

Adelaide H. P. Silva

**As fronteiras entre Fonética e Fonologia e a alofonia dos róticos
iniciais em PB: dados de dois informantes do sul do país**

Tese apresentada ao Curso de
Lingüística do Instituto de Estudos da
Linguagem da Universidade Estadual
de Campinas como requisito parcial
para a obtenção do título de Doutor
em Lingüística.

Orientadora:

Profa. Dra. Eleonora Cavalcante Albano

UNICAMP

Instituto de Estudos da Linguagem

2002

UNIDADE Bc
 CHAMADA T/UNICAMP
SI 38f
 EX
 BANCO BCI 50432
 PROC 16-837/02
 DI
 PREÇO R\$ 11,00
 DATA 20/08/02
 CPD

CM00172177-1

3 ID 252310

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA IEL - UNICAMP

Silva, Adelaide Hercília Pescatori

SI 38f
 As fronteiras entre Fonética e Fonologia e a alofonia dos róticos iniciais em PB / Adelaide Hercília Pescatori Silva. - - Campinas, SP: [s.n.], 2002.

Orientador: Eleonora Cavalcante Albano

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Dinâmica. 2. Produção de fala. I. Albano, Eleonora Cavalcante. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Este exemplar e a redação final da tese
defendida por Adelaide Azeite

Pescatori Sílvia

e aprovada pela Comissão Julgadora em
04/05/02.

Eleonora Albano

Banca Examinadora:

Profa. Dra. Eleonora Cavalcante Albano – Orientadora

Profa. Dra. Iara Bemquerer Costa

Profa. Dra. Mirian da Matta Machado

Profa. Dra. Maria Bernadete Marques Abaurre

Prof. Dr. Wilmar da Rocha D'Angelis

Suplentes:

Dra. Ana Luíza Pinto Navas

Prof. Dr. Plínio de Almeida Barbosa

*Para o Gabriel,
Dos olhos vivos e alegres,
Do riso maroto,
Das mãozinhas de doce afago,
Da voz terna que chama “mamãe”.*

*Para quatro grandes e velhos amigos:
Aglael Gama-Rossi, Agnaldo Moreira,
Patrícia Aquino & Rosiney Casadó.
Tão longe, tão perto...*

AGRADECIMENTOS

À profa. dra. Eleonora Cavalcante Albano, pela orientação precisa, pela leitura atenta dos meus textos, pelas críticas construtivas, nos momentos adequados. Agradeço-lhe também toda a minha formação na Fonética: a profa. Eleonora despertou-me o interesse pela área e acompanhou-me, desde muito cedo, nos primeiros anos da graduação, mostrando-me os caminhos de um campo ao mesmo tempo tortuoso e apaixonante. Cabe-lhe também um agradecimento especial por ter-me oferecido seu ombro amigo e seus conselhos em algumas horas muito difíceis.

Aos colegas todos que passaram pelo LAFAPE, nestes dez anos do laboratório, cujos nomes prefiro não citar, para não incorrer em eventual esquecimento: pelo companheirismo e por levarem a sério o sentido do trabalho em equipe.

Ao Max Guimarães, pelas intermináveis discussões lingüísticas e por ter-me auxiliado a desenvolver o *corpus* do experimento fonético-acústico da tese.

Ao padre Orestes Gomes Filho e aos colegas de departamento da UFPR, profs. Drs. Antônio José Sandmann e José Luís da Veiga Mercer, por terem aceitado servir como sujeitos para o experimento fonético-acústico.

Ao prof. dr. Edson Françoço, pelo convívio no LAFAPE e pelas sugestões no exame de qualificação em Psicolingüística.

Ao prof. dr. Plínio Barbosa, por ajudar-nos a compreender um pouco mais sobre o ritmo do PB e também pelo trabalho conjunto na elaboração do Aiuruetê, o sintetizador de fala desenvolvido no LAFAPE.

Ao prof. dr. Wilmar da Rocha d'Angelis, pelas sugestões várias ao texto desta tese, na ocasião do exame de qualificação, e pelas discussões que nos unem em torno de um mesmo problema: como representar adequadamente o modo de articulação, num modelo fonológico.

À dra. Ana Luíza Navas, pelas valiosas sugestões ao texto da tese, no exame de qualificação.

À dra. Aglael Gama-Rossi: pela parceria em vários trabalhos; por introduzir-me e guiar-me por questões audiológicas; pela leitura atenta dos originais desta tese e pelas suas pertinentes sugestões; pelo companheirismo e pelo carinho que me fizeram levantar num

momento em que eu já não tinha forças para conseguir me levantar sozinha. Enfim, por dividir comigo trabalhos, anseios e angústias.

Aos funcionários da Secretaria de Pós-Graduação do IEL: pela presteza em cuidar de todas as questões burocráticas.

Aos funcionários da biblioteca do IEL: pelos auxílios precisos e pela gentileza com que sempre prestaram esses auxílios, desde meus primeiros anos de graduação na UNICAMP.

Aos colegas do Departamento de Linguística, Letras Clássicas e Vernáculas da Universidade Federal do Paraná: pela maneira calorosa com que me receberam no Departamento e por me terem concedido dois anos e meio de afastamento para que eu terminasse a tese.

Dentre os colegas, agradeço especialmente aos profs. Anamaria Filizola, Iara Bemquerer Costa, José Borges Neto, Lígia Negri, Maria José Foltran, Marilene Weinhardt, Reny Gregolin e Teresa Wachowicz, pela amizade e pelo apoio – acadêmico e pessoal. Um agradecimento particular cabe ao prof. dr. João Alfredo Dal Bello, “agregado” do departamento e compadre: pela amizade e pelas palavras certas nas horas certas.

Às secretárias do DELIN/UFPR, Delma Maskow e Sirlei Ferraz da Anunciação, pelo auxílio para resolver questões burocráticas.

Aos meus alunos, que também me ensinam como ensinar.

À CAPES, pelo financiamento de parte do meu doutorado, através de uma bolsa PICDT.

Aos profs. drs. Iara Bemquerer Costa, Mirian da Matta Machado, Maria Bernadete Marques Abaurre, Wilmar da Rocha d’Angelis, banca examinadora de minha tese, pelas discussões enriquecedoras e pelas sugestões pertinentes, durante a defesa. Muitas dessas sugestões foram incorporadas a este texto, que é a versão final de minha tese de doutorado.

Fora da academia, cabe um agradecimento muito especial a meus pais, Décio e Maria José: pelo apoio constante e incondicional. E por terem deixado sua casa, por um longo período, para me auxiliarem nos cuidados com o Gabriel. Devo a conclusão desta tese, em parte, a essa valiosa ajuda.

A meu pai, Décio, agradeço também a cuidadosa revisão da versão final desta tese.

À Sílvia Lazarini, ao Ernesto, ao Lucas e à Olívia: pela velha amizade e pelo carinho de sempre.

À Vera Lúcia de Oliveira e Silva, por me ajudar a enxergar a terceira margem de um rio caudaloso e, por vezes, de difícil navegação.

Ao pessoal da Escola de Educação Infantil “Peixinho Dourado”, particularmente à diretora Márcia Canova e às professoras Maria Inês, Giselle, Carla, Luciana, Ângela, Vanessa e Cristiane, pela competência, pela dedicação e pelo carinho dispensados ao Gabriel. Isso tudo foi fundamental para que eu pudesse concluir este trabalho.

RESUMO

Esta tese discute uma questão em voga na literatura fonética atual, concernente à comensurabilidade de fatos presentes na fala: sabe-se que algumas alofonias tradicionalmente tidas como categóricas são, na verdade, gradientes, contínuas (para o inglês, vide Browman & Goldstein, 1990; Sproat & Fujimura, 1993; Huffman, 1997; Gick, 1999; para o português, vide Albano, Barbosa, Gama-Rossi, Madureira & Silva, 1998; Albano, 2001; Silva & Albano, 1999; ou o terceiro capítulo desta tese). Apesar de gradientes, tais alofonias não são aleatórias, mas condicionadas pela estrutura prosódica do enunciado no qual ocorrem. A estrutura prosódica, por sua vez e de alguma forma, mapeia informações sobre a estrutura sintática do enunciado. Assim sendo, as alofonias necessitam ser modeladas na gramática de uma língua.

Surge daí a questão: em que nível da gramática colocar esses fatos, na Fonética ou na Fonologia? Colocá-los num nível fonético, como concebido tradicionalmente, seria insuficiente para capturar a relação entre os processos alofônicos e o nível prosódico que os condiciona. Colocá-los num nível fonológico, por outro lado, não permitiria captar a natureza gradiente dos processos, já que os modelos fonológicos têm por primitivos unidades categóricas, como argumento no segundo capítulo.

A saída que se afigura viável é o tratamento desses processos à luz de modelos dinâmicos de produção de fala, como a Fonologia Articulatória (Browman & Godstein, 1986, 1990, 1992), porque a mesma se baseia numa unidade de tempo intrínseco, o gesto articulatório, o que torna direta a relação entre representação e implementação. Por conta dessa relação direta, nesse modelo, o nível fonético não está dissociado do fonológico; ao contrário, eles se fundem num único.

Assim, proponho, no quarto capítulo, uma representação dos róticos, considerando os gestos articulatórios que os constituem. Porém, diferentemente da Fonologia Articulatória, assumo, baseada em Sproat & Fujimura (1993) e Gick (1999) que um segmento pode ser constituído por mais de um gesto. E, à luz da Fonologia Acústico-Articulatória (Albano 2001) proponho uma representação para os róticos de início de palavra que considera que os gestos não se definem por conjuntos de articuladores, mas por regiões acústico-articulatórias e que as variantes dialetais são todas lexicalizadas, escolhendo o falante a variante mais adequada a um dado contexto prosódico.

SUMMARY

Recent phonetic literature shows that some allophonic variation processes, usually considered to be categorical are gradient, indeed (see Browman & Goldstein, 1986, 1989, 1990, 1992; Sproat & Fujimura, 1993; Gick, 1999, for English, and Albano *et alii*, 1998; Albano, 2001; Silva & Albano, 1999, for Portuguese and also chapter 2 of the thesis). Although these allophonic variations are gradient, they are not at random. On the contrary, they are conditioned by prosodic structure. Prosodic structure, on its turn, carries information about the syntactic organization of an utterance. So, allophonic variation processes of the kind referred to above, must be represented in the grammar of a specific language.

This raises the question about where in the grammar to put allophonic facts: inside Phonetics or inside Phonology? Treating those facts in the light of Phonetics, as it is traditionally conceived, would not capture the relationship between the gradient allophony and the prosodic level. On the other hand, treating gradient allophony inside Phonology would not capture the gradient nature of the allophonic variation processes, because phonological models have categorical units as their primitives.

So, the only way to model gradient allophony inside of a grammar is to treat them in the light of dynamic models of speech production, as Articulatory Phonology (Browman & Goldstein, 1986, 1989, 1990, 1992), since this model takes as primitives the articulatory gestures, that have intrinsic time. This turns direct the phonetic-phonology relationship.

