

UNIVERSIDADE ESTADUAL DE CAMPINAS

Reestruturações rítmicas da fala no português brasileiro

ALEXSANDRO RODRIGUES MEIRELES

Tese apresentada ao Programa de Pós-Graduação do Departamento de Lingüística do Instituto de Estudos da Linguagem, como requisito parcial para a obtenção do título de doutor em Lingüística.

INSTITUTO DE ESTUDOS DA LINGUAGEM

Campinas, fevereiro de 2007

Ficha catalográfica elaborada pela Biblioteca do IEL - Unicamp

M478r Meireles, Alexsandro Rodrigues.
Reestruturações rítmicas da fala no português brasileiro /
Alexsandro Rodrigues Meireles. -- Campinas, SP : [s.n.], 2007.

Orientador : Plínio Almeida Barbosa.
Tese (doutorado) - Universidade Estadual de Campinas, Instituto
de Estudos da Linguagem.

1. Ritmo da fala. 2. Taxa de elocução. 3. Sistemas dinâmicos. 4.
Fonética. 5. Fonologia. I. Barbosa, Plínio Almeida. II. Universidade
Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Título em inglês: Speech rhythmical restructurings in Brazilian Portuguese.

Palavras-chaves em inglês (Keywords): Speech rhythm; Speech rate; Dynamical systems; Phonetics; Phonology.

Área de concentração: Lingüística.

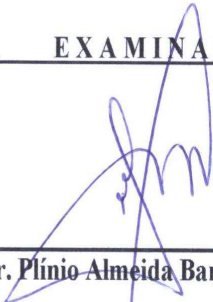
Titulação: Doutor em Lingüística.

Banca examinadora: Prof. Dr. Plínio Almeida Barbosa (orientador), Prof. Dr. José Olímpio de Magalhães, Profa. Dra. Maria Filomena Spatti Sândalo de Sá Porto, Profa. Dra. Sandra Madureira Fontes e Profa. Dra. Maria Bernadete Marques Abaurre.

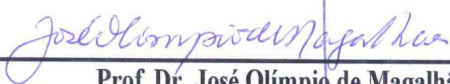
Data da defesa: 09/02/2007.

Programa de Pós-Graduação: Programa de Pós-Graduação em Lingüística.

BANCA EXAMINADORA



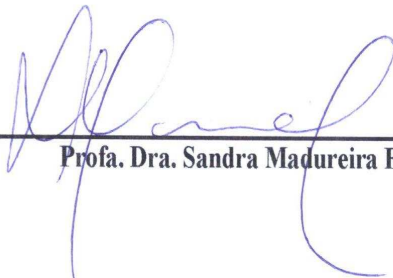
Prof. Dr. Plínio Almeida Barbosa



Prof. Dr. José Olímpio de Magalhães



Profa. Dra. Maria Filomena Spatti Sândalo de Sá Porto



Profa. Dra. Sandra Madureira Fontes



Profa. Dra. Maria Bernadete Marques Abaurre

Agradecimentos

Ao Professor Plínio Almeida Barbosa, meu sábio orientador, que, perturbando meu sistema nervoso instável, colaborou magnificamente na minha transição para um padrão dinâmico mais estável e eficiente de pensamento.

Às Professoras Maria Bernadete Abaurre e Sandra Madureira, pelas valiosas sugestões que contribuíram para o aperfeiçoamento deste trabalho.

Ao Professor José Olímpio de Magalhães, por me mostrar os rumos possíveis para uma nova abordagem lingüística.

Aos meus professores da UFMG, César Reis, Thaïs Cristófaró Alves da Silva, Seung Hwa-Lee, Marco Antônio de Oliveira, Mário Alberto Perini, por compartilharem um pouco de seu conhecimento.

Ao Professor Mike Dillinger, por me incentivar a trabalhar com programação de fenômenos lingüísticos.

Aos meus professores da UNICAMP, Maria Bernadete Marques Abaurre, Edson Françoso e Tânia Maria Alkmim, por me auxiliarem no meu percurso

acadêmico.

À Professora Eleonora Albano, por toda ajuda e compreensão nos meus anos de Lafape.

À Professora Dani Byrd e ao Professor Sungbok Lee da *University of Southern California* (USC), que me acolheram de braços abertos e me ajudaram a contribuir para a pesquisa com dados articulatórios no português brasileiro.

Aos meus amigos da USC, Rebeka Campos, Abe Kazemzadeh e Stephen Tobin, por me ajudarem na adaptação a Los Angeles e pelas discussões teóricas calorosas.

Aos pesquisadores do *Haskins Laboratories*, Louis Goldstein, Elliot Saltzman e Carol Fowler, pela hospitalidade e discussão teórica dos modelos prosódicos com o qual trabalhamos.

Ao Léo e à Flávia, que me acolheram calorosamente durante a minha estada no *Haskins Laboratories*.

Aos amigos do *Haskins Laboratories*, Hosung Nam, pela grande valorização de nosso trabalho com sistemas dinâmicos, e Minjung Son, pela cooperação e acolhimento.

À Ana Cristina e ao Cecílio, pelas nossas históricas “brigas” para desenvolver o programa *SilWeb* e pela grande parceria.

Aos velhos companheiros do Lafape, Laudino, Renato, Val, Maria Luiza,

Carmen, Léo, Sandra, Ricardo e Edmilson.

Aos meus amigos do ProFala, Luciana, Leandro e Sandra. Em especial, ao Pablo, pelas grandes discussões teóricas, e à Jussara, pela sua essencial colaboração e força no começo de minha jornada.

Aos funcionários da Secretária de Pós-Graduação, Cláudio Pereira Platero e Rosemeire Aparecida de Almeida Marcelino, pela ajuda prestimosa com a parte burocrática de meu curso.

À Capes, À Fapesp e ao CNPq, pelas bolsas que me foram oferecidas para me dedicar totalmente à pesquisa.

À música, por me acalmar em diversos momentos e me inspirar na busca de novos conhecimentos rítmico-melódicos.

Last but not least, minha família, minha sogra Lêda e D. Dolores, pelo suporte em todos os momentos, minhas sobrinhas Izabela e Milena, pela alegria, minha mãe Leila e meu pai Wallace, pela educação e, especialmente, minha esposa Joyce, calmaria para a minha turbulência, por tudo.

To Joyce, for letting me be.

Sumário

Primeira parte: Aspectos teóricos	1
1 Fundamentos Teóricos	2
1.1 Modelos dinâmicos	2
1.1.1 Flutuações críticas no deslocamento de gatos	8
1.1.2 Sinergia na fala	9
1.1.3 Coordenação motora na fala: produção e percepção . .	11
1.2 Modelos dinâmicos aplicados aos estudos lingüísticos	15
1.2.1 A gramática na TSD	15
1.2.2 Análises lingüísticas à luz da TSD	16
1.2.3 A abordagem dinamicista na produção da fala: dados e teoria	24
1.3 A Fonologia Articulatória	31
1.3.1 Motivação teórica	32
1.3.2 Gestos como estruturas articulatórias dinâmicas	33

1.3.3	O modelo dinâmico da tarefa	33
1.3.4	Constelações gestuais	39
1.3.5	Fala informal na FAR	42
1.4	Modelos dinâmicos prosódicos	45
1.4.1	A Fonologia Temporal	48
1.4.2	O gesto prosódico	60
1.4.3	O modelo dinâmico de produção do ritmo da fala . . .	66
	Segunda parte: Reestruturações rítmicas prosódicas	88
2	Ritmo: definição e aspectos teóricos	89
2.1	Ritmos biológicos	90
2.2	Ritmos espontâneos	93
2.3	Formação de grupos rítmicos	97
3	Estudos acústicos	101
3.1	Estudo preliminar	107
3.1.1	Metodologia	107
3.1.2	Análises e Resultados	113
3.2	Estudo complementar	120
3.2.1	Metodologia	120
3.2.2	Análises e Resultados	123

3.2.3	Frases 1 e 2	124
3.2.4	Frases 3 e 4	132
3.2.5	Frase 5	136
3.2.6	Frases 6 e 7	138
3.2.7	Frases 8 e 9	142
3.2.8	Frases 10 e 11	145
3.2.9	Análises quantitativas	149
3.2.10	Conclusão	161
4	Estudos articulatórios	163
4.1	Metodologia	165
4.1.1	O magnetômetro (EMMA)	167
4.1.2	Análises articulatórias no MAVIS	170
4.1.3	Desenho experimental	185
4.1.4	Sessão de gravação	187
4.1.5	Hipóteses e questões	187
4.2	Resultados	189
4.2.1	Detecção de grupos acentuais	189
4.2.2	Comparação entre a duração do VV articulatório e acústico	192
4.2.3	Efeitos gerais da taxa na duração e cinemática	195

4.2.4	Fontes cinemáticas de reestruturações rítmicas	201
4.3	Reorganizações rítmico-articulatórias	203
4.3.1	Duração da aceleração	210
4.3.2	Deslocamento vertical máximo	214
4.3.3	Deslocamento da constrição	217
4.3.4	Velocidade (pico ou vale)	222
4.3.5	Duração articulatória	225
4.3.6	Tempo-para-velocidade-de-pico proporcional	228
4.4	Conclusão	232
Terceira Parte: Reestruturações rítmicas no item lexical		233
5	Variações/mudanças rítmicas no léxico	234
5.1	Introdução	234
5.1.1	Diacronia x sincronia?	235
5.1.2	Mudança/variação histórica das proparoxítonas	239
5.1.3	Variação lexical na FAR	241
5.2	Estudos acústicos	244
5.2.1	Metodologia	246
5.2.2	Análise dos dados	252
5.2.3	Considerações sobre a influência da TE nas reestrutur- ações lexicais	271

5.2.4	Discussão geral	276
5.2.5	Conclusão	279
6	Variações articulatórias	281
6.1	Introdução	281
6.2	Metodologia	283
6.3	Análise dos dados	284
6.4	Considerações finais	292
7	Conclusões	294
A	Programas em Praat e Perl	299
A.1	Beat Extractor	299
A.2	Script do SGdetector	311
A.3	Duração	327
A.4	Rename	328
A.5	Formant Extractor	331
A.6	Dictionary	336

Lista de Figuras

1.1	Representação de um plano de fase	27
1.2	Variáveis de trato e seus articuladores	35
1.3	Equação de um sistema massa-mola amortecido	37
1.4	Pautas gestuais esquemáticas	40
1.5	Camadas rítmicas na FAR	47
1.6	Valores de notas musicais	52
1.7	Diferentes divisões rítmicas musicais	53
1.8	Relações harmônicas	55
1.9	Momento de ocorrência do pulso silábico	56
1.10	Ângulo de fase de <i>cards</i>	57
1.11	Ação do gesto prosódico	64
1.12	Modelo dinâmico do ritmo da fala	68
1.13	Segmentações em unidades VV	74
1.14	Ilusão de ótica na interpretação de imagens	76

1.15	Interação prosódia e segmentos	81
1.16	Hipótese gestual 1	84
1.17	Hipótese gestual 2	85
1.18	Hipótese gestual 3	86
3.1	Contornos duracionais em PB	105
3.2	SegmentaçãoVV	111
3.3	Metodologia dos estudos acústicos (taxa lenta)	112
3.4	Z-score em três taxas	117
3.5	Tendência à isocronia dos GA nas taxas rápidas	118
3.6	Contorno duracional da frase 1	127
3.7	Contorno duracional não-suavizado da frase 1	128
3.8	Contorno duracional da frase 2	129
3.9	Contorno de F0 de “abóbora nanica”	131
3.10	Contorno duracional da frase 3	133
3.11	Contorno duracional da frase 4	134
3.12	Contorno duracional não-suavizado da frase 4	135
3.13	Contorno duracional da frase 5	137
3.14	Contorno duracional da frase 6	139
3.15	Contorno duracional da frase 7	141
3.16	Contorno duracional da frase 8	143

3.17 Contorno duracional da frase 9	144
3.18 Contorno duracional da frase 10	146
3.19 Contorno duracional da frase 11	147
3.20 Espectrograma da frase 11	148
3.21 Duração do grupo acentual inter-taxas para a frase 11	158
4.1 Deslocamento mandibular hipotético	165
4.2 Magnetômetro (EMMA)	169
4.3 Traçados articulatórios no MAVIS	171
4.4 Funcionalidades do MAVIS	173
4.5 Gestos prontos para marcação	175
4.6 Velocidade derivada no eixo y	176
4.7 Sinal articulatório marcado	178
4.8 Deslocamento da mandíbula em “pavê”	180
4.9 Velocidades instantâneas	180
4.10 Delimitação dos gestos articulatórios	182
4.11 Medidas articulatórias para o experimento prosódico	188
4.12 VV articulatório x VV acústico	193
4.13 Fontes cinemáticas de reestruturações rítmicas	202
4.14 Reestruturação rítmica na sentença 1	205
4.15 Reestruturação rítmica na sentença 2	207

4.16	Evolução dos VVs articulatórios na sentença 3	209
4.17	Evolução dos VVs articulatórios na sentença 4	209
4.18	Duração da aceleração para a sentença 1.	211
4.19	Duração da aceleração para a sentença 2.	212
4.20	Duração da aceleração para a sentença 3.	212
4.21	Duração da aceleração para a sentença 4.	213
4.22	Deslocamento vertical máximo para a sentença 1	215
4.23	Deslocamento vertical máximo para a sentença 2	216
4.24	Deslocamento vertical máximo para a sentença 3	216
4.25	Deslocamento vertical máximo para a sentença 4	217
4.26	Deslocamento da constrição para a sentença 1	219
4.27	Deslocamento da constrição para a sentença 2	219
4.28	Deslocamento da constrição para a sentença 3	220
4.29	Deslocamento da constrição para a sentença 4	221
4.30	Velocidade articulatória para a sentença 1.	223
4.31	Velocidade articulatória para a sentença 2.	224
4.32	Velocidade articulatória para a sentença 3.	224
4.33	Velocidade articulatória para a sentença 4.	225
4.34	Duração articulatória dos gestos para a sentença 1.	226
4.35	Duração articulatória dos gestos para a sentença 2.	227
4.36	Duração articulatória dos gestos para a sentença 3.	227

4.37	Duração articulatória dos gestos para a sentença 4.	228
4.38	Tempo-para-velocidade-de-pico proporcional para a sentença 1.	229
4.39	Tempo-para-velocidade-de-pico proporcional para a sentença 2.	229
4.40	Tempo-para-velocidade-de-pico proporcional para a sentença 3.	230
4.41	Tempo-para-velocidade-de-pico proporcional para a sentença 4.	231
5.1	Diferenças lexicais na FAR	243
5.2	Realizações de [si] em três taxas de elocução	245
5.3	Segmentação acústica no Praat em três taxas de elocução	251
5.4	Taxas de elocução para o estudo lexical	254
5.5	Espectrogramas de “fôlego” e “cócega”	255
5.6	Formas de onda de “análise”	257
5.7	Reduções distintas nas três posições acentuais	260
5.8	Variação no F1 de “ó” em “cócega”	263
5.9	Representação fonológica (FAR) de abóbora	269
5.10	Forma de onda de um trecho de “abóbora”	272
5.11	Forma de onda de um trecho de “análise”	273
5.12	“Abobra” x “fósforo”	275
5.13	Espectrograma e forma de onda de “fosfro”	275
5.14	Espectrograma e forma de onda de um trecho de “fósforo”	276
6.1	Duração do intervalo consonantal em “bê b ada”	285

6.2	Duração do intervalo entre o início e o fim de “bêbada”	286
-----	---	-----

Lista de Tabelas

3.1	Duração das unidades VV: experimento preliminar	114
3.2	Duração dos grupos acentuais: experimento preliminar	115
3.3	Duração das unidades VV do estudo complementar	124
3.4	Desvio-padrão da duração das unidades VV do estudo com- plementar	151
3.5	Média e desvio-padrão dos VVs átonos do estudo complementar	152
3.6	Análises estatísticas para o desvio-padrão dos VVs átonos do estudo complementar	153
3.7	Duração dos grupos acentuais: experimento complementar . .	155
3.8	Análises estatísticas da duração dos GA: experimento comple- mentar	156
3.9	DVV/GA_{ig}	157
3.10	VV/GA para as frases do experimento complementar	161
4.1	Informações geradas pelo MAVIS	183

4.2	Sentenças do experimento articulatório prosódico	186
4.3	Diminuição do desvio-padrão da duração do VV articulatório .	194
4.4	Vogais utilizadas nas análises articulatórias	197
4.5	Deslocamentos verticais com o aumento da TE: deslocamento vertical máximo	197
4.6	Deslocamentos verticais com o aumento da TE: deslocamento da constrição	198
4.7	Análises estatísticas da duração da aceleração e AT	199
4.8	Variáveis articulatórias para a sentença 1	206
4.9	Fatores articulatórios para a sentença 2	208
4.10	Estatística da duração da aceleração	210
4.11	Estatística do deslocamento da constrição	218
4.12	Estatística da duração AT dos gestos	226
5.1	Média de F2 de algumas vogais do estudo acústico-lexical . . .	265
5.2	Taxa de decrescimento da duração das vogais	266
6.1	Duração da palavra com o aumento da TE	287
6.2	Intervalo consonantal com o aumento da TE	288
6.3	ICPs para todas as palavras	290
6.4	ICPs em função das palavras	291

Self-organizing systems have no deus ex machina, no ghost in the machine ordering the parts.

J.A. Scott Kelso, *Dynamic Patterns: the self-organization of
brain and behavior.*

Oscillation or rhythm is the cornerstone of life, of everything that isn't dead, of all the systems that are not in equilibrium.

J.A. Scott Kelso, *The complementary nature of coordination
dynamics: self-organization and the origins of agency.*

Resumo

Esta tese trata de reestruturações rítmicas da fala no português brasileiro ao se variar a taxa de elocução (*speech rate*). A reestruturação rítmica é considerada como uma reorganização entre grupos acentuais ao longo do enunciado ao se variar a taxa de elocução. A variação da taxa de elocução, por sua vez, é uma das maiores causas de modificação fonética, pois, perturbando a cadeia da fala dessa forma, pode-se avaliar as possibilidades de variação que revelariam novos padrões estáveis. A fim de explicarmos como a mudança da taxa de elocução pode modificar a estrutura rítmica da fala, dividimos a tese em três partes distintas.

Na primeira parte apresentamos os fundamentos teóricos que serviram de guia para os experimentos realizados. Nessa parte introduzimos a principal motivação teórica de nossa tese, a aplicação da Teoria dos Sistemas Dinâmicos à linguagem. Os pressupostos teóricos desse programa estão presentes na Fonologia Articulatória (Browman & Goldstein, 1992) e no Modelo Dinâmico do Ritmo de Barbosa (2006), utilizados como suporte teórico dos nossos experimentos.

Na segunda parte, após apresentarmos as noções básicas sobre o ritmo, relatamos um experimento acústico e um articulatório que revelam variações rítmicas nas frases estudadas com o aumento da taxa de elocução. Os prin-

cipais resultados para o estudo acústico, em trechos que se reestruturaram ritmicamente, foram: a) o número de unidades VV (vogal-a-vogal) por grupo acentual aumenta proporcionalmente ao crescimento da taxa de elocução; b) a duração do grupo acentual tende a se manter constante com o aumento da taxa de elocução; c) o desvio-padrão da duração das unidades VV, bem como da duração dos grupos acentuais é menor nas taxas rápidas. Já o estudo articulatório, realizado com a utilização de um magnetômetro (EMMA), revela que a taxa de elocução tende a afetar todos os gestos dos enunciados de maneira uniforme, independentemente dos mesmos pertencerem a uma unidade VV em posição de acento frasal ou não.

Na terceira parte, trabalhamos com variações rítmicas no item lexical. Especificamente estudamos a variação/mudança de proparoxítonas a paroxítonas com o aumento da taxa de elocução. Primeiramente, apresentamos aspectos históricos desta variação/mudança lexical, para, depois, apresentarmos uma análise à luz da Fonologia Articulatória. Para tanto, realizamos dois estudos: um acústico e um articulatório (EMMA). A principal conclusão para o estudo acústico foi de que a variação de proparoxítonas para formas paroxítonizadas é influenciada pela taxa de elocução, sendo que a probabilidade de encontrar formas consideradas como paroxítonas é maior nas taxas rápidas. No entanto, há fatores individuais/dialetais envolvidos, pois há falantes que nunca as produziram na taxa rápida. Já os estudos articulatórios serviram

para corroborar os resultados acústicos, pois, o aumento de coarticulação entre as consoantes pós-tônicas com o crescimento da taxa de elocução explicaria a percepção de proparoxítonas como paroxítonas, sobretudo nas taxas mais rápidas.

Concluindo, esta tese mostra que a explicação dinamicista de fenômenos lingüísticos de variação e mudança lingüística pode vir a conciliar aspectos fonéticos e fonológicos da linguagem. Especificamente, apresentamos aqui, como variações contínuas da fala, através do aumento da taxa de elocução, são capazes de modificar a estrutura rítmica da fala e atuar na variação/mudança lingüística.

Abstract

The present thesis deals with speech rhythmical restructurings due to speech rate variation in Brazilian Portuguese. Rhythmical restructuring is considered as a reorganization of stress groups along the utterance due to speech rate variation. Speech rate variation, on the other side, is one of the greatest causes of phonetic change, since this speech chain's perturbation makes possible to evaluate the possibilities of variation which would reveal new stable patterns. In order to explain how speech rate change modifies the rhythmic structure of speech, the thesis was split into three different parts.

The first part presents the theoretical background on which our experiments were based upon. In this part, it is introduced the thesis' main theoretical motivation, the application of Dynamical Systems Theory to language. Its theoretical basis is found in the Articulatory Phonology (Browman & Goldstein, 1992) and the Dynamical Model of Rhythm (Barbosa, 2006). Both models are used as our experiments' theoretical basis.

The second part, after presenting some basic notions about rhythm, presents an acoustical and an articulatory experiment which reveal rhythmic variations on sentences in which speech rate was varied. The acoustical study's main results in stretches with rhythmical restructurings are: a) the number of VV units (vowel-to-vowel) increases proportionally to speech rate

increase; b) the stress group's duration tends to be constant with speech rate increase; c) the standard deviation of VV units' duration, as well as the stress groups' duration, is smaller at fast rates. On the other hand, the articulatory study, through the use of a magnetometer (EMMA), reveals that speech rate tends to affect all gestures in a utterance, independently of the VV unit phrasal position.

The third part works with lexical rhythmic variation. Specifically, variation/change from antepenultimate stress words to penultimate stress words is studied. Firstly, historical aspects of this lexical variation/change is presented, then, an analysis by the light of Articulatory Phonology is presented. In order to do so, two studies were made: an acoustical and an articulatory one (EMMA). The acoustical study's main conclusion is that antepenultimate stress words variation to penultimate stress words is influenced by speech rate. Higher probability of penultimate stress words forms is to be expected at fast rates. Nevertheless, there are individual/dialectal factors involved, since there are speakers who never produced such forms at fast rates. On the other hand, the articulatory studies were used to corroborate the previous acoustical results, for coarticulation increase between post-stress consonants with speech rate increase would explain the perception of antepenultimate stress words as penultimate stress words, especially at fast rates.

Finally, this thesis shows how dynamical explanations of linguistic phe-

nomena like linguistic change and variation may come to reconcile phonetic and phonological aspects of language. Specifically, it presents how speech continuous variations through speech rate increase are able to modify speech's rhythmic structure and take part in linguist change/variation.

Primeira parte: Aspectos teóricos

Capítulo 1

Fundamentos Teóricos

1.1 Modelos dinâmicos

Como a teoria dos sistemas dinâmicos e seu respectivo vocabulário estará presente em vários momentos dessa tese, far-se-á necessária uma breve introdução sobre tal tema das ciências cognitivas. Ressalta-se, porém, que tal introdução não tem a finalidade de ser uma explanação exaustiva, mas sim de apresentar alguns de seus aspectos básicos, necessários à compreensão de sua aplicação à Lingüística feita neste trabalho, especificamente na área de fonética e fonologia. Para isto, faremos uma descrição baseada, principalmente, no livro *Dynamic Patterns: The Self-Organization of Brain and Behavior* de J.A. Scott Kelso, o qual aborda sucintamente a pesquisa na área dos sistemas dinâmicos nas últimas décadas.

Kelso (1995) apresenta os conceitos físicos de auto-organização e as ferramentas matemáticas não-lineares da teoria dos sistemas dinâmicos para entender como os humanos percebem, aprendem e coordenam comportamentos complexos. A grande contribuição de seu trabalho foi a acomodação de cérebro, mente e comportamento em um só programa de pesquisa, o dos sistemas dinâmicos auto-organizados.

O programa de pesquisa baseado nos sistemas dinâmicos auto-organizados é basicamente uma teoria de como os organismos vivos ou inanimados se coordenam para formar padrões distintos sobre determinadas forças externas ou internas e, conseqüentemente, sua utilização na ciência é chamada de abordagem dinamicista. Conceito fundamental aqui é o de auto-organização, que se refere à formação espontânea de padrões e mudança de padrões em sistemas abertos que não estão no estado de equilíbrio. Os sistemas são abertos pois interagem com o ambiente trocando energia, matéria ou informação com a sua vizinhança, e não estão no estado de equilíbrio pois sem essas fontes não podem manter sua estrutura ou função. A base teórica para tal abordagem dinamicista surgiu dos trabalhos de Hermann Haken, o qual introduziu o termo “sinérgica” em uma palestra na Universidade de Stuttgart em 1969 (Kelso, 1995, p. 291).

A sinérgica é um campo de investigação interdisciplinar que explica a formação e auto-organização de padrões e estruturas em sistemas abertos

fora do equilíbrio (vide Haken, 1983, 1984, para uma abordagem teórica introdutória). Tais “padrões dinâmicos”, uma vez identificados, são expressos, como o próprio nome sugere, através da linguagem matemática dos sistemas dinâmicos não-lineares, incluindo aspectos estocásticos (vide Verhulst, 2006, para uma introdução matemática às equações diferenciais não-lineares).

O programa de pesquisa dos sistemas dinâmicos é bastante distinto das idéias defendidas por Noam Chomsky (vide Chomsky, 1985, Chomsky & Halle, 1991), pesquisador que inaugurou no final dos anos 1950 o conceito de competência lingüística, um sistema internalizado de regras que gera as estruturas lingüísticas de uma língua. Esse conceito foi crucial para a emergência das ciências cognitivas modernas, baseadas no estudo da mente, as quais se opunham à teoria behaviorista de Skinner (vide Catania & Harnad, 1988, para uma visão retrospectiva de sua obra). No entanto o exarcebamento das idéias gerativas levou à difusão do cérebro como um sofisticado computador, cujo software é a mente. Essa concepção cria um problema citado pelos filósofos como “regressão infinita”, ou seja, o entendimento dos fatos científicos estaria sempre presente em um nível hierárquico superior (programa) ao da análise realizada.

No programa de pesquisa dos sistemas dinâmicos, a metáfora clássica dos cientistas cognitivos como Noam Chomsky, do cérebro ser um computador com conteúdos armazenados ou sub-rotinas a serem utilizadas por um pro-

grama, é radicalmente revista. Segundo Kelso (1995:34), “intelligent behavior may arise without intelligent agents — a priori programs and reference levels — that act intelligently”. Nesse programa não há uma dicotomia de estrutura e forma como na teoria chomskyana (competência e performance) e saussureana (língua e fala). Os conteúdos e representações emergem da tendência sistêmica dos sistemas abertos, fora do equilíbrio em formar padrões (vide Elman, 1995, para uma demonstração em redes naturais de tal processo na linguagem), i.e., os conteúdos emergem a partir dos componentes físicos do cérebro. Forma e conteúdo estão, pois, intrinsecamente ligados e não podem ser separados como nas teorias cognitivas tradicionais.

Outro aspecto distinto na abordagem dinamicista é a questão dos símbolos, elementos bastante ativos nas ciências cognitivas. Os mesmos assumem um papel bastante distinto na abordagem dinamicista, pois, apesar de existentes, são vistos como estruturas ou padrões estáveis duradouros, mas que, no entanto, são formados dinamicamente pelas atividades cerebrais, não-inatos, variantes no tempo e mutáveis.

Um exemplo prático da aplicação dos sistemas dinâmicos e de seu conteúdo explanatório está presente até em uma simples narração de um jogo de futebol, como descreveremos a seguir. Embora tal exemplificação do modelo dinamicista seja simples, abre-nos possibilidades para entender fenômenos complexos diversos, como mudanças meteorológicas, formação de nuvens e

turbulência em fluidos.

Imaginemos uma situação hipótetica em que um locutor esteja narrando pelo rádio uma partida de futebol entre o Cruzeiro e o São Paulo. Como podemos notar, este é um sistema dinâmico em que há acoplamento entre a narração e os eventos singulares no jogo, o que torna a taxa de elocução (*speech rate*, doravante TE) bastante acelerada nestas situações, pois, diferentemente da televisão, toda as jogadas devem ser transmitidas ao ouvinte simplesmente pelo recurso sonoro. Nesse sistema o deslocamento da mandíbula do locutor, principal responsável pela silabicidade, como veremos nos próximos capítulos, é chamado pelos físicos de *efeito coletivo* ou *cooperativo*, o qual surge sem nenhuma instrução explícita da motivação do deslocamento cíclico mandibular — a movimentação no campo. Como veremos na seção sobre a Fonologia Articulatória, um simples deslocamento mandibular é composto pela sinergia de diversos músculos, daí o efeito coletivo neste exemplo. Além disso, na locução, há uma sinergia de diversos articuladores, como a mandíbula, lábios e ponta de língua, para a realização de tarefas articulares. A TE do locutor é chamada nos sistemas dinâmicos de *parâmetro de controle*. A amplitude do deslocamento mandibular é chamada de *parâmetro de ordem* ou *variável coletiva*, definida por uma equação dinâmica não-linear (vide seção 1.3.3). Outros termos dinâmicos importantes aqui são o de *bifurcação* e *atrator*.

Uma bifurcação num sistema dinâmico é uma mudança qualitativa do atrator de um sistema dinâmico como consequência de uma mudança de um parâmetro do sistema. Já o atrator é definido como uma região ou ponto para o qual o sistema dinâmico tende, seja estático (atrator pontual), seja oscilatório (atrator cíclico). Para que uma estrutura seja considerada um atrator é necessário que as trajetórias próximas a ela continuem próximas, mesmo que levemente perturbadas. No jogo hipótetico citado, o atrator é algum evento singular no campo, o qual motiva a narração do locutor e estabelece momentos de mudança de estratégia/estilo na narração, como um grande drible, uma jogada rápida, um cruzamento de longa distância, etc. Nesse contexto, temos uma bifurcação no sistema quando ocorrem mudanças bruscas de estratégias de narração. Por exemplo, ao ocorrer um determinado evento, como um gol, ocorre uma mudança qualitativa do atrator. Tem-se, assim, uma bifurcação da amplitude do movimento da mandíbula, resultando, em um primeiro momento, na paralização do movimento (locutor gritando gooooooooool), e depois um deslocamento mais espontâneo, mais lento do que no decorrer da partida, como nos comentários proferidos pelo locutor tipo: “Que golaço fez o Fred, hein, Casagrande?” em um estilo próprio da fala cotidiana.

Após essa breve introdução à teoria dos sistemas dinâmicos, citaremos mais alguns exemplos de auto-organização em seres vivos. Para isto, pri-

meiramente daremos um exemplo de flutuações críticas no deslocamento de gatos, depois descreveremos um caso de sinergia na fala, e, por fim, mostraremos um experimento em que a produção e percepção da fala são ótimos exemplos da coordenação entre organismos nos sistemas dinâmicos.

1.1.1 Flutuações críticas no deslocamento de gatos

Flutuações críticas é um outro termo essencial na formação de padrões auto-organizados na natureza. Shik, Orlovskii & Severin (1966), por exemplo, em um experimento com um “gato mesencefálico”, através da estimulação elétrica da região mesencefálica do gato, fez com o que o mesmo modificasse seu padrão de deslocamento em velocidades críticas. Houve nesse experimento transições espontâneas de trote a galope em certas condições críticas de estimulação elétrica. No entanto, os autores relatam que em *regiões instáveis*, o animal às vezes trotava e às vezes galopava, ou seja, ocorreram *flutuações críticas* nesse tipo de experimento, fato que confirma a eficácia dos padrões dinâmicos em tal tipo de atividade biológica. Segundo Kelso (1995:74), “locomotion — the wondrous and numerous spatiotemporal orderings among the limbs — looks very much like a typical example of self-organization under the control of a single nonspecific parameter”.

Esse exemplo serve para mostrar que deslocamentos motores diversos —

mandíbula ou patas de animais — podem ser explicados segundo um mesmo princípio, ou seja, a auto-organização da teoria dos sistemas dinâmicos. Nas próximas seções, a fim de aprofundar nossa exemplificação com dados linguísticos, descreveremos casos de sinergia na fala.

1.1.2 Sinergia na fala

Sinergias são unidades funcionais que agem flexivelmente e temporariamente para o desempenho de uma determinada tarefa. Sendo assim, se perturbarmos um sistema, ele deverá se ajustar imediata e espontaneamente para preservar o alvo funcional da tarefa. A fala seria um local ideal para testar a hipótese de sinergia, pois para produzirmos uma simples sílaba como [ba] fazemos uso de aproximadamente trinta e seis músculos. O que ocorreria se perturbássemos uma produção silábica na fala? Continuaríamos a desempenhar a tarefa adequadamente? A fim de testar tal efeito na fala, Kelso, Vatikiotis-Bateson & Fowler (1984) realizaram o experimento descrito abaixo.

Nesse experimento a mandíbula foi perturbada por se mover constantemente para todos os tipos de sons e por sabermos que, se a bloquearmos com um lápis ou um cigarro, ainda podemos continuar a falar sem maiores problemas. Na observação de sinergias, os pesquisadores perturbaram a mandíbula

no decorrer da fala para observar se haveria uma reorganização instantânea dos articuladores para se adaptarem à tarefa de produzirem um [b] ou [z].

Sensores de luz infravermelhos foram utilizados para obtenção de movimentação dos lábios e mandíbula. Eletrodos foram inseridos nos músculos da fala como o genioglossus para monitorar a atividade eletromiográfica (EMG). Além disso, uma prótese de mandíbula feita de titânio foi utilizada em dois informantes. Os resultados revelaram que assim que a mandíbula foi bloqueada por alguns milissegundos no [b] final de [bæb], os lábios inferiores e superiores imediatamente compensaram tal perda de movimento para produzir o [b], embora não tenha havido nenhuma modificação nos movimentos da língua. Por outro lado, quando a mesma perturbação foi feita durante a produção de [z] em [bæz], a atividade muscular da língua aumentou rapidamente para atingir a configuração palatal necessária para a produção de um som fricativo, embora nenhuma compensação tenha ocorrido para a movimentação dos lábios.

Com esse experimento, os pesquisadores conseguiram provar que a forma de cooperação observada na fala, de acordo com Kelso (1995:41) “was not rigid and stereotypic; rather, it was flexible, fast, and adapted precisely to accomplish the task. The many components of the articulatory apparatus always cooperated in such a way as to preserve the speaker’s intent. The functional synergy, as it were, revealed”.

1.1.3 Coordenação motora na fala: produção e percepção

Kelso (1995:91) discorre em seguida sobre a questão de se encontrar variáveis coletivas em sistemas complexos como a fala, em que há um grande número de *graus de liberdade*, definidos como o conjunto de deslocamentos independentes possíveis responsáveis pela movimentação de um corpo ou sistema. Como vimos anteriormente, são necessários cerca de trinta músculos para se produzir uma simples sílaba como [ba]. Como os falantes humanos, a partir de um conjunto enorme de possibilidades, conseguiriam obter um padrão sonoro distintivo e bem formado? A questão crucial aqui é o da invariância perceptual, por um lado, e a variabilidade acústica e articulatória, por outro lado (vide Perkell & Klatt, 1986, para uma discussão abrangente sobre o tema). Apesar de tal invariância, todos sabemos que há enorme variação lingüística de acordo com diferentes dialetos, padrões acentuais, taxas de elocução, palavras/sons circundantes, etc. Sendo assim, Betty Tuller e Scott Kelso acreditam ser a interpretação da invariância lingüística, nos moldes da gramática gerativa, problemática para entender aspectos de produção e percepção da fala. Para esses autores, a invariância lingüística não ocorre somente na percepção da fala, mas também na produção da fala (vide Tuller & Kelso, 1984). Segundo Kelso (ibidem), “fluctuations are a conceptually

crucial part of the dynamics and an absolutely necessary part of the measurement process. Speech, as dynamic process, is no exception”. Um dos vários experimentos lingüísticos feitos pelos pesquisadores é explicitado abaixo.

Tuller & Kelso (1990) mostraram que a mudança de uma sílaba como “at” para “ta”, demonstrada por um experimento clássico de Stetson (1988), é explicada através de uma instabilidade, ou seja, uma *transição de fase*, definida como a mudança de fase (medida em graus) de um parâmetro do sistema. Nesse experimento foram medidas a abertura glotal, a pressão intra-oral, e a movimentação labial. A variável coletiva medida foi a fase relativa entre a abertura da glote e o fechamento dos lábios. Esta medida é bastante distinta para as sílabas [ip] versus [pi]. No entanto, em uma taxa crítica a relação de fase subjacente a [ip] mudou para [pi], a qual permaneceu estável para todas as taxas de elocução. Como podemos inferir, tal mudança de padrão silábico é resultado de uma instabilidade, e o sistema exibe histerese, propriedade dos sistemas dinâmicos em que, ao se mudar um parâmetro do sistema, uma variável do sistema pode permanecer em sua posição atual, ou voltar vagarosamente a um estado anterior. Por exemplo, ao esticarmos certos tipos de mola até um determinado limiar e depois soltarmos, elas podem não voltar mais ou demorar a voltar à sua posição inicial. No experimento mencionado, uma vez que um determinado atrator é atingido [pi], o sistema não volta mais à sua configuração anterior [ip].

Tuller & Kelso (ibidem) questionam que se a fase relativa entre os articuladores é uma variável coletiva importante para a comunicação humana, também deve haver conseqüências perceptuais de uma mudança de fase. Para testar esta hipótese, fizeram um experimento perceptivo com vários informantes, em que foi constatado que a percepção mudou de [ip] para [pi] exatamente no mesmo ponto em que a mudança de fase articulatória ocorreu. Dessa forma os autores concluíram que mudanças na fase dos articuladores resultam tanto em mudanças sonoras quanto semânticas. Kelso (1995:93) vai além e diz: “... using our phase transition methodology, such collective variables encompass production, acoustics, and perception, and offer new light on the time-honored invariance issue”.

Kelso (1995:26) resume a aplicação dos sistemas dinâmicos na natureza, comentada acima, da seguinte forma:

... the human brain is fundamentally a pattern-forming, self-organized system governed by nonlinear dynamical laws. Rather than compute, our brain “dwells” (at least for short times) in metastable states: it is poised on the brink of instability where it can switch flexibly and quickly. By living near criticality, the brain is able to anticipate the future, not simply react to the present. All this involves the new physics of self-organization in

which, incidentally, no single level is any more or less fundamental than any other.

O pensamento de Berthoz (Berthoz, 2000) se afina com a abordagem dinamicista de Kelso (*ibidem*), principalmente no que concerne à não dualidade entre o cérebro e o corpo. Para o pesquisador, o cérebro não é indissociado da cognição (como pensa, dentre outros, o filósofo e matemático Châtelet, que chegou a chamar os estudos cerebrais da cognição de “barbarismos mentais” (Châtelet, 1993)). Mais especificamente, para Berthoz, a cognição emerge da complexidade cerebral. Nesta perspectiva, o cérebro é usado para prever o futuro, antecipar as consequências de ações em geral e para poupar tempo. Segundo Berthoz (2000:2–3), “the most essential properties of human thought and sensibility are dynamic processes — ever changing, ever adapting relationships among the brain, the body, and the environment”.

Na próxima seção faremos uma abordagem lingüística mais abrangente de autores que aplicaram a teoria dos sistemas dinâmicos à linguagem.

1.2 Modelos dinâmicos aplicados aos estudos lingüísticos

Na seção anterior apresentamos alguns conceitos básicos da teoria dos sistemas dinâmicos (doravante TSD) e, ao final, demonstramos sua aplicação à Lingüística. Nesta seção ampliaremos as exemplificações da TSD à linguagem e também mostraremos as vantagens empíricas e teóricas de tal mudança de paradigma teórico. Para tanto, primeiramente, discutiremos aspectos lingüístico-teóricos da TSD e, posteriormente, apresentaremos uma proposta dinamicista de análise empírico-teórico da produção e percepção da fala de Kelso et al. (1986).

1.2.1 A gramática na TSD

A teoria lingüística tradicional (a partir de Chomsky, 1985), como vimos anteriormente, trata os processos lingüístico-cognitivos através da metáfora computacional, em que o complexo mente-cérebro funciona como um processador central de informações (um computador) que controla todo o nosso *desempenho lingüístico*. Nessa abordagem científica, nossa *competência lingüística* consiste em conhecimentos gramaticais inatos inerentes ao ser humano, os quais, ativados durante o período de aquisição da linguagem, geram todo nosso conhecimento lingüístico.

Em “Language as a Dynamical System”, Elman (1995) propõe que o processamento da língua natural é melhor explicado através de características peculiares aos sistemas dinâmicos não-lineares. O aspecto mais marcante desse artigo é a reconceitualização/reavaliação dinamicista de mecanismos lingüísticos como *competência* e *desempenho*, e símbolo lingüístico, reavaliados e reescritos através de conceitos da TSD.

Os trabalhos de Elman (Elman, 1995, 1991, dentre outros) revelam que a utilização de uma abordagem dinamicista à linguagem traz diversas vantagens, como: (i) podermos associar aspectos cognitivos do cérebro com outros órgãos corporais; (ii) respeitarmos a filogenia do sistema nervoso; e também por (iii) reconhecermos propriedades inerentes ao cérebro humano (vide Churchland & Sejnowski, 1992).

1.2.2 Análises lingüísticas à luz da TSD

Nesta subseção exemplificaremos alguns casos em que a TSD foi aplicada à linguagem (e cognição em geral), especificamente no campo da fala, ramo que obteve o maior número de adeptos por assimilar aspectos fonológicos e fonéticos de maneira bastante elegante, devido, sobretudo, à tese de comensurabilidade entre os níveis físicos e mentais da linguagem e à linguagem matemática utilizada na descrição de fenômenos lingüísticos. Como veremos

na seção 1.2.3, a noção de fase do gesto lingüístico da TSD, por exemplo, é capaz de explicar uma variada gama de fenômenos lingüísticos, como, por exemplo, a variação lexical entre [pas] e [pajs] de maneira parcimoniosa. Na próxima subseção faremos uma descrição detalhada de “The dynamical perspective on speech production: data and theory” de Kelso et al. (1986), o qual serviu de força-motriz a trabalhos teóricos posteriores/concomitantes, como a Fonologia Articulatória, a qual será retratada na seção 1.3 desta tese.

Em “Stress and Rate: Differential Transformations of Rate”, Tuller et al. (1982) trabalham a questão controversa de que variações em acentuação silábica e taxa de elocução resultam em transformações idênticas de atividade motora (vide Lindblom, 1963). Nesse artigo, os pesquisadores consideravam que cada fonema possuía um programa “invariante” imune aos efeitos de mudanças acentuais no léxico e à variação da TE. Além do mais, tal teoria previa que variações em acento e taxa afetariam o sinal acústico de forma idêntica, mas que não causariam efeito nos sinais fisiológicos. No entanto, estudos acústicos subseqüentes (Verbrugge & Shankweiler, 1977; Harris, 1978; Gay, 1978, dentre outros) e articulatórios (Gay & Hirose, 1973; Gay and Ushijima, 1975, Harris, 1978; Sussman & MacNeilage, 1978, dentre outros) mostraram que a variação do acento lexical e da TE geram efeitos acústicos e articulatórios distintos.

O objetivo dos experimentos de Tuller et al. era de se variar ortogo-

nalmente acento e TE e verificar se haveria efeitos acústico-articulatórios distintos nas enunciações. À época só havia um trabalho que havia feito procedimento semelhante, o de Harris (1978).

No primeiro experimento foram utilizados dois adultos do sexo feminino, falantes nativos de inglês americano. O corpus consistia de logatomas de quatro sílabas da forma /əpipipə/, /əpipibə/, /əpepepə/ ou /əpepebə/, com acento na primeira ou segunda sílaba medial. Os informantes foram instruídos a ler tais logatomas em duas taxas distintas: uma lenta (conversacional) e outra rápida. Os dados eletromiográficos foram obtidos dos músculos orbicularis oris, principais responsáveis pela produção das consoantes bilabiais, e genioglossus, responsável pela produção do [i]. Os dados acústicos foram gravados simultaneamente em uma fita FM multi-canal (vide Tuller et al., 1982, pp. 1535-6 para detalhes experimentais).

Os resultados acústicos e articulatórios para o primeiro experimento mostraram que variações no acento e TE produziram mudanças de duração acústica através de diferentes efeitos na atividade dos músculos genioglossus e orbicularis oris. No entanto, os autores acreditam serem tais resultados inconsistentes por lidarem com um conjunto limitado de músculos e contextos fonéticos, e por utilizarem informantes linguisticamente sofisticados. Para suprir essa deficiência, os pesquisadores realizaram um segundo experimento.

Nesse segundo experimento foram utilizados cinco informantes do sexo fe-

minino, quatro americanos e um neo-zeolandês. Destes, quatro desconheciam os objetivos do experimento. O corpus consistia de quatro logatomas de duas sílabas, a saber, /pipip, papap, pipap, papip/. Acento e TE foram variados exatamente como no primeiro experimento. Além dos músculos utilizados anteriormente, foram gravados os músculos pterigóide lateral, responsável pelo levantamento da mandíbula, o ventre anterior do digástrico e o ventre inferior do pterigóide lateral, responsáveis pelo abaixamento da mandíbula.

Novamente, foram obtidos resultados articulatórios distintos para acento e TE. Resumidamente, as variações de acento foram bastante consistentes entre músculos relacionados às vogais, mas as variações de TE não foram suficientemente robustas para serem significativamente diferentes através de análises de variância (ANOVA).

A conclusão teórica oriunda desse artigo é o fato de apesar da organização temporal (timing) da atividade intersegmental variar, a mudança temporal não é necessariamente concomitante à variação suprasegmental.

Em “The Timing of Articulatory Gestures: Evidences for Relational Invariance”, descrito a seguir, Tuller & Kelso (1984) mostram que a organização temporal intersegmental relativa é mantida ao se variar a estrutura suprasegmental. Nesse artigo é discutida uma questão central na pesquisa lingüística, a de achar aspectos invariantes na produção de uma unidade lingüística, seja ela traços distintivos, fonemas ou gestos articulatórios. Em

suma, encontrar invariâncias perceptivas que sejam independentes de variações supra-segmentais.

A procura por invariantes relacionais é motivada por três fatores distintos:

1) a organização temporal relativa de eventos articulatórios em tarefas motoras como coordenação manual, escrita à mão, digitação, controle postural e locomoção é mantida independentemente de mudanças nas taxas de produção e força (cf. Kelso et al., 1983); 2) o falante se baseia em durações acústicas relativas e não absolutas na percepção de distinções lingüísticas; o *timing* relativo da cinemática articulatória nos níveis silábico e segmental não é influenciado por variações suprasegmentais (cf. Kozhevnikov & Chistovich, 1965; Kent & Netsell, 1971; Kent & Moll, 1975; Löfqvist & Yoshioka, 1981).

Evidências para tal invariância relacional já foram encontradas em Tuller et al. (1982), descrito anteriormente, em que os resultados comprovam que a duração média do intervalo entre o início da atividade muscular para vogais sucessivas foi sistematicamente relacionado à latência média da atividade do músculo relacionado à produção da consoante média. Sendo assim, o experimento de Tuller & Kelso (1984) relatado a seguir serviu para ampliar/aperfeiçoar os resultados prévios encontrados.

Três informantes do sexo feminino e um do sexo masculino, falantes nativos de inglês, foram utilizados no experimento. Os enunciados possuíam a seguinte estrutura: /b/VCV/b/. A primeira vogal foi /a/ ou /æ/, a se-

gunda vogal /a/ e a consoante medial foi /b/, /p/, /w/ ou /v/. O acento foi colocado na primeira ou na segunda sílaba. Os informantes leram listas de palavras quasialeatórias em duas taxas de elocução — uma conversacional e outra mais rápida — inseridas na fraseveículo “It’s a _____ again”.

Os dados articulatórios foram colhidos através de diodos infravermelhos emissores de luz (LEDs) colocados nos lábios inferiores e superiores, mandíbula e nariz. Os dados acústicos foram colhidos simultaneamente em uma fita FM.

Os resultados gerais desse experimento mostraram que o começo do gesto consonantal inter-vocálico é fortemente relacionado ao período das vogais circundantes ao gesto consonantal. Além disso, tal invariância foi independente de variações em TE e acento, as quais, freqüentemente, geraram grandes variações na duração, deslocamento e velocidade dos articuladores individuais. Tais resultados corroboram os dados de Tuller et al. (1982).

Apesar de encontrarem invariância relacional em seus dados, Tuller & Kelso (1984) tiveram a preocupação de apontar algumas deficiências em sua abordagem, a saber: 1) a medida do início do movimento articulatório não foi necessariamente isomórfica à medida articulatória eletromiográfica do experimento anterior (Tuller et al., 1982); 2) o início do movimento gestual não implica, necessariamente, que este tenha um status privilegiado como variável de controle; 3) a movimentação da língua — o mais importante arti-

culador — não foi analisada; 4) o material de fala foi muito simples — apenas consoantes simples entre duas vogais plenas.

Apesar desses resultados serem ainda preliminares, Tuller & Kelso informam que eles se afinam com outras formas de atividade coordenada (vide Boylls, 1975; Fowler et al., 1980, Grillner, 1982; Kelso et al., 1983; Kelso et al., 1984; Turvey et al., 1978) e sugerem que a estabilidade temporal dos movimentos relativos a um ciclo e a independência da organização temporal relativa do movimento é uma das principais características do complexo músculo-articulação ao cooperarem para realização de uma determinada tarefa.

Em um outro artigo, Nitttrouer et al. (1988) aperfeiçoam os resultados sobre a invariância relativa utilizando a metodologia inaugurada em Kelso et al. (1986)¹ de se utilizar um plano de fase (um gráfico de deslocamento em função da velocidade) para análise do ciclo vocálico.

Dois informantes do sexo masculino e dois do feminino, falantes nativos de inglês americano, foram utilizados no experimento. O material de fala consistiu dos seguintes enunciados: (1) /pa#map/, (2) /pam#ap/, (3) /pa#pap/, /pap#ap/. Os estímulos foram aleatorizados com variação de sílaba acentuada, estrutura silábica e identidade da consoante inter-vocálica. Os informantes foram instruídos a produzir os enunciados nas taxas normal

¹A mesma será detalhada na próxima sub-seção desta tese.

e rápida na frase-veículo “It’s a _____ again”.

Como nos experimentos anteriores, o movimento articulatório foi obtido através de LEDs posicionados na fronteira vermillion dos lábios inferiores e superiores, na mandíbula e nariz. Os dados acústicos foram gravados simultaneamente em uma fita FM.

Os resultados encontrados nesse experimento, devido à sua forma lingüística mais complexa, revelaram, contrariamente ao reportado nos artigos citados anteriormente, que as relações de fase entre os gestos articulatórios usados na produção de segmentos adjacentes variaram sistematicamente em função da estrutura lingüística e não-lingüística, como TE, padrão acentual e identidade consonantal. No entanto, os autores afirmam que

this suggestion is not meant to refute the idea that phonological structure within a word may be instantiated by interarticulator phase relations. Instead, what the present data suggest is that the spatiotemporal relations among articulatory gestures vary systematically with the structure of the utterance at all levels, not just the phonological level (Nitttrouer et al., 1988, pp. 1659-60).

1.2.3 A abordagem dinamicista na produção da fala: dados e teoria

Kelso et al. (1986) apresentam um programa de pesquisa baseado na TSD para compreender regularidades na produção de padrões articulatórios na fala, aprofundando os estudos anteriores apresentados na última seção. A questão principal aqui é o fato de percebermos enunciados da mesma forma, mesmo que sejam variados aspectos como TE, ênfase, palavras circundantes. Por exemplo, sempre percebemos /'fɔSforo/ como uma palavra invariável, mesmo sendo pronunciada em extremos de TE — muito lenta e muito rápida. Portanto, os estudos articulatórios devem levar em conta tanto aspectos variantes quanto invariantes da produção cinemática.

Kelso, Saltzman & Tuller, juntamente a colegas do Haskins Laboratories, (vide Browman, Goldstein, Kelso, Rubin & Saltzman, 1984) se propuseram a responder tal questão. Para os pesquisadores as unidades lingüísticas são melhor explicadas como o produto de estruturas de controle invariantes no tempo para um sistema de articuladores do aparelho fonador. Uma noção importante nessa abordagem é a de estruturas coordenadas.

Na seção 2 de Kelso et al. (1986), os pesquisadores detalham o funcionamento das estruturas coordenadas e apresentam evidências empíricas para as mesmas, através de perturbação dos articuladores da fala. O mo-

delo matemático-teórico empregado em tal descrição é o do modelo dinâmico da tarefa (cf. Saltzman & Kelso, 1987; Saltzman & Munhall, 1989, vide Hawkins, 1992, para uma introdução ao modelo).

Na seção 3, Kelso et al. introduzem o conceito de fase como uma variável dependente chave para a compreensão dos aspectos invariantes na fala. Além disso, é mostrado que as estruturas coordenadas são definidas dinamicamente e unificam ambos os níveis articulatórios abstratos e de produção de fala.

Por fim, na seção final, os pesquisadores debatem os aspectos teóricos e empíricos da utilização da abordagem dinamicista na compreensão das unidades fonológicas da linguagem.

Estruturas coordenadas na fala

A questão central envolvida na coordenação gestual é entender como um número enorme de componentes neuro-musculares se combinam para formar o som — um padrão distintivo bem formado. Tal questão é um problema básico no controle de sistemas complexos, ou seja, descobrir quais parâmetros de controle podem manipular a quantidade enorme de graus de liberdade.

Fowler (1977, 1980) apresentou evidências especulativas para resolver esse problema. Segundo a lingüista, os componentes neuro-musculares cooperam para formar unidades de ação funcionais (ou estruturas coordenadas). Kelso et al. (1986), por sua vez, apresentam evidências empíricas para tais estru-

turas coordenadas.

Estudos anteriores (Kelso, Tuller & Fowler, 1982; Kelso, Tuller, Vatikiotis-Bateson & Fowler, 1984) já apontaram evidências para as estruturas coordenadas. Nesses trabalhos, os articuladores da fala (lábios, língua, mandíbula), ao serem perturbados, compensaram imediatamente para a produção da tarefa articulatória, conforme apresentamos na seção 1.1.2. Outros estudos (Munhall & Kelso, 1985; Abbs, Gracco & Cole, 1984) revelaram que a noção de fase, a qual será explicada abaixo, é importante para compreendermos a produção articulatória da tarefa em situações em que o sistema é perturbado (vide Kelso et al., 1986, pp. 33-39; Saltzman, 1987, para uma descrição da modelagem dos gestos articulatórios à luz do modelo dinâmico da tarefa. Aspectos de modelamento gestual estão fora do escopo dessa tese).

Orquestração gestual: a noção de fase

Como vimos anteriormente, Kelso e colegas em uma série de estudos anteriores encontraram evidências para estruturas articulatórias estáveis na fala (Tuller, Harris & Kelso, 1982; Tuller & Kelso, 1984; Harris, Tuller & Kelso, 1986), as quais partilham características em comum com outros estudos motores (Grillner, 1977; Nashner, 1977; Lestienne, 1979; Viviani & Terzuolo, 1980; Hollerbach, 1981, dentre outros). A questão empírica a ser colocada nesse panorama é a de como tal estabilidade articulatória deve ser melhor

racionalizada. Kelso et al. respondem a essa questão afirmando que os movimentos articulatorios devem ser reanalisados como trajetórias em um plano de fase.

No plano de fase temos a representação da posição dos articuladores (eixo x) junto com a sua velocidade (eixo y). Temos na figura 1.1 um exemplo de reescrita da movimentação cinemática da mandíbula em um plano de fase.

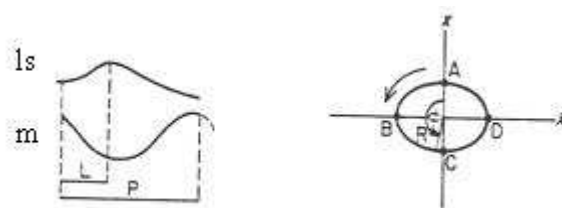


Figura 1.1: Representação de um dado articutório abstrato. À esquerda temos a movimentação cinemática dos lábios superiores (ls) e da mandíbula (m). L representa a latência para o começo do gesto vocálico [a] e P o período do ciclo consonantal. À direita temos a representação do mesmo gesto em um plano de fase. O intervalo A-B representa o movimento do final do gesto [b] até sua velocidade máxima negativa (v_{mn}). O intervalo B-C representa o movimento da v_{mn} até o início do próximo [b]. O intervalo C-D representa o movimento do início do segundo [b] até a sua velocidade máxima positiva (v_{mp}). O intervalo D-C representa o movimento da v_{mp} até o final do gesto do segundo [b]. ϕ representa o ângulo de fase dos lábios superiores em relação ao ciclo consonantal.

A utilização das trajetórias em um plano de fase, segundo Kelso et al., preserva, por exemplo, importantes diferenças entre sílabas acentuadas e não acentuadas. Na primeira o abaixamento máximo da mandíbula é maior do que na segunda e a velocidade máxima dos articuladores difere substancialmente entre elas (cf. MacNeilage et al., 1970; Stone, 1981; Tuller, Harris &

Kelso, 1982). Por outro lado, as “durações” em função do acento não são expressas explicitamente nesse sistema, embora possam ser inferidas deste. Por esta razão não utilizamos tal medida nas análises articulatórias de nossa tese.

As vantagens empíricas da análise em ângulo de fase são: (i) análises mais consistentes em função de condições específicas; (ii) a utilização de apenas um parâmetro em suas descrições. Já as vantagens teóricas são: (i) a normalização da duração em função do falante, acento, TE; (ii) o suporte às teorias de organização temporal intrínsecas (cf. Fowler, 1980; Fowler et al., 1980); (iii) a sobreposição dos gestos é observada de maneira bastante clara; (iv) “segmentos” no sentido lingüístico tradicional podem não ser unidades relevantes na produção da fala; (v) minimização da questão do problema da relação entre mente e corpo.

Transição de fase

A forma mais comum de descoberta de padrões na TSD é a utilização de um elemento perturbador do sistema para observar a emergência de regularidades. No caso da fala, o perturbador mais comum é a variação da TE, que segundo Stetson (1988) é uma das maiores causas de modificação fonética.

Os experimentos de Stetson (1988), apesar de não estarem inseridos na TSD, a qual veio a se consolidar muitos anos mais tarde, possuem elementos

facilmente relacionados à abordagem dinamicista. Em um de seus experimentos, o pesquisador relata, como vimos brevemente na seção 1.1.3., que uma sílaba como [ip] produzida com o aumento gradativo da TE muda para [pi] ao atingir um determinado limiar. O mesmo, porém, não acontece com a sílaba [pi], a qual se mantém inalterada com o aumento da TE. Tal resultado, dentre outros de Stetson, sugere a mudança de transição de fase em sistemas longe do equilíbrio, em que a TE seria um parâmetro de controle e o ângulo de fase o parâmetro de ordem. Para investigar tal especulação teórica, Kelso et al. elaboraram o experimento descrito a seguir.

Nesse experimento três informantes produziram a sílaba [pi] ou [ip] com aumento gradativo da TE, da lenta à rápida. LEDs foram posicionados nos lábios e mandíbula dos mesmos e a abertura do fechamento da glote foi monitorada por transiluminação (cf. Baer, Löfqvist & McGarr, 1983). Medidas prévias haviam mostrado que as relações de fase entre o início da movimentação dos lábios e a abertura do pico glotal para [pi] e [ip] eram estatisticamente distintas.

Os resultados mostraram que a relação de fase para [ip] se manteve estável até o parâmetro de ordem chegar a 4 sílabas/segundo, em que a relação de fase mudou para a averigüada anteriormente para [pi]. Tal transição de [ip] a [pi] além de uma taxa crítica reflete uma bifurcação no sistema. Além disso, o sistema exibiu histerese, pois, ao mudar a relação de fase, o

sistema não volta mais à configuração original [ip]. Novamente, como no experimento de Stetson, um fenômeno distinto foi observado para [pi], em que a relação de fase se manteve constante ao longo de todo o experimento. Tal resultado em termos de descrição lingüística é bastante significativo, pois mostra de maneira simples ser a forma CV um padrão mais estável do que VC, e conforme sabemos, CV é o padrão mais recorrente nas línguas do mundo e é o padrão utilizado pelas crianças na produção das primeiras sílabas.

Implicações teóricas

Kelso et al. sugerem que movimentos articulatorios tão díspares quanto movimentação das mãos, fala, e locomoção de quadrúpedes podem ser analisados dentro de um mesmo programa de pesquisa — o da TSD. Em todos esses sistemas ocorrem instabilidades quando os mesmos são afastados do equilíbrio, as quais podem ser elegantemente explicadas através de transições de fase. Além do mais, em sistemas complexos como a fala, flutuações no sistema nos permitem observar novos modos estáveis de acoplamento gestual.

Conforme afirmam Kelso et al (1986:54), “not only does the present theoretical perspective afford the potential of a principled analysis of pattern formation (for many more details see Haken et al., 1985) but, as we mentioned earlier, the nature of the pattern change itself may prove rather informative”. Além disso, “finding the articulatory system to perform at unusual rates may

reveal the primitive units of language and, more important, their interactions with other units” (p. 55). Por fim, “the uniqueness of the present scheme is that stability and change of speech action patterns are seen as different manifestations of the same underlying principles — the phenomenon observed depends on which region of the parameter space the system occupies” (p. 56).

Na próxima seção faremos uma descrição detalhada da Fonologia Articulatória, uma teoria fonológica que engloba elegantemente tanto aspectos mentais quanto físicos da produção lingüística. Além disso, tal teoria exemplifica uma das áreas em que a TSD foi mais bem sucedida na linguagem — a do controle motor da fala.

1.3 A Fonologia Articulatória

Catherine Browman e Louis Goldstein do Haskins Laboratories, baseados em trabalhos anteriores de controle motor na fala, conforme vimos na seção 1.2, foram pioneiros na aplicação da TSD à fonologia (cf. Browman & Goldstein, 1986, 1989, 1990b, 1992). O programa de pesquisa inaugurado por esses dois pesquisadores foi chamado de Fonologia Articulatória (doravante FAR). Na apresentação da FAR à comunidade científica (Browman & Goldstein, 1986) são mostradas as motivações teóricas para a elaboração de uma nova

abordagem fonológica para a linguagem.

1.3.1 Motivação teórica

Na teoria gerativa (a partir de Chomsky, 1985), aspectos mentais (considerados propriamente lingüísticos) e físicos da linguagem são incongruentes. Existe, assim, nas teorias gerativas (e também nas pós-gerativas) um vácuo a ser preenchido entre esses dois níveis da linguagem.

Apesar das inovações teóricas das fonologias não-lineares (Lieberman & Prince, 1977; Nespor & Vogel, 1986; Clements, 1985; Goldsmith, 1990) e dos trabalhos fonéticos que propuseram regras implementacionais *ad hoc* (Port, 1981; Port & O'Dell, 1984; Keating, 1985) resolverem vários problemas precedentes, os mesmos não foram capazes de fazer uma ponte formalmente elaborada entre a estrutura fonológica e a estrutura física da linguagem. Esta foi a principal inovação da FAR que, através da noção de gesto lingüístico, foi capaz de descrever teórico-abstratamente características espaço-temporais da linguagem. Assumir os gestos articulatórios como unidades fonológicas, segundo Browman & Goldstein (1989), tem ainda uma vantagem adicional: o estágio pré-lingüístico da aquisição da linguagem pode ser descrito utilizando-se as mesmas estruturas do sistema fonológico já estabilizado. Além disso, mostra evidências neuro-fisiológicas para o desenvolvimento fonológico da

criança.

Alguns anos mais tarde, após aperfeiçoarem aspectos teóricos e empíricos da FAR, Browman & Goldstein (1992) apresentam um panorama das principais características dessa teoria fonológica, relatado a seguir.

1.3.2 Gestos como estruturas articulatórias dinâmicas

De acordo com Browman & Goldstein (1992), gestos são caracterizações de eventos discretos, fisicamente reais, revelados durante o processo de produção da fala (p.156). Além disso, são caracterizações abstratas de eventos articulatórios com durações intrínsecas específicas (p.155). São também considerados unidades primitivas de contraste lexical. Tendo em vista que o gesto é definido tanto abstrata quanto fisicamente, ele fornece subsídios para a solução do problema da ponte entre estruturas fonológicas e físicas, as quais são formalizadas abstratamente através do modelo dinâmico da tarefa (Saltzman & Kelso, 1987; Saltzman & Munhall, 1989).

1.3.3 O modelo dinâmico da tarefa

O modelo dinâmico da tarefa (doravante MDT), baseado nas estruturas coordenadas vistas na seção 1.2.3, define os gestos discretos, abstratos, os quais geram trajetórias articulatórias no espaço e no tempo sem qualquer interpre-

tação adicional ou regras *ad hoc*.

Um aspecto importante do MDT é que o movimento das variáveis do trato vocal (doravante VT) é caracterizado abstratamente, não o movimento dos articuladores propriamente ditos (Browman & Goldstein, 1992, p.156). Uma VT é definida por Browman & Goldstein como uma dimensão de uma constrição no aparelho fonador, em que os articuladores necessários à formação e soltura desta constrição são organizados em uma estrutura coordenada (cf. Turvey, 1977; Fowler et al., 1980). Por exemplo, a VT local de constrição da ponta da língua é causada pela ação de dois articuladores: corpo de língua e mandíbula. Na figura 1.2 temos as VT e seus respectivos articuladores. O gesto na FAR é especificado através de um conjunto de VT. Por exemplo, local e grau de constrição da ponta da língua são duas dimensões da mesma constrição - ponta de língua. Dessa forma, temos cinco conjuntos de VT, a saber, lábios, ponta de língua, corpo de língua, véu palatino e glote.

O MDT têm a função de especificar o modo pelo qual um sistema dinâmico subjacente invariável gera uma variabilidade contextualmente condicionada na superfície e regida por regras na observação da movimentação de membros e articuladores na fala. Nessa modelação, dois níveis funcionais distintos mas interativos são necessários: (i) o nível de constrição, o qual é descrito tanto pelas coordenadas dos articuladores do modelo (ex. ponta e dorso de língua), bem como pelas coordenadas das variáveis do aparelho

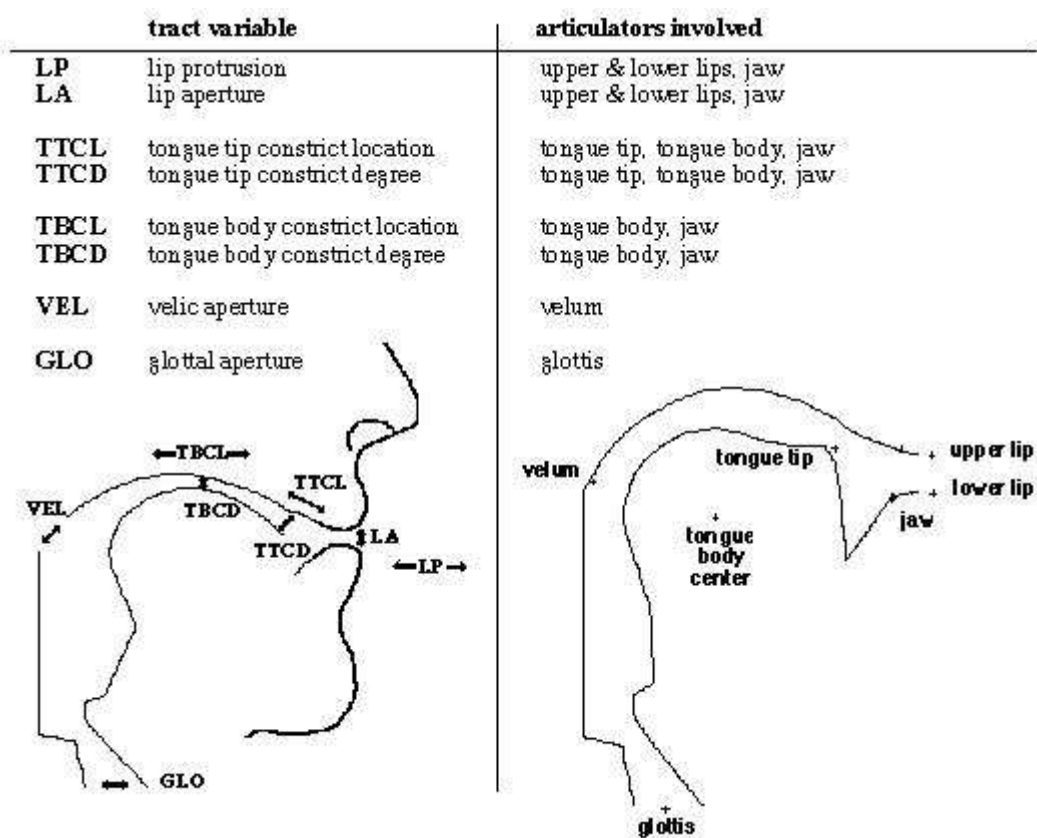


Figura 1.2: Variáveis de trato e seus articuladores (Browman & Goldstein, 1992, p. 157).

fonador (ex.: abertura labial); (ii) o nível de ativação, o qual é definido de acordo com um conjunto de coordenadas de ativação gestual. Este nível governa a evolução temporal das trajetórias de ativação dos gestos, especificados através de pautas gestuais, nos enunciados, bem como os padrões de relações de fase entre os gestos que participam do enunciado.

No nível de constrição, de acordo com Byrd & Saltzman (2003:153), os gestos são caracterizados qualitativamente como detentores de uma dinâmica de “atratores pontuais”. Os modelos de “atratores pontuais”, segundo tais pesquisadores, possuem propriedades importantes, as quais parecem ser características do comportamento na fala, a saber: (i) eles atingem seu alvo (ou “posição de equilíbrio”), mesmo quando são inesperadamente perturbados e independentemente de sua posição inicial; (ii) os atratores (ou tarefas fonológicas) em atrito uns com os outros se auto-organizam para realização de suas respectivas tarefas.

O atrator pontual utilizado no MDT foi o sistema massa-mola amortecido, o qual simula a formação e soltura de constrições no aparelho fonador. Na figura 1.3 temos a equação do movimento que descreve a dinâmica dos sistemas massa-mola amortecidos. Tais parâmetros serão necessários quando descrevermos os experimentos articulatórios desta tese. Nessa fórmula: x_o , representa a posição da VT na posição de equilíbrio ou descanso da mola; m , representa a massa, que é tida, nesse momento inicial com valor 1 para

todos os gestos; k , representa a rigidez da mola, a qual representa a elasticidade do sistema e é proporcional à “velocidade” gestual intrínseca. Seu valor é inverso à duração do gesto, ou seja, uma rigidez baixa indica gestos com grande duração (ex.: vogais), enquanto uma rigidez alta indica gestos com pequena duração (ex.: consoantes); b , representa o amortecimento do sistema e é considerado “crítico”, ou seja, o gesto alcança seu alvo de maneira assintótica.

$$\begin{array}{c}
 \text{distância do alvo} \\
 \text{(repouso)} \\
 \overbrace{m\ddot{x} + b\dot{x} + k(x-x_0)}^{\text{inércia} \quad \text{amortecimento} \quad \text{força elástica da mola}} = 0
 \end{array}$$

Figura 1.3: Equação de movimento de um sistema massa-mola amortecido (adaptado de Byrd & Saltzman, 2003, p. 154). Vide texto para maiores detalhes.

Os atratores pontuais têm sido usados para descrever sistemas físicos (Shadmehr, 1995; Mussa-Ivaldi, 1995; Flash & Sejnowski, 2001) e também, de acordo com Plaut (1995), nos sistemas cognitivos e comportamentais. Assim, Byrd & Saltzman (2003:154) afirmam que “the mass-spring model is used not as a model of physical system but of the cognitive control of abstract linguistic units”. Dessa forma, os atratores de ponto gestuais são capazes

de iluminar o caminho na transição da dinâmica subjacente invariante às variações contextuais no desempenho lingüístico. Por exemplo, estipulando o valor $x_o = 0$ para o atrator pontual labial, na produção de um [b], resultará na movimentação até que o alvo de abertura labial atinja o valor 0, ou até que eles se reorganizem na competição com outros gestos concomitantes, como na produção de [ba, bi, bu], os quais resultam em alvos distintos para [b], devido à sobreposição dos gestos consonantais com os vocálicos. Por sua vez, a sobreposição temporal dos gestos são determinadas pelas relações de fase nas trajetórias de ativação dos gestos, modeladas pelo nível de ativação.

As trajetórias de ativação no MDT são pré-planejadas segundo regras fonológicas e acopladas unidirecionalmente no nível de constrição. De acordo com Byrd & Saltzman (2003:155), “the dynamical system for activation can be conceptualized in terms of a “clock” whose time flow is mapped into a pattern of sequentially ordered and temporally overlapping gestural activations and which is sensitive to the ongoing state of the constriction-level variables”. Em outras palavras, a taxa do relógio interno (*clock rate*) é continuamente modificada/atualizada pela evolução dos gestos ao longo do enunciado (Saltzman, Löfqvist, Kay, Kinsella-Shaw & Rubin, 1998). Em outros trabalhos também verificou-se que a taxa do relógio articulatorio é modificada de acordo com as estruturas silábicas e segmentais do enunciado (Bailly, Laboissière & Schwartz, 1991; Barbosa & Bailly, 1994; Port & Cummins, 1992;

dentre outros). Em Byrd & Saltzman (2003), a hipótese adotada é a de que a taxa do relógio interno é modulada pela estrutura frasal do enunciado devido à ação de gestos prosódicos (gestos- π) (vide seção 1.4.2) .

1.3.4 Constelações gestuais

Os gestos são organizados entre si através de coordenações temporais, chamadas de constelações gestuais, para formar as estruturas conhecidas tradicionalmente como segmentos, sílabas, palavras e enunciados maiores. Na figura 1.4, temos alguns exemplos de combinações gestuais no PB.

Como podemos ver na figura 1.4, os itens lexicais são contrastados, basicamente, pela presença ou ausência de um determinado gesto (ex., “há” vs. “dá”, fig. 1.4a,b; “dá” vs. “pá”, fig. 1.4b,c; “si” vs. “sim”, fig. 1.4d,e). Outro aspecto das pautas gestuais é o fato de o vozeamento ser considerado valor *default* para todos os gestos², como notamos na representação da consoante vozeada [d] em 5b. Notamos também que “dá” difere de “há” pela presença de um “segmento”, enquanto “si” e “sim” diferem por um único “traço”, nasalidade. Contudo, tanto as diferenças lexicais, baseadas tanto em “segmentos” quanto em “traços distintivos” são feitas na FAR pela presença ou ausência de um único gesto. Vemos, pois, uma descrição parcimoniosa dos dados

²Lembremos que em Chomsky & Halle (1991) somente as soantes (vogais, semivogais, nasais e líquidas) possuíam vozeamento como valor *default*.

