professor do ensino superior A opção por essa região deve-se, unicamente, à facilidade de acesso entre o pesquisador e o falante. Além disso, o interesse desse estudo não é observar as diferenças entoacionais regionais, mas, por outro lado, correlacionar dados acústicos e auditivos, como mencionado.

Antes de darmos início à gravação, o leitor foi instruído a fazer, primeiramente, uma leitura silenciosa da obra. Em seguida, para dar início à gravação, ele também foi orientado a não fazer uma leitura lenta, rápida ou pausada e, sim, a mais natural possível. Optamos por selecionar somente um falante devido à extensão do trabalho, uma vez que o texto terá uma análise, como dito, a partir do modelo proposto por Halliday (1970) e Cagliari (1982/2007), sendo segmentado em trechos correspondentes aos grupos tonais, pausas, pés, sílaba tônica saliente e de acordo com o modelo de Pierrehumbert (1980), além da análise dos significados associados aos padrões entoacionais. Além disso, trata-se de um trabalho de natureza qualitativa, que não tem por finalidade enumerar os dados, mas sim descrever os fatos.

6.1.3 Dados da análise acústica

Com relação à análise acústica, o texto foi segmentado em trechos correspondentes aos grupos tonais, para facilitar a análise, que não comporta longos trechos gravados. Os dados foram processados através do programa PRAAT, desenvolvido por Paul Boersma and David Weenink, do Institute of Phonetic Sciences (University of Amsterdam), disponível gratuitamente em http://www.fon.hum.uva.nl/praat/. A gravação foi feita em uma taxa de amostragem de 22050 Hz.

A gravação do *corpus* foi processada em um ambiente acusticamente adequado, via microfone unidirecional, modelo Microsoft LifeChat LX-3000. Tanto a gravação quanto o teste perceptivo foram feitos na própria Universidade (UNESP / FCL-AR). Como não dispomos de um laboratório acusticamente isolado, o ambiente escolhido para a gravação foi o que apresentou o mínimo de ruído possível, para não interferir na qualidade do áudio.

Posteriormente, o texto foi segmentado em 15 enunciados correspondentes a 22 grupos tonais. Também foi feita a transcrição fonética[108] e ortográfica de cada um dos enunciados. Em seguida, as frases foram descompactadas e subdividas em três níveis: 1)

[108] A transcrição fonética seguiu o modelo do IPA.

segmentos, 2) sílabas e 3) palavras[109]. Esse método foi feito a partir do *TextGird* do PRAAT e através da observação das informações carreadas pelo espectrograma.

Para a análise do *pitch*, foi marcado o valor da frequência fundamental (F0) para cada sílaba, localizando-a no centro da duração das vogais, isto é, a partir do seu ponto medial. Para a realização dessa etapa, foi preciso trabalhar com dois métodos diferentes: 1) um referente às sílabas vozeadas e 2) outro para as sílabas desvozeadas. Para as sílabas vozeadas, utilizamos a função CTRL + 0 do PRAAT. Com esse comando, foi possível analisar, com pequenas variações, o ponto medial, que, selecionado, vai para o cruzamento de zero mais próximo do ponto médio do segmento. O ponto medial estabelecido depende de como o início e o fim do segmento foi determinado. Essa segmentação é feita por meio de dois procedimentos:

a) uma observação auditiva, levando em consideração as possibilidades articulatórias e as realizações fonéticas da língua;

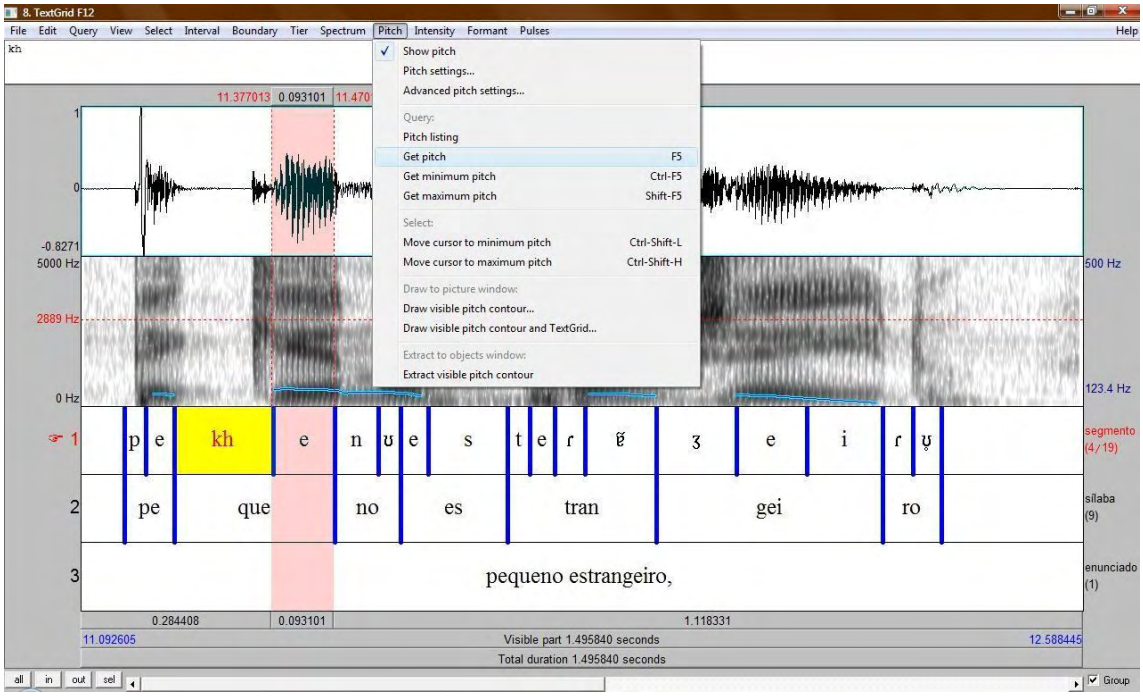
b) uma análise visual, que interpreta parâmetros acústicos atribuíveis ao segmento, às suas transições e aos vizinhos. (transições).

Para as sílabas desvozeadas, foi feita a seleção da sua parte vozeada e, em seguida, utilizamos a função *Picth > get picth* do software. Não foi possível utilizar o mesmo método porque, muitas vezes, o ponto médio mais próximo do cruzamento de zero estava na vogal desvozeada.

Veja abaixo um exemplo dos dois procedimentos utilizados. O primeiro ilustra o método seguido para as sílabas vozeadas e o segundo, as sílabas que apresentaram sons desvozeados e vozeados.

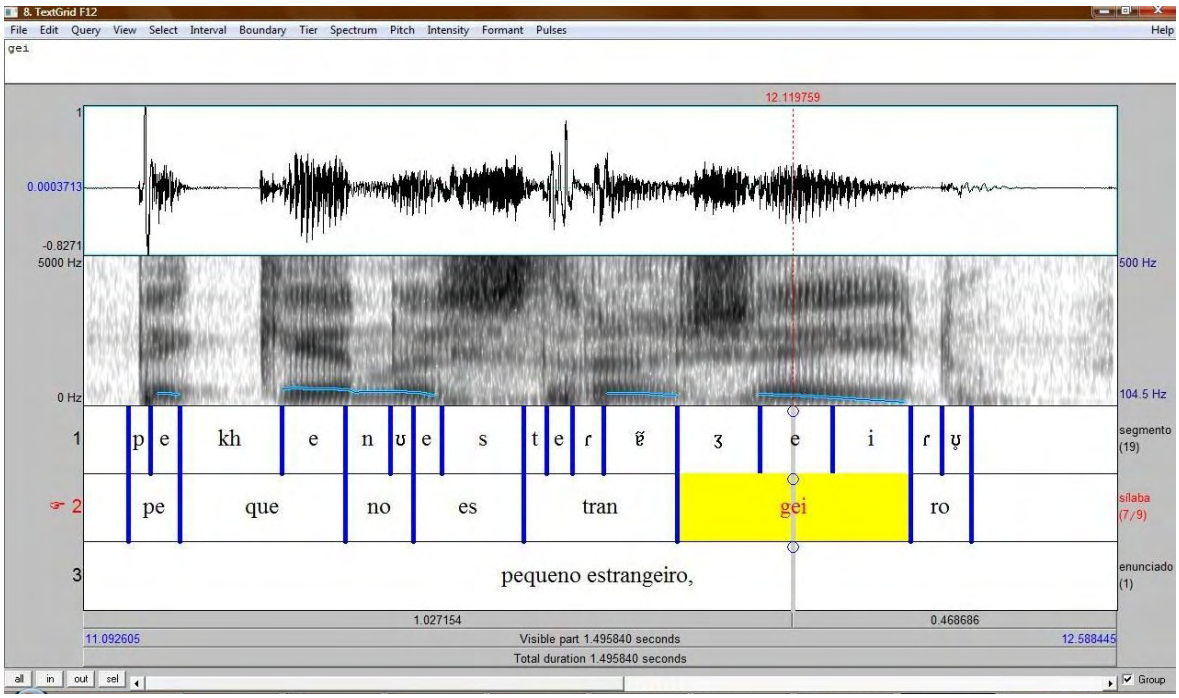
[109] Essas segmentações podem ser vistas nas figuras 44 e 45 apresentadas a seguir.

Figura 44- Representação do método de análise a partir do ponto medial das sílabas vozeadas do enunciado 12 – "pequeno estrangeiro"



Fonte: própria.

Figura 45 - Análise da frequência fundamental no ponto medial das sílabas desvozeadas, a partir da função: *pitch* > *get pitch*, no enunciado 12 – "pequeno estrangeiro"



Fonte: própria.

Na figura 44, mostra-se, na faixa rosa, a seleção do segmento vozeado na sílaba *que*. Seguido disso, selecionamos a função *pitch* e depois *get pitch* (mostrada na figura). Através desse comando, abre-se uma tela com a informação da frequência fundamental. Na figura 45, o ponto medial vem representado pela linha pontilhada, em vermelho, na sílaba *gei*, do enunciado. A partir dessa função, a frequência fundamental é mostrada, em azul, do lado direito da tela. Feito isso, os resultados das duas medições foram anotados em tabelas no Excel.

Vale lembrar que havia outros métodos para se medir a frequência fundamental (F0), como, por exemplo, por meio da fronteira do segmento. Porém, entendemos que, nesse momento, há alguns fatores que influenciam e dificultam a segmentação, como a co-articulação. De acordo com Cagliari (1982), o que ocorre nesse momento são modificações fonéticas de vizinhança, devido à interferência das características articulatórias de um som sobre o outro. Essas influências podem ocorrer tanto no início quanto no fim de um elemento, o que é natural da fala. Não se pode presumir quais segmentos fonéticos apresentam propriedades acústicas ou articulatórias independentes, uma vez que elas variam de acordo com o contexto. (FUJIMURA; ERICKSON, 1999). Por isso, optamos por não fazer anotações pelo início ou fim de fronteira, uma vez que poderia haver uma interferência nos resultados da frequência fundamental (F0).

Em relação à análise auditiva da F0, segundo o modelo de Halliday (1970), vale salientar que ela foi feita diretamente pelo pesquisador e conferida pelo orientador desta dissertação, sem que os resultados fossem modificados. Os dados não foram julgados por outras pessoas, já que esse tipo de análise requer um treinamento cuidadoso, procedimento que levaria muito tempo e seria inviável para o desenvolvimento desta pesquisa.

6.2 Tessitura: procedimentos para análise

A tessitura, neste trabalho, foi analisada a partir da maior e menor frequência nas sílabas, que foram anotadas a partir do recurso do Praat: 1) *pitch>get minimum pitch* (para a menor frequência) e 2) *pitch>get maximum pitch* (para a maior frequência). Tanto as sílabas tônicas quanto as sílabas átonas, de cada grupo tonal, foram marcadas, mas foram utilizadas somente as de maior e menor frequência. As sílabas desvozeadas não foram consideradas, já que não apresentam variação de F0.

Optamos por utilizar esse recurso do *software* porque, manualmente, existe uma grande possibilidade de variação, uma vez que os pontos anotados entre uma sílaba e outra, na maioria das vezes, não irão coincidir. Nesse sentido, é mais confiável utilizar a

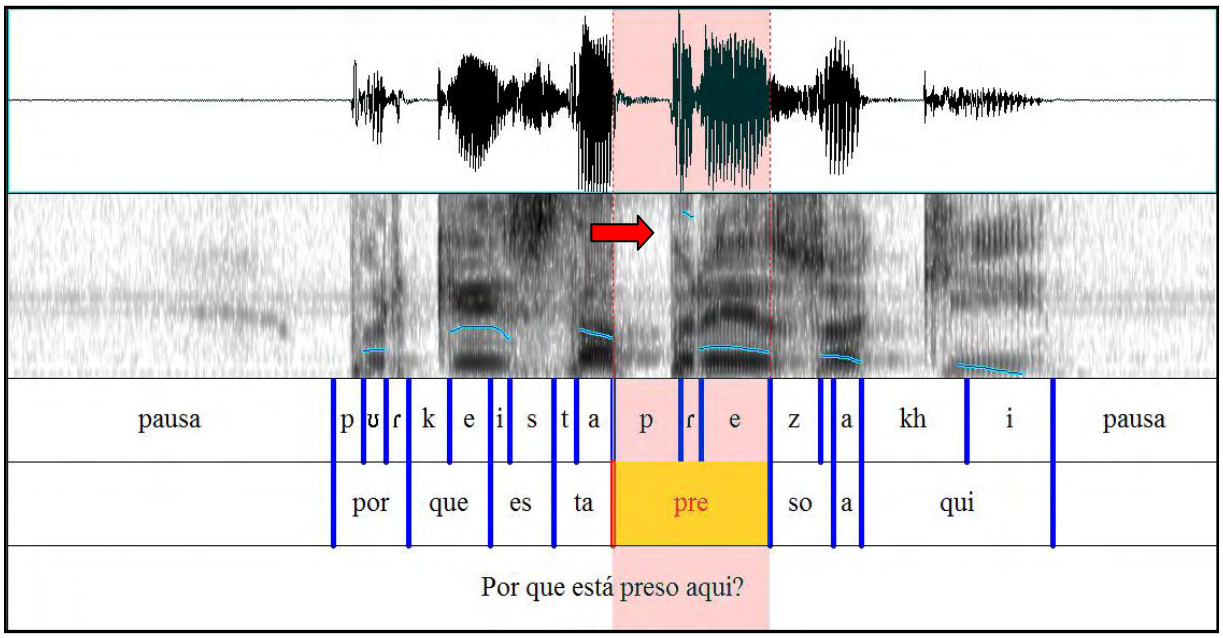
ferramenta fornecida pelo programa, além de ser uma forma de padronização. As opções manuais que podem ser feitas no *software* não foram completamente descartadas, mas, sim, utilizadas como um recurso.

Também vale salientar as razões que justificam o não uso das frequências fundamentais marcadas nas figuras para a análise acústica. Primeiramente, como descrito, essas frequências foram anotadas a partir de dois métodos: 1) ou pelo ponto medial (para as sílabas vozeadas) ou, através da função *Picth > get picth* (para as sílabas desvozeadas). Assim, se optássemos por essas frequências, não teríamos uma fidelidade quanto ao objetivo dessa etapa do trabalho, que é, justamente, anotar a maior e a menor F0 das sílabas em cada grupo tonal. No primeiro método, as sílabas de todos os grupos tonais foram anotadas, mas não com essa finalidade.

No que diz respeito às frequências (maior e menor) de cada sílaba elas foram anotadas em tabelas via Excel. Isso foi feito seguindo a divisão dos enunciados em grupos tonais, para a análise no modelo de Halliday (1970) e Pierrehumbert (1980).

Vale mencionar que, em algumas sílabas, ocorreram problemas de filtragem gerados pelo PRAAT. Esses problemas aconteceram somente para as marcações da maior F0, em que os valores gerados pelo programa eram impossíveis, como por exemplo, o caso da sílaba pre na palavra preso do enunciado E3, em que o valor da maior frequência dado pelo programa foi de 463,4, o que seria completamente improvável. Veja na figura abaixo o que gerou esse erro.

Figura 46- Exemplo de um problema de filtragem gerado pelo PRAAT do enunciado 3 – "Por que está preso aqui?"



Fonte: própria.

O problema vem indicado pela seta em vermelho, na própria figura. Isso pode ter ocorrido devido a uma interferência na gravação, como, por exemplo, um ruído. A solução encontrada para esse tipo de questão foi utilizar outra ferramenta do PRAAT: 3) *pitch>pitch listing*, que abre uma janela com todas as frequências possíveis para essa sílaba, permitindo, assim, a escolha da maior possível entre elas.

6.3 O padrão entoacional: procedimentos para análise

Para analisarmos quais foram os tipos de padrões entoacionais presentes no fragmento do texto *História sem fim*, procedemos da seguinte forma:

1. Primeiramente, o texto foi segmentado e dividido em 15 frases;

2. Seguido a isso, foi feita a análise auditiva no modelo de Halliday (1970) e Cagliari (1982/2007), com as respectivas informações sobre a formação dos GTs, dos pés, da tonicidade, da sílaba tônica saliente (STS) e do tom;

3. Posteriormente, foi feita uma análise acústica a partir do método descritivo de Pierrehumbert (1980);

4. Tendo em vista a análise nesses dois modelos teóricos, foi feita uma correlação entre eles, a fim de converter a forma fonológica dos tons de um modelo para o outro e estabelecer uma correspondência entre eles. Por meio desse procedimento, também

fizemos uma comparação com o trabalho de Cagliari (2011) [110], em que o autor apresentou, brevemente, a conversão para os tons primários (simples e compostos). A partir disso, nossa intenção é ampliar os resultados obtidos pelo autor, devido à grande incidência de tons secundários em nossos dados. A tabela do padrão entoacional proposta por Ladd (1996) [111] para o inglês britânico também será levada em consideração, para reforçarmos os resultados dessa comparação. Vale mencionar que não será possível transpor a forma fonológica de todos os tons propostos por Halliday e Cagliari para a de Pierrehumbert, já que estamos trabalhando apenas com os tons encontrados em nosso *corpus*.

Sendo assim, a partir dos procedimentos elucidados seguimos, no próximo item, com as análises dos dados e descrição dos resultados obtidos.

[110] Conferir item 4.3.2 "Semelhanças e diferenças".
[111] Para conferir a representação feita por Cagliari (2011) e a tabela comparativa proposta por Ladd (1996), veja as figuras (41, 42, 43) do item 5.3.1 "Semelhanças e diferenças" da dissertação.

7 Análise dos dados e resultados

7.1 Análise da tessitura

A variação da tessitura foi um fato esperado nesta dissertação, já que, por se tratar de uma história e de um processo de leitura, os acréscimos significativos foram muito importantes para a construção do discurso, além disso, o fragmento apresenta um contexto emotivo, o que permite ainda mais essas variações.

Na tabela 1 que segue, vêm descritos os valores de frequência fundamental mais alta e mais baixa das sílabas[112]. As colunas identificam, respectivamente, os enunciados, o grupo tonal, as sílabas de maior e menor frequência, bem como sua respectiva soma e média aritmética.

[112] As sílabas mostradas na tabela abaixo podem ser vistas no item 7.3.1.2 referente à análise dos enunciados em que apresentamos os 15 enunciados analisados.

Tabela 1- Descrição das sílabasque apresentaram maior e menor F0, com sua respectiva soma e média aritmética

Enunciado	Grupo tonal	Sílaba com Hz mais alta	Hz mais alta	Sílaba com Hz mais baixa	Hz mais baixa	Soma das frequências	Média
E1	1	quem	160,6	ce	109,7	270,3	135,1
E1	2	per	125,0	u	75,8	200,8	100,4
E1	3	mou	135,6	pa	80,1	215,7	107,8
E2	1	sou	154,8	gi	75,8	230,6	115,3
E2	2	bi	142,9	men	81,5	224,4	112,2
E3	1	que	195,1	qui	85,9	281,0	140,5
E4	1	se	164,7	es	112,8	277,5	138,7
E4	2	se	154,0	ram	124,9	278,9	139,4
E5	1	quem	170,3	quem	77,3	247,6	123,8
E6	1	les	191,4	rren	94,0	285,4	142,7
E7	1	on	193,0	ram	75,5	268,5	134,2
E8	1	gi	127,4	deu	76,1	203,5	101,7
E9	1	lhou	152,9	cha	75,6	228,5	114,2
E10	1	pois	148,0	de	104,1	252,1	126,0
E10	2	di	105,7	di	92,6	198,3	99,1
E11	1	qui	161,9	vo	98,2	260,1	130,0
E12	1	que	124,5	gei	84,5	209,0	104,5
E13	1	ci	157,0	não	111,8	268,8	134,4
E13	2	pa	140,0	is	85,4	225,4	112,7
E14	1	que	165,7	cu	94,1	259,8	129,9
E15	1	u	147,2	a	93,8	241,0	120,5
E15	2	ca	128,3	be	82,0	210,3	105,1

Fonte: própria

A partir dessas anotações foi obtido o valor da tessitura na fala do leitor durante a leitura do fragmento. Veja o gráfico a seguir:

Gráfico 1- Representação gráfica da tessitura, alcançada a partir da média da distribuição fundamental nas sílabas dos grupos tonais



Fonte: própria

O gráfico acima mostra a média aritmética da distribuição da frequência fundamental nas sílabas de cada grupo tonal. O eixo vertical indica a variação da F0, que aparece especificada em cada coluna da tabela 1. O eixo horizontal, por sua vez, mostra os grupos tonais.

Através da análise do gráfico, observa-se uma variação de tessitura em uma escala melódica que oscila de 100 a 140 Hz. Essa variação é típica de uma voz masculina adulta, como afirma Ladefoged (1973, p. 163): "*For a male voice the frequency of the vocal cord vibrations in speech may be between 80 and 200 Hz*"[113], caracterizando, assim, uma voz mais grave. A voz feminina adulta, por outro lado, segundo Cagliari (2007), pode chegar a uma frequência média de 200 Hz por segundo.

Pelo gráfico, observa-se também que o falante utiliza-se de uma tessitura mais alta logo no início do enunciado. Isso aconteceu porque o texto já começa com uma pergunta que é enfatizada pelo leitor: "Quem é você?". É como se o leitor quisesse, de imediato, mostrar o início de um conflito na narrativa, colocando em foco seus objetivos expressivos através do pronome interrogativo "Quem". Posteriormente, nos grupos tonais dois e três, há

[113] "Para a voz masculina a frequência de vibrações das cordas vocais na fala deve estar entre 80 e 200 Hz" [tradução nossa].

uma diminuição da tessitura. Isso ocorre porque o texto apresenta dois personagens, Atreiú e Gmork e mais um narrador, o que permite as frequentes incidências de diálogos entre eles e, outras vezes, a substituição da voz de Atreiú e Gmork para a entrada do narrador. Nesse momento da leitura, o falante tinha a função de narrar a história, precisava sair desse contexto de destaque, justamente para mostrar que já não era mais Atreiú quem falava e, sim, o narrador.

Nota-se também que o falante se utiliza das frequências mais baixas, não tão baixas quanto às do narrador, quando está no papel do personagem Gmork (o lobisomem). No contexto da história, Gmork é um personagem triste, esquecido e maltratado pelas pessoas que o deixaram de lado e isso pode ter levado o leitor a optar pelas tessituras mais baixas.

No corpo do fragmento, as tessituras caracterizam-se por estarem mais elevadas, com pequenas variações. Nessa parte da história, ocorre uma mudança de ação na narrativa, os diálogos são mais constantes e há uma omissão do narrador. A tessitura começa a se elevar na ocasião em que Atreiú está prestes a descobrir quem prendeu Gmork, com uma frequência de 123,8 Hz. A partir dessa descoberta, o leitor passa a dar um destaque maior para as falas de Gmork, a fim de enfatizar a indignação do animal diante da crueldade sofrida, havendo então uma elevação na tessitura.

Mais para o fim da história, o narrador reaparece e a tessitura começa novamente a diminuir. Agora, a ação não gira mais em torno de Gmork, mas, sim, de Atreiú. Há uma inversão nos papéis. Agora, é o lobo quem quer saber das intenções de Atreiú: “Você não é daqui [...] Não é desta cidade nem deste país. O que procura?”. Essa desconfiança de Gmork é identificada pela elevação das frequências de 130,0 Hz, 134,4 Hz e 129,9 Hz, respectivamente nos grupos tonais identificados. As frequências mais baixas que aparecem entre esses GTs dizem respeito aos enunciados encaixados que, no contexto do diálogo, não tiveram importância significativa, isto é, no conjunto dos acontecimentos, essas informações não foram tão relevantes para o leitor.

Com a chegada do fim do enunciado, diferentemente do seu início, a tessitura caracteriza-se por estar bem mais baixa. Esse fato, além de ser próprio, novamente, da entrada do narrador no discurso, também ajuda a caracterizar, para quem ouve a história, a atitude do personagem que talvez por medo ou precaução, não soube se deveria ou não responder o questionamento anterior, pois lobos são animais traiçoeiros. É como se o narrador quisesse evidenciar a desconfiança de Atreiú ao ser indiferente e deixar a pergunta sem resposta: “Atreiú baixou a cabeça” (105,1 Hz). A queda da tessitura no final do

enunciado pode ser relacionada ainda com uma intenção implícita do leitor de mostrar que a sua leitura está chegando ao fim.

Essas variações de tessitura no decorrer do texto nos mostram a sua relevância para a organização da estrutura discursiva, bem como sua relação com a entoação. Por meio da tessitura, é estabelecido um padrão de leitura, uma estruturação do texto, de modo que é possível evidenciar onde o texto começa e termina e, ainda, os seus conflitos. Além disso, outra evidência do papel organizacional desse elemento prosódico é a alternância de frequência para cada um dos integrantes da história. Pode-se dizer que o personagem Atreiú identifica-se por uma tessitura mais alta, enquanto Gmork apresentou uma tessitura que varia entre o tom médio e baixo, apresentando apenas pequenos picos, que estão relacionados à sua curiosidade em saber a origem do pequeno estrangeiro. Outro fato curioso é que essa tessitura identifica a condição do personagem na trama, desarmado, impossibilitado de agir diante das circunstâncias que o afligem. É importante deixar bem claro que Gmork é um personagem, que, nesse momento da narrativa, foge da sua condição de crueldade. A final de contas, ele é um "lobisomen", um animal, na maioria das vezes, traiçoeiro e perverso. Assim, pode-se dizer que a opção por uma tessitura mais baixa também seria uma forma de colocar em suspense a personalidade de Gmork.