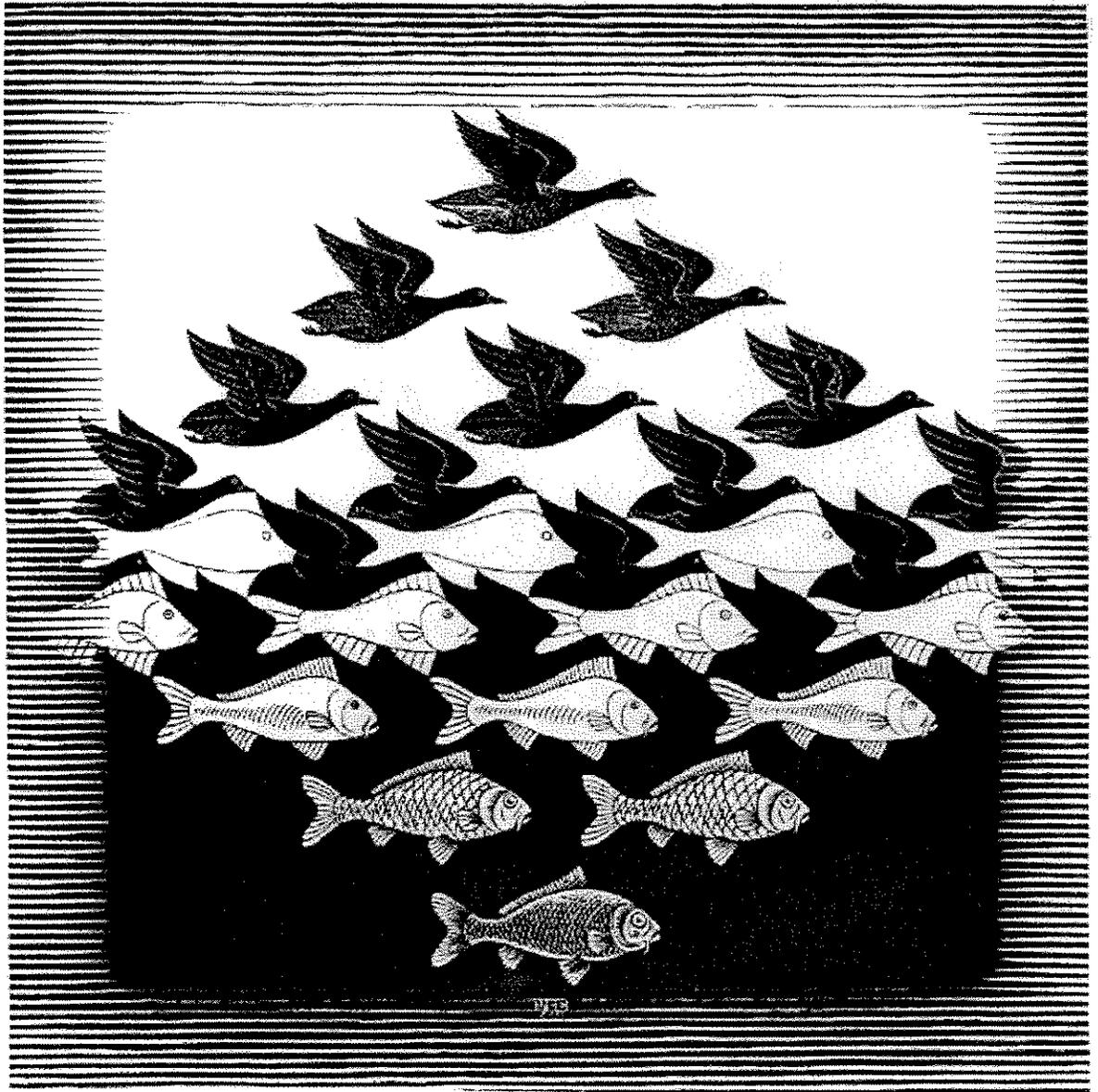
My proposal, then, is to represent rhotics in initial position in PB considering the articulatory gestures that constitute them. But, following Sproat & Fujimura (1993) and Gick (1999), I assume a segment to be constituted by more than one single gesture. And, in the light of Acoustic-Articulatory Phonology (Albano 2001) I say that articulatory gestures are not defined by sets of articulators, but by acoustic-articulatory zones and that dialectal variants are all put in the lexicon. The speaker, then, chooses which allophone is more adequate to a certain prosodic context.

SUMÁRIO

RESUMO	11
SUMÁRIO	13
INTRODUÇÃO	21
1 – O GESTO ARTICULATÓRIO COMO UNIDADE CONTRASTIVA	27
1.1 – O início do problema: a definição do objeto de estudo da Fonética e da Fonologia	29
1.1.1 – O fonema como unidade mínima de análise	29
1.1.2 – Os traços distintivos: unidades mínimas constitutivas do fonema	31
1.2 – A dissociação entre Fonética e Fonologia, numa outra perspectiva	33
1.2.1 – A Fonologia Gerativa	33
1.2.2 – A Geometria de Traços	37
1.3 – Uma nova unidade de análise: o gesto articulatório	38
1.3.1 – A Fonologia Articulatória	38
1.3.1.1 – Evidências para o estabelecimento do gesto articulatório como primitivo teórico e sua formalização inicial	38
1.3.1.2 – O gesto como unidade discreta e organizadora da Fonologia	50
1.3.1.3 – O gesto como unidade modeladora de alofonias gradientes	61
1.4 – A Fonologia Acústico-Articulatória	66
1.5 – Considerações finais	75
2 – OS LIMITES DE UMA REPRESENTAÇÃO DAS LÍQUIDAS, NA TEORIA DOS TRAÇOS DISTINTIVOS: DE <i>FUNDAMENTALS</i> À GEOMETRIA DE TRAÇOS	77
2.1 – Questões inerentes à representação fonológica das líquidas	81
2.2 – É sustentável a proposta de uma classe natural “líquidas”?	82
2.3 – Como distinguir laterais de róticos numa mesma representação?	87
2.4 – Como representar a classe “líquidas”?	90
2.5 – Como distinguir laterais de róticos numa representação fonológica?	94

	16
2.5.1 – A representação das laterais em <i>Fundamentals of Language</i> e SPE	94
2.5.1.1 – Proposta para a representação da oposição /r/~l/ via [contínuo]	96
2.5.2 – A representação de segmentos laterais nas geometrias de traços	99
2.5.3 – Como representar a oposição /r/~r/?	102
2.5.3.1 – Oposição /r/~r/ estabelecida entre um segmento geminado e um não-geminado	102
2.5.3.2 – Oposição /r/~r/ dada por traços distintivos	109
2.6 – Uma representação das líquidas dentro de uma geometria de traços	112
2.6.1 – Unificação de líquidas e nasais numa classe “soantes”	112
2.6.1.1 – O nó “vozeamento espontâneo” na estrutura interna dos segmentos proposta por Clements & Hume	115
2.6.1.2 – Problemas do nó “vozeamento espontâneo”	119
2.6.1.3 – Outra representação das líquidas numa geometria de traços	121
2.7 – Sumário e discussão	126
3 – A ALOFONIA CONTÍNUA DOS RÓTICOS INICIAIS EM PB: EVIDÊNCIAS PARA UM TRATAMENTO DINÂMICO DOS PROCESSOS FÔNICOS	131
3.1 – Indícios favoráveis à natureza contínua da variabilidade de pronúncia dos róticos	134
3.2 – Contínuo fonético na variabilidade de pronúncia dos róticos: prosodicamente condicionado?	142
3.2.1 – <i>Design</i> experimental	143
3.3 – Análise dos dados	147
3.3.1 – Contínuo físico na variabilidade de pronúncia dos róticos: análises visuais e espectros de FFT nos dados de JLM	148
3.3.1.1 – Alguns indícios de condicionamento prosódico sobre a realização do contínuo físico na variabilidade de pronúncia de /r/ para o informante JLM	152
3.3.1.2 – Quantificação do ruído fricativo em alguns dados de JLM	154
3.3.2 – Contínuo físico na variabilidade de pronúncia dos róticos: análises	158

	17
visuais dos dados de AJS	
3.3.2.1 – Indícios de condicionamento prosódico sobre a realização do contínuo físico na variabilidade de pronúncia de /r/: dados de AJS	162
3.3.3 – Considerações finais sobre os dados	164
4 – PROPOSTA DE REPRESENTAÇÃO DOS RÓTICOS À LUZ DE UM MODELO DINÂMICO DE PRODUÇÃO DA FALA	169
4.1 – Retomando alguns pressupostos da FAR	173
4.2 – Para além da FAR	175
4.2.1 – O modelo “bigestual” de Sproat & Fujimura para as laterais do inglês norte-americano	175
4.2.2 – O tratamento “bigestual” dos <i>glides</i> e líquidas do inglês norte-americano por Gick	177
4.2.3 – O gesto articulatório definido em função de regiões acústico-articulatórias: a contribuição da FAAR	180
4.3 – Proposta de pauta gestual para os róticos do PB	186
CONCLUSÃO	197
REFERÊNCIAS BIBLIOGRÁFICAS	201
ANEXO	207



M. C. Escher (1938) *Céu e água*, in Taschen Diary, 1997.

Introdução

O desenho de Escher, escolhido como “epígrafe” desta tese, traduz, em formato gráfico, um fato que se verifica também na fala e que será empregado como a evidência empírica, para sustentar toda a argumentação teórica que desenvolvo. Trata-se, em ambos os casos – o do desenho e o do fato fônico – da existência de um *continuum* entre duas categorias.

No desenho, reconhecem-se claramente duas “categorias animais”: a dos peixes e a das aves. Há, porém, um momento em que essas duas categorias se confundem, quando as características dos peixes se sobrepõem às das aves, e vice-versa.

A exemplo desse desenho, também na variabilidade de pronúncia dos róticos (sons de /r/) iniciais¹ de palavra em português brasileiro (doravante PB), é possível verificar um contínuo físico ao longo do qual ela se realiza: conforme mostro no terceiro capítulo, através de experimento acústico é possível reconhecer na fala de um informante (JLM) duas categorias que se alternam na posição inicial: a categoria das vibrantes apicais e a categoria das fricativas velares. Há, no entanto, sons intermediários a essas duas categorias, nos quais o ruído fricativo se sobrepõe, em maior ou menor grau, à estrutura acústica da vibrante. Na fala de um segundo informante (AJS), observa-se a coexistência da vibrante apical, do *tap* e também de sons intermediários a essas duas categorias, os quais são enfraquecidos, quando comparados a uma vibrante apical canônica, mas que são mais fortes que o *tap*. Como noto no terceiro capítulo, esses sons intermediários a vibrantes e *taps*, quando ouvidos isoladamente, i.e., sem comparação com vibrantes e *taps* canônicos, dão-nos a impressão de serem um *tap*. Porém, a comparação auditiva evidencia a diferença entre os sons que, em termos acústicos, se reflete na maior duração dos sons intermediários e, geralmente, na presença de dois fechamentos – o *tap* apresenta apenas um. A

¹ Deixo já clara, de antemão, minha opção por enfocar, nesta tese, os róticos iniciais: como se sabe, segmentos em posição inicial de sílaba/palavra, tendem a ser fortalecidos, i.e., produzidos com maior força articulatória (vide, e.g., Fougeron, 1999). Segmentos de final de sílaba/palavra, por outro lado, tendem a ser produzidos com menor força articulatória e são, por isso, enfraquecidos. Em decorrência, esperar-se-ia pouca ou nenhuma variabilidade de pronúncia de um segmento em posição inicial. Encontrar tal variabilidade e, mais, observá-la condicionada por fatos prosódicos (vide terceiro capítulo), é relevante para a gramática fônica de uma língua, que deverá prever a representação de processos de enfraquecimento – neste caso, condicionados prosodicamente – também em início de palavra.

comparação desses sons intermediários com uma vibrante apical canônica, por outro lado, mostra que a vibrante é mais longa e apresenta mais aberturas orais e fechamentos do que esses sons intermediários.