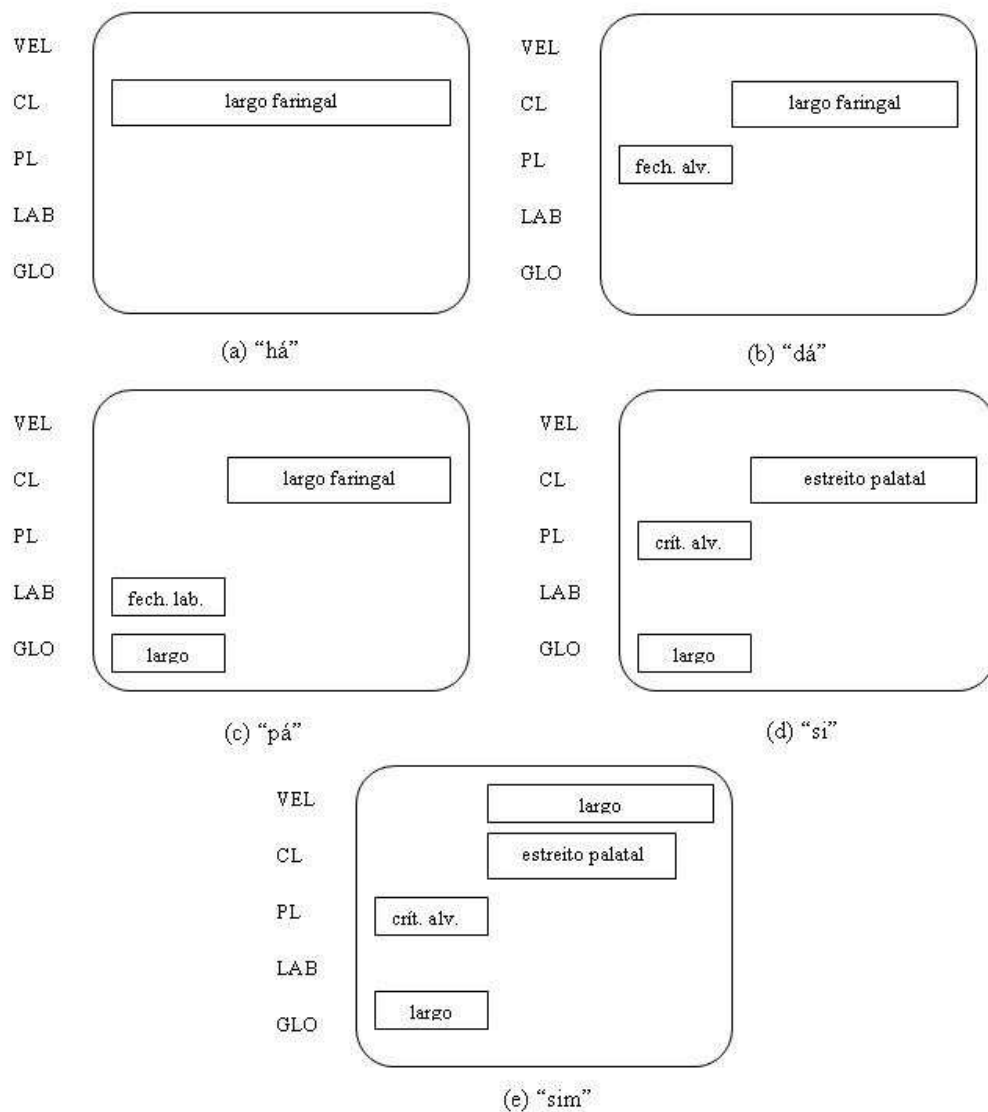


Figura 1.4: Pautas gestuais esquemáticas para o PB. VEL representa o véu palatino, CL, o corpo de língua, PL, a ponta de língua, LAB, os lábios e GLO a glote.

lingüísticos, pois com um único elemento lingüístico, o gesto, somos capazes de distinguir uma palavra da outra. Por outro lado, lembramos que nas teorias lingüísticas tradicionais, a distinção entre “pá” e “dá”, por exemplo, é sobrecarregada informacionalmente, pois implica na distinção de um segmento por outro, os quais, por sua vez, são formados por matrizes de traços distintivos. No entanto, as teorias não-lineares, como a geometria de traços, apesar de ainda partilharem dos pressupostos teóricos da gramática gerativa, aperfeiçoaram as análises fonológicas pós-SPE.

Clements & Hume (1995), criticando a análise fonológica oriunda de Chomsky & Halle (1991), em que os fonemas têm uma relação bijetora (de um para um) com os traços distintivos e os traços não têm estrutura interna, propõe uma nova abordagem empírico-teórica, a saber: (i) os valores dos traços são arranjados em camadas distintas, nas quais podem assumir relações não-lineares (não-bijetoras) entre si; (ii) os traços são organizados em arranjos hierárquicos, nos quais cada constituinte pode funcionar como uma unidade nas regras fonológicas (Clements & Hume, 1995:251). Essa nova abordagem foi chamada de Geometria de Traços. Apesar do avanço em relação às teorias fonológicas lineares e da tentativa de uma ponte entre a estrutura fonológica e a interpretação fonética (Clements & Hume, 1995:301), a geometria de traços não lida com dois aspectos inerentes na FAR: (i) a representação fonológica da duração; (ii) a importância dos valores contínuos

na explicação de fenômenos lingüísticos. Ambos aspectos estão presentes, em parte, nas pautas gestuais da figura 1.4.

Nas pautas gestuais esquemáticas da figura 1.4 ainda observamos alguns descritores dentro das caixas representantes dos gestos articulatórios, os quais não foram especificados na figura 1.2. Na FAR são utilizados cinco descritores para os graus de constrição: fechado (para oclusivas), crítico (para fricativas), e estreito, médio e largo, que, combinados com os descritores de local de constrição, são utilizados na descrição de vogais. Já quanto ao local de constrição são utilizados descritores distintos para lábios (protuso, labial ou dental), ponta de língua (labial, dental, alveolar, pós-alveolar ou palatal) e corpo de língua (palatal, velar ou uvular). Há ainda previsões não-implementadas no modelo para um articulador raiz de língua e para um descritor forma de constrição, que poderiam distinguir, por exemplo, as consoantes laterais entre si e das demais consoantes, como o [t].

1.3.5 Fala informal na FAR

Browman & Goldstein (1992) apresentaram estudos em que variações articulatórias, influenciadas, por ex., pela TE, não são percebidas pelos falantes e, portanto, não se enquadram nos casos tradicionais de alternância lexical. Há casos, porém, em que tais variações, provenientes de sobreposições gestuais,

resultam na percepção de diferentes formas lexicais. Este é o caso chamado por Browman & Goldstein de ocultamento gestual (gestural “hiding”) e na teoria lingüística tradicional, dentre outros, como “apagamento” (vide Guy, 1980), em que um segmento é tido como apagado. Um exemplo disso é a seqüência “perfect memory”, em que o [t] não é tido como pronunciado. Outro caso é o de “assimilação”, em que um segmento supostamente assimila traços de segmentos adjacentes. Um exemplo disso é o enunciado “seven plus”, em que o [n] é pronunciado como [m]. No entanto, estudos articulatórios como os de Browman & Goldstein (1990b) mostram que o movimento articulatório para /t/ no primeiro caso e a oclusão alveolar para [n] no segundo caso estão presentes na fala informal. Os falantes não percebem tais segmentos, pois variações gradientes de sobreposição gestual, ao atingirem um determinado limiar de sobreposição, fazem com que não mais percebam tais segmentos em sua forma plena.

O último processo a ser mencionado aqui é o de reestruturação silábica,³ fenômeno encontrado, por exemplo, no dialeto mineiro, principalmente, em que palavras como “abóbora, fósforo, xícara” são percebidas como “abobra, fosfro, xicra”. Processo semelhante a esse do PB ocorre no inglês na palavra “beret” que é pronunciada como [brej]. Estudos fonológicos tradicionais

³Na parte 3 desta tese abordaremos análises à la FAR de tais processos lingüísticos em PB.

explicam tal reestruturação silábica pelo apagamento da vogal da primeira sílaba “be”. Entretanto, Browman & Goldstein (1990b) mostraram que tal mudança silábica decorre da sobreposição do gesto labial de /b/ com um gesto de língua de /r/, ou seja, não houve apagamento da vogal entre tais consoantes, simplesmente uma aproximação maior entre as duas consoantes mencionadas, resultando, como consequência, numa mudança lexical ao nível perceptual. Baseado nesses fatos, Browman & Goldstein (1992) sugerem que casos de “apagamento” devem ser repensados como um fato articulatório mais consistente, ou seja, uma sobreposição maior dos gestos articulatórios na fala fluente. Os pesquisadores apontam duas razões principais para isso: (i) o “apagamento” é parte de um mesmo processo articulatório geral — o de sobreposição gestual; (ii) os constituintes fonéticos das palavras não mudam na fala fluente, e sim o grau de sobreposição gestual entre eles, que pode levar a mudanças perceptivas das palavras.

Browman & Goldstein (1992:172) resumem toda a argumentação teórico-descritiva aqui apresentada, a qual será abordada ao longo desta tese:

...factors associated with increased fluency (e.g. increased rate, more informal style) result in increasing the temporal overlap among gestures. Additionally, prosodic boundaries may influence the degree of overlap between neighboring gestures that be-

long to successive words. We have hypothesized that this kind of variation can result in changes that have traditionally been described as “fast speech” alternations of various sorts, and have presented articulatory evidence for this (Browman & Goldstein, 1990a, 1990b).

1.4 Modelos dinâmicos prosódicos

Apesar de os autores da FAR tentarem convencer a comunidade científica internacional de que uma fonologia baseada nos sistemas dinâmicos é capaz de explicar melhor os processos fônicos das línguas naturais, percebemos nos artigos de Browman & Goldstein uma grande dificuldade em romper com o paradigma lingüístico tradicional, baseado nas fonologias derivadas de Chomsky & Halle (1991). Em todos esses textos parece haver uma certa hesitação em mostrar à comunidade científica o que de fato os autores estavam introduzindo na Lingüística, um programa de pesquisa radicalmente diferente do tradicional, baseado nos sistemas dinâmicos não-lineares. Há, claramente, nesses textos, uma tentativa de aproximação com a fonologia tradicional, como notamos em Browman & Goldstein (1989) que associam a Fonologia Articulatória à Geometria de Traços (Clements, 1985). Isso se apresenta também, de forma bastante evidente, em Browman & Goldstein (1990b),

que aproximam as camadas da FAR às camadas C e V da fonologia CV (Clements & Keyser, 1983).

Essas manobras mostram uma grande incongruência teórica, pois o programa de pesquisa baseado nos sistemas dinâmicos, como o próprio nome diz, trata inerentemente da formação de padrões lingüísticos no tempo e espaço; abordagem radicalmente diferente do paradigma tradicional, que trata dos fenômenos lingüísticos como se a língua estivesse congelada no tempo. Como então podemos comparar paradigmas tão distintos? A questão crucial não é comparar e sim explicar os símbolos lingüísticos como padrões que emergem da variação do contínuo da fala.

A aplicação da TSD à fala tratava basicamente dos aspectos tradicionalmente chamados segmentais. Houve, no entanto, na FAR (Browman & Goldstein, 1990b), uma tentativa de se atribuir uma camada rítmica independente às pautas gestuais, como podemos notar na figura 1.5. Tais camadas rítmicas foram utilizadas a fim de se ter informações nas pautas referentes à acentuação dos gestos articulatórios. A acentuação dos gestos seria atribuída à constelação gestual formadora das sílabas das palavras. O grau de acentuação afetaria ou a rigidez da mola do MDT ou o grau de constrição gestual. Por exemplo, uma vogal acentuada diminuiria a rigidez da mola para gerar uma vogal mais alongada, típica em situações de acento lexical ou frasal. Já, para aumentarmos a duração de uma consoante em sílaba acentu-

ada, dois mecanismos são possíveis: redução da rigidez da mola ou aumento do percurso até o alvo articulatorio (variação da posição inicial da mola). Nessa abordagem investigativa, as camadas rítmicas seriam, pois, responsáveis pelas acentuações nos enunciados e a silabicidade emergiria da interação complexa entre os componentes das pautas gestuais. Estudos posteriores, entretanto, não deram prosseguimento às hipóteses rítmicas aventadas por esses pesquisadores. Apesar disso, como veremos na seção 1.4.2 sobre o gesto prosódico, várias dessas idéias foram abordadas, como a questão de redução da rigidez do sistema massa-mola do gesto articulatorio em posição de acento frasal.

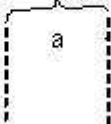
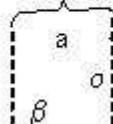
Camada	Gestos	
Rítmica	$x_{\text{secundário}}$	$x_{\text{primário}}$
Corpo de língua		
Ponta de língua		
Lábios		

Figura 1.5: Camadas rítmica e oral do sintagma “a paz” [a'pas]. a representa um gesto faringal, β , um gesto de fechamento bilabial e σ , um gesto de aproximação alveolar crítica. Por sua vez, o $x_{\text{primário}}$ representa um acento frasal e o $x_{\text{secundário}}$, uma proeminência acentual secundária.

Port, Cummins & Gasser (1995) propõem também uma aplicação dos sistemas dinâmicos para tratar dos aspectos prosódicos da fala. Este programa de pesquisa foi chamado de Fonologia Temporal (embora o nome não tenha repercutido tanto quanto o da FAR). Segundo esses autores, como na FAR,

essa área de pesquisa fonológica lidaria com a produção e percepção da fala no tempo e com a descrição de línguas naturais usando sistemas dinâmicos. Sua tese principal é a de que o sistema fonológico de uma língua seria um complexo dinâmico específico com base em sistemas dinâmicos não-lineares como osciladores silábicos e acentuais.

1.4.1 A Fonologia Temporal

A abordagem lingüística de Port, Cummins & Gasser (1995), diferentemente da exposição lingüística da FAR, assume, explicitamente, que uma revisão radical da natureza das teorias fonológicas e lingüísticas é necessária. Os autores, pois, baseados nos dados lingüísticos e não-lingüísticos (principalmente os dados de controle motor) da aplicação da TSD a campos distintos, propõem um ponto de vista lingüisticamente radical, em que o tempo e a estrutura temporal em todos os níveis, do fonético ao sintático, devam ser incluídos e ter papel fundamental nas análises lingüísticas (p. 2). Dessa forma, tanto os símbolos discretos quanto os gestos articulatórios fazem parte de um mesmo todo e estão envoltos em um sistema dinâmico complexo.

Após exemplificarem estudos lingüísticos em que a aplicação da TSD foi efetuada com sucesso (vide seção 1.3), Port e al. afirmam que a estrutura temporal da produção da fala é um forte candidato a ser explorada pela TSD.

Nesse momento, os autores estão se referindo, particularmente, às noções tradicionais de “acento” e “métrica” (Liberman & Prince, 1977; Hayes, 1985, 1995), as quais são modeladas com a utilização de símbolos hierarquicamente arranjados, mas internamente estáticos.

Considerar o estudo lingüístico com base no programa de pesquisa dinamicista, porém, segundo Port et al., apesar de sucessos nas áreas de controle motor, revela-nos certas questões: Será que a TSD é realmente relevante para teorias sobre línguas? Será que tudo isso tem alguma relevância para a Lingüística? Como lidar com um tema tão abstrato quanto uma gramática?

Os pesquisadores argumentam que, apesar de o controle motor da fala ser, em sua maior parte, um processo cognitivo, ele sempre resulta no deslocamento de objetos físicos reais (ex. língua e mandíbula). Por esse motivo, alguns lingüistas podem dizer que uma teoria do controle motor, a qual envolve sistemas físicos, não é adequada para a teorização lingüística. Entretanto, é sabido que a percepção da fala é um processo estritamente cognitivo e que não envolve deslocamento macroscópico de estruturas físicas. Baseados nessa informação, os autores se propõem a mostrar que a percepção da fala para estruturas prosódicas é modelada adequadamente por um sistema dinâmico abstrato, o qual se assemelha às estruturas dinâmicas empregadas no controle motor para a execução de tarefas. Dessa forma, mostrando que produção e percepção da fala são processos semelhantes, pode-se mostrar

a eficiência da abordagem dinamicista à linguagem e que, especificamente, “the phonology of any language is itself simply a particular complex dynamic system — a system based on non-linear dynamic systems such as oscillators which exist at levels like the syllable and foot” (Port et al., 1995, p.4).

A fim de provar sua hipótese, Port et al. estudam o ritmo lingüístico através de análises baseadas em durações lingüísticas relevantes, tais como VOT, durações vocálicas, intervalos inter-vocálicos (Port, 1986; Port, 1990; Port et al. 1995), e se propõem a mostrar que noções tradicionais como “pé”, “vogal breve”, “ritmo acentual”, dentre outras, podem ser reinterpretadas como manifestações de um modelo perceptual mais amplo. Nesse panorama, são reavaliadas noções simbólicas tradicionais e reinterpretadas através de noções dinâmicas como oscilador, velocidade, ângulo de fase, atratores, etc. Com isto a descrição lingüística “will require no additional clocking devices to perceive speech or produce it in real time since the model will already live in time”. (Port et al., 1995, p.4)

Os autores da Fonologia Temporal, mesmo reconhecendo a radicalidade de sua proposta, afirmam estar à procura de propriedades universais da estrutura temporal da fala, o que os aproxima da busca por “universais lingüísticos” na lingüística tradicional. No entanto, procuram mecanismos dinâmicos gerais para o ritmo e a métrica, que podem ser aplicados a qualquer língua e, inclusive, à música.

O ritmo na Fonologia Temporal

Vários estudiosos, falantes nativos do inglês, têm a impressão de que as sílabas acentuadas ocorrem a intervalos regulares na fala. Conseqüentemente, parece-nos que, para uma análise rítmica do inglês, devemos nos guiar pela organização das sílabas acentuadas, portanto pés métricos, ao longo dos enunciados (Port et al., 1995). No entanto, evidências da literatura fonética não comprovam tal regularidade acentual. Como então extrair uma regularidade na produção a partir da sensação regular na percepção? Port et al. sugerem que, para isto, devemos recorrer a um experimento chamado de tarefa cíclica da fala⁴ (doravante TCF). Com esse experimento eles conseguiram evidenciar uma dinâmica subjacente à fala e também explicar a sensação de isocronia nas moras do japonês e nos intervalos interacentuais do inglês.

A TCF, baseada nos experimentos de Stetson (1988), consiste na repetição contínua de frases em sincronia com um estímulo auditivo externo, como um metrônomo. Estudos como os de Classé (1939) e Cummins (1997), dentre outros, mostraram que em frases interconectadas, como na TCF, a fala se apresenta com um ritmo bastante regular.

Sabe-se que na música os intervalos rítmicos são bastante regulares⁵ obedecendo a fórmulas de compasso como $\frac{2}{4}$, $\frac{3}{4}$, $\frac{4}{4}$, em que o numerador representa

⁴Em inglês, “speech cycling task”.

⁵No entanto, mesmo na música, estudos revelam que os intervalos rítmicos não são absolutamente regulares, sujeitos a variações (Clarke, 1989; Church & Broadbent, 1990).

o número de batidas por compasso e o denominador o tipo de figura rítmica a ser utilizada. No exemplo acima temos duas ($\frac{2}{4}$), três ($\frac{3}{4}$) ou quatro ($\frac{4}{4}$) batidas de semicolcheia⁶ por compasso. Obviamente, como podemos ver na figura 1.7, outras subdivisões das notas podem ocorrer, mas as mesmas seguem a fórmula rítmica do compasso, o qual, funcionando como um atrator, estabelece o ritmo da música.

Especulando se essa regularidade rítmica na música pode ser encontrada na fala, Port et al. afirmam que, se o inglês é rítmico ao nível das sílabas acentuadas, deve existir uma métrica subjacente que coordene os períodos interacentuais (como as notas musicais) com um outro período maior (como um compasso), como o período da frase inteira. A relação entre os intervalos interacentuais com o período da frase agiria como um atrator na fala.



Figura 1.6: Representação das notas musicais a partir da nota semibreve de valor 1.

A literatura do controle motor de membros humanos evidencia tais atratores (Kelso & de Guzman, 1988; Treffner & Turvey, 1993). No experimento

⁶Divisões rítmicas na música se baseiam na nota chamada semibreve. Todas as outras notas são oriundas desta em divisões pela metade. Ex.: semibreve (1), mínima ($\frac{1}{2}$), e assim por diante. Vide figura 1.6.



Figura 1.7: Representação de duas composições que, apesar de utilizarem a mesma fórmula de compasso ($\frac{4}{4}$), apresentam divisões rítmicas distintas. Na primeira pauta (trecho de 2-Part Invention # 13 em A menor. B.W. III de Johann Sebastian Bach), temos subdivisões rítmicas da semínima em colcheias ($\frac{1}{8}$ da figura 1.6) e semicolcheias ($\frac{1}{16}$ da figura 1.6), além de pausas. Na segunda pauta (trecho do segundo piano de “Country Dance” para dois pianos de Ludwig van Beethoven), temos apenas uma figura rítmica, a semínima, que se alterna entre batidas fortes (cabeça do compasso) e fracas (restante do compasso). O compasso é delimitado por barras verticais.

de Treffner & Turvey, por exemplo, os sujeitos tinham a tarefa de tentar sincronizar o balanço de um pêndulo com a mão direita enquanto a mão esquerda balançava uma baqueta com uma outra mão. Cada oscilador tinha um quociente de frequência diferente, estabelecendo relações como de 2:1 (2 movimentos de pêndulo para 1 de baqueta), 1:1, 3:1. A instrução dada foi de que mantivessem a movimentação dos osciladores independentes. No entanto, houve uma tendência forte nesse experimento de ambos osciladores se acoplarem e, não importando o quociente, caírem em relações harmônicas mais simples como 1:2 ou 1:1⁷. Tais quocientes foram considerados como atratores do sistema de controle motor.

O experimento detalhado a seguir parte dos mesmos princípios dos trabalhos na área de controle motor, ou seja, da presença de atratores que favorecem certas relações rítmicas. Nesse experimento, esperava-se que os falantes produzissem quocientes em ângulos de fase $\frac{1}{2}$, $\frac{1}{3}$, $\frac{2}{3}$ em relação à frase toda, da mesma forma que, na música, as notas musicais ocorrem em frações harmônicas do compasso, como vemos na figura 1.8. Como nos compassos simples o compasso todo vale 1, no $\frac{4}{4}$ a terceira nota começa no ângulo de fase $\frac{1}{2}$ (0.5), no $\frac{3}{4}$, a terceira nota começa em $\frac{2}{3}$ (0.666...), e assim por diante. Relações harmônicas como essas são esperadas no experimento relatado a seguir.

⁷Obviamente, certos músicos conseguem produzir ritmos mais intrincados. Contudo, tais ritmos não são considerados naturais aos humanos e, portanto, devem ser aprendidos.

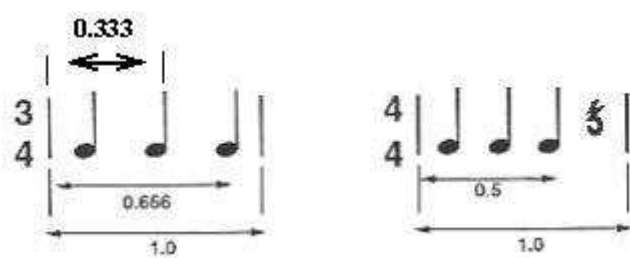


Figura 1.8: Representação de relações harmônicas na música nos compassos $\frac{3}{4}$ e $\frac{4}{4}$.

Relações harmônicas em inglês

No experimento descrito, os pesquisadores utilizaram a noção de *p-center* (Scott, 1993; Marcus, 1981; vide também Barbosa, Arantes, Meireles & Vi-eira, 2005 para uma versão resumida) como uma medida aproximada do local onde as sílabas seriam alinhadas com um outro estímulo. Especificamente, foi utilizado o programa *Lutear*, o qual gerou os contornos de amplitude suavizados vistos na figura 1.9. O momento de ocorrência do pulso silábico (*syllable beat*) foi definido como o ponto entre o máximo local e o mínimo local anterior (marcados com retas verticais na figura 1.9), que corresponde, aproximadamente, ao início acústico da vogal⁸.

Os estímulos consistiram de repetições do par *take* e *cards* produzidos consecutivamente com vários espaçamentos. O intervalo entre os pulsos de dois consecutivos *take* foi fixado em 1.5 segundos, correspondendo ao ciclo

⁸Posteriormente, veremos que tais noções foram aplicadas no programa *Beat Extractor* utilizado nas análises acústicas desta tese.

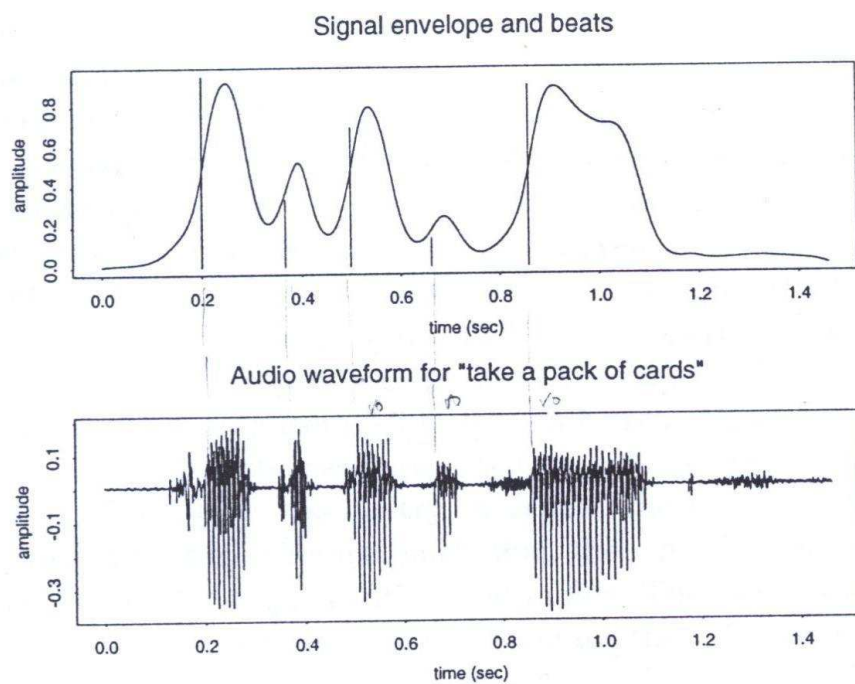


Figura 1.9: Representação do momento de ocorrência do pulso silábico, tirado através de uma suavização da forma de onda. As linhas verticais no envelope de energia do sinal representam tal momento.

completo (vide figura 1.10). O ângulo de fase de *cards* foi variado de 0.3 a 0.65 e foi chamado de Φ_{cards} (ângulo de fase de *cards*).

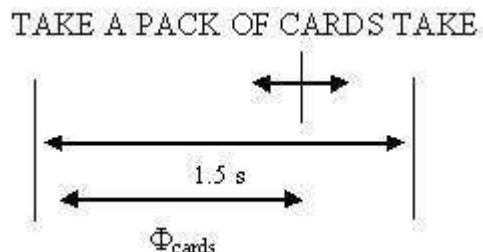


Figura 1.10: Representação do ângulo de fase de *cards* Φ_{cards} . O traço vertical em *cards* representa a movimentação possível da palavra para a direita ou esquerda, manipulada sinteticamente no experimento de Port et al. (1995).

A tarefa consistia na repetição contínua da frase *take a pack of cards* sincronizada com os estímulos sintetizados. O conjunto de dados consistiu de três grupos: 1) dados produzidos simultaneamente com as palavras sintetizadas; 2) dados produzidos logo após cessar o estímulo; 3) dados produzidos depois de uma pausa.

Os resultados mostraram, como no experimento de Treffner & Turvey (1993), uma preferência por certos ângulos de fase, a saber: $\frac{2}{4}$, $\frac{3}{4}$, $\frac{4}{4}$ ⁹, considerados atratores na TCF estudada. Apesar de vários padrões rítmicos terem sido estimulados, os falantes convergiram para determinados atratores de fase, localizados em torno dos harmônicos definidos pela repetição da frase inteira. Desta forma, Port et al. (1995:22) mostraram que há “strong

⁹Coincidentemente, esses são os compassos mais utilizados nas músicas populares, dentre elas: rock, blues, samba, bossa nova.

prima facie evidence for a dynamic process of entrainment, rather than mere concatenation of individual durations” e que, como na música, os sujeitos mostraram preferência por um certo tipo de unidade de medida — a frase toda — e uma unidade de pulsação — o início acústico da sílaba (p. 22). Além disso, os autores esperam, devido à observação de um padrão muito recorrente, que as mesmas ocorram na fala casual (p. 22).

A hipótese adotada por Port, Cummins & Gasser (1995) para explicar a sensação perceptual que os falantes têm, na fala informal, de isocronia moraica/acental é a de que existe um oscilador adaptativo (cf. McAuley, 1994, 1995), que é sincronizado com as batidas das moras/intervalos interacentuais individuais. Estes dois osciladores acoplados (o estímulo acústico e o oscilador cognitivo) fariam com que a média da duração das frases fosse percebida como isócrona. Como veremos na seção 1.4.3, idéias semelhantes são encontradas no modelo dinâmico de produção do ritmo da fala utilizado nesta tese.

Implicações teóricas

As implicações teóricas dos trabalhos de Robert Port e colegas são: (i) uma revisão radical nas teorias fonológicas e lingüísticas em geral, devido à aplicação da TSD à linguagem; (ii) uma unificação da teoria da linguagem com teorias do controle motor e percepção. Com isso, a ciência lingüística pode

novamente se integrar ao resto das ciências cognitivas (Port & Van Gelder, 1995).

Uma última indagação oriunda do trabalho de Port et al. (1995) é a seguinte questão: se as periodicidades encontradas na fala são meros aspectos naturais do mundo físico, então o ritmo da fala seria apenas uma característica universal da movimentação de membros do corpo humano? Se assim o fosse, os ritmos das línguas seriam uniformes, semelhantes, por exemplo, à locomoção de animais. No entanto, como sabemos, as línguas do mundo se encaixam em tipologias rítmicas distintas. Por exemplo, no ritmo em japonês, as moras exercem um papel mais relevante do que em inglês e português, em que os intervalos interacentuais são mais relevantes. Assim, Port et al. (1995:26) afirmam que: “these results support the hypothesis that the phonology of a language is a description of a motor and perceptual *skill* that may make opportunistic use of any mechanical or other constraints on coordination, but which is still obviously a cognitive function on par with other high level functions”. Além disso, os pesquisadores dizem que “psychological mechanisms for adaptive oscillators are consistent with known psychophysical data and with known properties of nervous systems. And they can be expressed in terms that permit specification of real time rather than the pseudo-time of serially ordered strings of symbols (p. 27)”.

Na seção seguinte faremos a descrição do gesto prosódico (π -gesture), um

aperfeiçoamento da FAR para lidar com aspectos prosódicos dinâmicos da linguagem, o qual compartilha muitas das idéias expressas pela Fonologia Temporal descrita acima.

1.4.2 O gesto prosódico

Byrd & Saltzman (2003), baseados em estudos articulatórios anteriores (Byrd & Saltzman, 1998; Byrd, Kaun, Narayanan & Saltzman, 2000; Byrd, 2000; Saltzman & Byrd, 2000), apresentam à comunidade científica o gesto prosódico. O mesmo é elaborado a partir dos princípios empírico-teóricos da FAR, porém se propõe a analisar a relação da estrutura prosódica frasal com o controle e coordenação da articulação à luz da TSD. A idéia principal do trabalho era investigar mudanças das estruturas articulatórias em função de diferentes fronteiras frasais. Propõe-se, para explicar tais mudanças, o gesto prosódico, o qual atua localmente em tais fronteiras, modificando a organização temporal dos gestos articulatórios.

A justificativa dada pelos pesquisadores do porquê de se usar a TSD como programa de pesquisa é o fato de a mesma possibilitar uma ponte nítida entre as unidades subjacentes de controle da fala e a variabilidade na superfície das mesmas. Além disso, abrem-se portas para a integração de outros aspectos da linguagem (sociais, psicológicos, biológicos e culturais) em um mesmo

programa de pesquisa.

A elaboração do gesto prosódico se baseia na FAR, como uma teoria formal de unidades fonológicas e suas respectivas organizações, e no MDT, como um modelo quantitativo que implementa gestos articulatórios nos sistemas de produção de fala. Recordaremos a seguir, sucintamente, as principais características da FAR.

Na FAR, as unidades fonológicas são os gestos articulatórios, os quais possuem a dupla função de servirem como unidades de informação, ou seja, atuam no contraste lingüístico, e de serem unidades de ação na produção da fala. Sendo assim, não há mediação entre a representação fonológica e sua implementação pelo sistema de produção da fala. As características principais dos mesmos são: (i) os gestos são abstratos, pois são definidos em termos de tarefas lingüísticas, como “realize uma constrição labial”, e não, “movimente os lábios superiores, os lábios inferiores, e a mandíbula em tantos milímetros em uma determinada direção”; (ii) os gestos possuem propriedades espaço-temporais inerentes; (iii) os gestos são coordenados uns aos outros em relações de fase bem definidas.

Antes de apresentarmos detalhes do gesto prosódico, resumiremos as noções de modelos de tempo intrínseco e extrínseco, para obstruir possíveis interpretações errôneas em relação ao relógio empregado no gesto prosódico. Nos modelos de tempo extrínseco, as propriedades temporais e o tempo não

são parte da representação simbólica das unidades lingüísticas. Por outro lado, nos modelos de tempo intrínseco, a temporalidade é incorporada na definição dessas unidades. Sendo assim, como a FAR trabalha à luz da TSD, ela se encaixa, necessariamente, nesse tipo de modelo. Da mesma forma, apesar de se utilizar um relógio para controlar o fluxo de ativação gestual, o mesmo é considerado como integrante de modelos de tempo intrínseco, pois as dinâmicas do nível de constrição e do nível de ativação são bidirecionalmente acopladas e, formam, pois, um coletivo dinâmico de alta ordem (Byrd & Saltzman, 2003, p. 156).

Definição do gesto prosódico

O gesto prosódico surgiu a partir do fato de haver alongamento gestual na vizinhança de fronteiras frasais. Dentre as características dos gestos em fronteiras frasais, apontamos: (i) aumento local de duração (Beckman, Edwards & Fletcher, 1992; Byrd & Saltzman, 1998; Byrd, 2000); (ii) diminuição da velocidade de pico (Cho, 2001); (iii) contato linguo-palatal estendido (Fougeron, 2001; Cho & Keating, 2001); (iv) diminuição no parâmetro rigidez (Beckman et al., 1992; Byrd & Saltzman, 1998; Byrd, 2000). Entretanto, em um outro estudo foi apontado que a modulação da rigidez sozinha é insuficiente para explicar modificações gestuais nas fronteiras prosódicas (Saltzman & Byrd, 2000; Cho, 2001). No trabalho descrito a seguir, o controle do alon-

gamento local, em vez de se basear apenas na manipulação do parâmetro rigidez, foi feito pela modulação da composição temporal do enunciado ao redor de fronteiras frasais.

O termo *gesto prosódico* foi inaugurado em Byrd, Kaun, Narayanan & Saltzman (2000) e Byrd (2000), onde se propôs que os gestos prosódicos atuam em fronteiras frasais, agindo transgestualmente, para lentificar todos os gestos de constrição ativos, de acordo com o nível de ativação do gesto prosódico (vide figura 1.11). Tal proposta foi feita como uma tentativa de se fazer uma implementação dinâmica da estrutura frasal, seguindo os parâmetros da FAR. Dentre as suas características principais, apontadas por Byrd & Saltzman (2003), estão: (i) os gestos prosódicos, tais como os gestos da FAR, possuem uma extensão temporal e se sobrepõem aos gestos de constrição; (ii) as pautas gestuais do gesto prosódico representam a atividade de um conjunto de atratores pontuais gestuais abstratos, a fim de se manter o modelo o mais abstrato possível; (iii) os gestos prosódicos não possuem uma realização articulatória independente. Sendo assim, eles só são realizados indiretamente pelo seu efeito na dinâmica articulatória.

O trabalho de Byrd & Saltzman (2003) difere dos trabalhos anteriores sobre o gesto prosódico. Nos trabalhos anteriores, o gesto prosódico era considerado como perturbador do parâmetro rigidez, lentificando, assim os gestos na vizinhança da fronteira frasal. Byrd & Saltzman (2003), por outro

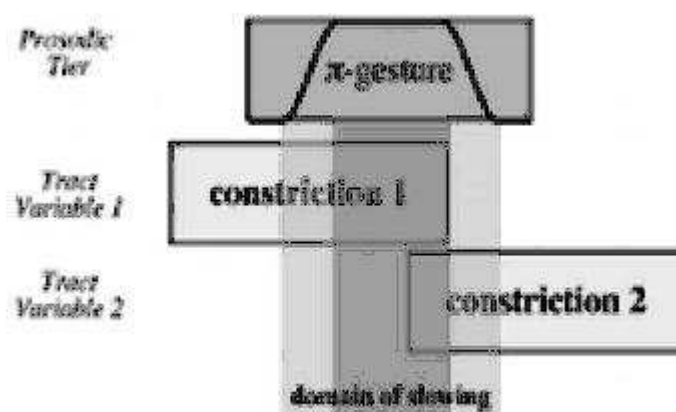


Figura 1.11: Gesto prosódico agindo sobre dois gestos adjacentes, como nas consoantes em negrito de “I can’t make it. **P**ay it yourself” (adaptado de Byrd & Saltzman, 2003, p. 160).

lado, verificam a incompletude de tal modificação de parâmetro para explicar os fenômenos estudados e propõem que o gesto prosódico lentifica localmente o relógio que controla o fluxo temporal de um enunciado.

Segundo Byrd & Saltzman (2003), a aplicação dos gestos prosódicos em um modelo dinâmico de produção da fala é capaz de fazer várias previsões lingüísticas, dentre elas: a) os gestos serão lentificados se ocorrerem durante o período de ativação do gesto prosódico (Beckman et al., 1992; Byrd & Saltzman, 1998; Byrd, 2000; Fougeron, 2001); b) o grau de lentificação será maior no grau de ativação máximo do gesto prosódico (Byrd & Saltzman, 1998; Byrd, 2000); c) os efeitos do gesto prosódico se limitarão aos gestos perto da fronteira frasal e não ocorrerão em gestos longe desta (Cambier-Langeveld, 1997; Edwards et al., 1991); d) a fonte dinâmica de duração intragestual e os

efeitos de organização temporal intergestuais serão, em princípio, os mesmos à direita e à esquerda dos gestos na vizinhança de uma fronteira frasal (Byrd & Saltzman, 1998; Byrd et al., 2000; Cho, 2001); e) diferentes tipos de fronteira, como sintagma entonacional, intermediário, diferirão apenas em grau, não no tipo de efeito de alongamento (Byrd & Saltzman, 1998; Cho, 2001).

As simulações¹⁰ feitas por Byrd & Saltzman (2003) mostraram que a modulação da taxa do relógio foi realmente mais eficaz do que a modulação da rigidez, a qual afeta apenas as durações gestuais, para explicar o comportamento articulatório em fronteiras frasais. Além do aumento de duração do gesto, tal modulação explica a redução de sobreposição intergestual e o aumento na magnitude espacial, ambos presentes nas fronteiras prosódicas.

Os autores, pois, propõem que tal relógio é capaz de explicar diferenças de TE associadas a fatores sociolingüísticos, dialetais e individuais. Por fim, Byrd & Saltzman (2003) concluem que seus resultados, apesar de estarem ainda em um estágio primitivo de análise, são promissores para uma confluência da prosódia com a dinâmica da fala.

Pelo relatado acima, notamos que diversos autores mostraram ser possível estudar a linguagem com base na TSD, a qual consegue abarcar num mesmo sistema aspectos tradicionalmente considerados segmentais quanto prosódi-

¹⁰Para detalhes do modelamento do gesto prosódico, consultem Byrd & Saltzman (2003:163-175).

cos da linguagem. No entanto, nenhum desses pesquisadores relacionou de forma clara e concisa os aspectos segmentais com os aspectos prosódicos. O modelo dinâmico do ritmo da fala, em contínuo desenvolvimento pelo Grupo de Estudos em Prosódia da Fala da UNICAMP, apresenta, por outro lado, de maneira bastante precisa, a interação prosódia-segmentos. Além disso, o modelo relaciona todos os componentes lingüísticos *hardcore* da gramática: fonética-fonologia, léxico, semântica e sintaxe. Tal modelo ainda abre possibilidades para inclusão de aspectos culturais, sociológicos, etnológicos, dentre outros, em uma etapa futura de seu desenvolvimento. Descreveremos a seguir tal modelo, o qual será usado como base empírico-teórica das análises lingüísticas de nossa tese.

1.4.3 O modelo dinâmico de produção do ritmo da fala

Os estudos de Plínio Barbosa (1996b, 2001, 2002a, 2006¹¹, dentre outros) foram fundamentais para o desenvolvimento de um modelo dinâmico do ritmo (doravante MDR). Acreditamos que a justificativa teórica fundamental de tal modelo é a própria utilização da TSD, pois, como vimos anteriormente, um sistema dinâmico muda de estado com a passagem do tempo, uma das propriedades essenciais do ritmo não abarcada na teoria lingüística tradicio-

¹¹O livro de Barbosa (2006) faz uma descrição ampla e objetiva de todos os aspectos empírico-teóricos da elaboração do MDR.

nal de maneira satisfatória. Além disso, a TSD é capaz de lidar com todos os aspectos interdisciplinares lingüísticos, extralingüísticos e paralingüísticos envolvidos na produção da fala. Outras justificativas, vistas anteriormente, para um MDR aplicado à fala são: (i) a incomensurabilidade entre os níveis discreto e contínuo. Matematicamente, inclusive, não é possível uma relação bijetora entre o conjunto dos números naturais (discretos) e o dos números reais (contínuos); (ii) a presença de sinergias na fala, através da presença de estruturas coordenadas para a realização de uma tarefa. Barbosa (2006:13) afirma que o MDR satisfaz todas as três propriedades das estruturas coordenadas citadas por Fowler (1980), a saber: (i) a sua implementação por uma equação de diferenças finitas define uma classe equivalente de movimentos; (ii) os osciladores do modelo geram intervalos temporais hierarquizados em escalas de diferentes ordens de magnitude; (iii) os osciladores acoplados geram a duração de unidades VV, definidas abaixo, ciclicamente. Dessa forma, Barbosa (2006:4) define o MDR, mostrado na figura 1.12, como “uma implementação matemático-computacional de um sistema dinâmico do ritmo da fala que exhibe todas as propriedades de um sistema auto-organizado¹²”.

Nesse modelo considera-se como os gradientes e as categorias fônicas dos níveis prosódicos podem ser modelados de forma unificada, através de um

¹²Para uma explicação mais detalhada de tais propriedades no MDR, vide Barbosa (2006).

esquema híbrido, que combina informação gramatical discreta com informação fônica que alia parâmetros contínuos a parâmetros discretos em vários níveis.

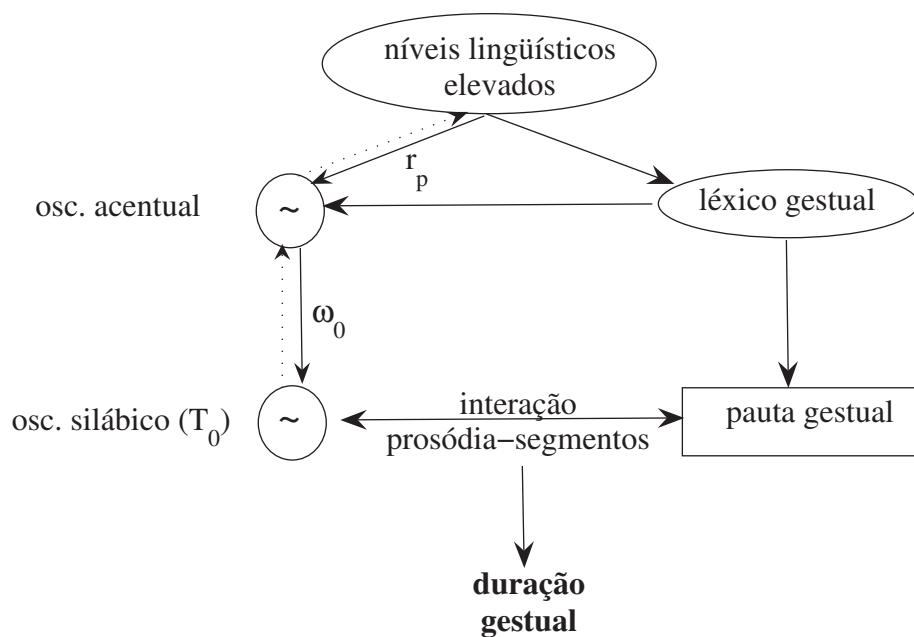


Figura 1.12: Diagrama do MDR. Vide texto para detalhes.

Nota-se, na figura 1.12, que o MDR se subdivide em dois ramos: (i) o prosódico e (ii) o segmental¹³. Os níveis linguísticos mais elevados (sintaxe, semântica, pragmática)¹⁴ exercem influência em ambos os ramos. Do ramo

¹³Apesar de não lidarmos com segmentos, como na visão tradicional, utilizaremos tal nomenclatura para facilitar a discussão teórica.

¹⁴No presente momento, apenas a interação entre a informação sintática e o oscilador acentual estão implementadas. Dessa forma, em nossa tese, referências a níveis linguísticos mais elevados devem ser lidas como informação sintática.

prosódico, a interação entre os níveis lingüísticos mais elevados e o oscilador acentual se dá através de um grau de acoplamento, r_p . O oscilador acentual, por sua vez, está bidirecionalmente acoplado ao oscilador silábico através de uma força de acoplamento, w_o . Do ramo segmental, o léxico gestual (Browman & Goldstein, 1989), influenciado pelos níveis lingüísticos mais elevados, serve de entrada para a geração das pautas gestuais (Browman & Goldstein, 1990b). A junção desses dois níveis, chamada no MDR de interação prosódia-segmentos, gera a duração acústica ou articulatória¹⁵ (Browman & Goldstein, 1986, 1989, 1990b, 1992).

Barbosa (2002a, 2006) faz uma descrição detalhada do modelo de produção do ritmo da fala do PB. Segundo o pesquisador, o ritmo da fala é concebido como uma consequência da variação da duração percebida ao longo da sentença. Postula-se, pois, que o parâmetro de controle prosódico para o estudo do ritmo em PB é a duração, a qual serve para assinalar o acento frasal. A duração no MDR divide-se em duas: 1) intrínseca, correspondente à duração abstratamente especificada (a partir de relações de fase entre gestos) em um léxico gestual. Tal abstração se faz necessária no modelo, a fim de se excluir a variância da duração intrínseca dos gestos na pauta gestual, reduzindo, assim, os graus de liberdade, e tornando a análise rítmica independente de tais variações duracionais; 2) extrínseca, responsável propriamente

¹⁵A implementação articulatória no modelo encontra-se em desenvolvimento.

pelo nível rítmico do controle da duração. Esta última é implementada por dois osciladores acoplados: um silábico e um acentual. A interação entre esses dois osciladores gera o acento frasal, definido como uma proeminência acentual obtida através de um pico local de duração ao longo do enunciado. Tais acentos frasais ocorrem sempre em sílabas lexicalmente marcadas como tônicas pelo léxico gestual. As equações principais de tal modelo são descritas abaixo.

Através da adaptação do modelo de percepção de intervalos de tempo de McAuley (1995) para a produção da fala, Barbosa (2002a) propõe uma equação de diferenças finitas, expressa abaixo, que modela a interação entre os osciladores silábico e acentual:

$$\Delta T = \alpha.T.s(n).i(m) - \beta.(T - T_0).i(m - 1) \quad (1.1)$$

Em que $s(n)$ é definida iterativamente por:

$$s(0) = w_0.e^{(-N+2)}, e \ s(N - 1) = w_0.e^{(-5,81+0,016.T_0)} \quad (1.2)$$

$$s(n) = (1 - w_0).s(n - 1) + w_0.e^{(-N+n+2)}, \text{ para } 0 < n < N - 1 \quad (1.3)$$

Na equação (1.1) acima, ΔT é o valor de variação de duração (positivo ou negativo) a ser acrescentado à duração do grupo vogal-a-vogal (VV) cor-

rente (do *onset* de uma vogal ao *onset* da imediatamente seguinte) no grupo acentual. Para tanto contribuem fatores para aumentar a duração (primeira parcela), e fatores para diminuir a duração (segunda parcela).

Na primeira parcela, temos os fatores α , T , $s(n)$ e $i(m)$, em que α é a taxa de indução, ou seja, modela o aumento, de maneira fina, da duração dos VVs até o acento frasal; T é o período do oscilador silábico atual; $s(n)$ é a função de sincronismo entre os osciladores acentual e silábico (que integra um parâmetro, w_0 , variável de língua para língua. Para o locutor estudado, segundo Barbosa (2002a), adotaremos o valor de 0,8); e $i(m)$ corresponde à amplitude do ciclo do oscilador acentual que culmina no acento frasal. Este último pode ser modificado de acordo com os diferentes tipos de estruturas sintáticas e prosódicas (no atual estágio da pesquisa foi atribuído um valor constante de 1 para esse parâmetro).

Na segunda parcela, temos os fatores β , $(T-T_0)$ e $i(m-1)$, em que β é a taxa de decaimento, i.e., simula um reinício do período do oscilador silábico; T_0 representa o período do oscilador silábico desacoplado. Por sua vez o termo $(T-T_0)$ mostra que quanto maior for a diferença entre os dois períodos (atual e desacoplado), mais forte será a queda na duração dos VV iniciais no grupo acentual seguinte.

Barbosa (2006) aponta que o MDR é lingüístico nos seguintes sentidos:

(i) o ritmo lingüístico é consequência da maneira pela qual os pulsos do

oscilador acentual se alinham com os inícios (*onsets*) de vogais acentuadas lexicalmente, i.e., com material lingüístico; (ii) a hierarquia entre as magnitudes dos pulsos do oscilador acentual gera uma métrica dinâmica, pela interação entre os níveis lingüísticos mais elevados e as restrições do sistema de produção/percepção da fala; (iii) a força de acoplamento relativa é variável de língua para língua; (iv) a interação prosódia-segmentos é distinta inter-línguas (Keating et al., 2003; Barbosa, 2001); (v) a informação sintática é crucial, embora não única, para a explicação da variabilidade da localização e magnitude dos acentos frasais de um enunciado; (vi) os acentos frasais são uma consequência superficial das proeminências expressas por picos de duração abstrata do oscilador silábico; (vii) o acento lexical é definido na pauta gestual abstrata.

Nas próximas seções serão apresentadas evidências científicas, apontadas por Barbosa (2006), que suportam a utilização dos osciladores silábico e acentual no MDR.

O oscilador silábico

O movimento oscilatório da mandíbula é o que mais se aproxima à especificação das sílabas nas línguas do mundo, ou seja, as sílabas são delimitadas, *grosso modo*, pelo deslocamento vertical desse articulador. *Grosso modo*, pois, como veremos no estudo articulatorio do capítulo 4, as fronteiras silábicas

nem sempre são bem demarcadas pelo deslocamento mandibular. Apesar de desvios na demarcação de sílabas, tal articulador é extremamente importante na caracterização das mesmas. Como se comportariam, pois, as vogais e consoantes nesse panorama oscilatório?

Estudos articulatórios (Keating et al., 1994; Lindblom, 1983; Rhardisse & Abry, 1995) mostram que as sílabas parecem se organizar através de fluxos vocálicos intermeados por consoantes. Rhardisse & Abry (1995), por exemplo, mostraram, através da observação da movimentação vertical da mandíbula, que as vogais influenciam muito mais a extensão vertical da consoante do que o oposto. Estudos como esse sugerem a existência de um oscilador silábico na produção da fala, o qual é utilizado no MDR. Convém ressaltar, entretanto, que a periodicidade de unidades do tamanho da sílaba nesse modelo não é coincidente, na maior parte, com a sílaba fonológica. No MDR a unidade prosódica básica é a unidade VV, definida do início de uma vogal até o início da próxima vogal subsequente, independentemente do número de segmentos intervenientes entre as mesmas e de fronteiras silábico-lexicais, conforme podemos observar na figura 1.13.

A idéia de uma produção contínua de vogais intercaladas por consoantes está presente na literatura desde o século XVIII, bem antes da inauguração da lingüística como ciência por Ferdinand de Saussure. Obviamente, devido à insuficiência de recursos tecnológicos à época, no princípio, tais observações

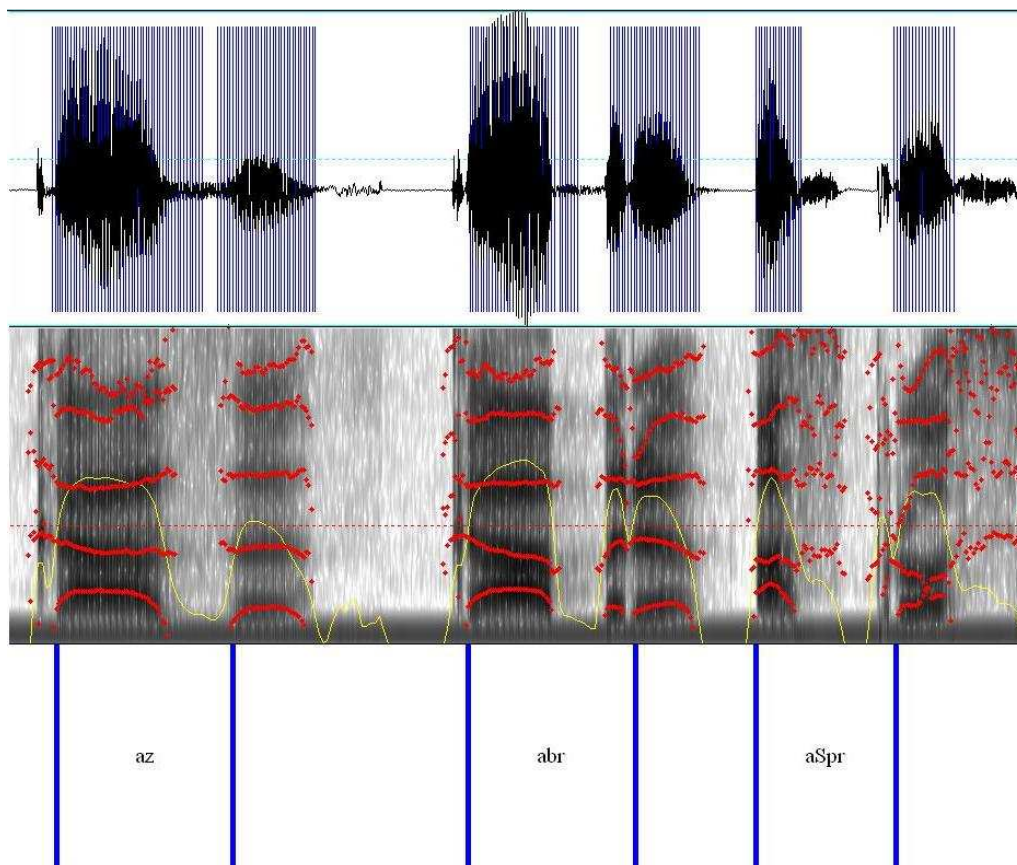


Figura 1.13: Exemplos de unidades VV no português brasileiro: (1) /az/ de “casa”, (2) /abR/ de “cabra” e (3) /aSpr/ de “as próximas”.

se baseavam apenas na observação visual da produção da fala.

Brosses (1765, *apud* Genette, 1976), por exemplo, apresenta em seu trabalho a hipótese de que a produção da fala se subdivide entre *figura* e *fundo*. As figuras corresponderiam às consoantes, as quais, perturbando o pano de fundo amórfico composto pelas vogais, atribuiriam significados à comunicação lingüística. Tal noção de vogais entrecortadas por consoantes pode, inclusive, ser estudada por teorias modernas como a TSD, através da noção de fase, vista na seção 1.2.3.

Até a invenção do espectrograma na década de 1940, estudos qualitativos prevaleceram nos estudos lingüísticos. A partir de então, pode-se vislumbrar em tais aparelhos evidências acústico-visuais de que, na produção da fala, as vogais são entrecortadas por consoantes. Na dimensão acústica podemos até preencher os vazios no espectrograma através de soluções de continuidade do nosso sistema perceptual (Berthoz, 2000), da mesma forma que enxergamos figuras em situações em que elas não estão objetivamente definidas (vide figura 1.14). Como enxergamos um triângulo na fig. 1.14, também podemos perceber a continuidade vocálica no espectrograma da fig. 1.13.

Com o advento do espectrograma, vários estudos proliferaram mostrando a solução de continuidade vocálica no sistema de produção da fala. Dentre eles, citamos os estudos acústicos de Öhman (1966), Stetson (1988) e, mais recentemente, os estudos articulatórios de Tuller & Kelso (1990, 1991).

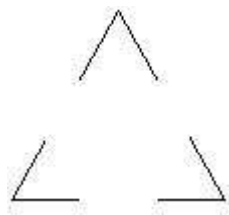


Figura 1.14: Ilusão de ótica na interpretação de imagens. No exemplo acima, observamos um triângulo mesmo ele não estando plenamente construído.

Conforme comenta Barbosa (2006:37), “a evidência experimental da maior estabilidade da sílaba CV é também uma evidência da preferência ou atração do mecanismo de produção por uma sílaba que contenha uma transição do tipo CV”. Tal afirmação se baseia no princípio de que as unidades CV são mais separáveis do ponto de vista acústico-articulatório, o que facilitaria o ancoramento de nossa percepção na transição CV. Por outro lado, conforme veremos abaixo, a localização dos pulsos silábicos delimita uma unidade estável na produção da fala, a unidade VV.

A homogeneidade das unidades VVs é respaldada por dados clássicos da literatura fonética (Peterson & Lehiste, 1960; Lehiste, 1970) de que há uma tendência nas línguas do mundo de vogais diante de consoantes curtas se alongarem e, diante de longas, encurtarem. Esse efeito foi comprovado para diversas línguas como o sueco, o francês e o PB. Tais resultados suportam a idéia de uma unidade VV subjacente homogênea. No entanto, estudos em línguas como o árabe (Mittleb, 1984, 1992, apud Barbosa, 2006), polo-

nês e tcheco (Keating, 1979, 1985, apud Barbosa, 2006) não observaram tal tendência.

Barbosa (2006), no entanto, argumenta que tais contra-exemplos necessitam de observações empíricas mais cuidadosas para se negar a existência de uma tendência universal das unidades VV. Acreditamos que os pesquisadores podem não estar olhando para os dados da forma certa. Pode ser que nesses contra-exemplos haja efeitos compensatórios ao longo dos enunciados e não apenas localmente na unidade VV. Além disso, mesmo que tal homogeneidade não seja universal, Barbosa (2006) afirma que as mesmas seriam tidas como fenômenos da fonética lingüística (Ladefoged, 1971), ou seja, a não-compensação duracional deveria ser aprendida pelos falantes das línguas em questão.

Concluindo, Barbosa (2006) afirma que a estabilidade da unidade VV serve como indício de que uma unidade periódica do tamanho da sílaba funciona como um atrator cíclico no sistema da fala — um oscilador silábico. As vantagens para tal atrator seriam: (i) a necessidade de um consumo mínimo de energia e planejamento para o sistema cognitivo (Fraisse, 1968); (ii) a possibilidade de sincronização com um estímulo externo (vide seção 1.4.1).

O oscilador acentual

O ritmo da fala, bem como outros tipos de ritmo, é composto, segundo Fraisse (1968:28) por “deux éléments apparaissant caractéristiques: la périodicité et la structure” e são “indissociables: la structure se trouve toujours coulée dans une périodicité et la périodicité est toujours organisation de structures”. Sendo, portanto, o ritmo composto por dois elementos indissociáveis — a periodicidade e a estrutura — ambos não poderiam deixar de ser incluídos em um modelo dinâmico do ritmo da fala. Dessa forma, no MDR, a periodicidade é representada pelo oscilador silábico e a estruturação dessa periodicidade é realizada pelo oscilador acentual. Além disso, por serem indissociáveis, ambos estão birecionalmente acoplados.

Como o MDR trabalha com o ritmo da fala, o oscilador acentual, especificamente, “assinala de tempos em tempos posições com informação lingüística relevante ao perturbar de maneira específica a regularidade silábica, tornando a seqüência de *onsets* de vogal anisócrona e, portanto, saliente perceptivamente” (Barbosa, 2006, p. 100). Tal como o oscilador silábico, esse oscilador possui um padrão cíclico, porém, influenciado por fatores lingüísticos de alto nível como sintaxe, semântica e pragmática.

No presente estágio de desenvolvimento, como já dissemos, o MDR trata apenas do acoplamento entre o oscilador acentual com o componente sintático

da gramática, o qual gera dois tipos de efeito nos pulsos do oscilador acentual — um temporal e outro espacial.

O efeito temporal é consequência de dois mecanismos atuantes na especificação dos pulsos do oscilador acentual. O primeiro mecanismo ocorre por sucessões periódicas desses pulsos, em condição de não-acoplamento com os fatores de alto nível, e essas sucessões são da ordem de grandeza de 1–2 segundos para parágrafos lidos e de 0,5–1 segundo para a frase isolada lida. O segundo mecanismo ocorre na condição de acoplamento do oscilador com o fator sintático, o qual atrai os pulsos deste para posições sintaticamente pertinentes, localizadas em certos acentos lexicais das enunciações.

O efeito espacial, segundo Barbosa (2006:101) é oriundo de dois fatores: “(1) da magnitude da projeção da força sintática entre as unidades lingüísticas à esquerda e à direita da posição do pulso do oscilador acentual; (2) da magnitude da carga semântica impressa na vizinhança do pulso do oscilador acentual correspondente”.

Os pulsos do oscilador acentual determinam isomorficamente a posição dos acentos frasais. O mesmo é considerado no MDR como “consequência de um mecanismo de acentuação engatilhado pelos pulsos do oscilador acentual” (Barbosa, 2006, p. 105). Os acentos frasais, por sua vez, delimitam os chamados grupos acentuais. Empiricamente, como veremos no capítulo 3, os pulsos do oscilador acentual são detectados por picos de duração das

unidades VV ao longo do enunciado. Convém ressaltar que tais picos de duração no modelo são considerados independentemente de uma contrapartida perceptiva, pois a mesma é condicionada por outros fatores ainda não implementados, como o sistema entonacional.

As relações hierárquicas entre acentos frasais, obtidas através das magnitudes dos pulsos do oscilador acentual, geram uma métrica dinâmica, a qual poderia vir a ser estudada segundo os procedimentos inovadores dos trabalhos de Port e colegas relatados na seção 1.4.1.

Na próxima seção apresentaremos como se dá o acoplamento entre prosódia e segmentos no MDR. Basicamente, como o modelo utiliza as pautas gestuais da FAR, o componente gestual tem a função de inserir perturbações duracionais e espaciais nas unidades VVs abstratas, a fim de gerar a duração final do enunciado.

O acoplamento entre prosódia e segmentos

No MDR o acoplamento entre prosódia e segmentos se dá bidirecionalmente. Por um lado, o oscilador silábico induzido pelo oscilador acentual gera a duração abstrata dos gestos, de acordo com sua posição no grupo acentual. Vale lembrar que o grupo acentual é composto por um crescendo duracional exponencial a partir da primeira unidade VV do mesmo. Por outro lado, a pauta gestual, a partir de informações oriundas do léxico gestual, interage

com as durações abstratas do oscilador silábico especificando aspectos intrínsecos dos gestos, interação esta (bidirecional) que resulta na delimitação acústica e espacial dos gestos. A fim de visualizarmos melhor esse panorama, observemos a figura 1.15.

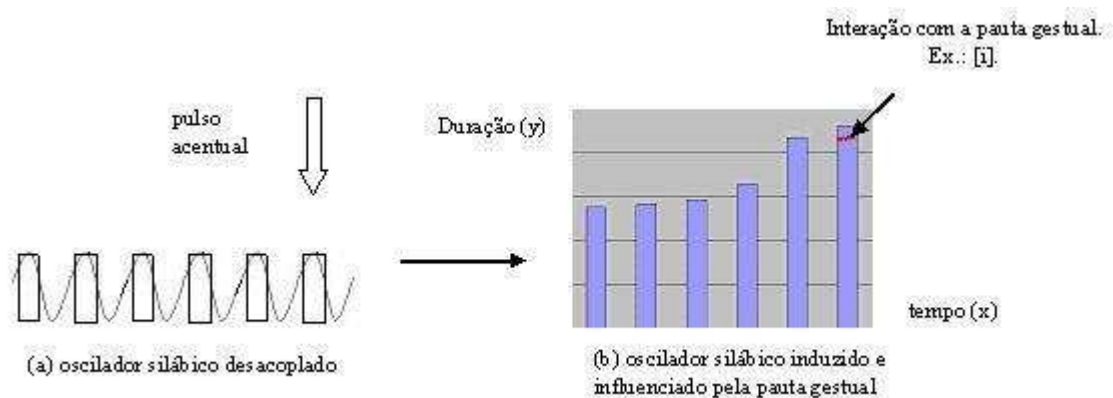


Figura 1.15: Representação de uma unidade VV abstrata (a) e de sua interação com a pauta gestual (b). Para mais detalhes vide texto.

Em 1.15(a) temos uma situação hipotética em que o oscilador silábico desacoplado possui um período de 138 ms. O mesmo, induzido pelo pulso do oscilador acentual (sexta unidade VV do oscilador silábico) reestrutura exponencialmente seus períodos (durações) através da equação $y = 113.66e^{0.1133x}$ (evolução em ms dos VVs: 138–141–146–164–217–230), que prevê um aumento gradual exponencial ao longo do grupo acentual. Convém lembrar que tal aumento exponencial foi verificado a partir de dados reais do PB (Barbosa, 2002a) e que ocorre da esquerda à direita do GA. Em 1.15(b) te-

mos o resultado do acoplamento entre os dois osciladores. Até o presente momento, as durações resultantes são abstratas, ou seja, independentes dos gestos envolvidos. A partir deste momento, a pauta gestual, através de informação do léxico gestual e da interação com o oscilador silábico induzido, especifica, fisicamente, o gesto duracional e espacialmente. Como exemplo, mostramos a duração para a constelação gestual de [i] (área pontilhada do último VV indicada pela seta na figura 1.15b) para uma frase como “você não vai cair” com os respectivos VVs [os-en-ẽwv-ajk-a-ih]. De acordo com os dados da literatura fonética, as vogais mais altas são mais curtas do que as baixas, pois, por percorrerem uma menor distância articulatória, possuem durações menores (Lehiste, 1970). Além disso um outro estudo mostrou que a duração da vogal está correlacionada à sua altura (Meyer, 1903, 1904), *apud* (Tillmann, 1994). Resultados semelhantes foram mostrados por Barbosa (2006) (vogal (média, desvio-padrão): [a] (134,31), [ɛ] (123,23), [ɔ] (130,25), [e] (98,17), [o] (108,15), [i] (69,25), [u] (89,24)). Sendo assim, como a vogal [i] é intrinsecamente curta, a duração do VV abstrato correspondente diminuirá, como mostrado pela seta em 1.15(b). Por outro lado, a presença de uma constelação gestual para [r] pouco ou nada modificará a duração do VV correspondente, dada à sua baixa variabilidade nas línguas do mundo.

Barbosa (2006) aponta alguns dos fatores que afetam a duração gestual resultante. Os fatores estudados até o momento foram a influência da TE

e de ênfase no enunciado. Os resultados mostraram que a taxa de elocução influencia todos os segmentos de um enunciado, enquanto a ênfase afeta apenas o item lexical enfatizado do mesmo.

Barbosa (*ibidem*), especifica abstratamente no MDR alguns resultados do PB obtidos em seu trabalho, os quais não foram previstos de antemão, a saber: (i) não houve efeito compensatório da unidade VV na duração de consoantes que seguem vogais com diferentes durações intrínsecas; (ii) houve efeito compensatório na duração da vogal precedendo consoantes com vozeamentos distintos; (iii) houve uma tendência à densidade gestual (número de segmentos/VV) afetar todos os gestos no interior de uma unidade VV.

Os resultados de Barbosa (2006) quanto à não-compensação de consoantes não possui contra-argumentos fortes o suficiente para descartar a unidade VV como uma unidade de planejamento prosódico mínima, visto que, no MDR, a interação entre os osciladores silábicos e acentuais não leva em conta a duração intrínseca dos segmentos. Além disso, (i) trabalhos como os de Nooteboom (1972) propõem que “o programa rítmico do enunciado não leva em conta as diferenças de duração intrínseca entre as vogais” (Barbosa, 2006), *apud* (Nooteboom, 1972), e (ii) a sobreposição de gestos articulatórios pode estar mascarando as durações acústicas finais das vogais e consoantes. Lembremos que os resultados de Barbosa (2006) se baseiam apenas em dados acústicos do PB. Sendo assim, o autor brasileiro propõe algumas hipóteses,

a serem confirmadas, em relação às pautas gestuais, para dar conta dos resultados acústicos encontrados em sua pesquisa.

A primeira hipótese estipula que a duração abstrata das unidades VV seja idêntica para todo e qualquer segmento. No entanto, no caso das vogais (figura 1.16), a especificação de diferentes graus de magnitude para as vogais nas pautas gestuais (eixo vertical), resultará, no âmbito acústico, em uma duração maior para [a] do que para [i] (cf. a hipótese “the farther-the-longer” de Lehiste, 1970).

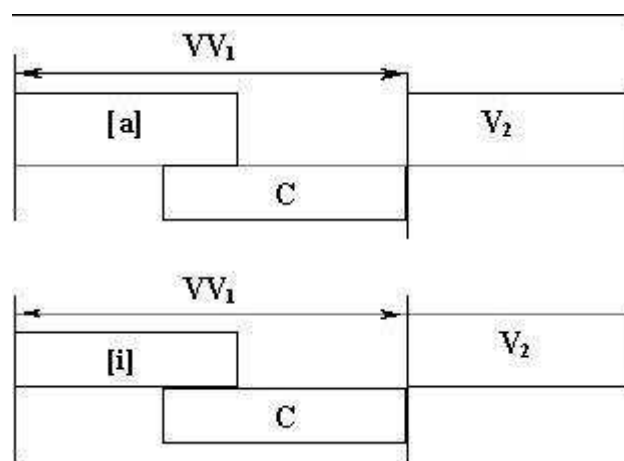


Figura 1.16: Especificação abstrata das pautas gestuais esquematizadas para as vogais (Barbosa, 2006, p. 347).

Em seus estudos, Barbosa (2006) verificou que, especificamente, há uma compensação da vogal precedendo consoantes vozeadas e desvozeadas. Diante de consoantes vozeadas, a duração das vogais é maior do que diante das desvozeadas. Como sabemos, as consoantes desvozeadas são, em média,

maiores do que as vozeadas. Em vez de estipular abstratamente tal diferença duracional, o pesquisador propõe que tais consoantes sejam definidas identicamente nos VVs abstratos (figura 1.17). A diferenciação duracional só será efetivada após a interação com a pauta gestual, que especifica que o gesto glotal seja ativo (aberto) mais cedo em [p] do que [b], gerando, como consequência, uma duração maior para [p] e a vogal que precede [p].

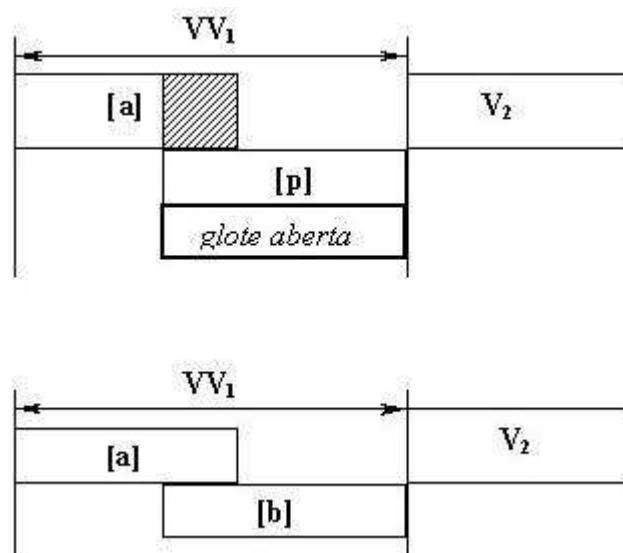


Figura 1.17: Representação abstrata das consoantes [p] e [b] e suas respectivas vogais precedentes (Barbosa, 2006, p. 347).

Em relação ao fato de uma maior densidade gestual, salvaguardada sua variabilidade intrínseca, afetar todos os gestos de uma unidade VV, Barbosa (2006) propõe uma maior sobreposição desses gestos no interior de uma unidade VV, resultando, fisicamente, em uma duração menor para cada um dos gestos que compõem o VV (figura 1.18).

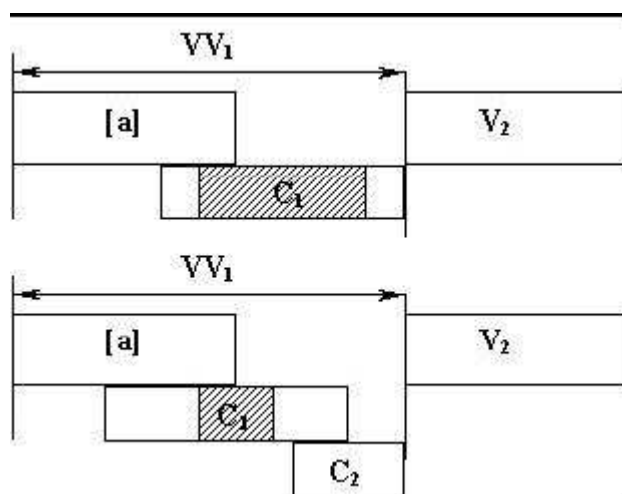


Figura 1.18: Representação abstrata da diminuição das durações segmentais com o aumento da densidade gestual.

Concluindo, esta seção termina a explanação do arcabouço teórico geral no qual nossa tese está inserida. Nos próximos capítulos apresentaremos noções gerais sobre o ritmo e resultados acústicos e articulatórios provenientes de nossos trabalhos concernentes a reestruturações rítmicas nos âmbitos da sentença e da palavra. Em ambos os casos, tais reestruturações foram resultado de perturbações dinâmicas no ritmo da fala através da variação da TE.