O narrador, por outro lado, durante toda a leitura, utilizou uma tessitura baixa, revelando seu distanciamento e neutralidade frente aos conflitos narrativos, deixando claro o seu papel de apenas notificar os acontecimentos. É como se o narrador colocasse as ideias entre parênteses, mais especificamente, em suspense, deixando as suposições por conta do leitor. Esse fato justifica a importância do narrador no texto, uma vez que ele é o responsável por revelar as atitudes e expressões dos personagens. No fragmento analisado, essa reflexão mostra que as tessituras mais baixas também são importantes, pois, mesmo colocando as ideias em uma posição marginal, podem revelar as intenções e comportamentos dos personagens, dando uma informação a mais sobre os seus atos. Sobre essa afirmação, Cagliari (1992a, p. 57) diz:

> Durante um texto, de repente, o falante começa a falar num tom baixo, para fazer ver ao seu interlocutor que aquilo que ele está dizendo com tessitura baixa deve ser interpretado com algo diferente em valor com relação ao resto que se está enunciando: às vezes se pretende colocar algo com valor diminuído, ou como se diz, parenteticamente; às vezes, o objetivo é justamente oposto: destacar um argumento que, por ser óbvio, vem falado em tessitura mais baixa, mas cujo valor argumentativo é justamente dar mais força ao que se está dizendo num tom normal de fala.

7.1.2 A tessitura no nível tonal a partir do modelo de Halliday

A partir da análise dos tons, também se pode chegar à variação da tessitura. Para isso, por meio da interpretação auditiva, os valores acústicos foram classificados de acordo com a descrição utilizada e distribuídos em: baixo, meio baixo, médio, meio alto e alto. Esses valores foram coletados tendo em vista a metodologia empregada neste trabalho: ou pelo valor na posição central de sua duração (para as sílabas vozeadas) ou pelo valor gerado pelo comando *get picth* do PRAAT (para as sílabas desvozeadas), a depender do contexto. Nesse caso, não foi preciso fazer uma diferenciação entre as sílabas de maior e menor F0, já que o objetivo dessa etapa consistiu em validar a eficiência da metodologia proposta e, em consequência, provar a correlação entre uma análise acústica e auditiva.

Abaixo, veja a comparação feita entre os valores acústicos da F0 para cada sílaba e a atribuição de um nível tonal (tom) a elas. Os níveis auditivos que aparecem com realce (em vermelho) na tabela referem-se aos valores auditivos que variaram (para mais ou para menos) em relação à classificação proposta, como discorreremos mais adiante.

Tabela 2- Distribuição da variação melódica das sílabas em Hz de acordo com a classificação auditiva

Baixo	Meio baixo	Médio	Meio alto	Alto
85,3	122,9	124,3	150,0	191,9
105,0	117,2	133,2	148,6	182,8
93,3	116,9	134,4	139,5	176,9
84,2	115,0	129,3	146,5	153,0
92,2	119,2	130,9	145,5	161,9
95,7	110,9	129,4	143,8	154,5
110,0	119,0	139,4	137,0	152,3
79,9	123,0	110,3	149,3	175,6
112,0	121,9	141,7	147,6	186,8
110,7	112,8	125,3	158,2	188,0
81,2	114,4	140,4	148,4	184,3
86,4	119,4	124,4	145,1	160,0
89,0	119,6	138,3	143,9	150,7
105,4	123,3	132,6	142,7	181,7
101,0	117,3	136,3	148,8	166,4
109,8	120,9	126,3	145,5	153,2
97,6	112,7	139,3	145,2	164,8
100,2	123,3	134,7	146,7	155,4
99,2	115,2	130,2	136,5	153,5
106,3	124,2	126,8		
101,6	115,5	131,1		
127,6	117,4	135,1		
	120,5	127,3		
	115,8	136,4		
	123,3	130,4		
	114,4	139,9		
	110,2	130,5		
	109,2	129,0		
	119,0	133,4		
	123,1	129,1		
	116,7	132,6		
	122,6	134,2		
	116,7	133,9		
	122,2	125,7		
	124,3	132,1		
		133,8		
		133,1		
		131,9		
		129,7		

Fonte: própria.

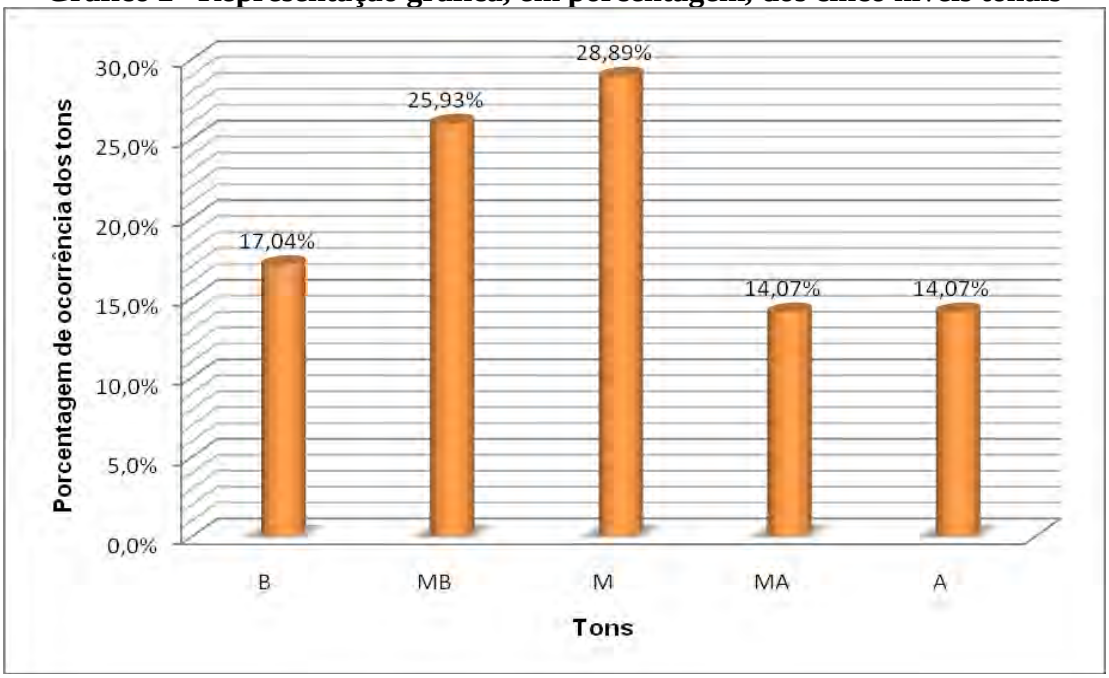
Por meio desses valores, foi possível quantificar o número de ocorrências para cada um desses tons, além de estabelecer o intervalo de frequência mais baixo e o mais alto (Hz ou cps), para a definição da tessitura desse informante. Em relação à primeira afirmação, observe a tabela 3 a seguir e, posteriormente, o seu gráfico ilustrativo.

Tabela 3- Quantidade e porcentagem de ocorrência dos tons

Baixo	Meio baixo	Médio	Meio alto	Alto	Total
23	35	39	19	19	135 Tons
17.04%	25.93%	28.89%	14.07%	14.07%	99.99 %

Fonte: própria.

Gráfico 2– Representação gráfica, em porcentagem, dos cinco níveis tonais



Fonte: própria.

Levando em consideração a tabela e o gráfico acima, verifica-se que, para 135 tons (100%), o tom alto e o tom meio alto foram os que apresentaram menor número de ocorrência (14,07%) cada um, seguido do tom baixo (17,04%), meio baixo (25,93%) e médio (28,89%).

Os tons de qualidade mais baixa (B e MB) aparecem com mais frequência depois de uma pausa e início de enunciado. Posteriormente, no decorrer da leitura, eles vão se ajustando ao longo do componente pretônico. Na leitura desse fragmento, houve a predominância desses dois tons (aproximadamente 43% de ocorrência para os dois). Isso ocorreu porque, em um contexto geral, a obra não apresentou momentos de muito conflito. Os personagens (Atreiú e Gmork) estavam apenas se conhecendo e os diálogos, em sua maioria, eram tranquilos. Outros dois fatores, mencionados no item anterior, também influenciaram a maior incidência desses tons. Um deles está relacionado à presença do narrador que, para se distanciar dos dois personagens, utilizou-se de um tom mais baixo. O

outro, tem relação com o desenvolvimento da história que, de um modo geral, é construída em um contexto de suspense.

Os momentos mais expressivos da obra são justificados pelos tons de qualidade mais alta (tom alto e meio alto), com uma ocorrência de aproximadamente 28% para os dois, o que representa uma diferença de 15% para os tons de qualidade mais baixa. Os tons baixo e alto também podem representar o ponto típico do final do componente tônico, ocorrendo, na maioria das vezes, na sílaba tônica saliente que representa o foco.

Por sua vez, os tons médios, em relação aos quatro tons analisados, foram os mais frequentes (28, 89% dos casos). Em geral, eles representam um tom de passagem, servindo de base para que o falante modifique a entoação, seja indo para cima ou para baixo. Essa característica é justificada se levarmos em consideração que, do tom baixo ao tom médio, a variação melódica foi ascendente e, do tom alto ao tom médio, foi descendente. Devido a esse caráter ascendente/descendente, os tons meio altos possuem uma distribuição aproximada de (14,07%) e os tons meio baixos de (25,93%).

O tom médio também é característico dos enunciados suspensivos, isto é, incompletos. Os enunciados suspensivos indicam ao interlocutor que o pensamento será completado com a frase que vem logo em seguida. No texto, foram exemplos desse tipo de tom: Esqueceram-se de mim (no enunciado E4); Depois de um longo silêncio (no enunciado E10); Não é desta cidade (no enunciado E13); Atreiú (no enunciado E15). Todos eles remetem a um fato suspenso, ou seja, a um pensamento que precisa ser completado, pois, no contexto da história, ficam sem sentido.

A tessitura também pode ser analisada com mais detalhes através da observação da distribuição da variação melódica da F0 em função dos cinco tipos de tons da pauta entoacional. Veja o resultado na tabela 4 a seguir:

Tabela 4- Intervalo de frequência que caracteriza os tons, dada a tessitura desse informante e a interpretação auditiva dos tons

Baixo	**Meio baixo**	**Médio**	**Meio alto**	**Alto**
entre 79,9 e 110,0 cps	entre 110,2 e 124,3 cps	entre 124,3 e 141,7	entre 142,7 e 150,0	entre 150,7 e 191,9

Fonte: própria.

Dentro desse intervalo, foram estabelecidas as faixas acústicas dos tons de acordo com os resultados obtidos na análise auditiva. Esses valores mostram que há uma

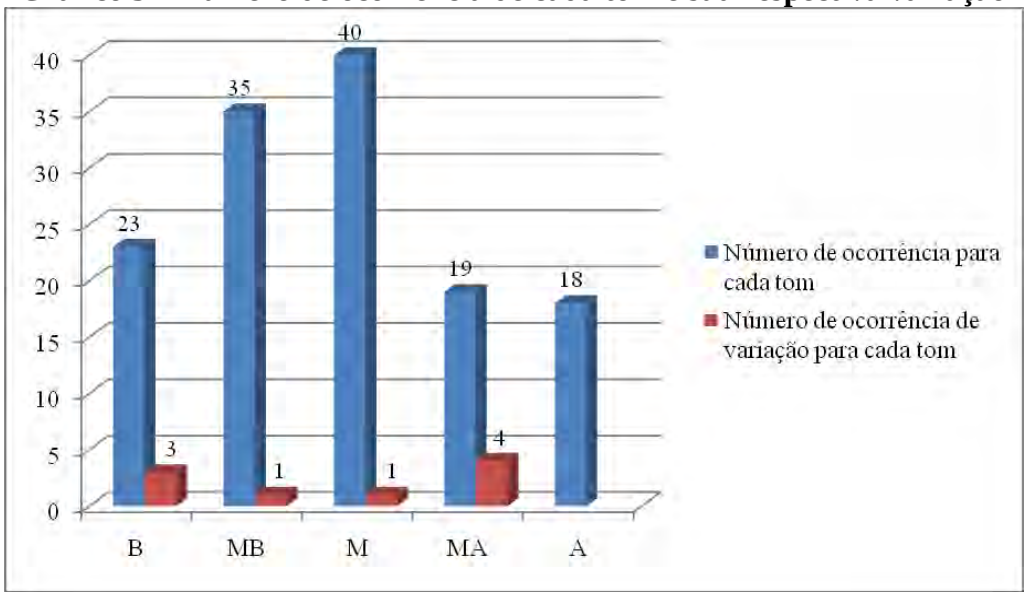
correspondência bem clara de frequências acústicas para cada faixa de tom, tendo em vista a interpretação auditiva. Para esse informante, a tessitura variou de 80cps a 192cps, aproximadamente.

Esse dado confirma a validade das análises auditivas, pois revelou, na maior parte das vezes, um paralelo entre os níveis de altura melódica e os valores acústicos. Em um trabalho anterior[114], esse resultado também foi comprovado, desmistificando a suposição errônea de que não é possível juntar duas análises tão "distantes" entre si. Além disso, mostrou também que é possível trabalhar apenas com um informante.

7.1.3 As divergências nos dados

Como pode ser visto nos resultados que aparecem em vermelho, na tabela 2 (p. 161), referente à distribuição das variações melódicas nas sílabas, há uma diferença entre os níveis melódicos e os intervalos de frequência estabelecidos. Observe o gráfico a seguir:

Gráfico 3 – Número de ocorrência de cada tom e sua respectiva variação



Fonte: própria.

O gráfico mostra que, para a quantidade total de cada tom, o número de variação das frequências acústicas foi praticamente irrelevante. Como se vê, os tons que mais apresentaram variações foram, primeiramente, o tom MA (meio-alto), com quatro ocorrências, e o tom B (baixo), com três ocorrências. Os tons MB (meio-baixo) e M (médio)

[114] Para a confirmação dessa asserção veja: CARVALHO, M; GEMENTI, M.M; CAGLIARI, L. C. *Uma análise interpretativa dos valores acústicos e auditivos dos padrões entoacionais de alguns dados do Português Brasileiro*. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/anais_coloquio/article/view/1204>.

tiveram um caso de variação e o tom A (alto), nenhum, o que não inviabiliza sua possibilidade de ocorrência para outros casos. Para o tom B, em 23 ocorrências, as três variações foram para mais (110,7 Hz, 112,0 Hz e 127,6 Hz), sendo as duas primeiras para o intervalo de frequência dos tons MB (de 110,2 e 124,3 cps) e a última para o intervalo de frequência do tom M (de 110,2 e 124,3 cps). Para o tom MA, tivemos uma divergência em quatro tons. Desses casos, três deles (136,5 Hz 137,0 Hz e 139,5 Hz) enquadraram-se no intervalo de frequência do tom M (de 124,3 e 141,7 cps) e um (158,2 Hz), na faixa de frequência do tom A (de 150,7 e 191,9 cps). Por sua vez, os tons MB e M apresentaram um caso cada: 109,2 Hz e 110,3 Hz, respectivamente. O tom MB enquadrou-se no intervalo de frequência do tom B (de 79,9 a 110,0 cps) e o tom M, no intervalo de frequência do tom MB.

Na tabela 5 que segue, mostra-se o número de ocorrências de tons em função das faixas de frequência.

Tabela 5- Número de ocorrência dos tons em função das faixas de frequência

Hz	Baixo	Meio baixo	Médio	Meio alto	Alto
75	1				
80	5				
90	6				
100	7	1			
110	3	21	1		
120	1	13	12		
130			24	3	
140			2	14	
150				2	6
160					5
170					2
180					5
190					1

Fonte: pória

Na tabela, as maiores ocorrências caracterizam os valores mais típicos dos tons. Por sua vez, os valores mais extremos, que podem ser vistos tanto na tabela acima quanto no gráfico 3 (na página anterior), revelam essa discrepância entre os dados auditivos e acústicos. Esses valores foram interpretados em função do contexto em que ocorrem, considerando a tendência da curva melódica. A observação da curva melódica como um todo no enunciado também permite a verificação imediata do padrão entoacional dentre os tons da língua.

Como foi brevemente comentado, esses casos não descaracterizam a possibilidade de correlacionar dados de uma análise auditiva com dados de uma análise acústica. Para

Cagliari (2007, p. 168), essas variações são perfeitamente possíveis, pois "Como trabalhamos com medidas relativas, devemos deixar bem claro que nem sempre os falantes usam medidas rigorosas, por exemplo, em ciclos por segundo, para o que linguisticamente equivale a um mesmo tom". O autor admite "[...] a possibilidade, não rara, de se ter um tom alto, por exemplo, que, em termos absolutos, tem menos ciclos por segundo do que um tom meio alto, ocorrido pouco antes ou pouco depois, isso porque o falante alterou sua escala entoacional". (CAGLIARI, 2007, p.168). Fato também observado em nosso caso com os tons B, MB, M e MA, que apresentaram mais ou menos ciclos por segundo, como, por exemplo, um dos casos do tom B, que apresentou mais ciclos (127,6) por segundo do que a faixa de frequência estabelecida para o tom MB (de 110 a 124,3 cps).

Os intervalos de frequência estabelecidos para a tessitura também foram outra prova de que essas variações não foram significativas para o que propomos em nosso estudo. Dessa maneira, se fizermos a diferença entre as faixas de frequência determinadas para cada tom (tabela 6), comparando, em seguida, com os tons que ficaram fora dos intervalos auditivos estabelecidos (tabela 7), teremos os seguintes resultados:

Tabela 6- Intervalo da frequência fundamental em função dos cinco níveis tonais e a diferença, em cps, para cada tom

Tons	Intervalo de frequência que caracteriza os tons	Diferença entre os intervalos de F0
B	79,9 – 110,0cps	30,1cps
MB	110,2 – 124,3cps	14,1cps
M	124,3 – 141,7cps	17,4cps
MA	142,7 – 150,0cps	7,3cps
A	150,7 – 191,9cps	41,2cps

Fonte: própria

Tabela 7- Tons que variaram em relação aos intervalos da F0 estabelecidos e sua respectiva diferença, dada a partir dos valores obtidos na tabela anterior

Tons que apresentaram variações nos intervalos determinados	Variação entre os tons a partir das faixas de F0
B→ MB	16cps
B→ M	12,7cps
MB→ M	3,3cps
M→ MA	10,1cps

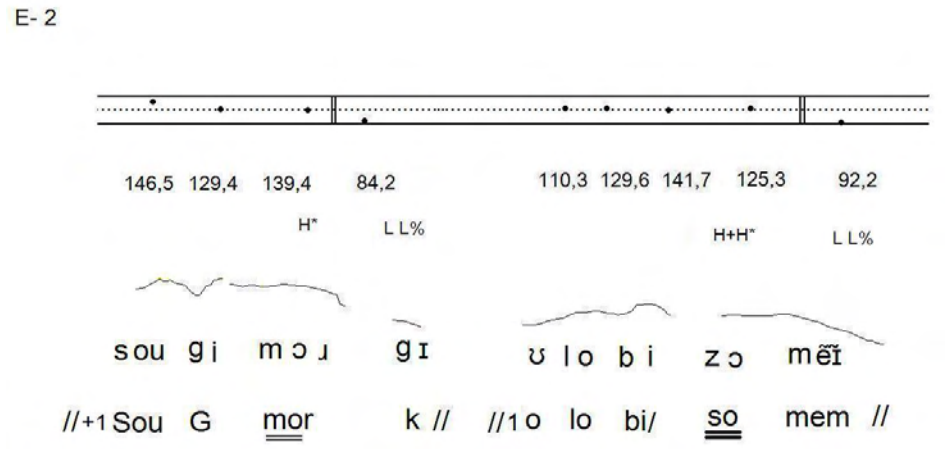
Fonte: própria.

Tendo em vista as tabelas 6 e 7 acima, notamos que se levarmos em consideração que, para o tom baixo, tivemos três variações (duas no intervalo do tom meio baixo e uma no do tom médio), constataremos a não relevância desse resultado, pois uma perda de 16 cps entre um tom e outro é muito irrelevante para o ouvido humano, assim como uma de 12,7 cps, como ocorreu na variação do tom baixo ao tom médio. Pode-se chegar a essa mesma conclusão com os outros tons, considerando que o tom meio baixo atingiu o intervalo de frequência do tom baixo (com uma diferença de 16 cps) e o tom médio, a faixa de frequência do tom meio baixo (com uma diferença de 3,3 cps). Para o tom meio alto, por sua vez, o intervalo foi mais significativo, tendo em vista os dois casos que alcançaram a frequência dos tons altos (com uma diferença de 33,0 cps). Mesmo assim, se tivermos em mente que essa alteração ocorreu com as frequências 153,2 Hz e 158,2 Hz e, ainda, que o intervalo de frequência máximo determinado para o tom meio alto foi de 150,0 cps, e o mínimo para o tom alto foi de 150,7cps, veremos que essa diferença entre os tons é auditivamente insignificante. Ainda para o tom meio alto, as outras duas ocorrências sucederam-se na faixa de frequência do tom M (com uma diferença de 10,1 cps).

Portanto, auditivamente, esses resultados que fogem do padrão são irrelevantes no conjunto da frequência. O ouvido não considera a diferença entre esses valores como sendo relevantes, uma vez que eles ocorrem como variação da curva melódica que não é *stead state*, ou seja, não tem valores fixos, mas pequenos altos e baixos, que são homogeneizados pelo ouvido. Assim, esses resultados são interpretados em função do contexto em que ocorrem, considerando a tendência da curva melódica.

Veja um caso que exemplifica esse fenômeno.

Figura 47- Exemplo de variação entre a análise acústica e a auditiva no segundo grupo tonal do enunciado: "Sou Gmork, o lobisomem"



Fonte: própria

A figura[115] acima faz parte do nosso *corpus* de análise e corresponde ao enunciado 2. Nela, vem identificada a análise auditiva no modelo de Halliday (1970) e Cagliari (1982/2007), a análise acústica no modelo de Pierrehumbert (1980), assim como a transcrição ortográfica e fonética. Também está descrita, nessa imagem, a F0 das sílabas e logo abaixo a curva melódica.

A variação que ocorreu entre a análise auditiva e a análise acústica pode ser evidenciada no segundo GT "O lobisomem", mais precisamente, na sílaba o. Como se vê, essa sílaba foi anotada acusticamente com uma F0 de 110,3Hz, o que, auditivamente, corresponderia a um tom MB. Contudo, ela foi percebida como sendo um tom M. Quando comparamos a sílaba o (110,3 Hz) com a sílaba lo (129,6), vemos que elas apresentam uma diferença de 19,3 cps, o que, estatisticamente, pode ser um fator relevante para a percepção mas, certamente, irrelevante fonologicamente. Isso se justifica porque levamos em consideração a tendência do movimento de variação da curva entoacional, como dito. Além disso, essas sílabas ocorrem em decorrência do que vem antes e depois e, nesse contexto, elas são tipicamente sílabas médias. Se prestarmos atenção na curva melódica da sílaba o, veremos que ela não começa em um nível baixo, assim como a sílaba gi, anterior, do primeiro GT.

[115] Na figura 47, do enunciado E-2, ocorreu um erro de leitura. O falante troucou a consoante (K) pela consoante (g) do final da palavra GmorK.

Além dessas considerações, essa diferença de percepção da altura melódica pode estar relacionada a uma questão pragmática, de modo que, nesse momento da leitura, o falante pode ter tido a intenção de enfatizar a apresentação do personagem e, por isso, mudou a tessitura, o que, consequentemente, levou à percepção de um tom mais alto do que o tom anterior.

Outro fato que também pode ser relacionado a essas variações diz respeito à natureza da melodia dos componentes. Sabe-se que cada componente (pretônico e tônico) tem um valor fonológico que é atribuído pelo falante, como uma forma de simplificação dos valores acústicos. No entanto, há variações acústicas que podem interferir de modo mais ou menos significativo na interpretação auditiva dos tons. Quando estamos analisando um componente pelo seu extremo é mais fácil porque as variações não são tão notáveis. Por outro lado, em um padrão como o mostrado pela figura abaixo, as interferências acústicas ocorrerem facilmente.

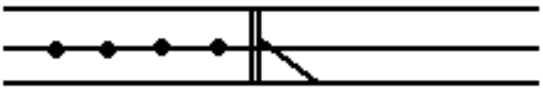
Figura 48- Representação de um padrão melódico que pode interferir de modo mais significativo na correlação entre uma análise acústica e auditiva



Fonte: própria.

Essa dificuldade se deve tanto à característica do componente pretônico (CPT), que é não nivelado e de altura variada, seguindo um padrão que desce/sobe, quanto ao componente tônico (CT), com um padrão sobe/desce. Diferentemente do que aconteceria se tivéssemos um padrão como o exemplificado abaixo, em que temos uma pretônica nivelada e uma tônica descendente, que não apresenta variações no CPT e no CT.

Figura 49- Representação de um padrão melódico que não interfere na correlação entre uma análise acústica e auditiva



Fonte: própria

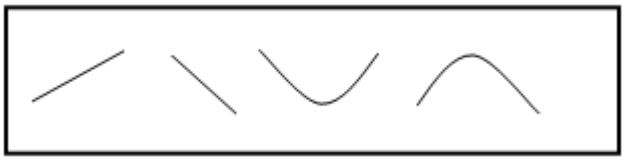
Além do que foi dito, a não identificação entre os valores acústicos e os valores auditivos atribuídos às sílabas pode estar relacionada a um problema de percepção em função da duração desses valores, que, sendo mais curtos, na maioria das vezes, colocam as análises (acústica e auditiva) em confronto, diferentemente das durações mais longas.

À parte dessas considerações, também foi nossa preocupação observar se a metodologia usada neste trabalho (CTRL+0 do PRAAT) para a identificação acústica das

sílabas poderia ter interferido na relação Hz/tom auditivo. Por isso, surge uma pergunta: Será que o centro da sílaba é representativo do padrão entoacional? A resposta mais adequada seria dizer que depende do caso. Às vezes, sim, mas, outras vezes, não.

Se observarmos as tabelas e gráficos descritos mais acima, veremos que, para esse caso, estatisticamente, é mais para sim do que para não, já que, em um total de 135 tons, apenas 9 deles apresentaram discrepância, o que dá uma porcentagem de 6,66% dos casos. Além disso, os valores obtidos por meio dessa função representam um alvo (*target*) articulatório, semelhante ao que ocorre com a estrutura formântica das vogais. Como a transição articulatória ocorre em uma fala contínua, a articulação procura um local ideal para a sílaba, vindo de outra e indo para a seguinte. Quando a curva melódica é mais constante, fica fácil determinar o ponto médio do segmento, pois as variações são menores. Agora, quando temos um padrão, como o exemplificado pela figura 50 abaixo, fica mais difícil a determinação do *target* ideal, ocorrendo, assim, as divergências nos dados, bem como o padrão representado pela figura 48.

Figura 50- Padrões que dificultam a representação do *target* ideal



Fonte: própria

Vale ressaltar, também, que os tons auditivos não mudam os *targets* auditivos, de modo que, tanto em uma frequência mais baixa ou em uma mais alta, o alvo articulatório será sempre o mesmo. As sílabas que apresentaram mais problemas para a determinação do ponto médio foram as que continham sons desvozeados. Normalmente, nesses casos, o *target* ideal estava no segmento desvozeado, o que impossibilitava a análise da F0. Por isso, quando tínhamos esse tipo de som, optamos pelo segundo método de análise, que consistia em selecionar a parte vozeada do segmento e, em seguida, pedir a função *Get pitch* do *software*.