Observe-se que não se trata de afirmar apenas a existência de categorias intermediárias a vibrantes apicais e *taps* ou a vibrantes apicais e fricativas: o estudo acústico mostra não só a existência dessas categorias intermediárias mas também, e principalmente, um contínuo físico entre elas. Nesse sentido, a alofonia dos róticos iniciais em PB mostra-se gradiente, e não categórica, como podem fazer pensar estudos fonológicos baseados numa investigação impressionística.

Este exemplo soma-se a outros, investigados no Laboratório de Fonética e Psicolinguística (LAFAPE), do Instituto de Estudos da Linguagem, na UNICAMP, para o português brasileiro (PB) e discutidos, e.g., por Albano, Barbosa, Gama-Rossi, Madureira & Silva (1998) ou Albano (2001). Soma-se também a exemplos de alofonias contínuas para outras línguas, como o inglês, e abordados por Keating (1985); Sproat & Fujimura (1993); Browman & Goldstein (1990) ou Gick (1999).

Como os fatos investigados pelos estudos supracitados, a alofonia contínua dos róticos iniciais em PB mostra-se condicionada pela estrutura prosódica dos enunciados onde /r/ ocorre: manipulando a força de fronteira antecedente aos róticos, o experimento acústico, descrito no terceiro capítulo, mostra – especialmente para os dados de AJS – que variantes fortalecidas de /r/ - i.e., aquelas que tendem a vibrantes – ocorrem adjacentes a fronteiras prosódicas fortes. Contrariamente, variantes enfraquecidas desse segmento, ou seja, tendendo a *taps*, ocorrem adjacentes a fronteiras prosódicas fracas. (Frise-se que a denominação variantes “fortes” e variantes “fracas” é empregada, nesta tese, no sentido, por exemplo, de Ohala & Kawasaki (1984), Vaissière (1986, a 1988), Fujimura (1990) ou Keating (1995), citados por Fougeron (1999). Ou seja, segmentos fortalecidos são aqueles que envolvem “maior força articulatória”, “maior energia” ou “maior tensão”.)

Dada essa característica das alofonias contínuas, faz-se necessário representá-las no léxico, de modo que parâmetros fonéticos quantitativos sejam controlados por parâmetros gramaticais qualitativos. Surge aí a questão central desta tese, concernente à maneira como se faz esse modelamento. Ou, dito de outro modo, como se “traduz” um fato gradiente, numérico, numa representação categórica, simbólica. Em última instância, trata-se mesmo

de questionar a dissociação estabelecida, desde o Estruturalismo, entre a Fonética e a Fonologia e que relega à Fonética a descrição acústico-articulatória dos sons da fala, ao tempo em que estabelece para a Fonologia o estudo dos sons distintivos numa língua.

Digo que este estudo questiona a dissociação entre Fonética e Fonologia porque colocar as alofonias contínuas num nível fonético, como se faria, caso considerássemos os limites tradicionais supracitados, seria insuficiente para capturar a relação entre os processos alofônicos e o nível prosódico que os condiciona, dado que, por essa dissociação, o nível fonético é meramente implementacional, mecânico. Colocar as alofonias contínuas num nível fonológico, por outro lado, não permitiria captar a natureza gradiente dos processos, já que os modelos fonológicos têm por primitivos unidades categóricas.

Para iniciar essa discussão, o primeiro capítulo traz uma espécie de percurso histórico, dentro da teoria fonológica, focalizando especialmente a unidade de análise tomada por alguns modelos, desde a Fonologia Estruturalista (Jakobson & Halle, 1956), passando pela Fonologia Gerativa Padrão (Chomsky & Halle, 1991 [1968]) e pela Geometria de Traços na versão de Clements & Hume (1995), até a Fonologia Articulatória (Browman & Goldstein, 1986, 1989, 1990, 1992) e a Fonologia Acústico-Articulatória (Albano, 2001).

Abordo, particularmente, a natureza dos primitivos de análise: observo que, desde o início da Fonologia, a unidade de análise tomada tem natureza categórica. Essa concepção se propaga até modelos fonológicos atuais, como a Geometria de Traços (Clements & Hume, 1995; Sagey, 1986), que tem o traço distintivo como primitivo teórico, ou a Teoria da Otimalidade (Prince & Smolensky, 1993), que toma o segmento como unidade de análise.

Observo que, contrapondo-se a esses modelos fonológicos todos – os quais assumem primitivos “estáticos” –, surgem os modelos dinâmicos de produção de fala, cujo expoente mais significativo é a Fonologia Articulatória (Browman & Goldstein, *op. cit.*).

Há que se considerar que o aparecimento dos modelos dinâmicos de produção da fala só foi possível graças ao avanço tecnológico, ocorrido a partir da segunda metade do século passado, o qual promoveu o desenvolvimento de ferramentas computacionais, como os oscilógrafos e os espectrógrafos – para análise acústica –, ou a cinefluorradiografia, a eletromiografia e os palatômetros – para análise articulatória –, os quais permitiram um

avanço considerável da Fonética Experimental. A partir de então, tornou-se possível realizar estudos articulatórios e acústicos acurados e, conseqüentemente, explorar a alternativa instrumental às análises impressionísticas que vigoram na literatura da área e, ao mesmo tempo, observar fatos fônicos, nunca antes documentados, como as alofonias contínuas .

A despeito da disponibilidade de ferramentas robustas de análise e de a Fonética trazer evidências empíricas contrárias a algumas análises fonológicas – como é o caso das alofonias contínuas – a Fonologia (que é, na maioria, constituída pelos modelos que adotam primitivos “estáticos”) insiste em desconsiderar essas informações e continua apoiando suas conclusões numa Fonética impressionística, típica de meados do século XX.

Decorre daí que os modelos fonológicos, em geral, continuam adotando as mesmas unidades de análise estáticas, o que contribui para que se mantenha a velha dissociação entre Fonologia e Fonética².

Visando a suprimir essa lacuna da teoria fonológica, Browman & Goldstein propuseram a Fonologia Articulatória, modelo que revê a unidade de análise a ser adotada a partir de informações articulatórias. Tomando um primitivo teórico de natureza dinâmica, como o gesto articulatório, a “tradução” do dinâmico no simbólico é direta, podendo prescindir da intermediação de regras derivacionais – às vezes inúmeras e *ad hoc*.

O primeiro capítulo traz – ainda como parte do que chamei acima “percurso histórico na teoria fonológica” – a resenha de alguns textos nos quais a Fonologia Articulatória é apresentada, chamando a atenção para a natureza dinâmica do gesto e para a possibilidade que ele enseja de representar fatos dinâmicos na gramática de uma língua.

Na mesma direção da Fonologia Articulatória, a Fonologia Acústico-Articulatória (Albano, 2001) também toma o gesto articulatório como unidade mínima de análise, mas introduz algumas modificações que contribuem para uma tradução mais parcimoniosa do dinâmico no simbólico. Dentre as modificações sugeridas, a definição do gesto articulatório, em termos de um espaço acústico-articulatório – e não apenas de grupos de articuladores, como preconizam Browman & Goldstein – e a lexicalização de variantes

² Por ser um modelo representacional, e não derivacional, talvez a Teoria da Otimalidade – dentre todos os modelos fonológicos que adotam primitivos estáticos – possa estabelecer algum diálogo com a Fonologia Articulatória. Entretanto, fuge do escopo desta tese abordar como se daria esse diálogo.

dialetais são as mais significativas. O primeiro capítulo também resenha esse modelo, focalizando suas principais contribuições à Fonologia Articulatória. Argumento aí que somente modelos como a Fonologia Articulatória ou a Fonologia Acústico-Articulatória podem dar conta da representação das alofonias contínuas, em particular a dos róticos de início de palavra, devido à natureza dinâmica do gesto articulatório, o que torna esses dois modelos representacionais e não derivacionais, como a maioria dos modelos que tomam unidades de análise estáticas.

O insucesso desses modelos fonológicos em lidarem com as alofonias gradientes, em especial a dos róticos iniciais de palavra no PB, é abordado no segundo capítulo. Nele, exploro a Geometria de Traços no seu limite, quanto ao tratamento de processos alofônicos. Um dos primeiros problemas que surgem – e talvez o principal – é que no referido modelo não se sabe como tratar o traço [lateral], isto é, não se tem certeza sobre o lugar em que ocorre na hierarquia da estrutura interna do segmento. O outro problema é que, curiosamente³, o modelo sequer prevê a possibilidade de tratamento dos róticos, não mencionando como seria a estrutura interna desses segmentos, a partir do conjunto de traços tomado para representar a estrutura interna de qualquer segmento.

Minha argumentação vai no sentido de mostrar que, com o conjunto de traços proposto, é impossível dar conta da representação dos róticos na Geometria de Traços. Exploro, então, a possibilidade de propor um traço específico para a estrutura desses segmentos, mas noto que igualmente não é possível representar os sons de /r/ de modo a capturar a identidade de classe que existe na alofonia desses sons, em várias línguas do mundo. Qualquer que seja a análise tentada, ela parecerá arbitrária, mesmo que se mantenha a coerência interna da representação, de acordo com o modelo.

Se modelos fonológicos tradicionais não dão conta de representar as alofonias contínuas, como é o caso da que envolve os róticos iniciais, espera-se que modelos dinâmicos dêem conta dessa tarefa, como já mencionei acima.

Por isso, no quarto capítulo, proponho uma representação dos róticos em termos de sua organização gestual. Embora tome a Fonologia Articulatória como ponto de partida, observo que a Fonologia Acústico-Articulatória pode oferecer uma representação mais

adequada desses sons, capturando a identidade de classe entre variantes vibrantes e fricativas com maior facilidade, especialmente por conta da definição do gesto articulatório em termos de regiões acústico-articulatórias. A representação dos róticos, proposta por Albano (2001), é considerada; contudo, proponho ainda a especificação da região glotal para os róticos. Há que se considerar também que minha análise – como a da Fonologia Acústico-Articulatória, aliás – prevê que os róticos sejam constituídos de mais de um gesto, contrariamente à proposta inicial de Browman & Goldstein (*op. cit.*) e com base em Sproat & Fujimura (1993).

Por fim, noto que, dado o caráter implementacional do gesto articulatório, sua definição, a partir de parâmetros dinâmicos, e a lexicalização das variantes dialetais, é possível não só propor um modelamento mais elegante para os róticos iniciais de palavra, mas também associar a Fonética e a Fonologia, num mesmo módulo de processamento fônico, onde a tradução do gradiente no simbólico é automática, embora não-mecânica.

³ Digo curiosamente porque os róticos estão presentes na maior parte das línguas do mundo (vide Ladefoged & Maddieson, 1996), o que faria esperar que qualquer modelo fonológico cuidasse de representá-los.

1 – O GESTO ARTICULATÓRIO COMO UNIDADE CONTRASTIVA

Fonética e Fonologia são duas disciplinas historicamente dissociadas: desde o advento da ciência lingüística, à primeira é atribuído o papel de investigar as características físico-fisiológicas de quaisquer sons da fala – sejam eles distintivos ou não, numa dada língua –, fato que confere à Fonética um caráter naturalista e, portanto, não humanístico. À Fonologia, por outro lado, é atribuída a investigação dos sons funcionalmente distintos, numa língua.