Especificamente, trabalharemos, em nossa tese, os limites do MDR a partir do estudo descritivo da estruturação rítmica com o aumento da taxa de elocução, a qual atuará como perturbador do sistema dinâmico da fala nos âmbitos prosódico e lexical. Tal modelo nos permitirá simular as seguintes possibilidades em relação à variação da TE: (1) interação entre o oscilador

acentual e silábico para a especificação da posição das batidas acentuais; (2) aumento da frequência do oscilador silábico, que pode ser efetuado variando-se T_0 ; (3) mudança nos parâmetros α e β , responsáveis por variações finas com conseqüências a longo termo nos enunciados; e (4) interação do modelo de osciladores acoplados com a pauta gestual.

Segunda parte: Reestruturações rítmicas prosódicas

Capítulo 2

Ritmo: definição e aspectos teóricos

A palavra ritmo (do latim *rhythmus*) vem do grego *ρυθμος*, que é derivada de *ρειν* (fluir). Devido à sua etimologia, muitos dicionários a exemplificam citando o fluxo das marés. No entanto, tal palavra nunca foi usada com tal sentido. Uma definição mais precisa da noção de ritmo foi utilizada, na filosofia iônica (fundada em 600 a.C. por Thales de Mileto), por Leucipo e Demócrito. Para eles o ritmo tinha o sentido de “forma”, juntamente com “ordem” e “posição”, um dos três critérios de diferenciação das coisas (Benveniste, 1951). Somente a partir de Platão a palavra adquire uma conotação próxima ao sentido atual empregado pelos estudiosos do ritmo. Em suas Leis (665a), o filósofo o definiu como “a ordem no movimento”. A inovação

de Platão consiste na atribuição da “forma” de seus predecessores aos movimentos humanos, como a dança, o canto, o trabalho. Nesse panorama, Fraisse (1974) afirma ser o conceito de ritmo oriundo da organização do movimento humano. Sendo assim, discutiremos aqui, com base nesse trabalho, alguns aspectos do ritmo que tenham relação com a língua. Antes, contudo, faremos uma breve excursão pelos ritmos biológicos, depois comentaremos sobre os ritmos espontâneos e, por fim, sobre a formação de grupos rítmicos.

2.1 Ritmos biológicos

Os ritmos biológicos não podem ser deixados de lado em uma introdução geral sobre o ritmo, pois os mesmos afetam todos os seres vivos, da célula ao organismo como um todo. Tais ritmos, segundo Fraisse (1974:15), têm uma tripla importância do ponto de vista da psicologia: (i) alguns deles são bem próximos aos da atividade humana e são facilmente percebidos, como o ritmo do coração e da respiração; (ii) há, geralmente, analogias sugestivas entre os ritmos humanos e biológicos sob o ponto de vista de suas estruturas e mecanismos; (iii) os ritmos biológicos têm uma grande repercussão na vida cotidiana.

A atividade rítmica é uma propriedade essencial de todos os seres vivos. Sendo assim, a ciência deve ser capaz de explicar uma questão clássica: são

os ritmos biológicos adquiridos ou inatos¹? Fraisse (1974:17) afirma que a ciência é incapaz de responder inteiramente a essas questões, mas que pode descrevê-las. O pesquisador, pois, apresenta os diversos tipos de classes rítmicas (pp. 17-18): a) ritmos espontâneos. Como o próprio nome diz, tais ritmos ocorrem espontaneamente. Ex.: ritmo cardíaco, ritmo do metabolismo celular; b) ritmos desencadeados. Diferenciam-se do espontâneo, pois necessitam de uma excitação não-periódica para acioná-lo. Ex.: um feijão cultivado em luminosidade constante necessita ser colocado uma única vez na escuridão para que ocorram os movimentos oscilatórios de suas folhas; c) ritmos induzidos. Tais ritmos possuem um movimento oscilatório espontâneo, mas o seu período pode se acoplar ao período de uma força externa. Ex.: o ritmo alfa das ondas cerebrais pode se sincronizar a um ritmo luminoso externo. Com o fim do indutor, o ritmo alfa retorna à sua frequência inicial; d) ritmos adquiridos: em tais ritmos, o período de resposta é provocado por um período externo. Tais fenômenos remontam a Skinner e seus experimentos comportamentais de reforço. Estudos comprovam que os animais podem adquirir ritmos bastante complexos. Se condicionarmos, por exemplo, uma pomba a um ritmo como: 15 segundos de nutrição, 15 segundos de pausa, 15 segundos de nutrição, 90 segundos de pausa, os animais continuam a mani-

¹Para a linguagem, conforme seção 1.4.1 (pp. 58-59) desta tese, a discussão mais coerente se pauta em torno dos universais rítmicos *versus* específicos a uma língua.

festar tais ciclos periódicos mesmo após o fim do experimento (Popov, 1950; Ferster & Skinner, 1957; *apud* Fraisse, 1974).

Os ritmos ainda se subdividem em três categorias principais: os lentos, os rápidos e os circadianos. Todos três estão presentes nos seres humanos. Os ritmos lentos possuem um período superior a 24 horas. Ex.: o ciclo ovariano da mulher possui um período de 28 dias, coincidentemente próximo a uma revolução da lua (29 dias). Os ritmos rápidos possuem um período compreendido de frações do segundo a vinte horas. Ex.: ritmos do encéfalo (ondas delta: 1–5 Hz; teta: 4–7 Hz; alfa: 8–13 Hz; beta: 14–30 Hz), do coração e da respiração. O ritmo respiratório é particularmente interessante para os lingüistas, pois Stetson (1988) definiu a sílaba a partir de pulsos torácicos, ou seja, ciclos respiratórios². Por fim, os ritmos circadianos possuem um período de cerca de 24 horas. Ex.: ritmos da temperatura central, da eliminação de sódio e potássio.

Todas as características rítmicas mencionadas acima estão presentes nos organismos vivos em geral. No entanto, o homem se distingue dos outros animais, pois, apesar de possuir características comuns, é capaz de modificar os estímulos externos através de roupas adequadas, moradia e controle da luz.

A abundância de ritmos biológicos na natureza nos faz supor a existência

²No entanto tal hipótese se mostrou infrutífera.

de um “relógio biológico”, o qual controlaria a duração das fases dos ciclos rítmicos, bem como sua organização (Fraisse, 1974, p.37). Segundo o pesquisador francês, no entanto, os mecanismos desse relógio se dividiriam em dois ramos: 1) regularizador, o qual daria conta dos aspectos periódicos de determinados ritmos, como os circadianos; 2) acoplador, o qual lidaria com o acoplamento entre os aspectos regulares do ritmo com aspectos estruturantes. Ex.: regulação da fotossíntese e floração das plantas.

Os mecanismos de um “relógio biológico” apontado por Fraisse estão presentes, como vimos, no MDR, no qual os osciladores silábico e acentual funcionam como elementos periódicos e seu acoplamento gera a estruturação rítmica do enunciado. No modelo dinâmico de Byrd & Saltzman (2003) também há elementos de um “relógio biológico”, pois os autores inclusive mencionam um “relógio interno” na explicação de alongamentos em fronteiras prosódicas. No entanto, nesse trabalho não há uma explicação muito clara de como tal “relógio” atuaria, biologicamente, na produção da fala.

2.2 Ritmos espontâneos

Os ritmos espontâneos, por sua vez, são evidências de um mecanismo controlador dos ciclos oscilatórios nos animais. Especificamente, para os objetivos de nossa tese, os ritmos espontâneos motores, como a sucção e a mastigação

são os mais interessantes no quadro de um MDR que faz uso dos conceitos da FAR.

Segundo Fraisse (1974:47), os períodos de sucção e mastigação, apesar de variações individuais entre os recém-nascidos, são, em média da ordem de grandeza de 600–1200 ms³. Resultados esses intrigantes se se considerar a aquisição da linguagem a partir de gestos primitivos (Browman & Goldstein, 1989, 1990b, pois tal faixa é similar à de outras atividades motoras, como a batida espontânea de um dedo sobre uma mesa. Segundo Fraisse (1982:153), uma duração de 600 ms é a mais representativa para tal tarefa. Surpreendentemente, em várias línguas foram encontrados períodos semelhantes em trabalhos que datam desde o começo do século XX. Wallin (1911), *apud* Fraisse (1974) mostrou que o intervalo de um acento a um outro em inglês é, em média, de 550 ms. Em Barbosa, Arantes, Meireles & Vieira (2005), o intervalo entre os inícios de vogais repetidas espontaneamente, em sílabas CV, continuamente, foi, em média, 556 ms para o PB. Nord et al (1990) mostraram que, na sincronização dos dedos com as sílabas fortes de poemas, o intervalo entre batidas foi, em média, 500 ms em sueco. Eriksson (1991), também para o sueco, relata que o intervalo entre acentos lexicais primários foi, em média, 580 ms. Fant & Kruckenberg (1989) relatam novamente

³Além disso, estudos como os de Sansavini et al. (1997), através de medidas de taxa de sucção, mostra que mesmo recém-nascidos são capazes de distinguir diferentes padrões rítmicos da fala (vide também Konopczynski, 1995, que observou semelhanças na aquisição do ritmo em línguas consideradas de padrões rítmicos distintos.)

para o sueco uma duração média dos GA de 548 ms e Dauer (1983), 530 ms para o inglês americano. Outro dado do PB mostra que a duração média dos GA, em frases isoladas lidas por um locutor pernambucano, é de 549 ms (Barbosa, 2006:182). Esses dados mostram a similaridade de tarefas motoras repetitivas com padrões da fala, bem como a similaridade entre repetição de acentos frasais na frase isolada ou de acentos melódicos (*pitch accents*) em inglês.

Outra característica a ser mencionada a respeito dos ritmos espontâneos é a variação individual em tarefas repetitivas. Tal característica é extremamente importante na comparação com dados lingüísticos em falantes distintos, a fim de se respeitar a condição *ceteris paribus* do experimento. Dessa forma, devemos ser bastante cuidadosos na comparação da TE entre diferentes dialetos e línguas. Os resultados podem ser mascarados devido a características individuais lentas ou rápidas dos falantes. Para se determinar o tempo individual ou natural dos falantes, segundo Fraisse (1974:48), basta fazê-los bater o dedo sobre uma mesa em uma taxa confortável. A média desses intervalos será considerada seu tempo espontâneo⁴. A gama de indivíduos lentos a muito rápidos, na tarefa de bater sobre a mesa, varia de 450 a 1500

⁴Em Barbosa, Arantes, Meireles & Vieira (ibidem) utilizamos um princípio semelhante para saber a partir de onde variaríamos a taxa do metrônomo. Para isso, utilizamos a repetição livre de uma sílaba como [pa]. A eliciação espontânea é utilizada nos experimentos relatados nesta tese, pois, na aquisição de nossos dados, sempre aceleramos ou lentificamos a taxa a partir da taxa espontânea do informante.

ms. Novamente, tais intervalos são encontrados nos intervalos entre GA em PB, em que a variação na leitura de um texto de dois parágrafos foi de 889 a 1518 ms (Barbosa, 2006:175). Outro resultado surpreendente em relação ao ritmo é uma relação entre esses tempos espontâneos com os batimentos do coração, da ordem do segundo, a ponto de o poeta Paul Claudel afirmar que “l’iambe fondamental est le battement du coeur”.

O acoplamento dos ritmos motores com outros tipos de ritmo como a música é um fato bastante conhecido. Qual de nós ao ouvirmos uma música com uma pulsação marcante, não começamos, espontaneamente, a bater o pé no chão em sincronia com o ritmo da música, independentemente de gostarmos ou não da mesma? Essa peculiaridade dos ritmos aparece notadamente em crianças pequenas, que, intuitivamente, começam a dançar ao ouvir uma música movimentada. Uma das propriedades do ritmo que nos permite tal comportamento é a nossa capacidade cognitiva de prever, através de experiências prévias, quando um evento futuro ocorrerá, propriedade exarcebada pelos bons músicos, que são capazes de manter um ritmo bastante regular, mesmo sem a presença de um metrônomo. Tal capacidade de sincronização rítmica é a base para uma noção muito usada na música, na metrificação poética e até na prosódia — o pé. Por analogia aos batimentos dos pés com uma música, há duas partes em um pé: uma correspondente aos tempos fortes — batida do pé — e outra aos tempos fracos — levantamento do pé. Dentre os

tipos de pés mais comuns em português estão o dáctilo (— U U), o anapesto (U U —), o iambo (U —) e o troqueu (— U). Essas propriedades naturais do ritmo, entretanto, são inibidas ao longo do caminho até a maturidade. Segundo Fraisse (1974:65) “les synchronisations spontanées apparaissent moins fréquemment chez l’adult occidental que chez l’enfant parce que les critères culturels de notre civilisation nous invitent à cette inhibition”.

2.3 Formação de grupos rítmicos

Deixando de lado as características principais dos ritmos, passemos a explicar os processos de ritmização, ou seja, de formação de grupos rítmicos. Há, basicamente, dois tipos de ritmização: (i) a subjetiva e (ii) a objetiva. A ritmização subjetiva corresponde à estruturação de unidades rítmicas estritamente regulares e idênticas. Por exemplo, ao ouvirmos uma seqüência de tons regularmente espaçados e de mesma intensidade, nossa audição cria uma ilusão de que determinados tons são mais longos que os outros ou que alguns deles são mais fortes ou agudos. Por outro lado, a ritmização objetiva elabora grupos rítmicos a partir de medidas duracionais e/ou tonais presentes objetivamente no sinal de fala. Dessa forma, pode-se supor que a criação de agrupamentos rítmicos é uma característica fundamental de nossa percepção do mundo. Quais seriam, pois, as unidades responsáveis pela delimitação de

grupos rítmicos?

Fraisse (1974:77) diz ser o ritmo composto por “*enchaînement de structures à intervalles isochrones*”, aos quais são associados “*patterns de mouvements simples (pendulaires) ou complexes*”. Sendo assim, temos, neste processo, as duas características indissociáveis do ritmo: o agrupamento perceptivo (ou aspecto estruturante) e a repetição dos grupos associados a eventos motores (ou aspecto repetitivo).

Os limites para os agrupamentos rítmicos apontados por Fraisse (1974:78–9) são: (i) temporais. Os limites inferiores são compreendidos entre 150–200 ms (ordem de grandeza de unidades VV) e os limites superiores entre 1500–2000 ms (ordem de grandeza de GA); (ii) número de elementos constituintes. A tendência geral observada é a de quanto maior o número de constituintes por unidade rítmica, menor o intervalo entre seus constituintes, o que é semelhante ao estudo reportado anteriormente, em que um grande número de segmentos por unidade VV resulta em durações menores dos segmentos devido à coarticulação.

Uma característica dos agrupamentos rítmicos relacionada à ritmização subjetiva é a tendência à acentuação de unidades rítmicas ao fim ou ao início do enunciado, o que explicaria as tendências prosódicas nas línguas do mundo de terem cabeça à esquerda ou à direita. Os estudos de Fraisse (1974) mostram que uma estrutura do tipo F(orte)-f(fraca)- f(fraca)- f(fraca) pode

ser percebida como Ffff ou fffF, mas nunca como fFff. Mais interessante é o caso em que uma sequência como FFFf é interpretada como fffF ou Ffff.

Vale a pena ressaltar que os estudos de Fraisse sobre os processos de ritmização guardam uma semelhança muito grande com os aspectos dinâmicos do ritmo discutidos anteriormente, a ponto de, estando à frente de seu tempo, dizer: “il y a forme puisque chaque structure rythmique correspond à une organisation dynamique. Tout intervalle dépend des autres et, si on change la durée de l’un d’eux, il se produit une réorganisation de l’ensemble” (Fraisse, 1974, p. 98). Uma reorganização de tal tipo é apresentada pelo pesquisador em um experimento em que uma sequência rítmica constituída por • • • • •, ou seja, um intervalo curto curtos seguido por dois longos se reestrutura nas taxas rápidas para • • • • •, dois intervalos curtos entremeados por um longo. Por outro lado, taxas muito lentas fizeram romper a percepção de intervalos longos e curtos devido ao longo espaçamento entre os intervalos (6570 ms). Lembremos que os limites entre os agrupamentos rítmicos devem ser respeitados para tornar possível a percepção rítmica. Tais tipos de reestruturações rítmicas influenciadas pelo aumento da taxa de elocução serão o objeto de estudo dos experimentos dessa tese.

Nos capítulos seguintes serão apresentadas evidências acústicas e articulatórias do efeito de perturbações dinâmicas na reorganização rítmica da fala em seus aspectos prosódicos, principalmente seus efeitos na sincronia entre os

osciladores silábico e acentual. Para isto serão discutidos três experimentos:

1) estudo acústico preliminar; 2) estudo acústico complementar; 3) estudo articulatorio.

Capítulo 3

Estudos acústicos

Estudos acústicos anteriores investigaram como a variação da TE atua na reorganização da estrutura prosódica da língua. Pasdeloup et al. (2006), à luz do programa de pesquisa da Gestalt (cf. Guillaume, 1979; Köhler, 1929), propuseram um modelo rítmico (cf. Pasdeloup, 1990, 1992) que analisa sílabas átonas como fundo da cena prosódica e sílabas tônicas como figuras que emergem desse fundo. Essa hipótese é corroborada pelo fato de sílabas átonas serem mais resistentes à variação da TE do que sílabas tônicas (cf. Pasdeloup, 2004, 2005; Pasdeloup et al., 2006). De acordo com Pasdeloup et al. (2006), “the unstressed syllable duration increases by an average of 35% between the fast and slow rate conditions, while the stressed syllable duration increases by an average of 85% between those conditions”. Convém ressaltar, entretanto, que, na percepção da TE, as sílabas átonas agem como

referência. Esses resultados revelam serem as sílabas afetadas de maneira distinta nas taxas rápida e lenta. Curiosamente, para o estilo de fala estudado, Pasdeloup (2005) relata que a duração das sílabas átonas na taxa lenta (158 ms) é, em média, menor do que as sílabas tônicas na taxa rápida (194 ms), o que mostra, apesar de as sílabas tônicas serem mais compressíveis/elásticas do que as átonas, ser a diferenciação entre sílabas átonas e tônicas ainda preservada na enunciação. As sílabas tônicas se reduzem estatisticamente mais, em termos duracionais, do que as átonas, mas não a ponto de se tornarem indistintas entre as taxas rápida e lenta.

De maneira semelhante, Barbosa (1996b, 2001, 2002a, 2002b) analisa as unidades VVs átonas abstratas como períodos de um oscilador silábico, ao qual é atribuída a silabidade lingüística. O oscilador acentual acoplado ao silábico age sobre o mesmo na atribuição dos acentos frasais. Como pode ser inferido, ao contrário dos estudos de Valérie Pasdeloup, as sílabas acentuadas e átonas dos enunciados são meras conseqüências da interação dos dois osciladores acoplados. A fim de suportar a idéia de osciladores acoplados na produção da fala, o autor realizou diversos experimentos com dados do português brasileiro.

Barbosa (2006:182) mostra, para um informante pernambucano, falante nativo do PB, com idade entre 20 e 30 anos, gravado em três TE, que a duração média dos grupos acentuais é de 549 ms para a leitura de frases

isoladas na taxa espontânea. Como vimos no capítulo 2, valores similares também foram achados para línguas distintas como o sueco (Fant & Kruckenberg, 1989, 548 ms) e o inglês (Dauer, 1983, 540 ms). Outro dado do PB mostra que a duração média dos GA, em um experimento com repetição espontânea de sílabas isoladas (ex. 10 repetições da sílaba [pa]), sincronizadas com batidas de um metrônomo em diversos andamentos, é de 556 ms (Barbosa, Arantes, Meireles & Vieira, 2005). De acordo com Barbosa, a similaridade desses dados suporta a idéia de um oscilador acentual que ocorre abstratamente a intervalos regulares. Apesar de serem encontradas durações médias de grupos acentuais próximas em diversas línguas, o mesmo não foi encontrado para a leitura de um texto longo em PB, conforme mostraremos no estudo preliminar relatado na próxima seção.

Barbosa também mostrou em seu trabalho reestruturações rítmicas devido à variação da TE em francês (1994, pp. 121-3) e em PB (2006, pp. 249-54). Uma reestruturação rítmica é considerada aqui como uma reorganização do GA ao longo do enunciado ao se variar a TE, que tende a gerar um número menor de GA nas taxas rápidas como veremos nos estudos acústicos relatados nesta seção. De acordo com o pesquisador, algumas fronteiras menores são apagadas devido ao aumento da taxa de elocução, pois não é possível desacelerar em algumas fronteiras para assinalar acento frasal e, ao mesmo tempo, manter um passo acelerado pela taxa ao longo de todo o enun-

ciado (pp. 248-9). A fim de avaliar probabilisticamente onde ocorrerão os GA, Barbosa (2006) propôs uma abordagem que atribui acentos frasais de acordo com restrições sintáticas e/ou de tamanho do GA, cujo acoplamento é parametrizado pelo parâmetro r_p , grau de acoplamento entre sintaxe e produção no modelo. Através da aplicação dessa abordagem a um texto extenso, o autor concluiu que, em geral, palavras lexicais (como substantivos e adjetivos) atraem o acento frasal, o verbo é neutro e palavras funcionais (como artigos, preposições e conjunções) o repelem.

Em um outro estudo, Wightman et al. (1992) investigou durações segmentais em torno de fronteiras frasais, a fim de observar quais unidades estariam sendo alongadas. Como no trabalho de Barbosa e em nossa tese, os autores utilizaram a duração normalizada para se evitar efeitos de duração intrínseca. Embora não tenham sido utilizadas TE variadas, mudanças de duração devido à variação da TE foram normalizadas por um fator de correção α (p. 1712-13). As análises acústicas revelaram que “lengthening associated with the perceived size of the boundary is confined to the rhyme (vowel nucleus and any coda consonants) of the final syllable before the boundary” (p. 1714). Esse resultado é interessante, pois a rima silábica é a unidade que encabeça a unidade VV nos trabalhos de Barbosa¹, sendo a sub-unidade que

¹Entretanto, como vimos no capítulo 1 (vide fig. 1.14), a unidade VV não se limita à rima silábica

mais se alonga devido à influência de uma proeminência frasal. No MDR, entretanto, o alongamento da unidade VV é uma consequência de um aumento de duração não-linear a partir do início de um GA, enquanto em Wightman et al. (ibidem) “no other part of the foot appears to lengthen in a way related to the perceived size of the boundary” (p. 1714). Como no modelo prosódico de Byrd & Saltzman (2003), o alongamento só ocorre localmente na fronteira frasal. Um outro resultado desse trabalho, que se assemelha ao reinício de período dos VV no MDR, é a ausência de alongamento para os fones seguintes à fronteira prosódica.

Barbosa também mostrou em Albano et al. (1998) que a estruturação rítmica do PB é influenciada por fatores extralingüísticos como a TE, conforme podemos observar na figura 3.1.

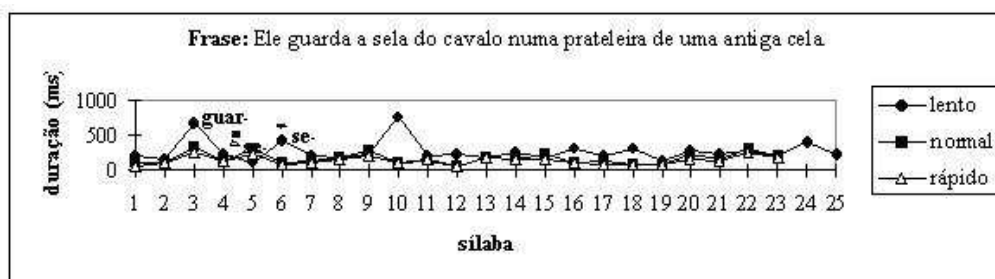


Figura 3.1: Contornos duracionais para a frase “Ele guarda a sela do cavalo numa prateleira de uma antiga cela”.

Nesta figura está representada uma frase gravada em três taxas distintas: lenta, normal e rápida. A estrutura rítmica desse enunciado foi analisada

através do contorno duracional das sílabas, medidas em milissegundos. Observamos nitidamente na figura 3.1 a influência da taxa de elocução nas reestruturações rítmicas, considerando-se padrões duracionais como abordagem descritiva. Como exemplo, examinemos a seqüência “ele guarda” nas três taxas, nas quatro posições iniciais. Como podemos ver, há uma mudança gradual de dois grupos acentuais na taxa lenta (cf. alternância duracional) para um grupo acentual na taxa rápida (cf. crescendo duracional monotônico). Há, aqui, uma variação gradativa do ritmo relacionada a aspectos quantitativos, i.e., fonéticos, da fala (resultado similar foi encontrado para o francês em Barbosa, 1994, pp. 121-3). Considerando as possibilidades de combinação para locutores distintos e para situações variadas, não há possibilidade de segmentação discreta dessa variação contínua de padrões de proeminência lingüisticamente pertinentes (Albano et al., 1998, p. 137).

Nos estudos acústicos descritos a seguir, investigaremos o efeito da variação da TE nas reorganizações rítmicas da fala em PB com dois *corpora* distintos. Os objetivos desses trabalhos são:

1. Descrever quantitativamente os processos fônicos relacionados com a variação da taxa de elocução, começando pela descrição da reorganização dos padrões duracionais, contribuindo, assim, para a constituição de uma metodologia para a análise de dados da variação prosódica;

2. Contribuir para um melhor entendimento do papel da taxa de elocução na constituição do ritmo da fala;
3. Contribuir para um melhor entendimento da acentuação frasal em português brasileiro;
4. Discutir à luz da TSD os resultados encontrados.

3.1 Estudo preliminar

Utilizando-se diferentes TE como perturbadores do sistema, procedimento clássico na teoria dos sistemas dinâmicos para se descobrir parâmetros subjacentes a um sistema, realizamos um estudo, à luz do MDR, a fim de se avaliar reestruturações rítmicas da fala com o aumento da TE a partir da interação entre os osciladores silábico e acentual.

3.1.1 Metodologia

Para verificar a estrutura rítmica dos enunciados em três taxas de elocução: lenta, normal e rápida, utilizamos o seguinte *corpus* (Lobato, 1920, p. 18):

Em seguida apareceu um papagaio real que tinha fama de orador.

Subiu a tribuna de um poleiro de ouro e fez um belo discurso a respeito da arte de falar. Nesse discurso provou que os homens

tinham aprendido a falar com os papagaios, e não os papagaios com os homens, como diz a ciência destes. Uma chuva de palmas acolheu suas palavras.

O mesmo não aconteceu, porém, com a poetisa Lagartixa, que principiou a recitar uma longa poesia e engasgou no meio, acabando o recitativo em choro e faniquito. Para destruir essa má impressão vieram três vagalumes mágicos que fizeram várias sortes, sendo muito apreciada a sorte de comer fogo.

Os quatro informantes gravados (AC, AP, DP, FA) são falantes nativos do PB, do sexo masculino, paulistas, sem qualquer deficiência fonatória ou audiológica. A gravação foi realizada em uma cabine acusticamente tratada do Laboratório de Fonética e Psicolinguística(Lafape)/IEL/Unicamp e os dados foram digitalizados a 22,05 kHz. O texto foi lido pelos informantes em três taxas de elocução distintas. Diferentemente de nossa dissertação de mestrado (Meireles, 2001), em que usamos um metrônomo na eliciação das taxas, as três taxas foram livremente escolhidas. Optamos, em nossa tese, pela utilização deste método, pois, se admitirmos por hipótese que o modelo de osciladores acoplados é efetivamente adequado à análise da fala, o acoplamento de seus componentes poderia interferir na produção linguística.

Por exemplo, utilizando-se um metrônomo, os falantes poderiam ora alinhar as batidas do mesmo com o oscilador silábico, ora com o acentual. Além disso, a utilização do metrônomo, como observamos em Meireles (2001), produz, muitas vezes, frases artificiais, diferentes de uma produção mais livre. As instruções dadas aos informantes, pois, para aquisição das taxas foram: 1) fale o mais lentamente possível preservando a estrutura prosódica da frase; 2) fale em uma taxa confortável; 3) fale o mais rapidamente possível sem cometer distorções. Lembramos que, antes da gravação, os falantes leram algumas frases extras para verificarmos se as instruções foram compreendidas pelos mesmos. Quaisquer dúvidas quanto às instruções foram prontamente esclarecidas. A ordem de gravação das taxas foi: 1) normal, 2) lenta e 3) rápida. Com estas instruções obtivemos as taxas lenta, normal e rápida pretendidas.

Utilizando o software livre Praat de Paul Boersma & David Weenink (<http://www.praat.org>), procedemos às tarefas metodológicas descritas abaixo, de acordo com Barbosa (2006:169-70) (três primeiros itens), apresentadas sinteticamente na figura 3.3:

1. Segmentação do texto oral em unidades VV e cálculo no Praat das respectivas durações. Por exemplo, o trecho “em seguid(a)” foi segmentado como: /eNs/, /eg/ e /id/, e depois calculadas suas durações

(vide figura 3.2);

2. Utilização da análise estatística z-score, com o intuito de eliminar os efeitos de duração intrínseca dos segmentos no cômputo das unidades VV (cf. Barbosa, 1996a; Campbell & Isard, 1991);

3. Suavização da evolução do z-score dessas unidades através da fórmula:

$$z_{suav}^i = \frac{5.z^i + 3.z^{i-1} + 3.z^{i+1} + 1.z^{i-2} + 1.z^{i+2}}{13}, \text{ a fim de minimizar os efeitos de}$$

oscilação local (acentos lexicais não marcados frasalmente e boa parte de efeitos de duração intrínseca) e determinar os limites dos grupos acentuais através dos pontos de máximo da curva de duração suavizada;

4. Cômputo da duração dos grupos acentuais e do número de unidades VV realizadas neles contidos para a comparação das distribuições entre duas taxas de elocução, lenta e rápida, para cada locutor (distribuição da duração e distribuição do número de unidades VV);

5. Realização de análises de significância: a) testes-t (para uma distribuição gaussiana): utilizados na avaliação das diferenças entre as médias da duração dos GA e unidades VV nas taxas lenta e rápida²; b) testes-F: utilizados na avaliação da igualdade de variâncias da duração dos GA e unidades VV nas taxas lenta e rápida; e 3) Mann-Whitney U

²Para este estudo, como ainda estávamos em fase investigativa, optamos por trabalhar apenas com dois extremos de TE: lenta e rápida

(para uma distribuição não-gaussiana) de variáveis independentes: utilizado na avaliação das diferenças nas médias e variâncias dos VV/GA nas taxas lenta e rápida.

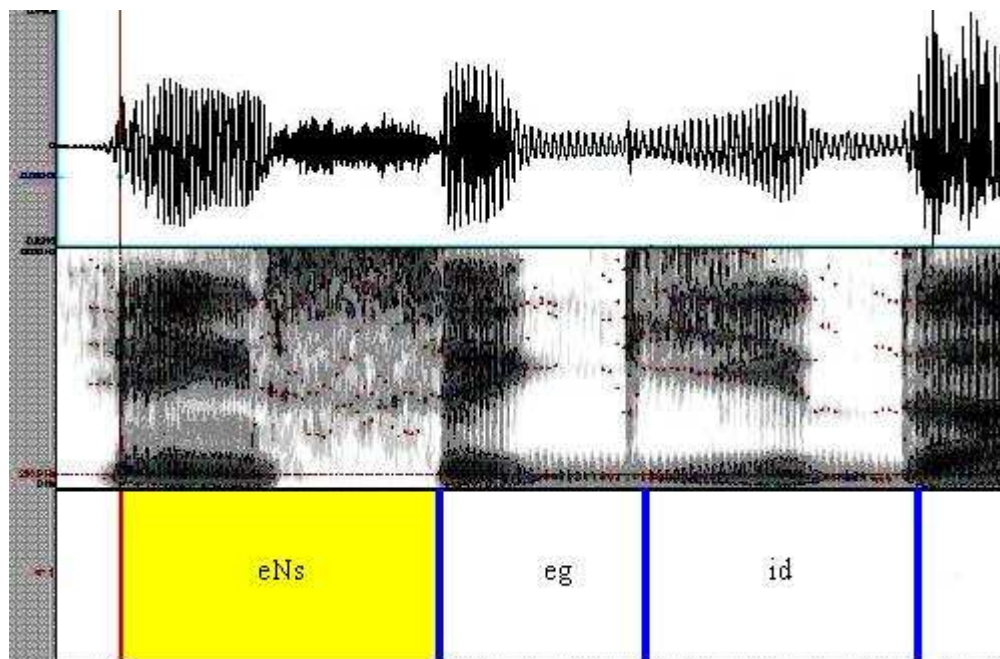


Figura 3.2: Exemplo de segmentação do texto em unidades VV. Trecho “Em seguida”, composto pelos VVs /eNs/, /eg/ e /id/. Informante: AP; taxa: normal.

VVrp	Duração	Z-score	Z-score suavizado	Derivada	Teste de máximos	Dur GA	VV/GA
			{5*C5+3*C6+3*C4+C3+C7}*13				
ês	0,181589	-1,241303	0	D5-D4	"SE(E(E30<0,E29>=0),1,0)"		
eg	0,107383	-1,400434	0				
id	0,085384	-1,528621	-1,370526757	-1,370527	0	0,374356	3
ap	0,14691	-1,067044	-1,257790143	0,1127366	1		
ar	0,07785	-1,530002	-1,35803568	-0,100246	0		
es	0,187091	-0,439747	-1,635588386	-0,277553	0		
eup	0,140004	-3,955461	-2,174070187	-0,538482	0		
sp	0,11803	-1,540484	-1,746213842	0,4278563	0		
ag	0,11803	-1,014915	-0,963753854	0,78246	0		
ajwã	0,277978	0,3525138	-0,211851419	0,7519024	0		
awk	0,338648	0,0651473	-0,049593103	0,1622583	1	1,404543	8

Figura 3.3: Exemplo dos procedimentos metodológicos utilizados em nossos experimentos. Falante: AP, taxa: rápida, trecho: “em seguida apareceu um papagaio real”. A primeira coluna representa o VV respectivo. A segunda, a duração bruta do VV em segundos. A terceira, o z-score da duração bruta do VV. A quarta, o z-score suavizado. A quinta, a derivada discreta do z-score suavizado, ou seja, a diferença entre um VV atual e o anterior. A sexta, o teste de máximos, que, pela fórmula lógica apresentada, atribui as proeminências acentuais do enunciado. A fórmula atribui tais proeminências da seguinte forma: se o VV atual for menor do que 0 e o VV anterior maior (ou igual) a 0, será atribuído o valor 1 (proeminência acentual); do contrário será atribuído 0 (nenhuma proeminência encontrada). Por sua vez, a sétima coluna soma a duração em segundos dos VVs presentes em um GA e a oitava, o número de VVs por GA.

Convém salientar que, para esse experimento, todas as segmentações e análises acústicas foram feitas, em sua maior parte, manualmente, excetuando-se a extração da duração das unidades VVs, na qual foi utilizado um *script*, Duração.psc, por mim elaborado, para extração dos VVs pré-etiquetados no Praat. Tal *script* encontra-se no apêndice A.3. Além desse *script* foram utilizadas fórmulas do programa *Excel* da *Microsoft*. No que se refere à detecção dos grupos acentuais, através do “Teste de máximos” (cf. figura 3.3), sempre que tal teste atribuíu um acento frasal a sílabas átonas, como podemos notar

nos VVs amarelos da figura 3.3, retornávamos ao texto e observávamos, de oitiva, em qual posição tônica deveriam, de fato, ocorrer tais acentos e os modificávamos nas planilhas eletrônicas (VVs marcados em negrito na figura 3.3). Lembremos aqui que o MDR atribui batidas do oscilador acentual somente a sílabas acentuadas das enunciações, visto que tal modelo ainda não possui um nível lingüístico semântico, que poderia explicar casos de atribuição de acentos a unidades átonas por questões de ênfase, como na seguinte “correção” gramatical hipotética: “Meu filho não diga “vou NA missa”, mas “vou À missa””, com ênfase nos sintagmas preposicionais “na” e “à”.

3.1.2 Análises e Resultados

As hipóteses iniciais, na ocorrência de reestruturações rítmicas, a serem investigadas nesse estudo foram: (i) uma maior presença de VV/GA nas taxas rápidas, devido à perda de algumas proeminências frasais; (ii) duração semelhante dos GA nas taxas lenta e rápida, consequência da compressão das unidades VV em tais taxas, associada ao maior número de unidades VV/GA.

Análises estatísticas de significância (testes-t e F) revelaram uma diferença significativa (para $\alpha = 0,05$) entre as taxas lenta e rápida para todos os informantes (cf. Tabela 3.1), avaliada pela distribuição das durações das unidades VV no texto. Comprovamos, dessa maneira, estarmos trabalhando

com taxas de elocução distintas (teste-t) e, além disso, serem as unidade VV menos variáveis nas taxas rápidas (teste-F, maiores detalhes sobre essa questão serão apresentados no próximo estudo acústico). Tendo-se, pois, assegurado estatisticamente a diferença entre as taxas, partimos para a comparação entre os contornos duracionais das taxas lenta e rápida.

Tabela 3.1: Duração média (desvio-padrão) em milissegundos, valor p do teste-t e do teste-F das unidades VVs para todos os informantes do estudo preliminar em duas TE, lenta (L) e rápida (R).

VV (duração em ms)				
	L	R	teste-t	teste-F
AC	244(205)	195(139)	0,007	$2,62 \cdot 10^{-7}$
AP	291(189)	204(146)	$1,29 \cdot 10^{-6}$	0,0007
DP	207(164)	167(122)	0,014	0,0004
FA	188(111)	151(74)	0,0004	$2,34 \cdot 10^{-7}$

Os contornos duracionais foram obtidos através da evolução da duração das unidades VVs nos GA, gerados semi-automaticamente no *Microsoft Excel* pelos seguintes procedimentos: 1) aplicação de uma derivada discreta (diferença) no z-score suavizado da duração das unidade VV; 2) aplicação de um teste de máximos, para detectar os pontos de máximo do gráfico do z-score suavizado da duração das unidade VV, através da seguinte fórmula lógica: $SE(E(\text{posição atual} < 0; \text{posição anterior} \geq 0); 1; 0)$. A fórmula atribui tais proeminências da seguinte forma: se o VV atual for menor do que 0 e o VV anterior maior (ou igual) a 0, será atribuído o valor 1 (proeminência

acentual); do contrário será atribuído 0 (nenhuma proeminência encontrada).

Depois de detectarmos os máximos dos contornos duracionais, fizemos um teste t de variáveis independentes para as durações dos grupos acentuais e o de Mann-Whitney U (MWU) para o número de VV por grupo acentual no texto lido, conforme observamos na Tabela 3.2.

Tabela 3.2: Média, desvio-padrão e teste-t para a duração dos grupos acentuais (GA), em milissegundos, e número de VV por grupo acentual (VV/GA) em duas taxas de elocução para todos os falantes. TL e TR são, respectivamente, as taxas lenta e rápida. ** significativo; * marginalmente significativo.

Duração do GA (ms)				Número de VV/GA		
	TL	TR	teste-t	TL	TR	teste-MWV
	μ (σ)	μ (σ)	p<	μ (σ)	μ (σ)	p<
AC	1630 (720)	1350 (440)	0,098	6,65 (1,98)	6,9 (1,82)	0,7
AP	1800 (820)	1400 (560)	0,05*	6,1 (2,5)	6,9 (2,6)	0,4
DP	1360 (500)	1050 (440)	0,03**	7,2 (2,25)	6,65 (1,9)	0,5
FA	1190 (480)	860 (430)	0,008**	6,7 (2,77)	5,93 (2,7)	0,4

Esta tabela sugere para metade dos locutores (AC e AP) uma tendência a não alterar demasiadamente a duração dos grupos acentuais, compensando-a em parte com um maior número de VV por grupo acentual na taxa rápida. Apesar desse maior número de VV por grupo acentual, a duração dos grupos acentuais ainda é marginalmente menor na taxa rápida para esses locutores, o que indica uma compensação parcial. Esta deve estar relacionada a grupos acentuais delimitados por fronteiras prosódicas fortes (para os quais o número de VV no grupo, independentemente da taxa, tende a permanecer como tal,

excetuando-se os casos de elisão) contrastados com grupos acentuais que podem ser reestruturados por estarem separados por uma fronteira fraca, como podemos observar na figura 3.4 para AP.

Com o intuito de verificarmos nossa hipótese de que as fronteiras prosódicas fortes são locais-limite para a ocorrência de reestruturações rítmicas na fala, retiramos os grupos acentuais que se manifestaram da mesma maneira nas taxas lenta e rápida, ficando apenas com os grupos reestruturados nessa mudança de taxas. Procedemos, então, com uma análise estatística dos dados de um falante (AP). Os resultados foram: a) média e desvio-padrão da duração em segundos dos grupos acentuais para a taxa lenta: 1,63 (0,87); e, taxa rápida: 1,79 (0,65); b) média e desvio-padrão do número de VV/GA para a taxa lenta: 5,5 (2,56); e, taxa rápida: 8,4 (2,99). Com esses dados efetuamos um teste-t na duração dos grupos acentuais e um teste MWV no número de VV/GA, que confirmou a hipótese de que as durações nas taxas rápida e lenta tendem a ser iguais ($p < 0.6$) e de que o número de VV por GA na taxa rápida tende a ser maior ($p < 0.02$).

Temos na figura 3.4 reestruturações do trecho “... orador. Subiu a tribuna de um poleiro de ouro...” [o.ra.'dox.su.'biw.a.tri.'bu.nɐ.ɕjũ.pu.'lej.rũ.'ɕjow.rũ]³, em que houve uma mudança de três grupos acentuais na taxa lenta (culmi-

³As transcrições fonéticas desta tese são baseadas no Alfabeto Fonético Internacional (IPA).

nando em [oxs], [un] e [jour]) para dois (culminando em [oxs] e [jour]) na taxa rápida. Nesse trecho percebemos também, como em diversos outros trechos, a importância de havermos utilizado o z-score para a normalização da duração intrínseca dos grupos VV e, com isso, acharmos mais coerentemente a duração extrínseca das frases. Essa frase, antes de calcularmos o z-score, tinha uma proeminência acentual em “de um” [d̥jũ], que foi eliminada após o cálculo de z-score. Com esse procedimento descobrimos que a proeminência acentual estava em “-bu-” de “tribuna” (cf. figura 3.4, taxa lenta).

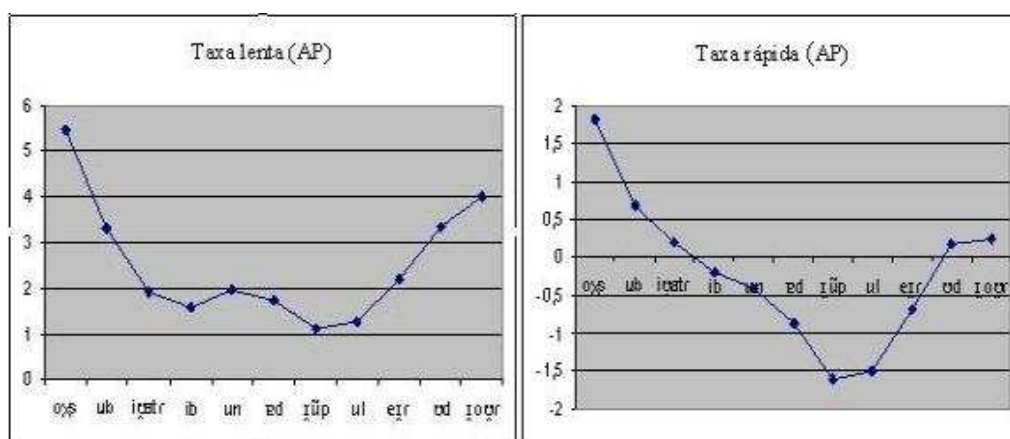


Figura 3.4: Taxas lenta e rápida de AP. Contorno duracional de um enunciado, a partir do z-score suavizado da duração VV, com picos em [oxs], [un] e [jour] (taxa lenta) e [oxs] e [jour] (taxa rápida).

Outro ponto que podemos observar na Tabela 3.2 é o fato de haver menor variação na duração dos grupos acentuais na taxa rápida (menor desvio-padrão), o que sugere uma tendência à isocronia acentual nessa taxa. Observamos esse fenômeno, nitidamente, através da análise da figura 3.5.

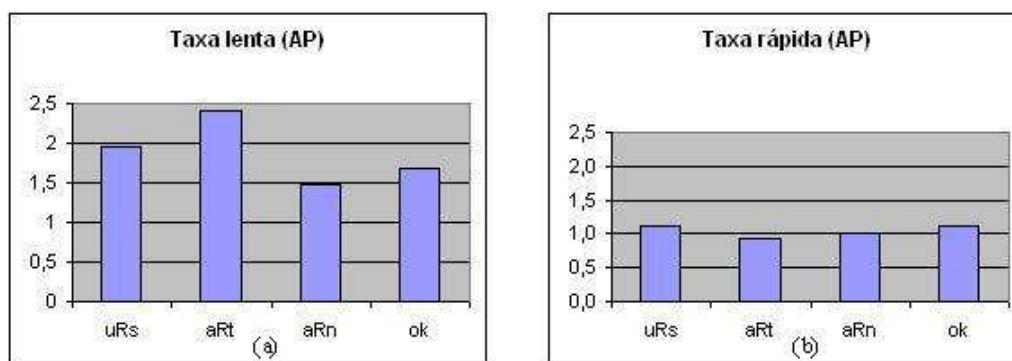


Figura 3.5: Durações dos grupos acentuais em segundos do enunciado “...discurso a respeito da **arte** de falar. Nesse discurso provo **que** ...” nas taxas lenta (a) e rápida (b) de AP. Os itens em negrito representam o VV no final do GA respectivo. A duração do GA culminando em **arn**, por exemplo, é compreendida pelo intervalo “...e de falar. N..”

Na figura 3.5a e b, obtivemos médias e desvios-padrão em segundos para os GA nas duas taxas, obtendo como resultado, (1.9; 0,4) para a taxa lenta e (1,0; 0,08) para a taxa rápida. Depois realizamos um teste-t, a fim de verificarmos se as diferenças entre as duas taxas nesse trecho eram significativas, ou seja, se havia evidências estatísticas para diferenciação entre essas taxas. Obtivemos, assim, um resultado significativo ($p < 0,006$, $\alpha^4=0,05$). Esses resultados mostram uma tendência clara a menores variações nas taxas rápidas (significativa ou marginalmente significativa em três sujeitos), o que está de acordo com menores valores de α , taxa de indução no MDR, a qual modela o aumento, de maneira fina, da duração dos VV até o acento frasal.

⁴Este α representa aqui o valor a partir do qual iremos considerar um valor como significativo (probabilidade menor do que 5% de serem iguais). Não confundir com o valor α no MDR.

Nossos resultados corroboram os fenômenos rítmicos observados em Barbosa (2002a), em que se propõe que os fatores α e β da equação $\Delta T = \alpha.T.s(n).i(m) - \beta.(T - T_0).i(m - 1)$ do MDR, vistos anteriormente, são distintos para diferentes taxas de elocução (na verdade uma variação contínua de taxas), como é demonstrado em Barbosa (2002a), a partir da otimização dos parâmetros do modelo. Para a taxa lenta temos $(\alpha, \beta) = (0.43, 1.04)$, para a taxa normal, $(0.38, 1.13)$ e para a taxa rápida, $(0.31, 1.13)$. Esses valores revelam que, à medida que a taxa de elocução aumenta, a taxa de decaimento (β) aumenta. No entanto, valores baixos de (β) podem mostrar uma certa quantidade de alongamento inicial para as taxas lentas. Por outro lado, à medida que a taxa de elocução aumenta, a taxa de indução⁵ (α) decresce, indicando, talvez, que um acoplamento perfeito seria mais difícil nas taxas rápidas, o que implicaria em uma diferenciação menor entre os períodos do oscilador silábico para as taxas rápidas (e portanto menores valores de variância estatística).

Como podemos notar pela análise da tabela 3.2, a atribuição de proeminências acentuais para esse experimento foi extremamente variável de falante para falante. Dessa forma, para aperfeiçoar os resultados encontrados, decidimos realizar um novo experimento relatado a seguir, no qual a variabilidade

⁵Este parâmetro mostra o efeito de acoplamento entre os dois osciladores. Por isso, se α é menor, temos menos efeito de acoplamento.

de acentuação frasal fosse menor que a aqui encontrada.

3.2 Estudo complementar

3.2.1 Metodologia

Optamos neste estudo por trabalhar com frases isoladas, em vez de um texto como no estudo preliminar, para evitarmos grandes variabilidades de atribuição de grupos acentuais. O *corpus* composto de 11 frases utilizado foi:

1. Há três tipos de abóbora no centro de Belo Horizonte.
2. Há três tipos de abóbora nanica em Belo Horizonte.
3. A análise de tantos dados parecia certa.
4. A análise computacional parecia certa.
5. Quando sinto cócegas, logo tenho falta de ar.
6. Dei um beijo de tirar o fôlego.
7. Fôlego de atleta foi exigido do competidor.

8. Meu fósforo terminou. Me empresta um fósforo.
9. Uso fósforo para acender o fogão.
10. Como carne de boi no sábado.
11. No sábado passado comi carne de boi.

Tal *corpus* foi utilizado com o intuito de corroborar/refutar os resultados preliminares anteriores, a saber, na ocorrência de reestruturações rítmicas:

- O número de VV/GA aumenta proporcionalmente ao crescimento da TE.
- A duração do GA tende a se manter constante com o aumento da taxa de elocução.
- O desvio-padrão absoluto e relativo (coeficiente de variação) é, em geral, menor nas taxas rápidas, o que implica em uma sensação de isocronismo mais acentuada nessas taxas.
- O aumento da taxa exacerba o caráter misto do ritmo em português, ou seja, tendências tanto ao ritmo silábico quanto ao acentual.

Além destas hipóteses oriundas do estudo preliminar, tal *corpus* também foi utilizado a fim de se observar a influência de fatores sintáticos na

atribuição de acentuações frasais. Por esta razão escolhemos frases pareadas (exceto 5) com conteúdos semânticos próximos, mas com variação nos sintagmas ou representação prosódica. Para as frases 1–2, hipotetizamos que o acréscimo de “nanica” ao sintagma preposicional “de abóbora” desloque o acento de “abóbora” (núcleo do sintagma) para “nanica”, devido às características de ritmo com cabeça à direita no PB. De forma semelhante, hipotetizamos que o acréscimo de “computacional” ao sintagma nominal “a análise” (frases 3–4) desloque o acento de “análise” para “computacional”; o acréscimo de “de atleta” (frases 6–7) ao sintagma nominal “fôlego” desloque o acento para “atleta”; e o acréscimo de “passado” (frases 10–11) ao sintagma preposicional “no sábado” desloque o acento para “passado”. Hipotetizamos ainda que, na frase 5, devido a fronteiras sintáticas fortes (marcadas pelo fim de orações subordinadas) o acento frasal recaia em “cócegas” e “ar”; e que nas fronteiras frasais das frases 8–9 recaiam os acentos frasais, ou seja, em “terminou”, “fósforo” e “fogão”. Obviamente, tais expectativas de acentuação não desconsideram que outros acentos frasais, por questões de ênfase, por exemplo, possam ser atribuídos a essas frases.

O informante gravado é falante nativo do PB, do gênero feminino, dialeto mineiro, sem nenhuma deficiência fonatória ou audiológica. A gravação foi feita em uma cabine acusticamente tratada do Lafape/IEL/Unicamp. Os dados foram gravados diretamente no disco rígido a uma taxa de amostragem

de 22,05 kHz. As frases foram lidas em três taxas de elocução distintas seguindo os mesmos procedimentos do estudo preliminar discutido acima.

Neste experimento, diferentemente do estudo anterior, os procedimentos metodológicos foram feitos semi-automaticamente. Primeiramente, etiquetamos no Praat os segmentos VVs e, depois, utilizamos o *script SGDetector*, desenvolvido por Barbosa (2006, vide apêndice A.2 desta tese), para realizar as tarefas metodológicas que descrevemos no estudo preliminar (vide figura 3.3 da seção anterior). Posteriormente, como veremos abaixo, caso considerássemos a atribuição dos acentos frasais errônea, recorriamos às gravações originais para dirimir dúvidas quanto à correta posição de tais acentos.

3.2.2 Análises e Resultados

Antes de começarmos nossa análise, efetuamos, para todas as frases, uma *Anova one-way*, no programa *Statistica* (www.statsoft.com), com a duração da unidade VV como variável dependente e a taxa de elocução como variável independente, com o intuito de verificarmos se estávamos trabalhando com TE distintas em nosso corpus. As análises estatísticas de variância verificaram, para todas as frases, diferenças significativas entre as taxas (as médias e desvios-padrão das unidades VV com suas referidas estatísticas encontram-se na tabela 3.3 abaixo). No entanto, um teste *post-hoc Scheffé* mostrou que

as frases 5, 6 e 9 se distinguíram apenas entre duas taxas, com o seguinte padrão: lenta $\neq$ (normal = rápida).

Tabela 3.3: Duração (média, desvio-padrão) e *Anova One-way* para todas as frases do estudo complementar em três taxas de elocução: lenta (L), normal (N) e rápida (R).

Frase	VV (Duração em ms)					
	L	N	R	Anova	p <	Scheffé
1	259 (118)	202 (99)	149 (81)	F (2,12) = 120,57	10^{-5}	L $\neq$ N $\neq$ R
2	230 (102)	179 (99)	163 (92)	F (2,22) = 200,37	10^{-5}	L $\neq$ N $\neq$ R
3	340 (146)	216 (97)	179 (83)	F (2,8) = 104,63	10^{-5}	L $\neq$ N $\neq$ R
4	218 (84)	176 (76)	145 (63)	F (2,7) = 61,136	0,00004	L $\neq$ N $\neq$ R
5	277 (126)	218 (91)	205 (63)	F (2,13) = 103,88	10^{-5}	L $\neq$ (N = R)
6	263 (103)	212 (88)	198 (96)	F (2,31) = 48,657	10^{-5}	L $\neq$ (N = R)
7	249 (75)	173 (64)	148 (63)	F (2,14) = 136,44	10^{-5}	L $\neq$ N $\neq$ R
8	220 (104)	179 (94)	164 (86)	F (2,25) = 54,810	10^{-5}	L $\neq$ N $\neq$ R
9	206 (70)	162 (67)	178 (68)	F (2,25) = 11,120	0,0004	L $\neq$ (N = R)
10	215 (65)	181 (60)	143 (40)	F (2,24) = 162,59	10^{-5}	L $\neq$ N $\neq$ R
11	233 (81)	189 (58)	161 (60)	F (2,19) = 106,42	10^{-5}	L $\neq$ N $\neq$ R

Ciente, pois, de estarmos lidando com pelo menos duas TE diferentes para cada frase, começaremos a seguir a fazer análises individuais de cada sentença do corpus, para, ao final, discutirmos aspectos gerais de reestruturações rítmicas encontrados neste estudo acústico.

3.2.3 Frases 1 e 2

Os grupos acentuais para as frases 1 e 2 foram divididos da seguinte forma, com os VVs em negrito marcando o fim dos mesmos:

Frase 1:

Taxas lenta-normal: Há três **tipos** | de ab**ób**bora | no centro de Belo | Ho-
rizonte.|

Taxa rápida: Há três tipos de ab**ób**bora | no centro de Belo | Horiz**onte**.|

Frase 2:

Taxa lenta: Há três **tipos** | de ab**ób**bora | nanica em Belo | Horiz**onte**.|

Taxas normal-rápida: Há três tipos de ab**ób**bora | nanica em Belo Horiz**onte**.|

Houve, portanto, para ambas as frases, reestruturações dinâmicas do ritmo sob o efeito do aumento da taxa de elocução, conforme grupos acentuais acima. Para a frase 1, houve a mudança de 4 grupos acentuais (taxas lenta e normal, culminando em “ip”, “ób”, “el” e “ont”) para 3 grupos acentuais (taxa rápida, culminando em “ób”, “el” e “ont”). Tal reestruturação rítmica é evidente em (1) da figura 3.6, em que o contorno duracional da sentença foi modificado gradualmente de uma curva descendente (taxa lenta) a uma curva ascendente (taxa rápida). Quanto aos outros picos acentuais, podemos notar que o pico de “ób” (2 na figura 3.6) não é evidente em nenhuma das três taxas; o pico de “el” (3 na figura 3.6) é evidente na taxa lenta e ambíguo nas demais; e o pico de “ont” (4 na figura 3.6) é evidente em todas as taxas. Para

resolver tais ambigüidades de atribuição de proeminências acentuais, conforme dissemos anteriormente, analisamos de oitiva as repetições das frases em questão e observamos em qual posição nas frases estariam culminando os grupos acentuais. Além disso, verificamos o gráfico de duração não-suavizada nas três taxas, o qual comprova, principalmente, a perda da proeminência em “ti” (1 da figura 3.7) na taxa rápida. Salientamos que, para todos os casos, a proeminência obtida automaticamente sempre se encontrava dentro da mesma palavra da proeminência efetiva analisada. Tal análise resultou nos grupos acentuais como descrito acima para as frases 1 e 2.

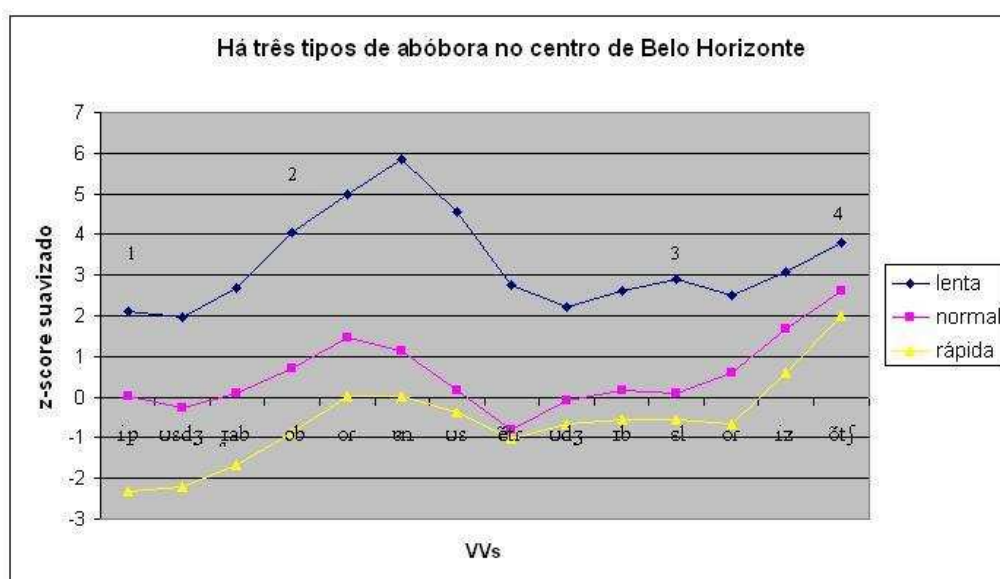


Figura 3.6: Grupos acentuais para a frase “Há três tipos de abóbora no centro de Belo Horizonte” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Lembremos que, devido aos procedimentos metodológicos empregados, só obtíamos GA a partir do terceiro VV das frases de nosso *corpus*, por isso o contorno duracional começando em “ip” aqui. Para maiores detalhes vide texto.

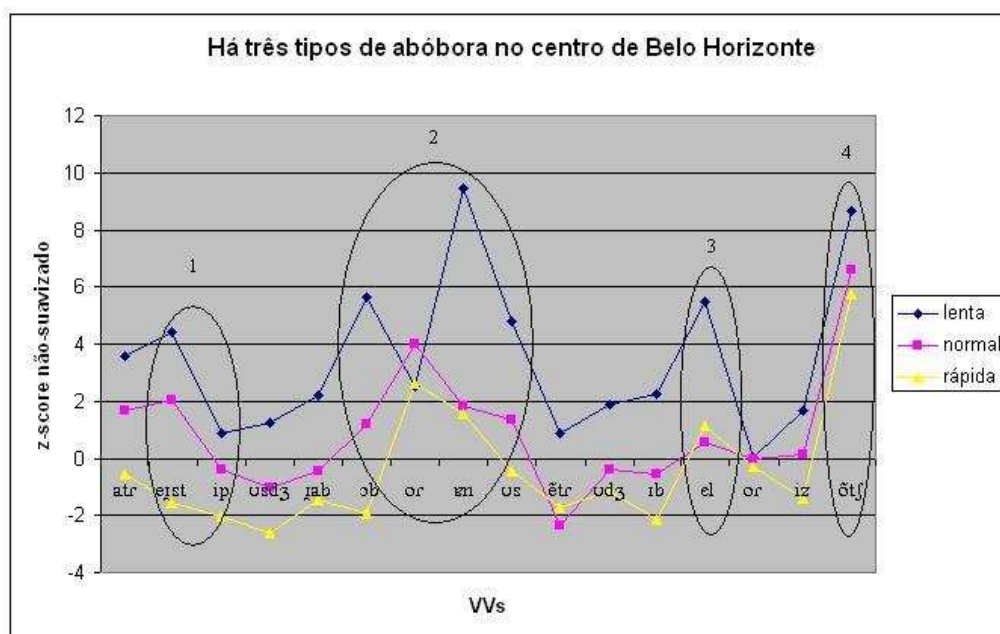


Figura 3.7: Grupos acentuais para a frase “Há três tipos de abóbora no centro de Belo Horizonte” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score não-suavizado. Com este gráfico podemos observar o pico duracional em “ti” nas taxas lenta e normal, ausente na rápida.

Para a frase 2, houve a mudança de quatro grupos acentuais (taxa lenta, culminando em “ip”, “ób”, “el” e “ont”) para dois grupos acentuais (taxa rápida, culminando em “ób” e “ont”). Em (1) da figura 3.8, como na figura 3.6, é evidente a mudança de padrão do contorno duracional de descendente (taxa lenta) para ascendente (taxa rápida). O pequeno movimento ascendente nas taxas normal e rápida não foi computado como pico pela análise automática nem pela de oitiva. Em (2) da figura 3.8, como na figura 3.6, o pico em “ób” foi comprovado de oitiva. Em (3) da figura 3.8 a mudança do contorno

duracional é mais suave. Para a taxa lenta há, apesar de bastante sutil, um pico em “el”, comprovado pela análise de oitiva, mas não nas outras taxas. Já para (4) da figura 3.8 é evidente o pico em “ont” para todas as taxas.

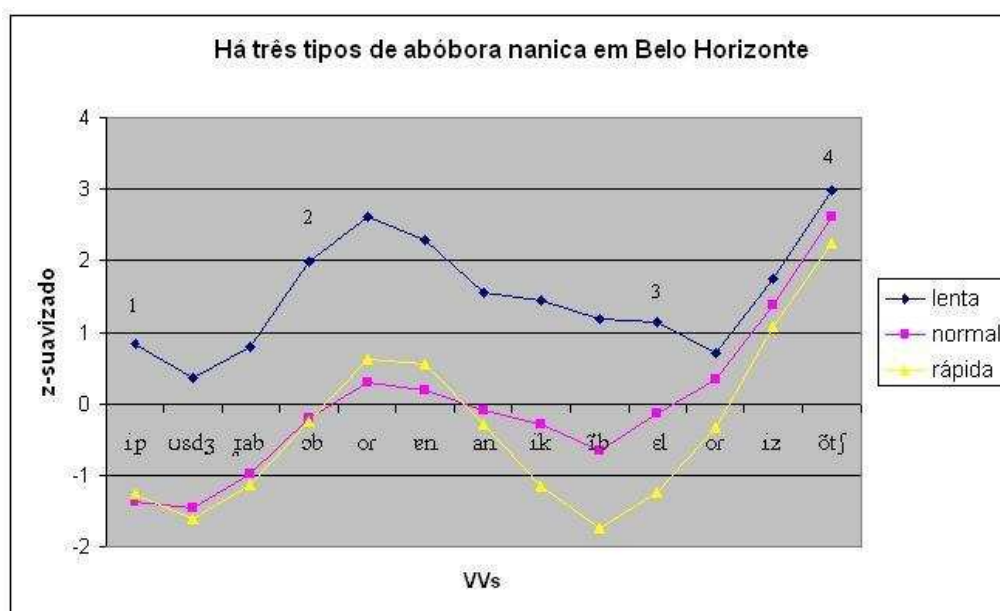


Figura 3.8: Grupos acentuais para a frase “Há três tipos de abóbora nanica em Belo Horizonte” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

A inspeção das figuras 3.6 e 3.8 ainda nos revela que o deslocamento acentual de “abóbora” para “nanica” não ocorreu de acordo com nossa hipótese (vide p. 122; observem o grande pico duracional em [ɔb] em 2 da figura 3.8), permanecendo a acentuação frasal no constituinte “abóbora”. Precisamente, das 10 repetições analisadas para a taxa lenta, houve apenas um caso de proeminência acentual em “nanica”, o que impossibilita qualquer análise

estatística, mas que, no entanto, demonstra ser possível tal deslocamento acentual. Especulamos, como ainda não temos dados de frequência lexical da língua falada e por considerarmos ser a frequência lexical da língua escrita distinta da falada, que, nesse caso, o falante optou por manter uma proeminência acentual em “abóbora”, palavra bastante freqüente no dialeto mineiro, em detrimento de “nanica”, devido à influência da frequência lexical. Hipotetizamos, pois, que palavras muito freqüentes em PB tenderiam, pelo menos para alguns falantes, a atrair proeminências acentuais para si. Ainda para esse caso, pode ser que o falante estudado tenha atribuído ênfase à palavra “abóbora” em todas as taxas, o que poderia explicar a manutenção do acento frasal em todas as três taxas. Evidências para isso são encontradas na análise da curva de F0 (frequência fundamental) no trecho “abóbora nanica” (figura 3.9), em que há um padrão ascendente em abóbora (seta ascendente da figura 3.9), sugerindo ênfase, e descendente (seta descendente da figura 3.9) em nanica, sugerindo apenas final de sintagma. Estudos acústicos com situações de foco distintas poderiam verificar se, de fato, palavras enfatizadas modificam proeminências acentuais padrão previstas pelo MDR.

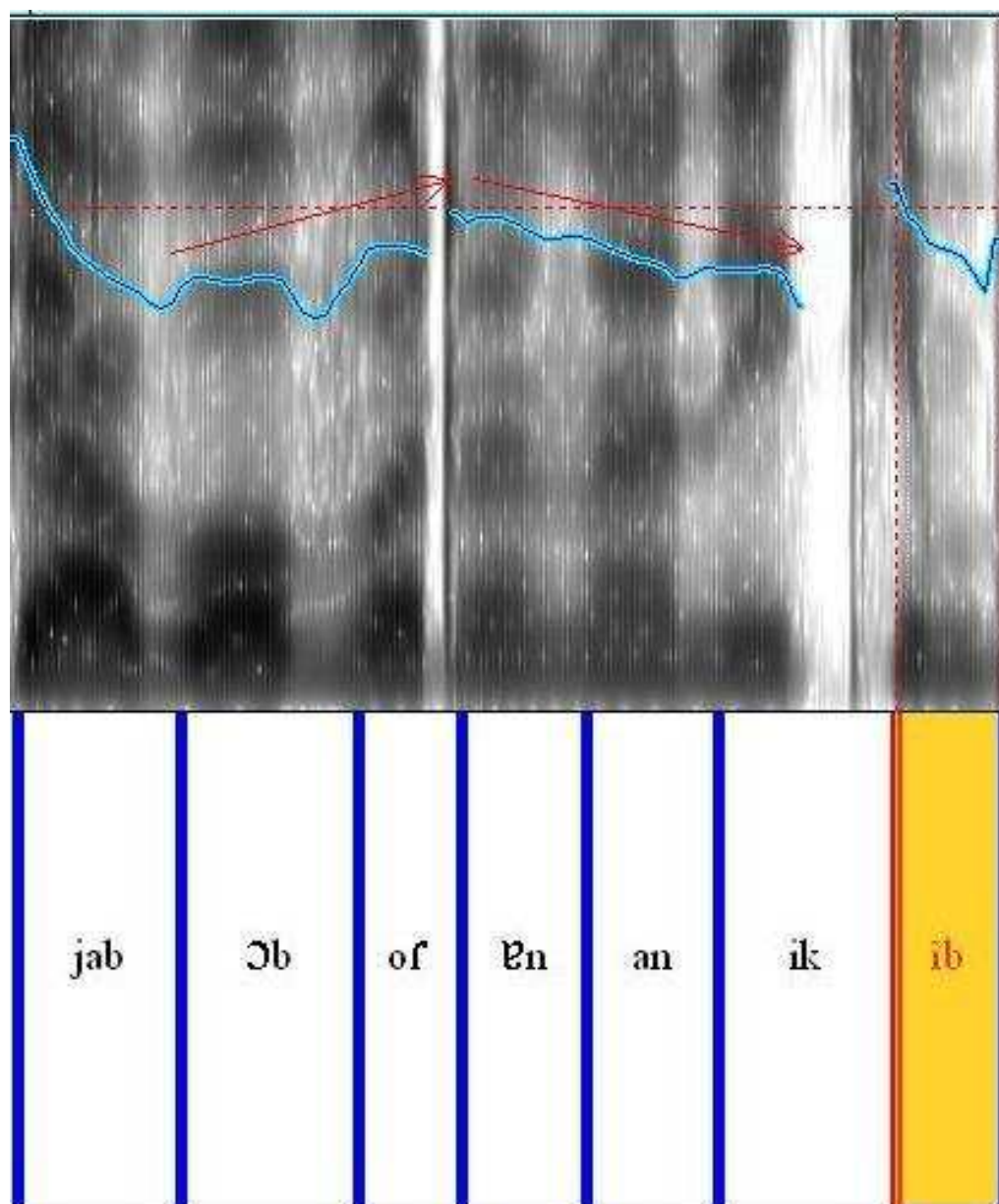


Figura 3.9: Contorno de F0 para o trecho “abóbora nanica” na taxa rápida. O eixo x representa a duração das unidades VVs e o eixo y, a curva de F0 medida em Hz.

3.2.4 Frases 3 e 4

Os grupos acentuais para as frases 3 e 4 foram divididos da seguinte forma:

Frase 3:

Taxa lenta: A **análise** | de tantos **dados** | parecia **certa**.|

Taxa normal: A **análise** | de tantos dados parecia **certa**.|

Taxa rápida: A análise de tantos dados parecia **certa**.|

Frase 4:

Taxa lenta: A **análise** | **computacional** | parecia **certa**.|

Taxas normal-rápida: A análise **computacional** | parecia **certa**.|

Para ambas frases houve reestruturações rítmicas, porém de forma distinta. Para a frase 3, os três grupos acentuais na taxa lenta (culminando em “ál”, “ad” e “ert”) se reorganizaram para dois grupos acentuais (culminando em “ál” e “ert”) na taxa normal e 1 GA na taxa rápida. Conforme (2) na figura 3.10, observamos um pico sutil em “ad” na taxa lenta, mas não nas outras taxas. Já análise do pico em “al” foi mais complicada, pois, de acordo com o método utilizado, as proeminências só são reveladas a partir da terceira unidade VV (vide derivada e teste de máximos na figura 3.3). Por isso, efetuamos uma análise de oitava. Tal análise nos mostrou que a proeminência

observada em (1) da figura 3.10 na verdade correspondia ao efeito de uma proeminência em “al” (um VV anterior). A terceira proeminência acentual desta frase está representada claramente em (3 da figura 3.10) para todas as taxas.

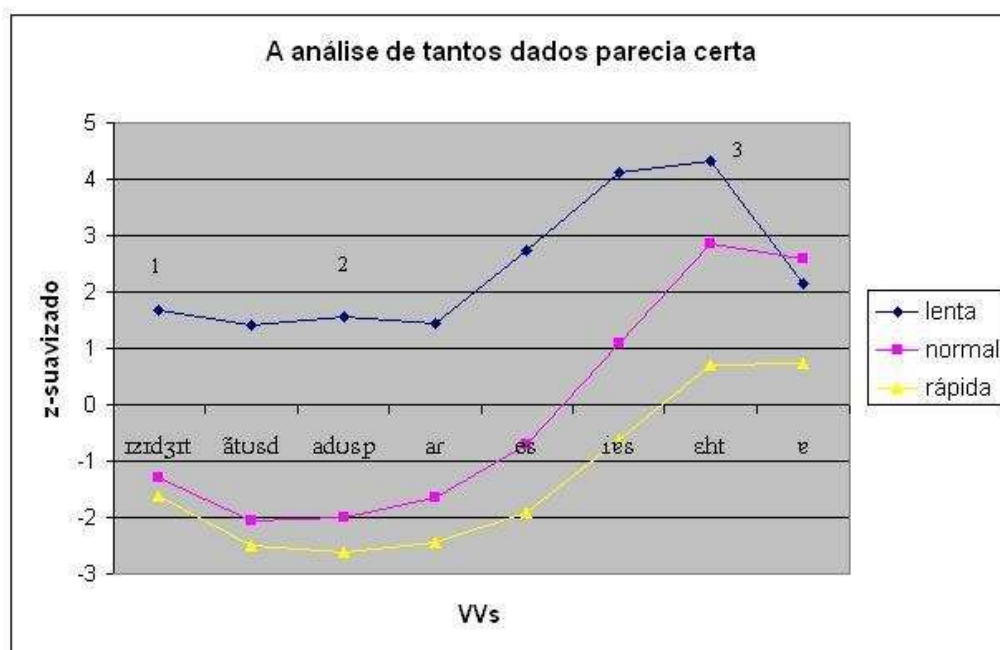


Figura 3.10: Grupos acentuais para a frase “A análise de tantos dados parecia certa” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

Na frase 4, diferentemente da frase 3, houve uma reestruturação rítmica de três grupos acentuais na taxa lenta (culminando em “ál”, “alp” e “ert”) para dois (culminando em “alp” e “ert”) nas taxas normal e rápida, mas, ao invés de se perder um acento frasal medial, o falante optou por não realizar a primeira proeminência acentual. Para esta frase a proeminência acentual em

“alp” (2 da figura 3.11) é evidente na taxa lenta e comprovada de oitiva nas outras taxas. Já em “ert” (3 da figura 3.11) é evidente a proeminência em todas as taxas. No entanto, como na frase 3, a atribuição da proeminência acentual em “ál” (1 da figura 3.11 representa um pico que foi deslocado para o VV “ál” de “análise”) foi obtida através da análise de oitiva. A análise do z-score não-suavizado também nos auxiliou, pois pudemos ver um pico em “iz” ou “ál” nas três taxas (1 da figura 3.12), indicando uma proeminência em “ál”, e um pico em “ac” ou “alp”, indicando uma proeminência acentual em “alp” para as três taxas.