Portanto, mesmo tendo em vista essas diferenças entre as faixas de frequência acústica e auditiva, acreditamos que a determinação da F0 pelo ponto medial da sílaba é o mais representativo do valor entoacional das sílabas do ponto de vista auditivo. A comparação da projeção gráfica com as medidas acústicas do centro das sílabas também mostrou ser um fator muito importante para as duas análises, pois pode revelar aspectos relevantes e irrelevantes da F0. Esses resultados mostraram ainda que é seguramente

possível estabelecer uma relação entre as análises acústicas e auditivas. Além disso, as duas medidas acústicas usadas (CTRL+0 e *Get Pitch*) são uma forma de estabelecer uma regularidade de análise para todas as sílabas dos enunciados.

7.2 Análise do padrão entoacional dos enunciados

7.2.1 Tipos de tons

Nas 15 frases analisadas, foram encontrados 22 grupos tonais[116]. Entre esses GTs, cinco deles foram classificados como uma sequência de dois (caso dos enunciados 2, 4, 10, 13 e 15) ou mais (caso do enunciado 1, com uma sequência de três GTs) grupos tonais. Uma sequência de GTs é muito diferente de um GT composto. O primeiro caso indica que temos, na estrutura de um discurso, duas ou mais unidades de informação, cada uma com uma STS. É o que ocorre, por exemplo, com o enunciado E1, do nosso *corpus*: //Quem /é você// ∧ pergun/tou Atrei/ú// ∧ e aproxi/mou se/ mais um /passo//. Por outro lado, um GT composto apresenta somente uma unidade de informação com duas sílabas tônicas salientes. Em nossos dados, não foi encontrado nenhum GT desse tipo. No entanto, uma frase característica desse tipo de GT é: //∧ Eu não /acho que /seja as/sim// (exemplo retirado de CAGLIARI, 2007, p.86). Os outros dez GTs que faltaram (3, 5, 6, 7, 8, 9, 11, 12, 14), por sua vez, foram classificados GTs simples, com apenas uma mudança notável do contorno melódico ou STS, como em: ∧ Por/que está /preso a/qui// (frase retirada do nosso *corpus*).

Cada GT vem identificado por um tipo de tom. Em nossos dados, tivemos a incidência de tons primários e secundários. Como foi visto, o primeiro é característico de uma enunciação mais neutra, enquanto o segundo refere-se a uma ideia semântica mais marcada, mais enfática do que o seu correspondente tom primário. Dessa maneira, foram identificadas para esses GTs as seguintes ocorrências de tons.

Tabela 8- Especificação dos tipos de tons e número de ocorrência de cada um deles

Tons	Número de ocorrências
Tom 1	17
Tom 3	5

Fonte: própria.

Esses tons estão assim distribuídos, em sequência, no texto analisado:

[116] Para que essas informações fiquem mais claras, observe no item 7.3.2.4 "Análise dos enunciados", as análises feitas para os 15 enunciados.

1+ 1 1 +1 1 1- 3 1-
1+ 1 :1 1 1 3 1-
3+ -1 3 1 +1 3 1

Nessa sequência, temos tons primários e secundários. Os tons secundários recebem um diacrítico, que pode vir antes ou depois do número que caracteriza um tipo de tom. Quando o diacrítico vier depois do número do tom, significa que o tom secundário é uma variação do componente tônico do tom primário correspondente. Por outro lado, se a marca estiver antes do número, quer dizer que esse tom é uma variação do componente pretônico do tom primário correspondente. (CAGLIARI, 2007). No fragmento utilizado, em uma ocorrência de 22 tons, dez deles apresentaram uma configuração de tons secundários e doze de tons simples, sendo que 77,27% dos enunciados foram analisados com o padrão do Tom 1 simples e suas variantes. A maior incidência dos tons primários simples em nosso *corpus* não é tão significativa, se levarmos em consideração que, praticamente, a outra metade dos tons foi classificada com um padrão entoacional de tons secundários. Mesmo assim, esse resultado mostra que, na maior parte das vezes, o leitor optou por uma leitura menos acalorada, com mais frases afirmativas e declarativas. Ainda assim, os tons secundários foram expressivos durante a leitura, mostrando que o falante também procurou enfatizar, quando necessário, determinadas passagens da história.

Os tons distribuídos acima referem-se, portanto, aos padrões entoacionais característicos de cada enunciado, para esse falante. Sendo assim, veja a seguir alguns exemplos de padrões entoacionais que ocorreram na gravação do texto.

Quadro 1- Interpretação dos tons que ocorreram no fragmento do texto *História sem fim*

Tom	Padrão	Significado	Exemplo
1	----||--._	asserção neutra	Gmork não respondeu
1+	¯¯¯¯||--._	forte, inesperado	Você não é daqui
+1	-.- - - |-- - -	preocupação, curiosidade	O que procura?
1-	----||___	polido, suave	Por que está preso aqui?
:1	_–¯ ⟍||	Insistência	E para onde foram?
3	--._||----	Incompleto	Depois de um longo silêncio,

Fonte: própria

Nesse quadro, estão descritos, na primeira coluna, os tons primários e secundários que ocorreram no trecho de leitura. Na segunda coluna, mostra-se a representação pictográfica dos tons, de acordo com a curva melódica do componente pretônico e tônico, que está divido pelas barras duplas ‖. Na terceira coluna, fazemos uma descrição prévia dos significados entoacionais obtidos para cada tom, bem como sua exemplificação na quarta coluna. A análise semântica detalhada aparece no item 7.3.1.1 "Análise dos enunciados".

Estabelecido o padrão entoacional característico de cada tom de acordo com o modelo de Halliday, nossa intenção foi converter a forma fonológica obtida para a de Pierrehumbert, como será descrito em seguida.

Quadro 2- Conversão do tom 1 primário para a forma fonológica de Pierrehumbert

Enunciados	Grupo tonal/tom primário	Forma fonológica no modelo de Pierrehumbert
E1	GT 2/ Tom 1	H+L* L L%
E1	GT 3/ Tom 1	H+L* L L%
E2	GT 2/ Tom 1	H+L* L L%
E6	GT simples/ Tom 1	H+L* L L%
E8	GT simples/ Tom 1	H+L* L L%
E9	GT simples/ Tom 1	H+L* L L%
E13	GT 2/ Tom 1	H+L* L L%
E15	GT 2/ Tom 1	H+L* L L%

Fonte: propria

7.2.1.2 Conversão da forma fonológica para os tons primários e secundários do texto

Em seu artigo, como dissemos, Cagliari (2011) mostrou que a forma fonológica convertida dos dois tons primários simples (1 e 3) do modelo de Halliday (1970) e Cagliari (1982/2007) para o de Pierrehumbert (1980) tem o seguinte padrão:

TOM 1 ➔ H* L L%
TOM 3 ➔ L*+H^{-} L^{-} H%

Já em nossa análise, os padrões obtidos para os tons simples e para os tons secundários tiveram o seguinte padrão, como está exemplificado, respectivamente, pelos quadros 2, 3 e 4 abaixo.

Quadro 3 – Conversão do tom 3 primário para a forma fonológica de Pierrehumbert

Enunciados	Grupo tonal/tom primário	Forma fonológica no modelo de Pierrehumbert
E4	GT 1/ Tom 3	L+H* H H^{L}%
E10	GT 1/ Tom 3	L+H* H H^{L}%
E13	GT 1/ Tom 3	L+H* H H^{L}%
E15	GT 1/ Tom 3	L+H* H H^{L}%

Fonte: própria

Quadro 4- Os tons secundários de Halliday analisados no modelo fonológico de Pierrehumbert

Enunciados	Grupo tonal/tom primário	Forma fonológica no modelo de Pierrehumbert
E1	GT 1/ Tom 1+	L+H* L L%
E5	GT simples/ Tom 1+	H+L* L L%
E2	GT 1/ Tom +1	H+L* L L%
E14	GT simples/ Tom +1	L+H* L L%
E3	GT simples/ Tom 1^{-}	L+H* L L%
E4	GT 2/ Tom 1^{-}	L* L L%
E10	GT 2/ Tom 1^{-}	L* L L%
E12	GT simples/ Tom -1	L* L L%
E7	GT simples/ Tom :1	L+H* L L%
E11	GT simples/ Tom 3+	H* H H^{L}%

Fonte: própria

De uma maneira geral, os três quadros caracterizam, na primeira coluna, os enunciados; na segunda, os respectivos grupos tonais e tons, conforme o modelo de Halliday (1970) e Cagliari (1980/2007) e, na terceira, fazem a correlação fonológica entre essas duas

abordagens, de acordo com o método de análise proposto por Pierrehumbert (1980). Os quadros 2 e 3 são representativas dos tons primários simples 1 e 3. O quadro 4, por sua vez, diz respeito aos tons secundários 1+, +1, 1˗ , ˗1, :1 e 3+.

É importante salientar que os padrões fonológicos obtidos para o modelo de Pierrehumbert (descritos na terceira coluna) tiveram como ponto de partida para a sua transcrição o *pitch accent* ou STS. Isso porque, para a autora, o que vem antes do acento nuclear não é relevante para a análise; o que interessa é o alvo, isto é, o final do movimento a partir da sílaba mais proeminente, servindo como ponto de ancoragem. Para esse modelo, é o fato fonético e não o fonológico que está em discussão. Para o modelo funcional, o CPT tem sua importância, uma vez que indica as variações de tessitura no enunciado. Contudo, como nossa intenção não foi fazer uma análise isolada, mas, sim, comparativa, levamos em consideração, nessa abordagem, somente a tendência da curva melódica a partir da STS.

7.2.2 Caracterização dos padrões fonológicos para o PB a partir do modelo de Halliday e Pierrehumbert

Nesse primeiro momento, faremos uma caracterização dos padrões entoacionais do texto analisado, tomando como base a configuração da curva melódica para o modelo gerativo e a anotação auditiva para o modelo funcional. Nossa intenção é estabelecer uma padronização entre os tons e verificar a aproximação ou não entre os dois modelos. Em relação à análise semântica e pragmática dos enunciados, faremos apenas os apontamentos necessários, deixando a análise mais específica para o item 7.3.1.2 "Análise: o significado dos enunciados de acordo com os diferentes padrões entoacionais", bem como a apreciação dos pontos específicos, quando necessário, de cada modelo teórico.

Em relação ao modelo teórico autossegmental, é importante fazermos uma consideração pontual. Em nossa análise, não anotamos o acento frasal e nem o tom do *pitch accent*, que não é o nuclear, pelo diacrítico sobrescrito (˗) como faz Peirrehumbert, pois entendemos que a análise por si só nos dá os caminhos necessários para identificá-los como tal, uma vez que, de acordo com a metodologia, um acento frasal só poderá vir anotado no final da frase e não pode ocorrer em sílabas tônicas. Quanto ao nuclear *pitch accent*, não há possibilidade de haver dois deles em um tom bitonal, como, por exemplo, H*+L*, já que um dos dois precisa, necessariamente, ser representativo de uma sílaba átona. Como a própria autora diz, *"The bitonal accents resemble bisyllabic feet in that they consist of two elements*

ordered in time on which a strength relationship is defined; the starred tone is the strong one, and the unstarred tone is the weaker one" [117]. (PIERREHUMBERT, 1980, p. 25).

7.2.2.1 Análise: tons primários

Tom 1

Observando o quadro 10, percebe-se que a maioria dos padrões fonológicos apresentou a mesma configuração entoacional quando transpostos para o modelo de Pierrehumbert. Eventualmente, o sistema também permite analisarmos o tom E1 (GT3) por meio de um tom simples L*LL%, devido à possibilidade da curva melódica, o que também não descaracteriza o padrão entoacional descendente do enunciado. No entanto, para efeitos de padronização, optamos pelo bitonal.

A configuração proposta pelo trabalho de Cagliari (2011) foi diferente da nossa, no sentido da configuração da STS, que, além de ser analisada por ele como monotonal (H*LL%), também apresentou como foco um pico (H*) e não um vale (L*). Mesmo assim, essa diferença não contradiz os nossos resultados, já que está relacionada a uma anotação de proeminência na curva melódica que não é fonologicamente contrastiva para os enunciados com esse padrão. Ou seja, todos os enunciados classificados com o padrão H+L*L% em nossos dados foram típicos de enunciados declarativos neutros. Além disso, não podemos julgar o padrão proposto por Cagliari como acusticamente válido, porque fica difícil para o autor estabelecer um padrão sem os valores da F0 e visualização da configuração da curva melódica. Mesmo assim, o padrão proposto por Cagliari evidencia um contorno entoacional descendente para o limite (final) da frase, assim como o nosso.

No modelo de Halliday, o tom 1 caracteriza-se por ter um movimento tônico descendente com tendência terminal baixa ou média baixa. No modelo de Cagliari (2007), esse mesmo tom é descrito por ter um componente tônico com um padrão descendente médio-baixo. Em nossos dados, ele também apresentou esse mesmo padrão entoacional descendente, tanto em comparação com o modelo de Halliday quanto com o de Pierrehumbert.

Além do trabalho de Cagliari (1982/2007), o padrão descendente dos enunciados declarativos também encontra respaldo nos trabalhos de Gebara (1976), que também segue o modelo de Halliday, na abordagem de Fernandes (1976), observando tanto os níveis tonais

[117] "Os acentos bitonais assemelham-se aos pés bissilábicos, já que eles consistem em dois elementos ordenados no tempo em que é possível estabelecer uma relação de força; o tom com asterisco é o mais forte e o tom que não leva essa anotação é o mais fraco". [tradução nossa].

quanto a curva melódica e, por fim, no trabalho de Morais (1998), destacando o padrão descendente em final de enunciado, especialmente para a sílaba tônica.

Portanto, segundo a nossa análise, o tom representativo das declarativas neutras finais para o PB, segundo o modelo de Pierrehumbert é H+L* L L%, iniciando-se com uma F0 mais alta e, em seguida, mantendo-se em queda na posição final do enunciado. Em todas as frases analisadas, esse padrão ocorreu nas palavras focais em final de enunciado, mais especificamente com tom L alinhado à penúltima (E1 – GT3/ E2 – GT2/ E6/ E9/ E15 – GT2) e última (E1 – GT2/ E8/E13 – GT2) sílabas.

Em relação ao padrão das declarativas, de acordo com o padrão fonológico alcançado a partir da descrição de Pierrehumbert (1980) para esta dissertação, Lucente *et alli* (2007) no artigo "Declarativas em PB: downstepping ou nova combinação bitonal? H+!H* e H+L*" discutem a adoção desse padrão tonal (H+L*) ao sistema ToBiPI (*Transcription* of *Brazilian Intonation*), que foi desenvolvido para a notação do português brasileiro (PB). Nesse trabalho, a intenção dos autores é observar se as sentenças declarativas em PB poderiam ser representadas por meio da alternância entre o tom H+!H*, previamente estabelecido para a marcação de *downstep* e sua nova proposta tonal H+L*. Segundo eles,

> A nova marcação até então não era utilizada na notação do PB. Para assinalar o movimento de tom descendente era utilizada a notação H+!H* (Barbosa, 2005), que marca o movimento de downstepping, em especial a declinação que ocorre em declarativas.
> (LUCENTE *et alli*, 2007, p. 3)

Contudo, os autores admitem esse padrão e confirmam a possibilidade de que, dependendo da função expressa nas sentenças, é possível que a configuração final das sentenças declarativas em PB seja representada pela alternância entre esses dois padrões entoacionais. A fronteira final também é marcada com L%, mas, diferentemente da nossa proposta, não há a notação do acento frasal, já que o sistema ToBiPI a eliminou, substituindo uma combinação do tipo L- L% somente por L%. [118] Dessa maneira, mesmo sem a sua marcação, o trabalho dos autores confirma a nossa análise, admitindo que as sentenças declarativas em PB podem ser representadas também por esse padrão H+L* L%.

Moares (2008), a fim de estabelecer o inventário geral do sistema entoacional do PB para a sua variante carioca, identificou dois tipos de padrões entoacionais para as declarativas neutras L+H* e H+L*L%. O primeiro, na posição de acento prenuclear, isto é,

[118] Não trouxemos para a discussão o porquê de Lucente *et alli* (2007) proporem mais uma forma de notação para as declarativas em PB, já que queremos apenas validar nossa análise;

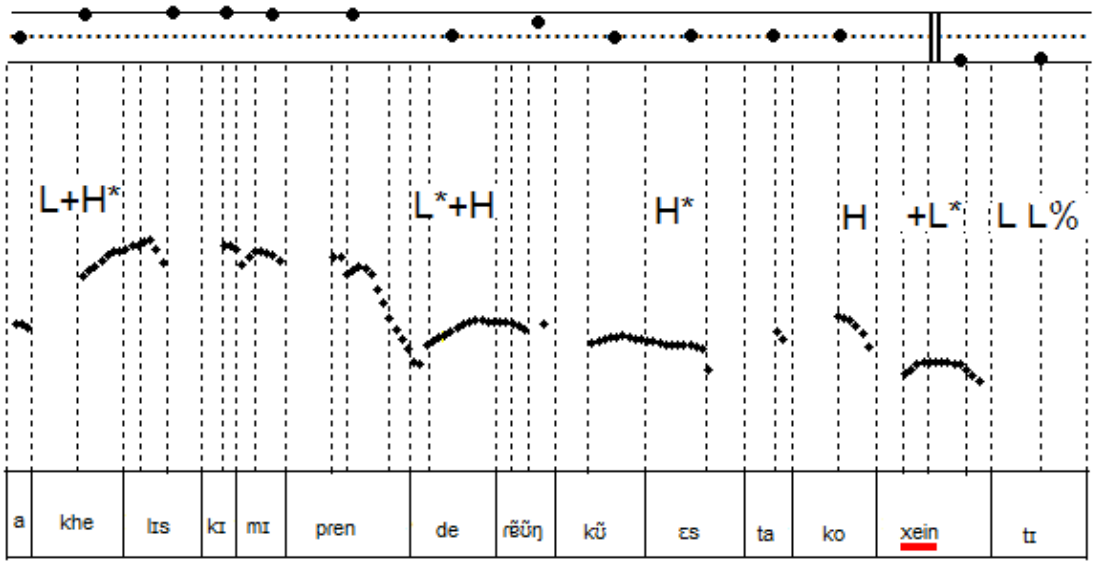
nas sílabas pretônicas e o segundo, na posição nuclear, sílabas tônicas. Essa última notação também confirma a análise obtida em nossos dados para o dialeto paulista.

Truckenbrodt *et alli* (2008), também propõem um tipo de correspondência entre o modelo auditivo e o modelo autossegmental. Para isso, os autores se baseiam no trabalho de Cagliari (1982), analisando somente os tons primários. Para o modelo autossegmental, seguem as anotações de Pierrehumbert (1980), Ladd (1996), Tenani (2002), Moraes (1998), entre outros. Feito isso, Truckenbrodt *et alli* (2008), chegam à conclusão de que os enunciados declarativos de tom 1 também apresentam padrão entoacional descendente, caracterizando-se, tomando como base o trabalho de Tenani (2002), pelo padrão H L*Li[119].

Vale lembrar, que os autores não objetivaram fazer esse tipo de comparação em seu artigo, estando interessados apenas em mostrar as duas possibilidades de análise. Mas, mesmo que implicitamente, eles mostraram a possibilidade desse tipo de correlação.

Agora, veja um exemplo da configuração da curva melódica no sentido da queda da Hz, para o modelo de Pierrehumbert e de acordo com os pontos auditivos propostos para a anotação de Halliday em nossos dados.

Figura 51- Exemplo de sentença declarativa segundo os padrões de Halliday (1970) /Cagliari (1982/2007) e Pierrehumbert no enunciado: "Aqueles que me prenderam com esta corrente"



Fonte: própria

[119] A notação mostrada por Truckenbrodt *et alli* (2008) toma como base o trabalho de Tenani (2002). No entanto, como vimos, em trabalhos mais atuais, a autora passa a considerar tons bitonais para o padrão das sentenças declarativas em PB, como o descrito nesta dissertação. No artigo de Truckenbrodt *et alli* , o (i) ao lado do tom baixo (Li) indica um tom de fronteira.

O exemplo mostra a tendência à declinação da curva melódica para esse tipo de sentença nos dois modelos teóricos. Os outros enunciados declarativos também seguiram esse mesmo padrão para o componente tônico.

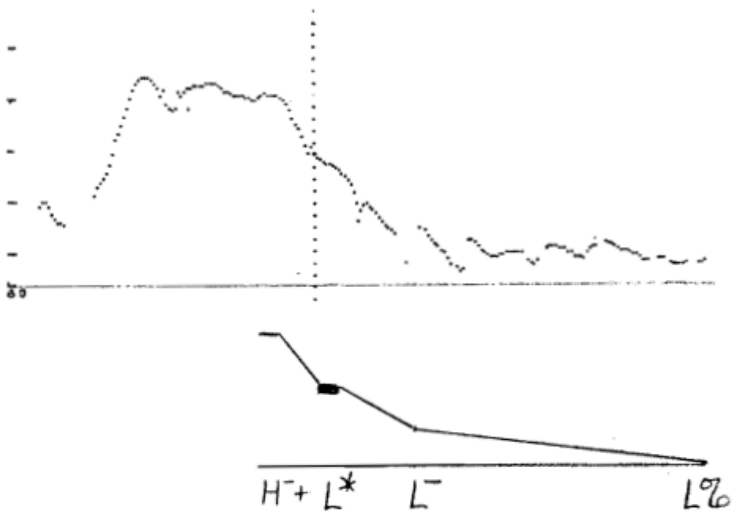
Sendo assim, o padrão proposto para as declarativas neutras em nossos dados é H+L* L L%, de acordo com a curva melódica. Já para o modelo funcional, o padrão resultante da análise corrobora o que foi proposto nos trabalhos de Halliday/Cagliari: de tônica descendente. Dessa maneira, segundo nossos dados, o tom 1 primário pode ser descrito linguisticamente, segundo o padrão auditivo, da seguinte forma:



Os CPTs variaram entre uma pretônica média nivelada e baixa nivelada[120]. Abaixo da descrição do exemplo acima, H+L* L orma fonológica obtida de acordo com a notação de Pierrehumbert.

No apêndice de sua tese de doutorado, Pierrehumbert ilustra todas as possibilidades de combinação entre o *pitch accent*, o acento frasal e o tom de fronteira para o final do enunciado. A combinação proposta por ela resultou em um total de 22 diferentes padrões entoacionais. Os exemplos usados para gerar esses contornos são variantes do enunciado *"The Uruguayan bulldozer drivers' union"*. Entre esses exemplos de variação, está representado o padrão descrito em nossa análise. Veja a seguir:

Figura 52- Representação acústica do esquema de contorno tônico do padrão entoacional H-+L* L- L% do enunciado "The Uruguayan bulldozer drivers' union"



Fonte: Pierrehumbert (1980, p. 399).

[120] Como havíamos dito, a configuração do componente pretônico não é essencial para a definição do padrão entoacional em um enunciado, mas, para efeito representativo, optamos por mencioná-los.

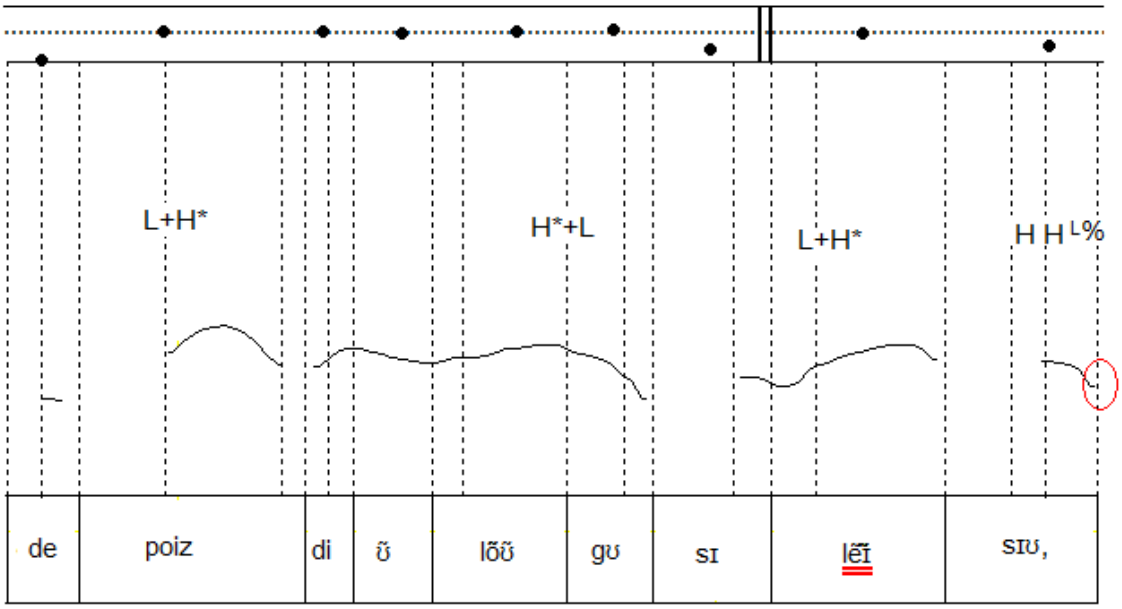
No exemplo mostrado, a linha vertical mostra o início da sílaba sobre a qual o foco recai, ou seja, a sílaba *bull-* da palavra *bulldozer*. Se compararmos a curva melódica do exemplo da figura 52 acima com essa esquematicamente representada por Pierrehumbert, observaremos um padrão entoacional descendente, que começa alto na sílaba átona e, em seguida, sofre uma queda em direção à sílaba tônica.

Ladd (1996), como foi dito, sumariza as 22 combinações de configuração dos acentos nucleares propostos por Pierrehumbert e faz uma comparação com o inglês britânico. De acordo com o trabalho do autor, o padrão mostrado para as declarativas nesta dissertação seria correspondente à descrição, em inglês britânico, de um tom do tipo *low fall (with a high head)*. De acordo com o autor, esse tipo de movimento mostra que a sílaba tônica acentuada (*starred*) é a central e a átona é a responsável, normalmente, por determinar o padrão da sílaba tônica anterior, quando estiver presente.