Essa é a maneira pela qual pioneiros, como Trubetzkoy, separavam as águas e também a maneira pela qual muitos fonólogos – seja por *ignorância, conveniência ou imediatismo*, como diz Albano (2001: 36) – continuam a estabelecer a fronteira entre Fonética e Fonologia, a despeito de todo o acúmulo de evidências acústicas e articulatórias pela Fonética que aponta a sua utilização de parâmetros gramaticais e, conseqüentemente, a sua proximidade com a Fonologia. Essa maneira unilateral de separar as águas impera também, infelizmente, em manuais introdutórios de Fonética e Fonologia (vide, e.g., Cristófaró Silva, 1999).

A dissociação entre as duas disciplinas, porém, tem sido alvo de questionamentos de vários foneticistas, desde a segunda metade do século passado, em especial, graças ao avanço tecnológico que possibilita maior acuidade nas análises do sinal acústico. Esses foneticistas tentam fazer um caminho inverso ao percorrido desde o Estruturalismo, ou seja, tentam buscar um ponto de intersecção, a chamada “interface” entre Fonética e Fonologia. Ohala (1990), contudo, argumenta no sentido de ser inócua a busca de uma interface, já que as duas disciplinas não são distintas, mas complementares, e a interface somente se estabelece entre áreas diferentes.

O argumento mais contundente e contrário à dissociação entre Fonética e Fonologia, porém, chega-nos pelas mãos de Browman & Goldstein (1986, 1989, 1990, 1992) com sua Fonologia Articulatória – doravante FAR. Tomando como primitivo o gesto articulatório⁴, a FAR dá um passo importante no sentido de unificar as duas faces de uma mesma moeda.

Isto porque, ao mesmo tempo em que os gestos são unidades de ação articulatória dinamicamente especificadas – uma característica física, portanto – assume-se que eles organizem a fonologia de uma língua, o que lhes confere uma base cognitiva. Essas características do gesto articulatório permitem uma “tradução” direta do físico no cognitivo. Permitem também explicar e fornecer uma representação para fatos que a literatura fonética tem mostrado serem gramaticais e que a Fonologia não consegue explicar (como mostrarei nos dois capítulos seguintes). Esse é o caso das alofonias contínuas, como a que envolve o alongamento da vogal, diante de consoante sonora, investigada por Keating (1985): a autora mostra que o alongamento de vogal, antes de consoantes sonoras, é revertido em algumas línguas, nas quais a vogal se torna breve diante de tais consoantes. Este, para a autora, é um argumento de que universais fonéticos – como considerados tradicionalmente por modelos como a Fonologia Gerativa Padrão (Chomsky & Halle, 1968) – são, na verdade, específicos de uma língua.

Este primeiro capítulo, então, trata de uma resenha visando a recuperar historicamente a dissociação entre Fonética e Fonologia, num primeiro momento para, em seguida, percorrer o caminho de Browman & Goldstein, na proposição do gesto articulatório como unidade de análise fônica e, conseqüentemente, na associação de Fonética e Fonologia.

Assim, para o panorama histórico da dissociação entre Fonética e Fonologia, focalizo especialmente a Fonologia Estruturalista, na versão de Jakobson & Halle (1956), a Fonologia Gerativa de Chomsky & Halle (1995[1968]) e a Geometria de Traços, na versão de Clements & Hume (1995)⁵. No momento seguinte do texto, em que chamo a atenção para a possibilidade de associação entre Fonética e Fonologia, através de um modelo dinâmico de produção da fala, resenho “três versões” da Fonologia Articulatória: a primeira, de Browman & Goldstein, 1986; a segunda, exposta em Browman & Goldstein

⁴ Um comentário mais detalhado sobre a natureza do gesto articulatório é feito no item 1.3, adiante.

⁵ Certamente há outros vários modelos dentro de cada corrente. Poder-se-ia citar a Fonêmica, de Pike (1947), no Estruturalismo, ou a Fonologia Natural, de Stampe (1973, *apud* Katamba, 1989), numa linha gerativa, ou ainda a Geometria de Traços, na versão de Sagey (1986). Entretanto, comento os modelos que se seguem, a partir do próximo item, ou por julgá-los marco dentro de um arcabouço teórico – como no caso de Jakobson

1989 e 1990 e a última, de Browman & Goldstein, 1992⁶. Termino o capítulo com uma resenha da Fonologia Acústico-Articulatória de Albano (2001), buscando enfatizar suas contribuições para a Fonologia Articulatória, já que a proposta de Albano (*op.cit.*) torna possível modelar a alofonia contínua, presente na variabilidade dos róticos de início de palavra no PB.

1.1 – O início do problema: a definição do objeto de estudo da Fonética e da Fonologia

1.1.1 – O fonema como unidade mínima de análise

O estudo dos sons da fala remonta a um período anterior ao estabelecimento da Lingüística como ciência: já em meados do século XIX, ele era realizado por “cientistas da linguagem” (filólogos, na verdade), sob uma perspectiva que procurava explicar a pronúncia das letras do alfabeto, fosse ela voltada aos aspectos históricos envolvidos na mudança das letras e da sua pronúncia, fosse ela voltada à busca de “leis” gerais que regulassem a mudança de pronúncia.

É nesse contexto que surge, pela primeira vez, a noção de fonema: no final do século XIX, o filólogo Baudouin de Courtenay nota que, a despeito da alternância Fonética, verificada na pronúncia de alguns sons, os falantes têm deles uma única representação. Assim, por exemplo, no caso do PB, os falantes teriam uma única representação para o “t” dental-alveolar de *toca* e o africado de *tipo*⁷.

& Halle (1956) ou Chomsky & Halle (1995[1968]) – ou por serem modelos largamente utilizados para a investigação da Fonologia do PB, como Clements & Hume (1995).

⁶ Devo esclarecer que o estabelecimento das três versões da Fonologia Articulatória é um recurso que emprego com fins meramente expositivos. Em momento algum, os autores reconhecem diferentes versões do modelo. Ainda como esclarecimento: embora publicado em 1990, *Tiers in articulatory phonology, with some implications for casual speech* é anterior a *Articulatory gestures as phonological units*, publicado em 1989. “*Tiers*” foi escrito em 1987 e circulou, primeiramente e nesse mesmo ano, no *Status Report* do Laboratório Haskins. Dada a grande complementaridade entre esses dois textos, abordo ambos no item 1.3.1.2, reconhecendo neles a segunda versão da Fonologia Articulatória.

⁷ Isto, frise-se, independentemente de sermos ou não alfabetizados.

Courtenay interpreta esse fato da seguinte maneira: os indivíduos têm uma espécie de “imagem mental”⁸ do som que objetivam produzir, no momento em que falam. Essa imagem é invariável e pode ou não ser realizada concretamente, no momento em que os falantes produzem um enunciado. No caso em que a imagem mental do som não é realizada concretamente, o som resultante sempre se assemelha ao visado, fazendo todos eles – som visado e som realizado – parte de uma mesma “família”. Voltando ao nosso exemplo: pela perspectiva de Courtenay, a imagem mental do som inicial de *tipo* e *toca* é /t/. Quando falam, porém, os indivíduos concretizam essa imagem em alguns casos – *toca* – mas não em outros – *tipo*, o que faz que tenhamos sons física e fisiologicamente distintos.

A representação, ou imagem mental que os falantes têm de um som da fala, é o que Courtenay denomina “fonema”, em oposição ao que denomina “fone”, que é a realização concreta do som visado. Ainda tomando o exemplo do PB, temos que [tʃ] e [t] são os fones, o fonema, mais abstrato, é /t/. É preciso ressaltar – como frisa Jones (1957) – que essa definição de fonema é estritamente Fonética (e nem poderia ser diferente, pois a Fonologia só viria a surgir no bojo do Estruturalismo Lingüístico).

Se, antes do estabelecimento da ciência Lingüística, temos uma unidade de análise fônica, definida em termos estritamente fonéticos, o surgimento da Lingüística impõe uma concepção radicalmente diferente a essa unidade. É possível, inclusive, dizer que Saussure passa a utilizar o mesmo rótulo “fonema” em seu modelo, mas dá a ele uma acepção distinta e que remete à oposição que tal unidade opera dentro de um sistema lingüístico⁹: embora a definição que Saussure dá para a unidade “fonema” considere os aspectos articulatórios e acústicos dessa unidade, o crucial na perspectiva saussureana é a concepção de que o fonema – como aliás, outras unidades lingüísticas, a exemplo do signo – deve operar distinções num sistema lingüístico. Nessa perspectiva, a Fonética é considerada uma

⁸ Courtenay não utiliza explicitamente o termo “imagem mental”, mas é possível antever, em sua perspectiva, um certo psicologismo o qual, aliás, será retomado mais tarde dentro da Teoria Lingüística, embora não exatamente da mesma maneira, por autores como Sapir (*A realidade psicológica do fonema*, 1981[1933]), ou Chomsky & Halle (*The sound pattern of English*, 1995[1968]).

⁹ O que, aliás, é consoante com a perspectiva saussureana de que os sistemas lingüísticos se organizam na base de oposições.

ciência diacrônica e seus objetivos confundem-se com os da Filologia, em certa medida, já que, para Saussure, a Fonética deveria se ocupar do estudo “evolutivo” dos sons.

A dissociação entre Fonética e Fonologia se institui com o Estruturalismo Lingüístico: o estabelecimento das bases de investigação da Fonologia por Trubetzkoy (1967)¹⁰, inicialmente, e por Jakobson & Halle (1956), num segundo momento, relega à Fonética uma perspectiva essencialmente descritivista, dentro da Lingüística.

Assim é que Trubetzkoy atribui à Fonética o estudo dos sons de uma dada língua, em suas propriedades acústicas e articulatórias. À Fonologia cabe estudar os sons que possuem caráter distintivo nessa língua. A Fonética, por seu turno, tem, digamos assim, apenas um papel acessório no estudo dos sons de uma língua. Trubetzkoy chega mesmo a dizer que a relação entre Fonética e Fonologia equivale à relação entre Numismática e Economia.

1.1.2 – Os traços distintivos: unidades mínimas constitutivas do fonema

Também Jakobson & Halle colocam a Fonética num segundo plano, já que, igualmente para esses autores, interessa investigar as unidades distintivas de uma língua. A diferença entre a perspectiva adotada por eles e por Trubetzkoy está no fato de que Jakobson & Halle observam que o fonema pode ser decomposto em unidades distintivas menores, os traços, os quais devem ter uma substância fonética. Para os autores, num par como “mala” e “bala”, por exemplo, a distinção entre as palavras não se dá pelos fonemas /m/ e /b/, mas por uma característica, ou traço, desses fonemas, presente numa palavra e ausente noutra; neste caso, o traço em questão é a nasalidade.

Os traços distintivos operam numa base binária e são definidos em termos acústicos e articulatórios – ou genéticos, como denominam os autores em *Fundamentals of Language*. São estabelecidos com base na descrição dos sons da fala, levada a cabo por Jakobson, Fant & Halle (1976) em seu *Preliminaries to speech analysis*. Tais unidades, organizadas em feixes, constituem os fonemas e, enquanto unidades mínimas da cadeia da fala, assume-se que devam dar conta de todos os aspectos fônicos nela presentes, quer dizer, eles devem

¹⁰ Refiro-me aqui à tradução inglesa dos *Grundzüge der Phonologie*, publicada em 1969.

marcar oposição não apenas entre sons, mas entre os aspectos que se distribuem por mais de um som, como os aspectos suprasegmentais.