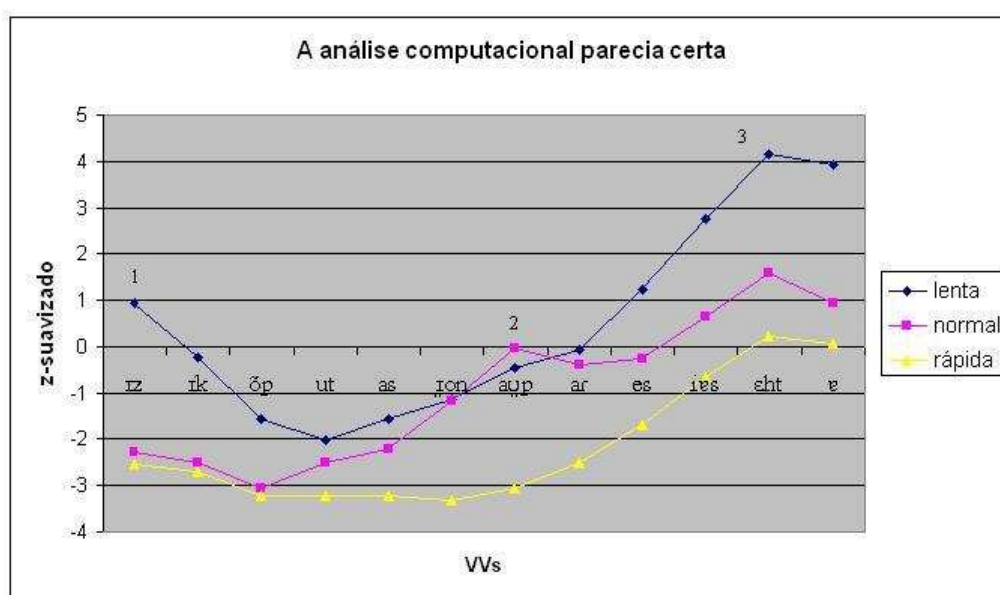


Figura 3.11: Grupos acentuais para a frase “A análise computacional parecia certa” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

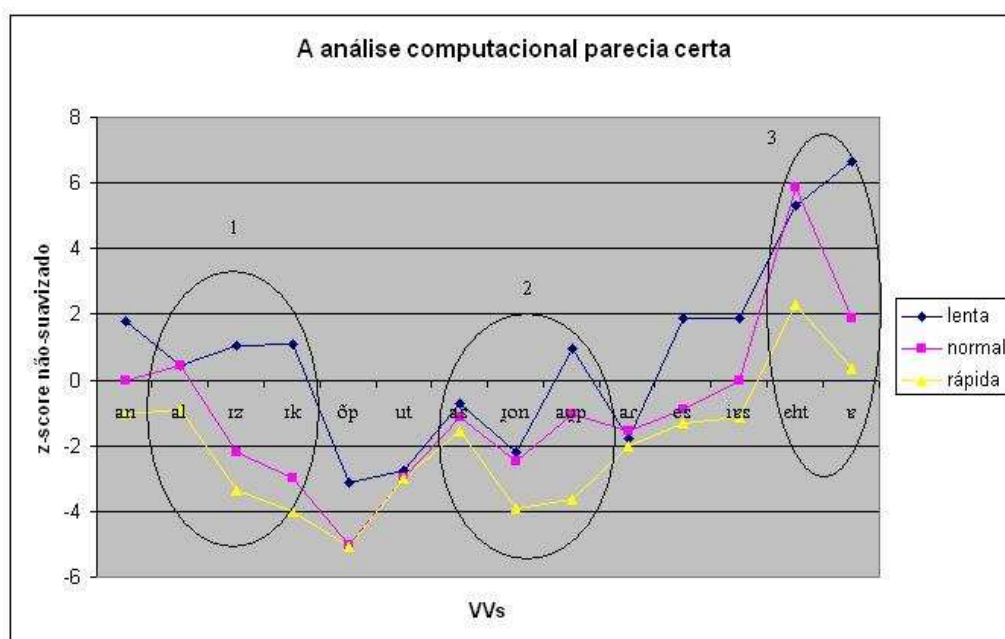


Figura 3.12: Grupos acentuais para a frase “Há três tipos de abóbora no centro de Belo Horizonte” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score não-suavizado.

Comparando a frase 4 com a frase 2, observamos que o deslocamento da proeminência acentual da esquerda para a direita ocorreu de acordo com nossa hipótese (vide p. 122). No entanto, acreditamos que a questão da frequência lexical na atração de proeminências acentuais não deve ser descartada, pois, mesmo na frase 4, sabemos que “computacional” é uma palavra altamente freqüente nos meios acadêmicos e, com certeza, muito comum para a informante graduada utilizada em nosso experimento.

3.2.5 Frase 5

A frase 5 foi realizada da seguinte maneira:

Frase 5:

Taxas lenta-normal-rápida: Quando sinto cócegas |, logo tenho falta de ar.|

Analisando a figura 3.13 pode-se perceber que as proeminências acentuais em “óc” e (1 da figura 3.13) “e ar” (2 da figura 3.13) não são evidentes pelo gráfico. Ressalto aqui que o método automático utilizado também não atribuiu as proeminências acentuais nos devidos locais. Como exemplo observemos a taxa normal em que a proeminência acentual foi marcada em “asl” e não “ós”. Para atribuir as proeminências acentuais, conforme acima apresentado, fizemos uso também aqui da análise de oitiva.

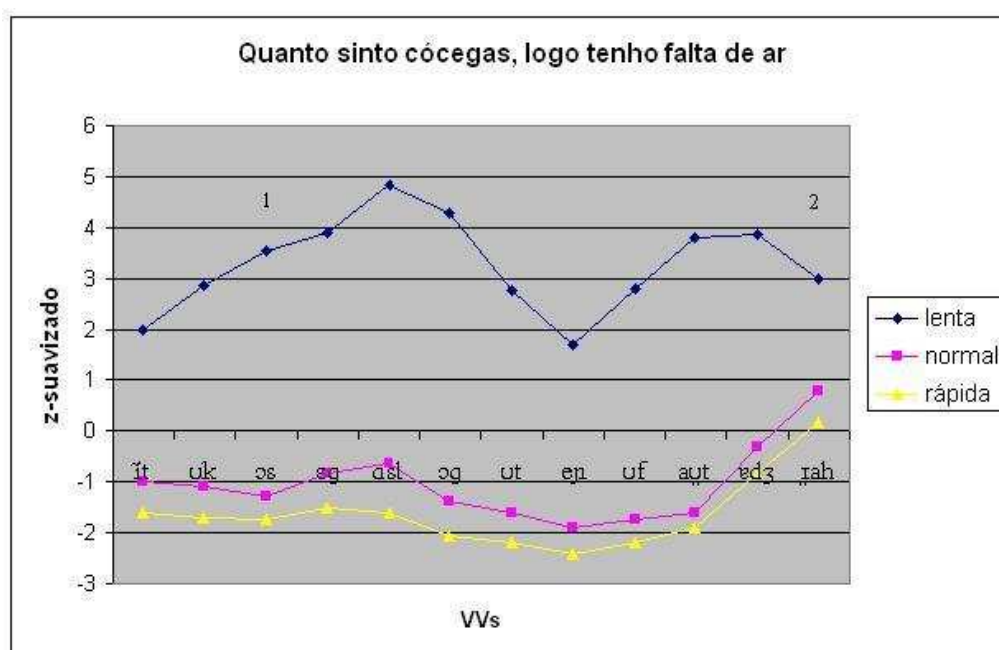


Figura 3.13: Grupos acentuais para a frase “Quando sinto cócegas, logo tenho falta de ar” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

Para essa frase, como hipotetizamos, observamos que as fronteiras sintáticas fortes (no caso fronteiras frasais indicadas por pausas) inibiram possíveis reestruturações rítmicas da frase. Sendo assim, a hipótese de existirem atratores sintáticos para a realização das proeminências acentuais não pode ser descartada. Contudo, há que se realizar experimentos mais detalhados para verificar qual fator é mais decisivo na atribuição das proeminências acentuais — fator sintaxe, frequência lexical ou semântico-pragmático, pois, como veremos na análise das frases 8–9, as fronteiras sintáticas fortes não implicam, necessariamente, em uma fronteira frasal nas taxas rápidas.

3.2.6 Frases 6 e 7

As frases 6–7 foram produzidas da seguinte forma:

Frase 6:

Taxas lenta: Dei um **beijo** | de tirar o **fôlego**.|

Taxas normal-rápida: Dei um beijo de tirar o **fôlego**.|

Frase 7:

Taxas lenta: **Fôlego** | de atleta | foi exigido | do competidor.|

Taxas normal-rápida: **Fôlego** | de atleta foi exigido | do competidor.|

Conforme os dados apresentados, a frase 6 mudou de dois grupos acen-
tuais (culminando em “eij” e “ôl”) na taxa lenta para um grupo acentual
(culminando em “ôl”) nas taxas normal e rápida. Mais uma vez verificamos
aqui a necessidade de fazermos uso da análise de oitiva para auxiliar a análise
automatizada, pois em (1) e (2) da figura 3.14 o acento recai sobre os VVs “o
de t” e “ego”, mas a observação de oitiva nos indica acentos como apontamos
acima.

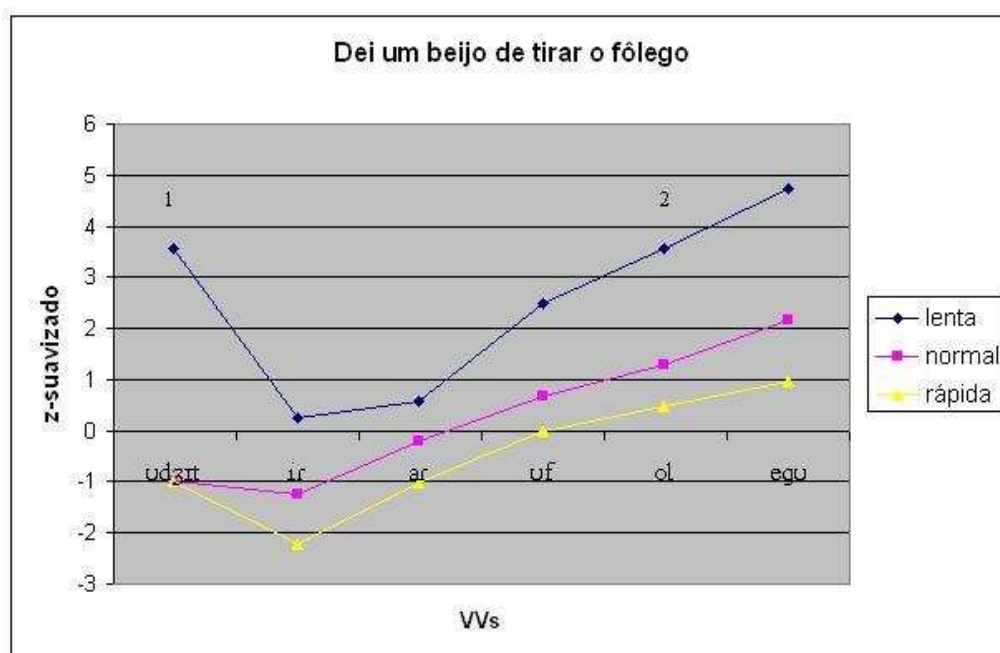


Figura 3.14: Grupos acentuais para a frase “Dei um beijo de tirar o fôlego” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

Para a frase 7 houve a reestruturação rítmica de quatro grupos acentuais na taxa lenta (culminando em “ôl”, “et”, “id” e “or”) para três grupos acentuais (culminando em “ôl”, “id” e “or”) nas taxas normal e rápida. Em (1) da figura 3.15 observamos um pico em “od” para todas as taxas, sugerindo uma proeminência acentual inicial em “ôl”, corroborada pela análise de oitiva. Esta análise foi essencial aqui, pois, de acordo com nossa metodologia (vide figura 3.3), só podemos visualizar as acentuações frasais a partir do terceiro VV. Contudo, esta proeminência foi descartada de nossas análises quantitativas subseqüentes por ser um GA acentual de apenas um VV, no qual não pode ser realizado o procedimento de suavização do z-score que requer cinco pontos (vide figura 3.3). Em (2) da figura 3.15 a proeminência acentual em “et” foi dada pela análise de oitiva. Já em (3) da figura 3.15. verificamos uma proeminência acentual nítida graficamente para todas as taxas. Por fim, em (4) da figura 3.15, para todas as taxas, o gráfico representa uma reestruturação rítmica com cabeça à direita, enfatizada pelo aumento da TE.

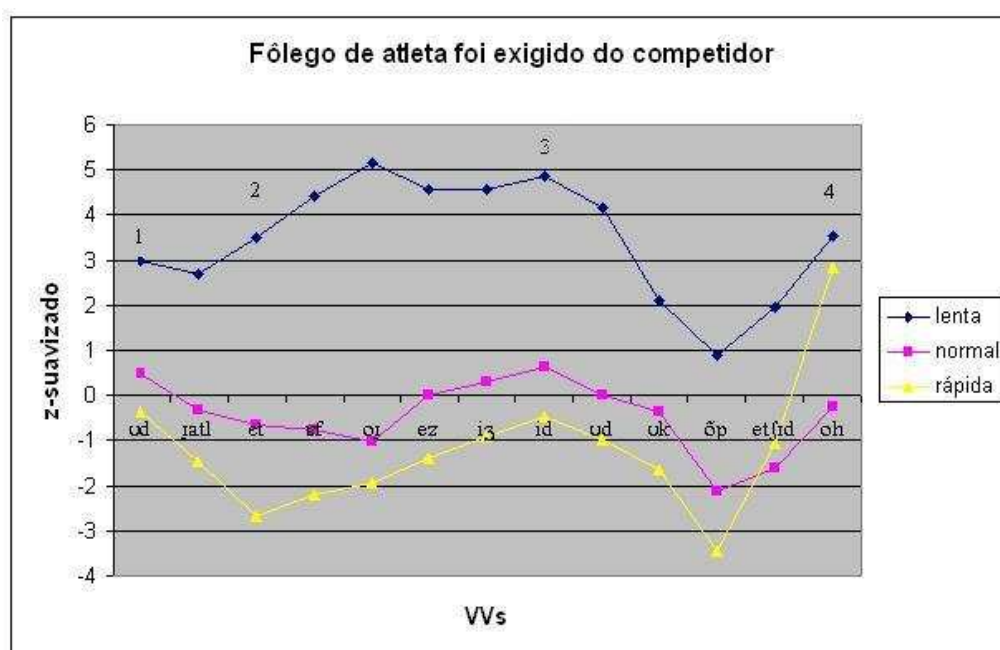


Figura 3.15: Grupos acentuais para a frase “Fôlego de atleta foi exigido do competidor” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

Para esse par de sentenças, como nas frases 3 e 4, houve o deslocamento da proeminência acentual da esquerda para à direita com o acréscimo do sintagma preposicional “de atleta” (cf. frase 7, taxa lenta) à palavra “fôlego”, como havíamos hipotetizado (vide p. 122). No entanto, tal mudança de proeminência, como nas frases 1 e 2, não ocorreu nas taxas normal e rápida.

3.2.7 Frases 8 e 9

As frases 8-9 foram realizadas da seguinte maneira:

Frase 8:

Taxa lenta: Meu **fósforo** | terminou. | **Me** empresta um **fósforo**. |

Taxas normal-rápida: Meu **fósforo** | terminou. Me empresta um **fósforo**. |

Frase 9:

Taxa lenta-normal: Uso **fósforo** | para acender o **fogão**. |

Taxa rápida: Uso **fósforo** para acender o **fogão**. |

Também aqui observamos que o aumento da taxa de elocução atua na reestruturação rítmica das frases. Na frase 8 houve a mudança de três grupos

acentuais (taxa lenta) para dois grupos acentuais (taxas normal e rápida). Em (1) da figura 3.16 aparece um pico duracional em “ot” (taxas lenta e rápida) “or” (taxa normal) de “fósforo” que, pela análise de oitiva, foi deslocado para “ósf”. Aqui, como comentado para a frase 7, só podemos visualizar as acentuações frasais a partir do terceiro VV. Neste caso os dois primeiros VVs são [ewf] e [ɔsf], mascarando assim possíveis acentos nos mesmos. Já (2) e (3) da figura 3.16 mostram claramente a perda de proeminência frasal em “terminou” e a proeminência frasal final em “fósforo”, excetuando-se a proeminência final na taxa normal, comprovada pela análise de oitiva.

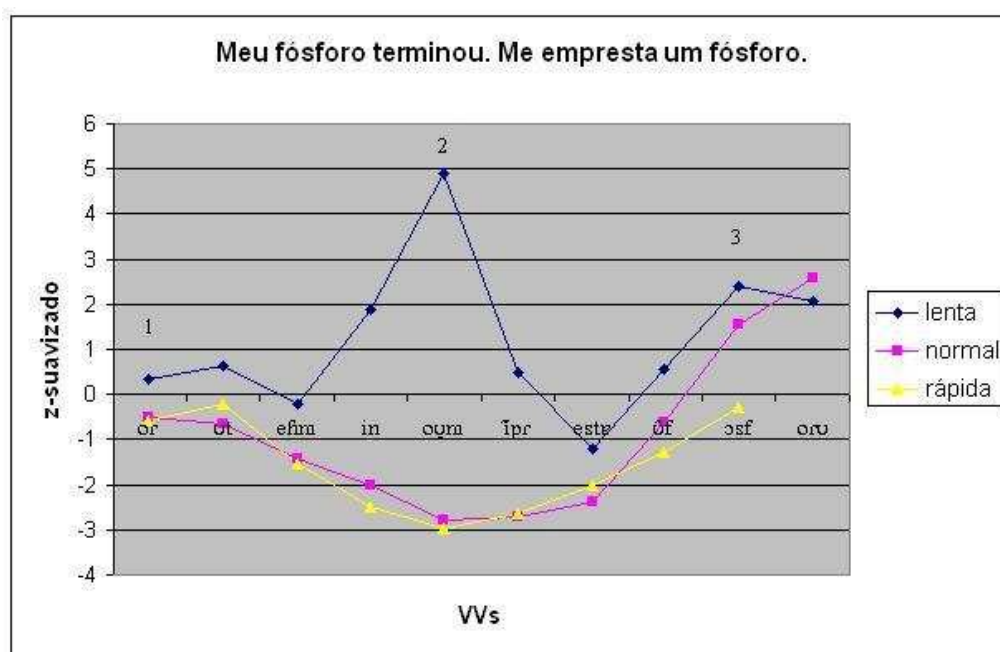


Figura 3.16: Grupos acentuais para a frase “Meu fósforo terminou. Me empresta um fósforo” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

Na frase 9 houve a reestruturação de dois grupos acentuais nas taxas lenta e normal (culminando em “ósf” e “ão”) para um (culminando em “ão”). Para essa frase fica evidente as proeminências acentuais em “ósf” (1 da figura 3.17) e “ão” (2 da figura 3.17).

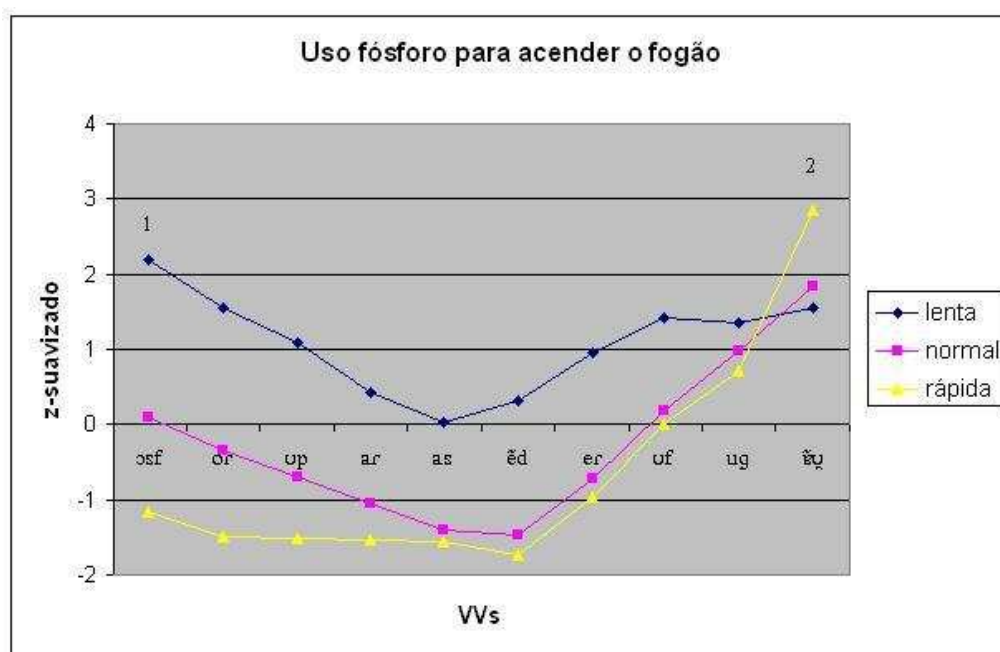


Figura 3.17: Grupos acentuais para a frase “Uso fósforo para acender o fogão” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, os valores do z-score suavizado. Para maiores detalhes vide texto.

De acordo com nossa hipótese de força do fator sintaxe na não alteração de acentuações frasais, vista anteriormente, esperávamos que os acentos nas fronteiras frasais da frase 8 fossem mantidos em todas as TE. No entanto, surpreendentemente, diferentemente da frase 5, houve, a partir da taxa normal, a perda da proeminência frasal na primeira frase (“ou” de “terminou”).

Sendo assim, conforme havíamos hipotetizado, pode haver forças em conflito na atribuição dos acentos frasais. Nos casos aqui estudados acreditamos serem: fator sintaxe x fator frequência lexical x fator semântico-pragmático. Ainda ritmizações distintas podem ser efetuadas pelos falantes, de acordo com a escolha de um ou outro fator. Por exemplo, o falante aqui estudado não atribuiu nenhuma pausa nítida nas taxas normal e rápida (cf. análise de oitiva) na realização da frase 8. Por esse motivo, não foram verificadas as proeminências acentuais hipotetizadas (vide p. 130). Se, no entanto, houvesse tais pausas, a ritmização da sentença seria feita de forma distinta.

3.2.8 Frases 10 e 11

As frases 10-11 foram produzidas da seguinte forma:

Frase 10:

Taxa lenta: Como carne de **boi** | no **sábado**.|

Taxas normal-rápida: Como carne de boi no **sábado**.|

Frase 11:

Taxa lenta-normal-rápida: No sábado **passado** | comi carne de **boi**.|

Nessas duas frases, só houve reestruturação rítmica para a frase 10, de dois grupos acentuais na taxa lenta (culminando em “oi n” e “áb”) para um nas taxas normal e rápida (culminando em “áb”). Para a taxa lenta, observamos que a proeminência é evidente em (1) da figura 3.18, mas não em (2) da figura 3.18, comprovada pela análise de oitiva. Para as outras taxas, é evidente a proeminência final em (2) da figura 3.18.

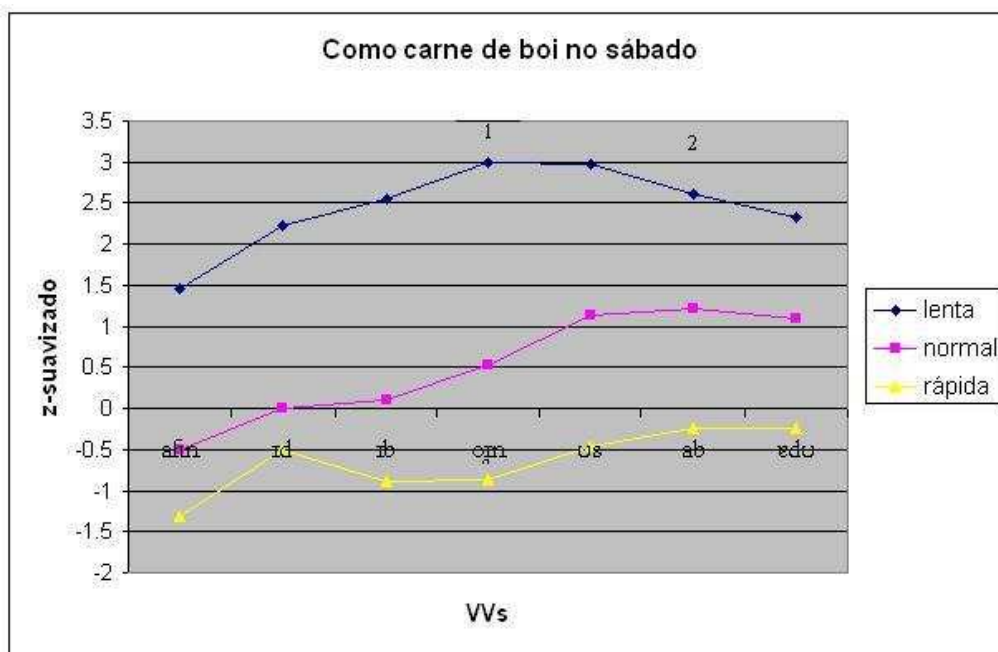


Figura 3.18: Grupos acentuais para a frase “Como carne de boi no sábado” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, as derivadas do z-score suavizado. Para maiores detalhes vide texto.

Já para a frase 11, conforme a análise do contorno duracional da figura 3.19, verificamos nitidamente padrões rítmicos semelhantes para todas as ta-

xas. Acreditamos ter ocorrido tal padrão, pois houve muitas possíveis sobreposições gestuais⁶ já na taxa lenta, conforme observamos no espectrograma da figura 3.20, que representa uma realização da frase 11 na taxa lenta. Observamos sobreposições gestuais em [aduk] e [afinid₃] no espectrograma da figura 3.20.

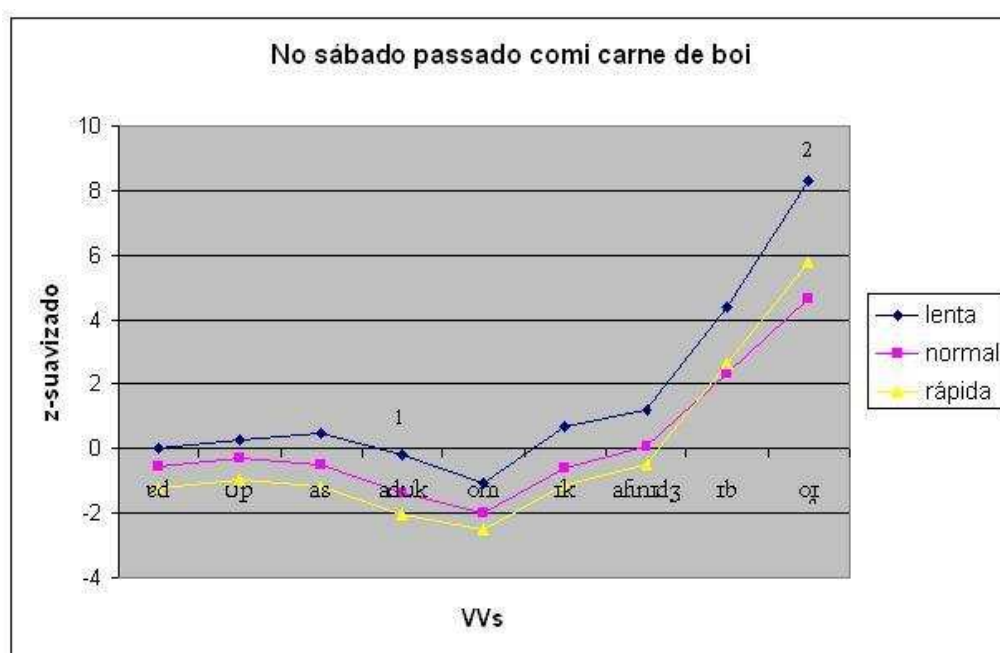


Figura 3.19: Grupos acentuais para a frase “No sábado passado comi carne de boi” em três taxas de elocução. O eixo x representa as unidades VVs e o eixo y, as derivadas do z-score suavizado. Para maiores detalhes vide texto.

Para o par de frases dessa seção, ocorreu, diferentemente das frases 1-2, como esperado, um deslocamento acentual ao se acrescentar “passado” ao sintagma preposicional “no sábado”. Acreditamos ter sido tal reestruturação

⁶Não temos dados articulatórios para confirmar essa afirmação. No entanto podemos fazer inferências a partir do sinal acústico analisado.

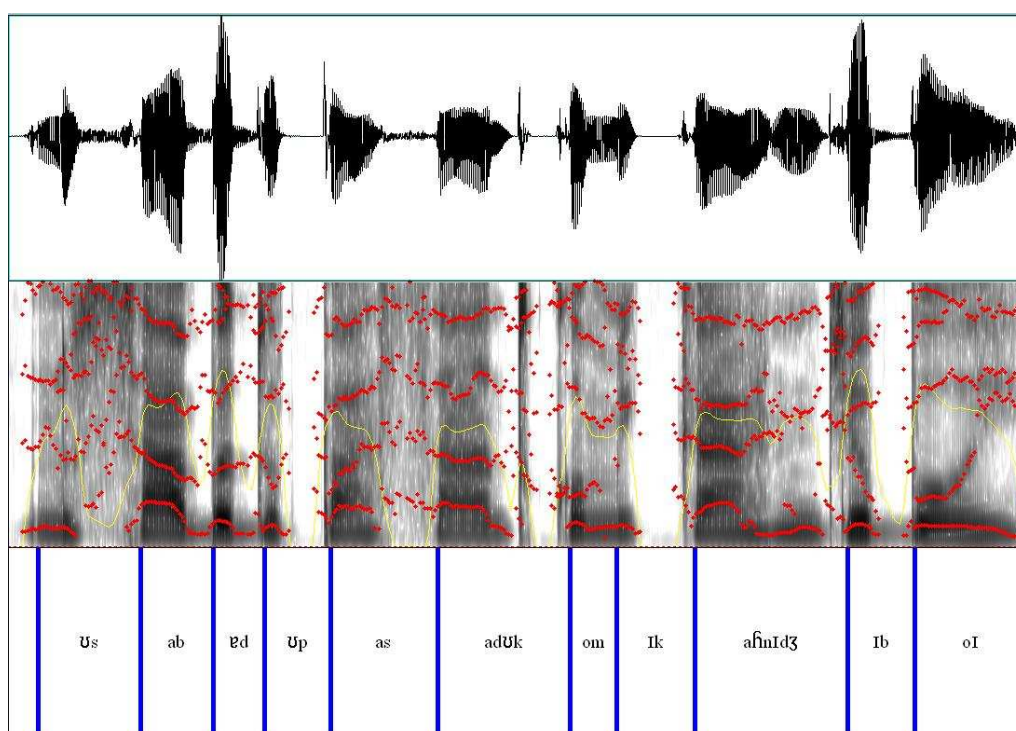


Figura 3.20: Espectrograma representando sobreposições gestuais articu-
tórias ([aduk] e [afnIdz]) na frase 11, inferidas a partir do sinal acústico.

rítmica novamente influenciada pelo fator sintaxe, pois “no sábado passado” pode estar funcionando aqui como uma fronteira sintática forte, o que faz com que o acento frasal recaia em “passado” (cabeça à direita), dividindo a frase em dois grandes blocos: “no sábado passado” (sintagma preposicional) e “comi carne de boi” (sintagma verbal). Essa possibilidade é tal que na língua escrita o sintagma preposicional comumente é separado por vírgulas, como na frase: “na semana passada, o Banco Central informou que a taxa do dólar iria subir”.

3.2.9 Análises quantitativas

Deixando de lado as análises qualitativas, partimos agora para as análises quantitativas, a fim de verificar as hipóteses mencionadas na metodologia deste estudo complementar.

Primeiramente, comentemos a questão da diminuição do desvio-padrão com o aumento da TE. Em relação ao desvio-padrão da duração das unidades VV, todas as frases apresentaram um padrão decrescente tomando-se como base o desvio-padrão da duração dos VV em função da TE. Conforme a tabela 3.4, no entanto, somente algumas frases apresentaram diferenças significativas entre as taxas através de uma *ANOVA One-way* (frases 1, 3, 4, 5, 10, 11), sendo que um teste *post-hoc Scheffé* relatou diferenças significativas

entre as três taxas somente para as frases 1 e 5. Esse teste ainda revelou que, para as frases 3 e 11, a taxa lenta (L) foi diferente da normal (N) e rápida (R) ($L \neq (N = R)$) e para as frase 4 e 10, a taxa lenta foi igual a normal e diferente da rápida ($(L=N) \neq R$). Mesmo não ocorrendo diferenças significativas de desvio-padrão da unidade VV em função da TE para todas as frases, os resultados indicam que a diminuição do desvio-padrão ao se aumentar a taxa é bastante sistemático, conforme indica a significância estatística ao se tomar os dados do desvio-padrão da duração da unidade VV de todas as frases em função da TE ($F(2,230) = 4,3419$, $p < 0,015$). Entretanto, um teste *post-hoc* *Scheffé* apresenta diferenças significativas apenas entre os dois extremos de TE: lenta e rápida.

Tabela 3.4: Desvio-padrão da duração das unidades VV e *Anova One-way* para todas as frases do estudo complementar em três taxas de elocução: lenta (L), normal (N) e rápida (R). n.s. significa valor não significativo e n.s.a., não se aplica.

Frase	VV (Duração em ms)					
	L	N	R	Anova	p <	Scheffé
1	118	99	81	F (2,12) = 30,610	0,00003	L $\neq$ N $\neq$ R
2	102	99	92	n.s.	n.s.	n.s.a.
3	146	97	83	F (2,8) = 52,687	0,00003	L $\neq$ (N = R)
4	84	76	63	F (2,7) = 12,271	0,006	(L = N) $\neq$ R
5	126	91	63	F (2,13) = 46,937	10 ⁻⁵	L $\neq$ N $\neq$ R
6	103	88	96	n.s.	n.s.	n.s.a.
7	75	64	63	n.s.	n.s.	n.s.a.
8	104	94	86	n.s.	n.s.	n.s.a.
9	70	67	68	n.s.	n.s.	n.s.a.
10	65	60	40	F (2,24) = 17,253	0,00003	(L = N) $\neq$ R
11	81	58	60	F (2,19) = 28,792	10 ⁻⁵	L $\neq$ (N = R)

A análise da tabela 3.4 acima levou em consideração todos os VVs presentes nos GA das sentenças de nosso corpus. Contudo, lembrando que os VV nos GA do MDR crescem exponencialmente culminando no acento frasal (vide p. 81), se eliminarmos os VVs presentes nas proeminências acentuais da fronteira, podemos observar o efeito da diminuição do desvio-padrão da duração das unidades VV com o aumento da TE mais facilmente. Além disso, como vimos nas páginas 101-2, as unidades VVs átonas funcionam como referência para a percepção da TE. Para comprovar essa hipótese de menor desvio-padrão para a duração das unidades VVs átonas, pois, retiramos todos os VVs em posição de acento frasal e realizamos uma outra Anova

One-way com o desvio-padrão da duração VV em função da TE (vide tabela 3.5). Os resultados desta análise estatística encontram-se na tabela 3.6 abaixo.

Tabela 3.5: Média e desvio-padrão da duração das unidades VV átonas para todas as frases do estudo complementar em três taxas de elocução: lenta (L), normal (N) e rápida (R). Itens em vermelho representam sentenças que tiveram muitas sobreposições gestuais.

Frases	Duração média (desvio-padrão) dos VVs átonos		
	L	N	R
1	203,3 (67,2)	177,4 (62,9)	138,0 (39,9)
2	217,6 (91,4)	156,9 (62,7)	141,1 (52,3)
3	340,4 (145,9)	216,5 (96,8)	179,4 (82,9)
4	198,8 (76,4)	157,7 (54,6)	130,5 (47,1)
5	251,2 (110,5)	194,4 (70,3)	182,2 (39,1)
6	265,9 (118,5)	212,0 (88,1)	198,2 (95,8)
7	235,7 (81,3)	163,6 (60,6)	141,9 (57,6)
8	180,3 (59,5)	151,3 (41,9)	140,8 (47,6)
9	190,3 (57,8)	145,9 (61,8)	173,4 (69,4)
10	203,3 (67,2)	177,4 (62,9)	138,0 (39,9)
11	209,5 (70,6)	175,9 (56,2)	142,4 (47,2)

Tabela 3.6: *Anova One-way* do desvio-padrão da duração das unidades VVs átonas para todas as frases do estudo complementar em três taxas de elocução: lenta (L), normal (N) e rápida (R). Itens em vermelho representam sentenças que tiveram muitas sobreposições gestuais. n.s. significa valor não significativo e n.s.a., não se aplica.

Frases	Análises Estatísticas		
	Anova	p <	Scheffé
1	F (2,12) = 77,656	10^{-5}	L $\neq$ N $\neq$ R
2	F (2,22) = 30,483	10^{-5}	L $\neq$ (N = R)
3	F (2,8) = 24,980	0,0004	L $\neq$ (N = R)
4	F (2,7) = 19,970	0,0013	L $\neq$ (N = R)
5	F (2,13) = 103,97	10^{-5}	L $\neq$ N $\neq$ R
6	F (2,31) = 7,1384	0,0029	L $\neq$ (N = R)
7	F (2,14) = 23,404	0,00004	L $\neq$ (N = R)
8	F (2,25) = 2,6796	0,04	L $\neq$ (N = R)
9	n.s.	n.s.a.	n.s.a.
10	F (2,24) = 13,483	0,00013	(L = N) $\neq$ R
11	F (2,19) = 24,075	0,00002	L $\neq$ N $\neq$ R

A análise das tabelas 3.5 e 3.6 corrobora nossa hipótese de que, excluindo-se os VVs em posição de acento frasal, obtemos menor desvio-padrão com o aumento da TE para as unidades VV átonas em comparação com as tônicas. Comparando estas tabelas com a tabela 3.4, obtivemos apenas um caso em que tal hipótese não foi confirmada (frase 9). Esse resultado é explicado pela maior sobreposição gestual para os gestos desta frase, bem como para os gestos das frases 6 e 8 (todos três marcados de vermelho nas tabelas 3.5 e 3.6). Essa sobreposição gestual, impossibilitando a divisão dos VVs acústicos, gera unidades muito grandes, as quais, conseqüentemente, terão valores duracionais maiores do que os VVs sem sobreposição gestual. Observando a tabela 3.6, verificamos que estas três frases tiveram os menores valores de significância estatística de todo o *corpus*.

Havendo-se comentado o desvio-padrão da duração dos VVs, passemos à análise do desvio-padrão da duração do grupo acentual, análise muito mais complexa do que a primeira. Em primeiro lugar, não foi possível fazer tal cômputo para todas as frases, pois, devido às suas extensões, as frases 3 (taxa rápida), 6 (taxas normal e rápida), 9 (taxas normal e rápida) e 10 (taxas normal e rápida) apresentaram apenas um GA nas taxas mais elevadas (vide tabela 3.9). Após excluí-las dessa análise, observamos dois processos distintos para as outras frases, conforme análise das tabelas 3.7 e 3.8: 1) valor decrescente da taxa lenta para a rápida (frases 1, 4, 5 e 11); 2) valor ascen-

dente da lenta para a normal (frases 2, 7 e 8). Para entendermos o porquê de não havermos encontrado uma diminuição do desvio-padrão ao se aumentar a TE para as demais frases, observemos na tabela 3.9 um valor que chamamos de desvio-padrão do número de VVs/GA intra-taxas (DVV/GA_{it}), ou seja, o desvio-padrão de VV/GA para cada uma das taxas isoladamente.

Tabela 3.7: Média e desvio-padrão da duração dos grupos acentuais (μ, σ) em três taxas de elocução: lenta (L), normal (N) e rápida (R). Itens em azul representam maior desvio-padrão com o aumento da TE; em preto, o oposto; e em vermelho, apenas 1 GA por repetição da sentença.

Frases	Taxas		
	L	N	R
1	1051 (300)	809 (174)	788 (169)
2	858 (200)	1353 (320)	1226 (321)
3	1094 (365)	1133 (876)	1733 (90)
4	912 (368)	1026 (249)	845 (172)
5	1944 (901)	1091 (516)	924 (362)
6	897 (224)	1153 (43)	991 (79)
7	1287 (42)	1199 (495)	939 (326)
8	818 (357)	982 (881)	891 (758)
9	1220 (472)	1807 (36)	1542 (53)
10	819 (498)	1206 (42)	973 (27)
11	1201 (98)	962 (54)	806 (57)

Tabela 3.8: Análises estatísticas da duração e do desvio-padrão dos grupos acentuais em três taxas de elocução: lenta (L), normal (N) e rápida (N). Dur GA representa a duração do grupo acentual e DP GA, o desvio-padrão da duração do GA.

Frases	Análises estatísticas ($\alpha = 0,5$)			
	Anova (Dur GA)	p <	Anova (DP GA)	p <
1	F (2,55) = 9,8390	0,0003	F (2,13) = 9,7256	0.003
2	F (2,45) = 16,676	10^{-5}	n.s.	n.s.
3	n.s.	n.s.	não se aplica (n.s.a.)	n.s.a.
4	n.s.	n.s.	F (2,7) = 49,814	0.00008
5	F (2,27) = 13,312	0,0001	F (2,12) = 112,76	10^{-5}
6	F (2,42) = 7,4013	0,002	n.s.a.	n.s.a.
7	F (2,35) = 6,4825	0,005	F (2,14) = 57,637	10^{-5}
8	n.s.	n.s.	F (2,25) = 387,67	10^{-4}
9	F (2,19) = 9,0147	0,002	n.s.a.	n.s.a.
10	F (2,30) = 6,9431	0,0034	n.s.a.	n.s.a.
11	F (2,39) = 129,31	10^{-4}	F (2,18) = 3,6690	0,046

A análise da tabela 3.9 nos revela que, para as frases em que a diminuição do desvio-padrão foi verificada, houve uma diminuição do DVV/GA_{it} (frase 5) ou uma invariância estatística desse valor (frases 1, 4 e 11), confirmada pela ANOVA Kruskal-Wallis. Por outro lado, houve um aumento do DVV/GA_{it} para as demais frases (2, 7, 8). Dessa forma, refinando nossa hipótese de diminuição do desvio-padrão da duração do GA com o aumento da TE, acrescentamos que, para que tal fato ocorra, o número de VV/GA não deve ser demasiadamente diferente, ou seja, o DVV/GA_{it} deve diminuir ou permanecer constante com o aumento da TE. Do contrário não encontraremos o padrão esperado. Na figura 3.21, temos um exemplo da diminuição

Tabela 3.9: DVV/GA_{ig} para as frases com desvio-padrão da duração do GA em todas as taxas. Números como 3-3-7-3 representam o padrão de VV/GA mais comum para cada TE e o número entre parênteses, o DVV/GA_{it} . L representa a taxa lenta; N, taxa normal; R, taxa rápida; K-W, análise não-paramétrica Kruskal-Wallis; p, o nível de significância da amostra. VVs em azul representam maior desvio-padrão com o aumento da TE, os VV em preto, o oposto e os VV em vermelho, frases sem desvio-padrão de VV/GA em algumas taxas.

Frases	Taxas			Estatísticas	
	L	N	R	K-W	p <
1	3-3-7-3 (2)	3-3-7-3 (2)	6-7-3 (2.1)	n.s.	n.s.
2	3-3-6-3 (1,5)	6-9 (2,1)	6-9 (2,1)	H (2, N=21) = 20.0	10^{-5}
3	2-3-4 (1)	2-8 (4,2)	9	n.s.a.	n.s.a.
4	2-7-4 (2,5)	7-4 (2,1)	7-4 (2,1)	n.s.	n.s.
5	5-9 (2,8)	3-7 (2,8)	3-6 (2,1)	H (2, N=15) = 14.0	0.001
6	2-5 (2,1)	6	5	n.s.a.	n.s.a.
7	5-5-5 (0)	9-5 (2,8)	9-5 (2,8)	H (2, N=17) = 11.9	0.003
8	2-5-4 (1,5)	2-9 (4,9)	2-9 (4,9)	H (2, N=28) = 23.1	10^{-4}
9	3-9 (4,2)	11	10	n.s.a.	n.s.a.
10	6-2 (2,8)	7	7	n.s.a.	n.s.a.
11	6-5 (0.7)	5-5 (0)	5-5 (0)	n.s.	n.s.

do desvio-padrão da duração dos GA com o aumento da TE, ou seja, as durações dos GA se tornam mais próximas em taxas mais elevadas.

Em segundo lugar, comentemos a questão da duração do GA se manter constante com o aumento da TE. Analisando-se as tabelas 3.7 e 3.8, observa-se que três das frases (3, 4, 8) confirmaram plenamente nossa hipótese, sendo que uma análise estatística *post-hoc Scheffé* mostrou que estas frases são indistintas duracionalmente entre as três taxas. Antes de analisar as demais frases, convém salientar que, para que tal hipótese seja efetivamente

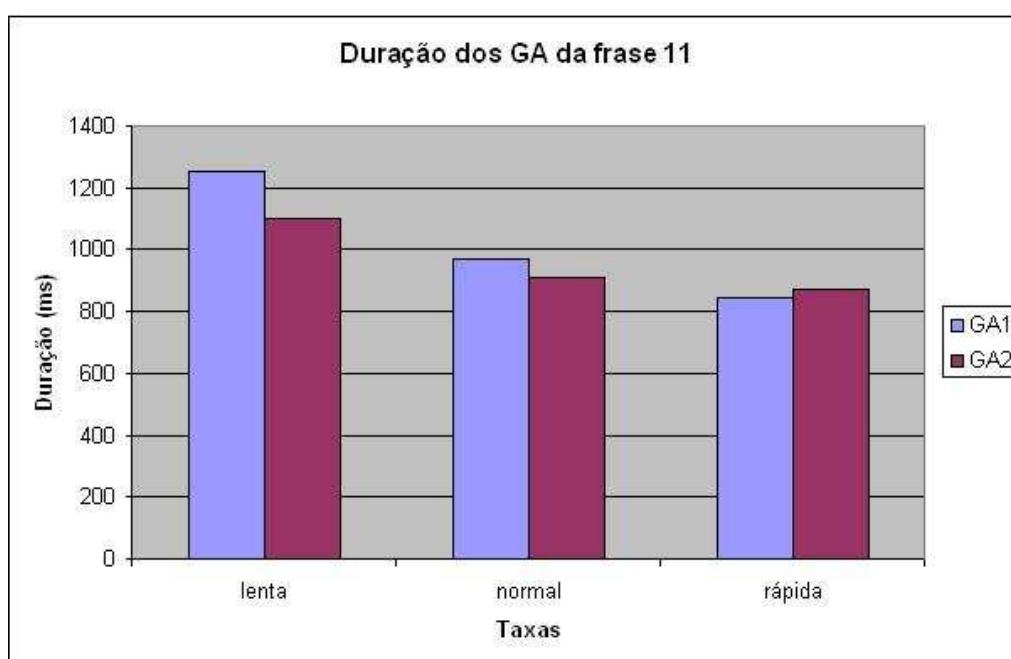


Figura 3.21: Exemplo da diminuição do desvio-padrão dos grupos acentuais com o aumento da taxa de elocução para a frase 11, ou seja, nas taxas mais rápidas a duração dos GA é mais semelhante entre si do que nas taxas lentas.

realizada, reestruturações rítmicas devem, necessariamente, ocorrer entre as taxas, tornando a duração das unidades VV nos GA menor, porém com um número maior de VV/GA, o que as torna estatisticamente idênticas inter-taxas. Dessa forma, as frases 5 e 11 (vide tabela 3.9), por não apresentarem reestruturações rítmicas entre as taxas, não mantiveram constante a duração do GA, conforme a alta significância estatística entre as taxas, maior em todo o corpus (excetuando-se a frase 2, vide Anova (Dur GA) na tabela 3.8).

A análise das tabelas 3.7, 3.8 e 3.9, confirmada através de um teste *post-hoc Scheffé*, ainda nos mostra que as demais frases também confirmam nossa hipótese, mas de uma maneira mais complexa. A frase 1 não manteve constante a duração da taxa lenta para normal (vide tabela 3.9), devido à ausência de reestruturação rítmica de uma taxa para a outra. Porém esta mesma frase manteve constante a duração do GA da taxa normal para a rápida, transição em que houve reestruturação rítmica de 4 GA para 3 GA (vide tabela 3.9). Semelhantemente, a frase 7 não manteve a duração do GA constante da taxa normal para a rápida (vide tabela 3.9), pois não houve reestruturações rítmicas nessa transição. Porém, a duração do GA foi constante entre as taxas lenta e normal, em que houve reestruturação rítmica de 3 GA para 2 GA. As frases 6, 9 e 10 se comportaram de maneira distinta, ou seja, apesar de haver reestruturação rítmica da taxa lenta para a normal (vide tabela 3.9), a duração do GA aumentou da taxa lenta para a normal, porém tal duração

não se diferenciou estatisticamente entre as taxas lenta e rápida. Isso mostra que para estas frases, foi necessário acelerar um pouco mais o andamento para que a duração do GA viesse a ser indistinta entre as taxas. Também para a frase 2 a duração do GA entre as taxas lenta e normal foi diferente, mas a taxa lenta ainda continuou diferente da taxa rápida. No entanto, entendemos que tal diferenciação entre as durações dos GA entre estas taxas ocorreu, pois, conforme confirmam os dados das tabelas 3.7 e 3.8, não houve grande aumento na transição da taxa normal para a rápida, diferentemente do que ocorreu para as frases 6, 9 e 10, elemento necessário para a manutenção da duração constante do GA inter-taxas. No entanto, mais dados são necessários para confirmar essa hipótese, pois, para todas as frases (2, 6, 9 e 10), não houve diferenças estatísticas entre as taxas normal e rápida.

Por fim, mediante análise da tabela 3.10, a hipótese do aumento de VV/GA proporcionalmente ao andamento da fala foi corroborada por nossos dados. Com exceção das frases 5 e 11 de nosso corpus, contrárias à nossa hipótese por não apresentarem reestruturações rítmicas (vide tabela 3.9), todas as demais frases apresentaram um aumento do número de unidades VV/GA com o aumento da TE, sendo que dessas 9 frases restantes, tal aumento foi confirmado estatisticamente pela *ANOVA Kruskal-Wallis* em 5 frases (2, 3, 6, 9 e 10).

Tabela 3.10: VV/GA (média, desvio-padrão) entre três taxas de elocução (lenta (L), normal (N) e rápida (R)), para todas as frases, com suas respectivas significâncias estatísticas através de uma *Anova Kruskal-Wallis* (K-W). Itens em azul representam maior desvio-padrão com o aumento da TE; em preto, o oposto; e em vermelho, apenas 1 GA por repetição da sentença.

Frases	Taxas			Estatísticas	
	L	N	R	K-W (VV/GA)	p <
1	4 (2)	4 (2)	5,3 (2)	n.s.	n.s.
2	3,8 (1,5)	7,5 (2,1)	7,5 (2,1)	H (2, N=48) = 23,62051	10^{-4}
3	3 (1)	5 (4,2)	9 (0)	H (2, N=19) = 8,600279	0,014
4	4,1 (2,1)	5,5 (2,1)	5,5 (2,1)	n.s.	n.s.
5	7 (2,8)	5 (2,8)	4,5 (2,1)	n.s.	n.s.
6	3,5 (2,1)	5,7 (0,7)	5 (0)	H (2, N=45) = 20,41318	10^{-4}
7	5,2 (0,3)	6,9 (2,9)	6,4 (2,9)	n.s.	n.s.
8	3,7 (1,5)	5,5 (4,9)	5,2 (4,5)	n.s.	n.s.
9	5,9 (4,1)	11,2 (0,4)	9,8 (0,8)	H (2, N=22) = 15,75148	0,0005
10	3,9 (2,7)	7 (0)	7 (0)	H (2, N=33) = 29,54934	10^{-4}
11	5,2 (0,2)	5 (0)	5 (0)	n.s.	n.s.

3.2.10 Conclusão

Concluindo, o experimento acústico complementar acima relatado serviu para corroborar hipóteses previamente traçadas no experimento acústico preliminar e inaugurar outras, tais como:

- O número de VV/GA aumenta proporcionalmente ao andamento da fala se ocorrerem reestruturações rítmicas com o aumento da TE, apesar dos processos de elisão;
- A duração do GA se mantém constante com o aumento da TE se ocor-

rerem reestruturações rítmicas com o aumento da TE;

- Fronteiras sintáticas fortes, em princípio, inibem reestruturações rítmicas prosódicas;
- Há três possíveis atratores para as proeminências frasais: 1) atrator lexical; 2) atrator sintático; 3) atrator semântico-pragmático. Os três atratores atuam no nível lingüístico mais elevado do MDR, acoplando-se ao oscilador acentual;
- O desvio-padrão da duração das unidades VV, bem como da duração dos GA é menor nas taxas rápidas (o que implica em uma maior sensação de isocronismo nestas taxas);
- Derivado dos resultados de menor desvio-padrão com o aumento da TE em nosso corpus, o aumento da taxa exacerba o caráter misto do ritmo em português, ou seja, tendências tanto ao ritmo silábico quanto ao acentual.

Capítulo 4

Estudos articulatórios

Os estudos articulatórios relatados nesta tese são oriundos de uma pesquisa de doutorado-sanduíche realizada na University of Southern California (USC), com apoio do CNPq (processo 200199/2004-8), sob a orientação da Profa. Dra. Dani Byrd. Esta pesquisa, intitulada “Investigação do grau de coarticulação dentro (em torno) de fronteiras maiores e menores em PB de acordo com o aumento da taxa de elocução”, tinha como objetivo principal observar como o aumento da TE modifica, dinamicamente, os padrões articulatórios envolvidos na estruturação rítmica da fala. Os resultados desta pesquisa complementam, articulatoriamente, os estudos acústicos realizados sob uma perspectiva dinamicista aqui relatados (vide também Meireles & Barbosa, submetido, cujos resultados serão detalhados no próximo capítulo). Com esse intuito, realizamos, pois, dois experimentos: 1) reestruturação rít-

mica ao nível da fronteira do GA; 2) reestruturações lexicais em palavras proparoxítonas. O primeiro experimento será relatado a seguir e o segundo na terceira parte desta tese.

Pretendíamos utilizar neste experimento as mesmas sentenças utilizadas no estudo acústico complementar anterior, porém percebemos que, devido à enorme variabilidade articulatória nesses dados, não poderíamos utilizar nem as mesmas sentenças nem as mesmas palavras usadas em estudos acústicos anteriores (estudo acústico complementar; Meireles & Barbosa, submetido).

Como exemplo desta impossibilidade ou impraticabilidade, cito a sentença 1 do estudo acústico complementar: “Há três tipos de abóbora em Belo Horizonte”. Comentando sobre as vogais, observamos que, para esta sentença, temos, dentre outras, [a, e, i, u, ɔ, o, ã, õ], ou seja [a] = vogal posterior-faringal, [e, i] vogais anteriores, [u, ɔ, o] = vogais posteriores arredondadas, [ã, õ] = vogais nasais. Para trabalhar com essa sentença, teríamos, pois, no mínimo, três sensores para a língua: um na parte posterior-faringal, outro na parte posterior, e outro na parte anterior da língua. Além disso, precisaríamos de um sensor no véu palatino, local problemático devido à ocorrência de ânsia de vômito.

Tendo em vista a necessidade de se ter um corpus bem controlado articulatoriamente, a fim de se evitar ruídos articulatórios posteriores, elaboramos o desenho experimental descrito na próxima seção.

4.1 Metodologia

O presente estudo examina a influência da taxa de elocução no deslocamento mandibular em fronteiras frasais do PB. A figura 4.1 mostra uma movimentação hipotética da mandíbula com o aumento da TE. Espera-se que proemiências frasais ao longo do enunciado sejam reestruturadas com tal aumento, e, conseqüentemente, as vogais acentuadas em uma fronteira frasal sejam realizadas com menor abertura da mandíbula em taxas rápidas. Além disso, hipotetizamos não haver reinício do movimento mandibular após fronteiras prosódicas fracas nas taxas rápidas, ou seja, diferentemente das taxas lentas, em que há reinício de movimento mandibular após uma fronteira fraca, nas taxas rápidas não haverá tal reinício nestas fronteiras.

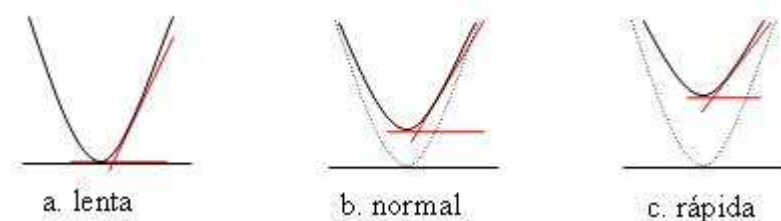


Figura 4.1: Apresentação esquemática mostrando, hipoteticamente, como o deslocamento mandibular (eixo-y) pode se modificar com o aumento da TE. Aqui estão representadas três TE: a) lenta; b) normal; e c) rápida. A linha contínua representa um suposto alvo para uma abertura de vogal plena acentuada. A linha pontilhada representa o gesto na taxa lenta (b e c), e a linha escura representa o gesto previsto. Além disso, da taxa lenta para a rápida, observamos, no cruzamento entre as duas setas vermelhas, uma diminuição da velocidade de pico (vide seção “Cálculo aplicado no Mavis: derivadas”).

Uma falante nativa do PB, 28 anos, foi gravada acústica (taxa de amostragem: 20,05 kHz) e articulatoriamente no Laboratório de Fonética da USC. A falante foi paga por sua participação no experimento e assinou um documento do National Institute of Health (NIH) explicando o propósito do experimento. Um magnetômetro 2-D AG-200 (www.articulograph.de) (cf. Perkel et al. 1992, sistema magnetômetro EMMA) foi usado para rastrear o movimento mandibular. Os dados de movimentação foram amostrados em 200 Hz, com correção de cabeça, deslocados ao plano de oclusão, e filtrados com passa-baixas em 25 Hz. Sensores foram colocados nos seguintes articuladores: língua (perto da região palatal), lábios inferiores, mandíbula (nos dentes incisivos inferiores). Dois outros sensores foram usados como referência para o sistema de aquisição do sinal: um no nariz (entre as duas sobrancelhas) e um na superfície central anterior dos incisivos maxilares. Apenas o movimento no eixo-y da mandíbula foi analisado neste experimento. Como em Erickson (2004), a abertura mandibular foi medida “in terms of the lowest vertical position of the mandibular pellet in the syllable from the maxillary occlusal plane” (p. c-134).

4.1.1 O magnetômetro (EMMA)

Como utilizamos o magnetômetro (ou articulógrafo) em todos os estudos articulatórios relatados nesta tese, explicaremos aqui como o aparelho funciona para a aquisição de dados articulatórios. Para uma informação técnica sucinta de como o magnetômetro funciona, citemos o resumo de MARINO (2002) abaixo:

Esta técnica, conhecida como articulografia eletromagnética (electromagnetic articulography EMA), utiliza dispositivos de campo magnético alternativo que permitem monitorar movimentos de pontos onde os sensores se encontram fixados nos articuladores em relação ao tempo. Dentre os sistemas comercialmente disponíveis no mercado têm-se o Articulógrafo AG 100 (Cartens, Alemanha). Este sistema é constituído por três bobinas geradoras que são fixadas em um capacete especialmente construído para esse fim. Quando dispostas em posição fixa ao longo do plano médio-sagital, cada uma das bobinas gera seu próprio campo magnético alternativo de alta frequência. Um conjunto de pequenas bobinas receptoras são aderidas temporariamente nos articuladores de interesse (lábios, língua, mandíbula, e/ou véu palatino) por meio de uma cola especial. Como em um transformador, os campos

magnéticos alternativos das bobinas geradoras induzem tensão elétrica nas bobinas receptoras. A tensão elétrica induzida em cada uma das bobinas receptoras é inversamente proporcional ao cubo de sua distância em relação às bobinas geradoras. Tais bobinas operam em frequências diferentes e geram campos magnéticos que se sobrepõem, portanto, o sistema é capaz de captar, simultaneamente, a distância entre cada bobina receptora e as bobinas geradoras (Baken e Orlikoff, 2000).

Além do Articulógrafo AG 100 citado por Marino (2002), ainda existem no mercado o Articulógrafo 2-D AG 200, e mais recentemente foi desenvolvido o Articulógrafo 3-D AG 500, com o qual pode-se extrair informações articulatórias em três dimensões. Informações detalhadas do produto podem ser obtidas no seguinte *website*: <http://www.articulograph.de/>.

Devido ao seu custo e aplicação específica no mercado, somente encontramos esse equipamento em algumas universidades ao redor do mundo. Mesmo nos Estados Unidos, encontramos esse equipamento em poucas universidades, tais como: *University of Southern California*, *University of California at Los Angeles*, *University of Connecticut*, *Haskins Laboratories*, *University of Texas at Dallas*. Para outras universidades do mundo que utilizam o sistema da Cartens, visitar <http://www.articulograph.de/user.htm>. No Brasil

não há nenhuma universidade que possua tal equipamento.

Seguem abaixo fotos de um articulógrafo do mesmo modelo usado em nosso experimento com vários sensores colocados na língua, nariz, e mandíbula. Tais fotos foram obtidas nos endereços www.phon.ox.ac.uk/~paula/emahome.html e <http://www.advance.uconn.edu/2004/041025/04102506.htm>.



Figura 4.2: Apresentação visual do magnetômetro utilizado em nossos experimentos.

Nesta tese, os dados gerados pelo magnetômetro foram analisados no programa MAVIS. Na seção seguinte faremos uma breve descrição desse programa e de como são feitas análises articulatórias no mesmo.

4.1.2 Análises articulatórias no MAVIS

Os dados articulatórios obtidos com o uso do magnetômetro não estavam prontos para análise. Para isso, deveríamos convertê-los, através de rotinas no programa Matlab. Antes desta conversão, porém, a fim de facilitar nossa análise, elaboramos ainda um *script* da linguagem de programação Perl (apêndice A.4), o qual renomeia os arquivos brutos originais. Após este processo, obtivemos os arquivos prontos para análise. A seguir, com o intuito de elucidar os métodos de análise utilizados, faremos uma descrição sucinta do Mavis.

O programa MAVIS

O programa foi elaborado pelo Dr. Mark Tiede (atual cientista sênior do *Haskins Laboratories*) no *ATR Research Laboratories* em Kioto, Japão. Como o próprio nome diz, MAVIS (*Multiple Articulator Visualizer*) é um visualizador de múltiplos articuladores em simultaneidade. Em outras palavras, tem a função de visualizar vários articuladores ao mesmo tempo, como ponta de língua, dorso de língua, mandíbula, etc.

Como consta nas informações do programa, MAVIS é um conjunto de funções integradas do Matlab úteis para visualizar dados de movimentos sincronizados, como os gerados por métodos de rastreamento da fonte de um

ponto, como o magnetômetro (EMMA), OPTOTRAK, ou sistema de microfios de raios-x. Provê apresentação simultânea de trajetórias variantes no tempo e sua distribuição espacial em um contrabalanceamento temporal, que pode ser incrementado automaticamente para vários efeitos de animação. MAVIS também permite a apresentação de informações estáticas como o traçado do palato ou dados de imagem, e fornece suporte para procedimentos tanto internos quanto requeridos pelo usuário (vide figura 4.3).

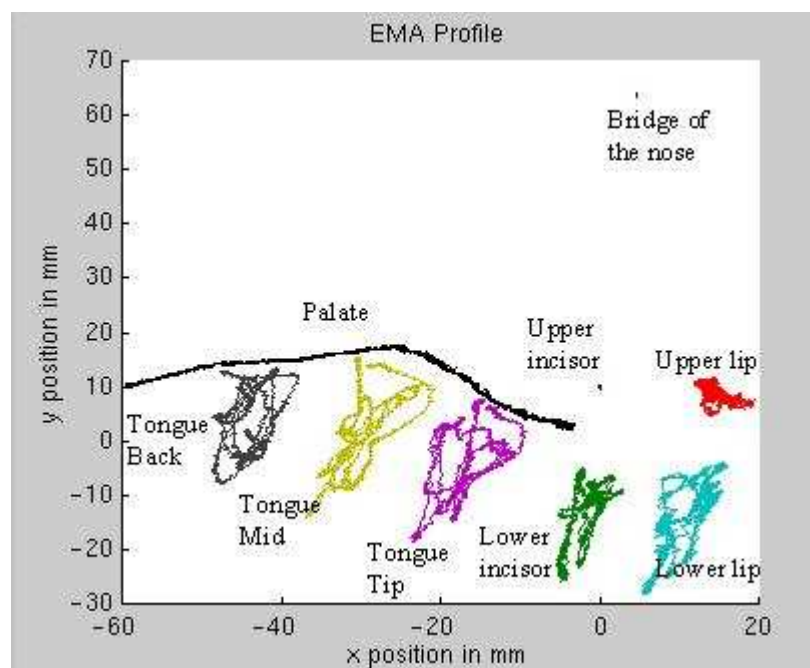


Figura 4.3: Representação de vários traçados possíveis com o magnetômetro: referência do nariz, palato, dentes incisivos superiores e inferiores, lábios inferiores e superiores, ponta de língua, parte central da língua, dorso da língua (www.phon.ox.ac.uk/~paula/emahome.html).

Para uma apresentação sucinta das funcionalidades do MAVIS, analise-

mos a figura 4.4 abaixo, obtida quando digitamos “mavis” dentro do programa MATLAB. Observamos que, o programa MAVIS, por não ser nativo ao MATLAB, deve estar em dos *PATHs* do MATLAB para que tal comando funcione.

Na janela maior acima, temos de cima pra baixo: a) onda sonora completa; b) onda sonora selecionada pelo usuário; c) movimento articulatório do corpo da língua; d) velocidade (eixo vertical) de deslocamento do corpo da língua; e) movimento articulatório da mandíbula; f) velocidade (eixo vertical) de deslocamento da mandíbula; g) movimento articulatório dos lábios inferiores; h) velocidade (eixo vertical) de deslocamento dos lábios inferiores.

Nas duas janelas à direita, temos: 1) parte superior (janela SPATIAL): disposição cartesiana dos três articuladores empregados no experimento 1: a) lábios (amarelo); b) mandíbula (roxo); c) corpo da língua (azul claro); 2) parte inferior (janela MAVIS): informações temporais de posicionamento dos gestos articulatórios (1 na figura 4.4), nome do arquivo aberto (2 na figura 4.4), e recursos adicionais (3 na figura 4.4), como “Events”, “Derive”, que serão explicados na próxima seção.

Etapas das análises articulatórias no MAVIS

Na análise dos dados, utilizamos os procedimentos descritos a seguir. Primeiramente, abrimos todos os arquivos no Mavis e selecionamos apenas a

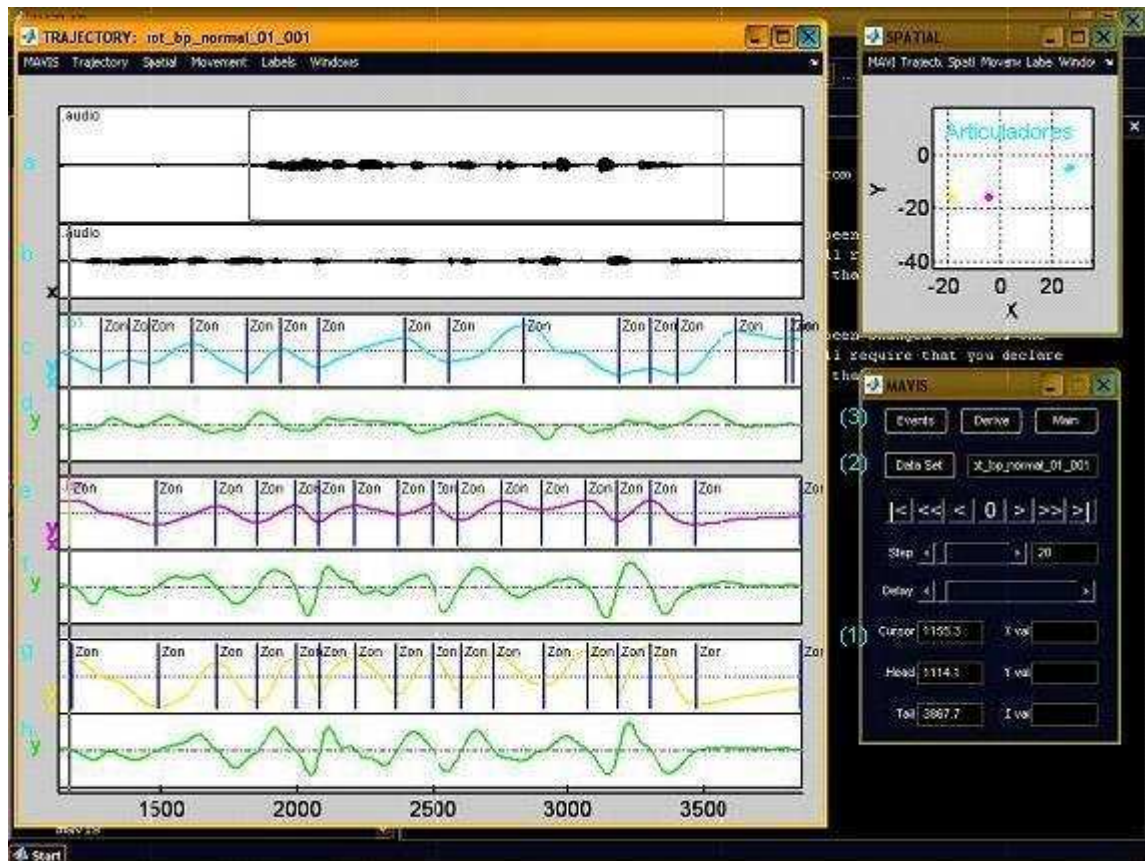


Figura 4.4: Janela do MAVIS com todas suas funcionalidades aberta no MATLAB: a) onda sonora completa; b) onda sonora selecionada pelo usuário; c) movimento articulatório do corpo da língua; d) velocidade (eixo vertical) de deslocamento do corpo da língua; e) movimento articulatório da mandíbula; f) velocidade (eixo vertical) de deslocamento da mandíbula; g) movimento articulatório dos lábios inferiores; h) velocidade (eixo vertical) de deslocamento dos lábios inferiores. Na janela SPATIAL, à direita em cima, temos os articuladores lábios (amarelo), mandíbula (roxo) e corpo da língua (azul claro). Na janela MAVIS, à direita embaixo, temos informações temporais e espaciais sobre os gestos.

sentença que seria analisada (janela maior, parte b da figura 4.4). Esse procedimento era necessário, pois o arquivo final gerado pelo magnetômetro consistia de partes pré-sentença e pós-sentença, as quais também continham movimentos articulatórios que poderiam confundir nossa análise.

Em segundo lugar, para cada uma das sentenças do *corpus*, foram modificados alguns parâmetros em “Trajectory -> Configure Trajectory Display” (segundo menu da janela maior (TRAJECTORY: rot_bp_normal_01_001), figura 4.4). Como necessitávamos apenas dos articuladores mandíbula, corpo de língua e lábios inferiores, os demais articuladores foram eliminados da janela de análise.

Em terceiro lugar, com base na movimentação dos articuladores, a qual pode ser feita, dentre outras maneiras, na janela “Spatial -> Movement -> Move backwards (ou Move forwards)” (terceiro menu da janela SPATIAL, figura 4.4), observamos que a movimentação principal no nosso corpus estava no eixo vertical, principalmente pelos movimentos da mandíbula e lábios (como veremos adiante, utilizamos vogais altas e baixas durante todo o enunciado). Sendo assim, escolhemos, para aparecer na janela principal apenas informações de deslocamento, velocidade e aceleração¹ no eixo y, juntamente com a informação da onda sonora. Após esta etapa ficamos com a janela da forma apresentada na figura 4.5.

¹A aceleração, contudo, não foi utilizada nas análises articulatórias posteriores.

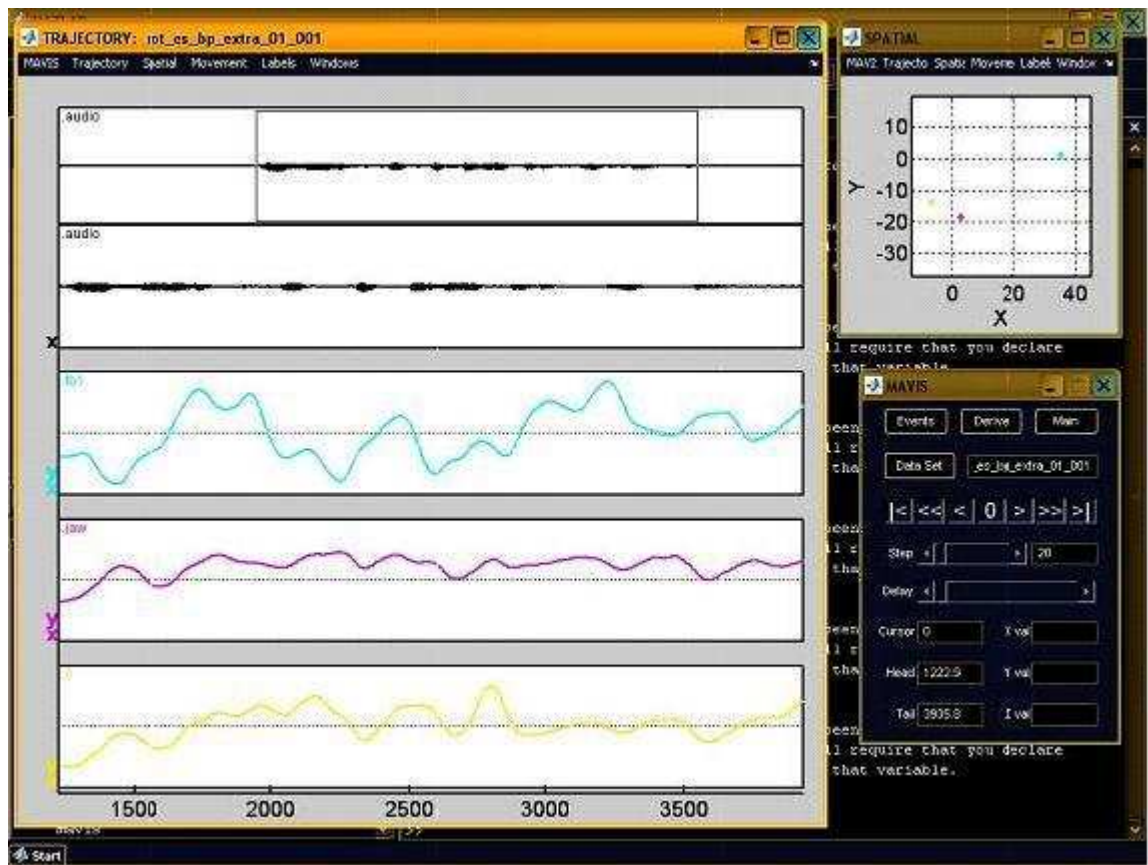


Figura 4.5: Janela Mavis preparada para marcação dos gestos articulatórios. Na janela principal, a linha amarela representa o movimento articulatório dos lábios inferiores; a linha roxa, o movimento articulatório da mandíbula; e a linha azul claro, o movimento articulatório do corpo da língua.

Em quarto lugar, fizemos marcações no sinal articulatório com base na informação de velocidade dos articuladores. Para isto derivamos sinais de velocidade no eixo y para todos os três articuladores. Isto foi feito clicando no botão “Derive” da janela “Mavis” (3 da figura 4.4) e depois em “Start” da janela “Mavis Signal Derivation”², obtendo-se o sinal derivado da figura 4.6.