Tom 3

Para todos os enunciados de tom 3 simples, encontramos o mesmo padrão L+H* H H^{L}%. Diferentemente do tom 1, nesse há uma tendência da curva melódica no sentido da subida de Hz, ocorrendo um pico local no início da sílaba acentuada, a partir de um nível bem mais baixo. Essa característica ascendente da curva melódica gera outra particularidade desses enunciados, isto é, a caracterização do tom final de fronteira final, que passa a ser anotado com um L sobrescrito. Adotamos essa forma de análise porque, em PB, é comum que toda a Hz ascendente termine com um valor de Hz um pouco mais abaixo, ou seja, com uma tendência ao abaixamento do ângulo, que está destacado, na figura que segue, por um círculo vermelho. Observe o exemplo:

Figura 53- Exemplo da análise do tom 3 na forma fonológica de Pierrehumbert no enunciado "depois de um longo silêncio,"

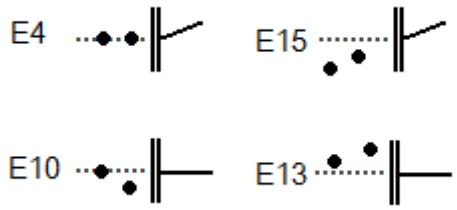


Fonte: primário

A figura identifica o tom 3 simples do enunciado E10, do GT "*depois de um longo silêncio,*". A STS está sublinhada em vermelho. A tendência em PB está identificada no final do enunciado, podendo ser descrita ainda, por uma configuração aproximada do tipo . Essa tendência foi encontrada para todos os outros enunciados com esse padrão. Por sua vez, o enunciado E13, diferenciou-se dos demais por ter a sílaba átona (ci-), que precede a tônica a STS (da-) mais alta (cf. item 7.3.2.4). De acordo com Cagliari (2002), essa também é outra tendência do PB, porém, mais característica dos dialetos nomeados genericamente de nordestinos, como, por exemplo, o dialeto baiano.

Halliday (1970) também identifica um padrão ascendente para o movimento da sílaba tônica de enunciados de tom 3. Cagliari (2007), por sua vez, a caracteriza como um movimento médio nivelado, por salto, o que não deixa de evidenciar um padrão ascendente, já que a tônica é mais alta, se comparada à pretônica (identificada pelo autor como descendente média baixa). Esse dado também não diverge de nossa análise, já que o padrão entoacional encontrado correspondeu ao que foi proposto por Cagliari para o sistema de tons do PB. No entanto, também encontramos para a tônica, além de um movimento médio nivelado (para E10 e E13), um movimento ascendente (E4 e E15).

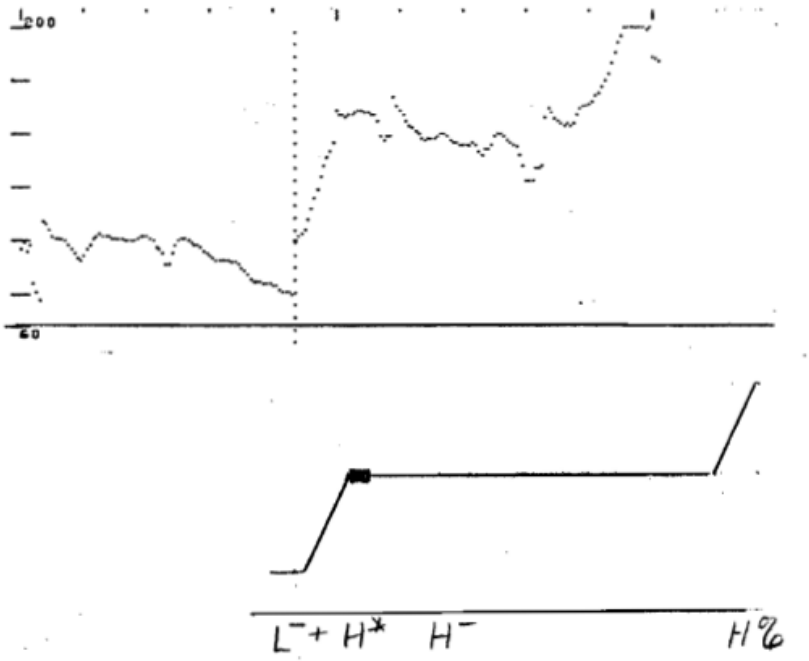
O componente pretônico variou entre médio nivelado (E4), descendente médio baixo (E10); meio-alto, alto (E13) e baixo/meio-baixo (E15). Esses padrões estão identificados a seguir:



Em relação ao modelo de Halliday e Cagliari, pode-se afirmar que houve uma correspondência com o padrão resultante desta dissertação, já que, para o primeiro, identificou-se o movimento ascendente da tônica e, para o segundo, o médio nivelado.

Acusticamente, o esquema entoacional proposto por Pierrehumbert (1980) também confirma o padrão tonal observado. Contudo, a curva melódica não apresenta essa tendência à queda, mesmo porque, como mencionamos, essa é uma característica típica do PB. A seguir, vemos a representação de Pierrehumbert.

Figura 54- Representação acústica do esquema de contorno tônico do padrão entoacional L-+H*H- H% do enunciado "The Uruguayan bulldozer drivers' union"



Fonte: Pierrehumbert (1980, p. 396).

Assim, como identificado para o tom 1, a representação do padrão entoacional para o tom 3, de acordo com a figura anterior, também mostra a correlação entre o modelo acústico e auditivo.

Para o inglês britânico, a correspondência feita por Ladd (1996) corrobora a afirmação acima, já que, para o autor, L+H* H H% corresponde a um tom *high rise* (*with*

low head). A correspondência feita para o PB por Truckendrodt *et alli* (2008) a partir do trabalho de Tenani (2002) também aponta a tendência final ascendente desse padrão. Contudo, para os autores, esse tom corresponde, em Tenani, a L* Hi, que, diferentemente do nosso, tem o foco em um tom baixo, mas que também termina com um padrão ascendente[121].

Em relação à descrição para o PB feita por Cagliari (2011) – L*+H$^-$ L$^-$ H% -, veremos que não há uma identificação absoluta com o nosso padrão, de modo que, para ele, o pico está em um nível mais baixo da curva melódica, havendo uma subida em direção à sílaba átona. A identificação entre os dois modelos ocorre somente no sentido final da curva de Hz, que acaba alta.

7.2.2.1 **Análise:** tons secundários

Tom 1+

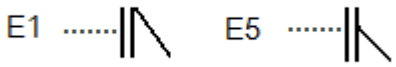
Foram encontradas duas ocorrências para o tom 1+: uma no enunciado E1[122], primeiro GT, e a outra no E5, GT simples. Na conversão para a forma fonológica de Pierrehumbert, os padrões obtidos para os dois enunciados foram: L+H* L L% para E1 e H+L* L L% para E5.

Em E1, a F0 se realizou com uma leve descida no início da sílaba, terminando com uma subida no final da mesma sílaba e, por isso, o nuclear *pitch accent* alinhado a um tom baixo. Já em E5, a curva melódica inicia-se por uma subida, indicando um tom alto logo no início da consoante, sendo o nuclear *pitch accent* no tom baixo logo no início da palavra, mais especificamente no começo da queda na curva melódica, caso contrário de E1, que começa com uma curva melódica levemente mais baixa na consoante, atingindo o pico no final da sílaba. Mesmo com a diferença nos dois enunciados quanto à localização do *pitch accent* (um no pico e o outro no vale), nota-se que o movimento final nas duas frases foi descendente. Isso pode ter ocorrido devido à diferença de tamanho entre os dois GTs. O GT de E5, se comparado ao de E1, é muito pequeno, fazendo com que haja uma compactação em um espaço de tempo muito curto. Mesmo assim, não há uma perda nas características de E5. Os dois padrões foram observados em sentenças interrogativas com foco na palavra interrogativa "quem".

[121] Truckendrodt *et alli* não fazem nenhuma consideração sobre a tendência ao abaixamento final mostrado aqui.

[122] A curva melódica que representa os enunciados E1 e E5 podem ser vistas no item 7.3.1.2 referente à análise dos enunciados.

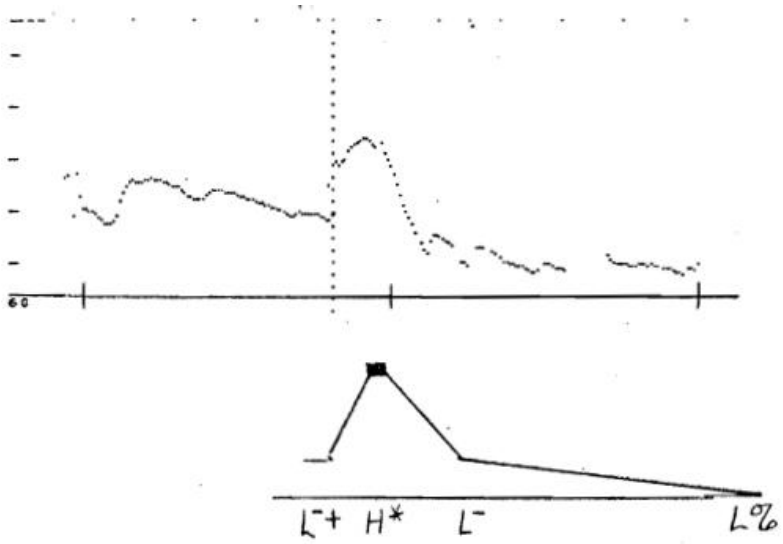
No modelo de Halliday (1970) e Cagliari (1982/2007), o tom 1+ é caracterizado por apresentar uma tônica descendente alta-baixa. Em E1, esse fato se confirmou, mas, em E5, a tônica foi anotada como descendente média-baixa. Observe abaixo:



Porém, esse fato, mais uma vez, não descaracteriza nossa análise, uma vez que o movimento final foi descendente para ambos. Por outro lado, seria interessante analisar mais enunciados desse tipo, para observar se é possível ou não uma padronização. Nos exemplos acima, o CPT não foi caracterizado, pois, em nossos dados, ele não estava presente, já que o foco ocorreu na primeira palavra da sentença.

No modelo de Pierrehumbert, há uma representação para esses dois padrões, ambos com movimento final descendente. A primeira combinação está representada na figura 54, na página anterior, para o tom 1, e a segunda possibilidade é mostrada a seguir:

Figura 55- Representação acústica do esquema de contorno tônico do padrão entoacional L-+H*L- L% do enunciado "The Uruguayan bulldozer drivers' union"



Fonte: Pierrehumbert (1980, p. 395).

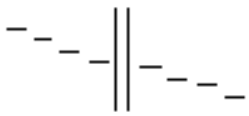
Na correspondência feita por Ladd (1996), esses dois padrões também estão caracterizados. O primeiro como *rise-fall* (L+H* L L%) para E1, considerando que o *rise* representa o pico entoacional. E, para E5, como *low-fall*, assim como mostrado para o tom 1 primário.

Tom +1

Acusticamente, o tom +1 foi analisado assim como o tom 1+ acima. Para esse tom, também foram identificados dois enunciados o E2[123] (primeiro GT) – H+L* L L% em sentença declarativa e E14 (GT simples) – L+H* L L%, em enunciado interrogativo.

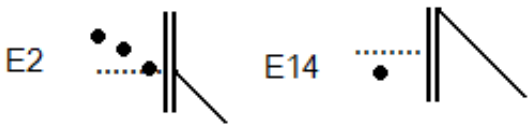
Em E14, o *pitch accent* é alto e está alinhado a um tom baixo e, em E2, ele é baixo e está ligado a um tom baixo. Em E2, o acento frasal L está na sílaba átona (MEM) da palavra "lobisomem", o foco do enunciado. Em E14, por sua vez, foi anotado no final da frase, mais especificamente na palavra "procura". Isso ocorreu porque a STS estava logo no início do enunciado, na palavra "que". Sendo assim, o acento frasal foi o responsável pela configuração final da curva melódica. Como pode ser observado, E14 tem o mesmo padrão fonológico de E1 e também é um enunciado com palavra interrogativa, como E5. Porém, nesse enunciado (E14), o CPT está presente, fazendo-se contrastivo entre E1 e E5. O mesmo pode ser dito para E2 em relação a E5. Nos dois tipos de tons, o movimento final foi descendente.

Tanto Halliday quanto Cagliari não descrevem o padrão auditivo para esse tom. Contudo, em conversa particular com o Prof. Cagliari, a representação mais adequada para ele seria:



TOM +1

Na descrição acima, a pretônica é descendente média-alta e a tônica, descendente média-baixa. Em nossos dados, E2 foi classificado com uma tônica descendente baixa e E14 com uma tônica descendente alta-baixa, por etapas. O CPT variou de médio-alto para o primeiro e baixo para o segundo. Portanto, a identificação completa com o modelo de Cagliari ocorreu apenas para E2. Veja os padrões da tônica e da pretônica em nosso *corpus*:



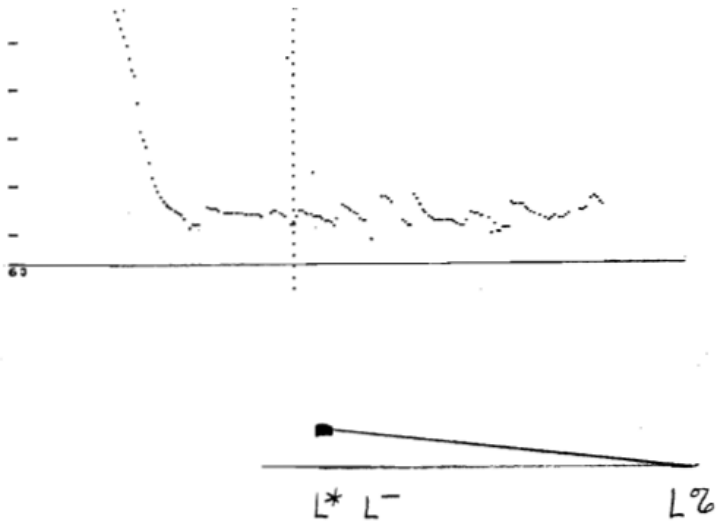
[123] Veja a curva melódica dos enunciados E2 e E14 no item 7.3.1.2 referente à análise dos enunciados.

Tom 1-

Foram encontradas três ocorrências (E4, E3 e E10)[124] para o Tom 1-. O E4 e E10 foram anotados com o padrão L* L L%, mostrando uma proeminência no nível mais baixo da curva melódica. Nos dois enunciados, o acento frasal é mais baixo do que o *pitch accent*, mostrando a tendência descendente da curva melódica no final do enunciado. Em E4, houve uma queda mais abrupta em relação ao CPT, enquanto, em E10, a queda foi mais gradual. Como o acento frasal já se iniciou de uma queda, ele não foi o responsável pelo abaixamento da F0 no final dos enunciados. O *pitch accent* e o acento frasal estão presentes no CT e ocorrem na mesma palavra, mas em posições diferentes, isto é, o primeiro em início de palavra, e o último no final da palavra, tanto para E4 quanto para E10. Em E4, o foco (nuclear *pitch accent*) recai sobre a segunda sílaba e, em E10, sobre a primeira. Mesmo E4 tendo um GT mais longo do que E10, não houve influência na configuração da curva melódica, uma vez que o foco estava na sílaba da última palavra do enunciado.

O padrão proposto para esse tom também foi reconhecido por Pierrehumbert, caracterizando-se por um abaixamento da curva melódica a partir do início da sílaba *bull-* onde cai o foco. A seguir, sua representação:

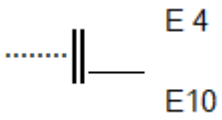
Figura 56- Representação do esquema de contorno tônico para o Tom 1-, no enunciado "The Uruguayan bulldozer drivers' union"



Fonte: Pierrehumbert (1980, p. 394).

[124] Veja a curva melódica dos enunciados E3 e E10 no item 7.3.1.2 referente à análise dos enunciados.

Em Ladd (1996), a correspondência também foi possível, sendo anotada por um padrão *low fall*. Para Halliday (1970) e Cagliari (1982/2007), o CT do tom 1- foi igualmente caracterizado de modo descendente, mais especificamente como descendente, meio-baixo, baixo. Contudo, em nossos dados, o contorno melódico do CT nos dois enunciados foi identificado como baixo nivelado, ainda assim com tendência à queda na curva melódica:



Diferentemente de E4 e E10, E3 foi caracterizado pelo padrão L+H* L L%. Nesse enunciado temos um bitonal. O tom alto foi marcado na sílaba tônica onde recai o foco e o tom baixo, na sílaba átona anterior. Ao observarmos a curva entoacional e a F0, vemos que a diferença entre a sílaba átona e a tônica é de 51,5 Hz, o que é consideravelmente grande. Depois disso, a queda passa a ser gradual, de modo que o acento frasal é anotado no final da sentença e de novo é o responsável por esse abaixamento. A outra diferença acontece no nível semântico, visto que E4 e E10 são enunciados declarativos e E3 interrogativo. Em E3, o foco está palavra interrogativa, logo no início da frase, enquanto em E4 e E10, foi identificado no final das sentenças. Além disso, em E10, o GT é novamente muito pequeno, levando à suposição de que, se estivesse expandido, poderia ter outra configuração, assim como ocorreu para E5. Dessa maneira, também podemos utilizar a representação do padrão auditivo de E4 e E10 acima para caracterizar E3, mais especificamente como descendente alto, meio alto, baixo. A diferença entre eles está no CPT, uma vez que, em E4, a pretônica é média nivelada e, em E10, não há caracterização, já que a STS ocorre logo na primeira sílaba do GT e, em E3, é meio-alta.

Tom -1

O tom -1 ocorreu apenas para E12, apresentando o mesmo padrão de E4 e E10 de tom 1- identificado, portanto, no sistema fonológico de Pierrehumbert e Ladd. No caso de E12, o acento frasal foi mais baixo do que o *pitch accent*, ocorrendo na última sílaba após o foco. Em E12, esse tom não foi responsável pela queda principal, mas, sim, o *pitch accent*, assim como em E4 e E10. Contudo, diferentemente daqueles, a queda da curva melódica não foi abrupta ou gradual, mas manteve-se nivelada ao longo da frase, com uma leve proeminência na sílaba tônica "que" de F0, mais alta entre as outras sílabas.

Auditivamente, a curva entoacional do CT desse tom também se igualou aos anteriores (E4 e E10), em nossos dados, mas, na análise de Halliday e Cagliari, ele foi

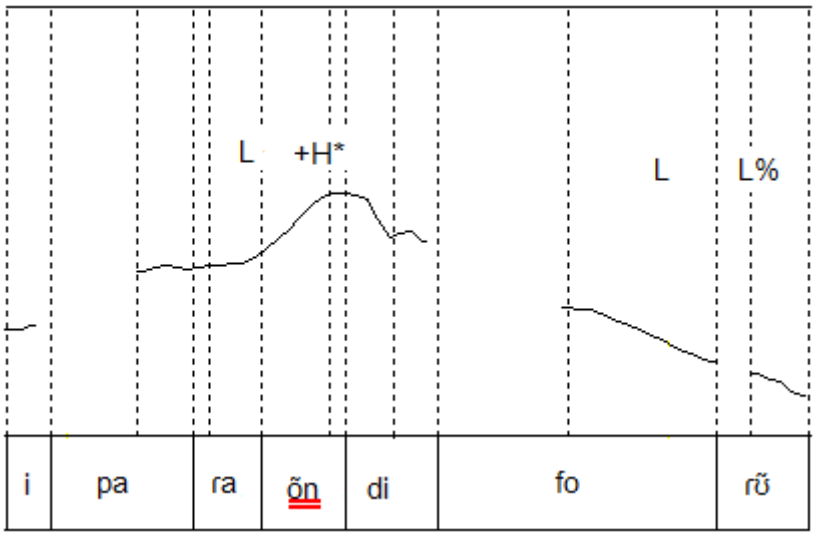
anotado com uma tônica descendente média baixa. Mesmo assim, como nos outros enunciados, esse resultado não desqualifica o nosso, apenas mostra que, para os autores, a queda parte de um nível médio e não de um baixo. A diferença entre esses tons no modelo auditivo, assim como ocorreu para o tom +1 e 1+, está na configuração do CPT e CT. Para o tom -1, Halliday e Cagliari descreveram a pretônica como desnivelada, variando, na maioria das vezes, entre baixa e meio baixa. Em nosso *corpus*, o CPT desse tom foi classificado como baixo nivelado. Por sua vez, o tom 1-, descrito pelos autores com uma pretônica nivelada meio-baixa, foi representado, neste trabalho, como baixa para E10 e, para E4, como nivelada meio-alta. Mesmo não sendo necessária para a descrição de Pierrehumbert, essa distinção é importante para os tons secundários no modelo funcional, pois, segundo Cagliari (2007), ajuda a evitar ambiguidades e a explicar melhor o contorno melódico dos CTs e CPTs.

Tom :1

O tom :1 ocorreu apenas em E7, sendo anotado como L+H* L L%, assim como o tom 1+ em E1 (primeiro GT). E7 e E1 são enunciados interrogativos, um com palavra interrogativa (E1) e o outro não (E7). Em E1, o *pitch accent* inicia-se mais baixo e está alinhado na mesma sílaba do foco. Por sua vez, em E7, o *pitch accent* está alinhado à sílaba anterior, que não é o foco do GT. Em E7, a F0 do tom H* foi bem maior do que a de E1, essa com 150 Hz e aquela com 181,7. Em E7, o acento frasal teve início a partir da sílaba átona da palavra que leva o foco e, em E1, na sílaba átona vo- da palavra "você", posterior ao foco.

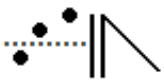
O movimento da curva melódica, mesmo com a configuração descendente final, é muito diferente nas duas frases. Em E7 é como se a curva melódica, com suas devidas variações, lembrasse uma parábola invertida, com uma proeminência bem no meio do enunciado. Essa configuração foi possível justamente pela posição da sílaba tônica. Por outro lado, em E1, o declive foi mais constante, uma vez que a STS estava no começo. Veja o exemplo da curva melódica para E7.

Figura 57- Exemplo da curva melódica para o enunciado E7 – "E para onde foram?" Com caracterização do *pitch accent*, destacado em vermelho, do acento frasal e do tom de fronteira



Fonte: própria.

Auditivamente, o tom :1 secundário não é descrito por Halliday (1970), somente por Cagliari (1982/2007). De acordo com esse autor, a tônica nesse tom é anotada como descendente alta-baixa e a pretônica, como ascendente, meio-baixa, alta. Para o único enunciado que apresentou esse tom em nossos dados, o resultado foi exatamente o mesmo descrito por Cagliari. Veja o padrão a seguir:



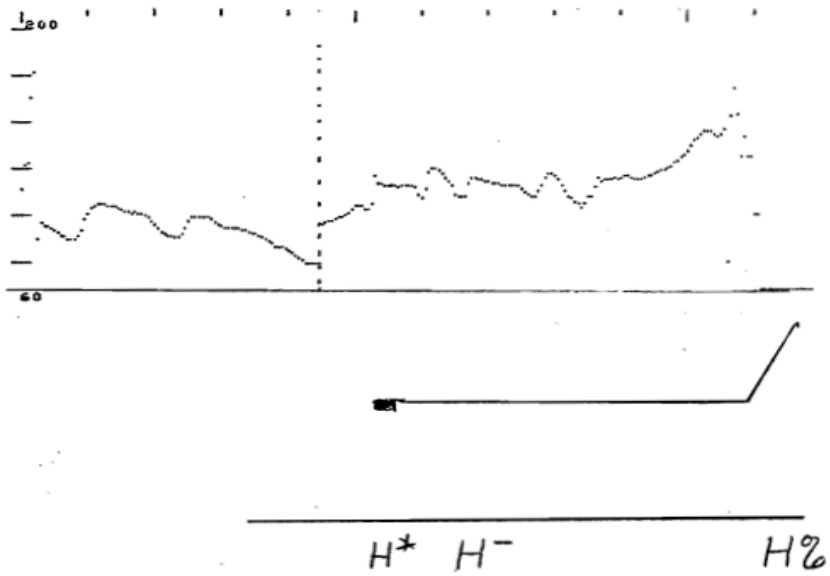
Tom 3+

Pela análise acústica de Pierrehumbert, o tom 3+ secundário foi descrito em nossos dados apenas para E11, com o padrão H* H H^{L}%. Assim como no tom 3 simples, a curva melódica caracterizou-se por terminar com uma subida, com uma leve tendência à queda. Contudo, diferentemente desse tom, o *pitch accent* não está associado a uma sílaba pretônica baixa (L+H*), já que a sílaba átona anterior é mais alta do que a STS.

Em E11, o *pitch accent* vem identificado na última sílaba do enunciado e é responsável por gerar uma elevação na curva entoacional, ou seja, um *upstep*, pois sua F0 é mais alta (136,5 Hz) do que a do tom H a sua esquerda (134,2 Hz).

Pierrehumbert também encontrou uma representação para esse tipo de padrão entoacional. Veja abaixo:

Figura 58- Contorno da F0 para o enunciado "The uruguayan bulldozer drivers' union"

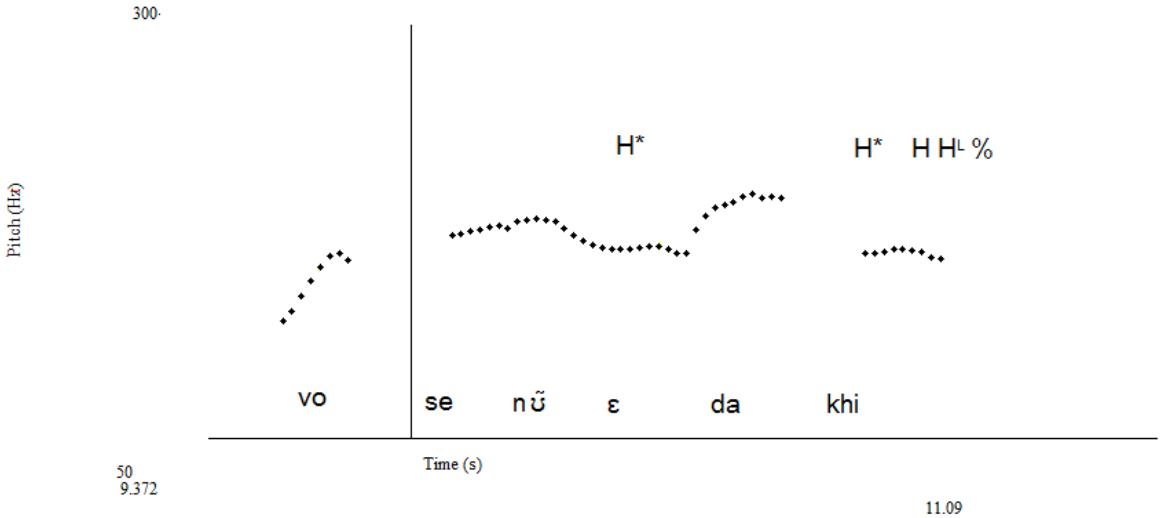


Fonte: Pierrehumbert (1980, p. 392)

Como se vê, diferentemente da caracterização feita para o tom 3 simples (veja a figura 54, página 182), não há o início da curva melódica mais baixa para o CPT, devido, como dissemos, à não associação com o tom L na sílaba precedente. Em Ladd (1996), esse padrão também é identificado e está descrito como *high rise*, com foco no tom H.

A correlação entre os dois modelos fica mais clara se observamos a curva entoacional representativa desse tom para esse enunciado (E11).