É preciso ressaltar que, no modelo estruturalista de Jakobson & Halle, o binarismo dos traços reflete o mesmo jogo de oposições presente em Saussure. Esse jogo define as quatro maiores classes de sons nas línguas do mundo, a partir das possibilidades combinatórias das especificações. Assim, as consoantes são definidas pelos traços [+consonantal/-vocálico]; as vogais são definidas pelos traços [-consonantal/+vocálico]; as líquidas, por sua vez, são [+consonantal/+vocálico] e as semivogais, [-consonantal/-vocálico].

A maneira como cada traço é caracterizado no modelo decorre do fato de que Jakobson & Halle recorrem às propriedades acústicas e articulatórias dos sons, a partir de uma descrição acústica e articulatória dos sons da fala. Desse *modus operandi* resulta o caráter universal dos traços distintivos, axiomáticamente capazes de dar conta de todas as distinções fonológicas, em todas as línguas do mundo.

Vê-se que, neste modelo, a tarefa da Fonética é distinta daquela que lhe é atribuída no modelo saussureano. Isto porque Saussure toma a Fonética como uma ciência histórica e que, por essa razão, não pode fazer parte da Lingüística. Para Jakobson & Halle, em contrapartida, cabe à Fonética o estudo das características acústicas e articulatórias dos sons da fala. À Fonologia caberá estudar como a língua utiliza o som, selecionando alguns de seus elementos e adaptando-os a diversos fins. É como se a Fonologia “lapidasse” os dados fonéticos, passando a considerar somente aqueles que operam distinções nos sistemas lingüísticos. Essa posição fica clara, na seguinte afirmação de Jakobson & Halle:

Enquanto a Fonética procura coleccionar o maior número de informações possíveis sobre o som em seu estado bruto, em suas propriedades fisiológicas e físicas, a Fonêmica¹¹, e a Fonologia em geral, aplicam critérios estritamente

¹¹ A distinção que Jakobson & Halle fazem, nesse momento, entre “Fonologia” e “Fonêmica” é a seguinte: “Fonêmica”, um termo empregado nos EUA, termo americano, refere-se à investigação sobre o modo como uma língua seleciona alguns de seus elementos e os adapta a vários fins, por exemplo, o de servir como um meio distintivo. O termo “Fonologia”, por sua vez, que Jakobson & Halle mencionam ter sido lançado em

lingüísticos à classificação do material registrado pela Fonética. (Fundamentals of Language p. 7)

Apesar de considerar a importância dos dados fonéticos para um modelo fonológico, Jakobson & Halle também estabelecem uma distinção entre Fonética e Fonologia: a primeira deve esgotar o tratamento dos sons da fala sob uma perspectiva concreta, recorrendo às características “visíveis” dos sons. A segunda, por sua vez, utiliza as informações fonéticas e, abstraindo-as, procura estabelecer as oposições fônicas que têm caráter funcional nos sistemas lingüísticos. A abstração dos dados fonéticos é uma exigência do modelo: no mesmo *Fundamentals*, Jakobson & Halle afirmam a necessidade de a Fonologia recorrer aos dados fonéticos, porém abstraindo-os ao máximo. Novamente, portanto, a Fonética é considerada um “acessório” nos estudos fonológicos.

1.2 – A dissociação entre Fonética e Fonologia numa outra perspectiva

1.2.1 - A Fonologia Gerativa

A Fonologia Gerativa de Chomsky & Halle (1995[1968]), diferentemente dos modelos que a antecederam, passa a se preocupar em estabelecer uma relação entre a competência fonológica do falante e o seu desempenho, i.e., entre a representação mental de processos fonológicos e a sua realização. Surge, então, a necessidade de se prever um conjunto de regras que gere uma saída a partir de uma representação abstrata. No referido modelo, a saída, ou *output*, trata-se da representação dos segmentos da cadeia da fala, como abordo adiante.

A preocupação com a representação mental de processos fônicos e com a realização dos mesmos torna possível começar a pensar a relação entre Fonética e Fonologia no

1923 com base em sugestões da Escola de Genebra, deve ser preferido, já que *a vantagem do termo “fonologia” poderia ser sua aplicação mais fácil a toda a variedade de funções lingüísticas desempenhadas pelo som, enquanto “fonêmica” sugere um confinamento às propriedades distintivas e, portanto, é uma*

arcabouço da Fonologia Gerativa. Tal relação era inconcebível na Fonologia Estruturalista, cujo objetivo único era fazer uma descrição dos sons de uma língua, em termos dos traços distintivos que os constituíam.

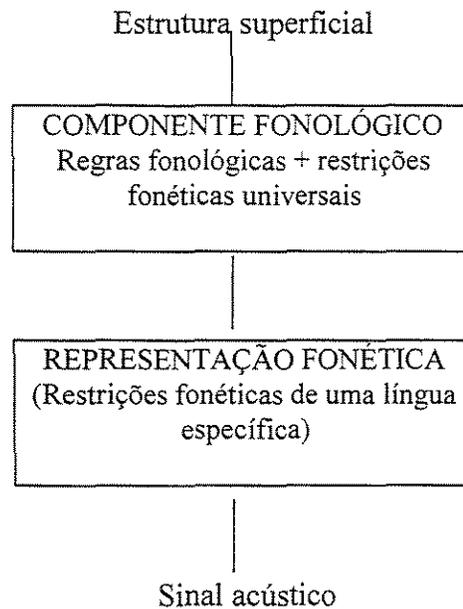
Obviamente, a possibilidade de pensar a relação entre as duas disciplinas decorre da concepção de linguagem subjacente ao modelo, que aliás herda muitos pressupostos do modelo sintático chomskyano, como o pressuposto de que a linguagem é um conjunto – finito ou infinito – de sentenças, as quais são geradas por regras de uma gramática internalizada na mente dos falantes. Em linhas gerais, essas regras fornecem uma representação fonética para cada sentença de uma língua, a partir de uma representação semântica. A preocupação central do modelo sintático chomskyano será, conseqüentemente, determinar como são geradas as sentenças no interior de uma língua particular, para isso descrevendo as regras que relacionam o significado de uma sentença a sua estrutura sonora.

Essas regras se aplicam em diferentes componentes da linguagem – como, e.g., o semântico ou o sintático –, de modo que um enunciado recebe informações referentes a cada um desses componentes, à medida que é processado. No que concerne à Fonologia, Chomsky & Halle (*op.cit.*) assumem que há um componente específico para seu processamento, o componente fonológico. Tal componente, como explicam os autores, faz a ponte entre uma estrutura superficial e o sinal acústico. Por estrutura superficial entende-se um enunciado no qual os itens lexicais são etiquetados com informações dos níveis nos quais já foram processados. Portanto, ao atingir o componente fonológico, os itens lexicais da estrutura superficial carregam informações morfossintáticas.

No componente fonológico, as regras fonológicas se aplicam sobre o enunciado, gerando uma representação abstrata, baseada em traços fonológicos. Esses traços, classificatórios, são binários. Regras da gramática de uma língua podem alterar o valor de um traço, além de inserir ou apagar segmentos, gerando daí fatos como a sonorização de /s/ de coda diante de soantes, no caso do PB.

Além de gerarem uma representação abstrata do enunciado, as regras fonológicas interagem com restrições fonéticas universais¹², gerando uma outra representação – anterior ao sinal de fala –, que é a transcrição, ou representação, fonética. Na representação fonética, o enunciado se constitui de uma seqüência de unidades discretas. Tais unidades são matrizes bidimensionais de traços fonéticos, os quais correspondem a escalas físicas que, segundo Chomsky & Halle, mapeiam aspectos controláveis da fala, tais como altura da língua, glotalização ou nasalização. Como decorrência, haveria tantos traços fonéticos quantos os aspectos passíveis de controle relativamente independente na fala.

A representação fonética, por seu turno, recebe informações fonéticas específicas de uma língua, donde resulta, finalmente, o sinal de fala. Portanto, é possível esquematizar a estrutura do processamento fonético-fonológico de Chomsky & Halle (*op. cit.*) da seguinte maneira:



É preciso ressaltar, no esquema acima, a diferença conceptual de transcrição fonética relativamente à Fonologia Estruturalista: como já mencionei, a Fonologia Gerativa considera a transcrição como uma representação, gerada por regras gramaticais específicas de uma língua. Conseqüentemente, comentam os autores, é natural haver discrepâncias

¹² Como, por exemplo, o alongamento de vogal diante de consoante sonora, cf. Albano *et al.* (1998).

entre a transcrição fonética e o sinal acústico, assim como é natural que a transcrição se constitua de símbolos discretos, diferentemente do sinal acústico, que é “quase-contínuo”.(p.294)

A Fonologia Estruturalista, porém, via a transcrição fonética como uma reprodução do sinal acústico. Chomsky & Halle (op.cit.) atribuem a essa concepção de transcrição fonética a impossibilidade de associar Fonética e Fonologia num arcabouço teórico estruturalista. Dizem eles:

“... mesmo que a transcrição fonética fosse um registro fiel da fala, como se poderia desejar, restam ainda questões sobre o interesse desse registro para os lingüistas, que estão preocupados primeiramente com a estrutura da linguagem e não com a acústica ou a fisiologia da fala. Por conta dessa questão, muitos lingüistas estruturalistas consideraram que a Fonética tem muito pouco a oferecer e por isso a colocaram num papel secundário, periférico.” (Chomsky & Halle, 1995[1968]: 293. A tradução é minha.)

O fato, portanto, de a Fonologia Gerativa prever uma representação fonética para a cadeia da fala permite incorporar o detalhe fonético à representação fonológica e à gramática. Isso não significa, entretanto, que seja possível contemplar todo o detalhe fonético lingüístico, dentro desse modelo: as alofonias contínuas, como as descritas por Albano *et alii* (1998) para o PB, ou como o *flapping* do inglês norte-americano, estudado por Fox & Terbeek (1977), para citar alguns exemplos, mostram-se gramaticais, porque condicionados por fatores gramaticais, bem como prosódicos. Mesmo considerando que os traços fonéticos são escalares, é impossível tratar o contínuo físico pelo qual se estendem as alofonias contínuas, em termos de uma escala física¹³. Ao fim e ao cabo, é ainda necessário prever, num modelo fonológico, como contemplar fatos que se distribuam num contínuo de fato numérico.

¹³ Sobre esse tópico farei um comentário mais detido no segundo capítulo.

1.2.2 - A Geometria de Traços

A Geometria de Traços, em qualquer de suas versões, representa um inegável avanço teórico, relativamente à Fonologia Gerativa de Chomsky & Halle (*op.cit.*): a proposta de uma organização tridimensional interna dos segmentos, na qual os traços se dispõem, hierarquicamente, em camadas e atrelados a nós de classe, permite representar e explicar fatos que um modelo linear – como, por exemplo, a Fonologia Gerativa ou a Fonologia Natural – não contemplariam. É o caso de processos não-lineares, como a assimilação a distância, em palavras como [mi.'ni.nu], no PB, onde a vogal pré-tônica sofre elevação, em função da natureza da vogal tônica.

Um modelo linear teria dificuldade em explicar esse fato,¹⁴ porque prevê que as regras fonológicas de assimilação se aplicam mais naturalmente entre segmentos adjacentes. Uma Geometria de Traços, em contrapartida, prevê que as duas vogais estejam adjacentes numa mesma camada, distinta daquela em que se encontra a consoante, o que permite a aplicação de uma regra de espalhamento do traço [anterior] da vogal tônica para a pretônica.