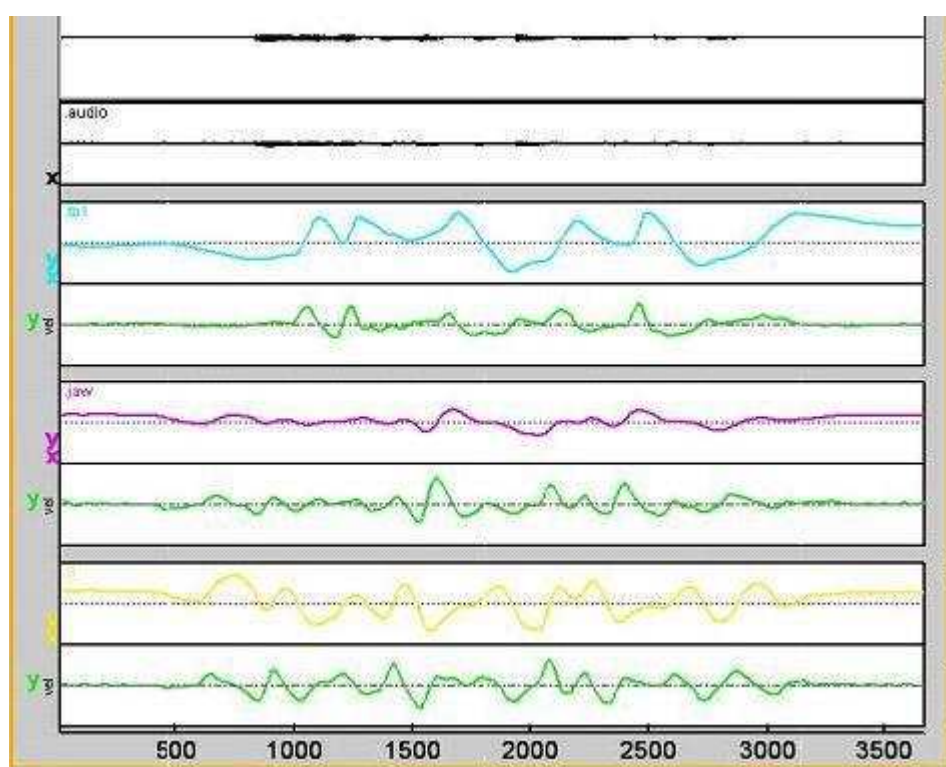


Figura 4.6: Sinal velocidade no eixo y (roxo), derivado e apresentado na janela principal do MAVIS. A linha amarela representa o movimento articulatório dos lábios inferiores; a linha roxa, o movimento articulatório da mandíbula; e a linha azul claro, o movimento articulatório do corpo da língua. As linhas verdes abaixo de cada gesto correspondem à velocidade no eixo vertical para os respectivos articuladores.

²Esta janela aparece quando clicamos no botão “Derive”.

Em quinto lugar, após solicitarmos manualmente as derivações de velocidade, para todas as 120 sentenças de nosso corpus, fizemos a demarcação dos gestos articulatórios. Para isto utilizamos a janela “Events” do MAVIS³. Nesta janela podíamos escolher qual articulador marcar, qual componente, e também qual evento. Nesta etapa, marcamos apenas os zeros da velocidade y para todos os articuladores, resultando nas marcações articulatórias apresentadas na figura 4.7.

Como podemos observar na figura 4.7, cada vez que a velocidade no eixo y cruza a linha cinza (velocidade zero) temos uma marcação. A velocidade zero aqui, relembrando conceitos básicos de Física, representa o momento exato onde o articulador muda de direção no eixo y. Por exemplo, para a mandíbula, representa o momento em que a mandíbula começa a mudar o movimento de descida para subida e o de subida para descida.

A fim de compreendermos como o processo de derivação, base para a marcação dos gestos articulatórios, é feito no Mavis, apresentamos na próxima seção alguns conceitos básicos de física e matemática.

Cálculo aplicado no MAVIS: derivadas

Conforme os fundamentos teóricos de “Cálculo I”, a inclinação de uma tangente a uma curva corresponde à derivada naquele ponto. Recordando a

³Esta janela surge ao clicarmos no botão “Events” (3 da figura 4.4).



Figura 4.7: Exemplo de marcação do sinal articutório no MAVIS. A linha amarela representa o movimento articutório dos lábios inferiores; a linha roxa, o movimento articutório da mandíbula; e a linha azul claro, o movimento articutório do corpo da língua. As linhas verdes abaixo de cada gesto correspondem à velocidade no eixo vertical para os respectivos articuladores. A linha azul escuro representa o momento em que a velocidade no eixo vertical é zero, ou seja, nos cruzamentos da linha verde com a linha pontilhada horizontal.

definição de velocidade instantânea da física, temos:

$$v = \frac{\Delta d}{\Delta t}, \quad (4.1)$$

em que v = velocidade instantânea, Δd = distância percorrida, e Δt = tempo decorrido.

Sem entrar em maiores detalhes, esse quociente em cálculo é chamado de taxa de mudança, que é igual à inclinação de uma reta tangente a uma curva (derivada naquele ponto). Em outras palavras, em nosso caso, tal inclinação mostra o quanto a distância muda em função do tempo ponto a ponto. Tal medida nos dá a velocidade instantânea de movimentação do gesto.

Para observar como o conceito de derivada é utilizado no MAVIS, observemos a figura 4.8. Nesta figura temos em azul escuro retas tangentes à curva de movimentação do gesto de mandíbula na palavra “pavê”. Como dissemos, a inclinação desta reta nos dá a velocidade instantânea do gesto, que nos dá informações importantes como velocidade máxima, mínima e zero do gesto, como vemos na figura 4.9.

Ao se fazer derivadas ponto a ponto no sinal de deslocamento da mandíbula, saber-se-á extrair informação, por exemplo, da velocidade zero, como medimos anteriormente. Dessa forma, se escolhermos derivação de velocidade zero para o programa MAVIS, após a computação ponto a ponto, o

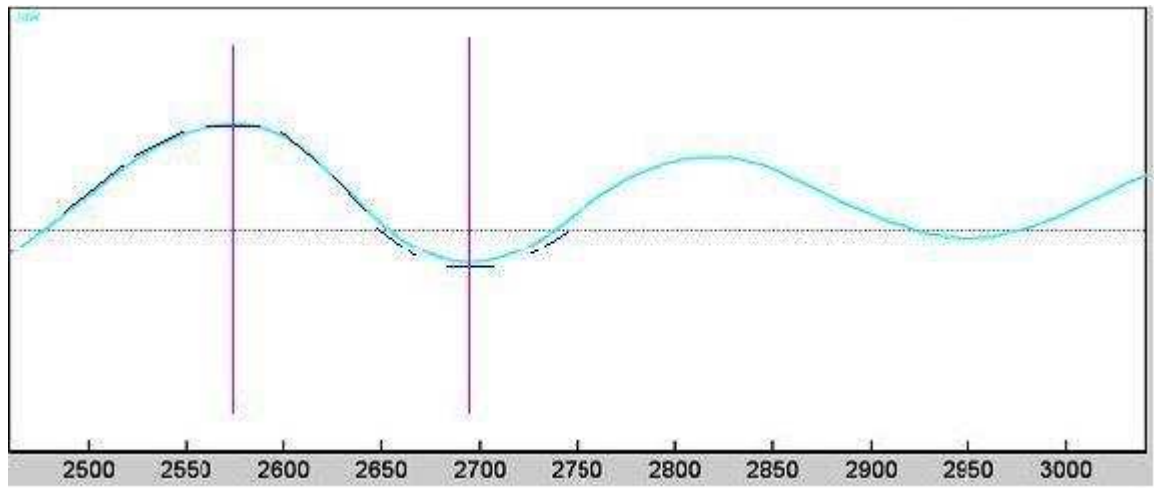


Figura 4.8: Deslocamento do gesto de mandíbula na palavra “pavê” na taxa lenta (traçado azul claro) e suas derivadas ponto-a-ponto (traçado azul escuro).



Figura 4.9: Inclinação 1 > Inclinação 2 > Inclinação 3. Sendo assim, para velocidade instantânea, temos velocidade 1 (positiva) > velocidade 2 (zero) > velocidade 3 (negativa).

programa marcará onde no sinal temos tais velocidades, como observamos na figura 4.7, em que temos marcados de amarelo pontos na curva em que a velocidade de deslocamento é zero.

Entendendo melhor como as derivadas funcionam⁴, podemos notar que para delimitarmos articulatoriamente a palavra “pavê” da figura 4.10, por exemplo, deveremos ter informações não somente da velocidade zero (correspondente, aproximadamente, ao ponto de máxima excursão do gesto), mas também de picos (velocidade máxima) e vales (velocidade mínima) dos mesmos. Com essas informações podemos delimitar o gesto, como fizemos na figura 4.10.

⁴Para informação de aceleração instantânea, usamos o mesmo procedimento detalhado para velocidade, ou seja, como tal aceleração corresponde a $\Delta v/\Delta t$, fazemos derivadas ponto a ponto na curva de velocidade do gesto em função do tempo.

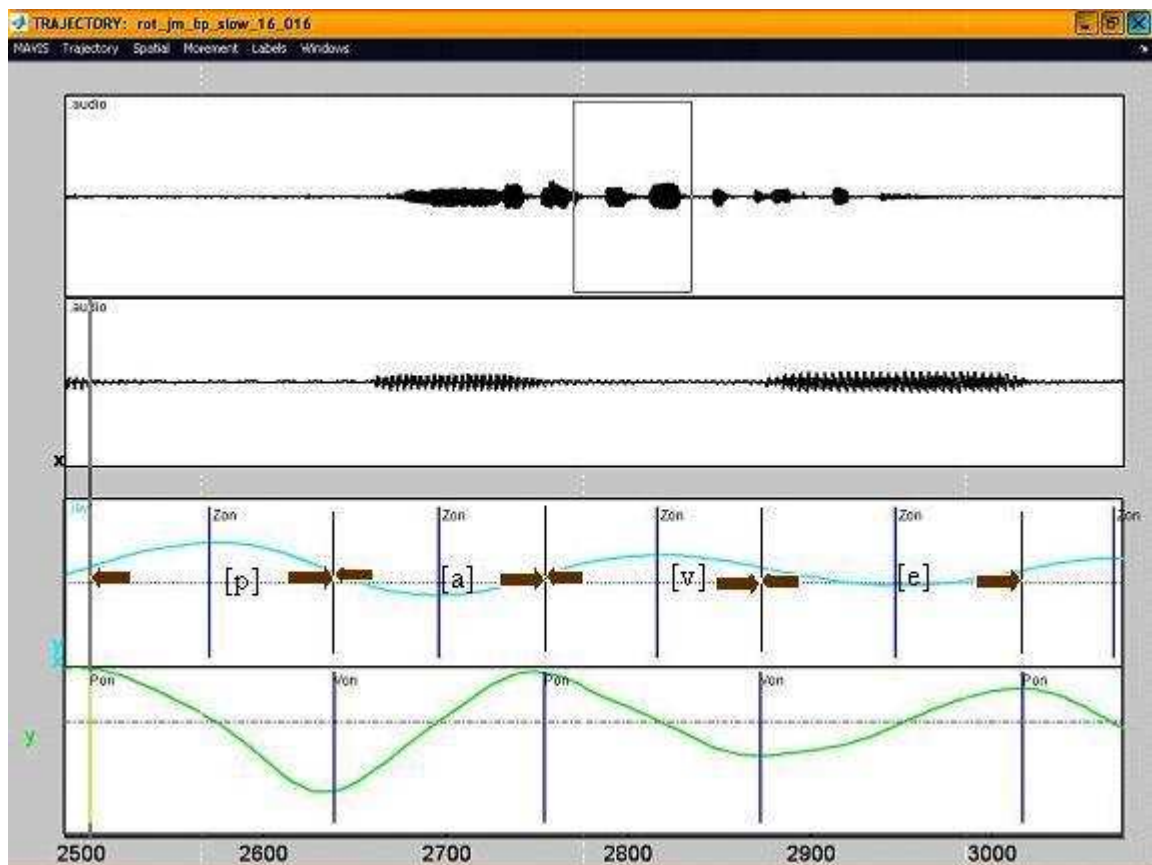


Figura 4.10: Delimitação dos gestos articulatórios para a palavra “pavê” medidas entre intervalos de velocidade no eixo vertical (y, traçado verde), conforme podemos ver entre as setas marrons.

Etapas das análises articulatórias no MAVIS: continuação

Retornando às etapas das análise articulatórias no Mavis, por fim, após delimitarmos todos os gestos articulatórios do *corpus*, salvamos seus respectivos valores e obtínhamos como resultado um arquivo .txt pronto para análise em planilhas eletrônicas como o *Excel*. Um exemplo de tal arquivo é apresentado na tabela 4.1. Terminado o trabalho no Mavis, ainda realizamos alguns procedimentos metodológicos adicionais, relatados abaixo.

Tabela 4.1: : Informações temporais (zeros da velocidade no eixo y) e espaciais (informações do deslocamento no eixo vertical) do gesto mandíbula, juntamente com a referência segmental.

SIGNAL 3:	rot-jm-bp-slow-16-016.jaw.VEL(:,2)	
Zero (time)	Info (displacement on the vertical axis)	
1590.39	-17.0645	
2005.53	-26.298	
2285.625	-17.7774	
2410.667	-22.4168	
2570.721	-17.2356	p
2695.763	-21.5551	a
2815.803	-18.2515	v
2945.847	-20.7704	e
3065.888	-18.533	
3180.926	-22.2437	
3335.979	-17.2551	
3481.028	-18.5296	
3666.09	-16.0394	p
3796.134	-21.3404	a
3906.171	-19.2303	p
4061.223	-24.4231	a

Em nossas análises utilizamos apenas dados com movimentos articulatórios mais precisos, sem pequenas perturbações dos articuladores. Para isto, foram observadas cada uma das 120 sentenças (4 sentenças x 3 taxas x 10 repetições), e anotadas uma a uma quais seriam as dificuldades para analisar cada uma destas sentenças. Além disso, observamos, para este experimento, qual articulador seria mais adequado para a análise dos dados, procedimento necessário devido às várias possibilidades combinatórias de dados articulatórios.

Na análise dessas sentenças, verificamos que, em geral, os movimentos de mandíbula e lábios inferiores eram bastante sincronizados (em fase), e que os dados de corpo de língua, muitas vezes eram imprecisos. Vale a pena lembrar que, de acordo com nosso desenho experimental, a movimentação da mandíbula e lábios inferiores estava bem controlada através da alternância de vogais altas e baixas durante todo enunciado. Por outro lado, o gesto articulatório para a vogal não estava bem controlado, pois para uma vogal como [a] podia haver uma articulação bem faringal, que não seria muito bem capturada pela posição do sensor na parte central do corpo da língua.

Com a informação de quais sentenças se comportavam melhor articulatoriamente e com a informação de maior controle para os gestos de mandíbula e lábios inferiores, optamos por analisar apenas dados de mandíbula. Lembremos que a análise dos lábios inferiores seria praticamente idêntica à da

mandíbula devido ao sincronismo de fase.

Tendo sido feita uma descrição sucinta dos equipamentos e programas utilizados para análise de sinais articulatórios, além da justificativa de se analisar apenas dados de mandíbula, passemos, pois, à descrição dos procedimentos metodológicos gerais utilizados neste experimento.

4.1.3 Desenho experimental

As sentenças utilizadas no experimento estão apresentadas na tabela 4.2. Essas sentenças foram desenvolvidas, a fim de se ter um padrão alternante (mandíbula aberta e fechada) durante todo o enunciado. Por essa razão utilizamos uma seqüência de vogais altas (ou média-altas) e vogais baixas do começo ao fim da sentença e também consoantes bilabiais nas fronteiras frasais previstas. Dez repetições de cada sentença (aleatorizadas em blocos) foram gravadas em três taxas de elocução, resultando em um total de 120 sentenças para análise (4 sentenças x 10 repetições x 3 TE).

Para se obter três taxas de elocução distintas, pedimos ao informante que lesse as sentenças de acordo com as seguintes instruções e ordem: (1) normal: fale de uma maneira confortável; (2) lenta: fale o mais lento que puder preservando o significado das sentenças e sem introduzir pausas entre as palavras; (3) rápida: fale o mais rápido que puder sem introduzir distorções

na fala. Exemplos também foram dados ao informante exemplificando cada uma das TE pretendidas.

Tabela 4.2: Sentenças utilizadas no experimento articulatório prosódico. TF significa transcrição fonética.

Sentença 1	Vou lá levar o pavê pra filha do papai.
TF	[vow.la.le.va.rũ.pa.ve.pra.fi.ʎe.dũ.pa.pai]
Sentença 2	Mamãe não quer mais qu'eu babe, mas qu'eu pape.
TF	[mẽ.mẽĩ.nẽũ.kefi.majs.kjew.ba.bi.mas.kjew.pa.pi]
Sentença 3	Ela diz mão de máfia no carro da moça do papai.
TF	[e.le.diz.mẽũ.ɕĩ.ma.fjẽ.nũ.ka.fũ.da.mo.sẽ.dũ.pa.paj]
Sentença 4	Foi bão gostar demais de papai, mas de máfia!?
TF	[foj.bẽũ.gos.tafi.ɕĩ.majz.ɕĩ.pa.paj.maz.ɕĩ.ma.fjẽ]

Conforme vimos na seção anterior, o programa MAVIS (Tiede et al. 1999), modificado na USC, foi empregado para medir o movimento dos sensores da mandíbula nas dimensões vertical (eixo y, utilizada em nossa análise) e horizontal (eixo x) (vide figuras 4.7 e 4.10 acima). As seguintes variáveis articulatórias foram medidas: (a) deslocamento da mandíbula: medido no cruzamento em zeros da velocidade-y; (b) deslocamento da constrição da mandíbula: medido como a diferença entre os cruzamentos em zero no início e no extremo da constrição; (c) duração (relacionada à acústica) do gesto mandíbula: medida pelo intervalo entre velocidades de máxima (positiva) e mínima (negativa); (d) duração do gesto mandíbula: medida entre os cruzamentos em zero da velocidade-y entre o início e o máximo da constrição; (e) velocidades de pico e vale da constrição do gesto mandíbula; (f) duração

da aceleração da mandíbula; (g) tempo-para-velocidade-de-pico proporcional: medido do quociente entre duração da aceleração da mandíbula (eixo y) e duração da formação da constrição. Na figura 4.11 temos uma descrição visual das medidas articulatórias aqui relatadas.

4.1.4 Sessão de gravação

Antes da gravação, a informante foi informada oralmente das regras do NIH para se realizar um experimento. Após a aceitação da informante dos termos do NIH, a mesma primeiro leu as instruções do experimento e, então, como um pequeno treinamento, leu a sentença de um experimento acústico prévio (experimento acústico complementar) em três TE distintas. Depois deste treinamento, um professor doutor da USC, dentista profissional, posicionou os sensores nos locais descritos no início da seção 4.1, e o experimento foi realizado.

4.1.5 Hipóteses e questões

A hipótese a ser investigada aqui é a diminuição do número de GA com o aumento da TE e a não realização do reinício do movimento mandibular nas fronteiras mediais das taxas rápidas. De acordo com esta perspectiva, uma fronteira prosódica fraca é inserida na realização dinâmica de um GA de uma

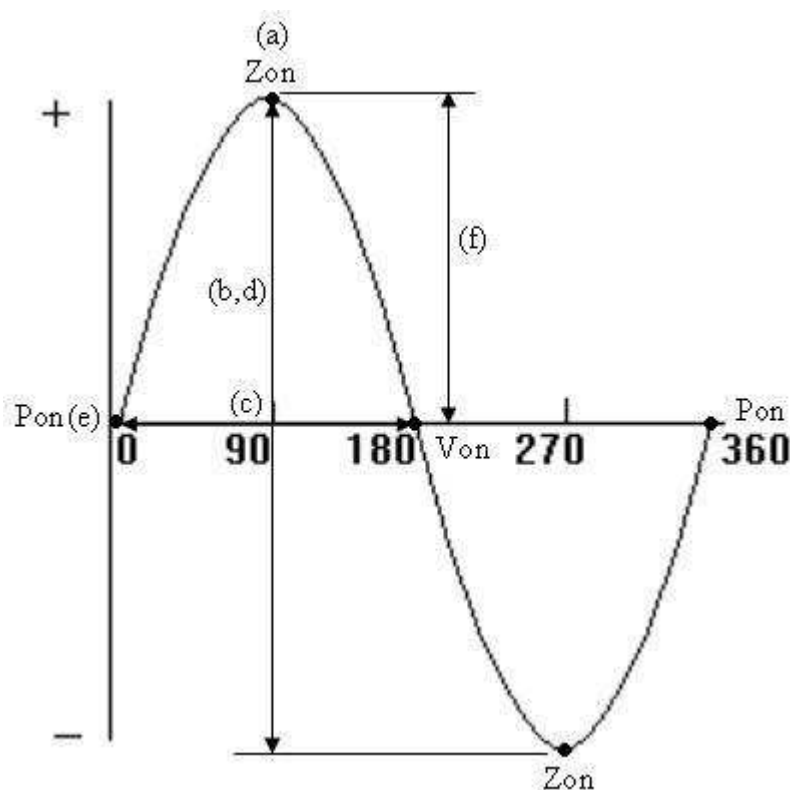


Figura 4.11: Representação abstrata das medidas articulatórias realizadas para o experimento acústico prosódico. Von representa, *onset* do vale de velocidade (velocidade negativa); Pon, *onset* do pico de velocidade (velocidade positiva); e Zon, *onset* dos zeros de velocidade. O intervalo de 0 a 90 graus representa um trecho do movimento de fechamento do gesto até a constrição máxima (a), considerando-se uma consoante. O intervalo de 90 a 270 graus representa a abertura de um gesto consonantal. Temos aqui representados: (a) deslocamento da mandíbula, medido no cruzamento em zeros da velocidade y (valor pontual); (b) deslocamento da constrição da mandíbula, medido como a diferença entre os cruzamentos em zero no início (a) e no extremo (270 graus) da constrição, considerando-se um gesto vocálico; (c) duração (relacionada ao acústico) do gesto mandíbula, medida pelo intervalo entre velocidades máxima (Pon) e mínima (Von); (d) duração do gesto mandíbula, medida entre os cruzamentos em zero da velocidade-y entre o início e o máximo da constrição; (e) velocidades de pico (Pon) e vale (Von) da constrição do gesto mandíbula (valor pontual); (f) duração da aceleração da mandíbula: positiva (270 a 360 graus) e negativa ((a) a 180 graus). O tempo-para-velocidade-de-pico proporcional é medido do quociente entre a duração da aceleração da mandíbula (f) e duração da formação da constrição (d).

fronteira prosódica forte e é apagada em função da restrição do aumento da TE. Ex.: || Subiu a tribuna | de um poleiro de ouro || é reorganizada para || Subiu a tribuna de um poleiro de ouro || nas taxas rápidas, como vimos no experimento acústico preliminar.

Há também uma questão subsidiária a ser respondida, relacionada à primeira hipótese: como processos de enfraquecimento prosódicos relacionados à variação da TE podem ser acomodados no MDR?

4.2 Resultados

Como métodos objetivos de se detectar GA ainda não existem para dados articulatórios, análises acústicas foram realizadas para se detectar fronteiras de grupos acentuais, aos moldes dos experimentos acústicos descritos anteriormente. Tal detecção foi crucial, pois necessitávamos garantir que as proeminências frasais fossem colocadas nos locais esperados no desenho experimental e, a partir daí, efetuarmos análises cinemáticas.

4.2.1 Detecção de grupos acentuais

Seguindo os mesmos métodos dos experimentos acústicos apresentados anteriormente⁵, as seguintes proeminências frasais (em negrito) foram obtidas

⁵As mesmas sentenças também foram ouvidas, através de fones de ouvido, a fim de se corrigir quaisquer incongruências de atribuição de acentos frasais (como em uma posição

nas três TE (lenta (L), normal (N) e rápida (R)): Sentença 1

L: Ela **diz** | **mão** de **máfia** | no **carro** | **da** **moça** | do **papai**. |

N: Ela **diz** | **mão** de **máfia** | no **carro** | **da** **moça** | do **papai**. |

R: Ela **diz** | **mão** de **máfia** | no **carro** | **da** **moça** do **papai**. |

Sentença 2

L: Foi **bão** | **gostar** | **demais** | **de** **papai**, | **mas** de **máfia**. |

N: Foi **bão** | **gostar** | **demais** de **papai**, | **mas** de **máfia**. |

R: Foi **bão** **gostar** **demais** | **de** **papai**, | **mas** de **máfia**.|

Sentença 3

L: Mamãe | **não** **quer** | **mais** qu'eu **babe**, | **mas** qu'eu **pape**. |

N: Mamãe **não** **quer** **mais** | **qu**'eu **babe**, | **mas** qu'eu **pape**. |

R: Mamãe **não** **quer** **mais** qu'eu **babe**, | **mas** qu'eu **pape**. |

Sentença 4

L: Vou **lá** | **levar** o **pavê** | **pra** filha do **papai**. |

N: Vou **lá** | **levar** o **pavê** | **pra** filha do **papai**.|

R: Vou **lá** | **levar** o **pavê** | **pra** filha do **papai**. |

átona) e de se observar se as proeminências frasais correspondiam a fronteiras frasais percebidas como tais.

As análises acústicas também confirmaram que o falante produziu três TE distintas, confirmadas por uma Anova one-way com a duração acústica do VV como variável dependente e taxa como fator. Os resultados estatísticos mostraram uma diferença consistente entre as TE para todas as sentenças (sentença 1: $F(2,12) = 41,847$, $p < 10^{-4}$; sentença 2: $F(2,25) = 93,860$, $p < 10^{-4}$; sentença 3: $F(2,18) = 194,07$, $p < 10^{-4}$; sentença 4: $F(2,22) = 21,883$, $p < 0,00002$). Contudo, um teste *post-hoc Scheffé* não mostrou diferenças entre as taxas lenta e normal.

Outro resultado acústico relacionado à duração do VV é a diminuição do desvio-padrão com o aumento da TE. Usando uma Anova one-way com o desvio-padrão do VV como variável dependente e TE como fator, os seguintes resultados foram obtidos: sentença 1, $F(2,12) = 4,9163$, $p < 0,03$; sentença 2, $F(2,25) = 28,274$, $p < 10^{-4}$; sentença 3, $F(2,18) = 33,631$, $p < 10^{-4}$; sentença 4, $F(2,18) = 156,60$, $p < 10^{-4}$. Um teste *post-hoc Scheffé* subsequente mostrou uma tripla diferenciação para as sentenças 2 e 4 e uma dupla diferenciação para as sentenças 1 ((lenta=normal) $\neq$ rápida) e 3 (lenta $\neq$ (normal = rápida)). Esse resultado corrobora os achados de nossos experimentos acústicos prévios.

Como uma análise final, para se avaliar a consistência da diferenciação inter-taxas e menores desvios-padrão com o aumento da TE, todas sentenças foram agrupadas e foi realizada uma Anova one-way com duração do VV

e desvio-padrão em função da TE. Os resultados mostram que, para esse conjunto de dados, as três TE são altamente distintas ($F(2,86) = 53,305$, $p < 10^{-4}$, com diferenças estatísticas entre as taxas, confirmadas pelo teste *post-hoc Scheffé*) e o desvio-padrão diminui com o aumento da TE ($F(2,286) = 13,399$, $p < 0,00002$, com a taxa lenta sendo diferente das taxas normal e rápidas, confirmadas pelo teste *post-hoc Scheffé*).

4.2.2 Comparação entre a duração do VV articulatório e acústico

As análises acústicas (*Anova one-way*) não mostraram diferenças estatísticas entre a duração articulatória (relacionada à acústica) e a acústica dos VV para todas as sentenças (sentença 1: $F(1,206) = 1,4195$, $p < 0,24$; sentença 2: $F(1,184) < 0,034$, $p < 0,86$; sentença 3: $F(1,214) < 0,2$; $p = 0,9$; sentença 4: $F(1,228) < 0,042$, $p = 0,84$). Um exemplo disso é dado na figura 4.12.

Da mesma forma, o desvio-padrão da duração do VV articulatório seguiu o mesmo padrão do desvio-padrão da duração do VV acústico, i.e., uma diminuição do desvio-padrão com o aumento da TE. Tal padrão, inclusive, com apenas uma exceção, seguiu a mesma diferenciação nas taxas do desvio-padrão acústico, de acordo com a tabela 4.3. Além disso, agrupando todas as sentenças, achamos o mesmo padrão do desvio-padrão acústico ($F(2,59)$

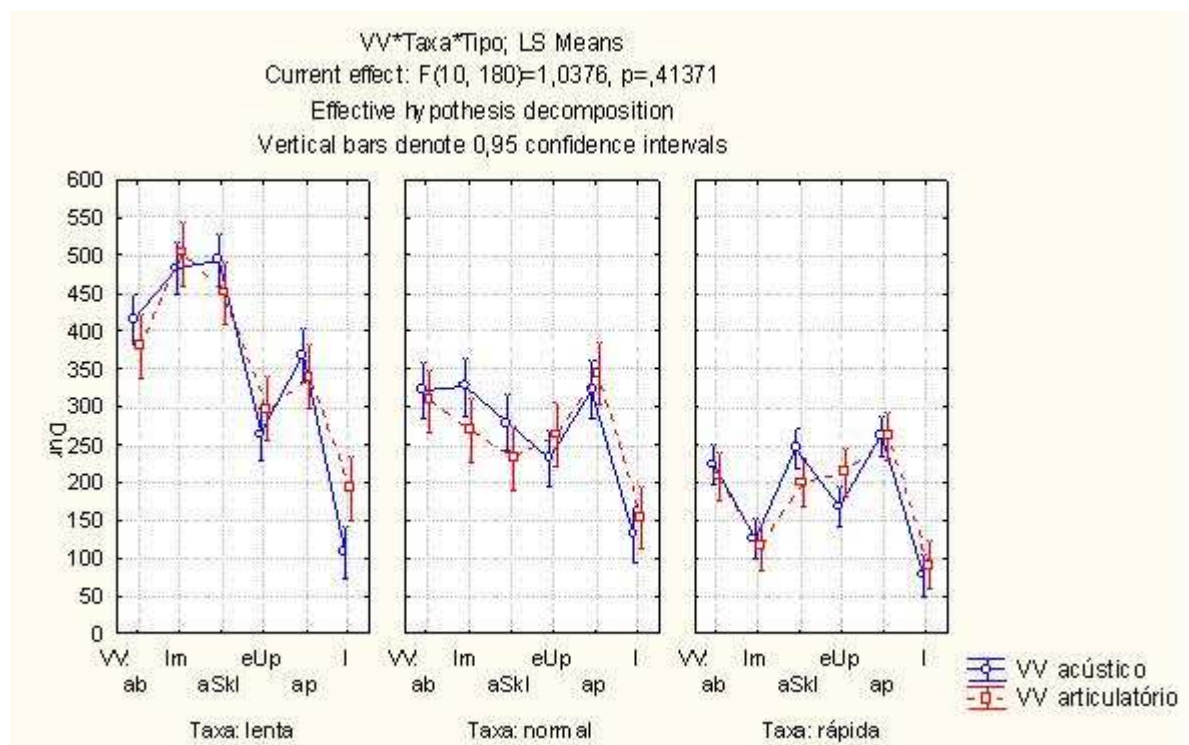


Figura 4.12: Anova Fatorial com duração em função da TE, unidade VV e tipo de VV para a sentença: “Mamãe não quer mais qu’eu **babe, mas qu’eu pape**”. As sílabas em negrito representam os VVs utilizados na análise.

= 9,4133, $p < 0,0003$ com lenta $\neq$ (normal = rápida), através de um teste *post-hoc Scheffé*). Por todas as similaridades entre a duração dos VVs articulatórios e acústicos, será considerado nesta tese que os GA encontrados na duração do VV acústico são refletidos nos VVs articulatórios (relacionados aos acústicos).

Tabela 4.3: Diminuição do desvio-padrão da duração do VV articulatório (relacionado ao acústico) para todas as sentenças do experimento complementar.

Sentença	Anova	$p <$	Scheffé
1	F (2,9) = 18,680	0.00064	(L = N) $\neq$ R
2	F (2,9) = 28,276	0.00014	L $\neq$ N $\neq$ R
3	F (2,12) = 10,481	0.0024	L $\neq$ (N = R)
4	F (2,20) = 14,394	0.00014	(L = N) $\neq$ R

Um outro resultado interessante é que até mesmo a duração do VV articulatório tradicional (medida, na literatura fonética, entre as velocidades zero ((d) na fig. 4.11), doravante AT) comparada à duração do VV acústico mostrou apenas uma diferença significativa (sentence 4, F (1,274) = 8,3741, $p < 0,0042$)⁶. Todas as outras sentenças foram estatisticamente iguais. Além disso, a duração AT⁷ em função da TE indicou uma diminuição da duração entre as taxas (sentença 1, F (2,90) = 10,769, $p < 0,00007$, sentença 2, F

⁶Esta é a única diferença entre a duração do VV articulatório e acústico, que é explicada pela diferença entre o VV articulatório final [a] comparado ao acústico [aj]. Não foi possível distinguir claramente [aj] no sinal articulatório, então escolhemos apenas [a].

⁷De agora em diante, toda referência à duração articulatória tradicional será indicada como AT. Do contrário, estaremos nos referindo à duração medida entre picos e vales de velocidade ((c) na fig. 4.11).

(2,57) = 5,2092, $p < 0,0084$, sentença 3, $F(2,72) = 48,665$, $p < 10^{-5}$, sentença 4, $F(2,135) = 32,219$, $p < 10^{-5}$). É importante salientar que nossa atenção foi focada em alguns momentos na duração articulatória entre os máximos (em módulo) de velocidade, porque ela corresponde mais diretamente à duração acústica.

4.2.3 Efeitos gerais da taxa na duração e cinemática

A presente seção analisa os efeitos do aumento da TE na duração articulatória e cinemática⁸. O objetivo desta seção é achar padrões articulatórios que correspondam a reestruturações do ritmo da fala. Estes achados irão também complementar as análises acústicas descritas nesta tese. Uma Anova de dois fatores foi usada em todas as análises estatísticas, com todos efeitos com $p < 0,05$ relatados.

Nossos resultados sugerem que o aumento da TE tende a aumentar a característica de cabeça à direita do PB. Para a sentença 1, com duração em função dos VVs articulatórios (relacionados ao acústico) [afj] e [aj], a maior proeminência frasal ocorreu em [afj] para as taxas lenta ($F(1,4) = 15,858$, $p < 0,017$) e normal ($F(1,10) = 72,681$, $p < 0,0002$), e em [aj] para a taxa rápida

⁸Cinemática é o estudo dos movimentos de objetos. A base da cinemática é a escolha de coordenadas que descrevem a posição e/ou orientação dos objetos. As derivadas destas coordenadas correspondem a velocidades e acelerações (vide seção 4.1: “Cálculo aplicado no Mavis: derivadas”).

(n.s.) Para a sentença 2, com duração em função dos VVs articulatórios [ap] e [afj], a maior proeminência frasal ocorreu em [ap] para as taxas lenta ($F(1,8) = 85,689$, $p < 0,00003$) e normal ($F(1,6) = 22,611$, $p < 0,0032$), e em [afj] para a rápida (n.s.). Para a sentença 3, com duração em função dos VVs articulatórios [ab] e [ap], nenhuma diferença estatística foi achada para a taxa lenta, e para as taxas normal ($F(1,6) = 6,0367$, $p < 0,0494$) e rápida ($F(1,12) = 15,042$, $p < 0,003$) a maior proeminência frasal ocorreu em [ap]. Para a sentença 4, com duração em função dos VVs articulatórios [epr] e [aj], não houve diferenças estatísticas entre as taxas.

Uma comparação das vogais das sentenças (cf. tabela 4.4) em função da taxa indica uma tendência geral para menores deslocamentos (deslocamento vertical máximo e deslocamento da constrição) com o aumento da TE. Vogais altas e média-altas tendem a ser menos altas e vogais baixas tendem a ser menos baixas da taxa normal para a rápida, conforme mostram as análises estatísticas das tabelas 4.5 (deslocamento vertical máximo, vide fig. 4.11(a)) e 4.6 (deslocamento da constrição, vide fig. 4.11(b)). Todas as outras vogais ausentes nesses gráficos não alcançaram significância.

Tabela 4.4: Vogais (em negrito) utilizadas em uma *Anova One-way* com deslocamento vertical máximo/deslocamento da constrição em função da TE.

Sentence 1:	Ela diz mão de máfia no carro da moça do papai .
Sentence 2:	Foi bão gostar demais de papai , mas de máfia !?
Sentence 3:	Mamãe não quer mais qu' eu babe , mas qu' eu pape .
Sentence 4:	Vou lá levar o pavê pra filha do papai .

Tabela 4.5: Significância do deslocamento vertical em função da TE. Nesta tabela temos representada a vogal com sua palavra respectiva, Anova e significância para todas as sentenças do *corpus*. * significa marginalmente significativo.

Deslocamento vertical máximo			
vogal	palavra	Anova	p <
Sentença 1			
[a]	['kafju]	$F(2,9) = 6,2189$	0.021
[a]	[da]	$F(2,9) = 7,4352$	0.013
[aj]	[pa'paj]	$F(2,9) = 3,9786$	0.06*
Sentença 2			
[a]	[pa'paj]	$F(2,9) = 6,1360$	0.021
[a]	[mas]	$F(2,9) = 12,116$	0.003
[a]	['mafje]	$F(2,9) = 18,273$	0.0007
[e]	['mafje]	$F(2,9) = 9,1656$	0.007
Sentença 3			
[a]	['babi]	$F(2,12) = 4,1435$	0.043
[i]	['babi]	$F(2,12) = 4,0615$	0.045
[a]	[mas]	$F(2,12) = 22,454$	0.0001
[a]	['papi]	$F(2,12) = 4,1749$	0.043
[i]	['papi]	$F(2,12) = 5,6012$	0.02
Sentença 4			
[a]	[le'var]	$F(2,20) = 6,0484$	0.009

Tabela 4.6: Significância do deslocamento da constrição em função da TE. Nesta tabela temos representada a vogal com sua palavra respectiva, Anova e significância para todas as sentenças do *corpus*.

Deslocamento da constrição			
vogal	palavra	Anova	p <
Sentença 1			
[a]	['mafjɐ]	$F(2,9) = 7,4418$	0.0189
[a]	['kafɨ]	$F(2,9) = 12,495$	0.003
[a]	[da]	$F(2,9) = 11,437$	0.0034
[aj]	[pa'paj]	$F(2,9) = 11,062$	0.0038
Sentença 2			
[a]	[pa'paj]	$F(2,9) = 4,5202$	0.044
[a]	[mas]	$F(2,9) = 9,2330$	0.007
[a]	['mafjɐ]	$F(2,9) = 9,0092$	0.008
[ɐ]	['mafjɐ]	$F(2,9) = 8,2842$	0.0092
Sentença 3			
[a]	['babɪ]	$F(2,12) = 22,306$	0.0001
[ɪ]	['babɪ]	$F(2,12) = 4,0261$	0.046
[a]	[mas]	$F(2,12) = 33,106$	0.00002
[a]	['papɪ]	$F(2,12) = 17,600$	0.00028
[ɪ]	['papɪ]	$F(2,12) = 15,114$	0.0006
Sentença 4			
[a]	[pa've]	$F(2,20) = 5,3657$	0.014
[e]	[pa've]	$F(2,20) = 18,887$	0.00003
[a]	[pra]	$F(2,20) = 20,979$	0.00002
[ɐ]	['fɪʌɐ]	$F(2,20) = 7,0721$	0.005
[aj]	[pa'paj]	$F(2,20) = 9,6657$	0.0012

Uma comparação da velocidade de pico (consoantes) ou de vale (vogais) em função da TE para todos os gestos sugere uma diminuição da velocidade máxima da taxa lenta à rápida. Embora não tenha sido achado estatisticamente tal padrão para gestos de algumas sentenças (SENTENÇA 1: 2 diminuições em 15, SENTENÇA 2: 5 em 15 possíveis, SENTENÇA 4: 7 em 12), tal padrão foi achado para todos os gestos da sentença 3 ($F(2,191) = 4,7334$, $p < .0099$, usando as velocidades máxima em valor absoluto). Portanto, estudos futuros são necessários a fim de suportar a hipótese de diminuição da velocidade máxima com o aumento da TE.

Efeitos da TE em ambas duração da aceleração e AT revelam um padrão de decaimento da taxa lenta à rápida, conforme mostram as análises estatísticas apresentadas na tabela 4.7.

Tabela 4.7: Análises estatísticas da duração da aceleração e AT em função da TE para todas as sentenças do *corpus*.

Duração AT		
Sentença	Anova	p <
1	$F(2,183) = 13,080$	10^{-5}
2	$F(2,141) = 9,1624$	0,0002
3	$F(2, 177) = 52,992$	10^{-4}
4	$F(2,273) = 76,072$	10^{-4}
Duração da aceleração		
1	$F(2,183) = 12,549$	0,00002
2	$F(2,141) = 12,382$	0,00002
3	$F(2,176) = 34,077$	10^{-5}
4	$F(2,273) = 30,905$	10^{-5}

A análise do tempo-para-velocidade-de-pico proporcional em função da TE não indicou um padrão comum para todos os gestos. Alguns gestos tiveram um padrão descendente, outros ascendentes. Portanto, vogais acentuadas e átonas foram separadas e colocadas em grupos distintos. Este agrupamento ainda não indica efeito da TE no tempo-para-velocidade-de-pico proporcional, mas indica que as vogais acentuadas tiveram menores tempo-para-velocidade-de-pico proporcional do que as vogais átonas (SENTENÇA 1, $F(1,91) = 9,6706$, $p < 0,003$; SENTENÇA 2, $F(1,70) = 5,4489$, $p < 0,023$; SENTENÇA 3, n.s. ; SENTENÇA 4, $F(1,136) = 10,051$, $p < 0,002$).

Resumindo, os dados indicam um aumento da rigidez⁹ da taxa lenta à rápida, e, conseqüentemente, um encolhimento dos gestos e menores movimentos espaciais. Por essas razões, reestruturações rítmicas são mais prováveis de ocorrer em taxas rápidas, pois os últimos VV de GA em fronteiras fracas podem ser modificados duracionalmente devido ao aumento da rigidez da mola. Convém lembrar que durações longas são necessárias para se delimitar fronteiras frasais em PB.

⁹A duração da aceleração (fig. 4.11(f)) representa o parâmetro rigidez no MDT (Saltzman & Munhall, 1989).

4.2.4 Fontes cinemáticas de reestruturações rítmicas

Com base nos resultados articulatórios apresentados acima, propomos as seguintes fontes cinemáticas de reestruturações rítmicas com o aumento da TE (cf. figura 4.13): a) diminuição da duração da aceleração, i.e., aumento da rigidez; b) diminuição do deslocamento vertical máximo; c) diminuição do deslocamento do gesto de constrição; d) diminuição (em módulo) das velocidade de pico e vale; e) diminuição da duração articulatória; e f) tempo-para-velocidade-de-pico proporcional permanece constante. Outra fonte cinemática de reestruturação rítmica (não relacionada à mudança da TE) é a diminuição de tempo-para-velocidade-de-pico proporcional de vogais tônicas comparadas com as átonas.

De acordo com essas fontes, espera-se uma diminuição geral de parâmetros articulatórios da taxa lenta à rápida. A figura 4.13a representa um encurtamento articulatório causado por uma mudança na rigidez (cf. Saltzman & Munhal, 1989), que é indicado por uma duração da aceleração mais curta. A figura 4.13b exibe deslocamento vertical máximos menores causados por uma posição menos aberta da mandíbula (vogais baixas) ou menos fechada (vogais altas ou meio-altas), que podem ser entendidas como um *undershoot* do gesto. Relacionada com a figura 4.13b está a figura 4.13c que representa uma menor diferença de deslocamento-y nas posições inicial e final do gesto.

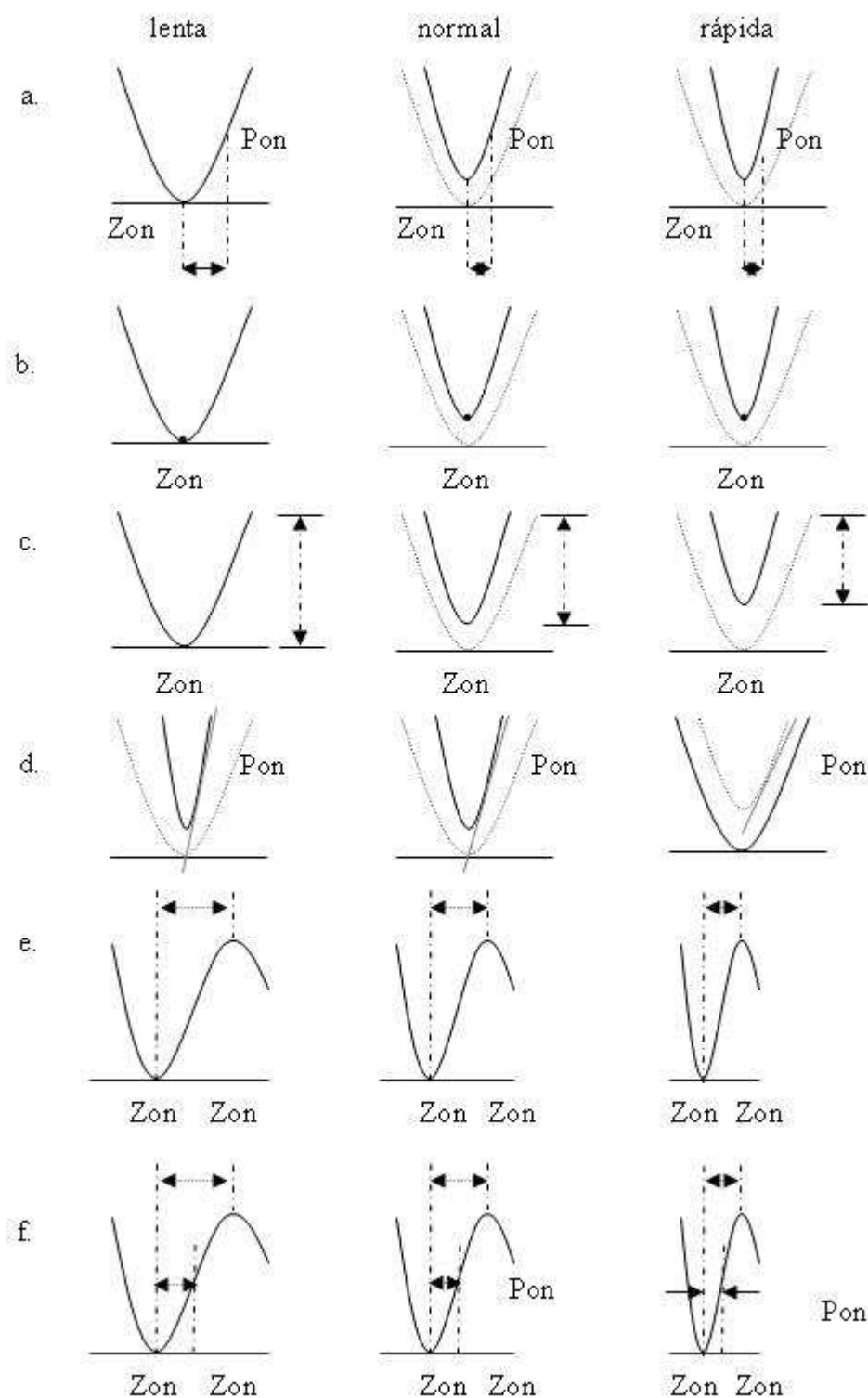


Figura 4.13: Fontes cinemáticas de reestruturações rítmicas com o aumento da TE: a. diminuição da duração da aceleração, i.e., aumento da rigidez; b. diminuição do deslocamento vertical máximo; c. diminuição do deslocamento do gesto de constrição; d. diminuição (em módulo) da velocidade de pico; e. diminuição da duração articulatória (tanto medida entre velocidades de pico quanto entre zeros); f. tempo-para-velocidade-de-pico proporcional permanece constante.

Se se espera menores deslocamento vertical máximos, também se espera menores deslocamentos da constrição nas taxas rápidas. A figura 4.13d mostra uma diminuição (em módulo) da velocidade de pico e/ou vale da taxa lenta à rápida. Esta hipótese pode ser explicada pelo fato de que, para se alcançar uma grande velocidade de pico e depois uma velocidade zero na mudança do movimento (a mandíbula sobe e desce), é necessária uma maior distância, que pode ser encontrada em taxas lentas, não rápidas. A figura 4.13e é meramente uma consequência das fontes cinemáticas anteriores, i.e., a diminuição da duração articulatória (tanto AT quanto medida entre velocidades de picos e vales). Por fim, a figura 4.13f mostra que o tempo-para-velocidade-de-pico proporcional permanece constante com o aumento da TE.

4.3 Reorganizações rítmico-articulatórias

Até o presente momento apresentamos mudanças articulatórias em geral relacionadas ao aumento da TE (cf. fig. 4.13, seção anterior), porém não fizemos a relação destas mudanças com as reestruturações rítmicas encontradas. Sendo assim, pretendemos nesta seção investigar se há algum fator articulatório capaz de explicar as reorganizações rítmicas dos GA com a aceleração da movimentação mandibular. Para tanto, utilizaremos os seis seguintes fatores: 1) duração da aceleração; 2) deslocamento vertical máximo; 3) des-

locamento da constrição; 4) velocidade (pico ou vale); 5) duração gestual; 6) tempo-para-velocidade-de-pico proporcional, analisados um a um para cada sentença.

Para a sentença 1, observamos, conforme figura 4.14, uma reestruturação rítmica do enunciado na taxa mais rápida. Nas taxas lenta e normal, o traçado da duração das unidades VV é praticamente o mesmo. No entanto, na taxa rápida, de acordo com a detecção dos grupos acentuais (seção 4.2.1), não há mais um pico no VV articulatório [os]¹⁰, indicado pela elipse da figura 4.14.

Concentrando-nos nos gestos componentes do VV [os] da sentença 1, analisamos se as hipóteses articulatórias (vide figura. 4.13, seção anterior) para as variáveis articulatórias descritas acima poderiam explicar tal reestruturação rítmica. Uma *Anova One-way* com variáveis dependentes articulatórias em função da TE foi efetuada e os resultados se encontram na tabela 4.8¹¹.

A tabela 4.8 nos revela que a hipótese de redução articulatória com o aumento da TE nas variáveis dependentes não foi achada para o gesto [o]. Para este gesto, a única hipótese confirmada foi a do valor constante do tempo para-velocidade-de-pico proporcional. Por outro lado, para o gesto

¹⁰O pico em [os] foi constatado na análise da evolução do z-score suavizado dos VVs acústicos.

¹¹Vide figs. 4.18 (seção 4.3.1), 4.22 (seção 4.3.2), 4.26 (seção 4.3.3), 4.30 (seção 4.3.4), 4.34 (seção 4.3.5) e 4.38 (seção 4.3.6), para uma visualização dos valores médios das variáveis desta tabela.

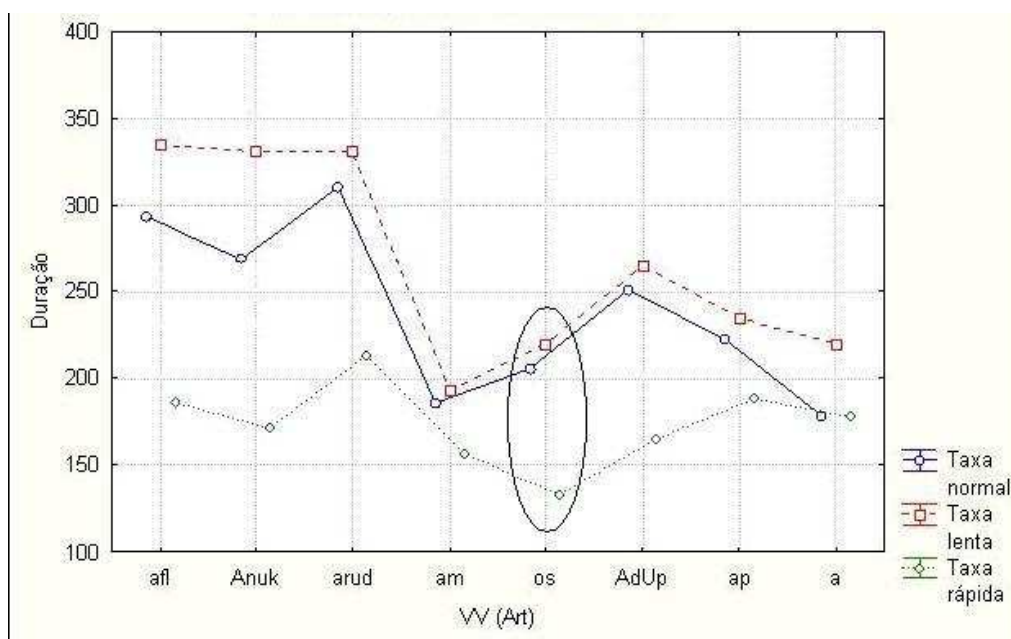


Figura 4.14: Reestruturação rítmica na sentença “Ela diz mão de máfia no carro da moça do papai”, em que a proeminência acentual no VV [os] nas taxas lenta e normal é apagada na taxa rápida (ressaltada pela elipse na figura).

Tabela 4.8: *Anova One-way* para as variáveis dependentes articulatórias relacionadas à perda da proeminência acentual em [os] na sentença “Ela diz mão de máfia no carro da moça do papai”.

Sentença 1				
	[o]		[s]	
Variáveis dependentes	Anova	p <	Anova	p <
1 - Duração da aceleração	n.s.	n.s.	$F(2,9) = 18,030$	0,0007
2 - Deslocamento vertical máximo	n.s.	n.s.	n.s.	n.s.
3 - Desl. da constrição	n.s.	n.s.	$F(2,9) = 13,543$	0,002
4 - Velocidade	n.s.	n.s.	$F(2,9) = 6,604$	0,02
5 - Duração articulatória	n.s.	n.s.	$F(2,9) = 22,282$	0,0004
6 - Tempo PVP proporcional	n.s.	n.s.	n.s.	n.s.

consonantal, [s], as hipóteses previstas foram confirmadas, com exceção da diminuição do deslocamento vertical máximo. Esses resultados sugerem que, nas reestruturações rítmicas da fala, reorganizações articulatórias relacionadas ao aumento da TE seguem as hipóteses descritas anteriormente (figura 4.13, seção anterior), contudo podem não se manifestar em todos os constituintes do VV em questão.

Para a sentença 2, observamos na figura 4.15, conforme análise da seção 4.2.1, uma reestruturação rítmica do enunciado na mudança da taxa lenta para a normal na unidade articulatória [ajzɔɣɪ], embora não manifestada na taxa rápida. Nesse caso, o traçado da duração do VV articulatório (medida entre picos e vales de velocidade) não nos ajuda na visualização da questão, pois, os dados articulatórios não foram submetidos ao *z-score* nem à

normalização, como nas análises acústicas anteriores.

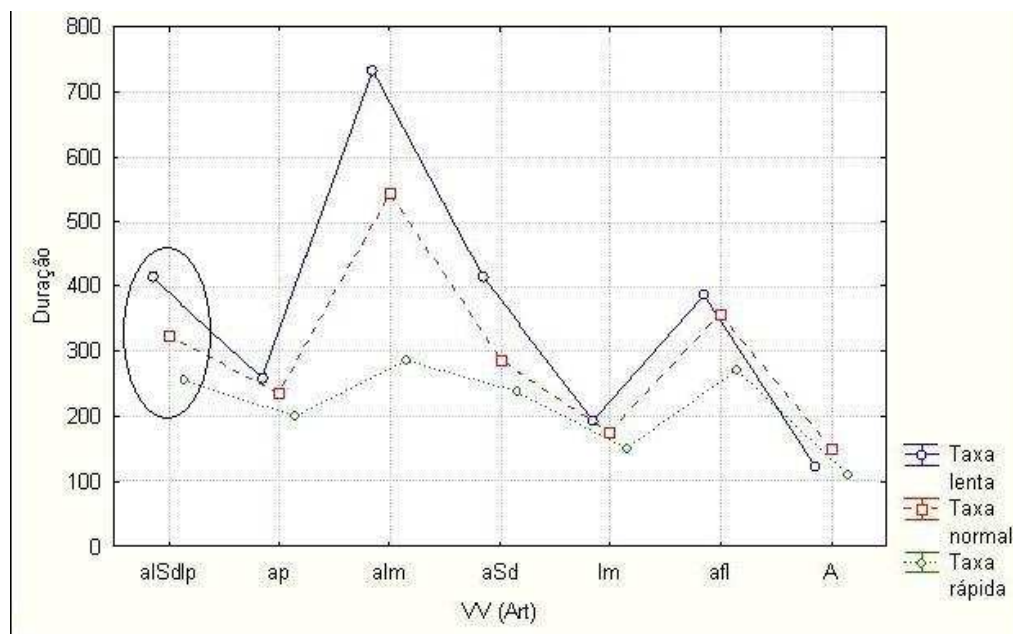


Figura 4.15: Reestruturação rítmica na sentença “Ela diz mão de máfia no carro da moça do papai”, em que a proeminência acentual na unidade articulatória [ajzɔ̃ɪ] nas taxas lenta e rápida é apagada na taxa normal (ressaltada pela elipse na figura).

Com base na unidade articulatória [ajzɔ̃ɪ] da sentença 2, analisamos, como na sentença 1, se as hipóteses articulatórias para as variáveis articulatórias descritas acima poderiam explicar tal reestruturação rítmica. Uma *Anova One-way* de variáveis dependentes articulatórias em função da TE foi efetuada e os resultados se encontram na tabela 4.9.

A análise da tabela 4.9 nos mostra que, diferentemente da sentença 1, houve confirmação da hipótese de redução articulatória para as variáveis articulatórias 3 e 4 e da hipótese de tempo-para-velocidade-de-pico proporcional

Tabela 4.9: *Anova One-way* para as variáveis articulatórias relacionadas à perda da proeminência acentual em [ajzɖɪ] na sentença “Foi bão gostar demais de papai, mas de máfia!?”.

Sentença 2				
	[ajzɖɪ]		[p]	
Variáveis dependentes	Anova	p <	Anova	p <
1 - Duração da aceleração	n.s.	n.s.	F (2,9) = 8,1382	0.0096
2 - Deslocamento vertical máximo	n.s.	n.s.	n.s.	n.s.
3 - Desl. da constrição	F (2,9) = 5,6169	0.027	F (2,9) = 4.2517	0.05
4 - Velocidade	F (2,9) = 5,5608	0.27	n.s.	n.s.
5 - Duração articulatória	n.s.	n.s.	F (2,9) = 40.683	0.00004
6 - Tempo PVP proporcional	n.s.	n.s.	n.s.	n.s.

constante (variável 6) para o gesto mandibular de [ajzɖɪ]. Para o gesto [p], por sua vez, as hipóteses foram confirmadas para todos os fatores, exceto os fatores 2 e 4. Conforme podemos notar, como na sentença 1, o fator deslocamento vertical máximo não demonstrou influência da TE no VV articulatório reestruturado. Esses resultados confirmam o fato de serem necessárias mudanças articulatórias, conforme hipóteses da figura 4.13 (seção anterior), em pelo menos alguns parâmetros articulatórios do VV a ser reestruturado.

Para as sentenças 3 (fig. 4.16) e 4 (fig. 4.17), por não haver reestruturações rítmicas no trecho estudado, não analisaremos os VVs. Ressaltamos que as figuras 4.16 e 4.17 representam a evolução dos VVs articulatórios sem *z-score* e normalização, por isso os picos duracionais não correspondem aos GA mencionados na seção 4.2.1 (marcados pela elipse nas figuras abaixo).

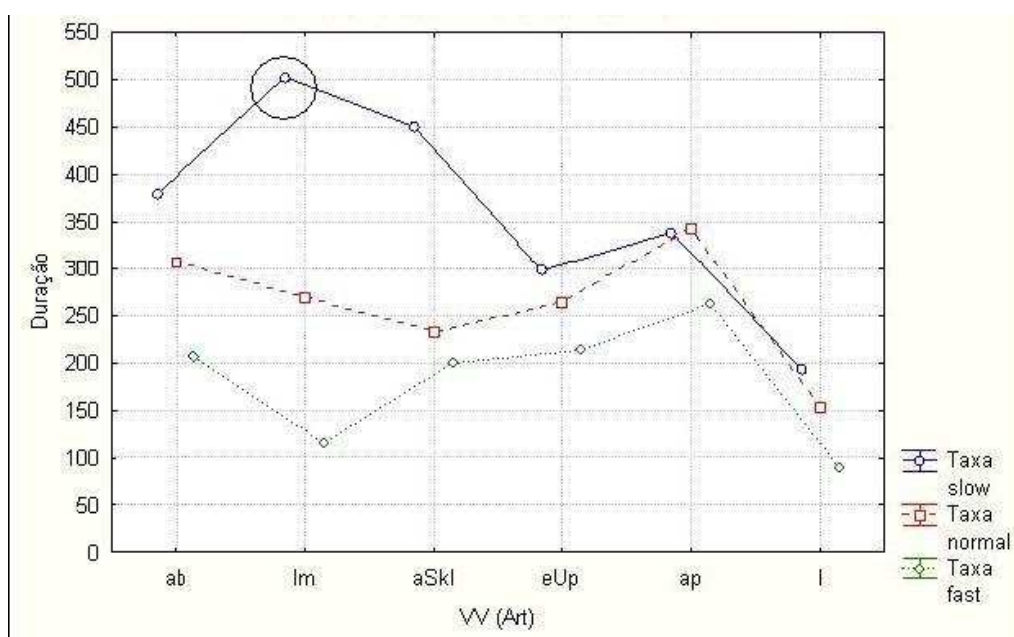


Figura 4.16: Evolução da duração dos VVs articulatórios na sentença 3.

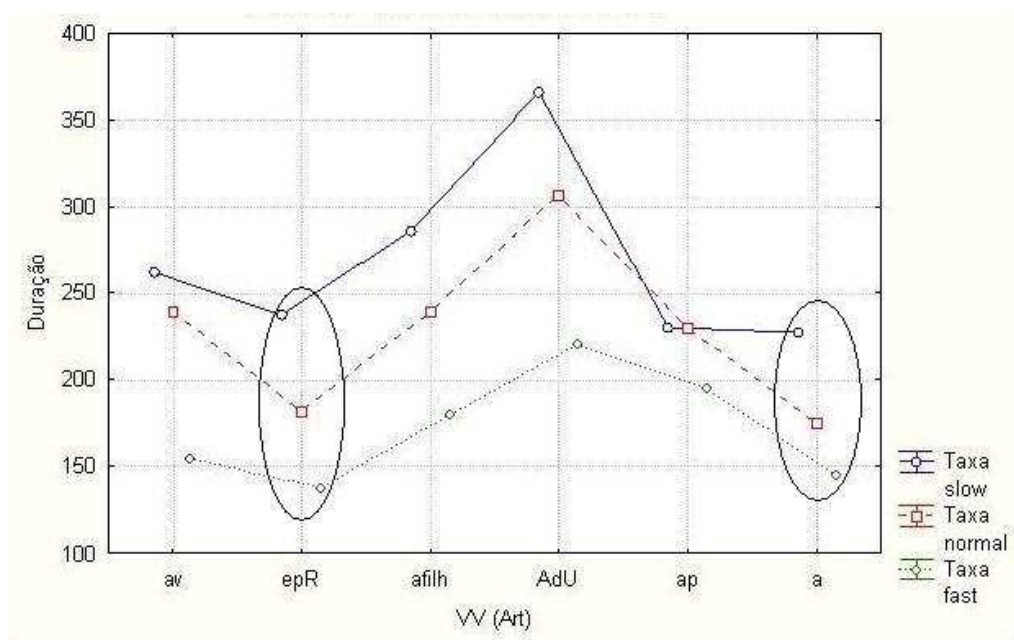


Figura 4.17: Evolução da duração dos VVs articulatórios na sentença 4.

Após verificarmos terem as hipóteses da figura 4.13 (seção anterior) se manifestado nos VVs reestruturados pelo aumento da TE, questionamo-nos se poderíamos atribuir diminuições significativas nos fatores articulatórios estudados apenas para as situações em que os VVs fossem reestruturados. Para verificar esta questão fizemos *Anovas One-way*, para cada variável dependente articulatória e sentença, com o gesto articulatório em função da TE, descritas a seguir.

4.3.1 Duração da aceleração

A tabela 4.10 nos mostra que todas as sentenças apresentaram uma diminuição da duração da aceleração, ao se fazer uma *Anova One-way* com duração da aceleração em função da TE. Além disso, tal padrão ocorreu para todos os gestos, independente de os mesmos pertencerem a VVs em posição de acento frasal.

Tabela 4.10: Estatística para a duração da aceleração para todas as sentenças do experimento articulatório prosódico, todos os gestos (consonantais e vocálicos) considerados.

Duração da aceleração x TE		
	Anova	p <
Sentença 1	F (2,183) = 12,549	0.00002
Sentença 2	F (2,141) = 12,382	0.00002
Sentença 3	F (2,176) = 34,077	10^{-5}
Sentença 4	F (2,273) = 30,905	10^{-5}

Para a sentença 1, dos 16 gestos mandibulares possíveis de diminuição da duração da aceleração com o aumento da TE, apenas 4 ([ɐ] em “máfia”, [m] e [o] em “moça”, e [a] em “papai”) não apresentaram tal padrão. Para a sentença 2, dos 12 gestos, 6 não apresentaram tal padrão ([a] e [aj] em “papai”, [m], [a] e [z] em “mas” e [m] em “máfia”). Para a sentença 3, dos 12 gestos, apenas 2 ([a] e o segundo [p] em “papai”) não apresentaram diminuição da duração da aceleração. Por fim, para a sentença 4, todos os 12 gestos apresentaram o efeito articulatorio. Tal fenômeno é claramente observável nos gráficos da duração da aceleração, para cada uma das sentenças, apresentados abaixo.

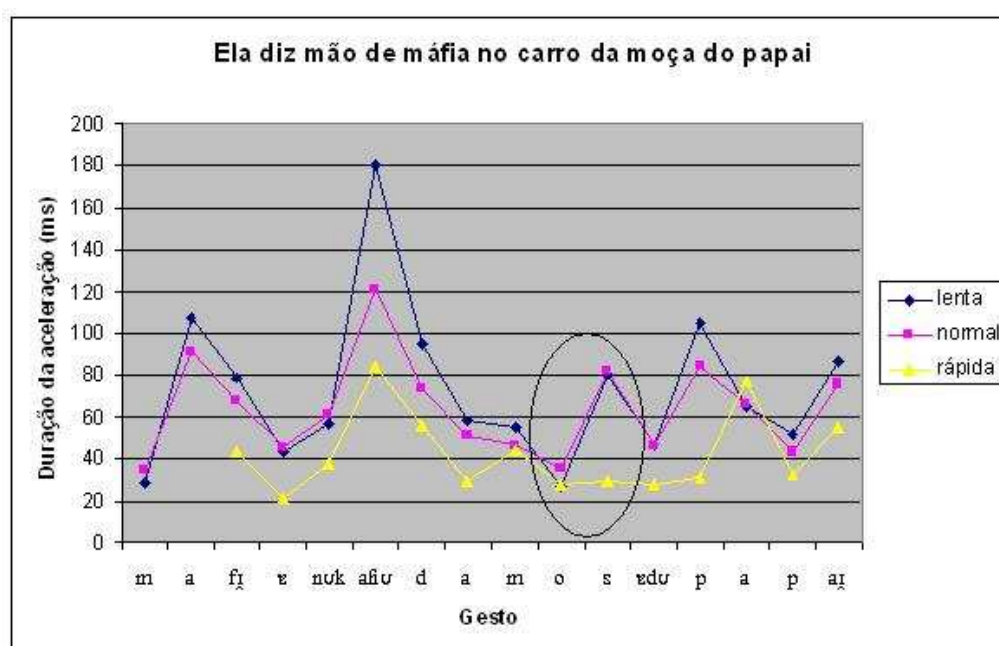


Figura 4.18: Duração da aceleração para a sentença 1 em três TE.

A análise das figuras 4.18–4.21 nos mostra que, na taxa rápida, a dura-

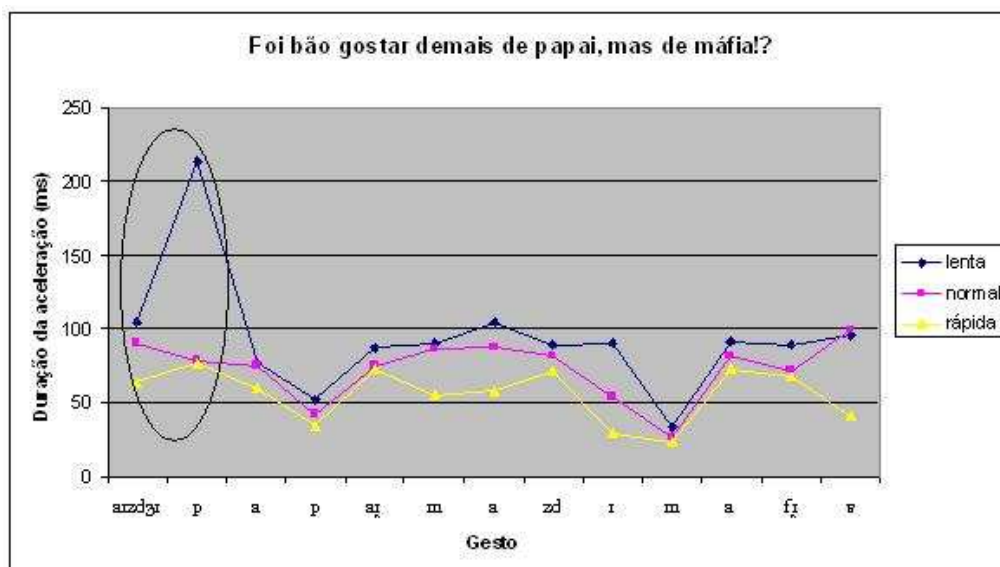


Figura 4.19: Duração da aceleração para a sentença 2 em três TE.

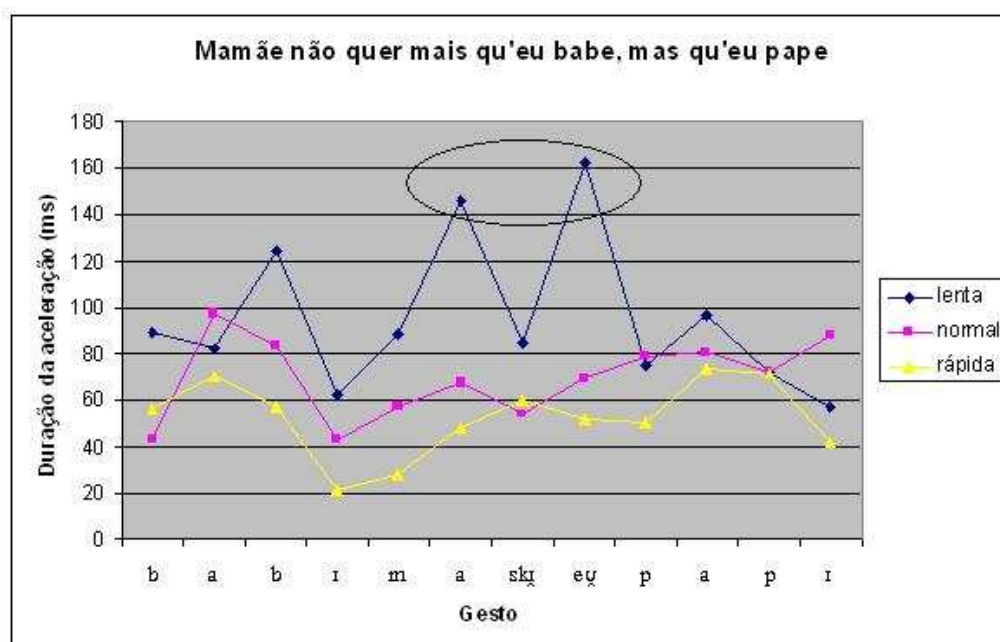


Figura 4.20: Duração da aceleração para a sentença 3 em três TE.

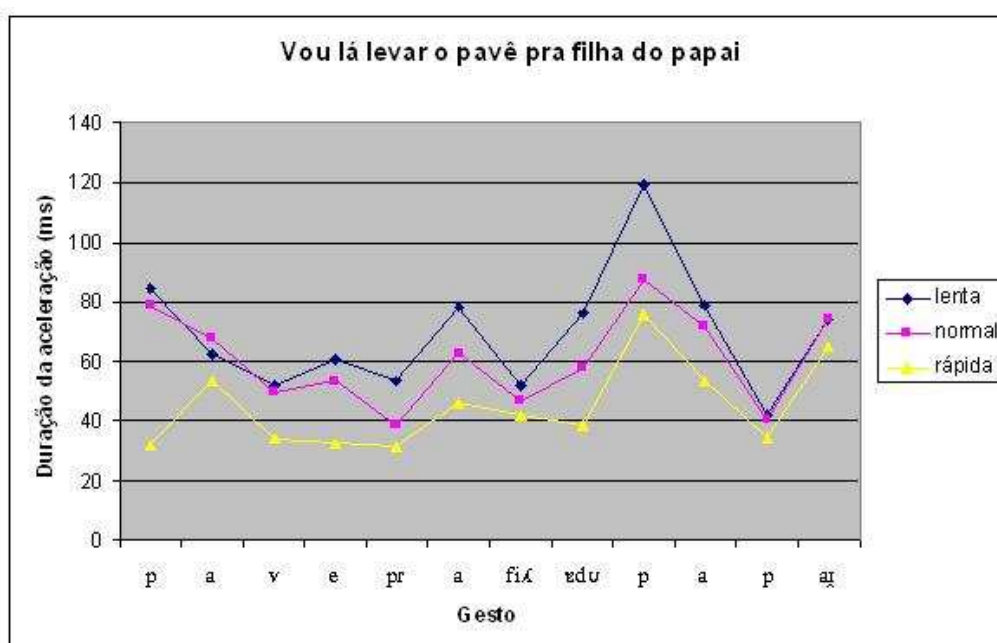


Figura 4.21: Duração da aceleração para a sentença 4 em três TE.

ção foi, em média, menor do que nas outras taxas. Além disso, tal padrão apresenta-se claramente o mesmo para todos os gestos dos enunciados (os grandes picos na figura 4.20 foram causados por grandes pausas silenciosas introduzidas pela locutora na taxa lenta). No entanto, um fato nos chamou a atenção: o contorno duracional da aceleração para os gestos [o] e [s] da sentença 1 (marcados pela elipse na figura 4.12) e [ajzdʒɪ] e [p] da sentença 2 (marcados pela elipse na figura 4.13) foi diferente dos demais gestos. Com exceção de [ajzdʒɪ] todos esses gestos apresentaram significância estatística. Além disso, justamente nos VVs correspondentes a esses gestos, ocorrem reestruturações rítmicas. Na sentença 1, a reestruturação rítmica ocorre na

mudança da taxa normal à rápida e na sentença 2, a mudança ocorre da taxa lenta à normal e desta à rápida (observem o diferente padrão de duração da aceleração nas elipses das figuras 4.18 e 4.19). Isso nos leva a crer que, apesar de não encontrarmos evidência com nossas medidas, ao se fazer certas derivações nos contornos articulatórios poderemos vir a observar alguma influência acentuada do fator duração da aceleração nas posições de acento frasal.