Figura 59- Contorno da F0 do enunciado E11 "Você não é daqui"



Fonte: própria

Como se vê, os dois contornos melódicos são bem parecidos. A maior diferença está na tendência final do enunciado, que se mostra mais alto para o modelo de Pierrehumbert,

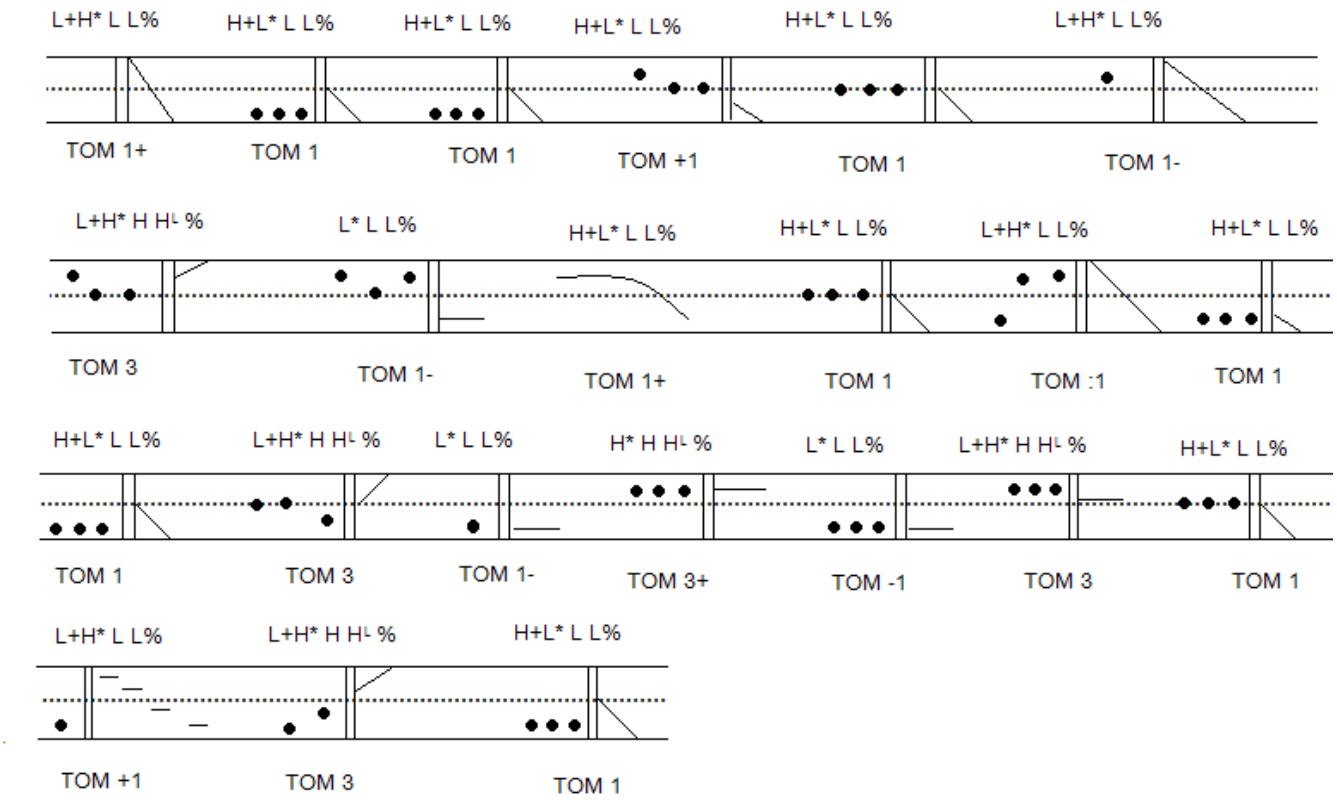
A linha vertical, nessa figura, divide o enunciado em CPT e o CT. A regra de alinhamento fonológico está claramente identificada a partir do *pitch accent*, na sílaba qui- da palavra daqui. A saliência anterior a esse acento tonal mostra que a sílaba átona é mais alta do que a tônica.

Quanto ao modelo auditivo, também houve equivalência entre os resultados. Halliday e Cagliari descrevem a tônica como nivelada alta, assim como ela ocorre em nossos dados. Em relação à pretônica, Cagliari (2007) a classifica como descendente alta, meio alta e, neste trabalho, ela começou mais baixa, mantendo-se entre alta e meio-alta.

Por fim, tendo em vista as análises anteriores, a distribuição dos tons pode ser resumida para os enunciados, tanto auditivamente quanto acusticamente, da seguinte forma:

Sendo assim, por meio das análises feitas, chegamos a uma correlação entre o sistema fonológico de Pierrehumbert (1980), Ladd (1996), Halliday (1970)/Cagliari (1982/2007) e o nosso, que pode ser visto no quadro (5) comparativo que segue:

Figura 60- Esquema de representação, em sequência, de todos os tons encontrados no fragmento de texto analisado



Fonte: própria

Quadro 5- Correspondência entre as formas fonológicas acústica e auditiva para os modelos teóricos que as ilustram, bem como os resultados obtidos para cada um deles nesta dissertação

Sistema fonológico de Pierrehumbert	**Sistema fonológico de Ladd**	**Tons: Halliday/ Cagliari**	**Padrão auditivo da tônica: Halliday/Cagliari**	**Resultados da análise acústica**	**Resultados da análise auditiva/padrão da tônica**
H-+L*L-L%	*low fall (with high head)*	TOM 1 – primário	descendente media-baixa	H+L*L L%	1) descendente media-baixa
L-+H*H-H%	*high rise (with low head)*	TOM 3 – primário	media nivelada, por salto	L+H*H H^{L}%	1) alta, por salto 2) média nivelada
1) H-+L*L-L% 2) L-+H* L-L%	1) *low fall* 2) *rise fall*	TOM 1+ secundário	descendente alta-baixa	1) H+L*L L% 2) L+H* L L%	2) descendente baixa 1) descendente alta-baixa
1) H-+L*L-L% 2) L-+H* L-L%	1) *low fall* 2) *rise fall*	TOM +1 Secundário	descendente média-baixa	1) H+L*L L% 2) L+H* L L%	1) descendente baixa 2) descendente alta-baixa, por etapas
1) L* L-L% 2) L-+H* L-L%	1) *low fall* 2) *rise fall*	TOM 1- Secundário	descendente meio-baixa, baixa	1) L*L L% 2) L+H* L L%	1) baixa nivelada 2) descendente alta, meio alta, média-baixa
L* L-L%	*low fall*	TOM -1 Secundário	descendente média-baixa	L*L L%	descendente baixa nivelada
L-+H* L-L%	*rise fall*	TOM :1 Secundário	descendente alta-baixa	L+H* L L%	descendente alta-baixa
H* H- H%	*high rise*	TOM 3+	nivelada alta	H* H H^{L}%	nivelada alta

Fonte: própria.

O quadro acima tem a intenção de apresentar e descrever resumidamente o que foi feito durante esse momento da dissertação. Dessa maneira, trazemos os padrões auditivos e acústicos que foram primeiramente observados em cada modelo teórico. Para dar conta dessa caracterização, foram destacados o modelo acústico de Pierrehumbert (1980) e Ladd (1996) e o modelo auditivo proposto por Halliday (1970) e adaptado para o PB por Cagliari (1982/2007). Para a última metodologia, mostramos ainda a configuração do CT, essencial para a comparação entre os dois métodos entoacionais. Esses resultados aparecem, respectivamente, nas quatro primeiras colunas da tabela. Seguidamente, nas duas últimas colunas, foi possível fazer uma comparação entre o modelo acústico e o auditivo, a partir dos resultados alcançados nessa análise para o PB. Essa foi uma das etapas mais importantes do nosso trabalho, já que permitiu confirmar ou não a possibilidade de correlação entre uma análise auditiva e acústica. Isso foi feito tendo em vista o CT (ou nuclear).

Sendo assim, para cada tom primário ou secundário descrito neste texto, foi gerada a forma fonológica para o modelo acústico. A forma fonológica auditiva também precisou ser analisada, para sabermos se estava em conformidade com as que foram propostas por Halliday e Cagliari. Sendo assim, no que concerne ao modelo auditivo, chegamos a uma correspondência entre o padrão fonológico de Halliday/Cagliari e o resultante de nossa análise.

Para o tom 1 primário, o CT seguiu a mesma descrição proposta pelos autores, assim como os tons secundários :1 e 3+. Com os outros tons, as alterações foram pequenas, às vezes, variando apenas em relação ao início de um tom, como, por exemplo, o Tom -1, que começou a queda com um tom baixo e não como um tom médio. Outras vezes, além do sistema proposto por Halliday/Cagliari, também foi identificada uma outra opção para a representação desses tons, como ocorreu para os tons 3, primário e 1+, +1 e 1- secundários. Isso ocorreu devido à configuração do CPT, que varia de pessoa para pessoa, de acordo com a tessitura. Contudo, esse aspecto não interferiu na tendência do CT que, em geral, teve um padrão descendente para o tom 1 (sendo ele primário ou secundário) e ascendente para o tom 3 (primário e secundário). Assim sendo, para o tom 1 e suas variações, a tendência da curva melódica foi caracterizada no sentido da queda da Hz, podendo vir de H>L ou de L>L. Já para o tom 3 e suas variações, a tendência da Hz foi no sentido da elevação da Hz vindo sempre de um H>H.

Acusticamente, a nossa análise também corroborou os padrões propostos por Pierrehumbert no apêndice de sua tese de doutorado. Para chegar a esse resultado, foi feita, primeiramente, uma análise de acordo com a sua teoria e, em seguida, os resultados foram

confrontados, mostrando a identificação entre eles. Em outras palavras, os sistemas fonológicos encontrados em nossa análise também foram contemplados na abordagem da autora.

Posteriormente, partindo desses dois resultados, foi possível confrontar as duas análises e verificar o quanto uma análise acústica se aproxima ou não de uma auditiva. A tabela mais acima confirmou esse paralelo. Ou seja, quando confrontamos o CT a partir da tendência final da curva melódica nos dois sistemas fonológicos, chegamos a um resultado comum, seja descendente para o tom 1 ou ascendente para o tom 3 (primário e secundário). Por exemplo, auditivamente, o tom 1+ teve duas possibilidades para o CT: 1) descendente baixo e 2) descendente alto baixo. Acusticamente, também propomos duas possibilidades para esse tom: 1) H+L*LL% e 2) L+H*LL%. Observando mais de perto, vemos o quanto essa identificação é clara. Note que o primeiro padrão auditivo estabelecido para o CT é descendente baixo e, no sistema fonológico acústico, o *pitch accent* (de acordo com a teoria de Pierrehumbert) também parte de um tom baixo (L*), estando somente alinhado a um tom H. Por outro lado, o segundo padrão fonológico foi anotado para o primeiro com uma tônica que parte de um tom alto, assim como no padrão acústico (H*), o que também foi verificado para os outros tons do tipo 1. O mesmo acontece com o tom 3 primário e secundário, com uma tendência ascendente nas duas metodologias. No entanto, para esse tom, propomos também um tom L sobrescrito ao lado do tom H de fronteira.

O esquema trazido por Ladd (1996) e mostrado durante as análises foi mais uma maneira de afirmarmos o paralelo entre o modelo acústico e o auditivo. Assim como ele encontrou uma correspondência entre o sistema fonológico do inglês britânico e as 22 possibilidades de combinação entre o *pitch accent*, acento frasal e o tom de fronteira em Pierrehumbert, nós também conseguimos sistematizar o seu padrão de acordo com a análise acústica e auditiva obtida aqui. Por exemplo, o tom -1 secundário, de padrão auditivo descendente baixo nivelado, foi classificado de acordo com Pierrehumbert como L*LL% e, por Ladd, como *low fall*, ou seja, descendente. O mesmo acontece para uma das possibilidades do tom +1, descrito acusticamente como L+H* L L%, auditivamente como descendente alto-baixo, por etapas e, em Ladd, como *rise fall*. Como se vê, nas três descrições, a tendência à queda é a mesma, assim como o início da STS, que, nos três modelos, parte de um tom mais alto.

Esses foram apenas alguns comentários, mas essa reciprocidade foi reconhecida em todos os tons dos enunciados. Portanto, a comparação entre as duas abordagens mostra uma correspondência bem clara entre os modelos fonológicos aqui apresentados.

Feita a comparação entre o ponto de vista perceptivo e o caráter acústico dos enunciados, descrevemos a seguir os significados associados aos padrões entoacionais, bem como a estrutura sintática dos mesmos (oração declarativa, interrogativa, etc.). Com isso, observaremos o modo como a entoação relaciona-se com outros mecanismos linguísticos, ou seja, sua relação com a gramática da língua.

7.3 Análise: o significado dos enunciados de acordo com os diferentes padrões entoacionais

7.3.1 Observações gerais

Como visto, nossa análise toma como base um texto lido que, por sua vez, é diferente do discurso oral e da fala espontânea. Na fala espontânea, "[...] os elementos prosódicos ocorrem em grande quantidade com uma gama enorme de variação, o que mostra o alto grau de envolvimento do falante com seu discurso". (CAGLIARI, 1992, p. 60). Por outro lado, na leitura, segundo Cagliari (1992), o falante procura ler mais neutramente, com um ritmo mais constante e procurando conter suas emoções.

No texto escrito, o leitor pode ver, pela escrita, qual é a estrutura sintática das frases e, ainda, como elas devem ser lidas, de modo que a entoação é evocada pela leitura e deduzida pelo conjunto e pela pontuação do texto que dão margem ao contexto emotivo da história. (CAMARA, 1980). Os sinais de pontuação ajudam o leitor, assim como a escolha das palavras do texto e o conteúdo da história. Mesmo assim, isso não impede que ocorram problemas de leitura gerados, por exemplo, pela hesitação da fala, como repetição, omissão e substituição de palavras, erros de pontuação, entre outros.

Na gravação estudada, não tivemos uma leitura monótona e neutra, nem erros significantes de leitura. A forma escrita do texto representou muito bem os elementos prosódicos, indicando a maneira fonética como eles deveriam ser realizados na linguagem oral, por meio de pausas bem marcadas e palavras que caracterizam as atitudes e emoções dos personagens. Além disso, o falante soube interpretar o texto de maneira adequada, não se influenciando por uma leitura escolarizada.

Em relação à análise sintática, semântica e pragmática dos enunciados, é necessário levarmos em consideração algumas questões. Em primeiro lugar, as figuras referentes às análises dos enunciados serão apresentadas de acordo com a análise acústica e auditiva feita.

Por isso, é importante lembrarmos algumas questões metodológicas das duas teorias. As figuras serão agrupadas, no decorrer da análise, de acordo com o tipo de padrão entoacional e, a partir disso, apontaremos os seus valores sintáticos e semânticos.

Praticamente, não houve necessidade de reajustes na análise auditiva, a não ser na marcação de um ou outro nível tonal, em frequência baixa. Essa dificuldade com tons baixos é encontrada tradicionalmente quando se comparam as análises auditivas com as acústicas.

7.3.1.1 Apontamentos metodológicos

Nos enunciados apresentados a seguir, em cada uma das quinze figuras, a tonicidade foi atribuída a partir da análise auditiva da fala como um todo e não por meio da leitura individual de cada palavra. Os enunciados foram segmentados em sílabas, de modo que o intervalo entre uma sílaba tônica e outra constitua um pé rítmico (ABERCROMBIE, 1964, p. 96-98). A divisão em pés é marcada com barras inclinadas simples (/) e cada pé apresenta pelo menos uma sílaba acentuada (ou tônica) e um número relativo de sílabas não acentuadas.

As figuras mostram também a análise prosódica. O grupo tonal (GT) representa a unidade de informação semântica, ou seja, o que o locutor deseja transmitir com seu discurso. Os GTs são marcados com barras inclinadas duplas (//) em seu início e fim. Um GT caracteriza-se por apresentar um ou mais pés (P) e por ter uma sílaba tônica saliente (STS), que representa o foco. A STS está destacada por dois traços que a sublinha e ocorre imediatamente após as barras verticais duplas (‖). Durante a análise, os GTs serão analisados sequencialmente.

As sílabas tônicas silenciosas aparecem representadas pelo diacrítico (◌̭), mais especificamente, por um acento circunflexo subscrito, que ocorre imediatamente após os números que indicam o tipo de tom do enunciado. Os pés que iniciarem um grupo tonal sem apresentar uma sílaba acentuada no início "[...] terão uma sílaba tônica silenciosa, marcando o início do pé". CAGLIARI (2007, p. 163).

As três linhas horizontais, duas cheias e uma pontilhada (assim como o padrão utilizado por Cagliari, 1982/2007), representam os limites da altura melódica para a marcação dos tons (alto, meio-alto, médio, meio-baixo e baixo), ou seja, o valor auditivo relativo da variação melódica que foi atribuído a cada sílaba. Essa representação auditiva é feita por meio de bolinhas cheias. Abaixo dessas três linhas, mostra-se o valor obtido para a F0 em números.

Feita a anotação da F0, segue a análise acústica a partir do modelo autossegmental de Pierrehumbert (1980). Com base nesse modelo teórico, foram feitas anotações para os eventos tonais, identificados por H (*high*) ou L (*low*). Esses tons foram anotados como simples, quando monotonais (H ou L), ou complexos, quando bitonais (H+L ou L+H). Os tons com asterisco caracterizam o *pitch accent*, a sílaba tônica e os tons que não levam esses diacríticos, as sílabas átonas. Em um tom bitonal, temos os dois tipos de sílaba. Quando, em um enunciado, aparecem mais de um *pitch accent*, devemos levar em consideração que o *nuclaer pitch accent* (o foco do enunciado) é aquele que ocorre mais à direita do final do enunciado, de modo que as outras sílabas são os *prenuclear pitch accents*.

Também foram marcados os acentos frasais e os tons de fronteira. Os dois eventos tonais estão representados no final do enunciado. Nas anotações, o primeiro tipo de tom não recebe nenhuma marcação com diacrítico. Esse evento tonal ocorre sempre entre as marcações do último *nuclear pitch accent* e o acento de fronteira final que, por sua vez, traz informações sobre o final acústico dos enunciados. Esse evento tonal é anotado pelo símbolo de % e pode ser H ou L. O símbolo de % também aparece para separar as frases entoacionais, marcando as pausas em alguns enunciados.

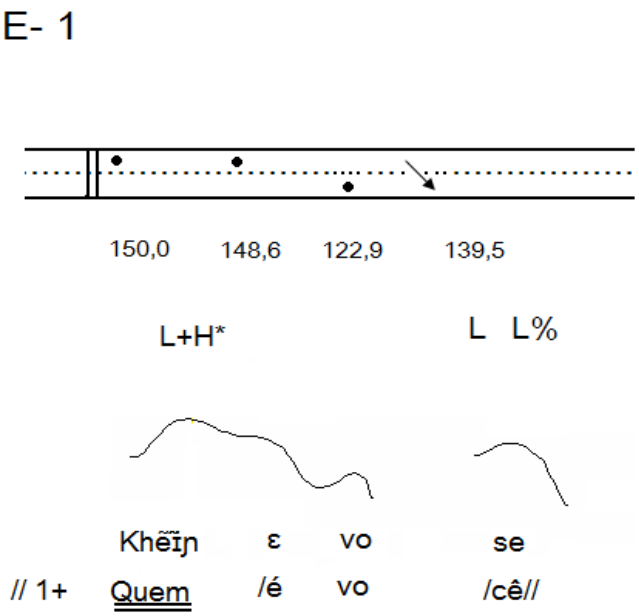
Abaixo dessa transcrição, destacamos o contorno da curva melódica de F0 para cada enunciado. Esse processo foi feito a partir do *praat picture*. Para isso, o primeiro passo que deve ser feito é abrir o enunciado e, pelo comando *edit*, gera-se uma janela. Nessa janela, por meio dos comandos *pitch* > *draw visible pitch countour*, abre-se a janela do *picture* com a respectiva configuração do contorno entoacional. Seguido a isso, essa figura é selecionada (pelo comando *print screen* do computador) e aberta no *software paint* do computador, para juntar-se aos outros componentes da análise.

Logo abaixo da representação do contorno entoacional, aparece a transcrição fonética acompanhada da transcrição ortográfica. Os tons dos padrões entoacionais vêm assinalados com os respectivos números, seguindo os tipos de tons propostos por Halliday (1970) e Cagliari (1982/2007). A escolha dos tons relacionou-se com os tipos sintáticos de sentença, como declarativas, interrogativas, exclamativas, entre outras; com os atos de fala, ou seja, se as frases indicam ordem, pedido, sugestão, etc.; e com as atitudes do falante, segundo seu valor semântico e pragmático, como expressões de surpresa, polidez, indiferença, destaque, entre outras. Os números que aparecem do lado esquerdo de cada figura, acima das três linhas horizontais, referem-se ao número de enunciados obtidos a partir da descompactação do texto, por exemplo, E-1, E-2, e assim por diante.

Feitos esses apontamentos, seguimos com a análise.

7.3.1.2 Análise dos enunciados

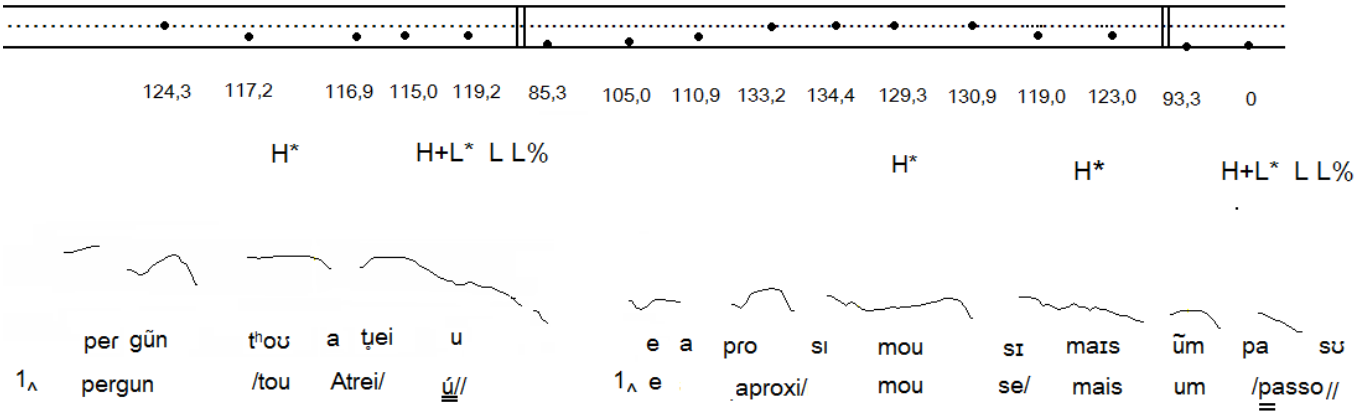
A figura E1 caracteriza-se por apresentar três GTs, respectivamente de tons 1+, 1 e 1. Como o enunciado é muito longo, por questões de espaço e visibilidade, o enunciado foi divido em dois momentos. No primeiro momento, temos o GT de tom 1+ e no segundo os GTs de tom 1.



Na primeira parte do enunciado 1, acima, o GT "Quem é você?" apresentou o tom 1+ secundário. Esse tom caracteriza-se por ter uma altura melódica (meio-alta) mais alta do que a do tom 1 primário, que se apresenta no nível médio constante. Nesse enunciado, a variação melódica começa em um nível meio-alto e, depois, segue com um padrão descendente até acabar em um nível baixo, com uma pequena ênfase na sílaba tônica (se), assim como indicado pela flecha. Não há variação do CPT, somente do CT.

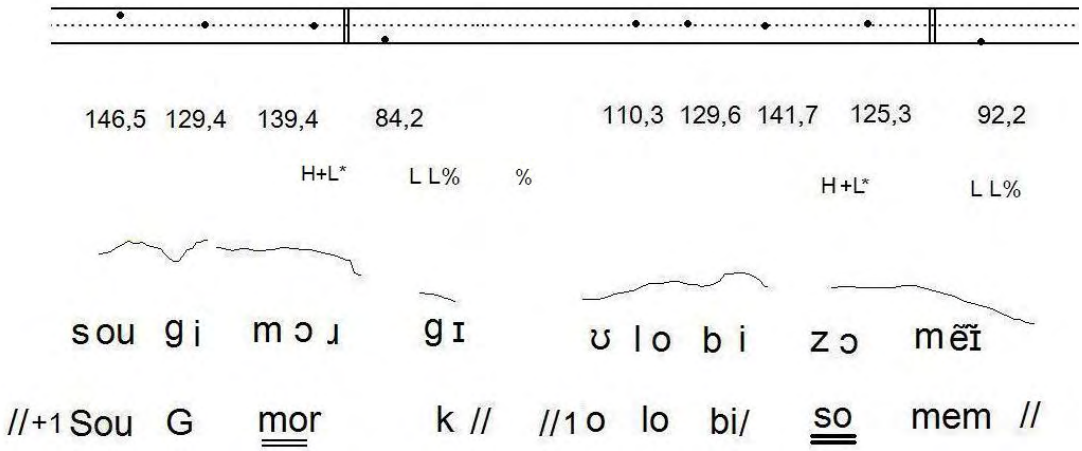
Esse padrão entoacional caracteriza-se por ter uma estrutura sintática interrogativa com palavra interrogativa, *Quem*, que, por sua vez, é o foco do enunciado. Como visto em Gebara (1976) é um enunciado não-polar típico, com palavra interrogativa em posição temática. Entoacionalmente, também segue um dos dois padrões mencionados pela autora, isto é: de tonicidade marcada, tom descendente e altura inicial alta, com tendência terminal baixa. O fato de o foco estar no início do GT acarreta um valor de curiosidade à pergunta, ou seja, o personagem está curioso, deseja saber quem é essa pessoa desconhecida e que surgiu do nada. É uma curiosidade que vem marcada também por uma precaução, por uma desconfiança. Por isso, uma tessitura inicial que se inicia mais alta e termina mais baixa.

E- 1



No segundo momento, representado acima, tivemos dois GTs com tom 1 simples: "perguntou Atreiú" e "aproximou-se mais uma passo", como mostrado acima. As duas orações são assertivas neutras e caracterizaram-se por apresentarem uma tessitura mais baixa e uma leitura pouco enfática em relação ao primeiro GT. Mesmo assim, o tom de curiosidade observado no primeiro GT continua, mas agora mais reservado, com um valor não tão forte e previsível quanto o anterior. No segundo GT, o CPT inicia-se com uma altura melódica mais alta (média) do que o terceiro GT (baixo), mas ambos iniciam-se com uma sílaba tônica silenciosa. O CT, nos dois GTs, é baixo e o foco está na última palavra do enunciado, sendo, no segundo GT, na última sílaba e, no terceiro, na primeira sílaba da palavra.