A despeito desses avanços, porém, esse modelo continua tomando o traço distintivo como primitivo. Disso decorre que toda a representação fonológica se baseia em símbolos e, portanto, assume um caráter discreto. Além disso, apesar de Clements & Hume (1995) advogarem uma “motivação fonética” para a estrutura interna dos segmentos, sua versão considera apenas uma descrição articulatória impressionística para a definição de traços e nós de classe. (Pode-se dizer, ao fim e ao cabo, que a concepção de Fonética subjacente a essa versão é um tanto “arcaica”, ou até mesmo equivocada.) Como decorrência, tem-se uma relação muito tênue entre Fonética e Fonologia.

Essa relação tênue entre Fonética e Fonologia, somada à representação fonológica discreta de quaisquer processos fônicos, impossibilita a Geometria de Traços, em qualquer versão e, em particular, na versão de Clements & Hume (*op.cit.*), de tratar processos gradientes de variação alofônica, como os descritos por Albano (2001), ou como o que

¹⁴ Não utilizo aqui o termo “harmonia vocálica” como referência a esse fato porque a sua existência, em PB, pode ser questionada, cf. D’Angelis (comunicação pessoal). O questionamento se dá com base na ocorrência de alçamento de vogais em palavras como [bu.'nɛ.kɐ] ou [bu.'ʃe.ʃɐ], casos que claramente não envolvem o alçamento da pretônica por influência da tônica, i.e., uma “harmonia vocálica”, propriamente.

tratarei mais adiante, no terceiro capítulo¹⁵, envolvendo a variabilidade de pronúncia dos róticos do PB.

Albano (*op.cit*) mostra, por exemplo, que indivíduos “não-chiantes”, i.e., que aparentemente não produzem uma variante africada de /t/ diante de /i/ apresentam, na verdade, uma gradação na produção de [tʃ], havendo ambientes favorecedores dessa alofonia.

Ora, explicar a alofonia contínua da africada em PB é impossível no arcabouço de qualquer Geometria de Traços, posto que todas as suas versões trabalham com regras de alteração de representações, prevendo o apagamento de algum traço ou a ligação de traços latentes a nós terminais. Para explicar fatos contínuos como o que cito acima, é necessário repensar a natureza da unidade de análise adotada: não é possível que a mesma tenha caráter estático, como o traço distintivo. Como se verá abaixo, para explicar processos gradientes são necessárias unidades de natureza dinâmica.

1.3 – Uma nova unidade de análise: o gesto articulatório

1.3.1 – A Fonologia Articulatória

1.3.1.1 – Evidências para o estabelecimento do gesto articulatório como primitivo teórico e sua formalização inicial

A unidade de natureza dinâmica à qual me refiro, ao final do item anterior, é o gesto articulatório¹⁶, proposto por Browman & Goldstein (1986, 1989, 1990, 1992), no arcabouço da Fonologia Articulatória.

¹⁵ Nesse capítulo tratarei mais detidamente das limitações da Geometria de Traços em lidar com a variabilidade de pronúncia de /r/.

¹⁶ O enfoque sobre aspectos articulatórios da fala é preferido aos aspectos acústicos, nesse modelo. O motivo dessa decisão é explicitado no texto de 1990. Dizem Browman & Goldstein, na página 341: *Em nossa opinião, a estrutura fonológica é uma interação de organizações acústica, articulatória, e outras (e.g. psicológica e/ou linguística). Estamos enfocando a organização articulatória porque acreditamos que a natureza inerentemente multidimensional da articulação pode explicar vários fenômenos fonológicos, particularmente aqueles que envolvem a sobreposição de gestos articulatórios. Portanto, representamos as*

Na primeira apresentação do modelo, de 1986, Browman & Goldstein observam que a distância estabelecida entre Fonética e Fonologia se deve, mormente, à natureza estática dos primitivos, tomados pelos modelos fonológicos, sejam eles traços, segmentos ou fonemas. Argumentam que somente um modelo baseado numa unidade de natureza dinâmica é capaz de fazer naturalmente a ponte entre o fonético (ou gradiente) e o fonológico (ou simbólico), prescindindo de regras transformacionais que tentam “traduzir” um noutro nível. A tradução “automática” do fonético no fonológico, por conseguinte, torna desnecessário prever que os aspectos fonético e fonológico sejam processados em dois níveis, distintos e organizados hierarquicamente, como preconiza a Fonologia Gerativa de Chomsky & Halle (1995[1968]), coerente com o modelo sintático chomskyano.

Foi preciso, nessa primeira versão da FAR, definir seu primitivo teórico. Para isso, Browman & Goldstein partem da observação de que inexistente um estado estacionário dos articuladores¹⁷. Tal observação se origina da medida da trajetória do lábio inferior na oclusão de /b/, numa seqüência simples, [abə]¹⁸. A trajetória mostra um movimento contínuo do articulador: ao atingir um máximo de deslocamento, durante a oclusão, o lábio inferior move-se novamente para baixo. Diante disso, os autores argumentam que melhor do que buscar um padrão articulatório estático para /b/ é definir esse som em termos de um gesto de oclusão labial.

Deve ficar claro, a partir do que foi exposto acima, que não há uma correspondência biunívoca entre segmentos e gestos. De fato, um mesmo gesto pode constituir mais de um segmento: o gesto de oclusão labial, por exemplo, não está presente apenas no segmento

estruturas lingüísticas em termos de movimentos articulatórios coordenados, chamados 'gestos', os quais são organizados numa 'pauta gestual' que se assemelha a uma representação autosegmental. (Tradução e grifo meus.)

¹⁷ A afirmação de que inexistente um estado estacionário dos segmentos é empregada em caráter retórico por Browman & Goldstein (*op.cit.*). Note-se que imagens cineradiológicas de consoantes articuladas com a língua (cf. Machado, comunicação pessoal) evidenciam a existência de estados estacionários que duram de 40 a 60 ms.

¹⁸ Frise-se que o movimento do lábio inferior durante a oclusão de /b/ será a base para toda a formalização do gesto articulatório, desde a observação da inexistência de um estado-estacionário para os sons da fala, até a demonstração de que o emprego do gesto articulatório como unidade de análise fônica proporciona um tratamento mais parcimonioso e menos controverso para fatos que a literatura fonológica não consegue resolver de maneira elegante.

/b/, mas também em /p/ e /m/. Tampouco o gesto corresponde a um traço: um traço pode caracterizar mais de um gesto.

A argumentação de que um som não deve ser definido por um estado estacionário, mas por um gesto articulatório, definido no espaço e no tempo¹⁹, abre precedente para que Browman & Goldstein demonstrem que dois fatos controversos da literatura fonológica segmental²⁰ – os encontros /s/-oclusivas em inglês e as prenasalizadas – podem ser explicados com maior parcimônia, caso se tome o gesto articulatório como unidade de análise.

Sobre os encontros /s/-oclusiva, Browman & Goldstein mencionam que haveria, ao menos, dois fatos que requerem observações específicas numa fonologia estática: 1) a fonotática da língua inglesa deve determinar que não há contraste entre oclusivas surdas e sonoras, seguindo /s/ inicial de palavra; 2) a realização dessa oclusiva como surda não-aspirada deve ser especificada por uma regra fonética ou fonológica separada. Note-se que quaisquer das observações mencionadas requerem a aplicação de uma regra pós-lexical, num modelo fonológico cujo primitivo tem natureza estática. Segundo ressaltam Browman & Goldstein, fonólogos como Ewen (1982, *apud* Browman & Goldstein, 1986) frisam a ambigüidade desse encontro, que pode ser analisado como uma seqüência de dois segmentos ou como um único segmento, com o argumento de que reconhecer a existência de dois segmentos levaria à violação da escala de sonoridade.

O argumento de Browman & Goldstein vai no sentido de afirmar que as duas observações mencionadas, acerca do encontro /s/-oclusiva, decorrem de restrições mais gerais sobre a estrutura articulatória do inglês. Tal restrição se aplicaria sobre o comportamento da glote, nesses encontros. Línguas germânicas em geral, segundo os autores, exibem contraste entre oclusiva surda aspirada e oclusiva sonora, mas neutralizam esse contraste, quando a oclusiva sucede /s/. E a neutralização ocorre de tal modo que se

¹⁹ A incorporação da coordenação espaço-temporal ao gesto tem duas funções no modelo, de acordo com Browman & Goldstein: a) definir uma classe fonológica, pela estreita relação entre dois gestos; b) capturar aspectos sistemáticos, próprios de uma língua, relativos à organização temporal dos gestos.

²⁰ Utilizo aqui a mesma nomenclatura que Browman & Goldstein empregam para se referir ao que, seguindo Albano (2001), eu denomino em outros lugares “modelos fonológicos estáticos”, ou seja, modelos – lineares ou não – que tomam o traço, o fonema ou o segmento como unidade de análise.

verifica apenas um gesto glotal no encontro /sC/. Isto é diferente, frisam os autores, de casos onde /s/ precede fronteira de palavra iniciada por consoante oclusiva: em casos como esses, Browman & Goldstein reportam que é possível notar dois gestos glotais, um presente na articulação de /s/ e, outro, na articulação da consoante.

Há, portanto, uma restrição articulatória para a ocorrência de dois gestos glotais contíguos em início de palavra, no inglês. Isso não significa que esse gesto não possa co-ocorrer com outros: de fato, Browman & Goldstein apontam para a ocorrência solitária do gesto de abertura/oclusão glotal, em início de palavra, como em /ha/, para a coocorrência desse gesto com um gesto de constrição oral, como em /sa/, ou com a coocorrência desse gesto com dois outros gestos de constrição oral, como em /spa/. Essa generalização dá conta, inclusive, da falta de contraste entre /sp/ e /sb/, no inglês.

Note-se, enfim, que a análise em termos gestuais resolve a questão da ambigüidade, envolvendo a análise de grupos /sC/, já que o grupo deixa de ser considerado uma seqüência de segmentos e passa a ser analisado como uma organização gestual. Diferenças nessa organização é que resultam em diferentes seqüências de sons, como as exemplificadas no parágrafo acima. Não se colocam, também, as duas questões com as quais um modelo fonológico estático teria de lidar, concernentes à determinação fonotática da ausência de contraste entre oclusivas surdas e sonoras, sucedendo /s/ inicial de palavra, e à necessidade de sinalizar a realização da oclusiva surda como uma não-aspirada, via regra específica.

Com relação às prenasalizadas, Browman & Goldstein observam que as análises fonológicas estáticas, especialmente as lineares, não captam adequadamente a diferença entre prenasalizadas e encontros nasais-oclusivas, dado que tais modelos não conseguem prever as regularidades temporais desses encontros.

Os autores fazem menção a uma análise linear das prenasalizadas, realizada por Anderson (1976, *apud* Browman & Goldstein, 1986) e segundo a qual haveria duas alternativas de tratamento para esses sons: a) prever uma matriz na qual cada traço fosse especificado para cada segmento. Como decorrência, não haveria diferença formal entre prenasalizadas e encontros do tipo nasal-oclusiva; b) prever que as prenasalizadas são um

segmento simples que compartilham alguns traços, como [nasal], na sua matriz. Assim, para a segunda alternativa teríamos (cf. Browman & Goldstein, 1986: 229):

	m	b
consonantal	+	
nasal	+	-
anterior	+	
coronal	-	

No caso desta segunda alternativa, o domínio de atribuição do valor do traço não coincide com as fronteiras dos segmentos²¹.