4.3.2 Deslocamento vertical máximo

As figuras 4.22–4.25 apresentadas abaixo mostram nitidamente não haver diferenciação dos deslocamento vertical máximos em função do gesto pertencer a um VV em posição de acento frasal. Além disso, para as sentenças 1 (fig. 4.22) e 2 (fig. 4.23), há uma diminuição (em módulo) do deslocamento vertical máximo e, para as sentenças 3 (fig. 4.24) e 4 (fig. 4.25), um aumento (em módulo) do mesmo. No entanto, uma Anova *One-way* com deslocamento vertical máximo em função da TE mostrou apenas um resultado significativo: sentença 2 ($F(2,165) = 3,5520$, $p < 0,031$). No entanto, acreditamos que a não observação de menores deslocamento vertical máximos em função da TE (semelhante a estudos como os de Erickson citados anteriormente) pode-se dever ao fato de termos poucos dados articulatórios para esta variável. Lem-

bremos que a variação do deslocamento mandibular é muito pequena com o aumento da TE.

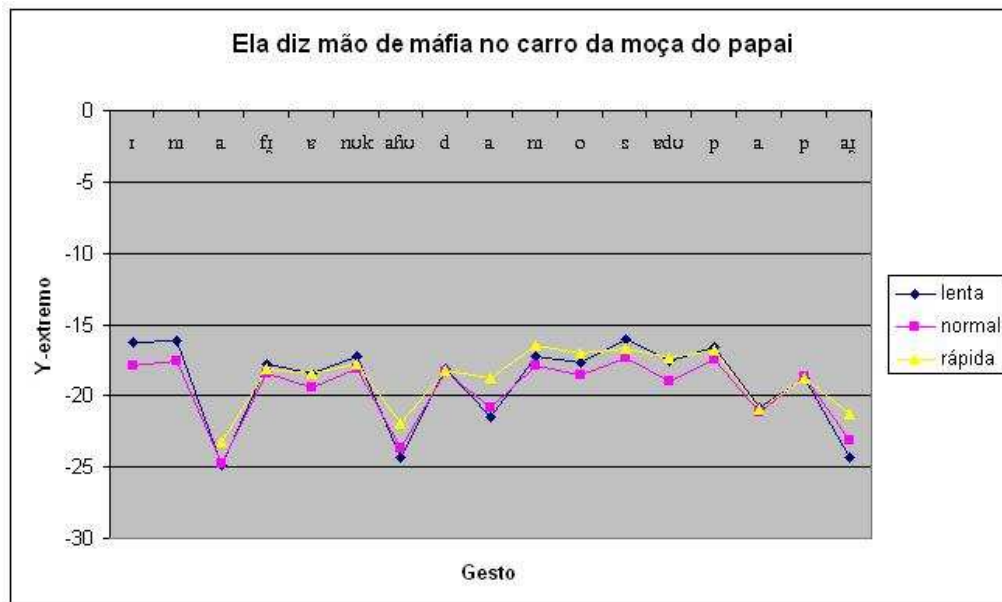


Figura 4.22: Deslocamento vertical máximo para a sentença 1 em três TE.

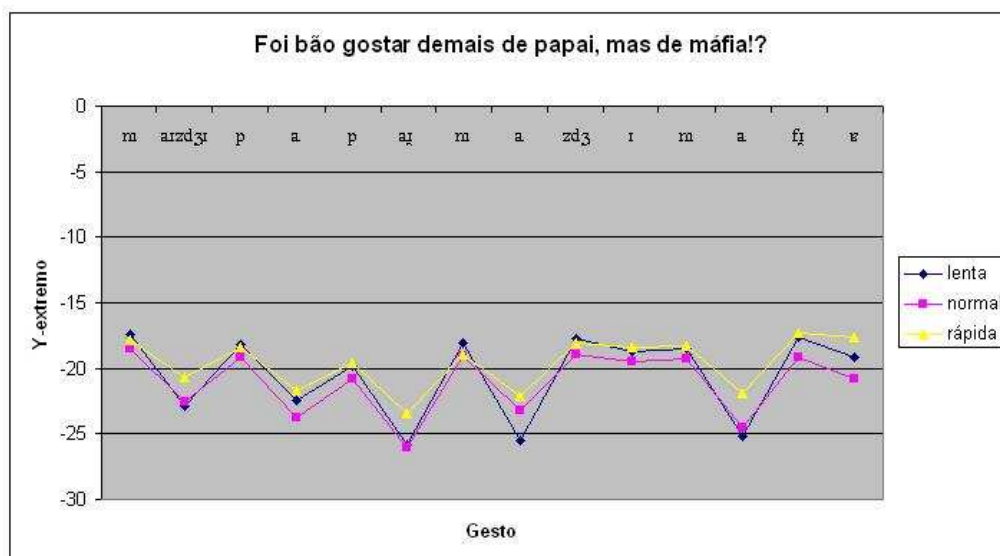


Figura 4.23: Deslocamento vertical máximo para a sentença 2 em três TE.

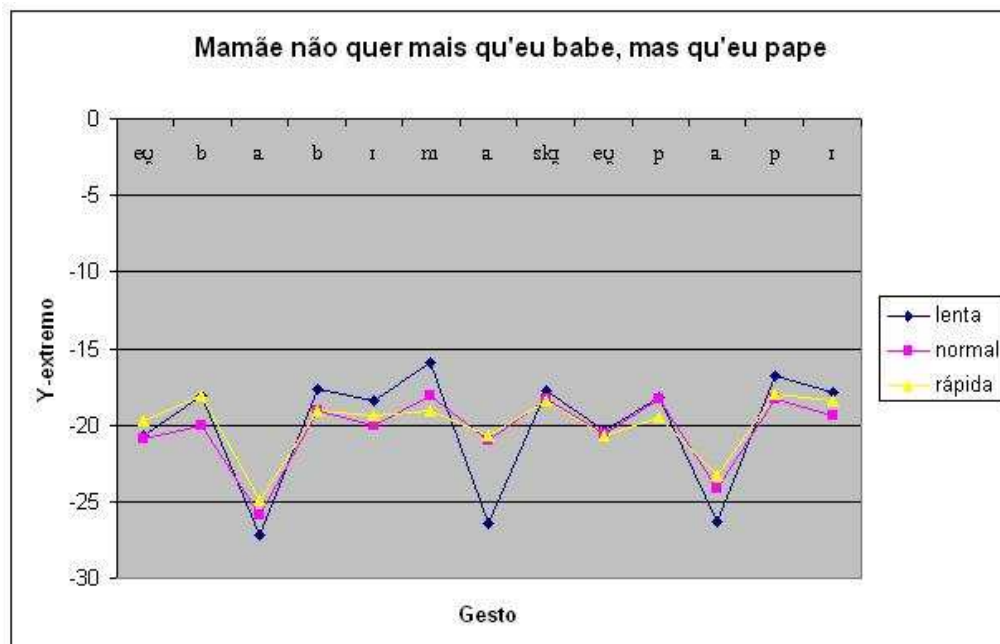


Figura 4.24: Deslocamento vertical máximo para a sentença 3 em três TE.

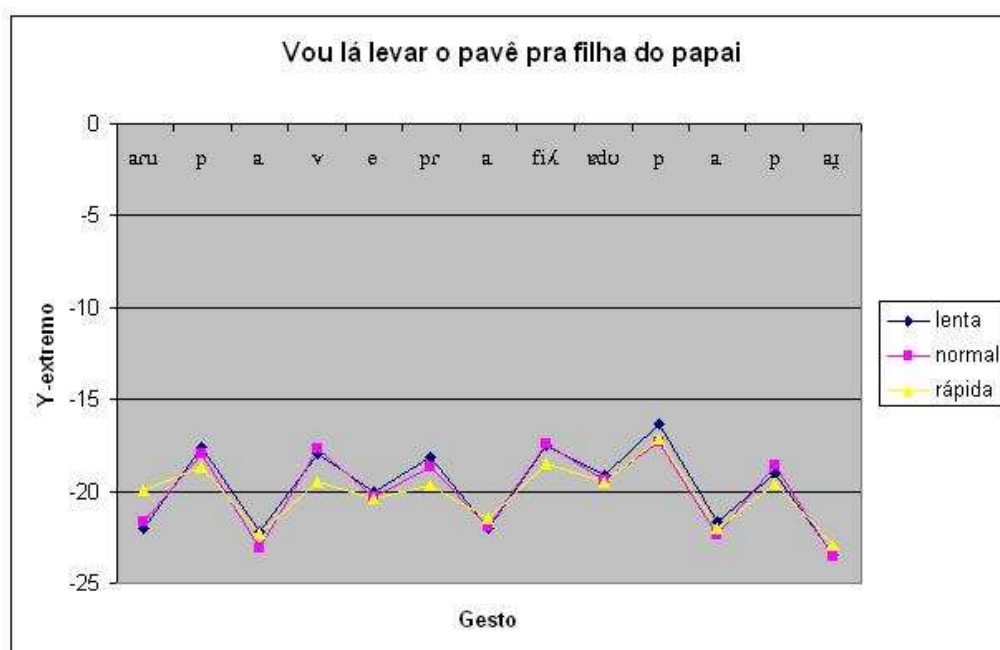


Figura 4.25: Deslocamento vertical máximo para a sentença 4 em três TE.

4.3.3 Deslocamento da constrição

As figuras 26–29 apresentam claramente que o deslocamento da constrição possui o mesmo padrão de variação independente da TE e do gesto envolvido e que as taxas rápidas tendem a apresentar um menor deslocamento da constrição nas taxas rápidas. Uma Anova *One-way* efetuada com o deslocamento da constrição em função da TE, para todas as sentenças, com todos os gestos considerados, confirma esta tendência, como mostra a tabela 4.11. Além disso, como para a duração da aceleração, tal padrão ocorreu para todos os gestos, seja em posição tônica ou átona.

Para a frase 1, dos 16 gestos mandibulares possíveis de diminuição do

Tabela 4.11: Estatística para o deslocamento da constrição para todas as sentenças do experimento articulatório prosódico, todos os gestos (consonantais e vocálicos) considerados.

Deslocamento da constrição x TE		
Sentença	Anova	p <
1	F (2,273) = 24,377	10^{-5}
2	F (2,177) = 17,340	10^{-5}
3	F (2,184) = 3,9812	0,021
4	F (2,153) = 10,418	0,00006

deslocamento da constrição com o aumento da TE, 8 não apresentaram tal padrão ([m] e [v] em “máfia”; [m], [o] e [vɔ] no trecho “moça do”; [a] e os dois [p] de “papai”). Para a frase 2, dos 12 gestos, 5 não apresentaram esse padrão (o segundo [p] e [aj] em “papai”; [m] em “maz”; [i] em “de”; [m] em “máfia”). Para a frase 3, dos 12 gestos, apenas 1 ([ew] do trecho “eu pape”) não apresentou esse padrão. Por fim, para a frase 4, dos 12 gestos apenas 1 ([a] em “papai”) não apresentou esse padrão.

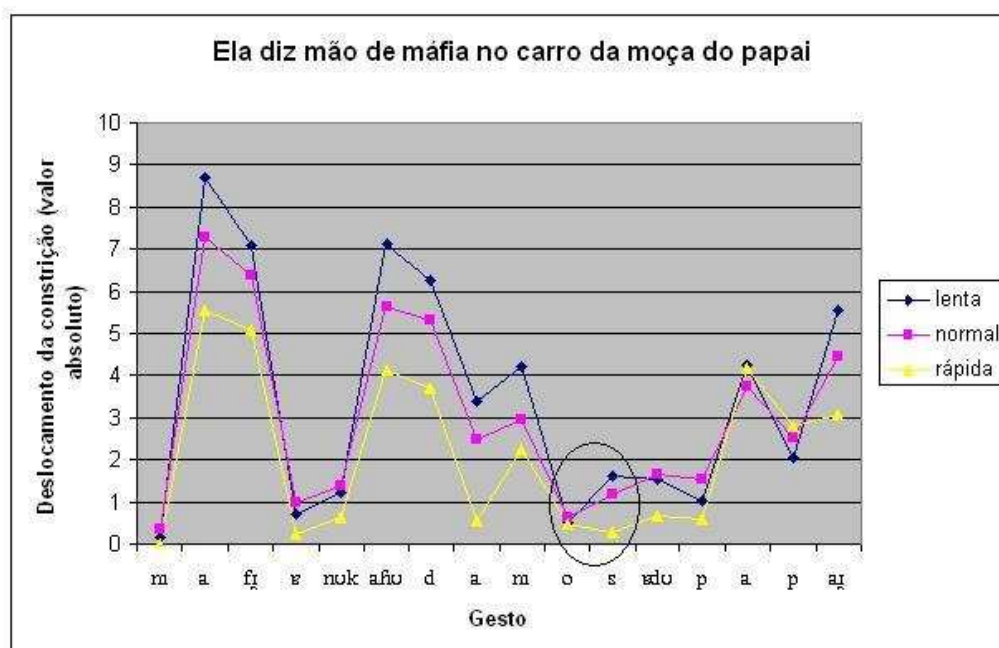


Figura 4.26: Deslocamento da constrição (valor absoluto) para a sentença 1 em três TE.

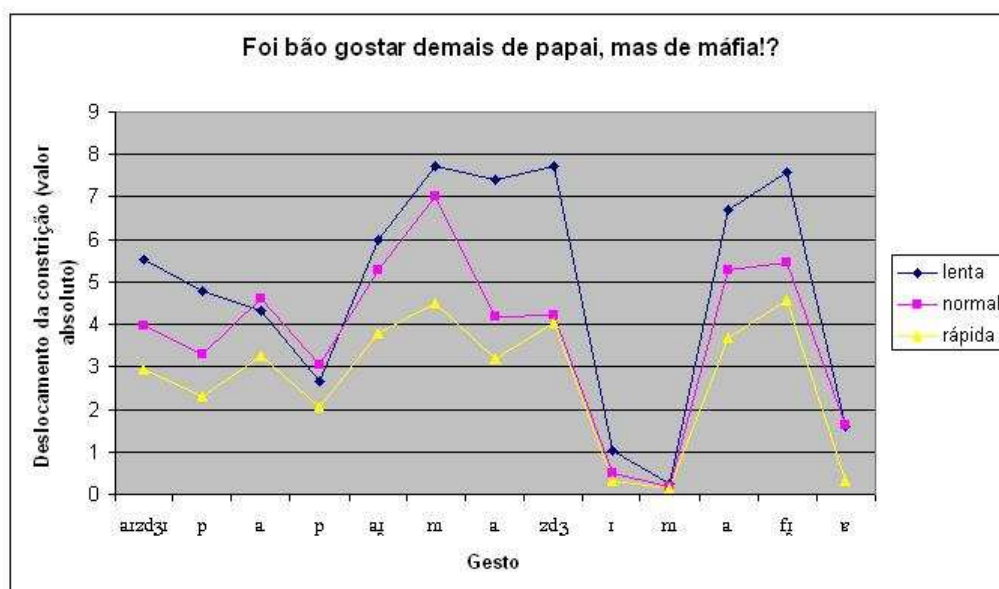


Figura 4.27: Deslocamento da constrição (valor absoluto) para a sentença 2 em três TE.

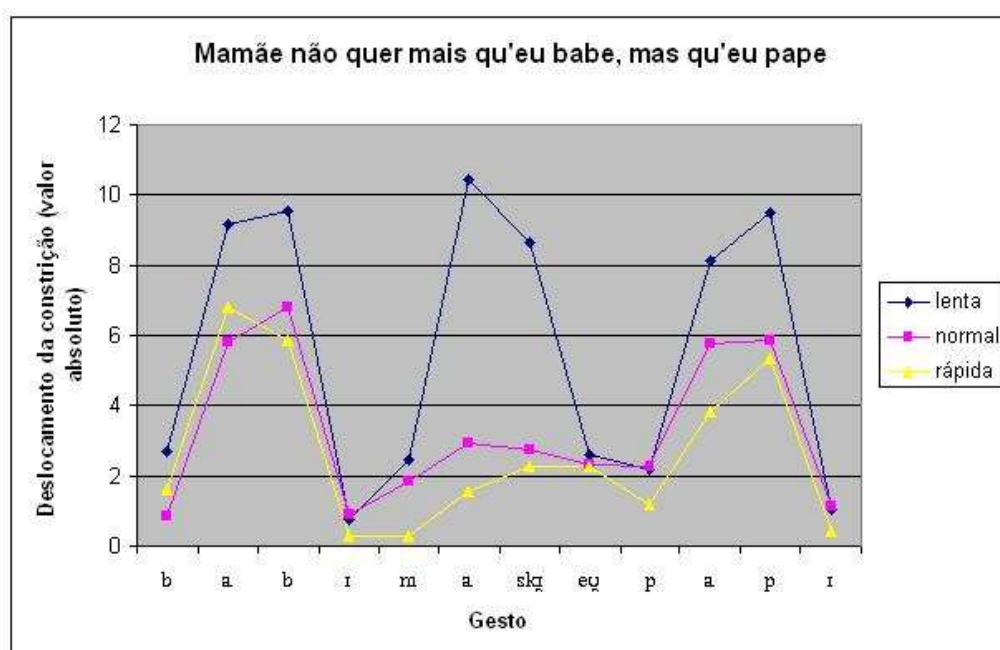


Figura 4.28: Deslocamento da constrição (valor absoluto) para a sentença 3 em três TE.

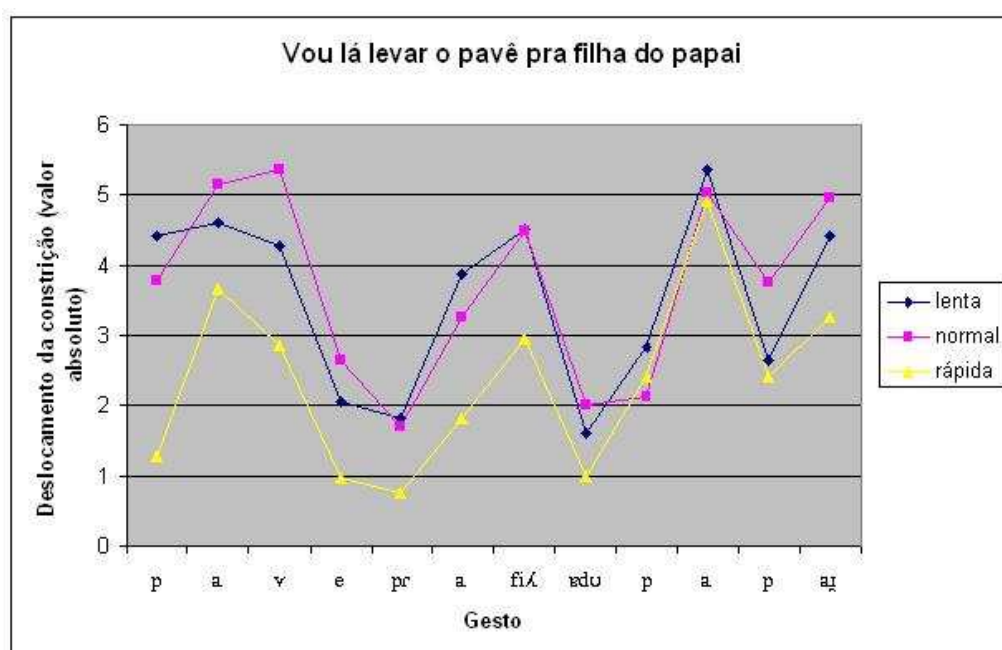


Figura 4.29: Deslocamento da constrição (valor absoluto) para a sentença 4 em três TE.

Na figura 4.26, um outro fato ainda nos chama atenção. A evolução do deslocamento da constrição (valor absoluto) em função do tempo mostra que para todos os gestos há um padrão semelhante de deslocamento nas três TE, exceto para o VV [os], em que para o gesto consonantal [s] não há mais um pico no gráfico (marcado pela elipse na fig. 26). Lembramos que é justamente neste local, na mudança da taxa normal à rápida, que ocorre uma reestruturação rítmica do enunciado (vide seção 4.2.1). Sendo assim, também o deslocamento da constrição poderá fornecer informações, se o sinal articulatório for devidamente normalizado, de onde no sinal articulatório ocorrem reestruturações rítmicas na fala.

4.3.4 Velocidade (pico ou vale)

A análise das figuras 4.30–4.33 abaixo sugere não haver diminuição da velocidade de pico/vale apenas nos gestos componentes dos VVs em acento frasal, pois, apesar de, nas figuras 4.30 e 4.31, ocorrer tal diminuição da taxa lenta à rápida nos VVs reestruturados [os] e [ajzɔɣɪ]; em outras sentenças, como na 3, tal padrão foi encontrado em 9 casos dentre 12 possíveis. Esses dados sugerem que a velocidade diminui para todos os gestos com o aumento da TE. No entanto mais dados articulatórios são necessários, pois obtivemos resultados significativos desse padrão, com velocidade em função da TE, apenas para a

sentença 3 ($F(2,192) = 4,7334$, $p < 0,0099$).

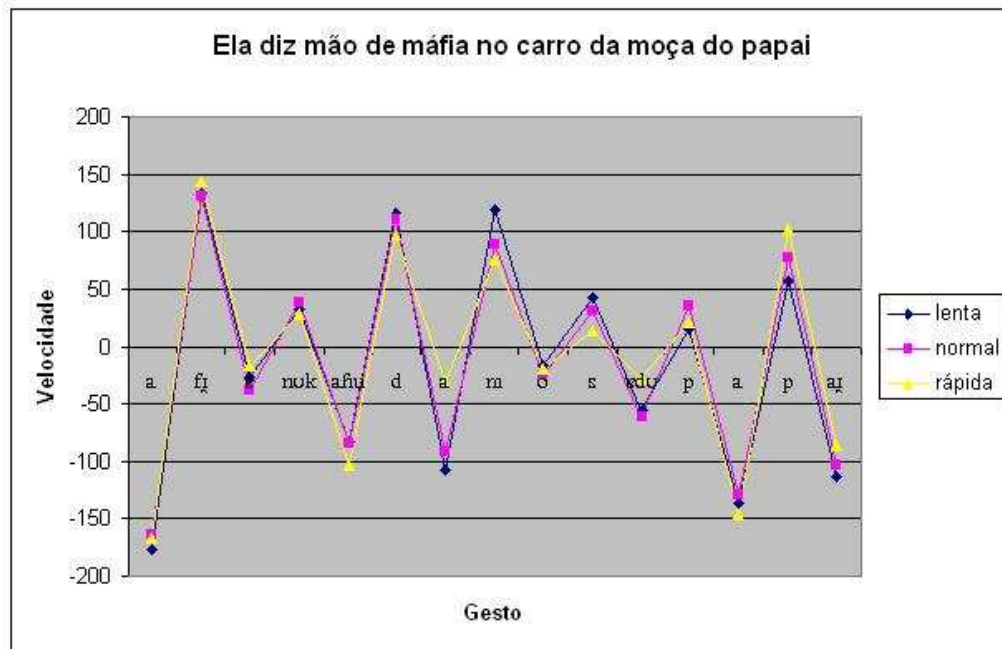


Figura 4.30: Velocidade articulatória para a sentença 1 em três TE.

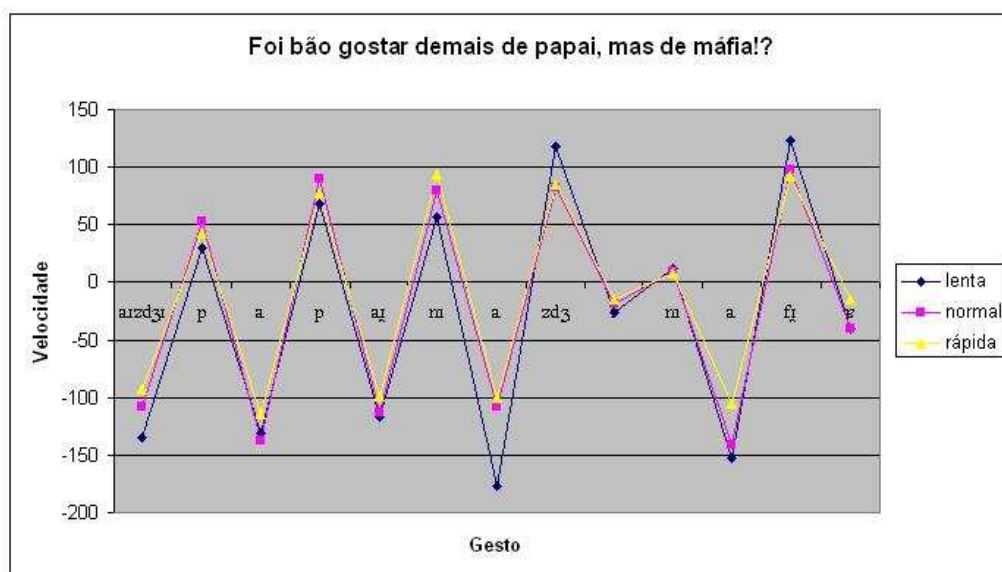


Figura 4.31: Velocidade articulatória para a sentença 2 em três TE.

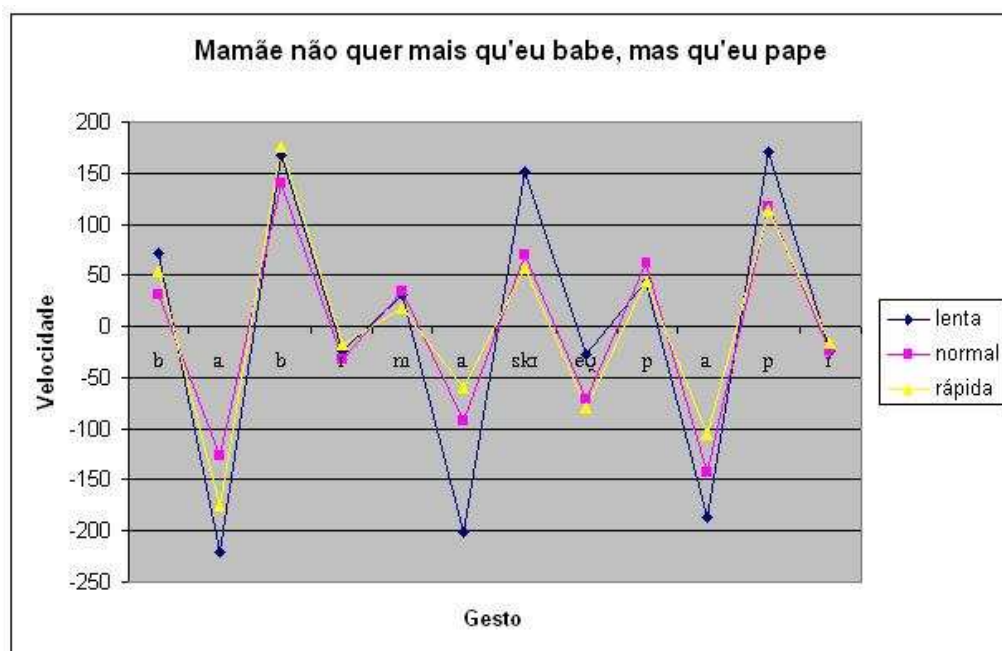


Figura 4.32: Velocidade articulatória para a sentença 3 em três TE.

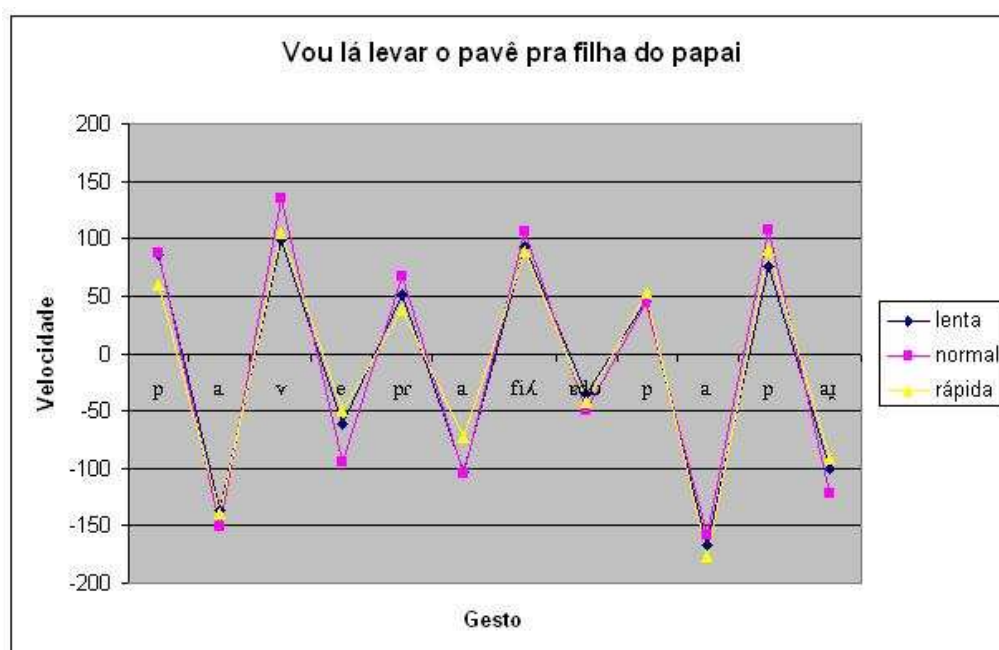


Figura 4.33: Velocidade articulatória para a sentença 4 em três TE.

4.3.5 Duração articulatória

Como era de se esperar, a duração AT dos gestos em função da TE, foi significativa para todas as sentenças, conforme mostra a tabela 4.12. Além disso as figuras 34–37 mostram diminuição da duração articulatória em todos os gestos das sentenças envolvidas, devido aos traçados duracionais das sentenças semelhantes em todas as TE. No entanto, observamos um padrão distinto para os gestos [o] e [s] da sentença 1 (fig. 4.28, marcados pela elipse), em que, na taxa rápida, houve uma mudança de padrão articulatório, justamente o VV que foi reestruturado (vide seção 4.2.1). Sugerimos, portanto, que uma análise da duração articulatória dos gestos com *z-score* e suavização, como

nos estudos acústicos anteriores, pode nos fornecer evidências articulatórias da posição dos acentos frasais. Lembremos que os dados articulatórios analisados aqui não foram normalizados como nos estudos acústicos.

Tabela 4.12: Estatística da duração AT dos gestos articulatórios para todas as sentenças do experimento articulatório prosódico.

Duração AT x TE		
	Anova	p <
Sentença 1	F (2,183) = 13,080	10 ⁻⁵
Sentença 2	F (2,141) = 9,1624	0.0002
Sentença 3	F (2,176) = 52,992	10 ⁻⁴
Sentença 4	F (2,273) = 76,072	10 ⁻⁴

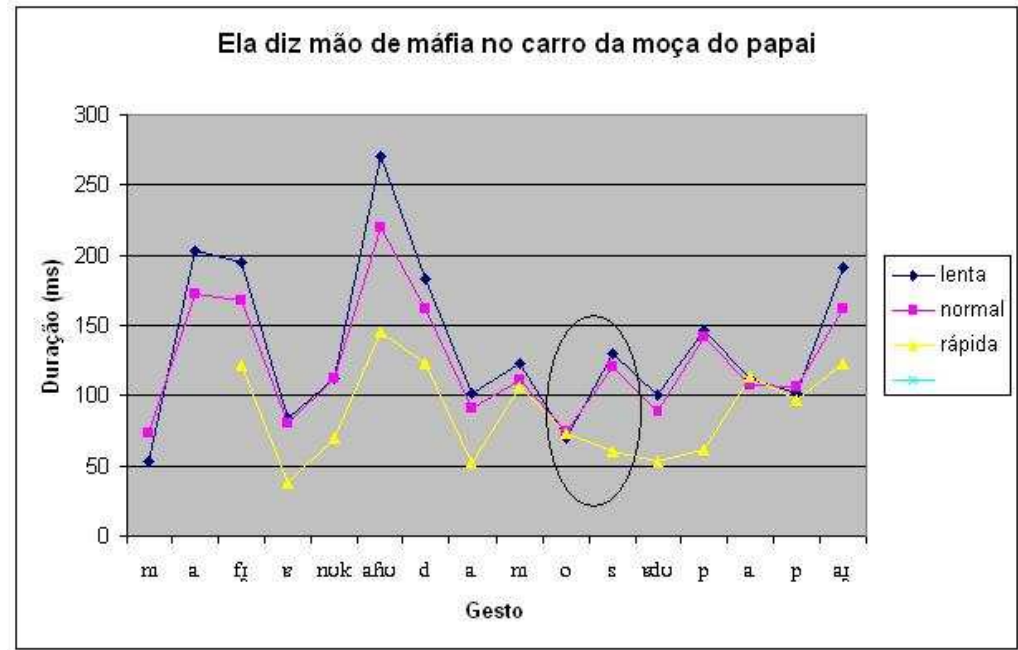


Figura 4.34: Duração articulatória dos gestos para a sentença 1 em três TE.

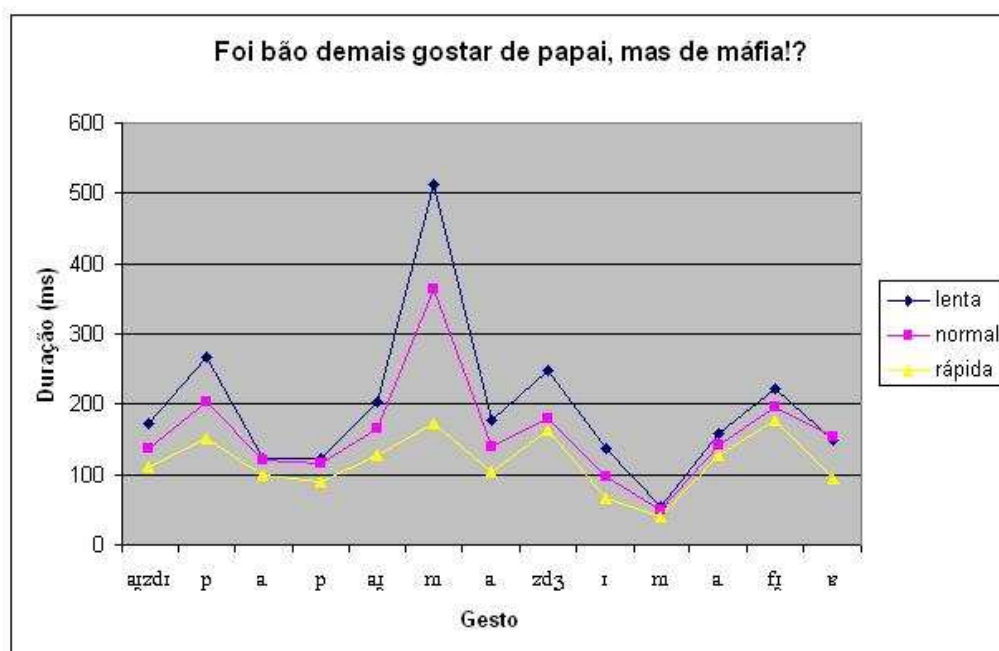


Figura 4.35: Duração articulatória dos gestos para a sentença 2 em três TE.

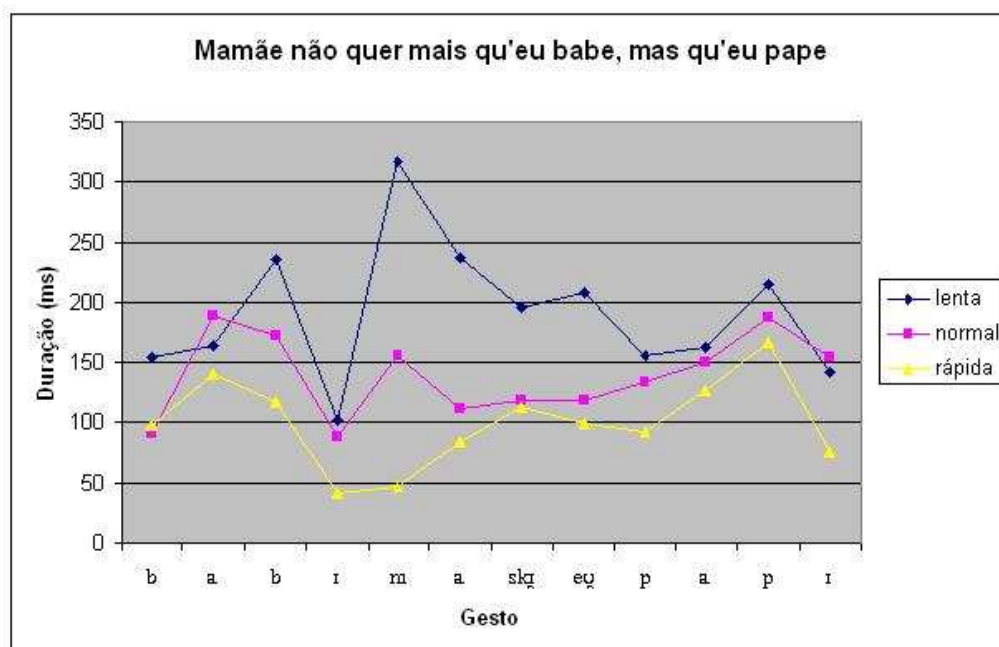


Figura 4.36: Duração articulatória dos gestos para a sentença 3 em três TE.

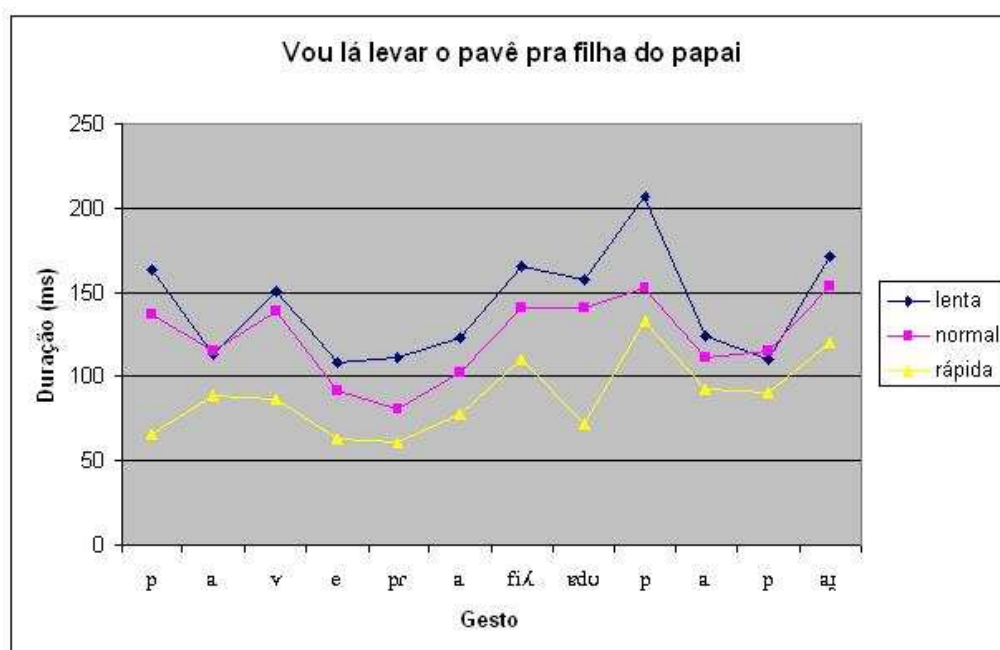


Figura 4.37: Duração articulatória dos gestos para a sentença 4 em três TE.

4.3.6 Tempo-para-velocidade-de-pico proporcional

Uma Anova *One-way* com o tempo-para-velocidade-de-pico proporcional em função da TE mostra ser tal fator invariável para todas as sentenças. Além disso, como mostram as figuras 4.38–4.41, tal fator é indiferente à questão do gesto pertencer a um VV em posição de acento frasal.

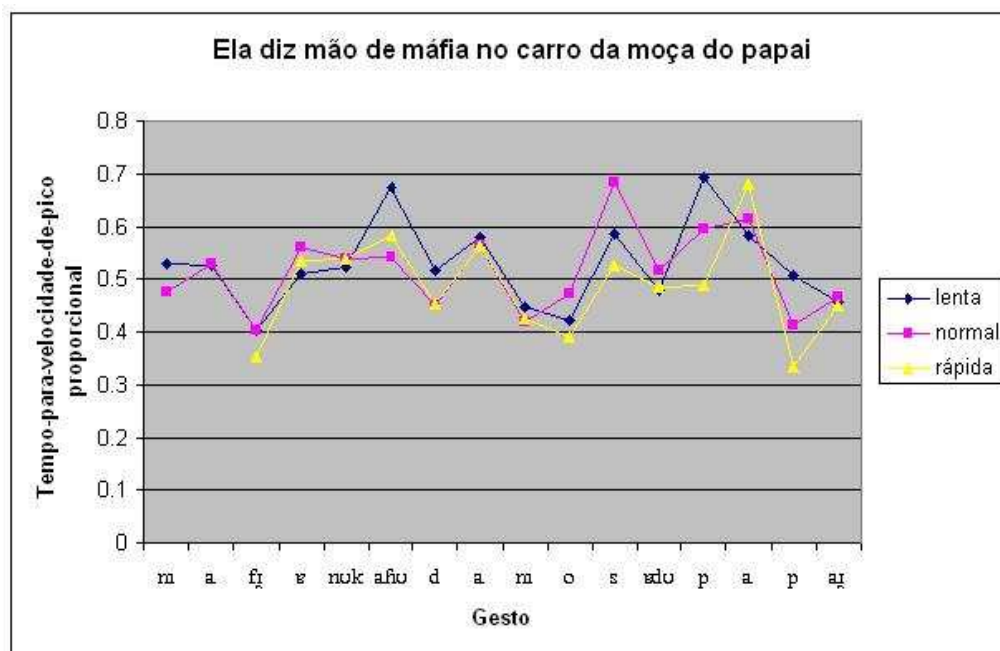


Figura 4.38: Tempo-para-velocidade-de-pico proporcional para a sentença 1 em três TE.

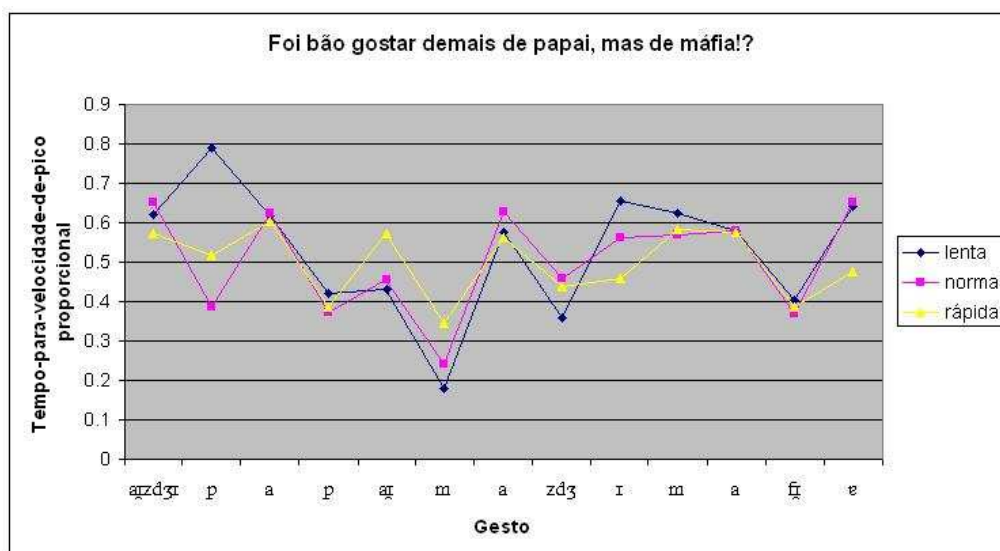


Figura 4.39: Tempo-para-velocidade-de-pico proporcional para a sentença 2 em três TE.

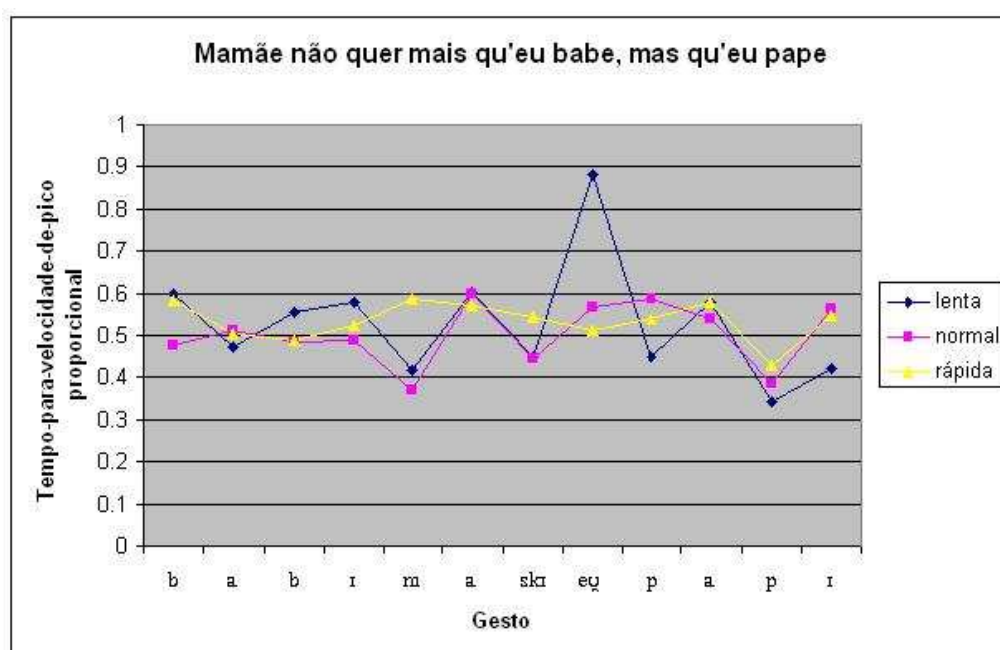


Figura 4.40: Tempo-para-velocidade-de-pico proporcional para a sentença 3 em três TE.

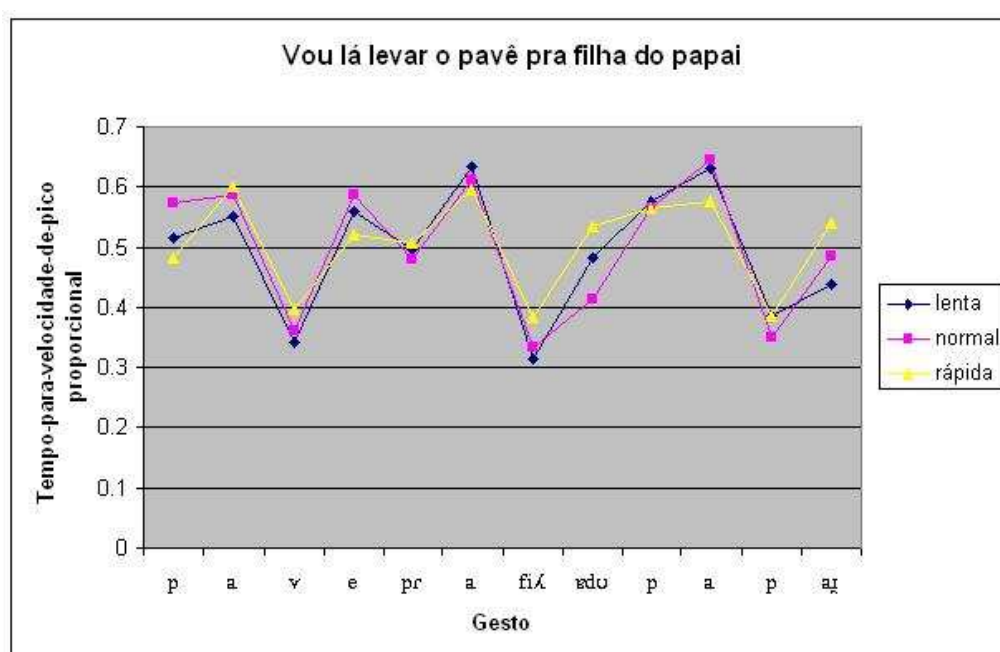


Figura 4.41: Tempo-para-velocidade-de-pico proporcional para a sentença 4 em três TE.

4.4 Conclusão

As análises articulatórias apresentadas neste capítulo mostram que a taxa de elocução afeta todos os gestos dos enunciados de maneira uniforme, independente do mesmos pertencerem a VVs em posição de acento frasal. No entanto, há evidências de que os fatores duração da aceleração (rigidez da mola no MDT), deslocamento da constrição e duração dos gestos articulatórios podem, se devidamente normalizados com, por exemplo, *z-score* e suavização, fornecer subsídios articulatórios para as reestruturações rítmicas do enunciado, analisadas no MDR com base em princípios metodológicos aplicados à duração acústica de unidades VV.

Relembrando, os resultados gerais encontrados neste estudo foram:

- Diminuição da duração da aceleração;
- Diminuição do deslocamento vertical máximo;
- Diminuição do deslocamento do gesto de constrição;
- Diminuição (em módulo) da velocidade de pico/vale;
- Diminuição da duração articulatória (tanto AT quanto medida entre velocidades de pico e vale);
- Tempo-para-velocidade-de-pico proporcional permanece constante.

Terceira Parte: Reestruturações rítmicas no item lexical

Capítulo 5

Variações/mudanças rítmicas no item lexical

5.1 Introdução

Nesta parte da tese serão apresentados estudos acústicos e articulatórios que apontam para reestruturações lexicais oriundas de reorganizações microrrítmicas na estrutura lexical dos enunciados em PB. Especificamente, apresentaremos exemplos de variação/mudança de proparoxítonas a paroxítonas em PB. Para isso, no entanto, primeiramente apresentaremos uma discussão sobre os aspectos diacrônicos e sincrônicos da linguagem (seção 5.1), seguidos de uma resenha bibliográfica sucinta da variação histórica de proparoxítonas a paroxítonas (seção 5.2) e de como a FAR lida com a variação lexical (seção

5.3). Por fim, apresentaremos dois estudos: um acústico e um articulatório (cap. 6), apresentando evidências à luz da TSD de como tal abordagem é também perfeitamente adequada à análise de aspectos sociolingüísticos da linguagem.

5.1.1 Diacronia x sincronia?

Com o advento da Lingüística científica, introduziu-se uma distinção entre Lingüística Sincrônica ou Estática e Lingüística Diacrônica ou Evolutiva (cf. Saussure, 1916). A partir daí o objeto de estudo da ciência lingüística passou a ser majoritariamente a língua, considerada como um sistema lingüístico estático e homogêneo. Segundo Saussure (1916:115), “todas as inovações na fala não têm o mesmo êxito e, enquanto permanecem individuais, não há porque levá-las em conta, pois o que estudamos é a língua; elas só entram em nosso campo de observação no momento em que a coletividade as acolhe.” Nesta abordagem, todos os fatores exteriores à língua foram colocados em segundo plano, como podemos notar abaixo:

a primeira coisa que surpreende quando se estudam os fatos da língua é que, para o indivíduo falante, a sucessão deles no tempo não existe; ele se acha diante de um **estado** (grifo meu). Também o lingüista que queira compreender esse estado deve fazer

“tabula rasa” de tudo quanto produzir e **ignorar a diacronia**

(grifo meu)(Saussure, 1916).

A Gramática Gerativa desenvolvida por Noam Chomsky, a partir da década de 1960, apesar de mudar substancialmente os conceitos lingüísticos introduzidos por Ferdinand de Saussure, conservou a idéia de uma lingüística estática, através dos conceitos *competência* e *desempenho*, correspondentes, respectivamente, grosso modo, a *langue* e *parole*. No entanto, nesta mesma década, principalmente através do trabalho seminal de Weinreich, Labov & Herzog (1968), começou a ser questionada a exclusão da diacronia dos estudos lingüísticos *per se*. Comentando sobre as visões gerativistas do fenômeno da mudança lingüística, esses autores afirmaram que:

(...) Long before predictive theories of language change can be attempted, it will be necessary to learn to see language — whether from a diachronic or a synchronic vantage — as an object possessing orderly heterogeneity. (...) The more linguists became impressed with the existence of structure of language, and the more they bolstered this observation with deductive arguments about the functional advantages of structure, the more mysterious became the transition of a language from state to state. After all, if a language has to be structured in order to function effi-

ently, how do people continue to talk while the language changes,
that is, while it passes through periods of lessened systematicity?
(pp. 99-100).

Em um trabalho subsequente, Kroch (1978) chegou inclusive a propor uma teoria sobre a variação lingüística. Dentre suas principais propostas estão: 1) mudanças fonológicas comuns inconscientes não são de forma alguma arbitrárias, mas, em geral, processos foneticamente motivados; 2) prestígio é um fator secundário na propagação de mudanças fonológicas foneticamente motivadas. O fator principal é o caráter lingüístico, que é a base original da difusão; 3) a atuação principal do prestígio social é inibir processos foneticamente condicionados, tanto de mudança em progresso, quanto de variações estáveis inerentes na fala dos grupos de status alto e daqueles influenciados por estes (p. 21). Entretanto, como o próprio autor afirma, essa teoria “concerns only the phonological aspect of social dialect variation and, therefore, it cannot hope to present a comprehensive theory of variation” (p. 17). Além disso, apesar de ser uma teoria inovadora, o lingüista não a desenvolveu em trabalhos futuros.

Apesar de esses autores considerarem a importância da heterogeneidade lingüística nos estudos lingüísticos, não houve a elaboração de uma teoria lingüística nova que pudesse suplantar a tradicional. A nosso ver, a maior

contribuição desses autores foi a de comprovar que aspectos sociais e lingüísticos *per se* caminham, muitas vezes, juntos. Apesar dessa contribuição, continuaram a trabalhar com os aspectos estritamente gramaticais da língua da mesma maneira que o fizeram seus predecessores.

No entanto, na década de 1980, surgiu uma versão radicalmente diferente da Lingüística Tradicional. Como vimos na primeira parte desta tese, a TSD, com toda sua bagagem empírico-teórica, passa a ser aplicada à linguagem. Nesta perspectiva a dicotomia entre *língua* e *fala* (ou *competência* e *desempenho*) não se coloca, pois tanto aspectos contínuos quanto simbólicos são objetos de estudos da linguagem. Nesse programa de pesquisa a língua passou a ser considerada como um sistema dinâmico auto-organizado (cf. Kelso, Saltzman & Tuller, 1986; Port, 2002), cujas regiões estabilizadas atuariam como símbolos passíveis de controle por parte do sujeito falante. Uma filha desse novo paradigma é a Fonologia Articulatória (vide seção 1.3), estendida mais recentemente para dados do português brasileiro por Albano (2001).

Procuraremos mostrar, aqui, à luz da FAR, como uma lingüística baseada em sistemas dinâmicos é capaz de explicar elegantemente processos de variação que podem vir a constituir mudança lingüística. Tarallo (1990), apesar de não se inserir no programa de pesquisa dinâmico, se afina com o mesmo, pois, segundo o autor, uma teoria de mudança lingüística deve romper com a dicotomia sincronia/diacronia, e tratá-las como parte de um mesmo todo.

Além disso, tal teoria deve constatar o vínculo necessário entre variação e mudança, o que implica na aceitação da história e do passado como reflexos do presente, dinamicamente se estruturando e funcionando (p. 25).

De maneira análoga a Tarallo (*ibidem*), demonstrarei como fenômenos atuais têm relações históricas com fenômenos encontrados no passado e como eles podem lançar luz a fenômenos de mudança lingüística posteriores. Especificamente, nosso objetivo será mostrar como o aumento da taxa de elocução pode vir a engatilhar um processo de mudança de proparoxítonas em paroxítonas no português brasileiro. Com isso pretendemos contribuir para esse complexo problema histórico remanescente de um processo lingüístico iniciado já no latim clássico.

5.1.2 Mudança/variação histórica das proparoxítonas

É amplamente conhecido pelos estudiosos da língua portuguesa o fato de proparoxítonas tenderem ou virem a se constituir como paroxítonas. Este fenômeno está relacionado ao processo lingüístico conhecido como síncope¹ (cf. Câmara Jr.; 1988:220). Na mudança de proparoxítona a paroxítona, ocorre, a longo termo, a síncope da vogal pós-tônica medial dos vocábulos proparoxítonos, como observamos em: a) latim clássico: *saeculum* > *saeclum*, *calidus* > *calmus*, *eremus* > *ermus* (cf. Quednau, 2002); b) latim vulgar:

¹Síncope é a perda de um fonema no interior de um vocábulo.

speculum > *speclum*, *angulus* > *anglus*; *oculus* > *oclos* (cf. citações do *Appendix Probi* comentadas em Silva Neto, 1946:140).

Houve, historicamente, a partir do latim clássico (cf. Nunes, 1969:13), uma mudança gradual de proparoxítonas para paroxítonas, culminando no português arcaico (oriundo do latim vulgar) em raríssimas realizações de proparoxítonas (cf. Nunes, 1973:361; Michaelis de Vasconcelos, 1956:62). Sendo que, mesmo essas formas raras, segundo Haug (1994:42), se tornavam paroxítonas pelo uso. Resumindo, houve uma redução crescente de uso das proparoxítonas da seguinte forma: Latim Clássico > Latim Vulgar > Português Arcaico. No entanto, a partir do século XV, com uma profunda valorização da cultura latina, as proparoxítonas foram reintroduzidas no léxico da língua portuguesa (cf. Michaelis de Vasconcelos, 1956:62), chegando novamente à alternância na língua entre proparoxítonas e paroxítonas (ex.: óculos vs. oculos; aurículo vs. orelha).

A variação entre proparoxítona ~ paroxítona seguiu dois caminhos distintos na mudança do latim ao português: a) lexicalização da forma paroxítona, conforme podemos notar em: *speculum* > *speclum* > espelho; *calidus* > *calmus* > calmo; *apicula* > *apicla* > abelha; b) variação histórica ainda em andamento, como em: século ~ seculo (latim: *saeculum* ~ *saeclum*); ângulo ~ anglo (latim: *angulus* ~ *anglus*); óculos ~ oculos (latim: *oculus* ~ *oclos*).

Pela abordagem lingüística tradicional, considera-se, na variação propa-

roxítona ~ paroxítona, a ocorrência de uma supressão (p. ex. Amaral, 2002:102), perda (p. ex. Câmara Jr., 1988:220; Quednau, 2002: 79) ou queda (p. ex. Williams, 1961:66) da vogal pós-tônica medial. Daí a representação simplificada dessas “paroxítonas” como, por exemplo, “abobra, anális, cosca”. Ressaltamos, contudo, que a análise de tais autores foram feitas, na maioria das vezes, apenas de oitiva. Além disto, esses autores trabalhavam com teorias lingüísticas que privilegiavam o aspecto intuitivo do falante, não levando em consideração o detalhe fonético nas análises. Sendo assim, o trabalho dos mesmos, em seu contexto, tem um valor para os estudos lingüísticos. No entanto, demonstraremos aqui, pela via da FAR, que o que ocorre de fato não é a “perda” de um segmento fônico, e sim uma redução de magnitude dos gestos que compõem a vogal pós-tônica medial. Em outras palavras, não há queda de segmentos, mas uma produção acústico-articulatória reduzida do referido segmento que pode vir a ser percebida/reinterpretada como segmento transicional dos segmentos adjacentes, estes últimos menos reduzidos, como veremos na próxima seção.

5.1.3 Variação lexical na FAR

Como vimos anteriormente, na FAR, não há dicotomia entre nível fonético e fonológico e a unidade mínima de pesquisa não é mais o traço distintivo,

e sim o gesto articulatório, definido abstratamente pelo modelo funcional da tarefa (Kelso, Saltzman & Tuller, 1986). Esse modelo fonológico lida, elegantemente, i.e., segundo um mesmo princípio, com fenômenos como a coarticulação, a variação alofônica e fenômenos subsegmentais diversos encontrados na aquisição da linguagem e nas patologias de fala.

Além disso, os gestos têm extensão temporal e espacial (cf. fig. 5.1), sugerindo que os mesmos não desaparecem (ou aparecem) nas reestruturações silábicas, ou seja, eles estão sempre presentes, e a percepção de perda ou inserção de um determinado “segmento” é devida a: (i) uma redução (ou aumento) da magnitude de um certo gesto (vide Albano, 2001, para uma explicação da vogal epentética [i] em português brasileiro, como em “pa|i|z”, devido ao atraso do gesto de constrição da língua, em sequência de vogal seguida de /s/ em final de palavra); (ii) um aumento (ou redução) da sobreposição entre gestos, se imaginarmos, por exemplo, o aumento da sobreposição dos gestos de “para” (fig. 5.1a), que passaria a ter uma organização semelhante a do início de “prata” (fig. 5.1b). Em (5.1b) percebemos claramente a sobreposição dos gestos “amplo faringal” e “fechado alveolar”, resultando na mudança de “para” para “pra”, comum em muitos dialetos e até realizado ortograficamente por diversos autores da língua portuguesa. A magnitude desses gestos, em nosso ponto de vista, como indicaremos nos próximos capítulos, é condicionada por fatores sociais, psicológicos, individuais e fonéticos, que podem ser eliciados

pela variação da taxa de elocução. Sendo assim, conforme Byrd (2003), “the postulation of abstract, dynamically-specified, and underlying stable phonological units, constrains the possibilities for accounting for variability in word forms.”

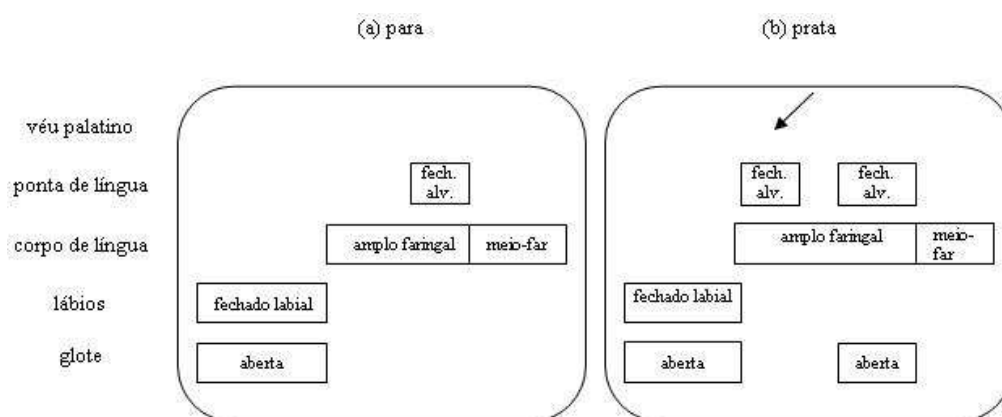


Figura 5.1: Diferenças gestuais entre as palavras “para” e “prata” segundo o modelo da FAR. Temos nesta figura: (a) pauta gestual de “para”; (b) pauta gestual de “prata”.

Por este motivo acreditamos que a Sociolingüística poderia se beneficiar em muito de uma Lingüística baseada nos sistemas dinâmicos, como a FAR, pois, diferentemente da Lingüística tradicional, a variação lingüística é inerente ao paradigma de pesquisa e a dicotomia entre nível fonético e fonológico não se dá.

Nos próximos capítulos, mostraremos como a FAR dá conta da variação ou caráter dúbio entre proparoxítonas e paroxítonas, e pode explicar a mudança da primeira forma para a segunda.

5.2 Estudos acústicos

Neste trabalho mostraremos como a variação/mudança lingüística pode ser disparada ou iniciada a partir da variabilidade fônica/lexical condicionada pela perturbação da fala através da TE.

Abaurre-Gnerre (1981:340) diz ser a TE desencadeadora de processos fonológicos. Em Abaurre-Gnerre (1979:14) temos um exemplo da influência da taxa na seguinte regra fonológica: “[-baixo, -acento] > [+alto, -tenso]; condição: tempo é [-largo]”. Em outra parte deste trabalho (p. 16), a autora cita realizações fonéticas de “teatrinho”, dialeto capixaba, dependentes da taxa, como: a) largo: [tʃiatrĩ̃o]; b) andante: [tʃiatrĩ̃u]; c) allegreto: [tʃjatrĩ̃ʷ]; d) presto: [tʃatrĩ̃:]².

Nos dois estudos citados acima, a lingüista diz que há formas lexicais específicas para cada taxa de elocução. No entanto, ela admite casos de “apagamento” de “i”, na fala de crianças de 7-12 anos, classe média/alta, em palavras como [pkə'lɛ, 'psinɐ], em taxa lenta (chamada no texto de fala cuidadosa) (p.328).

Meireles (2001) relata em sua dissertação, semelhantemente ao trabalho de Abaurre-Gnerre, a influência da TE na variação/mudança lingüística.

Como exemplo desta influência, observemos a figura 5.2.

²As taxas “largo, andante, allegreto, presto” correspondem, consecutivamente, a “lento, normal, rápido, muito rápido”.

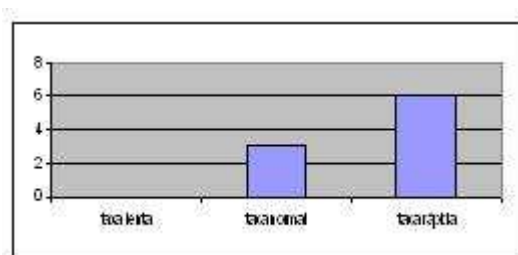


Figura 5.2: Número de realizações de [sʲ] ([s] palatalizado, em **negrito** na sentença), em três taxas de elocução, da sentença: “Fale numa **velocidade** moderada”.

Nesta figura vemos o aumento gradativo da realização de /si/ como [sʲ]. Na taxa lenta não tivemos nenhuma realização, na taxa normal, três, e na taxa rápida, seis³. Ou seja, aumentando-se a TE, as palavras são pronunciadas mais coarticuladamente e sofrem, por isso, influência de restrições articulatorio-temporais espontâneas. No entanto, como já havia sido previsto por Abaurre-Gnerre (1979:328), formas reduzidas, como [sʲ], não ocorrem somente nas taxas rápidas. Apenas ocorrem com uma frequência maior do que em outras taxas.

Esses são apenas alguns exemplos de análises qualitativas, baseados em teorias lingüísticas tradicionais, de ser a realização lexical influenciada pela TE. Entretanto, no estudo acústico descrito a seguir, mostraremos como a FAR, baseada na TSD, é capaz de explicar de maneira bastante simples e evidente a influência da TE em um processo de mudança/variação clássica

³Neste estudo, como trabalhos com a Fonologia Métrica, e também por usarmos o metrônomo na aquisição das taxas, as análises foram feitas de oitiva, e, portanto, são apenas qualitativas.

da língua portuguesa: a mudança de proparoxítonas a paroxítonas. Como vimos na seção anterior, esse processo tem suas origens na mudança do latim clássico para o vulgar e deste para o português (ex.: “apicula > apicla > abelha”, “teneru > tenru > tenro”) e continua até hoje (ex.: “córrego > corgo”, “fósforo > fosfro”, “música > musca”).

5.2.1 Metodologia

Os informantes escolhidos foram quatro do dialeto paulista e quatro do dialeto mineiro, e representavam inequivocadamente as características peculiares de seus dialetos, ou seja, não possuíam influência direta de dialetos de outros estados. Sendo assim, utilizamos quatros informantes nascidos e criados em São Paulo - SP: dois do gênero feminino (MM, LM) e dois do gênero masculino (GB e FN). Os outros informantes foram de regiões que são consideradas como falantes típicas do dialeto mineiro. Um desses informantes foi de Belo Horizonte (DF, gênero feminino), um de Mariana (JP, gênero masculino) e outros dois de Conselheiro Lafaiete (JM, gênero feminino e RM, gênero masculino). Todos os informantes são estudantes universitários de graduação e/ou pós-graduação na Universidade Estadual de Campinas (UNICAMP), exceto RM com 2º grau completo. A faixa etária foi de 20–30 anos, e não apresentavam nenhum comprometimento fonatório, neurológico ou mental

aparente.

As gravações foram realizadas em uma sala isolada acusticamente do LAFAPE/IEL/UNICAMP. Os dados foram gravados com um microfone dinâmico Sennheiser diretamente em disco rígido através do programa CSL da Kay Elemetrics e digitalizados a uma taxa de 22.050 Hz.

Um corpus composto pelas palavras “abóbora, análise, cócega, fôlego, fósforo e sábado”, inscritas na frase veículo “Digo _____ baixinho”, foi utilizado, a fim de se observar como o aumento da taxa de elocução poderia estar influenciando possíveis reestruturações destas palavras de proparoxítonas para paroxítonas. As palavras que foram analisadas foram escolhidas a partir de nossa intuição de que são bastante freqüentes na fala coloquial dos brasileiros⁴. A hipótese inicial seria de que encontraríamos (principalmente em taxas rápidas) formas como [a.'bɔ.brɐ, a.'na.lɨs, 'kɔs.kɐ, 'fɔs.frɯ], pois são formas lexicais reduzidas que já são percebidas como tais em diversos dialetos do português brasileiro (PB). Acreditamos que tais formas reduzidas são encontradas após ressilabificações decorrentes do grau de coarticulação maior em taxas mais rápidas. As outras duas palavras restantes (fôlego, sábado) foram também utilizadas, por não podermos prever de antemão suas realizações fonéticas na forma reduzida. Diferentemente das outras pala-

⁴Não poderíamos usar um critério objetivo para a freqüência, pois, até onde vai o nosso conhecimento, não existe um corpus em que se tenha feito tal computação para a fala. Sabemos apenas que pesquisadores do LAFAPE têm o objetivo de fazer tal computação baseados nas palavras do projeto NURC.

vras, hipotetizamos que suas possíveis formas paroxítonas ['fol.gʊ, 'sab.dʊ] não ocorreriam devido à violação de restrições fonotáticas dos dialetos estudados. Não estamos dizendo serem tais formas impossíveis de existir. Nos dialetos do sul, por exemplo, é pronunciado o [l] em posição de coda; e, em formas cultas, como “abdicar”, são pronunciadas duas consoantes oclusivas em seqüência. Em outras palavras, afirmamos, portanto, que as formações silábicas de [l] em coda e a seqüência [b] em coda seguida de [d] em ataque são muito raras, talvez inexistentes, na fala dos dialetos estudados.⁵

O objetivo desse *corpus* seria, pois, o de controlarmos palavras com possíveis reestruturações lexicais (as quatro primeiras) e palavras sem reestruturações esperadas (as duas últimas). Além disso, pretendíamos observar se o tamanho da palavra influenciava as reestruturações das palavras. Por este motivo também foram utilizadas palavras trissilábicas e quadrissilábicas em nosso *corpus*.

Utilizamos também um outro corpus composto por palavras sem forma ortográfica oficial: [a.'bɔ.brɐ, a.'na.lis, a.'na.liz⁶, 'kɔs.kɐ, 'fɔs.frʊ], gravadas pelos mesmos informantes, a partir de forma ortográfica explicitando a forma reduzida. Estas palavras poderiam servir como parâmetro de controle para

⁵Além disso, lembremos que formas mais extremas, como ['sa.bʊ], também seriam possíveis (vide Amaral, 1955, para mais exemplos de “pronúncia popular”; e Amaral, 2000, para um estudo sociolinguístico da variação/mudança de proparoxítonas a paroxítonas.)

⁶Gravamos [a.'na.lis] e [a.'na.liz], pois poderia ser que “análise” fosse pronunciada como [a.'na.liz], ao invés de [a.'na.lis].

compararmos as durações dos segmentos, bem como a estrutura formântica das vogais.

As palavras escolhidas foram lidas em três taxas de elocução distintas. As instruções para obtenção das três taxas foram as seguintes: a) lenta: fale o mais lentamente possível preservando a estrutura prosódica das frases e sem inserir pausas entre as palavras; b) normal: fale da forma mais confortável para você; c) fale o mais rapidamente possível sem cometer distorções na fala. Foram apresentadas aleatoriamente aos informantes em uma folha de papel e lidas dez vezes seguidas continuamente em cada uma das três taxas. Por exemplo, lia-se “abóbora” na frase-veículo 10 vezes seguidas na taxa normal, 10 na lenta e 10 na rápida, e assim por diante. Esta ordem de taxas foi a mais eficiente para obtenção de diferentes taxas, pois começava-se pela taxa mais confortável para o informante, e, tomando esta como referência, lentificava-se e ou acelerava-se a taxa⁷. As únicas observações para os informantes eram que fizessem uma pequena pausa entre as frases, a fim de facilitar segmentações futuras e de que não se preocupassem em contar o número de frases (tarefa nossa). A repetição contínua da mesma frase auxiliou a diminuir um pouco a formalidade do experimento e incitar uma fala um pouco mais espontânea, pois, na repetição, os informantes, muitas vezes, perdiam a concentração na

⁷Tentamos gravar também na ordem lenta, normal e rápida, mas, desta forma, alguns informantes tiveram dificuldades em executar a tarefa proposta.

tarefa proposta e falavam desvinculando-se um pouco da língua escrita⁸, ou seja, passavam a repetir as palavras de memória. Antes de iniciarmos as gravações, ainda fazíamos um treinamento para que tivéssemos certeza de que taxas perceptualmente distintas fossem realizadas.

Após a gravação, segmentamos, no programa Praat, as frases e palavras através dos seguintes procedimentos. As frases foram segmentadas do início da vogal [i] de “digo” até o início de [i] de “baixinho”⁹, como podemos ver na figura 5.3, para que pudéssemos comprovar estatisticamente a diferença entre as taxas. Aplicamos, então, através do programa Statistica da Statsoft, um teste *ANOVA multi-factorial* de três fatores para ver os resultados gerais entre as três taxas e um teste *post-hoc Scheffé* para verificarmos diferenças internas entre as taxas. Já as palavras foram delimitadas entre o início da primeira vogal e o fim da última vogal (cf. fig. 5.3).

As delimitações acima foram realizadas semi-automaticamente em três etapas:

1. utilizamos o script “Beat Extractor” do Praat (apêndice A.2), desenvolvido por Barbosa (2006), com base no programa homônimo de Fred Cummins (Cummins & Port, 1998). Este programa marca automa-

⁸Alguns informantes relataram, espontaneamente, ao final do experimento, que eles perdiam, devido à repetição, a concentração na tarefa proposta.

⁹Optamos por esta solução, pois, na maioria dos casos, o [u] final não era identificado nos espectrogramas devido a sua pequena intensidade.