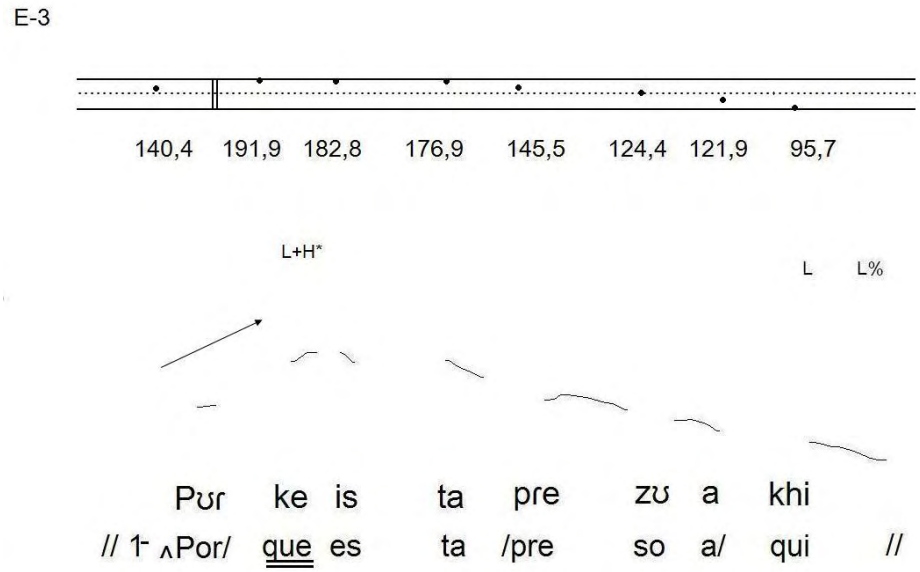
E- 2



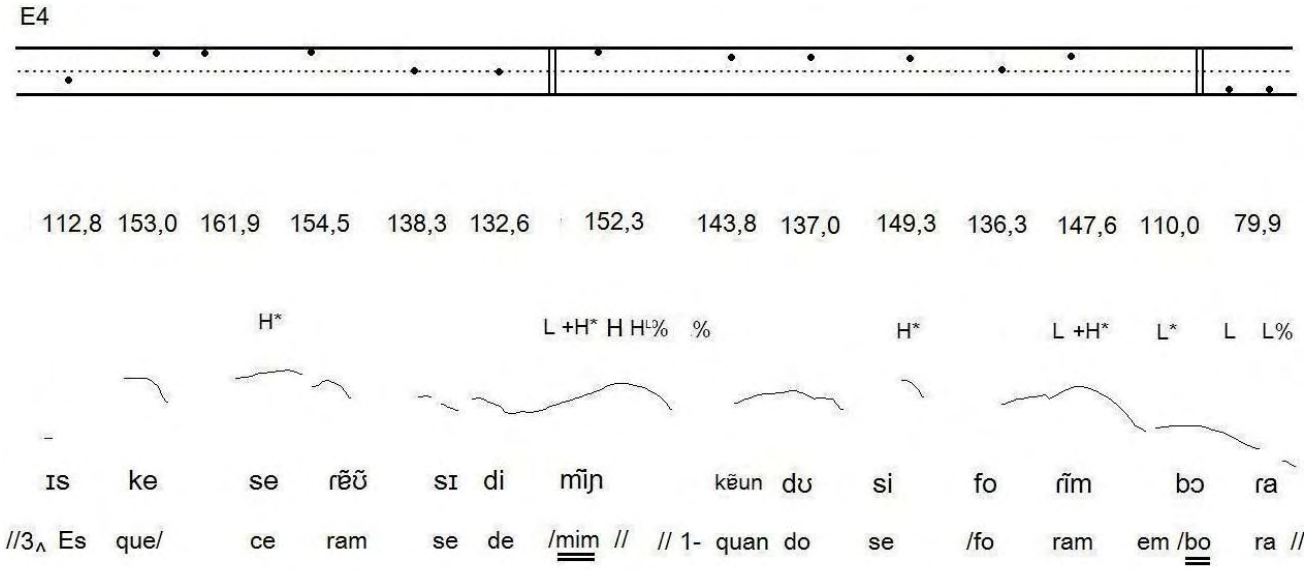
Em E2, os GTs foram dois, um de tom +1 secundário, “Sou Gmork”, e o outro de tom 1 simples, “o lobisomen”, os dois tons em sentenças declarativas. O tom +1 começa mais alto (meio-alto) do que o tom 1 simples do segundo GT, que se inicia em um nível médio. Além disso, esse tom caracteriza-se por ser uma variação do CT do seu tom primário correspondente. Nesse GT, houve um erro de leitura, mostrado na transcrição fonética. O leitor trocou a consoante (k) do final da palavra GmorK pela consoante (g), o que ocorre, às vezes, em leituras.

A tessitura começa com uma altura melódica mais alta no início do enunciado e depois começa a diminuir. Isso é próprio dos valores semânticos observados nos dois tons. No primeiro GT, a resposta é mais enfática. É como se o falante, interpretando a fala do personagem, quisesse sanar a curiosidade gerada pela pergunta anterior além de dar um valor irônico a sua resposta. É como se o falante já esperasse que o seu interlocutor soubesse quem era ele: “Como assim, quem sou eu? Sou Gmork, o lobisomem”. Depois disso, a expectativa do primeiro momento passa, caracterizando a tessitura mais baixa e a leitura mais neutra e pouco enfática do segundo GT.

Como se pode ver, a curva melódica acaba com uma altura melódica bem baixa no primeiro GT e, no segundo GT, começa em um nível médio. Acusticamente, o final do primeiro GT e o início do segundo são baixos, mas, auditivamente, tiveram classificações diferentes, um anotado como tom baixo, assim como os outros tons desse tipo, e o outro, como médio. Isso ocorreu no segundo GT porque há uma estrutura típica de platôs dando a sensação auditiva de que todos os tons estão na mesma altura, ou seja, auditivamente, não houve uma queda na F0 ou na tessitura. Esse padrão também prova a neutralidade da fala nesse GT. O fato de a curva melódica não ter se mantido constante entre um GT e o outro se deve à pausa (que está representada pelo símbolo de %) entre eles, o que também mostra a mudança no padrão entoacional e, consequentemente, no diferente significado semântico das duas unidades. Em outras palavras, revela a intenção do falante em fazer com que o seu interlocutor veja o que ele diz como uma declaração mais forte no primeiro GT e mais neutra no segundo.

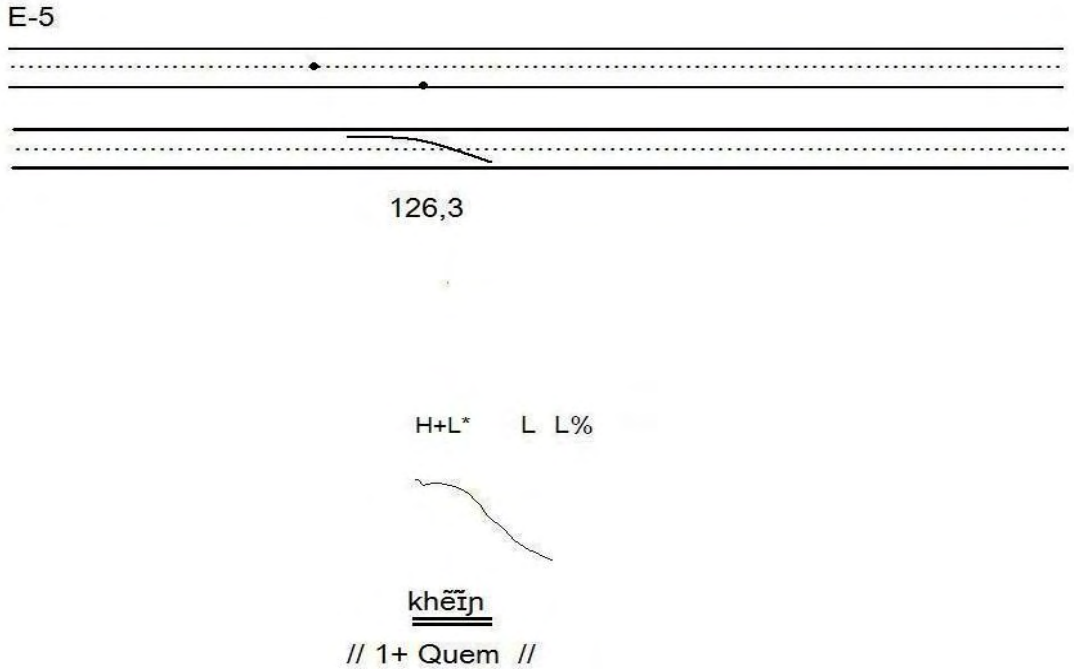


E3, por sua vez, foi classificado pelo tom 1˗. Esse tom caracteriza-se por apresentar, no componente tônico, uma variação melódica baixa, fato comprovado por meio do padrão descendente da curva entoacional. Esse padrão entoacional representa uma estrutura sintática de oração interrogativa com palavra interrogativa (que), com a mesma estrutura entoacional do primeiro GT de E1. Assim, por ser o foco do enunciado, a palavra interrogativa é responsável pelo aumento da tessitura (como indicado pela flecha) e pelo valor de curiosidade (como em E1) que a pergunta tem, ou seja, o falante quer saber a razão da prisão e não, por exemplo, o lugar em que ele está preso. É uma pergunta enfática, mas feita de maneira polida e suave.



Em E4, temos duas unidades informacionais que estão separadas por uma pausa. Os dois GTs têm uma estrutura sintática de frase assertiva, mas o primeiro, de tom 3 simples primário e o segundo, de tom 1^{-} secundário. O primeiro GT, "Esqueceram-se de mim", apresenta uma variação melódica que começa baixa, mantém-se em um nível alto ao longo do CPT e, no final do enunciado, mais especificamente na STS, passa a ter uma F0 mais alta e, novamente uma tendência final mais baixa. Essa característica é representada acusticamente pela notação H H^{L}% e é típica desse padrão entoacional. Esse tom também se caracteriza por ser suspensivo, ou seja, refere-se a um pensamento incompleto, suspenso, que sozinho é vazio de significação e, por isso, precisa de um complemento, que virá no GT seguinte do mesmo enunciado. E4 é uma resposta à pergunta feita em E3 e tem um valor semântico mais cortês e suave, de tom mais melancólico.

No segundo GT, "quando se foram embora", a expectativa do falante não mudou. Ele continua triste diante do ocorrido e mantém a polidez. No entanto, concede um destaque maior ao acontecimento, caracterizado pelo padrão descendente em nível alto que, no CT, passa a baixo.



Assim como o primeiro GT de E1, E5 também foi anotado com o tom 1+. Esse enunciado não tem uma estrutura sintática como a dos outros, mas é, por outro lado, um pronome interrogativo que funciona como uma pergunta. Assim, a função estabelecida é textual, e o enunciado deve ser interpretado de acordo com o contexto, ou seja, tanto em função do GT anterior quanto do que o segue. Halliday e Greaves (2008) nomeiam essa função de *textual metafunction*, definindo-a como: "*[...] is that aspect of the meaning of language that is concern with the creation of text, the ongoing flow of discourse that is systematically related to the context of the situation".* (HALLIDAY; GREAVES, 2008, p. 67-68) [125].

Reconhecemos que, nos enunciados anteriores, bem como nos subsequentes, a relação entre o que vem antes e o que vem depois é essencial para a interpretação do texto. Mas, a função estabelecida por eles não é textual e sim interpessoal[126], ou seja, uma função que, segundo Halliday e Greaves (2008), é obtida a partir da relação falante *vs* ouvinte, estabelecendo então relações sintáticas e semânticas no contexto: discurso/falante/ouvinte. De acordo com os autores, é uma relação tipicamente entoacional, prosódica. Por outro lado,

[125] "[...] o aspecto do significado da linguagem que está relacionado à criação do texto, com o fluxo contínuo do discurso, que está sistematicamente relacionado com o contexto da situação". [tradução nossa].

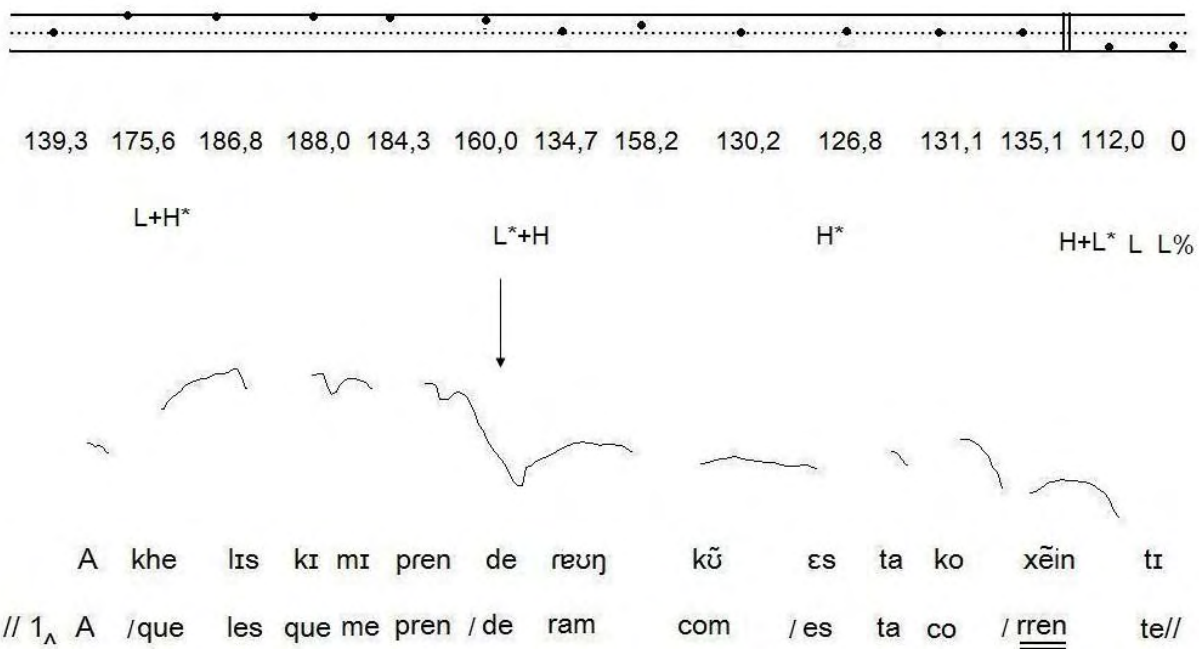
[126] Halliday e Greaves chamam essas funções de *metafunction* (metafunções) e as classificam em quatro tipos: *experiential*, *interpersonal*, *textual e logical*, respectivamente. Não entramos em discussão sobre a função de cada uma delas no texto. Os argumentos apresentados aqui foram muito sucintos e a proposta feita pelos autores é muito mais complexa do que a apresentada. Mas, para saber mais sobre o assunto veja (HALLIDAY; GREAVES, 2008, p. 67-79).

em E5, a relação é simplesmente textual e só pode ser interpretada em função do que vem antes e do que vem depois e não a partir de sua função entoacional e seu conteúdo sintático e semântico.

Mesmo assim, ao levarmos em conta o contexto em que o texto se insere, reconhecemos que a variação melódica descendente em E5 também apresenta um significado semântico, pois, assim como afirmam os autores, "*Language does not 'contain' meaning. Language consists of meaning*".[127] (HALLIDAY; GREAVES, 2008, p. 65). Assim sendo, interpretamos que nessa situação, bem como em E1 e E3, o falante está curioso em saber quem foi o responsável por tal ato e mostra-se indignado frente a essa situação tão cruel.

No que diz respeito à estrutura entoacional, o padrão segue as mesmas características dos enunciados anteriores com palavra interrogativa.

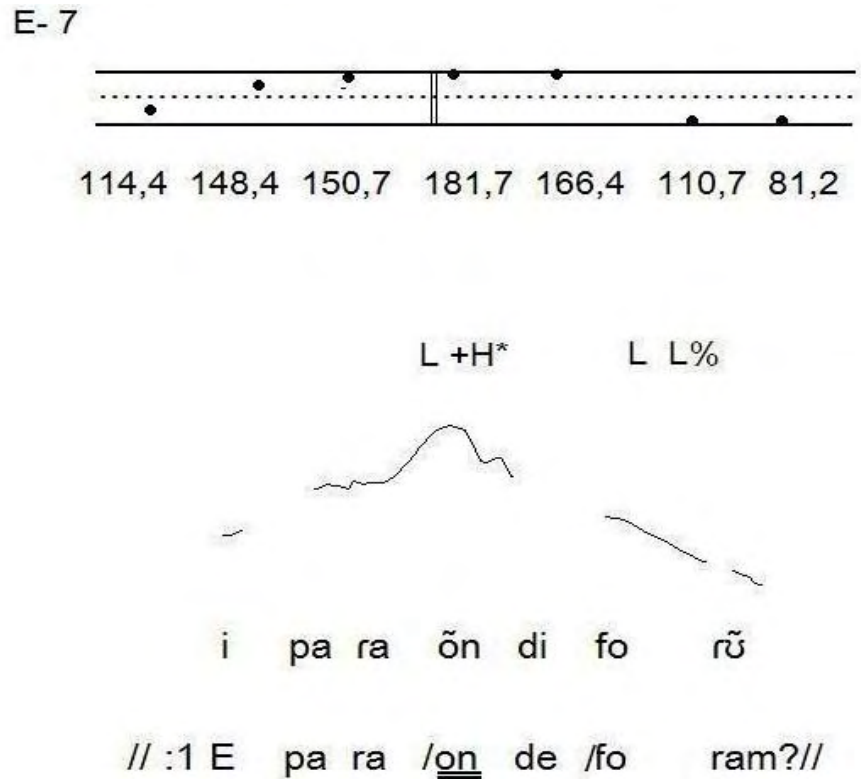
E-6



Assim como E1 e E2 (segundo GT), E6 foi anotado com o Tom 1 primário. Como esses GTs, esse também tem um padrão sintático de sentença assertiva. A leitura caracteriza-se por ser bem neutra e pouco enfática. Nesse caso, diferentemente dos outros, não há nenhuma indicação de curiosidade, mas, por outro lado, um tom de desapontamento. A variação melódica, no CPT, começa em um nível médio, tem um aumento ao longo do

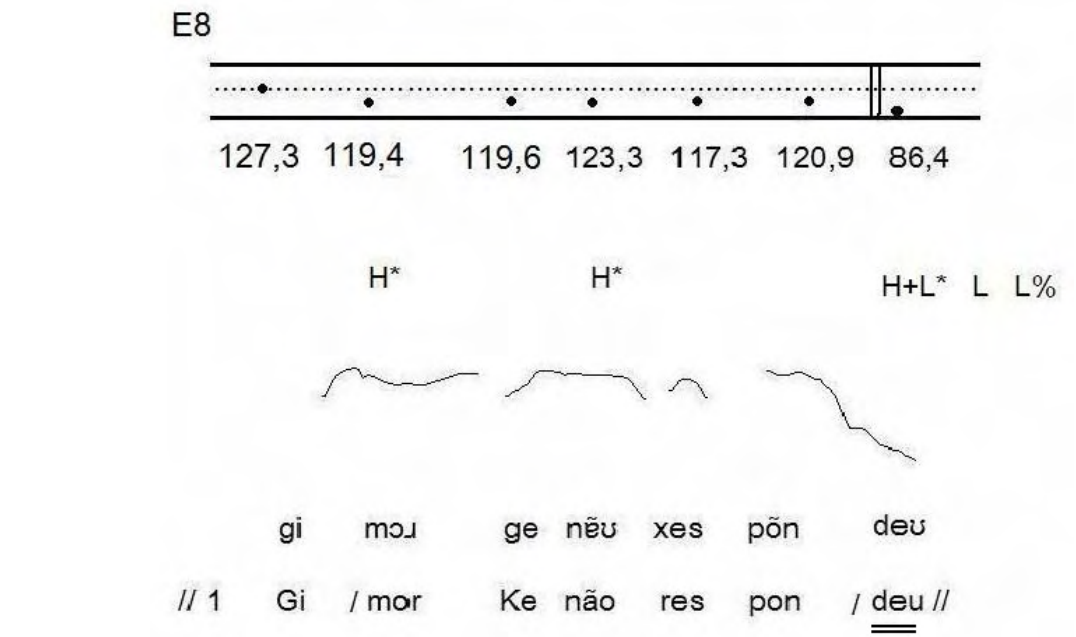
[127] "Linguagem não 'contem' significado. Linguagem consiste de significado". [tradução nossa].

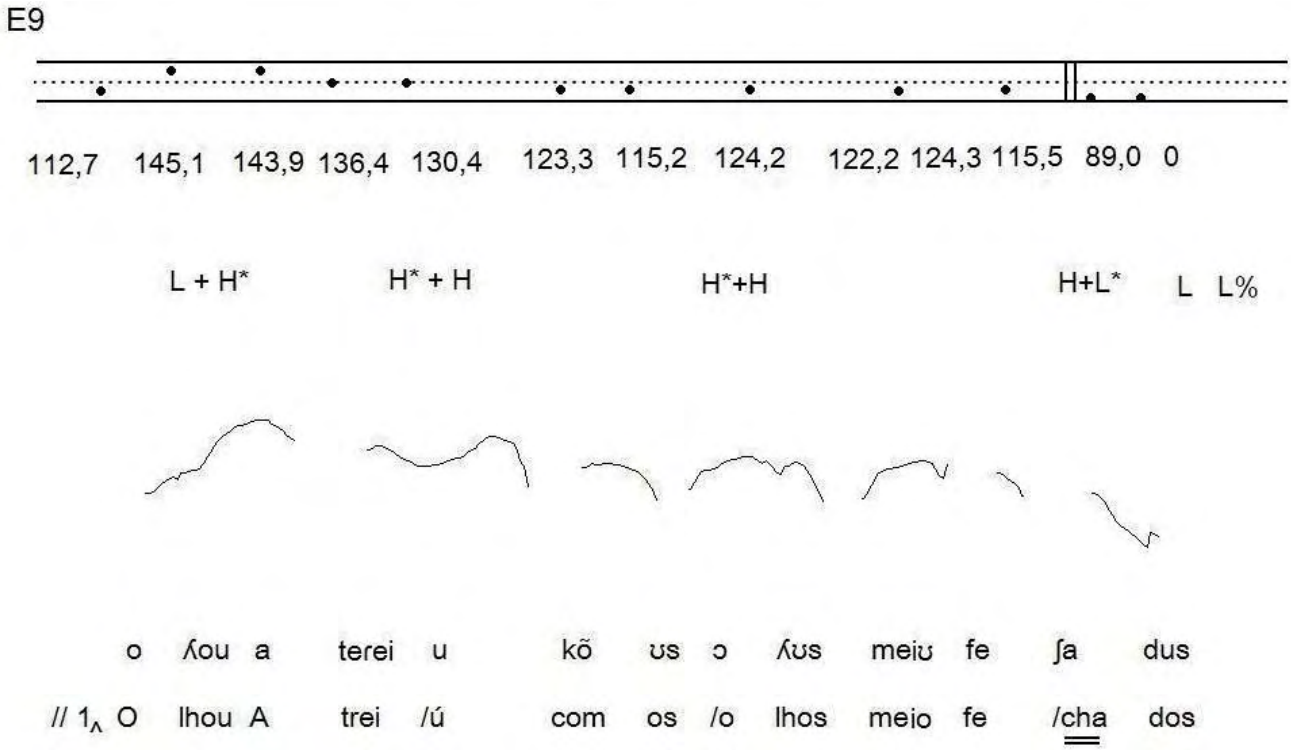
enunciado e, ao aproximar-se do CT, começa novamente a diminuir, passando do tom alto ao médio e, finalmente, no CT, passa de médio a baixo, o que caracteriza, ainda mais, o tom de decepção no decorrer da leitura. A flecha mostrada na figura acima identifica a queda da tessitura na sílaba tônica.



E7 foi anotado com o tom :1 secundário, variante do tom 1 normal. A variação melódica em E7 começa com um padrão ascendente no CPT e descendente no CT, sendo o pico de intensidade na palavra que representa o foco do enunciado (onde).

O padrão entoacional representa um significado sintático de oração interrogativa com palavra interrogativa *onde*, assim como em E3 e E5. O valor semântico atribuído a esse padrão, de acordo com o contexto, é o de preocupação. O falante quer saber quem foi o responsável pela situação. E7 também denota um valor de insistência. É como se o leitor não abrisse mão de uma resposta, chamando a atenção do interlocutor de uma maneira mais enfática.

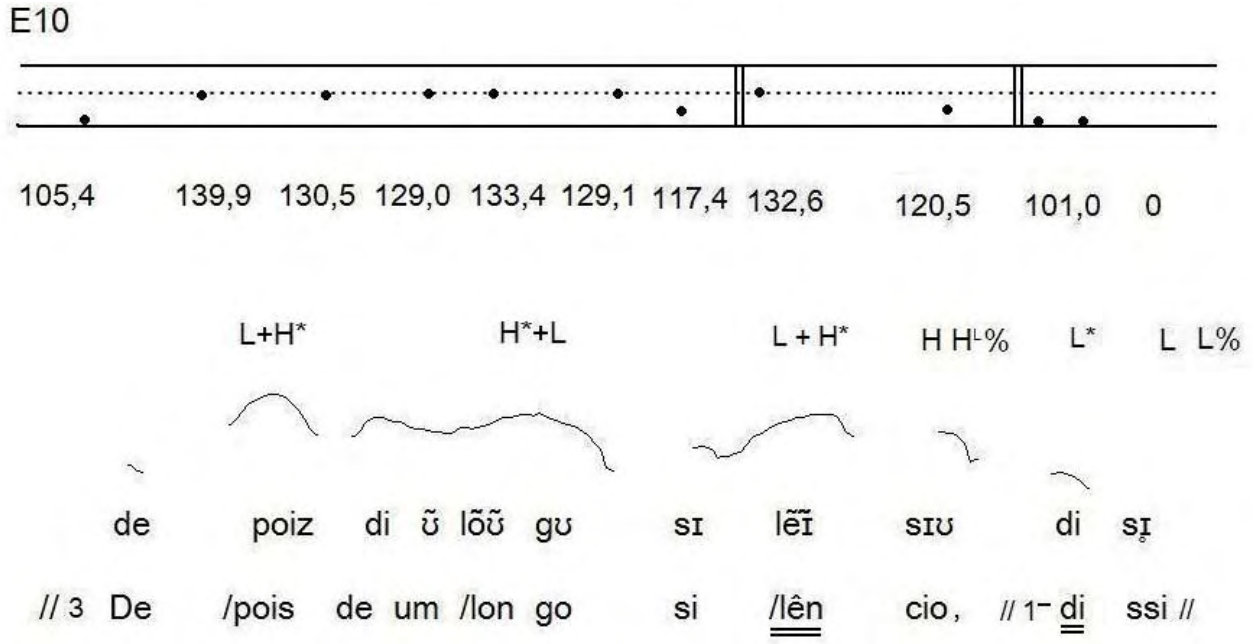




E8 e E9 foram classificados pelo tom 1 simples. No primeiro enunciado, a variação melódica começa mais alta do que no segundo, mas os dois seguem, no CPT, com um padrão meio baixo constante, terminando em um nível baixo no CT. Isso faz com que as duas orações tenham uma característica de platô, com leitura mais constante e, consequentemente, a incidência de tons baixo e meio-baixo.