Uma interpretação gestual, conforme Browman & Goldstein, faria a previsão de que a oclusiva prenasalizada apresenta um único gesto de oclusão bilabial, ao passo que o encontro nasal-oclusiva deve ter dois gestos de oclusão labial ou um único gesto de oclusão labial, mais longo do que esse mesmo gesto nas prenasalizadas. Para verificar qual das alternativas acima é mais plausível, comparam a trajetória articulatória dos lábios em palavras do inglês, contendo o encontro nasal-oclusiva – como *camper* e *camber* – e palavras contendo apenas uma das duas consoantes – como *capper* e *cabber*.

Num momento seguinte, comparam as trajetórias do lábio inferior, nessas palavras do inglês, com dados do chaga, língua bantu, que contrasta oclusivas com prenasalizadas. De acordo com Browman & Goldstein, tal comparação se justifica porque, em inglês, nasais em grupos apresentam ambigüidade acústica, quando comparadas a segmentos solitários²².

O estudo articulatório conduzido por Browman & Goldstein sinaliza, segundo eles, que, contrariamente às expectativas de Anderson (1976), todas as categorias fonológicas (exceto a seqüência nasal silábica + oclusiva) são constituídas por um único gesto de

²¹ Essa análise me parece metodologicamente duvidosa, pois, levando em conta Chomsky & Halle (1995[1968]), um traço deve ser especificado, positiva ou negativamente, para cada segmento. No máximo, um traço pode permanecer subespecificado, o que não é o caso da análise de Anderson (*op. cit.*).

oclusão labial. Isso os leva a afirmar a inexistência de diferença sistemática entre gestos labiais, associados a uma única consoante, a uma consoante prenasalizada ou a um grupo, tanto no inglês como no chaga.

Browman & Goldstein notam, porém, uma diferença no grupo /mp/ do chaga: comparando /mpaka/ com /paka/, os autores observam que a duração geral do gesto do lábio inferior é maior em /mpaka/. Tal diferença, argumentam, resultaria da sobreposição de dois gestos labiais.

Para testar sua hipótese, comparam a trajetória do lábio inferior em /maka/ e /mpaka/, observando ser mais longo o gesto deste último. Em seguida, comparam a trajetória do lábio inferior para /paka/ e /maka/, observando não haver diferença de duração entre as trajetórias. Finalmente, sobrepõem parcialmente a trajetória do lábio inferior de /maka/ à trajetória do lábio inferior em /paka/, obtendo resultado semelhante à trajetória do lábio inferior em /mpaka/.

Diante disso, sustentam que, embora fosse possível sugerir simplesmente a maior duração do gesto de oclusão bilabial, na seqüência nasal silábica + oclusiva, há evidências da presença de dois gestos de oclusão bilabial. Essas evidências seriam a amplitude e os envelopes do *onset* e do *offset* em /mp/, os quais não apresentam diferença, relativamente à consoante solitária.

Browman & Goldstein concluem que o que está em jogo na diferenciação entre consoantes solitárias – nasais ou oclusivas bilabiais – consoantes prenasalizadas e a seqüência nasal silábica + oclusiva é o número de gestos de oclusão bilabial aí presentes: dois no caso da seqüência nasal silábica + oclusiva e apenas um, nos demais casos. Uma descrição como esta é seguramente mais parcimoniosa do que uma análise como a de Anderson (*op.cit.*): seria necessário, para representar as prenasalizadas do chaga, assim como os encontros nasal-oclusiva do inglês, seguir a primeira alternativa, proposta pelo fonólogo, i.e., considerar que se tem, nesses casos, segmentos simples e que o traço [nasal] deveria ser especificado para um e outro segmento. Mas a seqüência nasal silábica +

²² De acordo com Vatikiotis-Bateson (1984, apud Browman & Goldstein, 1986), há encurtamento de nasais e oclusivas, quando ocorrem em grupo, comparando-se suas durações, quando ocorrem sozinhas.

oclusiva só poderia ser representada através da segunda alternativa, i.e., considerando-se que o traço [nasal] seja compartilhado tanto pela nasal como pela oclusiva. Ora, tomar duas matrizes de traços para representar um mesmo fenômeno, além de pouco parcimonioso, dificulta o postulado de regras derivacionais. Talvez fosse preciso propor regras *ad hoc* para derivar a seqüência nasal silábica + oclusiva (/mp/).

Dadas as evidências em favor do gesto articulatório como um primitivo que proporciona um modelamento mais parcimonioso e generalizante dos fatos mencionados, Browman & Goldstein passam à formalização do gesto articulatório. Essa unidade de análise é definida a partir da especificação: 1) de uma equação dinâmica; 2) do movimento de uma ou mais variáveis, correspondentes à posição instantânea do objeto na equação dinâmica; 3) dos valores dos parâmetros dinâmicos na equação; 4) dos pesos para articuladores individuais.

Explico melhor: a primeira noção de gesto articulatório, explicitada nesta versão inicial da FAR, toma o gesto como um padrão característico de movimento de um articulador – ou de um subsistema articulatório – no espaço, em função do tempo. Para definir os padrões espaço-temporais, dizem Browman & Goldstein que é preciso levar em conta as variações das trajetórias dos articuladores, devidas, por exemplo, a mudanças na taxa de elocução ou fatores prosódicos. Por isso, afirmam, é inviável modelar, ponto por ponto, as trajetórias dos articuladores, sendo melhor entender a mudança de posição no tempo como o resultado de sistemas mais abstratos, como sistemas dinâmicos, que são capazes de especificar padrões característicos do movimento, organizados no tempo e no espaço.

Embora originalmente utilizados para modelar a coordenação de movimentos em sistemas biológicos, os sistemas dinâmicos teriam duas características que, de acordo com Browman & Goldstein, os tornam propícios para a descrição dos gestos articulatórios: 1) eles geram um número infinito de trajetórias diferentes, mas relacionadas, partindo de um número fixo de especificações dos seus parâmetros, em função das condições iniciais dos articuladores e em função de outros sistemas dinâmicos – como, por exemplo, gestos contíguos – que possam estar ativados; 2) embora os articuladores se movam para realizar

um gesto, a equação dinâmica, derivada pelos sistemas, não varia em função do tempo, mas caracteriza todo o padrão do movimento.

O exemplo de sistema dinâmico, arrolado pelos autores, é um sistema massa-mola, no qual uma massa, presa a uma mola, desencadeia a oscilação do sistema, ao ser puxada – esticando a mola – e solta, em seguida. O movimento oscilatório é descrito pela equação²³

$$m\ddot{x} + b\dot{x} + k(x - x_0) = 0,$$

onde, m = massa do objeto,

b = amortecimento do sistema

k = rigidez da mola,

x_0 = comprimento da mola na posição de repouso,

x = deslocamento instantâneo do objeto,

\dot{x} = velocidade instantânea do objeto,

\ddot{x} = aceleração instantânea do objeto.

Alterados os valores relativos à massa do objeto, rigidez da mola e comprimento da mola na posição de repouso, obtêm-se diferentes trajetórias do sistema. Também a mudança no movimento inicial da posição instantânea do objeto e na sua aceleração instantânea pode provocar diferenças nas trajetórias, segundo Browman & Goldstein. Nesse sentido, mudanças na rigidez da mola introduzem alterações na frequência de oscilação da massa. Mudanças na posição de repouso da mola ou na sua posição inicial, por sua vez, afetarão a amplitude da oscilação.

Na equação acima – frisam os autores – a posição instantânea do objeto representa, axiomáticamente, a posição vertical do lábio inferior, para o modelamento de oclusivas bilabiais. No entanto, um gesto de oclusão bilabial não pode ser totalmente descrito pelo

²³ É preciso ressaltar que transcrevo aí a equação em sua versão definitiva, tal qual aparece nos textos de Browman & Goldstein, a partir de 1989. No texto de 1986, não há qualquer menção às variáveis “amortecimento do sistema” e “velocidade instantânea do objeto”. Não consegui recuperar o percurso dos autores, no sentido de inserir essas duas variáveis. Acredito, no entanto, que o amortecimento do sistema evite um *overshoot* articulatório, i.e., a variável “amortecimento” impede que um articulador ultrapasse seu alvo, durante a produção de um gesto articulatório.

movimento de um único articulador, pois envolve o movimento do lábio superior, do lábio inferior e da mandíbula.

Para coordenar a ação dos gestos na FAR, Browman & Goldstein recorrem à Dinâmica de Tarefa (Saltzman & Kelso, 1983, *apud* Browman & Goldstein, 1986), a qual prevê que um conjunto de articuladores, anatomicamente relacionados, tem uma tarefa a realizar, como o gesto de oclusão labial²⁴.

A relação entre os articuladores, envolvidos na produção de um gesto, garante o deslocamento de todo o conjunto, mesmo quando um dos articuladores é bloqueado – como, e.g., em experimentos de *bite block*. No caso da oclusão bilabial, conforme Browman & Goldstein, isto garante sua realização de maneiras diferentes: movendo-se apenas a mandíbula ou movendo-se somente o lábio inferior com um pequeno movimento de mandíbula. A flexibilidade articulatória, prevista pela Dinâmica de Tarefa, pode também dar conta de aspectos coarticulatórios, como a produção da oclusão de /b/ com mandíbula mais alta, diante de [i] que de [a]. Nesse modelo, a participação de cada articulador, na realização de um gesto, é especificada em termos de pesos característicos desses articuladores²⁵.

Além de prever como se dá a coordenação entre os articuladores, para a produção de um gesto, Browman & Goldstein necessitam, ainda, modelar a coordenação entre os gestos – em estruturas que denominam “constelações”. A Fonologia Articulatória prevê que a coordenação dos gestos ocorre pela especificação das fases de cada um desses gestos, a qual é, concomitantemente, espacial e temporal, já que a especificação temporal varia em função da frequência da oscilação, envolvida em um dado gesto (ou do parâmetro “rigidez” na equação dinâmica), e a especificação espacial é intrinsecamente relacionada à mudança de amplitude do gesto (ou do parâmetro “posição de repouso”, na equação dinâmica).

Uma representação simbólica dos gestos projeta-os numa pauta gestual²⁶. A dimensão vertical dessa pauta organiza os gestos nos subsistemas articulatórios oral,

²⁴ Como menciono no início deste item, o gesto de oclusão labial é o fio condutor de toda a argumentação de Browman & Goldstein (1986).

²⁵ Desconsiderando-se efeitos coarticulatórios.

²⁶ Tradução para “*gestural score*”, cf. Albano (2001).

laríngeo e vélico. Os gestos do subsistema oral são dispostos no alto da pauta, os do subsistema vélico são dispostos na porção inferior da pauta e os gestos do subsistema laríngeo se encontram entre os dois primeiros. De acordo com Browman & Goldstein, tal representação visa a relacionar a organização gestual a uma organização rítmica da fala, de modo que, quanto mais próximo do topo da pauta gestual estiver um gesto, mas relevante ele é para a organização do ritmo. Assim, declaram os autores,

(...) gestos vocálicos são encontrados na linha mais alta, já que parecem ser os mais importantes na determinação do ritmo da fala, com outros gestos sendo coproduzidos com eles (...). Os gestos vélicos, em contrapartida, são dispostos bem abaixo, porque contribuem muito pouco para a estrutura rítmica. (Browman & Goldstein, 1986: 244. A tradução é minha.)

O contraste estabelecido entre gestos vocálicos e gestos vélicos faz pensar que a importância para o ritmo não está relacionada à menor rigidez (duração) dos gestos, mas à maior ou à menor periodicidade dos gestos vocálicos, hipótese que se reforça diante do postulado de que a cadeia da fala é organizada pelos gestos vocálicos, aos quais se sobrepõem os gestos consonantais, conforme Browman & Goldstein (1990).