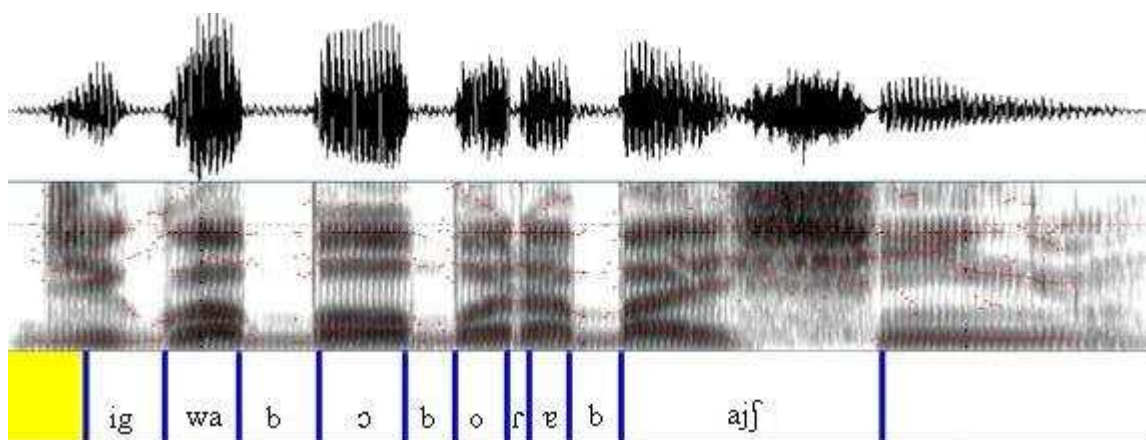


Figura 5.3: Exemplo de segmentação da frase “Digo abóbora baixinho” na taxa lenta no programa Praat. A primeira janela representa a forma de onda; a segunda, o espectrograma; a terceira, a segmentação da frase para que fosse extraída a duração da frase e dos segmentos. No espectrograma, temos: eixo x, duração dos segmentos (ms); eixo y: frequência dos segmentos (Hz).

ticamente o início das vogais. Avaliado por seus autores, observou-se uma economia de aproximadamente 70 % no trabalho de segmentação. Eventualmente, o programa marcava alguns inícios erroneamente e deixava de marcar outros. Sendo assim, algumas vezes, tínhamos de corrigir manualmente a segmentação. Tínhamos também de marcar manualmente os finais dos segmentos, pois o programa não o faz, já que foi concebido para delimitar inícios de vogais;

2. depois de devidamente segmentado, utilizamos o *script* “Duração” (apêndice A.3), que extrai o valor em segundos das durações dos intervalos de uma camada (*tier*) da janela de edição de som do Praat, previamente

etiquetada, e gera um arquivo .txt, que pode ser usado em programas de planilha eletrônica, como *Microsoft Excel* ou *OpenOffice Math*;

3. com o arquivo .txt, gerado pelo script “Duração”, contamos o número de sílabas fonológicas das frases e dividimos pela duração total da frase, obtendo o resultado em sílabas/segundo, que foram, posteriormente, submetidas a análises estatísticas no programa *Statistica* da *Statsoft*.

5.2.2 Análise dos dados

Antes de iniciarmos a análise era preciso verificar, de fato, se as três taxas eliciadas eram distintas. A análise estatística *ANOVA One-Way* de variáveis independentes revelou uma diferença significativa (para $\alpha = 0,05$) entre as taxas, para todas as palavras de cada informante, que foram confirmadas através do teste *post-hoc Scheffé*.

A fim de observarmos se a taxa de elocução poderia exibir um grau de co-articulação inter-gestual que sugerisse uma mudança de proparoxítonas para paroxítonas, fizemos transcrições fonéticas de todas as palavras estudadas (inclusive repetições numa mesma taxa), para todos os informantes, resultando em um total de 720 transcrições (8 informantes x 6 palavras x 3 taxas x 5 repetições). Estas transcrições permitiram o estudo inicial de processos de redução, que foram verificados inicialmente pelo cálculo da duração de

cada vogal da palavra proparoxítona.

Depois de transcrevermos todas as palavras, para cada um dos informantes, observamos que não havia como prevermos como uma palavra seria falada, ou seja, não conseguíamos, na maioria dos casos, dizer, por exemplo, que a palavra “abóbora” seria falada por um dado informante como [a'bɔbɔrɐ] na taxa lenta, e [a'bɔb^ɐrɐ] nas taxas normal e rápida. Em outras palavras, não podíamos dizer, categoricamente, que em determinada taxa uma palavra foi produzida como paroxítona, mesmo porque o grau de redução que cada taxa incita depende de seu valor numérico, isto é, depende de quantas unidades lingüísticas por unidades de tempo (e.g. sílabas/s) estão sendo efetivamente produzidas. Isso demonstra os limites da transcrição fonética.

Como estamos analisando como taxas mais rápidas influenciam reestruturações rítmicas com possibilidade de reestruturação lexical, primeiramente veremos se há relação entre taxas rápidas e as variáveis informante, gênero e palavra.

Diferenças de duração na taxa rápida inter-informantes

Através da inspeção da fig. 5.4, verificamos que os informantes mais rápidos foram consecutivamente: 1) JM (MG); 2) JP (MG); 3) RM (MG); 4) GB (SP); 5) FN (SP); 6) DF (MG); 7) MM (SP); 8) LM (SP) com diferença significativa, tomando-se informantes em função da TE ($F(7,40) = 27,3$;

$p=10^{-5}$). Através de um teste *post-hoc* *Scheffé*, observamos que JM (MG) não se diferencia estatisticamente de JP (MG), RM (MG) e GB (SP), e que MM (SP), LM (SP) e DF (MG) também não se diferenciam estatisticamente entre si. Isso separa três grupos de falantes: a) os mais rápidos: JM (MG), JP (MG), RM (MG), GB (SP); b) um intermediário: FN (SP); c) os mais lentos: DF (MG), MM (SP) e LM (SP). Esperamos, pois, que os falantes mais propensos à paroxitonização sejam, de acordo com nossa hipótese, os mais rápidos (75% de mineiros).

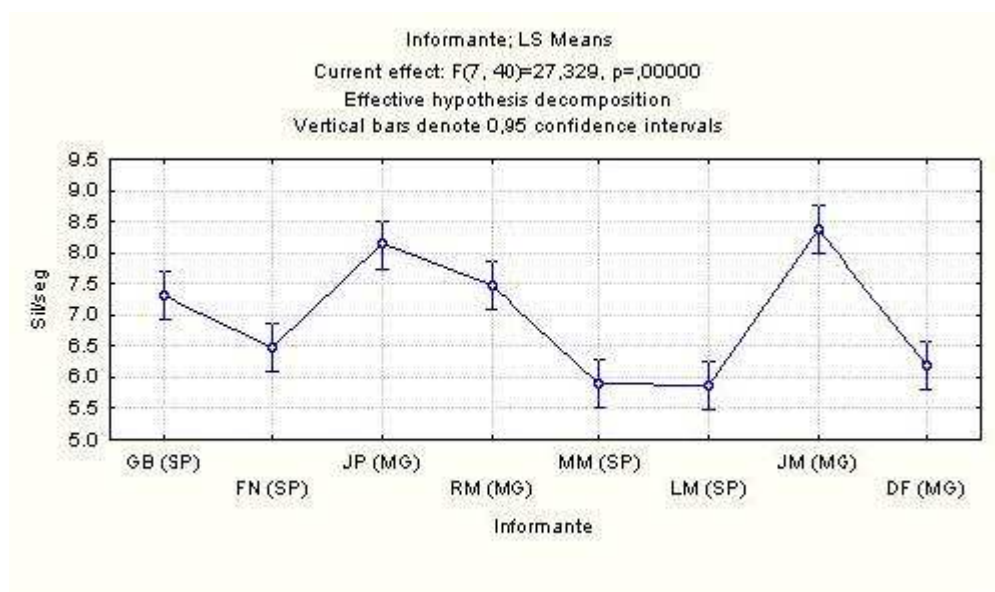


Figura 5.4: Taxa de elocução dos informantes na taxa rápida.

Nossa hipótese foi confirmada posteriormente, através de análises da duração, pois os falantes que obtiveram as formas lexicais mais reduzidas na taxa rápida foram justamente os dois mais rápidos — JM (MG) e JP (MG),

conforme podemos observar na figura 5.5. JM (MG) produziu, por exemplo, “fôlego” como [folegʷ], e JP (MG) produziu, por exemplo, “cócega” como [kʷsgə].

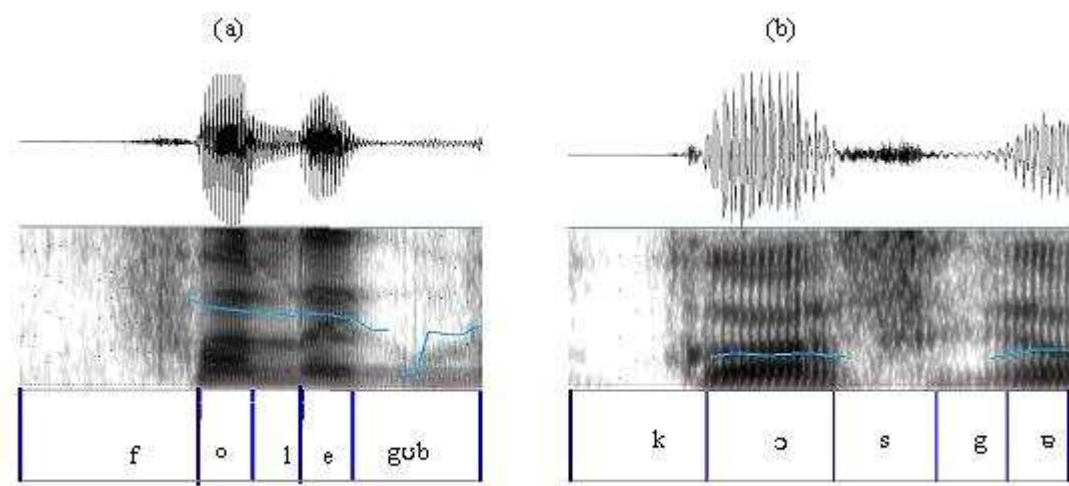


Figura 5.5: a) Espectrograma de [folegʷ], taxa rápida, informante JM (MG); e b) Espectrograma de [kʷsgə], taxa rápida, JP (MG).

Entretanto em informantes como GB (SP), um dos falantes mais rápidos, não encontramos formas tão reduzidas. Isso nos mostra que há estratégias diferentes a serem adotadas para se falar em taxas mais rápidas, ou seja, os falantes utilizam reduções na taxa rápida com diferentes graus de sobreposição dos gestos articulatórios.

A fim de avaliarmos se os diferentes graus de redução se diferenciam também inter-dialetos prosseguimos com a análise relatada a seguir.

Diferenças de duração na taxa rápida inter-dialetos

Houve uma diferença extremamente significativa entre os dialetos mineiro e paulista na taxa rápida [$F(1,70) = 25,3; p < 10^{-5}$]. O mesmo, porém, não se pôde dizer da taxa normal e lenta, que não mostraram diferenças significativas entre os dois dialetos. No entanto, o dialeto mineiro, perceptualmente, parece mais rápido que o dialeto paulista, devido às realizações fonéticas mais reduzidas para o primeiro.

Os falantes do dialeto paulista realizaram, por outro lado, como poderíamos esperar, as mesmas palavras com muito menos variações e reduções do que as do dialeto mineiro. Houve casos, por exemplo, como os de GB (SP) e LM (SP), que produziram praticamente apenas uma realização fonética semelhante para cada uma das taxas. Por exemplo, GB (SP) produziu, em todas as taxas, “abóbora” como [a'bɔbɔɾɐ], e LM (SP) produziu, em todas as taxas, “cócega” como [kɔseɡɐ, 'kɔseɡɐ]. Esses informantes, além de produzirem menos reduções, variavam menos do que os mineiros as suas formas lexicais (em torno de duas ou uma produção lexical por taxa), como podemos notar através da fig. 5.6, que mostra aparente variação livre na produção de /z/ na palavra “análise” como fricativa (a) e como aproximante (b).

A evidência de uma grande complexidade nas estratégias de que cada informante se serve para falar mais rápido é visível no fato de que a ordem

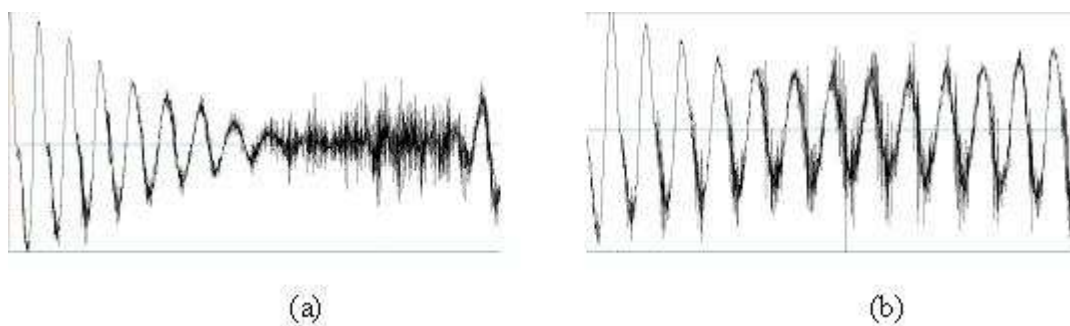


Figura 5.6: Forma de onda (a): representação de [z] vozeado da palavra “análise” [a'nalɪzɪ] de MM (SP), taxa normal. Forma de onda (b): representação de [z] aproximante da palavra “análise” [a'nalɪzɪ] de MM (SP), taxa normal.

de mais rápido para mais lento não implica, necessariamente, em categorias com menos reduções, como é o caso de RM (MG), que na taxa lenta foi o mais lento de todos os informantes (2,98 sílabas/s), na taxa normal foi o segundo mais lento (4,9 sílabas/s), mas, na taxa rápida foi um dos mais rápidos (7,48 sílabas/s). Esse foi mais um dos motivos para considerarmos a taxa rápida como um índice para avaliar qual o dialeto que mais tende a produzir taxas numericamente mais elevadas. Afinal de contas, a instrução para esta taxa, como vimos na metodologia, era “fale o mais rapidamente possível sem cometer distorções na fala”.

Outra pergunta que se pode fazer é se a estratégia de aumento de taxa de elocução é a mesma entre homens e mulheres, investigada a seguir.

Diferenças de duração intergêneros entre as taxas

O gênero masculino foi significativamente mais rápido que o gênero feminino, levando em consideração todas as três taxas juntas ($F(1,142) = 6,8$; $p < 0,01$). Nossos dados mostraram que, em média, os homens (5,8 sílabas/s) falaram 11,5 % mais rápido do que as mulheres (5,2 sílabas/s). Resultado já relatado para o inglês americano por Byrd (1994:43).

O fato de o gênero masculino ser mais rápido do que o gênero feminino já estava inclusive implícito na análise da fig. 5.4, pois, dos cinco informantes mais rápidos, quatro eram do gênero masculino; e todos os três informantes mais lentos eram do gênero feminino. A única exceção a esses dados foi JM (MG), que, em termos gerais, contrariamente ao averiguado por nossas análises, foi o informante mais rápido de todos. Acreditamos, por esse motivo, mais uma vez, que além de fatores condicionadores de mudança ativados por taxas rápidas, gênero e dialeto, devemos também considerar fatores individuais ao estudarmos aspectos relacionados à variação e/ou à mudança lingüística (cf. Oliveira, 1992:39).

Questionamo-nos ainda se o tamanho da palavra (trissilábica ou quadrisilábica) poderia interferir nos processos de redução gestual, conforme a próxima análise.

Diferenças de duração inter-palavras

Não houve diferença significativa de duração entre as palavras em nenhuma das três taxas analisadas. Isso implica que, hipoteticamente, todas as palavras estudadas estariam, em princípio, sujeitas aos mesmos efeitos relacionados à mudança da taxa de elocução. Analisamos esta questão na próxima seção.

Propensão a reduções distintas nas três posições acentuais

Em um primeiro momento de nossa análise procuramos verificar se a vogal pós-tônica medial se reduzia mais do que a pós-tônica final em todas as palavras de nosso corpus, fato que poderia explicar o porquê de a mudança de proparoxítonas a paroxítonas ocorrer a partir de um processo de redução maior (mais acentuado) na posição medial das palavras. Uma *Anova One-Way* com as palavras e respectivas sílabas em função da TE não mostra, entretanto, resultados estatisticamente significantes entre as pós-tônicas mediais e finais, quando todas as palavras foram consideradas para análise.

Entretanto, pela figura 5.7, verificamos haver padrões distintos de redução vocálica de acordo com o tipo da palavra. Percebemos, aqui, claramente a ocorrência de dois padrões distintos: 1) pós-tônica medial (PTM) reduzindo mais do que pós-tônica final: abóbora, cócega e fósforo; 2) pós-tônica final

(PTF) reduzindo mais do que pós-tônica medial: análise, fôlego e sábado.

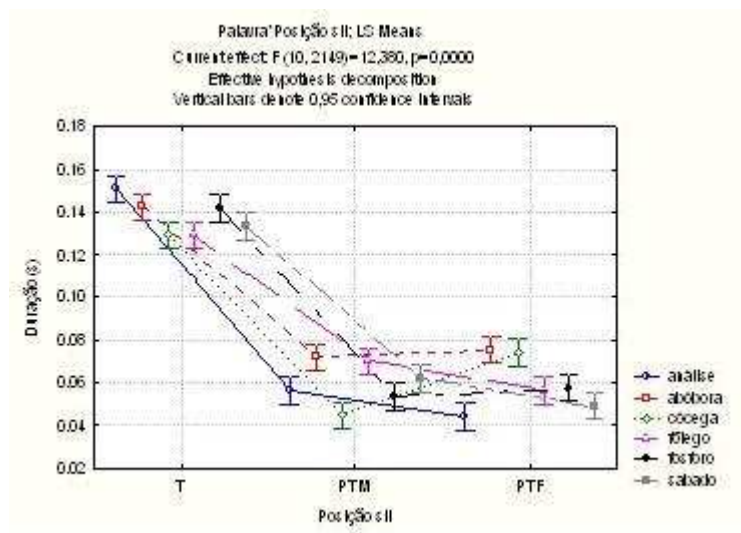


Figura 5.7: Duração de vogais em posição tônica (T), pós-tônica medial (PTM) e pós-tônica final (PTF). Todas as taxas consideradas.

Realizando uma *ANOVA One-way* separadamente para cada subconjunto, obtivemos os seguintes resultados: 1) primeiro padrão: $[F(2,1106)=516,92, p<10^{-4}]$; 2) segundo padrão: $[F(2,1055)=672,69, p<10^{-4}]$. Isso indica que a PTM é de menor duração do que a PTF no primeiro grupo e que, por outro lado, a PTF é de menor duração do que a PTM no segundo grupo.

Tais padrões explicariam porque as formas “abóbora, cócega, fósforo” variam com as formas sincopadas “abobra, cosca, fosfro”, em alguns dialetos, e também, porque a forma “análise” varia com a forma “anális/análiz”, em alguns dialetos. Além disso, obviamente, explicaria também a lexicalização das formas reduzidas (paroxítonas) dessas palavras em uma fase posterior

da evolução da língua. No entanto, sabemos que, em princípio, esses dados fonéticos não explicam o porquê de não tenderem a lexicalização formas como “foleg, sábad”, se adotarmos a hipótese de que a vogal final tende a cair em um determinado conjunto de palavras. Outro fato interessante nesse contexto é a presença de formas, em alguns dialetos, como “sabo”, em que houve a completa supressão da oclusiva em posição de coda (/d/) (cf. Amaral, 2002:100).

A realização em alguns dialetos de “sábado” como “sabo” nos questiona se, além de fatores duracionais, também haveria mudanças formânticas com o aumento da TE, as quais poderiam explicar variações dialetais nas proparoxítonas do PB. Para tanto, realizamos a análise descrita a seguir.

Estruturas formânticas intertaxas

Utilizamos as seguintes vogais das palavras para realizar as análises formânticas: a) abóbora = /ɔ, o, a/; b) análise = /a, i, e/; c) cócega = /ɔ, e, a/; d) fôlego = /o, e, o#/; e) fósforo = /ɔ, o, o#/; f) sábado = /a, (b)a, o/. Isso resulta em um total de 4320 segmentações (3 vogais x 2 formantes x 6 palavras x 8 informantes x 3 taxas x 5 repetições) e um total de 1540 análises estatísticas (4320 segmentações/ 3). Realizamos um teste ANOVA para cada uma das vogais, tendo como variável dependente F1 e F2 e independente a taxa de elocução. Só consideramos ser a estrutura formântica da vogal variá-

vel de acordo com a taxa de elocução se: a) p do teste ANOVA fosse menor do que 5% e se b) uma reta decrescente/crescente ocorresse da taxa lenta para a taxa rápida, como podemos observar na fig. 5.8. Com os dados dos formantes das vogais podemos especular, pela teoria acústica das vogais (cf, por exemplo Kent & Read, 1992), embora de maneira inferencial¹⁰, como os falantes realizaram sua articulação. Relembrando, F1 alto implica em uma abaixamento da vogal, F1 baixo em alçamento da vogal, F2 alto implica em anteriorização da vogal, e F2 baixo implica em posteriorização da vogal.

¹⁰Para sabermos, com certa precisão, as articulações vocálicas e consonantais, precisaríamos ter feito análises de ressonância magnética (MRI) e/ou medidas com um magnetômetro (ambas indisponíveis nos laboratórios de fonética do Brasil).

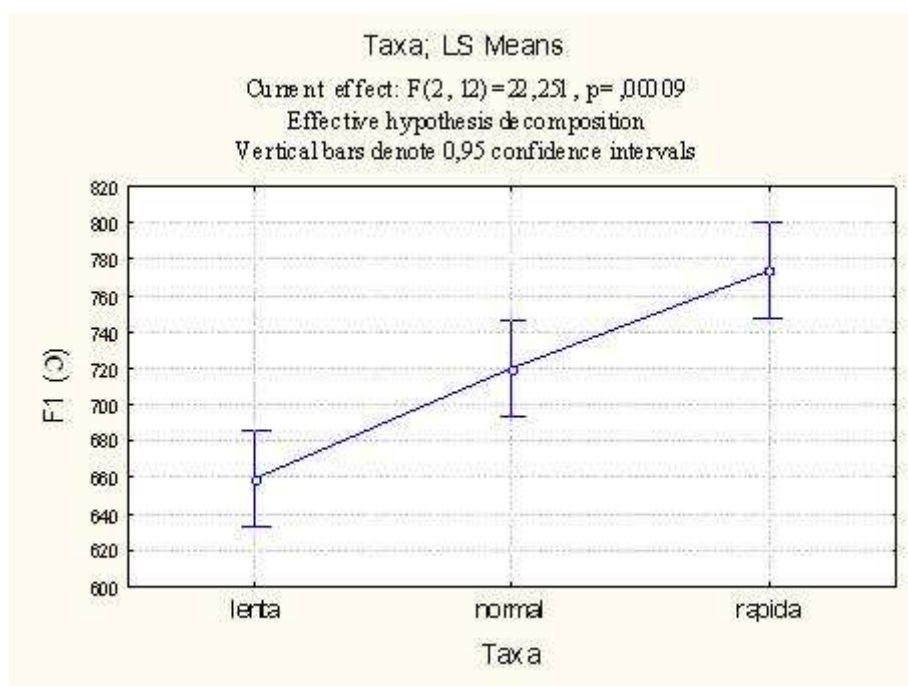


Figura 5.8: Variação de F1 de /ɔ/ de “cócega” em três taxas de elocução de JM.

O *script* do Praat “Formant Extractor” (apêndice A.6), por mim elaborado, foi utilizado na análise formântica das vogais. Tal *script* extrai informações formânticas de vogais pré-etiquetadas no Praat, a partir de seu ponto médio, e gera dois arquivos: (i) .txt com os quatro primeiro formantes e larguras de banda das vogais; (ii) .PraatPic com FFT e LPC sobrepostos, que podem ser abertos no Praat. No entanto, apesar desta automação da análise, observamos o espectrograma para cada uma das frases utilizadas, a fim de observar se o ponto médio das vogais correspondia, de fato, ao melhor ponto para a extração dos formantes. Se observássemos ser o ponto médio

inefcaz para a análise, modificávamos, manualmente, o ponto, e marcávamos à mão seus formantes, larguras de banda, LPC e FFT.

Os resultados obtidos revelaram que a taxa de elocução também tem um papel primordial em alterar a estrutura formântica das vogais, e, assim, mudar sua qualidade vocálica. Lembremos ainda que o contexto fonético também mostrou ser essencial na modificação das qualidades vocálicas devido ao aumento da taxa de elocução.

Tomemos como exemplo, JP (MG), que obteve diferenças significativas entre as taxas, através de uma ANOVA de variáveis independentes, dentre outras palavras, em F2 de: /o/ de “abóbora” ($F(2,22) = 8,5616$, $p < 0,00489$), /a/ de “análise” ($F(2,13) = 13,496$, $p < 0,00067$), /e/ de “cócega” ($F(2,12) = 9,5256$, $p < 0,00333$), /o/ de “fôlego” ($F(2,12) = 76,502$, $p < 0,00000$), conforme observamos na tabela 5.1¹¹. Observamos nesses casos uma adaptação gradual da articulação da vogal com a consoante com o aumento da taxa de elocução, ou seja, se a consoante fosse “anterior”, como /r, l/, a vogal precedente se anteriorizava. Se a consoante fosse “posterior” como /g/, a vogal se posteriorizava.

Este aumento da coarticulação com o aumento da taxa de elocução foi observado em todos os outros informantes, como GB (SP), que obteve di-

¹¹Notemos, entretanto, que muitas vezes, confirmadas por um teste *post-hoc* Scheffé, não houve distinção na média das taxas lenta e normal. Porém sempre houve da taxa lenta para a rápida.

Tabela 5.1: Média do F2 das vogais [o] de “abóbora”, [a] de “análise”, [e] de “cócega” e [o] de “fôlego” de JP.

F2	lenta	normal	rápida
[o]	1178	1207	1260
[a]	1519	1509	1570
[e]	1895	1932	1758
[o]	890	896	1035

ferenças significativas no F2 de: /o/ de “abóbora”, /a/ de “análise”, /ɔ/ de “cócega”, /e/ de “fôlego”, /o/ de “fósforo”.

Até o momento não tínhamos nenhuma evidência dinâmica de como as vogais nas palavras proparoxítonas se reduziam proporcionalmente umas em relação às outras. No intuito de confirmar nossas hipóteses acima mencionadas, e também de reforçar os aspectos dinâmicos na mudança de proparoxítonas a paroxítonas, efetuamos uma análise estatística não-paramétrica que mostra a proporção de diminuição relativa da vogal em relação às taxas.

Evidências dinâmicas do processo de paroxitonização

Através do teste *Anova* não-paramétrico de *Friedman*, pudemos observar aspectos dinâmicos na variação de proparoxítonas a paroxítonas. Neste teste verificamos qual a taxa de decrescimento da duração de cada uma das vogais (T, PTM, PTF), em uma determinada taxa (lenta e rápida), em relação à taxa normal. Em outras palavras, quanto, em termos percentuais, se reduzia

cada vogal com o aumento da taxa de elocução, conforme podemos observar na tabela 5.2.

Tabela 5.2: Valores percentuais negativos médios da taxa de decrescimento da duração das vogais de DF (MG), obtidos através da seguinte fórmula: [Vogal: $\frac{(R-N)}{N}.100\%$, $\frac{(N-L)}{N}.100\%$], em que R = duração da vogal na taxa rápida, N = duração da vogal na taxa normal e L = duração da vogal na taxa lenta.

Palavra	T	PTM	PTF
abóbora	[ɔ: 28, 53]	[o: 54, 34]	[ɐ: 22, 26]
análise	[a: 18, 43]	[i: 48, 20]	[ɪ: 38, 146]
cócega	[ɔ: 25, 39]	[e: 39, 9]	[ɐ: 13, 29]
fôlego	[o: 19, 42]	[e: 2, 15]	[ʊ: 7, 83]
fósforo	[ɔ: 13, 35]	[o: 5, 8]	[ʊ: 2, 30]
sábado	[a: 31, 32]	[ɐ: 10, 81]	[ʊ: 9, 71]

Foram feitas duas análises: uma com a unidade experimental o informante e a outra com a unidade experimental a palavra.

Na primeira análise foram calculadas, para cada informante, as médias entre as proporções de duração relativa da taxa normal para a rápida, para cada posição (T, PTM e PTF), considerando todas as palavras juntas. Obtivemos significância para o conjunto de sílabas ($p < 0,002$; $\alpha = 0,05$) e para pós-tônica medial vs. final ($p < 0,004$; α efetivo = 0,017). Obtivemos também significância para pós-tônica final e tônica, mas não para pós-tônica medial e tônica. Esses resultados implicam que a pós-tônica medial tende a se reduzir, dinamicamente, mais do que a pós-tônica final ao passar da taxa normal para a rápida. Isso complementa nossa análise anterior, a qual mostra

uma redução bruta maior na pós-tônica medial do que nas outras posições, e mostra também como aspectos dinâmicos relacionados ao aumento da taxa de elocução atuam na paroxitonização de palavras como “abóbora, cócega, fósforo”.

Na segunda análise foram calculadas, para cada posição (T, PTM e PTF), as médias entre as proporções de redução relativa da taxa normal para a rápida para todos os informantes juntos. Os resultados foram semelhantes aos encontrados na primeira análise, ou seja, houve significância para as três posições juntas ($p < 0,006$; $\alpha = 0,05$) e também entre a pós-tônica medial e a final ($p < 0,014$; α efetivo = 0,017). Aqui, no entanto, percebemos, novamente, uma diferenciação nos graus de redução da vogal pós-tônica medial, pois “fôlego, sábado” foram as palavras que menos reduziram suas vogais pós-tônicas mediais. Fato este condizente com a análise feita anteriormente, em que separamos dois grupos distintos de paroxitonização: 1) redução da vogal pós-tônica medial (abóbora, cócega, fósforo); 2) redução da vogal pós-tônica final (análise, fôlego, sábado). Também aqui percebemos, uma vez mais, aspectos dinâmicos, influenciados pelo aumento da taxa de elocução, relacionados ao processo lexical de paroxitonização.

Com o intuito de representar esquematicamente como a taxa de elocução pode estar atuando articulatoriamente na variação entre proparoxítonas e paroxítonas, a qual pode resultar em uma mudança lexical, observemos a

figura 5.9 abaixo. É importante ressaltar, entretanto, que a representação abaixo serve apenas como ilustração do processo de paroxitonização, visto que na FAR, como vimos anteriormente, os gestos são representados abstratamente através das pautas gestuais, ou seja, a implementação fonética dos gestos, apesar de ser facilmente derivada das pautas gestuais, não é representada. Em outras palavras, na FAR também não se representa explicitamente os processos de coarticulação resultantes da implementação fonética dos gestos abstratos. Esta característica dos gestos, como explicitado pela FAR, é importante para que se tenha uma representação abstrata de qualquer palavra de uma determinada língua. Mas, como podemos observar pela análise da figura 5.9, abre margem para explicação dos aspectos variacionistas das línguas naturais.

Na figura 5.9, indicamos, na pauta gestual (b), a redução maior da vogal pós-tônica medial na taxa de elocução rápida. Aqui percebemos uma aproximação do gesto de corpo de língua de [o]¹² (4) com o gesto labial [b] (8) e o gesto de ponta de língua [r] (1). A forma da pauta (5.9b) é a que poderia ser percebida pelos falantes como uma forma paroxítona, mesmo contendo uma vogal curta¹³. O que, em termos dinâmicos, é importante ressaltar nessa

¹²É mister lembrar que as transcrições fonéticas apresentadas na pauta gestual da figura 5.9 são apenas uma maneira de facilitar a leitura da mesma, visto que na FAR não faz sentido tal representação segmental.

¹³Esta vogal curta acontece foneticamente mesmo em palavras com encontros consonantais com [r] e, portanto, passaria a ser “interpretada” como parte inerente da produção de uma sequência /Cr/, em que C é a primeira consoante de um ataque complexo.

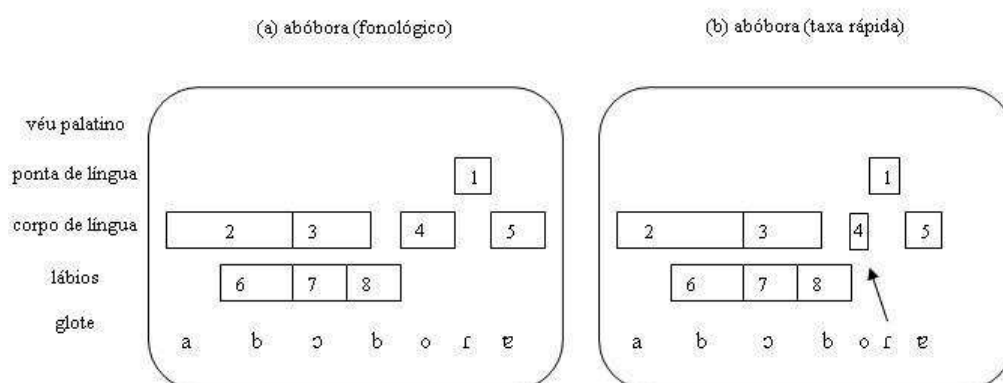


Figura 5.9: a) Representação fonológica (FAR) de abóbora; b) Implementação gestual da redução da vogal pós-tônica medial de abóbora na taxa rápida. Os números na pauta representam: 1) fechado alveolar; 2) amplo faringal; 3) meio-ampla velar; 4) meio-estreito velar; 5) meio-ampla faringal; 6) fechado labial; 7) amplo protuso; 8) fechado labial.

pauta é, como dissemos anteriormente, que os gestos não surgem ou desaparecem, eles estão sempre presentes. O que de fato ocorre é a redução do grau de magnitude de um determinado gesto ou a sobreposição de gestos que resulta nas variantes lingüísticas, que, em um dado momento da língua, pode se tornar uma mudança lingüística.

A partir desses dados aventamos duas hipóteses dos processos de mudança lingüística de proparoxítonas a paroxítonas.

1. Hipótese PB

- Padrão 1: A vogal pós-tônica medial gradualmente se reduz até ser percebida como inexistente pelos falantes e a sua respectiva consoante formar o ataque da sílaba seguinte. Restrição: não

formar moldes silábicos fonotaticamente inexistentes. Esquemáticamente: $CV_tCV_{ptm}CV_{ptf} > CV_tCCV_{ptf}$;

- Padrão 2: A vogal pós-tônica final gradualmente se reduz até ser percebida como inexistente pelos falantes e a sua respectiva consoante formar a coda da sílaba anterior. Restrição: não formar moldes silábicos fonotaticamente inexistentes. Esquematicamente: $CV_tCV_{ptm}CV_{pf} > CV_tCV_{ptm}C$.

2. HIPÓTESE LATIM:

- No latim, como nos é impossível fazer medidas acústicas, somente podemos hipotetizar o que haveria ocorrido de fato, ou seja, como já havia sido dito por diversos autores, a vogal penúltima pós-tônica breve (PTB) ($\check{a}$, $\check{e}$, $\check{i}$, $\check{o}$, ou $\check{u}$ reduziu-se gradualmente até ser percebida como inexistente entre certos pares de consoantes em latim vulgar: $alt\check{e}rum > outro$; $dom\check{i}num > dono$; $lep\check{o}rem > lebre$; $oc\check{u}lum > olho$; $pos\check{i}tum > posto$; $uir\check{i}dem > verde$ (cf. Williams, 1961 (1938)). Esquematicamente: $CV_tCV_{ptb}CV_{ptf} > CV_tCCV_{ptf}$.

Obviamente também aqui pode ter havido uma restrição fonotática que impediu que algumas vogais breves pós-tônicas mediais fossem reduzidas gradualmente até serem percebidas como inexistentes, como em $ang\check{e}lum > an-$

geo > anjo; debĩtam > dívida. No entanto, não temos mais recursos fonéticos para observar empiricamente se esta hipótese é verdadeira.

5.2.3 Considerações sobre a influência da TE nas reestruturações lexicais

Apesar de haver uma enorme variação de realizações fonéticas em relação ao dialeto, gênero, aspecto individual, podemos especular como poderiam estar sendo efetuadas reestruturações, nas palavras que analisamos, de acordo com a taxa de elocução.

Para as palavras “abóbora” e “fósforo” houve para alguns falantes, como DF (MG), a redução gradual, da taxa lenta para a rápida, da vogal imediatamente antes da vibrante simples, culminando em uma forma bastante reduzida, muitas vezes, percebida como inexistente para os falantes do PB, pois é associada à vocalização inerente do [ɾ], mesmo em encontro consonantal. Tais formas reduzidas foram transcritas como [a'bɔb^vrɐ] e [fɔsf^vɾɔ], conforme podemos observar pela vogal /o/ reduzida (21 ms) de “abóbora” na fig. 5.10.

Para “análise” houve em alguns informantes, a sobreposição do gesto articulatório de [ɪ] gradual, da taxa lenta para a rápida, com a consoante precedente, fazendo com que as consoantes [l, s] soassem como silábicas. Isso

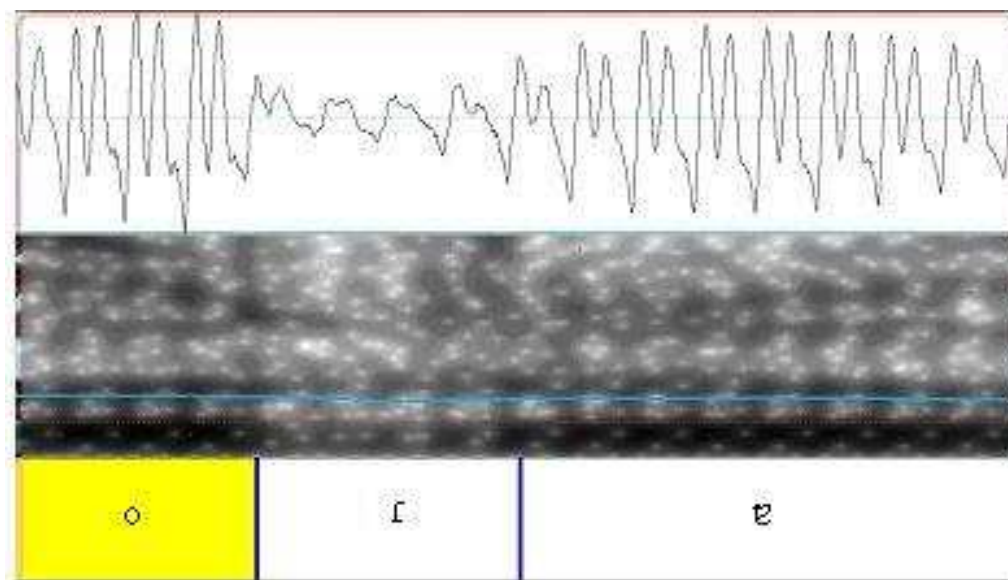


Figura 5.10: Representação da forma de onda e do espectrograma de [orɐ] de “abóbora”, DF, taxa rápida.

ocorreu, por exemplo, em “análise” de JM (MG), pronunciada como [a'nal'zɪ], conforme podemos observar na fig. 5.11.

Para “cócega”, apesar de ter ocorrido apenas uma forma reduzida como “cosga”, encontramos indícios de como a mudança de “cócega” para “cosca” poderia estar sendo operada no PB. Isso ocorreu com o informante JP (MG), que mudou de [kɔseɣɐ], na taxa lenta, para [kɔsɪɣɐ], taxa normal, alternando com [kɔseɣɐ], e desta para, na taxa rápida, [kɔs'ɣɐ, 'kɔsgɐ], também em alternância com as outras duas formas lexicais presentes nas outras duas taxas. Com o desvozeamento do [g], ocorreria, pois, [kɔskɐ].

Para “fôlego” encontramos apenas um exemplo de redução lexical apa-

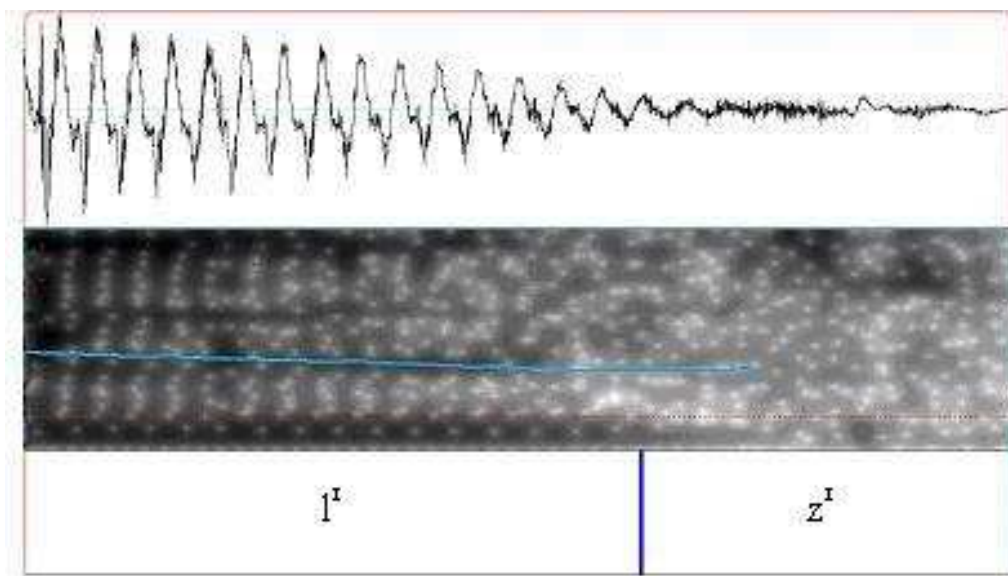


Figura 5.11: Representação da forma de onda e do espectrograma de [l'z'] de “análise”, JM, taxa rápida.

rente. Isso ocorreu com o informante JM (MG), que produziu essa palavra com o [v] final sobreposto ao [g], como podemos observar na fig. 5.5 (seção 5.2.2).

Para “sábado” não encontramos para nenhum dos oito informantes reduções lexicais aparentes. O único indício de mudança lexical encontrado foi o fato de termos percebido o /a/ de /ba/ dos mineiros como mais posterior do que o dos paulistas.

Concluindo, tivemos evidências em termos de produção, para alguns falantes, na maioria dos casos, de como a mudança de proparoxítonas para paroxítonas em “abóbora” e “fósforo” poderia, engatilhada pela TE, ser efe-

tuada. A consideração de “abóbora” como paroxítona está exibida nas figuras 5.10 e 5.12. Na figura 5.12 mostramos uma ANOVA representando a não diferença estatística entre o /o/ de “abóbora” e o [ɐ] que antecede a vibrante simples em “abobra”. Para “fósforo” não foi possível fazer uma análise estatística semelhante à da fig. 5.12, pois os falantes que produziram “fósforo” como paroxítona desvozearam o [ʊ] antes da vibrante simples de “fosfro”, o que impediu a separação desta consoante no encontro consonantal “fr”, como podemos ver na fig. 5.13. Apesar deste fato, consideramos “fósforo” como paroxítona nos casos em que a vogal antes da vibrante simples era bem menor do que a vogal final, como podemos observar na fig. 5.14. Para “cócega” tivemos apenas um exemplar de paroxítona, como pudemos observar na fig. 5.5 (seção 5.2.2). Para as outras palavras não tivemos evidências em termos de produção para dizer que fossem consideradas como paroxítonas. Aventamos, então, a hipótese, a ser investigada em trabalhos futuros, de que a consideração dessas mesmas palavras como paroxítonas se pauta em dois processos: a) ou essas formas só ocorrem em falas espontâneas, informais; b) ou são percebidas como tais devido à sobreposição de gestos vocálicos com consonantais.

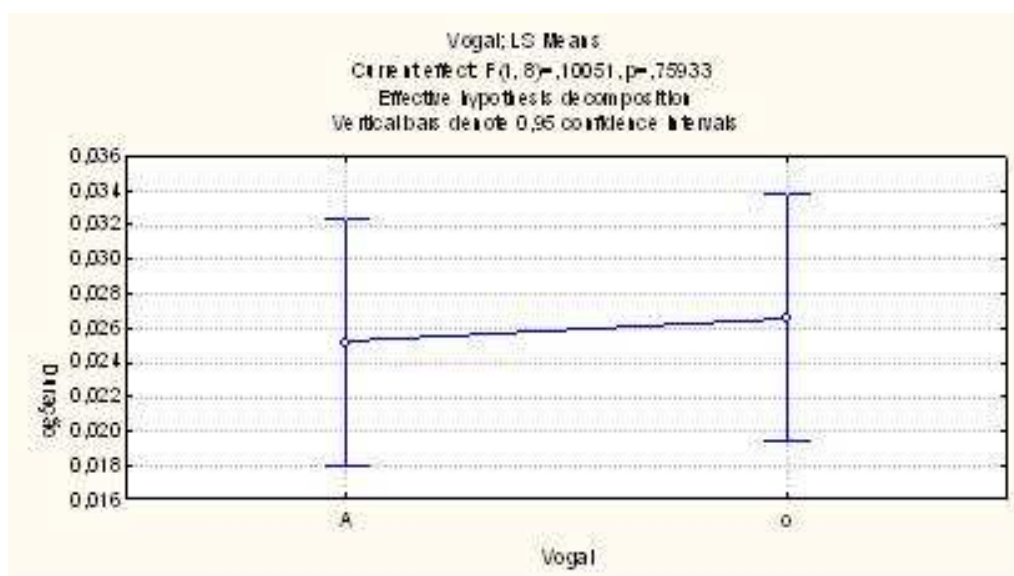


Figura 5.12: Duração do [v] antes da vibrante simples de “abobra” vs. o /o/ medial de “abóbora”, taxa rápida, DF.

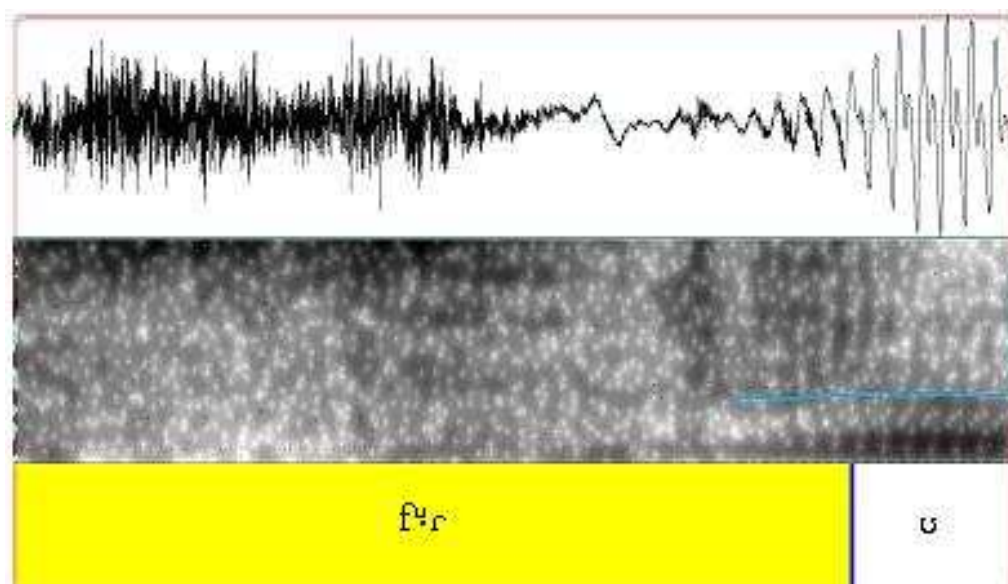


Figura 5.13: Representação da forma de onda e do espectrograma de [fʊr] de “fosfro”, DF.

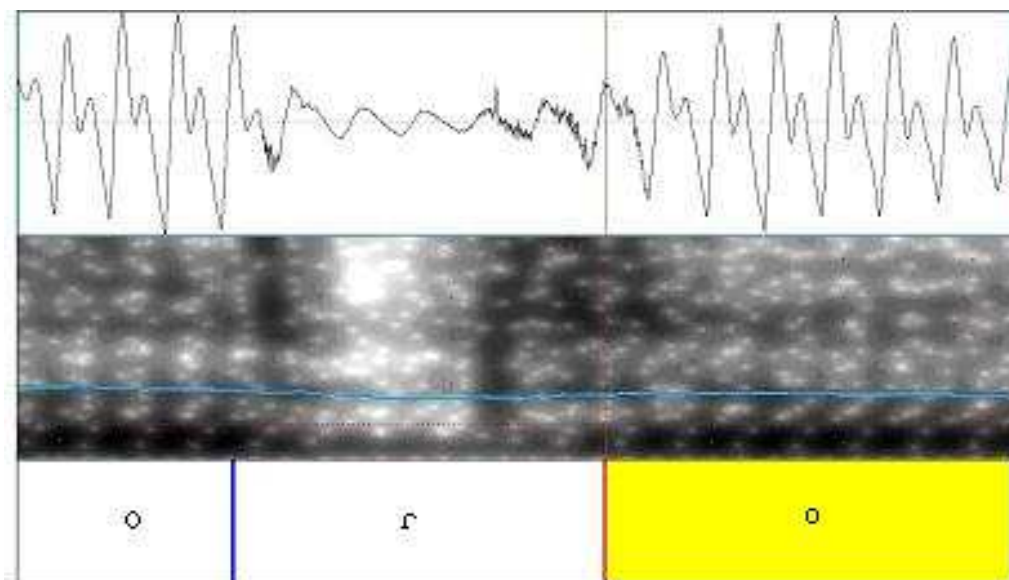


Figura 5.14: Representação da forma de onda e do espectrograma de [oro] de “fósforo”, DF.

5.2.4 Discussão geral

Em nosso trabalho, observamos que não dá para dizer categoricamente, como em Hooper (1976) e Abaurre-Gnerre (1979), que uma taxa “elege” uma determinada forma lexical para sua produção. Em FN, por exemplo, houve variação na palavra “fósforo” na taxa lenta, com inclusive a presença da forma [ˈfɔsfʊrɔ], que ocorria tanto na taxa normal quanto na rápida. Esses resultados confirmam a afirmação de Farnetani (1999:379) de que

the system response to high rate commands does not automatically result in reduced movements (Kuehn and Moll, 1976; Gay, 1978), and that reduction can occur also at slow rates (cf. Nord,

1986). This indicated that speaker can adapt to different speaking situations and choose different production strategies to avoid reduction/coarticulation or to allow it to occur.

Percebemos essa estratégia de evitar redução no informante GB (SP) que, apesar de ser um dos informantes mais rápidos de todo o grupo analisado, não realizou tanta redução quanto outros informantes na mesma taxa. Este mesmo fenômeno encontramos em locutores futebolísticos que produzem falas extremamente rápidas, porém, sem os processos de redução extremos encontrados aqui para os falantes mineiros.

Devido a essa enorme variabilidade lexical encontrada em nosso corpus, é difícil lidar com fenômenos lexicais variacionistas com base em segmentos discretos e atemporais. Daí a necessidade de um novo paradigma de pesquisa, que dê conta dessa variabilidade, como mostramos nas pautas gestuais de “para” (fig. 5.1a, seção 5.1.3), “prata” (fig. 5.1b, seção 5.1.3) e “abóbora” (fig. 5.9, seção 5.2.2).

Também notamos em nosso trabalho fortes evidências de influência de fatores fonéticos de base neogramática na variação lingüística, pois encontramos formas reduzidas, como “abobra”, “análiz”, “cosca” e “fosfro”, que surgiram a partir do aumento gradual da taxa de elocução. Sendo assim, acreditamos ter mostrado que determinadas mudanças lingüísticas tendem

a ter como mola-mestra (*actuation*) de iniciação fenômenos coarticulatórios impulsionados pela taxa de elocução, que podem ocorrer sem exceção, e também ser graduais e não abruptos.

Tal influência de fatores fonéticos para iniciar uma mudança lingüística já foi citado por Chen & Wang (1975:275), que disseram serem as restrições inerentes do aparato fisiológico e perceptual do falante os principais determinantes de uma mudança lingüística. Em outra parte desse artigo, os pesquisadores ainda dizem que “as for the actuation aspects of sound change, we have reason to believe that it is mainly the concrete, phonetic properties of speech sounds that trigger or allow changes to take place in the sound system, and determine their subsequent development (p.278)”.

Também Philips (1984:320), apesar de mostrar também influência de outros fatores na mudança lingüística, cita fatores fonéticos na mudança lingüística. Segundo a autora, “changes affecting the most frequent words first are motivated by physiological factors, acting on surface phonetic forms; changes affecting the least frequent words first are motivated by other, non-physiological factors, acting on underlying forms”.

Nosso trabalho também mostra que as mudanças sonoras podem ser foneticamente graduais, como mostram os autores da FAR em diversos artigos como Browman & Goldstein (1989, 1990b, 1992). Há também evidências de fenômenos graduais de inserções e apagamentos para o português brasileiro

em Albano (2001), evidência contrária à observação da impossibilidade de todas as mudanças sonoras serem foneticamente graduais e de que “cancelamentos, inserções e metáteses não podem ser graduais e devem ser concebidos como fenômenos discretos” (cf. Oliveira, 1991:1).

5.2.5 Conclusão

Em nosso trabalho encontramos evidências de que o fenômeno de variação de proparoxítona para formas paroxitonizadas é influenciado por fatores fonéticos como a taxa de elocução, sendo que a probabilidade de encontrar formas consideradas como paroxítonas é maior em taxas rápidas. No entanto há fatores individuais envolvidos, pois há falantes que nunca as produziram na taxa rápida.

Os resultados principais de nosso trabalho foram:

- a) Cada falante possui um tempo intrínseco próprio que faz com que não haja uma taxa de fala única para cada dialeto. Esta taxa depende de cada um dos falantes. Há falantes, que, por natureza, falam mais rápido, e falantes que falam mais lento;
- b) Falantes do dialeto mineiro (7,5 sílabas/s), no experimento realizado, falaram, em média, 17% mais rápido que os do dialeto paulista (6,4 sílabas/s);

- c) Falantes do gênero masculino (5,8 sílabas/s) falaram, em média, 11,5% mais rápido do que os do gênero feminino (5,2 sílabas/s) neste estudo;
- d) Há pelo menos dois processos lingüísticos distintos na variação/mudança de proparoxítonas a paroxítonas:
- diminuição da duração acústica/articulatória da vogal pós-tônica medial. Ex.: abóbora, cócega e fósforo;
 - diminuição da duração acústica/articulatória da vogal pós-tônica final. Ex.: análise, fôlego e sábado.

No próximo capítulo, como complemento ao presente trabalho, elaboraremos um teste articulatório, que pode servir de respaldo à idéia de algumas proparoxítonas poderem ser percebidas como paroxítonas, devido à sobreposição de gestos consonantais e vocálicos no aumento da TE.

Capítulo 6

Variações articulatórias em proparoxítonas

No capítulo 5 investigamos dados acústicos na variação/mudança de proparoxítonas a paroxítonas. Neste capítulo apresentaremos dados articulatórios, que mostram sobreposição dos gestos consonantais e vocálicos nesse processo lingüístico. Para tanto realizamos o experimento descrito abaixo.

6.1 Introdução

Nosso intuito inicial era poder utilizar o mesmo corpus de nossos estudos acústicos precedentes. No entanto, devido à grande variabilidade articulatória presente, optamos por utilizar um corpus mais bem controlado, a fim de

podermos utilizar um mesmo conjunto de articuladores-alvo para todas as palavras. Optamos, pois, por utilizar um corpus com os articuladores lábios inferiores (LI) e ponta de língua (PL) nas consoantes das sílabas pós-tônicas. Como exemplo, temos “abób**o**ra”, em que as consoantes em negrito possuem articulação labial [b] e alveolar [r]. Após definirmos nosso conjunto de articuladores, elaboramos um *script* (*Dictionary* no apêndice A.6) para busca automática no Dicionário Eletrônico Aurélio.

Em primeiro lugar, procuramos no dicionário uma palavra para usar na frase-veículo. A frase-veículo normalmente utilizada para dados acústicos, “Diga _____ baixinho”, não era adequada, devido ao fato de terminar em uma vogal, a qual pode ser produzida concomitante com a primeira vogal de “abóbora”, por exemplo. Como, para o dialeto de nossa informante, só podemos ter, foneticamente, /R/ ou /S/ em posição de coda¹, e como para usar /R/ teria de usar um verbo no infinitivo (ou imperativo: bastante incomum no português brasileiro), optamos por usar um verbo que terminasse em /S/. Para isso procurei no dicionário infinitivos que terminassem em “ser” ou “zer”, os quais podiam nos dar uma forma com /S/ na terceira pessoa do singular. No entanto, após testar todos os verbos possíveis, por problema de adaptação semântica dos verbos com as palavras a utilizar, optamos por usar

¹Excluimos aqui o arquifonema nasal /N/ devido a dificuldades de aquisição de dados em EMMA.

a frase-veículo: “Digas _____ baixinho”. Apesar de ser uma forma verbal muito pouco usada no Brasil, ao menos nos facilitaria a análise articulatória posterior.

Em segundo lugar, procurei no dicionário palavras proparoxítonas semelhantes a “abóbora, bêbado, fósforo, sábado”. Encontrei uma lista enorme de palavras, mas, para limitarmos o nosso corpus, optamos por trabalhar com um corpus bem controlado do ponto de vista articulatório e também do ponto de vista de número de gestos articulatórios. Com isso chegamos ao desenho experimental abaixo.

6.2 Metodologia

O corpus utilizado no experimento foi: “a víbora, abóbora, a bêbada, abóbada”. Estas palavras foram controladas do ponto de vista articulatório. Além destas também foram gravadas as palavras precedidas de artigo: “o fósforo” e “o sábado”, menos controladas fonético-articulatoriamente, mas escolhidas por terem sido usadas em nossos experimentos acústicos e perceptivos anteriores.

O objetivo deste trabalho era observar a coarticulação entre as vogais pós-tônicas e as consoantes circundantes. Para isso utilizamos quatro sensores nos respectivos articuladores: corpo de língua (CL), ponta de língua (PL), lábios

inferiores (LI) e lábios superiores (LS). Além destes, como no experimento do capítulo 4, foram utilizados sensores de referência no nariz e no maxilar.

A informante foi uma falante nativa do português brasileiro, dialeto mineiro, idade 30 anos, gênero feminino, a qual leu os enunciados inseridos na frase-veículo “Digas _____ baixinho” em três taxas de elocução, com dez repetições cada, aleatorizadas no *Microsoft Excel*. A ordem das taxas, como nos outros experimentos, foi: 1) normal; 2) lenta; 3) rápida.

6.3 Análise dos dados

Para medir o grau de coarticulação nas palavras estudadas, medi a distância articulatória entre as consoantes pós-tônicas (marcadas em **negrito**), a saber: abób**o**ra, abób**o**da, ví**b**ora, bê**b**ada, fós**f**oro e sáb**a**do. Como todas as consoantes do corpus consistiam de consoantes labiais e alveolares, tal medida foi feita a partir da velocidade zero no eixo y (LI) até a velocidade zero no eixo y (PL), como demonstra a figura 6.1 abaixo.

Além da medida do intervalo consonantal, também analisamos o intervalo articulatório entre o início e o final da palavra-alvo, como demonstra a figura 6.2 abaixo. Esta medida nos permitiu avaliar se estávamos trabalhando com TE distintas.

Uma *ANOVA One-way* com duração articulatória da palavra em função

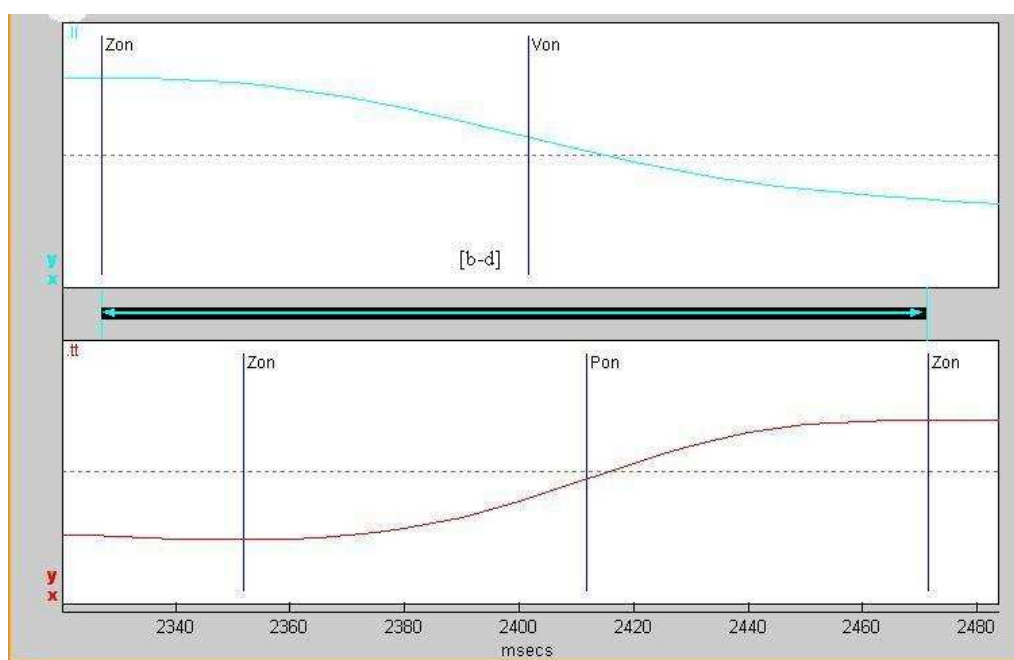


Figura 6.1: Exemplo da medida da duração do intervalo entre consoantes pós-tônicas. Aqui temos representado pela seta em azul claro a distância articulatória entre [b] (velocidade zero, eixo y, gesto lábios inferiores, linha azul claro) e [d] (velocidade zero, eixo y, gesto ponta de língua, linha vermelha).

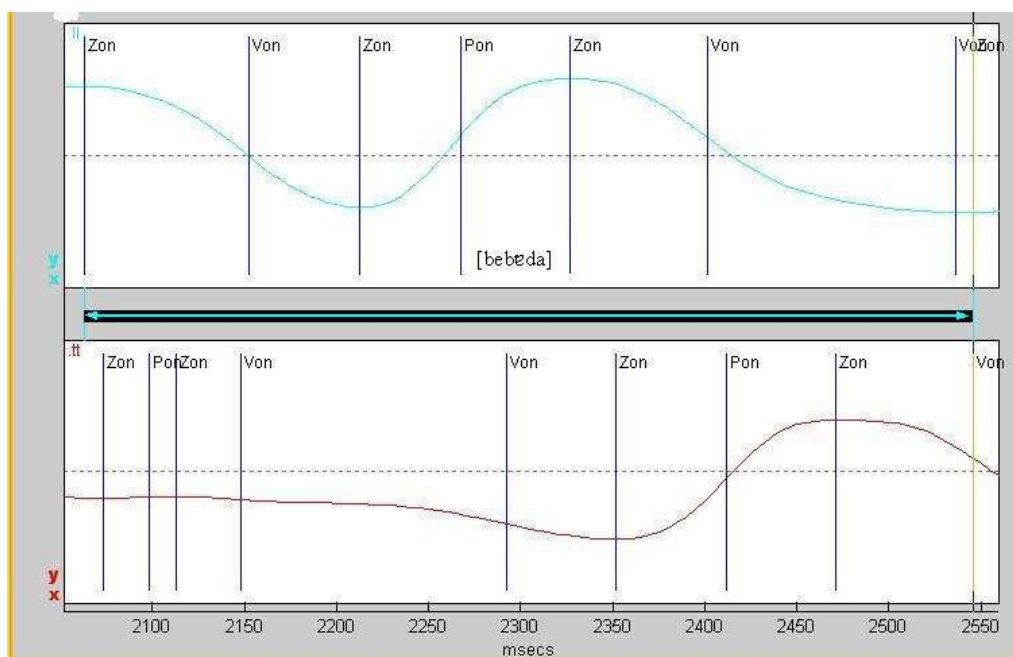


Figura 6.2: Exemplo da medida da duração do intervalo entre o início e o final de uma palavra. Aqui temos representado pela seta em azul claro a distância articulatória entre [b] (velocidade zero, eixo y, gesto lábios inferiores, linha azul claro) e [a] (velocidade zero, eixo y, gesto lábios inferiores, linha vermelha).

da TE nos indicou que as taxas foram estatisticamente diferentes entre si, como mostra a tabela 6.1. Além disso, houve uma diminuição gradativa da duração com o aumento da TE, comprovada estatisticamente por um teste *post-hoc Scheffé*.

Tabela 6.1: Dados estatísticos da duração da palavra com o aumento da taxa de elocução.

Palavra	Anova	p <
abóbora	$F(2,27) = 323.39$	10^{-4}
abóboda	$F(2,27) = 272.64$	10^{-4}
víbora	$F(2,26) = 177.74$	10^{-5}
bêbada	$F(2,25) = 147.90$	10^{-5}
fósforo	$F(2,27) = 189.37$	10^{-5}
sábado	$F(2,26) = 85.135$	10^{-5}

Comprovada estatisticamente a diferenciação entre as taxas no corpus, passamos a investigar o grau de coarticulação entre as consoantes pós-tônicas. O raciocínio empregado foi de que uma distância articulatória menor entre tais consoantes representariam um maior grau de coarticulação entre essas consoantes e a vogal envolvida pelas mesmas. Esperava-se que tal grau de coarticulação aumentasse gradativamente com o aumento da TE. Para observar tal hipótese, efetuamos uma *ANOVA One way* com a duração articulatória do intervalo consonantal em função da TE. Os resultados estatísticos estão representados na tabela 6.2 abaixo.

A análise da tabela 6.2 mostra que há de fato um aumento de coarticula-

Tabela 6.2: Dados estatísticos da duração do intervalo consonantal com o aumento da taxa de elocução.

Palavra	Anova	p <	Scheffé
abóbora	$F(2,27) = 61.803$	10^{-5}	$L \neq N \neq R$
abóboda	$F(2,27) = 78.661$	10^{-5}	$L \neq N \neq R$
víbora	$F(2,26) = 75.084$	10^{-5}	$L \neq N \neq R$
bêbada	$F(2,25) = 86.468$	10^{-5}	$L \neq N \neq R$
fósforo	$F(2,27) = 26.632$	10^{-5}	$L \neq (N = R)$
sábado	$F(2,26) = 52.209$	10^{-5}	$L \neq N \neq R$

ção nos segmentos pós-tônicos com o aumento da TE, ou seja, as consoantes pós-tônicas estão mais próximas entre si. Além disso, uma análise *post-hoc* *Scheffé*, revela que, com exceção de “fósforo”, todas as palavras se diferenciaram entre as taxas. Esse aumento da coarticulação é reforçado tomando-se a duração de todas as palavras juntas em função da TE ($F(2,173) = 168.52$, $p < 10^{-4}$). Pode-se questionar, entretanto, ser tal resultado óbvio, pois, ao se aumentar a TE, a duração dos segmentos tenderão a diminuir gradativamente. Contudo, pela análise das tabelas 6.1 e 6.2, observamos não ser tal diminuição linear, visto que, tomando-se os valores de F nas tabelas, as palavras se reduziram muito mais (valores mais altos) do que o intervalo consonantal estudado. Esse fato nos fez indagar se haveria uma possível diferenciação no intervalo consonantal com o aumento da TE entre as palavras. Para verificar tal questão, elaboramos uma *ANOVA One-way* com a duração do intervalo consonantal em função da palavra. O resultado estatístico mos-

trou, porém, não haver influência da palavra na duração estudada, ou seja, todas as palavras se comportaram da mesma forma para a análise efetuada.

Ainda procurando investigar se haviam diferenças articulatórias relativas ao intervalo consonantal estudado, efetuamos uma nova medida, a qual chamamos de intervalo consonantal proporcional. Tal intervalo refere-se ao quanto as consoantes se reduziram percentualmente dentro da palavra em função do aumento da TE e foi medido através da seguinte fórmula:

$$ICP = \frac{IC}{DP}, \quad (6.1)$$

em que ICP refere-se ao intervalo consonantal proporcional, IC ao intervalo consonantal e DP à duração da palavra. Por exemplo, se a duração de um IC é de 195.5 ms e a de um DP é 431 ms, temos que o ICP é de 45%, ou seja, o intervalo ocupa 45% da duração total da palavra.

A hipótese nula era de que o ICP se reduzisse (ocupasse uma parte menor da palavra) com o aumento da TE. Para investigar nossa hipótese, efetuamos uma *ANOVA One-way* com o ICP em função da TE. Os resultados estão apresentados na tabela 6.3 abaixo.

A análise da tabela 6.3 revela, contrário às nossas expectativas, que, proporcionalmente à redução na palavra, não houve para a maioria das palavras (4 dentre 6) efeitos da TE no ICP . Até mesmo nas palavras que apresentaram

Tabela 6.3: *Anova One-way* dos *ICPs* em função da taxa de elocução.

Palavra	Anova	p <	Scheffé
abóbora	n.s.	n.s.	n.s.
abóboda	n.s.	n.s.	n.s.
víbora	n.s.	n.s.	n.s.
bêbada	n.s.	n.s.	n.s.
fósforo	F(2,27) = 5,8319	0,008	(L = N) $\neq$ R
sábado	F(2,26) = 4,2452	0,03	L $\neq$ (N = R)

tal efeito (“fósforo” e “sábado”), não houve uma distinção para todas as taxas. Além disso, fazendo-se uma *Anova One-way* do *ICP* de todas as palavras juntas em função da TE, não houve diferença estatística do *ICP* com o aumento da TE. Tais dados nos remontam ao estudo de Tuller & Kelso (1984), onde, na busca por invariantes lingüísticos, verificou-se que a organização temporal intersegmental relativa é mantida ao se variar a estrutura suprasegmental. O mesmo parece ter ocorrido em nossos resultados, pois, apesar de haver-mos comprovado que a duração do intervalo consonantal estudado diminuir com o aumento da TE, tal intervalo se apresentou constante em relação ao tamanho total da palavra (*ICP*). Em outras palavras, o mesmo se reduziu proporcionalmente em função da diminuição da duração da palavra-alvo.

Como não encontramos, em geral, influência da TE no *ICP*, procuramos investigar se havia influência do tipo de palavra no *ICP*. Para isso efetuamos uma *Anova One-way* com o *ICP* em função do tipo de palavra.

A *Anova One-way* mostrou que o fator palavra é relevante para a atri-

Tabela 6.4: Média (μ) e desvio-padrão (σ) dos *ICPs* em função do tipo de palavra com seu respectivo grupo.

Palavra	μ (σ)	Grupo
víbora	49 (4)	1
bêbada	36 (4)	2
abóbora	31 (4)	2
fósforo	32 (5)	2
sábado	33 (6)	2
abóboda	27 (3)	3

buição do *ICP* ($F(5,170) = 83,733$, $p < 10^{-4}$). Além disso um teste *post-hoc* *Scheffé* dividiu as palavras em três grupos. Na tabela 6.4 podemos ver esses grupos resultantes com a média e desvio-padrão dos *ICPs* respectivos, compostos por: Grupo 1: víbora; Grupo 2: bêbada, abóbora, fósforo e sábado; Grupo 3: abóboda. Ressaltamos que um valor de *ICP* mais alto implica que a duração do *IC* é menos vulnerável à variação da TE, ou seja, apresenta menos coarticulação. Dessa forma, a palavra do grupo 1 seria produzida com menos coarticulação dos que as do grupo 2 e 3 consecutivamente. Esse resultado explica em parte, em nível articulatorio, o porquê de não termos na língua popular uma forma como “vibra” para “víbora” na língua popular e, por outro lado, de termos formas como “abobra”, “fosfro” e “sabo”. No entanto, os resultados são obscuros na explicação do porquê de “abóboda” ter sido a palavra com o menor *ICP*. Se por um lado esse menor valor pode ser explicado pela maior compressão da consoante [d] em relação à [r], por

outro lado, não encontramos na língua coloquial uma forma paroxítona para “abóboda”. Tais resultados sugerem haver outros fatores atuando na variação/mudança de proparoxítonas a paroxítonas (como frequência lexical, por exemplo), pois fatores exclusivamente articulatórios não foram capazes em nosso corpus de explicar plenamente a variação entre as formas proparoxítonas e paroxítonas de nosso *corpus*.

6.4 Considerações finais

Os resultados deste capítulo apresentam evidências articulatórias de como as proparoxítonas tendem a paroxítonas, avaliando-se o aumento da coarticulação com o aumento da taxa de elocução. Além disso, estudos perceptivos, em andamento, mostram que, a partir de dois cortes no ciclo acústico da vogal [o], temos uma ambigüidade na identificação da palavra como “abóbora” ou “abobra” (em torno de 50%). Nesse caso, medindo-se o *IC* acústico, tem-se uma duração de, aproximadamente, 115 milissegundos. Verificando a duração articulatória do *IC* no *corpus* deste capítulo temos que o mesmo possui média de 102 milissegundos com desvio-padrão de 13 ms. Sendo assim, um *IC* nesta faixa poderia, como sugerem nosso estudo-piloto perceptivo e os dados articulatórios (cap. 6) e acústicos (cap. 5) desta tese, resultar na percepção dessa palavra como “paroxítona”, mesmo tendo sido pronunciada

a vogal pós-tônica medial.

Capítulo 7

Conclusões

Nesta tese mostramos como aspectos contínuos da fala podem explicar a discretização de elementos lingüísticos. Para isso fizemos uso do programa de pesquisa da teoria dos sistemas dinâmicos. A explicação de seu uso advém do fato de que nesse programa não há dicotomia entre corpo e mente. Aplicado à linguagem, subte-se, pois, que a dicotomia clássica entre *língua* e *fala* (ou *competência* e *performance*) não se dá. Dessa forma, aspectos físicos/concretos convivem harmonicamente com os aspectos mentais/abstratos da linguagem. Esta é uma abordagem inovadora na análise de fenômenos lingüísticos, pois, desde Saussure (1916), considerado o pai da ciência lingüística, a Lingüística Tradicional não atribui a devida importância a aspectos contínuos/dinâmicos da fala na explicação de fenômenos gramaticais. Sendo assim, apresentamos aqui uma nova proposta para a análise de fenômenos

lingüísticos — a proposta dinamicista. Argumentos empíricos para isso foram apresentados através dos experimentos desta tese, os quais serão resumidos a seguir.

Na teoria dos sistemas dinâmicos, a forma mais comum de descoberta de novos padrões é utilizar um elemento perturbador do sistema e verificar a emergência de regularidades. Em nosso caso, utilizamos a variação da taxa de elocução, uma das maiores causas de modificação fonética, a fim de observarmos a emergência de novas estruturas rítmicas. Foram perturbadas tanto a estrutura rítmica de sentenças quanto de itens lexicais.

Quanto à reestruturação/reorganização rítmica de sentenças, observamos, utilizando o Modelo Dinâmico do Ritmo (Barbosa, 2006), que, no estudo acústico, o número de grupos acentuais das sentenças diminuiu gradativamente ao serem perturbados pelo aumento da taxa de elocução. Além disso, a variação da taxa de elocução fez com que, proporcionalmente ao aumento da taxa, o número de unidades VV (vogal-a-vogal) por grupo acentual aumentasse, a duração do grupo acentual se mantivesse constante e o desvio-padrão da duração das unidades VV, bem como da duração dos grupos acentuais, fosse menor nas taxas rápidas. Houve, portanto, uma variação gradativa do ritmo, observada através do contorno duracional de unidades VV, relacionada a aspectos quantitativos/fonéticos da linguagem.