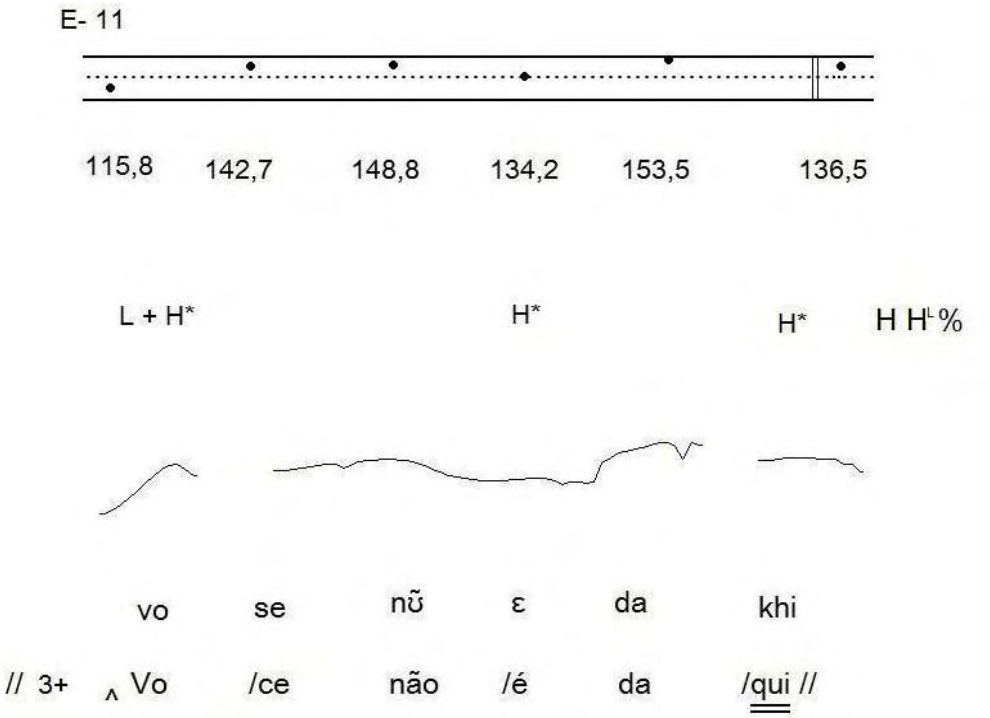
Sintaticamente, as duas orações são assertivas neutras. Nesses dois casos, a tonalidade mantém-se mais estável, não havendo sugestão de alguma característica

discursiva ou situacional mais relevante. Talvez, essa situação deva-se à atitude do narrador nesse momento da história, dando a impressão de querer distanciar-se dos conflitos. Mesmo assim, não podemos deixar de lado o valor de desapontamento em E8, mostrado pela leitura mais pausada e suave.

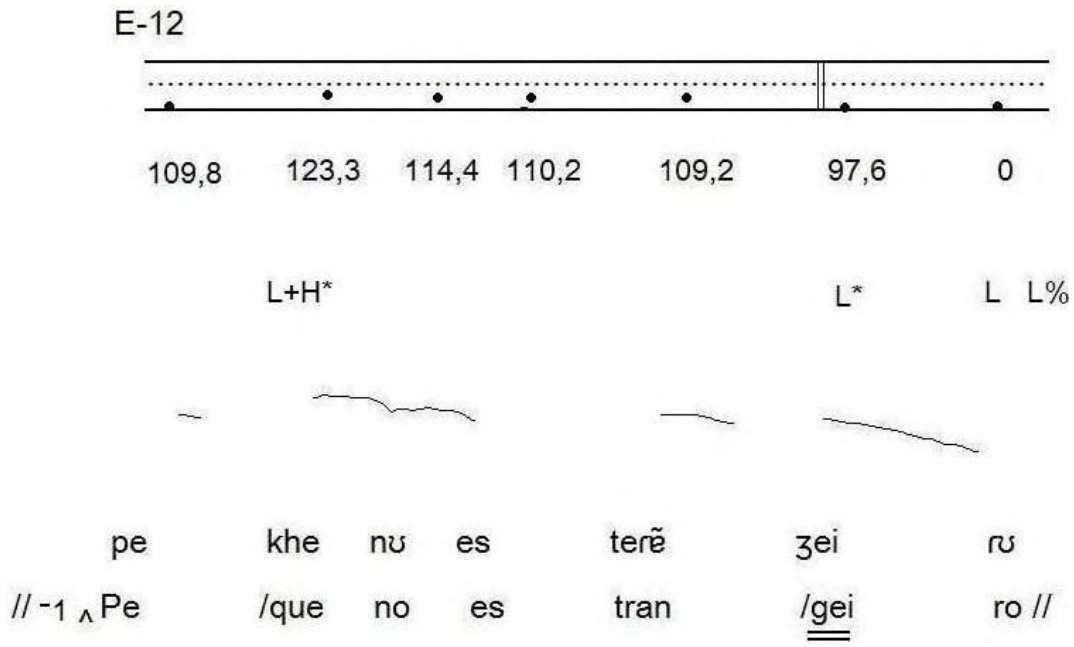


E10 dividi-se em dois GTs, um de tom 3 normal, “Depois de um longo silêncio”, e o outro, de tom 1- secundário, “disse”, assim como ocorrido em E4. No primeiro GT, bem como em E4, esse tom caracteriza-se por se referir a um pensamento incompleto. É como se o primeiro GT dependesse do segundo para ter sentido. No segundo GT, a variação melódica acaba no nível baixo e indica uma atitude mais suave e polida. Entretanto, diferentemente do segundo GT de E3, a atitude expressa pela entoação não tem um tom de melancolia, mas o sentido de fazer uma observação, esclarecer algo.

Tendo em vista nossos dados, quando há dois GTs em um mesmo enunciado e um deles é de tom 3, o outro GT será anotado, na maioria das vezes, com o tom 1 normal ou secundário. Tal fato foi observado em nossa análise nos enunciados E4 e E10 e também em E13 e E15 (como será visto). De acordo com as anotações de Cagliari (2007), esse resultado é válido e o autor acrescenta que essa relação de subordinação também pode ocorrer com o tom 2 normal. Contudo, Cagliari não observou essa relação para o tom 1 secundário.

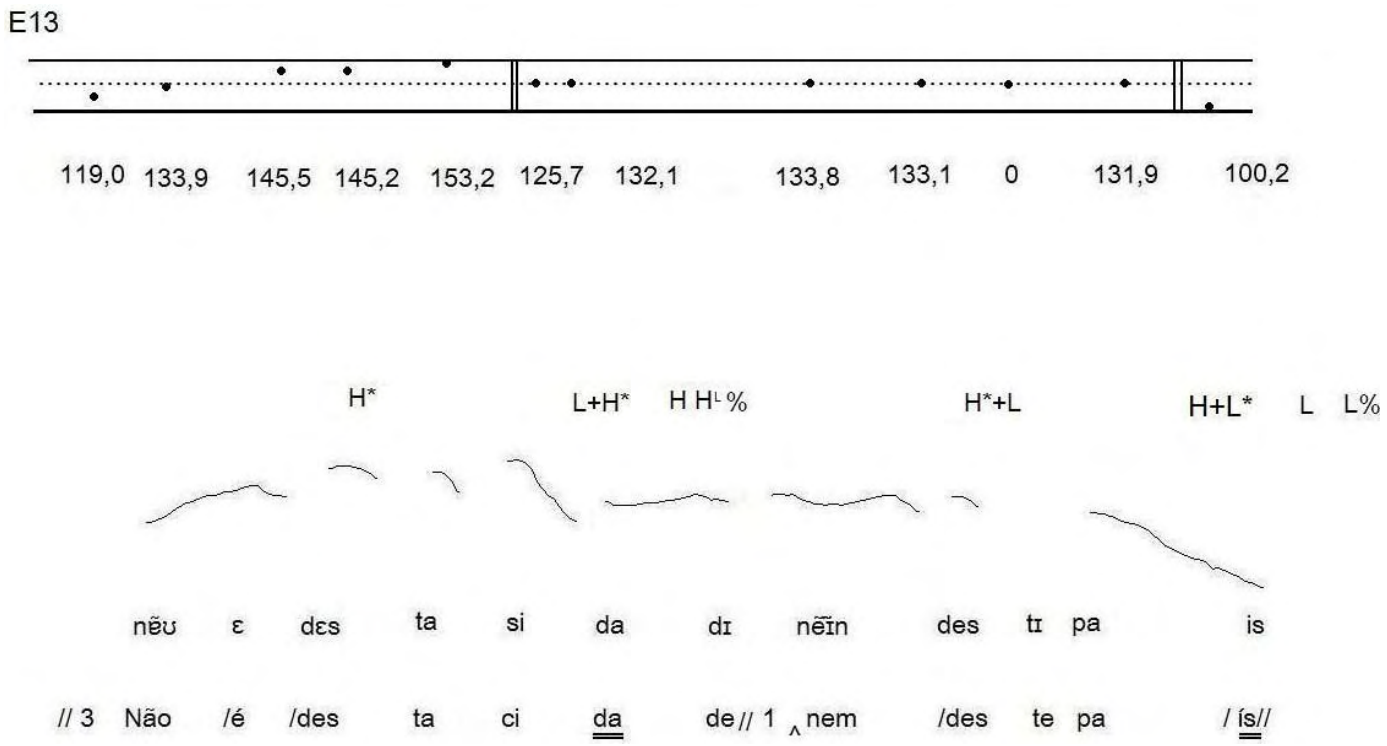


E11 apresentou o tom 3+ secundário, variante do tom 3 normal que, assim como ele, também caracteriza um pensamento suspenso. Acusticamente, o padrão final da curva entoacional com tendência à queda também teve a mesma caracterização do tom 3. No entanto, esse padrão melódico traz a ideia de uma afirmação categórica. É como se o leitor estivesse certo do que diz e por isso insiste em seu argumento.

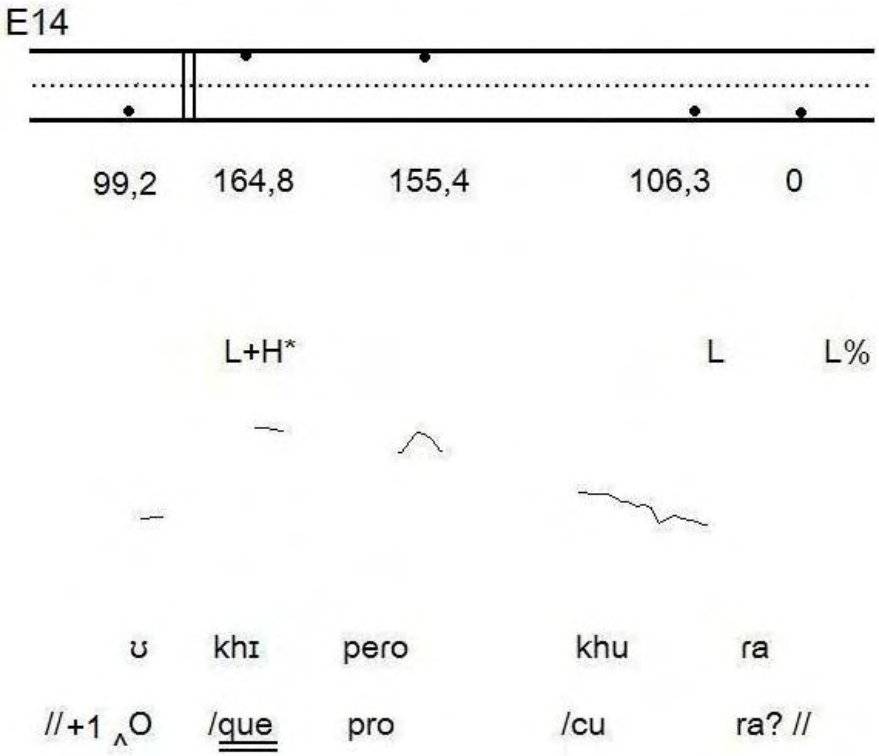


E12 foi classificado com o tom -1 secundário. Nesse enunciado, a variação melódica mantém um nível meio-baixo no CPT e decai, chegando a um nível baixo no CT. Esse padrão confere um valor mais vigoroso e ameaçador ao enunciado, gerado por uma

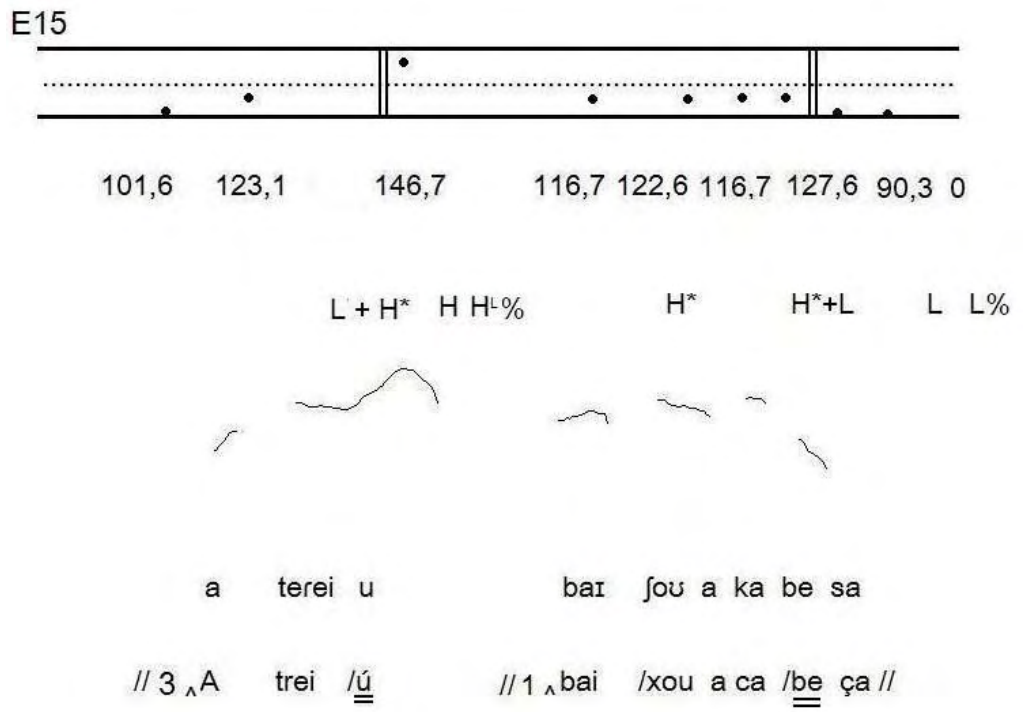
desconfiança, pois o falante não sabe, desconhece e, por isso, tem dúvidas em relação ao caráter dessa pessoa.



Novamente, em E13 (figura acima), temos uma relação de dependência entre as orações. O primeiro GT anotado com o tom 3 (incompleto) e o segundo, com o tom 1 declarativo, ambos simples. Nesse caso, o tom mais forte identificado na oração anterior (E12) diminui, gerando um valor mais neutro e pouco enfático aos dois GTs de E13. É como se o personagem quisesse distanciar-se do seu interlocutor, mantendo o tom de suspeita de E12.



E14 foi analisado com o tom +1. Assim como E1, E3, E5 e E7, classifica-se como uma frase interrogativa, com palavra interrogativa "que", ocupando a posição temática e, bem como os outros enunciados, apresenta padrão entoacional com tendência final descendente. Entretanto, nessa frase, diferentemente das outras, a atitude do falante, expressa pela entoação, indica uma pergunta requisitiva com um valor mais enfático e vigoroso. Nesse caso, é como se o falante exigisse do seu interlocutor uma resposta imediata, caso contrário ao das outras orações, que, mesmo querendo obter uma informação, apresentaram um padrão mais suave e curioso e não tão exigente e preocupado.



Por fim, o E15 foi anotado, em seus respectivos GTs, pelo tom 3 (suspenso) e 1 normal, assim como em E13. Discursivamente, no primeiro GT, temos um caso de sujeito que se torna tópico em Português, de modo que o que era uma informação dada (tema) torna-se o novo (rema). No segundo GT, a variação melódica vai do tom meio-baixo ao tom baixo e expressa neutralidade na declaração do leitor, além de indicar, implicitamente, a finalização da leitura.

Tendo em vista as relações sintáticas e semânticas propostas para esses 15 enunciados, descrevemos fonologicamente e sintaticamente, por meio do quadro 6 abaixo, um resumo das relações obtidas entre os diferentes padrões entoacionais e o significado produzido em cada um deles. Essa caracterização tem o intuito de ampliar a tabela 8, em que foi feita uma interpretação geral e sucinta dos padrões entoacionais do texto analisado e, ainda, um recorte simplificado da forma acústica mostrada no quadro 1, referente à segunda etapa da análise.

No quadro que segue, a primeira coluna caracteriza cada enunciado (de um a 15), a segunda, os tons respectivos de cada um deles e, a terceira traz a indicação acústica, a partir da tendência final da curva entoacional por meio do *nuclear pitch accent* (ou STS).

Quadro 6- Resumo dos significados produzidos pelos diferentes padrões entoacionais do texto analisado

Enunciado	**Tom**	**Padrão acústico do nuclear *pitch accent***	**Significado**
E1	1+ 1 1	L+H* L L% H+L* L L% H+L* L L%	pergunta com valor de curiosidade curiosidade com tom de reserva em sentença declarativa neutra curiosidade com tom de reserva em sentença declarativa neutra.
E2	+1 1	H+L* L L% H+L* L L%	declaração enfática declaração neutra.
E3	1^{-}	L+H* L L%	pergunta enfática suave com valor de curiosidade.
E4	3 1^{-}	L+H*H H^{L}% L* L L%	incompleto com tom de melancolia polido/suave
E5	1+	H+L* L L%	pergunta enfática com valor de curiosidade e indignação
E6	1	H+L* L L%	declaração neutra com tom de tristeza
E7	:1	L+H* L L%	frase interrogativa com valor de insistência
E8	1	H+L* L L%	declarativa neutra com valor de desapontamento
E9	1	H+L* L L%	declaração neutra
E10	3 1^{-}	L+H*H H^{L}% L* L L%	incompleto suave e polido com valor de esclarecimento
E11	3+	H*H H^{L}%	declaração categórica
E12	- 1	L* L L%	vigoroso e ameaçador com valor de desconfiança
E13	3 1	L+H*H H^{L}% H+L* L L%	incompleto declaração neutra
E14	+1	L+H* L L%	frase interrogativa com valor requisitivo e vigoroso
E15	3 1	L+H* H H^{L}% H*+L L L%	incompleto neutralidade

Fonte: própria.

Sendo assim, a tabela acima mostra, a partir do padrão entoacional e do significado semântico, sintático e pragmático atribuídos aos enunciados, que a fala, assim como explicam Halliday e Greaves (2008), não se constroi sozinha, *"[...] but always as (in some sense) the 'carrier' of meaning"*[128]. (HALLIDAY; GREAVES, 2008, p. 9). Consideração que não é nova, pois Jones (1917) e Pike (1945), em outras palavras, já afirmavam que a entoação, na linguagem, tem significados que se sobrepõem aos significados do dicionário, sendo estabelecidos de acordo com o contexto em que se inserem. Fato confirmado por meio das análises, uma vez que foi possível estabelecer significados sintáticos e semânticos para os 22 grupos tonais analisados. Além disso, também comprovamos o que já havíamos afirmado previamente, ou seja, que a entoação contribui para a construção dos significados, como, por exemplo, o tom 1+ secundário que, em nossos dados, foi um padrão típico de sentenças interrogativas com valor de curiosidade. Por sua vez, os enunciados de tons do tipo 1 e 3 simples não foram anotados com um significado atitudinal especial ligado a eles, o que é natural, pois, como vimos, tons primários são típicos de sentenças menos enfáticas. Portanto, nossos dados provam a viabilidade de correlacionar um tom a determinado padrão entoacional e, consequentemente, estabelecer um significado, estruturando o discurso em unidades de informação ou, conforme Halliday e Greaves (2008), por uma unidade gramatical realizada por meio de um tom. Para os autores, todo padrão sintático corresponde a um tipo de tom e, em decorrência, a um significado semântico, algo natural e comum na língua, segundo eles.

[128] "[...] mas sempre (de alguma forma) 'leva' um significado". [tradução nossa].

8 Conclusão

Conclusão

Esta dissertação tinha como objetivo fazer uma análise comparativa entre os dados acústicos e auditivos a partir de alguns dados do Português do Brasil. Nossa hipótese era a de que, quando correlacionadas, as análises acústicas e auditivas (e vice-versa) complementavam-se. Para isso, este estudo tomou como base dois modelos teóricos, um de base acústica (modelo Autossegmental) e outro de base auditiva (modelo Funcional), além da observação dos padrões entoacionais, acentuais e rítmicos e dos aspectos semânticos, sintáticos e pragmáticos dos enunciados.

No desenvolvimento da dissertação, nossa hipótese foi confirmada, mostrando que não só é possível como também viável estabelecer um paralelo entre as duas metodologias (acústica e auditiva), já que a análise acústica acoplada à análise auditiva ajuda a conciliar os detalhes da fala com as unidades fonológicas do sistema da língua.

Assim, para que pudéssemos chegar a esse resultado, seguimos os seguintes passos:

- Primeiro, fizemos uma comparação entre as duas abordagens teóricas, salientando suas semelhanças e diferenças;
- Depois, para a análise acústica, e com o auxílio do programa Praat, anotamos o valor da F0 para cada sílaba, para compará-la, em seguida, com os dados obtidos na análise auditiva. Para isso, desenvolvemos uma metodologia própria, suficiente para esse tipo de análise;

A partir disso, foi possível:

- chegar à análise dos enunciados e, consequentemente, comparar a forma fonológica dos dois modelos descritivos;
- comparar os níveis de altura melódica dos tons com a variação da frequência fundamental;
- relacionar os fatos acústicos da fala com interpretações linguísticas do valor da entoação na língua.

Portanto, tendo em vista esses procedimentos, conseguimos responder às propostas levantadas nos objetivos específicos desta dissertação, de modo que:

- ✓ A comparação entre o modelo auditivo (HALLIDAY, 1970 e CAGLIARI, 1982/2007) e acústico (PIERREHUMBERT, 1980), mostrou que mesmo sendo distantes teoricamente um do outro, apresentam aspectos formais que permitem estabelecer uma relação entre eles. Em outras palavras, a análise proposta por Halliday revelou-se um bom modelo para ser comparado com análises acústicas, enquanto a abordagem de Pierrehumbert mostrou que a interpretação dos tons, que vem de análises acústicas, também pode ter uma avaliação auditiva. Dessa maneira, os dois modelos teóricos ajudam a esclarecer os fatos que a outra abordagem não consegue tratar adequadamente;
- ✓ A análise da tessitura também revelou a possibilidade dessa correlação, pois estabelecidas as faixas de frequência, auditiva (alturas melódicas) e acústica (análise da F0), observamos a equivalência entre os dois modelos. As variações foram pouco expressivas e não foram suficientes para inviabilizar os resultados. Elas apenas mostraram os ajustes que o ouvido faz durante a percepção da curva entoacional;
- ✓ A conversão da forma fonológica dos dois modelos descritivos também ajudou na comparação entre eles e justificou, mais uma vez, a sua semelhança e proximidade;

  No que diz respeito aos fenômenos sintáticos, semânticos e pragmáticos, foi possível estabelecer uma relação com os tons propostos pelo modelo de Halliday/Cagliari, mostrando a viabilidade em se definir linguisticamente os padrões entoacionais a partir da gramática da língua. Os padrões entoacionais revelam as intuições do falante e do ouvinte, constituindo-se como parte do componente fonológico da gramática da língua, uma vez que a fonologia define o sistema oral da língua, revelando o que é e o que não é importante. Portanto, o significado entoacional decorre e se constroi a partir da união dos tons aos padrões de tonalidade (organização do discurso em unidades informacionais) e tonicidade (organização das unidades tonais) (HALLIDAY; GREAVES, 2008).

De um modo geral, vimos que há várias maneiras de realizar análises acústicas. O que as diferencia são as formas de aplicação, uma vez que alguns modelos são descritos do ponto de vista acústico (LADD, 1980; PIERREHUMBERT, 1980, entre outros), que leva em consideração as análises acústicas para revelar a frequência fundamental das ondas sonoras. Outros são descritos do ponto de vista auditivo (PIKE, 1945; ABERCROMBIE, 1967; HALLIDAY, 1970; CAGLIARI, 2007), levando em consideração a capacidade do ouvido humano em definir as alturas melódicas da fala e interpretá-las dentro do sistema

fonológico da língua. Nesse sentido, como feito nesta dissertação, chegamos à conclusão de que o mais viável é realizar trabalhos que façam, concomitantemente, uma comparação entre dados acústicos e auditivos, como foi feito, isto é, a partir da interpretação da entoação levando em consideração o ponto de vista perceptivo (fonologia funcional) e o ponto de vista acústico (fonologia autossegmental). Atualmente, os trabalhos referentes aos estudos entoacionais estão mais voltados para a questão estatística. Não há uma preocupação com a natureza auditiva, já que há uma dúvida quanto à sua confiabilidade. No entanto, este trabalho mostrou que esse tipo de pesquisa é possível, de modo que se voltarmos aos resultados obtidos veremos que em 135 tons, apenas em 6,66% dos casos houve uma variação entre o auditivo e o acústico. Fato que, como dissemos, pode ocorrer, uma vez que para o ouvido esses resultados são irrelevantes no conjunto da frequência. Além disso, também vimos na discussão referente aos estudos acústicos e auditivos (item), mais especificamente no item 3.3.4, que alguns foneticistas reconheceram na análise auditiva uma condição para a interpretação acústica, assim como Winckel, 1968; Cruz-Ferreira, 1983; Currie, 1979; Martin, 1975, entre outros.

Portanto, a união dessas duas análises, seja em que metodologia for, é essencial, pois, mesmo sendo o ouvido humano capaz de distinguir diferenças sonoras ao ouvir a linguagem humana, o cérebro processa um sistema linguístico de natureza fonológica e não leva em conta certas variações físicas da frequência fundamental ao observar a variação melódica da fala. O ouvido, segundo Cagliari (2002, p. 16), capta as variações melódicas em uma sequência contínua e ininterrupta como se apresentam na análise acústica; o cérebro, porém, age fonologicamente, interpretando a realidade física da variação melódica em termos de padrões linguísticos, chamados tons ou padrões entoacionais.

É por esse motivo que quem faz análise auditiva espera que os dados acústicos venham apoiar o que o ouvido processou. Se isto não acontecer, será preciso rever a interpretação auditiva ou as leituras dos dados acústicos.