No estudo articulatório das sentenças, também foram observadas reestru-

turações rítmicas, tais como no estudo acústico. Especificamente, observamos que os gestos articulatórios modificam dinamicamente suas trajetórias cinemáticas ao se aumentar gradativamente a taxa de elocução. Para este estudo foram relatadas as seguintes fontes cinemáticas (relativas aos gestos articulatórios) de reestruturações rítmicas, medidas através da movimentação da mandíbula: diminuição da duração da aceleração, diminuição do deslocamento vertical máximo, diminuição do deslocamento do gesto de constrição; diminuição (em módulo) das velocidades máximas de deslocamento; diminuição da duração articulatória e tempo-para-velocidade-de-pico proporcional constante. Além disso, verificou-se que a taxa de elocução afeta todos os gestos dos enunciados de maneira uniforme, independentemente dos mesmos pertencerem a um VV em posição de acento frasal ou não.

Quanto à reestruturação/reorganização rítmica de itens lexicais, analisamos, através da Fonologia Articulatória (Browman & Goldstein, 1992), a variação entre proparoxítonas e paroxítonas. Os resultados mostraram que as palavras proparoxítonas, perturbadas dinamicamente pelo aumento da taxa de elocução, reestruturam seus segmentos a ponto de poderem ser reinterpretadas perceptivamente como paroxítonas. Nesse contexto, dois processos lingüísticos dinâmicos foram encontrados: 1) diminuição da duração acústica/articulatória da vogal pós-tônica medial; 2) diminuição da duração acústica/articulatória da vogal pós-tônica final.

No estudo articulatório das proparoxítonas, corroboramos a hipótese de que algumas proparoxítonas podem ser percebidas como paroxítonas, devido à sobreposição de gestos consonantais e vocálicos. Neste estudo, através de medidas da duração articulatória entre consoantes pós-tônicas, verificamos haver um aumento gradativo do grau de coarticulação destas consoantes. Em outros termos, tais gestos foram gradativamente pronunciados mais pertos entre si em taxas mais rápidas. Além disso, verificamos graus distintos de coarticulação de acordo com o tipo de palavra, resultado que pode vir a explicar o porquê de não encontrarmos formas paroxitonizadas para todas as proparoxítonas do português brasileiro.

Como podemos notar, necessitamos ainda realizar estudos perceptivos para complementar as hipóteses aventadas ao longo desta tese de ocorrerem reestruturações rítmicas com o aumento da taxa de elocução. No entanto, de acordo, com um estudo perceptivo em andamento, já temos indicações de palavras proparoxítonas virem a ser reinterpretadas como paroxítonas devido ao aumento da coarticulação entre consoantes pós-tônicas. Esse estudo necessita, porém, ser aperfeiçoado para termos evidências mais contundentes desse fenômeno.

É necessário também, conforme os resultados encontrados nos estudos articulatórios de reestruturações rítmicas em sentenças, elaborar uma metodologia que nos permita, a partir dos dados articulatórios (como duração da

aceleração e deslocamento da constrição), identificar proeminências frasais ao longo dos enunciados.

Por fim, esperamos que esta tese contribua para um novo percurso na análise dos fenômenos lingüísticos — a aplicação da teoria dos sistemas dinâmicos à linguagem. Com isso, podemos conciliar aspectos cognitivos com aspectos fonéticos da linguagem e verificar a importância do tempo na análise das estruturas gramaticais de uma língua.

Apêndice A

Programas em Praat e Perl

A.1 Beat Extractor

```
# BeatExtractor.psc

# Script implemented by Plinio A. Barbosa, IEL/Unicamp, Brazil,

# plinio@iel.unicamp.br

# based originally on Fred Cummins' beat extractor with some

# modifications of the default parameters and some additions

# (an additional filter, and another technique for searching

# for beats).

# Please, DO NOT DISTRIBUTE WITHOUT THE README FILE

# BEATEXTRACTOR_RDM.TXT

# Credits: Fred Cummins, for tips and suggestions,
```

```
# Sophie Scott, for support on her p-centre predictor model
# Paul Boersma, for crucial tips/suggestions on programming
# in Praat, and Pablo Arantes, Jussara Vieira, Alexsandro
# Meireles, and Ana C. Matte, for comments during a debugging
# phase .

# Copyright (C) 2003 Barbosa, P. A.

# # This program is free software; you can redistribute it
# and/or modify it under the terms of the GNU General Public
# License as published by the Free Software Foundation;
# version 2 of the License.

# This program is distributed in the hope that it will be
# useful, but WITHOUT ANY WARRANTY; without even the implied
# warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR
# PURPOSE. See the GNU General Public License for more details.

# # Parameters' input

form Parameters'specification

sentence Path C:\windows\desktop\pline\

word File_(with_extension) apred.wav

choice Speaker_sex 1

button Male

button Female
```

```
choice Filter 1

    button Butterworth

    button Hanning

integer Filter_order 0 (= auto)

real left_Cut_off_frequency_(Hz) 0 (= auto)

real right_Cut_off_frequency_(Hz) 0 (= auto)

real Smoothing_cut_freq_(Hz) 0 (= auto)

choice Technique 2

    button Amplitude

    button Derivative

positive Threshold1_(0.05..0.50) 0.15

positive Threshold2_(0.05..0.15) 0.12

endform

##

# mindur is the minimum duration allowed between two
#consecutive boundaries.

# fcut is the cut-off frequency of the low-pass filters
# used here, and fe/male default are the default cut-off
# frequencies according to speaker sex

mindur = 0.040
```

```
male_default_left = 1000

male_default_right = 1800

female_default_left = 1150

female_default_right = 2100

if left_Cut_off_frequency = 0 ; automatic

    left_Cut_off_frequency = if speaker_sex$ = "Male" then
... 'male_default_left' else 'female_default_left' fi
endif

if right_Cut_off_frequency = 0 ; automatic

    right_Cut_off_frequency = if speaker_sex$ = "Male" then
... 'male_default_right' else 'female_default_right' fi
endif

if filter_order = 0 ; automatic

    filter_order = if filter = 1 then 2 else 0 fi
endif

if smoothing_cut_freq = 0 ; automatic

    smoothing_cut_freq = if technique$ = "Amplitude" then 40
... else 20 fi
endif

fcut = smoothing_cut_freq

##
```

```

fil$ = path$ + file$

Read from file... 'fil$'

filename$ = selected$ ("Sound")

centerf = ('right_Cut_off_frequency' +
... 'left_Cut_off_frequency')/2

w = ('right_Cut_off_frequency' - 'left_Cut_off_frequency')/2

select Sound 'filename$'

# The sound file is filtered according to the preceding choices

if filter = 1

    Filter (formula)... sqrt(1.0/(1.0 +
... ((x-centerf)/w)^(2*'filter_order')))*self; butterworth

... filter elif filter = 2

    Filter (pass Hann band)... 'left_Cut_off_frequency'
... 'right_Cut_off_frequency' 100

endif

Copy... temp

# Filtered sound file's rectification

Formula... abs(self)

w2 = 'smoothing_cut_freq'/10

# Rectified file is low-pass-band filtered producing the beat

# wave file

```

```
Filter (pass Hann band)... 0 'smoothing_cut_freq' w2

max = Get maximum... 0.0 0.0 None

# Beat wave is normalised

Formula... self/max

beatwave$ = filename$ + "_beatwave"

Rename... 'beatwave$'

select Sound 'beatwave$'

derivbeatwave$ = filename$ + "_drvbeatwave"

Copy... temp3

# The derivative of beat wave file is computed and low-pass

# filtered

Formula... (self[col+1] - self[col])/dx

Filter (pass Hann band)... 0 fcut fcut/10

Rename... 'derivbeatwave$'

max = Get maximum... 0.0 0.0 None

Formula... self/max

select Sound temp3

Remove

select Sound 'beatwave$'

begin = Get starting time

end = Get finishing time
```

```
beginindex = Get index from time... 'begin'

beginindex = round(beginindex)

endindex = Get index from time... 'end'

endindex = round(endindex)

fileout$ = filename$ + ".TextGrid"

# Start writing of the TextGrid file

filedelete 'fileout$'

fileappend 'fileout$' File type = "ooTextFile short"
... 'newline$'

fileappend 'fileout$' "TextGrid" 'newline$'

fileappend 'fileout$' 'newline$'

fileappend 'fileout$' 'begin' 'newline$'

fileappend 'fileout$' 'end' 'newline$'

fileappend 'fileout$' <exists> 'newline$'

fileappend 'fileout$' 1 'newline$'

fileappend 'fileout$' "IntervalTier" 'newline$'

fileappend 'fileout$' "VowelOnsets" 'newline$'

fileappend 'fileout$' 'begin' 'newline$'

fileappend 'fileout$' 'end' 'newline$'

i = beginindex

t = begin
```

```
cpt = 0

# Choice of technique

### Technique = 1

# This technique takes the values of the beatwave around
# threshold 1, within the rising parts (derivative > 0)

if technique = 1

    epsilon = 'threshold1'/5

    repeat

        select Sound 'beatwave$'

        value = Get value at index... 'i'

        value = round(1000*value)/1000

        select Sound 'derivbeatwave$'

        valuederiv = Get value at index... 'i'

        if (value < ('threshold1' + epsilon) and value >
        ... ('threshold1' - epsilon)) and (valuederiv > 0.01)

            time'cpt' = Get time from index... 'i'

            if cpt <> 0

                delayedcpt = cpt -1

                if (time'cpt' - time'delayedcpt') <= mindur

                    cpt = cpt -1

            endif
```

```
endif

cpt = cpt + 1

endif

t = t + 0.001

i = Get index from time... 't'

i = round(i)

until (i >= endindex-1)

### # Technique = 2

# This technique takes the values of the maxima of the

# derivative of the beatwave

# greater than threshold 2, where the amplitude of the

# beatwave is greater than threshold 1

elif technique = 2

select Sound 'derivbeatwave$'

drv2beatwave$ = filename$ + "_drv2beatwave"

Copy... temp2

Formula... (self[col+1] - self[col])/dx

Filter (pass Hann band)... 0 fcut fcut/10

Rename... 'drv2beatwave$'

max = Get maximum... 0.0 0.0 None

Formula... self/max
```

```
repeat

  select Sound 'drv2beatwave$'

  drvvalue = Get value at index... 'i'

  drvvalue = round(drvvalue)

  select Sound 'derivbeatwave$'

  value = Get value at index... 'i'

  select Sound 'beatwave$'

  valuebeat = Get value at index... 'i'

  if (drvvalue = 0) and (value > 'threshold2') and (valuebeat
... > 'threshold1') and (valuebeat < 0.3)

    time'cpt' = Get time from index... 'i'

    if cpt <> 0

      delayedcpt = cpt -1

      if (time'cpt' - time'delayedcpt') <= mindur

        cpt = cpt -1

      endif

    endif

    cpt = cpt + 1

  endif

  t = t + 0.001

  i = Get index from time... 't'
```

```
i = round(i)

until (i >= endindex-1)

select Sound  'drv2beatwave$'

plus Sound  temp2

Remove

endif

#####

tmp = cpt+1

fileappend 'fileout$' 'tmp' 'newline$'

temp = 0

for i from 0 to cpt-1

    fileappend 'fileout$' 'temp' 'newline$'

    temp = time'i'

    fileappend 'fileout$' 'temp' 'newline$'

    fileappend 'fileout$' "" 'newline$'

endfor

fileappend 'fileout$' 'temp' 'newline$'

fileappend 'fileout$' 'end' 'newline$'

fileappend 'fileout$' "" 'newline$'

fil$ = path$ + filename$ + "integr"

# Creates a long sound file containing the original sound
```

```
# and the beat wave. Select the created TextGrid file

# containing the detected boundaries

filext$ = fil$ + ".wav"

temp$ = filename$ + "integr"

select all

nb = numberOfSelected ("LongSound")

if nb <> 0

    select LongSound 'temp$'

    Remove

endif

select Sound 'filename$'

plus Sound 'beatwave$'

filedelete filext$'

Write to stereo WAV file... 'filext$'

Open long sound file... 'filext$'

tmp$ = filename$ + "_filt"

select Sound temp

#plus Sound 'beatwave$'

plus Sound 'derivbeatwave$'

plus Sound 'tmp$'

plus Sound 'filename$'
```

```
Remove Read from file... 'fileout$'
```

```
plus LongSound 'temp$'
```

```
Edit
```

A.2 Script do SGdetector

```
# SGdetector.psc
```

```
# Script implemented by Plinio A. Barbosa(IEL/Unicamp) for
```

```
# detecting stress group boundaries from
```

```
# production criteria,
```

```
# namely VV durations. Input: previously
```

```
# segmented VV intervals
```

```
# (TextGrid).
```

```
# plinio@iel.unicamp.br
```

```
# Please, do not distribute without the author's previous
```

```
# authorisation. The sound, TextGrid and Reference-statistics
```

```
# (zaldo.TableOfReal) files need to be in the same directory!
```

```
# # Copyright (C) 2004 Barbosa, P. A.
```

```
# # This program is free software; you can redistribute it
```

```
# and/or modify it under the terms of the GNU General Public
```

```
# License as published by the Free
```

```
# Software Foundation;
# version 2 of the License.

# This program is
# distributed in the hope that it will be
# useful, but WITHOUT ANY
# WARRANTY; without even the implied
# warranty of MERCHANTABILITY or
# FITNESS FOR A PARTICULAR
# PURPOSE. See the GNU General Public
# License for more details. #

form Aquisição dos arquivos

  text Caminho_do_arquivo

  ... c:\windows\desktop\pline\Corpora\StressShift\Locuteur\

  word Arquivo_(com_extensão) bordochines.wav

  integer Camada_de_extracao 1

  choice Referencia: 1

  button Zaldo

endform

# Lê o arquivo de referencia com as triplas (segmento,
# média, desvio-padrão) do locutor
```

```
# Referencia. A variável nseg contém o número total de
# segmentos do arquivo de referência

Read from file... 'referencia$'.TableOfReal

nseg = Get number of rows

# # Lê arquivo e TextGrid (desde q tenha o mesmo nome do
# arquivo de som

arq$ = caminho_do_arquivo$ + arquivo$

Read from file... 'arq$'

nomearq$ = selected$("Sound")

begin = Get starting time end = Get finishing time

arqgrid$ = nomearq$ + ".TextGrid"

arqgrid$ = caminho_do_arquivo$ + arqgrid$

Read from file... 'arqgrid$'

# Extrai todos os intervalos não-vazios do arquivo de som,
```

```
# segundo a segmentação feita no

# arquivo de extensão TextGrid. A variável nselected é o

# número de intervalos extraídos (ou seja, o número de unidades

# VV.

select Sound 'nomearq$'

plus TextGrid 'nomearq$'

Extract non-empty intervals... 'camada_de_extracao' yes

nselected = numberOfSelected ("Sound")

select Sound 'nomearq$'

plus TextGrid 'nomearq$'

Remove

arqout$ = nomearq$ + "dur" + ".txt"

filedelete 'arqout$'
```

```
fileappend 'arqout$' % Segmentos acusticos, duracao (ms) , z,
... z suav.,  fronteira 'newline$'

select all

soundID = selected ("Sound", 1)

select 'soundID'

initialtime = Get starting time

for i from 1 to nselected

  select all

  soundID = selected ("Sound", 'i')

  select 'soundID'

  nome$ = selected$ ("Sound")

  dur = Get duration

  dur = round(dur*1000)

  call zscorecomp 'nome$' 'dur'

  dur'i' = dur

  z'i' = z

  nome'i'$ = nome$

endfor

smz1 = (2*z1 + z2)/3
```

```
deriv1 = smz1

smz2 = (2*z2 + z1)/3

deriv2 = smz2 - smz1

i = 3

if smz1 < smz2

    minsmz = smz1

    maxsmz = smz2

else

    minsmz = smz2

    maxsmz = smz1

endif while i <= (nselected-2)

    del1 = i - 1

    del2 = i - 2

    adv1 = i + 1

    adv2 = i + 2

    smz'i' = (5*z'i' + 3*z'del1' + 3*z'adv1' + z'del2' + z'adv2')

.../13

deriv'i' = smz'i' - smz'del1'

if smz'i' < minsmz

    minsmz = smz'i'

endif
```

```
if smz'i' > maxsmz

    maxsmz = smz'i'

endif

i = i + 1

endwhile


tp1 = nselected -1

tp2 = nselected -2


smz'tp1' = (3*z'tp1'+ z'tp2' + z'nselected')/5


deriv'tp1' = smz'tp1' - smz'tp2'


if smz'tp1' < minsmz

    minsmz = smz'tp1'

endif


if smz'tp1' > maxsmz

    maxsmz = smz'tp1'

endif


smz'nselected' = (2*z'nselected' + z'tp1')/3


deriv'nselected' = smz'nselected' - smz'tp1'
```

```
if smz'nselected' < minsmz

    minsmz = smz'nselected'

endif

if smz'nselected' > maxsmz

    maxsmz = smz'nselected'

endif

tempfile$ = "temp.TableOfReal"

filedelete 'tempfile$'

fileappend 'tempfile$' File type = "ooTextFile short"
... 'newline$'

fileappend 'tempfile$' "TableOfReal" 'newline$'

fileappend 'tempfile$' 'newline$'

fileappend 'tempfile$' 2 'newline$'

fileappend 'tempfile$' columnLabels []: 'newline$'

fileappend 'tempfile$' "position" "smoothed z" 'newline$'
```

```
tpp = nselected + 2

fileappend 'tempfile$' 'tpp' 'newline$'

time = initialtime

fileappend 'tempfile$' row[1]: "0" 0.0 0.0 'newline$'

boundcount = 0

for i from 1 to nselected

  tempsmz = smz'i'

  tptime$ = nome'i'$

  adv1 = i + 1

  btime'i' = 0

  time = time + dur'i'/1000

  fileappend 'tempfile$' row['adv1']: "'tptime$" 'time'
... 'tempsmz' 'newline$'

if i <> nselected

  adv1 = i + 1

  if (deriv'i' >= 0) and (deriv'adv1' < 0)

    boundary = 1

    boundcount = boundcount + 1

    btime'i' = time

    btime'boundcount' = time
```

```
else

    boundary = 0

endif

else

    del1 = i -1

    if smz'i' > smz'del1'

        boundary = 1

        boundcount = boundcount + 1

        btime'i' = time

        bctime'boundcount' = time

    else

        boundary = 0

    endif

endif

tempz = z'i'

tempdur = dur'i'

fileappend 'arqout$' 'tpnome$' 'tempdur' 'tempz:2'

... 'tempzmz:2' 'boundary' 'newline$'

endfor

tp = i+1
```

```
fileappend 'tempfile$' row['tp']: "X" 'end' 0 'newline$'

select all

Remove

tp$ = caminho_do_arquivo$ + tempfile$

Read from file... 'tp$'

Draw scatter plot... 1 2 0 0 0.0 0.0 0.0 0.0 12 no + yes

select all

Remove

Red

for i from 1 to nselected - 1

    if btime'i' <> 0

        bt = btime'i'

        Draw line... 'bt' 'minsmz' 'bt' 'maxsmz'

    endif

endfor

Black

# Write a TextGrid with the stress group boundaries

fileout$ = nomearq$ + "2.TextGrid"

filedelete 'fileout$'
```

```
fileappend 'fileout$' File type = "ooTextFile short"
```

```
... 'newline$'
```

```
fileappend 'fileout$' "TextGrid" 'newline$'
```

```
fileappend 'fileout$' 'newline$'
```

```
fileappend 'fileout$' 'begin' 'newline$'
```

```
fileappend 'fileout$' 'end' 'newline$'
```

```
fileappend 'fileout$' <exists> 'newline$'
```

```
fileappend 'fileout$' 1 'newline$'
```

```
fileappend 'fileout$' "IntervalTier" 'newline$'
```

```
fileappend 'fileout$' "StressGroups" 'newline$'
```

```
fileappend 'fileout$' 'begin' 'newline$'
```

```
fileappend 'fileout$' 'end' 'newline$'
```

```
tmp = boundcount + 2

fileappend 'fileout$' 'tmp' 'newline$'


fileappend 'fileout$' 0.00 'newline$'


fileappend 'fileout$' 'initialtime' 'newline$'


fileappend 'fileout$' "" 'newline$'

temp = initialtime


for i from 1 to boundcount

    fileappend 'fileout$' 'temp' 'newline$'

    temp = bctime'i'

    fileappend 'fileout$' 'temp' 'newline$'

    fileappend 'fileout$' "" 'newline$'

endfor


fileappend 'fileout$' 'temp' 'newline$'


fileappend 'fileout$' 'end' 'newline$'
```

```
fileappend 'fileout$' "" 'newline$'
```

```
##
```

```
arqgrid1$ = caminho_do_arquivo$ + nomearq$ + ".TextGrid"
```

```
arqgrid2$ = caminho_do_arquivo$ + fileout$
```

```
Read from file... 'arqgrid1$'
```

```
Read from file... 'arqgrid2$'
```

```
select all
```

```
Merge
```

```
##
```

```
procedure zscorecomp nome$ dur
```

```
    sizeunit = length (nome$)
```

```
    sumofmeans = 0
```

```
    sumofvar = 0
```

```
    cpt = 1
```

```
    while cpt <= sizeunit
```

```
        nb = 1
```

```
        terminate = 0
```

```

k = 1

seg$ = mid$(nome$,cpt,1)

if cpt < sizeunit
    if mid$(nome$,cpt+1,1) == "h" or mid$(nome$,cpt+1,1)
... == "N"

        nb = nb + 1

        seg$ = seg$ + mid$(nome$,cpt+1,1)

    endif

    if (cpt+nb <= sizeunit)

        tp$ = mid$(nome$,cpt,1)

        call isvowel 'tp$'

        if ((mid$(nome$,cpt+nb,1) = "I") or (mid$(nome$,
... cpt+nb,1) = "U"))and truevowel

            seg$ = seg$ + mid$(nome$,cpt+nb,1)

            nb= nb+1

        endif

    endif

endif

j = 1

select all

tableID = selected ("TableOfReal")

```

```

select 'tableID'

while (j <= nseg) and not terminate

    label$ = Get row label... 'j'

    if seg$ = label$

        terminate = 1

        mean = Get value... 'j' 1

        sd      = Get value... 'j' 2

        sumofmeans = mean + sumofmeans

        sumofvar= sd*sd + sumofvar

    endif

    j = j+1

endwhile

cpt= cpt+nb

endwhile

z = (dur - sumofmeans)/sqrt(sumofvar)

endproc

procedure isvowel temp$

    truevowel = 0

    if temp$ = "i" or temp$ = "e"  or temp$ = "a"  or temp$ = "o"

... or temp$ = "u" or temp$ = "I" or temp$ = "E"

```

```
... or temp$ = "A"

or temp$ = "O"  or temp$ = "U"

    truevowel = 1

endif

endproc
```

A.3 Duração

```
# Duração.psc

# Script implemented by Alexsandro R. Meireles


form Aquisição dos arquivos

    text Caminho_do_arquivo E:\Sociolinguistica\MG\

    word Arquivo sabadolento

    integer Camada_de_extracao 1

endform


arq$ = caminho_do_arquivo$ + arquivo$

arqgrid$ = arq$ + ".TextGrid"

arqtext$ = caminho_do_arquivo$ + arquivo$ + ".txt"

Read from file... 'arqgrid$'
```

```
select all

Extract tier... 1 Down to TableOfReal (any)

select TableOfReal VowelOnsets

Write to headerless spreadsheet

file... 'arqtext$'

select all

Remove
```

A.4 Rename

```
# Rename.pl

# Script implemented by Alexsandro R. Meireles


#!/usr/bin/perl -w


print "Enter prt file name: "; chomp($lstfile = <STDIN>);


print "Enter speech rate (normal/slow/fast/normal): ";

chomp($rate = <STDIN>);
```

```
$subj = "jm";
```

```
$old = "$subj" . "_bp" . "_04_25_2005_" . "$rate" ;
```

```
open(LST, $lstfile); while(<LST>) {
```

```
    chomp;
```

```
    s/^\s+//;
```

```
    s/\s+$//;
```

```
    ($dum1, $emaid, $code, $uid, $sent) = split(/\s+/, $_, 5);
```

```
    if( $emaid < 10 ) {
```

```
        $emaid = "00$emaid";
```

```
    }
```

```
    else {
```

```
        $emaid = "0$emaid";
```

```
    }
```

```
$oldname = $old . ".$semaid";
```

```
if( $uid < 10 ) {
```

```
    $uid = "0$uid";
```

```
}
```

```
$newname = "$subj" . "_bp_" . "$rate" . "_$uid".$semaid";
```

```
print "\nOld= $oldname  New= $newname\n";
```

```
system("copy $oldname $newname");
```

```
$oldname =~ s/\.0/\.M/; $newname =~ s/\.0/\.M/;
```

```
system("copy $oldname $newname");
```

```
$oldname =~ s/\.M/\.A/; $newname =~ s/\.M/\.A/;
```

```
system("copy $oldname $newname");
```

```
$oldname =~ s/\.A/\.R/; $newname =~ s/\.A/\.R/;
```

```
system("copy $oldname $newname");

$oldname =~ s/\.R/\.T/; $newname =~ s/\.R/\.T/;

system("copy $oldname $newname");

}
```

A.5 Formant Extractor

```
# FormantExtractor.psc

# Script implemented by Alexsandro R.

# Meireles (LAFAPE/IEL/Unicamp) for obtaining

# FFT and LPC of non-empty, non-?

# previously segmented intervals (TextGrid)and

# writing them together on a praat picture.

# Please, do not distribute without

# the author's previous authorization form

Aquisição dos arquivos

text Caminho_do_arquivo E:\Sociolinguistica\MG\
```

```
word Arquivo aboboralento

choice Som 1

        button wav

        button nsp

integer Camada_de_extracao 2

positive Filter_order 24 (=sampling rate in kHz + 2)

endform


arq$ = caminho_do_arquivo$ + arquivo$


arqgrid$ = arquivo$ + ".TextGrid"


arqgrid$ = caminho_do_arquivo$ + arqgrid$


Read from file... 'arq$'.'som$'


samprate = Get sample rate

nyquistf = samprate/2


nformants = round(nyquistf/1000)
```

```
To Formant (burg)... 0.0 'nformants' 'nyquistf' 0.025 50
```

```
Read from file...
```

```
'arqgrid$'
```

```
select Sound 'arquivo$'
```

```
plus TextGrid 'arquivo$'
```

```
Extract non-empty intervals... 'camada_de_extracao' yes
```

```
nselected = numberOfSelected ("Sound")
```

```
select Sound 'arquivo$'
```

```
plus
```

```
TextGrid 'arquivo$'
```

```
Remove
```

```
arqout$ = arquivo$ + "LPC" + ".txt"
```

```
filedelete 'arqout$'
```

```
fileappend "'caminho_do_arquivo$'\arqout$'" Segmento tempo F1 B1
```

```
... F2 B2 F3 B3 F4 B4
```

```
'newline$'

for i from 1 to nselected

select all

    soundID = selected ("Sound", 'i')

    select 'soundID'

    nome$ = selected$ ("Sound")

    tpesq = Get starting time

    tpdire = Get finishing time

    tp = (tpesq + tpdire)/2

    tp = round(tp*1000)/1000

    To LPC (burg)... 'filter_order' 0.025 0.005 50

    To Spectrum (slice)... 0 20 0 50

        Red

        Draw... 0 0 0 0 yes

    select 'soundID'

    To Spectrum (fft)

        Grey

        Draw... 0 0 0 0 yes

        Write to praat picture file... 'caminho_do_arquivo$'\nome$'.prapic

    Erase all
```

```
select all
```

```
formantID = selected ("Formant")
```

```
select 'formantID'
```

```
f1 = Get value at time... 1 'tp' Hertz Linear
```

```
f1 = round(f1)
```

```
b1 = Get bandwidth at time... 1 'tp' Hertz Linear
```

```
b1 = round(b1)
```

```
f2 = Get value at time... 2 'tp' Hertz Linear
```

```
f2 = round(f2)
```

```
b2 = Get bandwidth at time... 2 'tp' Hertz Linear
```

```
b2 = round(b2)
```

```
f3 = Get value at time... 3 'tp' Hertz Linear
```

```
f3 = round(f3)
```

```
b3 = Get bandwidth at time... 3 'tp' Hertz Linear
```

```
b3 = round(b3)
```

```
f4 = Get value at time... 4 'tp' Hertz Linear
```

```
f4 = round(f4)
```

```
b4 = Get bandwidth at time... 4 'tp' Hertz Linear
```

```
b4 = round(b4)
```

```
fileappend "'caminho_do_arquivo$'\arqout$" 'nome$' 'tp' 'f1' 'b1' 'f2'
```

```
... 'b2' 'f3' 'b3' 'f4' 'b4' 'newline$'  
endif  
endfor  
select all  
Remove
```

A.6 Dictionary

```
#!/usr/bin/perl -w
```

```
#use strict;
```

```
#Script developed by Alexsandro Meireles for looking patterns in a dictionary.
```

```
#Send e-mail to meirelesalex@gmail.com to ask permission for use.
```

```
print "File to search: ";
```

```
my $arquivo=<STDIN>;
```

```
chomp $arquivo;
```

```
print "Pattern to look for: ";

my $pat=<STDIN>;

chomp $pat;

print "Output file: ";

my $output=<STDIN>;

chomp $pat;

my($file);

open(ARQ, $arquivo) || die "Cannot open $arquivo: $!";
open(OUT, ">$output") || die "Cannot open $output: $!";

$length = 0;

@words = '';

$match = 0;

$i=0;

while (<ARQ>){($line= <ARQ>)

{

# chop($line);

#$_ = $line;    # string is stored in special storage

foreach $word (@words = split(/ /)) { # field separator
```

```
is any number of " " (space)

print "$word\n";

$_ = $word; # GREAT! I had to specify the input to the pattern matching

if (/$pat/) {

    $match++;

    $i++;

    #print OUT "Word $i = $word[$i]";

    print $word;

    print "yEAH\n";

    print "We have found $match match(es).\n";

    print OUT "Word ($i) = $word\n";

}

$length++; }

}
```

```
close(ARQ);
```

Referências Bibliográficas

ABAURRE-GNERRE, M. B. *Phonostylistic aspects of a Brazilian Portuguese dialect: implications for syllable structure constraints*. Tese (Doutorado) — Faculty of the Graduate School of State University of New York at Buffalo, 1979.

ABAURRE-GNERRE, M. B. Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. *Caderno de Estudos Lingüísticos*, v. 2, p. 23–44, 1981.

ABBS, J. H.; GRACCO, V. L.; COLE, K. J. Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *Journal of Motor Behavior*, v. 16, p. 195–231, 1984.

ALBANO, E. *O gesto e suas bordas: esboço de fonologia acústico-articulatória*. São Paulo, Campinas: Mercado de Letras: Associação de Leitura do Brasil - ALB, Fapesp, 2001.

ALBANO, E.; BARBOSA, P.; GAMA-ROSSI, A.; MADUREIRA, S.; SILVA, A. A interface fonética-fonologia e a interação prosódia-segmentos. In: *Estudos Lingüísticos XXVII, Anais do XLV Seminário do GEL - GEL '97, Unicamp*. Campinas, SP: [s.n.], 1998. p. 135–143.

AMARAL, A. *O dialeto caipira*. São Paulo: Anhembi, 1955.

AMARAL, M. P. *As proparoxítonas: teoria e variação*. Tese (Doutorado) — PUCRS, 2000.

AMARAL, M. P. A síncope em proparoxítonas: uma regra variável. In: BISSOL, L.; BRESCANCINI, C. (Ed.). *Fonologia e Variação: recortes do português brasileiro*. Porto Alegre: EdPUCRS, 2002.

BAER, T.; LÖFQVIST, A.; MCGARR, N. S. Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. *J. Acoust. Soc. Am.*, v. 73, p. 1304–1308, 1983.

BAILLY, G.; LABOISSIÈRE, R.; SCHWARTZ, J. L. Formant trajectories as audible gestures: an alternative for speech synthesis. *Journal of Phonetics*, n. 19, p. 9–23, 1991.

BAKEN, R.; ORLIKOFF, R. *Clinical measurement of speech and voice (2nd edition)*. San Diego, CA: Singular Publishing Group, 2000.

BARBOSA, P. A. *Caractérisation et génération automatique de la structuration rythmique du français*. Tese (Doutorado) — Institut National Polytechnique de Grenoble, França, 1994.

BARBOSA, P. A. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration. In: *Proceedings of the First ESCA Tutorial Research Workshop on Speech Production Modeling and Fourth Speech Production Seminar*. Autrans, França: [s.n.], 1996. p. 85–88.

BARBOSA, P. A. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on segmental duration generation. *Cadernos de Estudos Lingüísticos*, n. 31, p. 33–53, 1996.

BARBOSA, P. A. É possível integrar o discreto e o contínuo em um modelo de produção do ritmo da fala? *Cadernos de Estudos Lingüísticos*, n. 40, p. 29–38, 2001.

BARBOSA, P. A. Explaining Brazilian Portuguese resistance to stress shift with a coupled-oscillator model of speech rhythm production. *Cadernos de Estudos Lingüísticos*, v. 43, p. 71–92, 2002.

BARBOSA, P. A. Explaining cross-linguistic rhythmic variability via a coupled-oscillator model of rhythm production. In: *Proceedings of the Speech*

Prosody 2002 Conference. Aix-en-Provence, França: Keikishi Hirose Laboratory, University of Tokyo, 2002. p. 163–166.

BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas, São Paulo: Pontes Editores, Fapesp, 2006.

BARBOSA, P. A.; ARANTES, P.; MEIRELES, A.; VIEIRA, J. Abstractness in speech-metronome synchronisation: P-centres as cyclic attractors. In: *Proceedings of the Ninth European Conference on Speech Communication and Technology (Interspeech 2005)*. Lisboa, Portugal: [s.n.], 2005. p. 1441–1444.

BARBOSA, P. A.; BAILLY, G. Characterisation of rhythmic patterns for text-to-speech synthesis. *Speech Communication*, n. 15, p. 127–137, 1994.

BECKMAN, M. E.; EDWARDS, J.; FLETCHER, J. Prosodic structure and tempo in a sonority model of articulatory dynamics. In: DOCHERTY, G. J.; LADD, D. R. (Ed.). *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press, 1992. p. 68–86.

BENVENISTE, E. La notion de “rythme” dans son expression linguistique. *J. Psychol. Norm. Path.*, v. 44, p. 401–411, 1951.

BERTHOZ, A. *The brain's sense of movement*. London, England: Harvard University Press, 2000.

BOYLLS, C. C. *A theory of cerebellar function with applications to locomotor behavior in the cat*. 1975. U. Mass., Dept. of Computer and Information Science. COINS Tech. Rep. 76-1.

BROSSES, C. de. *Traité de la formation mécanique des langues*. Paris: Saillant, Vincent et Desaint, 1765.

BROWMAN, C.; GOLDSTEIN, L. Towards an articulatory phonology. *Phonology Yearbook*, v. 3, p. 219–252, 1986.

BROWMAN, C.; GOLDSTEIN, L. Articulatory gestures as phonological units. *Phonology*, v. 6, p. 201–251, 1989.

BROWMAN, C.; GOLDSTEIN, L. Gestural specification using dynamically-defined articulatory gestures. *Journal of Phonetics*, v. 18, p. 299–320, 1990.

BROWMAN, C.; GOLDSTEIN, L. Tiers in articulatory phonology with some implications for casual speech. In: KINGSTON, J.; BECKMAN, M. E. (Ed.). *Papers in Laboratory Phonology I*. Cambridge, Reino Unido: Cambridge University Press, 1990. p. 341–376.

BROWMAN, C.; GOLDSTEIN, L. Articulatory phonology: an overview. *Phonetica*, v. 49, p. 155–180, 1992.

BROWMAN, C.; GOLDSTEIN, L.; KELSO, J. A. S.; RUBIN, P.; SALTZMAN, E. Articulatory synthesis from underlying dynamics. *J. Acoust. Soc. Am.*, v. 75, p. S22–S23, 1984.

BYRD, D. Relations of sex and dialect to reduction. *Speech Communication*, v. 15, p. 39–54, 1994.

BYRD, D. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, v. 57, n. 1, p. 3–169, 2000.

BYRD, D. Frontiers and challenges in articulatory phonology. In: *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: [s.n.], 2003. p. 89–92.

BYRD, D.; KAUN, A.; NARAYANAN, S.; SALTZMAN, E. Phrasal signatures in articulation. In: BROE, M. B.; PIERREHUMBERT, J. B. (Ed.). *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge, Reino Unido: Cambridge University Press, 2000. p. 70–87.

BYRD, D.; SALTZMAN, E. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, v. 26, p. 173–199, 1998.

BYRD, D.; SALTZMAN, E. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, v. 31, p. 149–180, 2003.

CÂMARA-JR., J. M. *Dicionário de lingüística e gramática*. Petrópolis: Vozes, 1988. 18ª edição.

CAMBIER-LANGEVELD, T. The domain of final lengthening in production of Dutch. In: COERTS; HOOP de (Ed.). *Linguistics in the Netherlands*. Amsterdam: John Benjamins, 1997. p. 13–24.

CAMPBELL, W. N.; ISARD, S. D. Segment durations in a syllable frame. *Journal of Phonetics*, v. 19, p. 37–47, 1991.

CATANIA, C. A.; HARNAD, S. *The Selection of Behavior*. Cambridge: Cambridge University Press, 1988.

CHÂTELET, G. *Les Enjeux du mobile — Mathématiques, physique, philosophie*. Paris: Le Seuil, 1993.

CHEN, M.; WANG, W. S.-Y. Sound change: Actuation and implementation. *Language*, v. 51, p. 255–81, 1975.

CHO, T. *Effects of prosody on articulation in English*. Tese (Doutorado) — University of California, Los Angeles, 2001.

CHO, T.; KEATING, P. A. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, v. 29, p. 155–190, 2001.

CHOMSKY, N. *Syntactic Structures*. Berlin and New York: The Hague: Mouton, 1985. [1957], Data de publicação original.

CHOMSKY, N.; HALLE, M. *The sound pattern of English*. Cambridge, Estados Unidos: The MIT Press, 1991. [1968], Publicado originalmente por Harper & Row, Nova York.

CHURCH, R. M.; BROADBENT, H. A. Alternative representations of time. *Cognition*, v. 37, p. 55–81, 1990.

CHURCHLAND, P. S.; SEJNOWSKI, T. J. *The computational brain*. Cambridge, MA: The MIT Press, 1992.

CLARKE, E. F. The perception of expressive timing in music. *Psychological Research*, v. 51, p. 2–9, 1989.

CLASSÉ, A. *The Rhythm of English Prose*. Oxford: Blackwell, 1939.

CLEMENTS, G. N. The geometry of phonological features. *Phonology Yearbook*, v. 2, p. 225–252, 1985.

CLEMENTS, G. N.; HUME, E. The internal organization of speech sounds. In: GOLDSMITH, J. (Ed.). *The Handbook of Phonological Theory*. London: Blackwell, 1995.

CLEMENTS, G. N.; KEYSER, S. J. *CV phonology: a generative theory of the syllable*. [S.l.]: MIT Press, 1983.

CUMMINS, F. *Rhythmic coordination in English speech: An experimental study*. Tese (Doutorado) — Indiana University, 1997.

CUMMINS, F.; PORT, R. Rhythmic constraints on “stress-timing” in English. *Journal of Phonetics*, v. 26, p. 145–171, 1998.

DAUER, R. M. Stress-timing and syllable-timing re-analyzed. *Journal of Phonetics*, v. 11, p. 51–62, 1983.

EDWARDS, J.; BECKMAN, M. E.; FLETCHER, J. The articulatory kinematics of final lengthening. *J. Acoust. Soc. Am.*, v. 89, p. 369–382, 1991.

ELMAN, J. L. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, v. 7, p. 195–225, 1991.

ELMAN, J. L. Language as a dynamical system. In: PORT, R. F.; GELDER, T. van (Ed.). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge-Massachusetts, Inglaterra: The MIT Press, 1995. p. 195–225.

ERICKSON, D. On phrasal organization and jaw opening. In: *From Sound to Sense: 50+ Years of Discoveries in Speech Communication, MIT, June 11-13*. [S.l.: s.n.], 2004. p. S15.

ERIKSSON, A. *Aspects of Swedish speech rhythm*. Tese (Doutorado) — University of Göteborg, 1991. Gothenburg Monograph in Linguistics 9.

FANT, G.; KRUCKENBERG, A. Preliminaries to the study of Swedish prose reading and reading style. *STL-QPSR*, v. 2, p. 1–80, 1989.

FARNETANI, E. Coarticulation and connected speech processes. In: HARD-CASTLE, W. J.; LAVER, J. (Ed.). *The Handbook of Phonetic Sciences*. [S.l.]: Blackwell Pub, 1999.

FERSTER, C. B.; SKINNER, B. F. *Schedules of reinforcement*. New York: Appleton-Century Crofts, 1957.

FLASH, T.; SEJNOWSKI, T. Computational approaches to motor control. *Current Opinion in Neurobiology*, v. 11, p. 655–662, 2001.

FOUGERON, C. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, v. 29, p. 109–135, 2001.

FOWLER, C. A. *Timing Control in Speech Production*. Bloomington, Estados Unidos: Indiana University Linguistic Club, 1977.

FOWLER, C. A. Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, v. 8, p. 113–133, 1980.

FOWLER, C. A.; RUBIN, P.; REMEZ, R. E.; TURVEY, M. T. Implications for speech production of a general theory of action. In: BUTTERWORTH, B. (Ed.). *Speech Production, v. 1*. Nova York: Academic Press, 1980. p. 373–420.

FRAISSE, P. Les rythmes. *Journal Français d'Oto-Rhino-laryngologie*, n. Supplément numéro 7, p. 23–33, 1968.

FRAISSE, P. *La psychologie du rythme*. Paris: Presses Universitaires de France, 1974.

FRAISSE, P. Rhythm and tempo. In: DEUTSCH, D. (Ed.). *The Psychology of Music*. Nova York, Estados Unidos: Academic Press, 1982. p. 149–180.

GAY, T. Effect of speaking rate on vowel formant movements. *J. Acoust. Soc. Am.*, v. 63, p. 223–230, 1978.

GAY, T.; HIROSE, H. Effect of speaking rate on labial consonant production: A combined electromiographic/high speed motion picture study. *Phonetica*, v. 27, p. 44–56, 1973.

GAY, T.; USHIJIMA, T. Effect of speaking rate on stop consonant-vowel articulation. In: *Proceedings of the Speech Communication Seminar 1974*. Estocolmo: [s.n.], 1975. p. 205–209.

GENETTE, G. *Mimologiques. Voyage en Cratylie*. Paris: Éditions du Seuil, 1976.

GOLDSMITH, J. *Autosegmental and metrical phonology*. Oxford: Blackwell Publishers, 1990.

GRILLNER, S. On the neural control of movement: a comparison of different basic rhythmic behaviors. In: STENT, G. S. (Ed.). *Function and formation of neural system*. Berlim: Dahlem, 1977.

GRILLNER, S. Possible analogues in the control of innate motor acts and the production of sound in speech. In: GRILLNER, S.; LINDBLOM, B.; LUBKER, J.; PERSSON, A. (Ed.). *Speech Motor Control*. Nova Iorque: Pergamon, 1982.

GUILLAUME, P. *La Psychologie de la forme*. Paris: Flammarion, 1979.

GUY, G. R. Variation in the group and the individual. In: LABOV, W. (Ed.). *Locating language in time and space*. Nova Iorque: Academic Press, 1980.

HAKEN, H. *Synergetics: An Introduction*. Berlim: Springer-Verlag, 1983.

HAKEN, H. *The Science of Structure: Synergetics*. Nova Iorque: Van Nostrand Reinhold, 1984.

HAKEN, H.; KELSO, J. A. S.; BUNZ, H. A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, v. 51, p. 347–356, 1985.

HARRIS, K. S. Vowel duration change and its underlying physiological mechanisms. *Lang. Speech*, v. 21, p. 354–361, 1978.

HARRIS, K. S.; TULLER, B.; KELSO, J. A. S. Temporal invariance in the production of speech. In: PERKELL, J. S.; KLATT, D. H. (Ed.). *Invariance and variability in speech processes*. Hillsdale, NJ: Erlbaum, 1986.

HAUY, A. B. *História da língua portuguesa - séculos XII, XIII e XIV. 2. ed.* São Paulo: Ática, 1994.

HAWKINS, S. An introduction to task dynamics. In: DOCHERTY, G.; LADD, D. R. (Ed.). *Papers in laboratory phonology I: gesture, segment, prosody*. Cambridge: Cambridge University Press, 1992. p. 9–25.

HAYES, B. *A metrical Theory of Stress Rules*. New York/London: Garland Publishing, 1985.

HAYES, B. *Metrical Stress Theory - Principles and Case Studies*. [S.l.]: University of Chicago Press, 1995.

HOLLERBACH, J. An oscillator theory of handwriting. *Biological Cybernetics*, v. 39, p. 139–156, 1981.

HOOPER, J. B. *An introduction to natural generative phonology*. New York: Academic Press, 1976.

KEATING, P. *A phonetic study of a voicing contrast in Polish*. Tese (Doutorado) — Department of Linguistics, Brown University, 1979.

KEATING, P. Universal phonetics and the organization of grammars. In: FROMKIN, V. A. (Ed.). *Phonetic Linguistics. Essays in honor of Peter Ladefoged*. Orlando, Estados Unidos: Academic Press, 1985. p. 115–132.

KEATING, P.; CHO, T.; FOUGERON, C.; HSU, C.-S. Domain-initial articulatory strengthening in four languages. In: LOCAL, J.; OGDEN, R.; TEMPLE, R. (Ed.). *Phonetic Interpretation. Papers in Laboratory Phonology VI*. Cambridge, Reino Unido: Cambridge University Press, 2003. p. 145–163.

KEATING, P.; LINDBLOM, B.; LUBKER, J.; KREIMAN, J. Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, v. 22, p. 407–422, 1994.

KELSO, J. A. S. *Dynamic patterns: the self-organization of brain and behavior*. Cambridge, Estados Unidos: MIT Press, 1995.

KELSO, J. A. S. The complementary nature of coordination dynamics: self-organization and the origins of agency. *Nonlinear Phenomena in Complex Systems*, v. 5, p. 364–371, 2002.

KELSO, J. A. S.; GUZMAN, G. C. de. Order in time: how the cooperation between the hands informs the design of the brain. In: HAKEN, H. (Ed.). *Neural and synergetic computers*. Berlin: Springer-Verlag, 1988. p. 271–287.

KELSO, J. A. S.; SALTZMAN, E. L.; TULLER, B. The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, v. 14, p. 29–59, 1986.

KELSO, J. A. S.; TULLER, B.; FOWLER, C. A. The functional specificity of articulatory control and coordination. *J. Acoust. Soc. Am.*, v. 72, p. S103, 1982.

KELSO, J. A. S.; TULLER, B.; HARRIS, K. S. A dynamic pattern perspective on the control and coordination of movement. In: MACNEILAGE, P. (Ed.). *The Production of Speech*. Nova Iorque: Springer-Verlag, 1983.

KELSO, J. A. S.; TULLER, B.; VATIKIOTIS-BATESON, E.; FOWLER, C. A. Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, v. 10, p. 812–832, 1984.

KENT, R.; READ, C. *The Acoustic Analysis of Speech*. San Diego, London: Singular Publishing Group, INC., 1992.

KENT, R. D.; MOLL, K. L. Articulatory timing in selected consonant sequences. *Br. Lang.*, v. 2, p. 304–323, 1975.

KENT, R. D.; NETSELL, R. Effects of stress contrasts on certain articulatory parameters. *Phonetica*, v. 24, p. 23–44, 1971.

KOEHLER, W. *Gestalt Psychology*. New York: Liveright, 1929.

KONOPCZYNSKI, G. A developmental model of acquisition of rhythmic patterns: results from a cross-linguistic study. In: *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Estocolmo, Suécia: [s.n.], 1995. v. 4, p. 22–29.

KOZHEVNIKOV, V. A.; CHISTOVICH, L. A. Speech: Articulation and perception. In: *Joint Publications Research Service*. [S.l.: s.n.], 1965. p. 543.

KROCH, A. S. Toward a theory of social dialect variation. *Language in Society*, v. 7, n. 1, p. 17–36, 1978.

KUEHN, D.; MOLL, K. A cineradiographic study of vc and cv articulatory velocities. *Journal of Phonetics*, v. 4, p. 303–320, 1976.

LADEFOGED, P. *Preliminaries to Linguistic Phonetics*. Chicago, Estados Unidos: The University of Chicago Press, 1971.

LEHISTE, I. *Suprasegmentals*. Cambridge, Massachussets: MIT Press, 1970.

LESTIENNE, F. Effects of inertial load and velocity on the braking process of voluntary limb movements. *Experimental Brain Research*, v. 35, p. 407–418, 1979.

LIBERMAN, M.; PRINCE, A. On stress and linguistic rhythm. *Linguistic Inquiry*, v. 8, n. 2, p. 249–336, 1977.

LINDBLOM, B. Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.*, v. 35, p. 1773–1781, 1963.

LINDBLOM, B. Economy of speech gestures. In: MACNEILAGE, P. F. (Ed.). *The production of speech*. Nova York: Springer-Verlag, 1983. p. 217–246.

LOBATO, J. B. M. *A Menina do Narizinho Arrebitado*. São Paulo: Revista do Brasil, 1920.

LÖFQVIST, A.; YOSHIOKA, H. Interarticulatory programming in obstruent production. *Phonetica*, v. 38, p. 21–34, 1981.

MACNEILAGE, P. F.; HANSON, R.; KRONES, R. Control of the jaw in relation to stress in English. *J. Acoust. Soc. Am.*, v. 48, p. 120A, 1970.

MARCUS, S. M. Acoustic determinants of Perceptual center (p-center) location. *Perception and Psychophysics*, v. 30, n. 3, p. 247–256, 1981.

MARINO, V. C. C. Avaliação instrumental da fala: articulografia. In: *IX Jornada Fonoaudiologica - Prof. Mariza Ribeiro Feniman, 2002, Bauru. IX Jornada Fonoaudiologica - Prof. Mariza Ribeiro Feniman*. [S.l.: s.n.], 2002. p. 24-24.

MCAULEY, J. D. Time as phase: A dynamic model of time perception. In: *Proceedings of the 16th Annual Conference of the Cognitive Science Society*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1994.

MCAULEY, J. D. *Perception of time as phase: toward an adaptive-oscillator model of rhythmic pattern processing*. Tese (Doutorado) — Indiana University, Estados Unidos, 1995.

MEIRELES, A. *Processos fonético-fonológicos decorrentes do aumento da velocidade de fala no português brasileiro*. 2001. Dissertação de mestrado defendida na Universidade Federal de Minas Gerais.

MEIRELES, A. R.; BARBOSA, P. A. O papel da taxa de elocução nos processos dinâmicos de mudança lingüística. *D.E.L.T.A.*, submetido.

MEYER, E. A. *Englische Lautdauer*. Leipzig: Harrassowitz, 1903.

MEYER, E. A. Zur Vokaldauer im Deutschen. In: *Nordiska Studia tillegnade Adolf Noreen*. Uppsala: Appelberg, 1904.

MITLEB, F. M. Vowel length contrast in Arabic and English: a spectrographic test. *Journal of Phonetics*, v. 12, p. 229–235, 1984.

MITLEB, F. M. Some aspects of Arabic duration. *Journal of the International Phonetic Association*, v. 22, n. 1/2, p. 27–34, 1992.

MUNHALL, K. G.; KELSO, J. A. S. *Phase dependent sensitivity to perturbation reveals the nature of speech coordinative structures*. 1985. Paper presented at the meeting of the Acoustical Society of America, Nashville, TN, U.S.A.. November 1985.

MUSSA-IVALDI, F. A. Geometrical principles in motor control. In: ARBIB, M. (Ed.). *The handbook of brain theory and neural networks*. Cambridge, MA: MIT Press, 1995. p. 434–438.

NASHNER, L. M. Fixed patterns of rapid postural responses among leg muscles during stance. *Experimental Brain Research*, v. 30, p. 13–24, 1977.

NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.

NITTROUER, S.; MUNHALL, K.; KELSO, J. A. S.; TULLER, B.; HARRIS, K. S. Patterns of interarticulator phasing and their relation to linguistic structure. *J. Acoust. Soc. Am.*, v. 84, p. 1653–1661, 1988.

NOOTEBOOM, S. G. Temporal patterns in Dutch. In: *Proceedings of the VIIth International Congress of Phonetic Sciences*. Montreal, Canadá: Mouton, 1972. p. 984–989.

NORD, L. Acoustic studies of vowel reduction in Swedish. *KTH-QPSR Stockholm 4*, p. 19–36, 1986.

NORD, L.; KRUCKENBERG, A.; FANT, G. Some timing studies of prose, poetry and music. *Speech Communication*, v. 9, p. 477–483, 1990.

NUNES, J. J. *Compêndio de gramática histórica portuguesa - fonética e morfologia*. Lisboa: Livraria Clássica Editora, 1969.

NUNES, J. J. *Cantigas d'amigo dos trovadores galego-portugueses*. Lisboa: Centro do Livro Brasileiro, 1973.

ÖHMAN, S. E. G. Coarticulation in VCV utterances: spectrographic measurements. *J. Acoust. Soc. Am.*, v. 39, n. 1, p. 151–168, 1966.

OLIVEIRA, M. A. de. A controvérsia neogramática reconsiderada. *International Journal of the Sociology of Language*, v. 89, p. 93–105, 1991. Tradução do autor do original em inglês.

OLIVEIRA, M. A. de. Aspectos de difusão lexical. *Revista de Estudos Lingüísticos - jul/dez*, v. 1, p. 31–41, 1992.

PASDELOUP, V. *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Tese (Doutorado) — Université de Provence Aix-Marseille 1, 1990.

PASDELOUP, V. A prosodic model for French text-to-speech synthesis: a psycholinguistic approach. In: BAILLY, G.; BENOÎT, C. (Ed.). *Talking Machines: Theories, Models and Designs*. Amsterdam, Holanda: Elsevier Science Publishers B.V., 1992. p. 335–348.

PASDELOUP, V. Le rythme n'est pas élastique: étude préliminaire de l'influence du débit de parole sur la structuration temporelle. In: *Actes des 25èmes Journées d'Etude sur la Parole, 19-22 Avril*. Fès, Maroc: [s.n.], 2004. p. 397–400.

PASDELOUP, V. Figures et fond dans la scène prosodique: leur résistance face aux variations du débit de parole. In: *IDP 2005 8-9 Septembre*. [S.l.: s.n.], 2005.

PASDELOUP, V.; ESPESSER, R.; FARAJ, M. Rate sensitivity of syllable in French: a perceptual illusion? In: *CD-ROM proceedings. Speech Prosody, 2006*. Dresden, Alemanha: [s.n.], 2006.

PERKELL, J.; COHEN, M.; SVIRSKY, M.; MATTHIES, M.; GARABITA, I.; JACKSON, M. Eletromagnetic midsagittal articulometer (EMMA)

systems for transducing speech articulatory movements. *J. Acoust. Soc. Am.*, p. 3078–3096, 1992.

PERKELL, J.; KLATT, D. H. (Ed.). *Invariance and Variability in Speech Processes*. Ann Arbor: Erlbaum Hillsdale, 1986.

PETERSON, G.; LEHISTE, I. Duration of syllabic nuclei in English. *J. Acoust. Soc. Am.*, v. 25, p. 175–184, 1960.

PHILIPS, B. S. Word frequency and the actuation of sound change. *Language*, v. 60, n. 2, p. 320–42, 1984.

PLAUT, D. C. Lesioned attractor networks as models of neuropsychological deficits. In: ARBIB, M. (Ed.). *The handbook of brain theory and neural networks*. Cambridge, MA: MIT Press, 1995. p. 540–543.

POPOV, N. A. *Études de psychologie*. Paris: Ed. du Cèdre, 1950.

PORT, R. Linguistic timing factors in combination. *J. Acoust. Soc. Am.*, v. 69, p. 262–274, 1981.

PORT, R. Invariance in phonetics. In: PERKELL, J.; KLATT, D. (Ed.). *Invariance and Variability in Speech Processes*. Hillsdale, New Jersey: Erlbaum Associates, 1986.

PORT, R. Representation and recognition of temporal patterns. *Connection Science*, v. 2, p. 151–176, 1990.

PORT, R. The dynamical systems hypothesis in cognitive science. In: *The MacMillan Encyclopedia of Cognitive Science*. [S.l.]: L. Nadel, Associate Editor, 2002. v. 1, p. 1027–1032.

PORT, R.; CUMMINS, F. The English voicing contrast as velocity perturbation. In: PRINTING, E. P. (Ed.). *Proceedings of the 1992 International Conference on Spoken Language Processing (ICSLP '92)*. Banff, Alberta, Canada: [s.n.], 1992. v. 2, p. 1311–1314.

PORT, R.; CUMMINS, F.; GASSER, M. A dynamic approach to rhythm in language: Toward a temporal phonology. In: LUKA, B.; NEEDS, B. (Ed.). *Proceedings of the Chicago Linguistics Society*. Chicago, Estados Unidos: [s.n.], 1995. p. 375–397.

PORT, R.; GELDER, T. van (Ed.). *Mind as motion. Explorations in the dynamics of cognition*. Cambridge, Estados Unidos: The MIT Press, 1995.

PORT, R.; O'DELL, M. L. Neutralization of syllable final voicing in german. In: *Research in Phonetics, Department of Linguistics*. [S.l.]: Indiana University, 1984. v. 4, p. 93–133.

QUEDNAU, L. R. A síncope e seus efeitos em latim e em português arcaico.

In: BRESCANCINI, L. B. e C. (Ed.). *Fonologia e Variação: recortes do português brasileiro*. Porto Alegre: EdIPUCRS, 2002.

RHARDISSE, N.; ABRY, C. Mandible as syllable organizer. In: *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Estocolmo, Suécia: [s.n.], 1995. v. 3, p. 556–559.

SALTZMAN, E.; BYRD, D. Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, v. 19, p. 499–526, 2000.

SALTZMAN, E.; KELSO, J. A. S. Skilled actions: a task-dynamic approach. *Psychological Review*, v. 94, n. 1, p. 84–106, 1987.

SALTZMAN, E.; LÖFQVIST, A.; KAY, B.; KINSELLA-SHAW, J.; RUBIN, P. Dynamics of intergestural timing: a perturbation study of lip-larynx coordination. *Experimental Brain Research*, n. 123, p. 412–424, 1998.

SALTZMAN, E.; MUNHALL, K. G. A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, v. 1, n. 4, p. 333–382, 1989.

SANSAVINI, A.; BERTONCINI, J.; GIOVANELLI, G. Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, v. 33, p. 3–11, 1997.

SAUSSURE, F. *Curso de Lingüística Geral*. São Paulo: Cultrix, 1916.

SCOTT, S. K. *Perceptual centres in speech: an acoustic analysis*. Tese (Doutorado) — University College London, 1993.

SHADMEHR, R. Equilibrium point hypothesis. In: ARBIB, M. (Ed.). *The handbook of brain theory and neural networks*. Cambridge, MA: MIT Press, 1995. p. 370–372.

SHIK, M.; ORLOVSKII, G. N.; SEVERIN, F. V. Organization of locomotor synergism. *Biophysics*, v. 13, p. 127–135, 1966.

SILVA-NETO, S. *Fontes do latim vulgar (o Appendix Probi)*. Rio de Janeiro: Imprensa Nacional, 1946.

STETSON, R. H. *R. H. Stetson's motor phonetics: a retrospective edition*. Boston, Estados Unidos: College-Hill Press, 1988. Reimpressão. KELSO, J. A. S.; MUNHALL, K. G. (Ed.). Publicado originalmente: Stetson, R. H. [1951] *Motor Phonetics*. 2ª edição.

STONE, M. Evidence for a rhythm pattern in speech production: observations of jaw movement. *Journal of Phonetics*, v. 9, p. 109–120, 1981.

SUSSMAN, H. M.; MACNEILAGE, P. F. Motor unit correlates of stress: Preliminary observations. *J. Acoust. Soc. Am.*, n. 64, p. 338–340, 1978.

TARALLO, F. *Tempos lingüísticos: itinerário histórico da língua portuguesa*. São Paulo: Ática, 1990.

TILLMANN, H. G. Phonetics, early modern, especially instrumental and experimental work. In: ASHER, R. E.; SIMPSON, J. M. Y. (Ed.). *The Encyclopædia of Language and Linguistics*. Oxford, Reino Unido: Pergamon Press, 1994. p. 3082–3095.

TREFFNER, P. J.; TURVEY, M. T. Resonance constraints on rhythmic movement. *Journal of Experimental Psychology: Human Perception and Performance*, v. 19, n. 6, p. 1221–1237, 1993.

TULLER, B.; HARRIS, K. S.; KELSO, J. A. Stress and rate: differential transformations of articulation. *J. Acoust. Soc. Am.*, v. 71, p. 1534–1543, 1982.

TULLER, B.; KELSO, J. A. S. The timing of articulatory gestures: Evidence for relational invariants. *J. Acoust. Soc. Am.*, v. 76, p. 1030–1036, 1984.

TULLER, B.; KELSO, J. A. S. Phase transitions in speech production and their perceptual consequences. In: JEANNEROD, M. (Ed.). *Attention and Performance XIII*. Hillsdale, Estados Unidos: Erlbaum, 1990. p. 429–452.

TULLER, B.; KELSO, J. A. S. The production and perception of syllable structure. *Journal of Speech and Hearing Research*, v. 34, p. 501–508, 1991.

TURVEY, M. T. Preliminaries to a theory of action with reference to vision. In: SHAW, R.; BRANSFORD, J. (Ed.). *Perceiving, acting, and knowing: toward an ecological psychology*. Hillsdale, Estados Unidos: Lawrence Erlbaum Associates, 1977. p. 211–265.

TURVEY, M. T.; SHAW, R. E.; MACE, W. Issues in the theory of action: Degrees of freedom, coordinative structures, and coalitions. In: REQUIN, J. (Ed.). *Attention and Performance (VII)*. Hillsdale, Estados Unidos: Erlbaum, 1978.

VASCONCELOS, C. M. de. *Lições de filologia portuguesa*. Lisboa: Nova Edição da Revista de Portugal - série A - Língua Portuguesa, 1956.

VERBRUGGE, R. R.; SHANKWEILER, D. P. Prosodic information for vowel identity. *J. Acoust. Soc. Am.*, S39, 1977.

VERHULST, F. *Nonlinear Differential Equations and Dynamical Systems*. Berlim: Springer-Verlag, 2006.

VIVIANI, P.; TERZUOLO, V. Space-time invariance in learned motor skills. In: STELMACH, G. E.; REQUIN, J. (Ed.). *Tutorials in motor behavior*. North Holland: Amsterdam, 1980.

WALLIN, J. E. W. Experimental studies of rhythm and time. *Psychol. Rev.*, v. 18, p. 100–131 e 202–222, 1911.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. In: LEHMANN, W.; MALKIEL, Y. (Ed.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968.

WIGHTMAN, C. W.; SHATTUCK-HUFNAGEL, S.; OSTENDORF, M.; PRICE, P. J. Segmental durations in the vicinity of prosodic boundaries. *J. Acoust. Soc. Am.*, v. 91, n. 3, p. 1707–1717, 1992.

WILLIAMS, E. *Do latim ao português: fonologia e morfologia históricas da língua portuguesa*. Rio de Janeiro: Instituto Nacional do Livro, 1961.
Traduzido por Antônio Houais.

Índice Remissivo

- ângulo de fase, 28, 57
- abordagem dinamicista, 3, 50, 295
- acento frasal, 69, 70, 72, 79, 232
- acento lexical, 72
- acentuação dos gestos, 46
- acoplamento dos ritmos motores, 96
- acoplamento gestual, 30
- alvo funcional da tarefa, 9
- amplitude do ciclo do oscilador acen-
tual ($i(m)$), 71
- análise rítmica, 51, 69
- aspectos contínuos, 294
- aspectos dinâmicos das proparoxíto-
nas, 265
- ativação gestual, 62
- atividade coordenada, 22
- atividade rítmica, 90
- atrator, 7, 52, 54
- atrator cíclico, 77
- atratores (ou tarefas fonológicas), 36
- atratores de fase, 57
- atratores pontuais, 36, 37
- aumento da TE, 107
- aumento exponencial das unidades VV,
81
- auto-organização, 3
- Beat Extractor, 55
- bifurcação, 7, 29
- camadas rítmicas, 46, 47
- características indissociáveis do ritmo,
98
- categorias fônicas, 67

- ciência lingüística, 58
- ciências cognitivas, 59
- cinemática, 195
- coarticulação, 292
- cognição, 14
- coletivo dinâmico de alta ordem, 62
- complexo dinâmico, 48
- conjuntos de VT, 34
- constelações gestuais, 39
- conteúdo e forma, 5
- continuidade vocálica, 75
- contorno duracional, 106, 114
- contraste lexical, 33
- controle motor da fala, 31, 49
- cooperação, 10
- densidade gestual, 85
- derivada, 177, 179
- derivada discreta, 114
- descritores gestuais, 42
- deslocamento da constrição, 196, 203, 217, 232
- deslocamento da constrição da mandíbula, 186
- deslocamento da mandíbula, 186
- deslocamento do gesto de constrição, 201
- deslocamento mandibular, 165
- deslocamento vertical máximo, 196, 201, 203, 214
- diacronia, 236
- dicotomia, 238, 241, 294
- dinâmica articulatória, 63
- dinâmica da fala, 65
- dinâmica subjacente invariante, 38
- DP (duração da palavra), 289
- duração, 69
- duração (relacionada à acústica) da mandíbula, 186
- duração articulatória, 192, 201, 203
- duração AT, 194, 199, 225
- duração da aceleração, 199, 201, 210,

- 232
- duração da aceleração da mandíbula, 187
- duração do gesto mandíbula (AT), 186
- duração dos gestos articulatórios, 232
- duração extrínseca, 69
- duração intrínseca, 69
- duração média dos grupos acentuais, 95
- efeito coletivo (ou cooperativo), 6
- efeito espacial, 79
- efeito temporal, 79
- elementos lingüísticos, 294
- emergência, 5, 295
- EMMA, 166
- equação de diferenças finitas, 70
- espaço, 46
- esquema híbrido, 68
- estabilidade articulatória, 26
- estado de equilíbrio, 3
- estrutura lexical, 234
- estrutura rítmica, 78, 105
- estrutura temporal da fala, 48, 50
- estruturacão rítmica, 86, 93, 105
- estruturas coordenadas, 24
- estruturas dinâmicas, 49
- estruturas gramaticais, 298
- FAR (Fonologia Articulatória), 31
- fase, 25
- fase relativa, 13
- fatores extralingüísticos, 105
- fatores fonéticos neogramáticos, 277
- fenômenos gramaticais, 294
- fenômenos lingüísticos, 298
- figura e fundo, 75
- flutuações críticas, 8
- fluxos vocálicos, 73
- fonética lingüística, 77
- Fonologia Articulatória, 31, 296
- Fonologia Temporal, 47

- fontes cinemáticas de reestruturações
 rítmicas, 201
 força de acoplamento (w_o), 69, 71, 72
 forma de constrição, 42
 frações harmônicas, 54
 fronteiras fracas, 116
 fronteiras frasais, 62–64, 165
 fronteiras prosódicas, 62, 65
 fronteiras prosódicas fortes, 115, 116
 função de sincronismo ($s(n)$), 71
 gesto articulatório, 33, 242
 gesto lingüístico, 32
 gesto prosódico (gesto π), 60–63
 gestos articulatórios, 32, 33, 48, 61, 64, 232, 242, 296
 gestos de constrição, 63
 gestos primitivos, 94
 gradientes, 67
 grau de acentuação, 46
 grau de acoplamento entre sintaxe e produção (r_p), 69, 104
 grau de ativação, 64
 grau de constrição gestual, 34, 42, 46
 grau de magnitude gestual, 269
 graus de liberdade, 11, 25, 69
 heterogeneidade lingüística, 237
 hierarquia métrica, 72
 histerese, 12, 29
 IC (intervalo consonantal), 289
 ICP (intervalo consonantal proporcional), 289
 implementação do MDR, 67
 implicações teóricas, 58
 início acústico da sílaba (*vide* unidade VV), 58
 incomensurabilidade, 67
 informação sintática, 72
 instabilidade, 12, 30
 interação prosódia-segmentos, 69, 72,

- intervalos rítmicos, 51 66–68, 78
- invariância lingüística, 11 MDT (modelo dinâmico da tarefa),
- invariâncias perceptivas, 20 33
- invariantes relacionais, 20 mecanismo acoplador, 93
- isocronia, 51, 58 mecanismo regularizador, 93
- isocronia acentual, 117 medidas articulatórias, 187
- léxico gestual, 69 metrônomo, 51, 96, 108
- língua, 16, 48, 298 modelo dinâmico da tarefa, 33
- limites do número de elementos cons- modificação fonética, 295
- tituintes do ritmo, 98 modos estáveis, 30
- limites temporais do ritmo, 98 modulação da rigidez da mola, 62
- local de constrição, 34, 42 movimentação mandibular, 203
- métrica, 50 movimento articulatorio, 172
- métrica dinâmica, 72, 80 movimento oscilatório da mandíbula,
- magnetômetro (ou articulógrafo), 166, 72
- 167 mudança de fase, 12, 13
- magnitude dos gestos, 241 mudança lingüística, 236, 238, 239,
- magnitude espacial, 65 245
- MAVIS, 169–171, 186 mudanças sonoras, 278
- MDR (Modelo Dinâmico do Ritmo), níveis funcionais, 34

- nível de ativação, 36, 63
- nível de constrição, 34
- ocultamento gestual (*gestural "hiding"*),
43
- organização temporal (*timing*), 19, 65
- oscilador, 54, 67
- oscilador acentual, 68
- oscilador adaptativo, 58
- oscilador silábico, 69, 73, 77
- osciladores acoplados, 58, 67, 70, 102
- padrão cíclico, 78
- padrões articulatórios, 24, 195
- padrões dinâmicos, 4, 28
- padrões duracionais, 106
- padrões lingüísticos, 46
- padrões rítmicos, 57
- parâmetro de controle, 6, 25
- parâmetro de ordem (ou variável co-
letiva), 6
- parâmetros subjacentes, 107
- pauta gestual abstrata, 63, 72
- período do oscilador silábico atual (T),
71
- período do oscilador silábico desaco-
plado (T_0), 71
- períodos interacentuais, 52
- percepção da fala, 49
- percepção rítmica, 99
- periodicidade, 73, 78
- perturbador, 10, 24, 28, 63, 86, 99,
107, 295
- picos de velocidade, 181
- plano de fase, 22, 27
- processos de ritmização, 97, 99
- produção da fala, 49
- programa de pesquisa, 45
- propriedades do ritmo, 96
- propriedades universais da estrutura
temporal da fala, 50
- prosódia, 65, 96
- pulso silábico, 55

- pulsos do oscilador acentual, 79
- ramo prosódico do MDR, 68
- ramo segmental do MDR, 68
- ramos lingüísticos mais elevados, 68
- reestruturação silábica, 43, 44, 242
- reestruturações lexicais, 234
- reestruturações rítmicas, 99, 103, 106,
107, 200, 203, 204, 206–208
- regiões instáveis, 8
- regularidade acentual, 51
- regularidade rítmica, 52
- relógio articulatório, 38
- relógio biológico, 93
- relógio interno, 38, 62, 64, 93
- relação bijetora, 67
- relações de fase, 23, 29, 36
- relações harmônicas, 54
- relações hierárquicas, 80
- relações rítmicas, 54
- reorganização articulatória, 10
- reorganização rítmica, 99, 106, 295,
296
- reorganizações microrrítmicas, 234
- representação simbólica, 62
- rigidez da mola, 37, 46, 63, 200, 201,
232
- ritmização objetiva, 97
- ritmização subjetiva, 97, 98
- ritmo, 50, 51, 66, 89, 90, 98
- ritmo biológico, 90
- ritmo circadiano, 92
- ritmo da fala, 69, 78, 86
- ritmo desencadeado, 91
- ritmo espontâneo, 91, 93, 95
- ritmo induzido, 91
- ritmo lento, 92
- ritmo lingüístico, 50, 71
- ritmo rápido, 92
- sílaba CV, 76
- sílabas átonas, 101, 102

- sílabas tônicas, 101, 102
 símbolos, 5, 16, 48
 síncope, 239
 sequência rítmica, 99
 simulações articulatórias, 65
 sincronia, 51, 99
 sincronização rítmica, 96
 sinérgica, 3
 sinergias, 6, 9, 67
 sistema de controle motor, 54
 sistema dinâmico abstrato, 49
 sistema dinâmico subjacente, 34
 sistema massa-mola amortecido, 36
 sistemas abertos, 3
 sistemas cognitivos, 37
 sistemas dinâmicos, 2, 3, 6, 13, 46, 48, 86, 238, 243, 298
 sobreposição articulatória, 297
 sobreposição gestual, 43, 269, 281
 Sociolingüística, 243
 tarefa cíclica da fala (*speech cycling task*), 51
 tarefas metodológicas, 109
 taxa crítica, 12, 29
 taxa de decaimento (β), 71
 taxa de elocução (*speech rate*), 6, 83, 99, 165, 232, 239, 295
 taxa de indução (α), 71
 taxa de mudança, 179
 taxa do relógio interno, 65
 TCF (tarefa cíclica da tarefa), 51
 TE (taxa de elocução), 6, 105, 107, 245
 tempo, 46, 48, 66, 298
 tempo-para-velocidade-de-pico proporcional, 200, 201, 203, 228
 tempo-para-velocidade-de-pico proporcional da mandíbula, 187
 tese de comensurabilidade, 16
 teste de máximos, 112, 114
 tipologias rítmicas, 59

- tipos de fronteira frasal, 65
- trajetórias cinemáticas, 296
- transição CV, 76
- transição de fase, 12, 29, 30
- TSD (teoria dos sistemas dinâmicos),
 - 15, 238
- unidade de pulsação, 58
- unidade VV, 67, 73
- unidade VV abstrata, 82, 84
- unidades de ação funcionais (ou estruturas coordenadas), 25
- unidades fonológicas, 32, 61
- unidades lingüísticas, 62
- unidades primitivas, 33
- universais lingüísticos, 50
- vales de velocidade, 181
- variáveis coletivas, 11
- variáveis do trato vocal, 34
- variação da TE, 28, 86, 103, 104, 106,
 - 295
- variação lingüística, 11, 238, 245, 277
- variações contextuais, 38
- variabilidade contextual, 34
- variabilidade intrínseca, 85
- velocidade do gesto, 172
- velocidade instantânea do gesto, 179
- velocidades de pico e vale, 199, 201,
 - 203, 222
- velocidades de pico e vale da mandíbula, 186
- VT (variável(is) do trato vocal), 34
- VV articulatório, 194
- VV articulatório reestruturado, 208
- VV articulatório tradicional (AT), 194
- z-score, 110
- z-score suavizado, 110
- zeros da velocidade y, 177