Sendo assim, as conclusões a que chegamos podem ser tomadas como hipótese para que pesquisas posteriores desenvolvam mais análises que façam esse tipo de correlação, seja a partir de uma nova base de dados ou não. Pois, em geral, os trabalhos ou tratam a entoação como um fenômeno auditivo ou como um fenômeno acústico. Com isso, será possível ampliar os trabalhos desse tipo e trazer aos estudos linguísticos um novo método de análise entoacional.

REFERÊNCIAS

ABAURRE, M. B. M. Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. *Cadernos de estudos linguísticos*. Campinas, v. 2, p. 23-43. Semestral. p. 23-43, 1991.

ABERCROMBIE, D. *Elements of general phonetics.* Edinburgh: Edinburgh University Press, 1967.

ANTUNES, L. B. *O papel da prosódia na expressão de atitudes do locutor em questões*. 2007. 306f. Tese (Doutorado em Letras: Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

AUBERGÉ , V. A gestalt morphology of prosody directed by functions: the example of a step by step model developed at ICP. In: BEL, B.; M ARLIEN, I. (Org.). *Proceedings of Speech Prosody 2002*. Aix-en-Provence: Laboratoire Parole et Langage, Université de Provence, 2002. p. 151-155.

AUSTIN, J. L. *Quand dire, c'est faire*. Paris: Le Seuil, 1970.

BALL, M. J.; RAHILLY, J. *Phonetics: the science of speech*. London: Arnold. 1999.

BECKMAN, M.; PIERREHUMBERT, J. Intonational structure in Japanese and English. *Phonology Yearbook*, n. 3, p. 255-310, 1986.

BLOCH, B.; TRAGER, G. L. *Outline of linguistics analysis*. Baltimore, MD: Waverly Press, 1942.

BLOOMFIELD, L. *Language*. New York: Holt, Rinehart & Winston, 1933.

BODOLAY, A. N. *Pragmática da entoação*: a relação prosódia/contexto em atos diretivos no português. 2009. 300f. Tese (Doutorado em Letras: Estudos Linguísticos). Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2009. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/1843/ALDR-7R5NRD/1/3d.pdf>. Acesso em: 12 jun. 2012.

BOERSMA, P.; WEENINK, D. Praat: doing phonetics by computer (Version 5.1) [Computer Program]. Disponível em: <http//: www.praat.org/>, 2009.

BOIX, J. L. *Introducción a la fonética: el método experimental*. Barcelona: Anthropos, 1991.

BOLINGER, D. A theory of pitch accent in English. *Word*, n. 14, p. 100-149, 1958.

_____. Intonation: level vs configurations. *Word*, n. 7, p. 199-210, 1951.

BOLLELA, M. F. F. P. A prosódia como instrumento de persuasão. In LOUZADA, M. S. O.; NASCIMENTO, E. M. F. S.; OLIVEIRA, M. R. M. (Org.). Processos enunciativos em diferentes linguagens. Franca: Unifran, 2006. p. 113-128. Coleção Mestrado em Linguística, 1.Disponível em: <http://publicacoes.unifran.br/index.php/colecaoMestradoEmLinguistica/article/viewFile/386/313>. Acesso em 04 de abr. de 2012. p. 113-128.

BRUCE, G. *Swedish word accents in sentence perspective.* Travaux de L'Institut de Linguistique de Lund, CWK Gleerup: Mamö, 1977.

CAGLIARI, L. C. A entoação no componente fonológico da gramática: Halliday (1970) e Pierrehumbert. 2012. p. 1-18 (ms).

_____. A entoação no componente fonológico da gramática. 2011, p.1-3 (ms).

_____. *Comparação entre interpretação auditiva e dados acústicos na análise entoacional: estudo baseado na leitura de um trecho de uma história.* 2010. p. 1-25 (ms).

_____. *Elementos de fonética do Português Brasileiro*. São Paulo: Paulistana, 2007.

_____. Prosódia: ontem e hoje. In: FONSECA-SILVA, M. C.; PACHECO, V.; LESSA-DE-OLIVEIRA, A. S. C. (Org.). *Em torno da lingua(gem): questões e análises*. 1 ed. Vitória da Conquista: Edições UESB, 2007b, v. 1, p. 15-40.

_____. *Dossiê Prosódia*, 2002 (ms).

_____ Prosódia: Algumas Funções dos Supra-segmentos. CADERNO DE ESTUDOS LINGUÍSTICOS. Campinas: Ed. UNICAMP, v. 23, 1992-. Semestral. p. 137-151.

_____Da importância da prosódia na descrição de fatos gramaticais. In: *Gramática do Português falado: níveis de análise lingüística*, ILARI, R. (Org.). Campinas: Ed. UNICAMP, v. 2, 1992a-. p. 39-64.

_____. Elementos de fonética do Português Brasileiro. 1982. 192f Tese (Livre docência em Linguística) - Universidade Estadual de Campinas, Campinas, 1982.

_____. A entoação do Português brasileiro. In: *Estudos Lingüísticos*, v. 3. Araraquara: GEL: UNESP. 1980. p. 308-329.

CAGLIARI, L. C.; CARVALHO, M.; GEMENTI, M. M. A questão da fonética acústica e auditiva dentro e fora dos laboratórios. *GSCP*, Belo Horizonte, 2012, p. 1-24 (ms).

CAGLIARI, L. C; MASSINI-CAGLIARI, G. O papel da tessitura dentro da prosódia portuguesa. In: CASTRO, I.; DUARTE, I. *Razões e emoção*: miscelânea de estudos em homenagem a Maria Helena Mateus, 2003, v.1, p. 67-85. Disponível em: <http://www.fl.ul.pt/dlgr/mateus/mateus.htm>: Acesso em 18 jun. 2010.

CAMARA JUNIOR, J. M. *Princípios de linguística geral*. 6. ed. Rio de Janeiro: Padrão, 1980.

CATFORD, J. C. *A practical introduction to phonetics*. Oxford: Clarendon Press. 1988.

_____. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press. 1977.

CHOMSKY, N. *Aspects of the Theory of Syntax*. Cambridge: The MIT Press, 1965.

COOK, V. J. *Active intonation*. London: Longmans, 1968.

CÓRDULA, M. S. M. Questões metodológicas para o estudo da interface entoação e semântica. In: ANAIS DO VII CONGRESSO INTERNACIONAL DA ABRALIN. 2011, Curitiba, p. 2733-2741. Disponível em: <http://www.abralin.org/abralin11_cdrom/artigos/Maira_Cordula.PDF> Acesso em: 11 jun. 2011.

COUPER-KUHLEN, E. *An introduction to English Prosody*. Tübingen: Niemeyer, 1986.

CLARK, M. M. *A dynamic treatment of tone with special attention to the tonal system of Igbo*. Bloomington: Indiana University Linguistics Club, 1978.

CLEMENTS, G. N. The Geometry of Phonological Features. *Phonology Yearbook 2*, Great Britain. p. 225-252. 1985.

CRISTÓFARO SILVA, T. C. Dicionário de fonética e fonologia. São Paulo: Contexto, 2011. 239 p.

CRUZ-FERREIRA, M. M. F. Non-Native comprehension of intonation patterns in Portuguese and in English.1983. 421f. Tese (Doutorado em Filosofia) – Faculdade de Artes, Universidade de Manchester, Manchester, 1983.

CRYSTAL, D. *Dicionário de linguística e fonética.* Rio de Janeiro: Jorge Zahar, p. 275, 2000.

_____. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press. 1969.

CRYSTAL, D.; QUIRK, R. *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton, 1964.

CURRIE, K. L. *Intonation systems in Scottish English*. Tese (PhD thesis), University of Edinburgh, 1979.

ENDE, M. *História sem fim*. São Paulo: Martins Fontes, [19--].

ENDE, M.; FUCHSHUBER, A. *O Pequeno Papa-Sonhos***.** 5. ed. São Paulo: Ática, 1998, p. 19.

FANT, G. Descriptive analysis of the acoustic aspects of speech. In: LEHISTE, I. (Ed.). *Reading in acoustic phonetics*. Cambridge: The M.I.T. Press. 1967. p. 93-107.

FERNANDES, F. R. *Ordem, focalização e preenchimento em português*: sintaxe e prosódia. 2007. 445f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2007.

FERNANDES, N. H. *Contribuição para uma análise instrumental da acentuação e intonação do Português.* 1976. Dissertação (Mestrado em Linguística). - Universidade Estadual de São Paulo, São Paulo, 1976.

FOWLER, C. A.; GALANTUCCI, B. The relation of speech perception and speech production. In: PISONI, D.; REMEZ, R. (Ed.). *The handbook of speech perception*. Oxford: Blackwell, 2006. p. 633-652.

FROTA, S.; VIGÁRIO, M. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. In: APL, v. 1, 1999, Faro. *Actas do XV encontro nacional da APL,* Faro: APL,1999. p. 533-555.

FRY, D. B. Linguistic theory and experimental research. In: JONES, W. E.; LAVER, J. (Ed.). *Phonetics in linguistics: a book of readings.* London: Longman. 1973. p.66-86.

FUJIMURA, O.; ERICKSON, D. Acoustic phonetics. *The handbook of phonetic sciences.* Oxford: Blackwell, 1999. p. 65-115.

GEBARA, E. S. *Alguns aspectos da intonação no Português*. 1976. 109f. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Campinas, 1976.

GIMSON, A. C. *An introduction to the pronunciation of English.* London: Edward Arnold, 1962.

GOLDSMITH, J. *Autosegmental Phonology*. New York: Garland, 1976

GONÇALVES, C. A. V. Uma abordagem autossegmental para a morfologia. *Cadernos de Letras da UFF – Dossiê:* Difusão da língua portuguesa, n. 39, p.211-232, 2009.

HALLE, M.; VERGNAUD, J. R. *Metrical Structures in Phonology*. Cambridge: MIT, 1978.

HALLIDAY, M.A.K.; GREAVES, W, S. *Intonation in the grammar of English*. London: Equinox, 2008.

HALLIDAY, M.A.K. The Tones of English. In: JONES, W. E.; LAVER, J. (Ed). *Phonetics in Linguistics*. London: Longman, 1973. p. 103-126.

_____. *A course in spoken english*: Intonation. London: Oxoford University Press, 1970.

_____. Intonation and Grammar in British English. *Janua Linguarum*, Series Practica 48. The Hague: Mouton, 1967.

HAYES, B. *A metrical theory of stress rules.* Indiana University Linguistics Club Bloomington: Indiana, 1981.

HIRST, D.; DI CRISTO, A. *Intonation system: survey of twenty languages*. Cambridge: Cambridge University Press. 1998.

JOHNSON, K. *Acoustic and auditory phonetics.* 6. ed. Oxford: Blackwell, 2003.

JONES, D. *An Outline of English Phonetics.* 9. ed. Cambridge: Cambridge University Press, 1960. (1st ed. 1917).

KOHLER, K. Three trends in phonetics: the development of the discipline in Germany since the nineteenth century. In: ASHER, R. E.; HENDERSON, J. A. (Ed.). *Towards a history of phonetics.* Edinburgh: Edinburgh University Press, 1981, p. 161-178.

KINGDON, R. *The roundwork of English intonation*. London: Longmans, 1958.

KIPARSKY, P. Metrical structure assignment is cyclic. *Linguistic Inquiry*, n, 10, p. 421-441, 1979.

LACERDA, A. *Análise de expressões sonoras de compreensão*. Coimbra, 1950.

LADD, D. R. *Intonational phonology*. Cambridge: Cambridge University press, 1996.

_____. *The structure of intonational meaning*: evidence from English. Bloomington: Indiana University Press, 1980.

LADD, D.R. Stylized intonation. *Language* 54, 1978, p. 517-540.

LADEFOGED, P. The respiratory mechanism. In: EPSTEIN, M.; HACOPIAN, N. LADEFOGED, P. (Ed.). *Dissection of the speech production mechanism*. California: UCLA, 2002. p. 1-11.

_____. The value of phonetic statements. In: JONES, W. E. and LAVER, J. (Ed.). *Phonetics in linguistics: a book of readings.* London: Longman, 1975. p. 218-228.

_____. Units in the perception and production of speech. *Three areas of experimental phonetics*. London: Oxford University Press.1967, p. 143-172.

LADEFOGED, P.; JOHNSON, K. *A course in phonetics*. Canada: Wadsworth, 2011.

LAVER, J. *Principles of phonetics*. Cambridge: Cambridge University Press, 1994.

LEMLE, M. *Phonemic System of the Portuguese of Rio de Janeiro*. Dissertação (Mestrado). Universidade da Pensilvânia, 1965.

LIBERMAN, A. M. *Speech*: a special code. Cambridge: Bradford Books, 1996.

_____. *The intonation system of English*. PhD thesis, MIT, 1975.

LIBERMAN, A. et al. Perception of speech code. *Psychological Review*. New York: New York, v.74, n.6, p.431-461, 1967.

LIBERMAN, M.; PRINCE, A. On stress and linguistic rhythm. *Linguistic Inquiry,* n. 8, p. 249-336, 1977.

LIEBERMAN, P.; MICHAEL, S. Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. In: BOLINGER, D. (Ed.). *Intonation*. London: Penguin Books, 1972. [First published in Journal of the Acoustic Society of America, 1962].

LUCENTE, L. *DaTo*: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea. 2008. 72f. Dissertação (mestrado) – Universidade Estadual de Campinas, Campinas, 2008.

LUCENTE, L., SILVEIRA, L. S., BARBOSA, P. A. Declarativas em PB: downstepping ou nova combinação bitonal? H+!H* e H+L*. *IX Congresso Nacional de Fonética e Fonologia III Congresso Internacional de Fonética e Fonologia*. Publicação Eletrônica. p. 1-15, 2007. Disponível em:

<http://unicamp.academia.edu/LucianaLucente/Papers/289074/Declarativas_Em_PB_Downstepping_Ou_Nova_Combinacao_Bitonal_H_H_E_H_L_>. Acesso em 20 abr. 2012.

LUCENTE, L.; BARBOSA, P. A. *The role of alignment and height in the perception of LH contours.* In: Speech Prosody 2010, 2010, Chicago. Proc. of the Speech Prosody 2010, 2010. p. 1-4.

_____. Sistema DaTo de notação entoacional do português brasileiro: teoria e funcionamento. *Cadernos de Pesquisas em Linguística* (PUCRS), v. 4, p. 41-66, 2009.

LYONS, J. *Lingua(gem) e linguística:* uma introdução. Rio de Janeiro: LTC, 1897, p. 161-175.

MADUREIRA, S. Entoação e síntese de fala: modelos e parâmetros. In: SCARPA, E. *Estudos de Prosódia*. Campinas: Editora da Unicamp, 1999, p. 53-68.

_____. *A qualidade de voz como recurso fonoestilístico.* Intercâmbio, São Paulo, p. 99-120, 1991.

_____ Em torno da entoação: alguns problemas teóricos. *Cultura Lingüística*, n. 1, 1982, p. 63-78.

MALMBERG, B. The linguistic basis of phonetics. In:_____. *Manual of Phonetics*. Amsterdam: North-Holland Publishing Company, 1968, p.1-16.

MARTIN, J. G. Rhythmic expectancy in continuous speech perception. In: COHEN, A.; NOOTEBOOM, S. C. (Ed.). *Structure and process in speech perception.* New York: Springer-Verlag, 1975.

MASSINI-CAGLIARI, G.; CAGLIARI, L. C. Fonética. In: BENTES, A. C.; MUSSALIM, F. (Org.). *Introdução à linguística: domínios e fronteiras*. São Paulo: Cortez, 2001. p. 105-146.

MORAES, J. The Pitch Accents in brazilian portuguese: analysis by synthesis. *Fourth Conference on Speech Prosody*. Proceedings of the Speech Prosody. Campinas: Unicamp, 2008, p. 389-397.

MORAES, J. Intonation in Brazilian Portuguese. In: HIRST, D.; DI CRISTO, A. (Ed.) *Intonation Systems*. Cambridge: Cambridge University Press, 1998. p. 179-194.

MORAES, J. A.; STEIN, C, C. Attitudinal patterns in brazilian portuguese intonation: analysis and synthesis. *Speech Prosody. Dresden*, maio, p. 2-5, 2006. Disponível em: <http://www.isca-speech.org/archive/sp2006/papers/sp06_223.pdf>. Acesso em: 22 set. 2009.

NESPOR, M.; VOGEL, I. *Prosodic Phonology:* with a new foreword. Mounton de Gruyter. Berlin, New York, 2007.

_____. *Prosodic Phonology*. Dordrecht: Foris. 1986.

NOOTEBOOM. S. The prosody of speech: Melody and rhythm. 1999 In: HARDCASTLE, W. J.; LAVER, J. (Org.). *The handbook of phonetic sciences*. Disponível em:

<http://www.blackwellreference.com/subscriber/tocnode?id=g9780631214786_chunk_g978063121478621>. Acesso em 10 nov. 2009.

O'CONNOR, J.D.; ARNOLD, G.F. Intonation of colloquial English. London: Longman, 1973.

PACHECO, V. *O efeito dos estímulos auditivo e visual na percepção dos marcadores prosódicos lexicais e gráficos usados na escrita do português brasileiro*. 2006. 349f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2006.

PALMER, H.E. *English intonation, with systematic exercises.* Cambridge: Heffer, 1922.

PEREIRA, M. C. C. *A expressão da emoção em atos de fala do Português do Brasil.* Disponível em: <http://www.letras.ufrj.br/posverna/mestrado/PereiraMCC.pdf.> Acesso em: 15 jan. 2010.

PERES, D. O.; CONSONI, F.; NETTO, W. F. A influência da cadeia segmental na percepção de variações tonais. *LL Journal*. v. 6, n. 1, s/p. 2011. Disponível em: <http://ojs.gc.cuny.edu/index.php/lljournal/article/view/652/895>. Acesso em 10 jan. 2012.

PIERREHUMBERT, J. *The Phonology and phonetics of english intonation*. 1980. 402f. Tese (Doutorado em Linguística) – Department of linguistics and philosophy, Massachusetts Institute of technology, Indiana University Linguistics Club, 1980.

PIERREHUMBERT, J.; BECKMAN, M. *Japanese tone structure*. Cambridge: The MIT Press, 1988.

PIKE, K. L. *The intonation of American English*. Ann Arbor: The University of Michigan Press. 1945.

PRIETO, P. *Tonal alignment*. p. 1-15. Disponível em: <http://prosodia.upf.edu/home/arxiu/publicacions/prieto/prieto_tonal-aligment companionphonology.pdfhttp://prosodia.upf.edu/home/arxiu/publicacions/prieto/prieto_tonal-aligment-companion-phonology.pdf> Acesso em 10 de jan. 2012.

RAMEH, C. *Contrastive analyses of English and Portuguese intonation.* Dissertação (Mestrado). Washington: Georgetown University, 1962.

REIS, C. A. Analyse acoustique de la question en portugais. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 3, p. 135-162. 1995.

ROUSSELOT, P. J. *Les modifications phonétiques du langage étudiees dans le patois d'une famille Celefrouin*. Paris: H. Welter, 1982.

SANTOS, R. S.; SOUZA, P. C. Fonética. In: FIORIN, J. L. (Org.). *Introdução à linguística II*: princípios de análise. São Paulo: Contexto, 2005. p. 9-31.

SAPIR, E. A linguagem: introdução ao estudo da fala. Tradução de J. M. CAMARA JUNIOR. Rio de Janeiro: Sedegra Sociedade, 1954.

SCARPA, E.; FERNANDES-SVARTMAN, F. A estrutura prosódica das disfluências em Português Brasileiro. *Cadernos de estudos linguísticos*. Campinas, v. 54, n. 1, p. 25-40.

2012. Disponível em: <http://www.iel.unicamp.br/revista/index.php/cel/article/view/2571/2004>. Acesso em: 16 ago. 2012.

SEARLE, J. R. *Os actos de fala:* um ensaio de filosofia da linguagem. Tradução de C. VOGT. Coimbra: Almedina, 1984.

_____. *Les actes de langage*. Paris: Hermann, 1972.

SELKIRK, E. O. The interaction of constraints on prosodic phrasing. In: HORNE, M. (Ed.). *Prosody*: *Theory and Experiment*. Netherlands: Kluwer Academic Publishers, 2000, p. 231-261.

_____. On derived domains in sentence phonology. *Phonology Yearbook*, n. 3, p. 371-405. 1986.

_____. *Phonology and Syntax:* The relation between sound and structure. Cambridge: The M.I.T. Press, 1984.

_____. *On prosodic structure and its relation to syntax structure*. Indiana University: Linguistic Club, p. 1-31, 1980.

SERRA, C. R. *Realização e percepção de fronteiras prosódicas no português do Brasil:* fala espontânea e leitura. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, 2009.

SILVERMAN, K.; et al. ToBi: a standard for labeling English prosody. *Proceedings of the 1992 International Conference on spoken Language Processing.* Banff: Canada, 1992.

SOUZA, K. K. *Análise do fenômeno da declinação na entoação de sentenças declarativas isoladas dos falantes do Português Brasileiro*. 2007. 144f. Dissertação (mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte. Disponível em: <http://dspace.lcc.ufmg.br/dspace/bitstream/1843/VCSA77SQT3/1/dissertação_KarineKelvia.pdf>. Acesso em: 20 jan. 2010.

STEVENS, K. Quantal theory, enhancement and overlap. *Journal of phonetics*. Massachusstts, Cambridge, n. 38 p. 10-19. 2010.

_____. *Acoustic Phonetics*. Cambridge: MIT, 2000.

_____. On the quantal nature of speech. *Journal of phonetics*. Cambridge: Academic Press, n. 1, p. 3-45. 1989.

_____. The quantal nature of speech: evidence from articulatory-acoustic data. In: DAVID, E. E.; DENES. P. B. (Ed.). *Human Communication: a unified view*. New York: McGraw-Hill, p. 51-66. 1972.

STRAKA, G. *Album phonétique*. Quebec: Les Presses de l'Université Laval, 1965.

TENANI, L. E. *Domínios prosódicos no Português do Brasil*: implicações para a prosódia e para a aplicação de processos fonológicos. 2002. 331f Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, 2002.

TENANI, L. E.; FERNANDES-SVARTMAN, F. R. Prosodic phrasing and intonation in neutral and subject-narrow-focus sentences of Brazilian Portuguese. Fourth Conference on Speech Prosody, *Proceedings of Fourth Conference on Speech Prosody 2008*. Campinas, p. 445-448.

't HART, J.; COLLIER, R.; COHEN, A. *A perceptual study of intonation: an experimental phonetic approach to speech melody*. Cambridge: Cambridge University Press. 1990.

TRAGER, G. L.; SMITH, H. L. *An outline of English structure 41*. Oklahoma: Battenburg Press.

TRUCKENBRODT, H.; SANDALO, F.; ABAURRE, M. B . Elements of Brazilian Portuguese Intonation. *Journal of Portuguese Linguistics*, v. 8, p. 75-114. 2009. Disponível em: <http://www.iel.unicamp.br/docentes/fsandalo/texto02.pdf.> Acesso em: 28 de maio de 2012.

VASCONCELOS, C. A. et al. A expressão da atitude de certeza em indivíduos com perda auditiva bilateral: análise prosódica. *Anais do III Colóquio Brasileiro de Prosódia da Fala*. Belo Horizonte. Jun 6-8, 2011. Disponível em: <http://www.experimentalprosodybrazil.org/0632011.pdf>. Acesso em 12 jun. 2012.

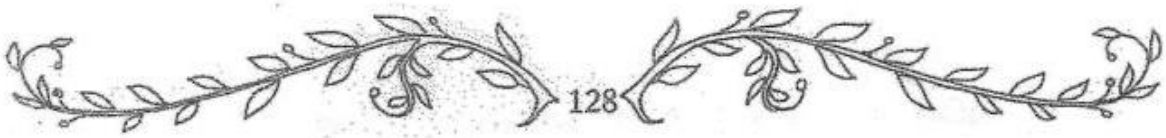
VIGÁRIO, M.; FERNANDES-SVARTMAN, F. R. A atribuição de acentos tonais em compostos no português do Brasil. In: BRITO, A. M.; SILVA, F.; VELOSO, J.; FIÉIS, A. (Org.). *XXV Encontro da associação portuguesa de Linguística – Textos selecionados*. v. 1, p. 769-786. 2010. Porto: Tip. Nunes, Ltda – Maia.

WINCKEL, F. Acoustical foundations of phonetics. In: MALMBERG, B. (Ed.). Amsterdam: North-Holland Publishing Company, 1968, p. 17-44.

YI, XU. In defense of lab speech. *Journal of phonetics*, London, v. 38, p.329-336, Apr. 2010.

ANEXOS

ANEXO A – Fragmento do texto *História sem fim*



— Quem é você?, perguntou Atreiú, e aproximou-se mais um passo.

— Sou Gmork, o lobisomem.

— Por que está preso aqui?

— Esqueceram-se de mim quando se foram embora.

— Quem?

— Aqueles que me prenderam com esta corrente.

— E para onde foram?

Gmork não respondeu. Olhou Atreiú com os olhos meio fechados. Depois de um longo silêncio, disse:

— Você não é daqui, pequeno estrangeiro, não é desta cidade nem deste país. O que procura?

Atreiú baixou a cabeça.

ANEXO B – Transcrição ortográfica e fonética do texto analisado

Quem é você?, perguntou Atreiú, e aproximou-se mais um passo
Khẽĩɲ ɛ vose pergũnthoʊ atᵻeiu, e aprosimousɪ maɪs˜um pasʊ

Sou GmorK, o lobisomem
sou gimɔɹgɪ ʊ lobizɔmẽĩ

Por que está preso aqui?
pʊr ke ista prezʊ akhi

Esqueceram-se de mim quando se foram embora
ɪskeseɾ̃ẽ̃ʊ sɪ di mĩɲ kẽundʊ si foɾ̃im bɔɾa

Quem?
khẽĩɲ

Aqueles que me prenderam com esta corrente
akhelɪs kɪ mɪ prendeɾ̃ẽ̃ʊŋ kũ̃ ɛsta koxeintɪ

E para onde foram?
i para õndi foɾ̃ʊ

Gmork não respondeu
gimɔɹge nẽʊ xespõndeʊ

Olhou Atreiú com os olhos meio fechados
oʌou atereiu kõ ʊs ɔʌʊs meiʊ feʃadus

Depois de um longo silêncio, disse:
depoiz di ũ lõũgʊ sɪlẽĩsɪʊ disᵻ

Você não é daqui, pequeno estrangeiro,
vose nũ̃ ɛ dakhi pekhenʊ esteɾ̃ɐʒeiɾʊ

não é desta cidade nem deste país
nẽʊ ɛ dɛsta sidadɪ nẽĩn destɪ pais

O que procura?
ʊ khɪ perokhura

Atreiú baixou a cabeça.
atereiu baɪʃoʊ a kabesa.