Na pauta gestual, a dimensão horizontal é constituída de grades que, assume-se, captam as relações de fase entre os gestos: as linhas das grades representam um intervalo de fases de 90 graus e dois gestos alinhados na mesma grade. A FAR pressupõe que dois gestos estão sincronizados, caso se encontrem na mesma grade. Assim, a pauta gestual para a primeira sílaba da palavra /mpaka/, do chaga, com o encontro nasal silábica + oclusiva, seria (cf. transcrito de Browman & Goldstein, 1986: 244)

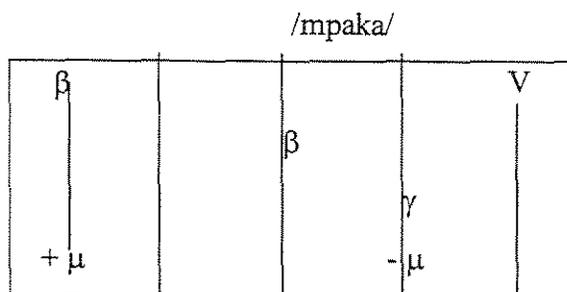


Figura 1.1 – Exemplo de pauta gestual da primeira sílaba da palavra /mpaka/, do chaga, de acordo com a versão inicial da Fonologia Articulatória, e inspirada em Browman & Goldstein, 1986: 244.

Uma interpretação possível para essa pauta gestual é a seguinte: há dois gestos de oclusão bilabial (na primeira e terceira grades), o que garante a distinção entre o encontro nasal silábica + oclusiva do chaga e o encontro nasal-oclusiva do inglês. A relação de fase entre os dois gestos de oclusão bilabial é de 270 graus – pois ambos estão separados por três intervalos de grades e o intervalo entre grades é de 90 graus²⁷. Vê-se, além disso, que o primeiro gesto de oclusão bilabial está sincronizado com um gesto de abertura do véu (+ μ), o que garante a representação da nasal /m/. Ao segundo gesto de oclusão bilabial segue-se um gesto de atividade glotal (γ), sincronizado a um gesto de fechamento do véu (- μ), representando, assim, a oclusiva surda. Nota-se ainda que ao gesto vocálico (V) não se sobrepõe nenhum outro gesto.

Ressalte-se que o posicionamento do primeiro gesto de oclusão bilabial no topo da pauta visa a capturar a função silábica desse gesto, função mais comumente ocupada por vogais, como se pode inferir pelo posicionamento do gesto vocálico também na linha mais alta da pauta – medida que, aliás, vai ao encontro da tese de Browman & Goldstein, enunciada *supra*.

A FAR assume que a pauta gestual é a própria representação de cada item lexical. Há, por conseguinte, uma diferença considerável entre esse e os modelos fonológicos estáticos, porque a FAR considera que a estrutura física de um item deva ser lexicalizada. A relação entre Fonética e Fonologia, então, muda substancialmente, em comparação com os modelos fonológicos estáticos, pois a Fonética é trazida para dentro da Lingüística: são as

²⁷ Considera-se que os gestos sejam constituídos por um ciclo abstrato de 360 graus.

informações fonéticas que organizam e implementam as unidades representacionais, fonológicas, do modelo.

Apesar dessa inovação, é curioso observar um forte apego de Browman e Goldstein aos modelos fonológicos estáticos, nesta primeira apresentação da FAR: a sessão final do texto de 1986 é dedicada a frisar a possibilidade de se empregar o gesto articulatório como substituto do traço, em qualquer modelo fonológico, em especial na Fonologia Autossegmental. Os autores chegam, inclusive, a defender uma proximidade entre os dois modelos: segundo eles, as relações de fase entre os gestos são reminiscências das linhas de associação da Fonologia Autossegmental. Além disso, os gestos poderiam ser organizados hierarquicamente, como os autossegmentos – a proposta de uma organização gestual hierárquica virá na versão seguinte da FAR, com a “Geometria do Tubo”.

A insistência em preservar os modelos fonológicos estáticos talvez se deva à formação de Browman & Goldstein: a influência da Fonologia Gerativa de Chomsky & Halle (1995[1968]) leva Browman & Goldstein a considerarem a estrutura gestual uma alternativa para a transcrição fonética – no sentido de SPE²⁸, que comento no item 1.2.1²⁹. Dizem Chomsky & Halle:

Na transcrição fonética, um enunciado é representado como uma seqüência de unidades discretas, sendo cada uma delas um complexo de traços fonéticos, como vozeamento, sonoridade, altura da língua, etc. A transcrição fonética pode, portanto, ser considerada uma matriz bidimensional, na qual as colunas correspondem a unidades consecutivas e as linhas, a diferentes traços. (1991[1968]: 294. A tradução é minha.)

²⁸ Abreviação para *The sound pattern of English*.

²⁹ Essa consideração é explícita no final do texto de 1986.

1.3.1.2 – O gesto como unidade discreta e organizadora da Fonologia

A insistência de Browman & Goldstein em preservar modelos estáticos continua na segunda versão FAR (1989, 1990)³⁰, com a proposta de uma “Geometria do Tubo”. A despeito disso, Browman & Goldstein buscam consolidar o gesto articulatório como a unidade atômica dos modelos fonológicos em geral, enfatizando seu caráter discreto e estabelecendo-o como unidade pré-lingüística e organizadora da Fonologia³¹. Este é o avanço desta nova versão com relação à anterior e é sobre esse tópico que me debruço neste item.

Uma das premissas da FAR, como já mencionei, é que o gesto articulatório traduz a ação dos articuladores do trato vocal, rumo à produção de uma constrição, num dos subsistemas – oral, laríngeo ou vélico – do trato. Essa ação é a “tarefa” que os articuladores têm de executar, a partir da especificação de parâmetros da equação dinâmica, que transcrevi no item 1.3.1.1.

Subjacente a esta afirmação, parece estar a questão acerca de como emerge o gesto articulatório, questão que toca, em última instância, na aquisição dos gestos. Browman & Goldstein argumentam que os gestos articulatórios surgem na fala de uma criança, antes do desenvolvimento lingüístico – são unidades pré-lingüísticas, portanto. Dizem eles que, mesmo no balbúcio – fase durante a qual a criança começa a produzir as primeiras sílabas – já é possível identificar os gestos articulatórios. Citando um exemplo de Ferguson & Farwell (1975, *apud* Browman & Goldstein, 1989), sobre várias tentativas de uma criança para produzir a palavra inglesa *pen*, Browman & Goldstein notam que, embora a criança não produza a palavra em si, os gestos articulatórios presentes nos “enunciados-tentativa” são os mesmos presentes em *pen*.

Para que a criança produza *pen*, como na fala adulta, são necessários dois passos, segundo os autores: a) diferenciação e ajuste dos gestos individuais. Ou seja, a partir de um único gesto, mais geral, que serve para produzir tanto oclusivas, como fricativas ou

³⁰ Vide nota 4 *supra*, sobre minha decisão de denominar “segunda versão da FAR” aos aspectos do modelo, discutidos por Browman & Goldstein, nos textos de 1989 e 1990.

³¹ Logo, é ainda visível nesse texto uma certa ambigüidade do modelo, a qual se desfaz apenas no texto de 1992.

aproximantes, a criança, em fase de aquisição, deve aprender a distingui-los e contrastá-los também no grau de constrictão³²; b) coordenação dos gestos individuais, pertencentes a uma palavra, quer dizer, não basta que a criança consiga produzir todos os gestos de uma palavra, como no exemplo trazido de Ferguson & Farwell (*op. cit.*). É preciso também que a criança os coordene, para produzir a própria palavra.

Os aspectos da coordenação – observam Browman & Goldstein – podem emergir como resultado da diferenciação de um número crescente de palavras perceptualmente distintas, utilizando-se recursos articulatórios limitados. (1989: 206. A tradução é minha.)

Pode-se dizer que vai de encontro a esse postulado a idéia do “aprender fazendo”, enunciada independentemente por Albano (1990): a criança aprende a realizar as constrictões do trato para produzir um determinado gesto, bem como aprende a coordenar esse com outros gestos, para produzir uma palavra, articulando. A Fonologia, por conseguinte, começa a se estabelecer a partir do momento em que a criança se torna capaz de simbolizar, i.e., de associar alguns gestos coordenados a um significado, formando uma palavra.

Note-se que a proposta da FAR é de que o primitivo, tomado para a descrição e análise, organiza a fala infantil e também a adulta. Metodologicamente, esta proposta é vantajosa, pois permite – como afirmam Browman & Goldstein – que a construção da Fonologia seja monitorada acuradamente, pela observação do desenvolvimento das primeiras estruturas gestuais das crianças em direção à estrutura elaborada dos adultos³³.

Fica claro, na argumentação acerca do gesto como unidade pré-lingüística e organizadora da Fonologia de uma língua, que a sua natureza é distinta da natureza de traços, fonemas ou segmentos. Mas Browman & Goldstein voltam a esse ponto, apenas

³² Frise-se que, para Browman & Goldstein, a aquisição dos gestos articulatórios acompanha a maturação neuromotora da criança. Portanto, essa diferenciação entre gestos e entre aspectos gestuais, como a constrictão, é concomitante ao desenvolvimento neuromotor.

tocado na primeira apresentação da FAR (1986). Eles reafirmam o caráter dinâmico do gesto articulatório, definido no tempo e no espaço. Voltam a observar que os gestos se organizam numa estrutura maior, a constelação, e que a constelação é representada numa pauta gestual. Mas introduzem, agora, as “variáveis do trato”, dimensões de tarefas independentes que especificam o objetivo de um gesto articulatório. Cada variável do trato é associada a um conjunto específico de articuladores. Assim:

Variável do trato	Articuladores envolvidos
PL – protrusão labial	lábios inferior e superior, mandíbula
AL – abertura labial	lábios inferior e superior, mandíbula
LCPL - local de constrição da ponta da língua	ponta e corpo da língua, mandíbula
GCPL – grau de constrição da ponta da língua	ponta e corpo da língua, mandíbula
LCCL – local de constrição do corpo da língua	corpo da língua, mandíbula
GCCL – grau de constrição do corpo da língua	corpo da língua, mandíbula
VEL – abertura do véu palatino	véu palatino
GLO – abertura da glote	glote

Quadro 1 - Variáveis do trato e articuladores envolvidos na realização da tarefa especificada pelas variáveis.

Ressalte-se que, de acordo com Browman & Goldstein, o quadro acima não está completo: faltariam ainda uma variável de raiz de língua, variáveis para controlar a forma das constrições de ponta e corpo de língua³⁴ e variáveis laringeas, para controlar *pitch*, e o movimento vertical da laringe, responsável pela produção de ejetivas e implosivas. Apesar de reconhecerem essa lacuna, entretanto, Browman & Goldstein nunca inseriram tais variáveis na FAR, nem mesmo em versões posteriores do modelo.

Às variáveis, elencadas no Quadro 1, os autores atribuem descritores gestuais para representar os contrastes entre elas, seja de grau, seja de local de constrição. São previstos

³³ Essa é a perspectiva adotada por Gama-Rossi (1999), em sua tese de doutorado, na qual discute o estabelecimento do ritmo em PB, pela comparação entre a fala infantil e a fala adulta.

³⁴ Talvez essas variáveis contribuíssem para o modelamento das líquidas que, como tentarei mostrar no quarto capítulo, é problemático mesmo no arcabouço da FAR.

cinco descritores gestuais para grau de constrição: [fechado], [crítico], [estreito], [médio], [largo]. O descritor [fechado] caracteriza oclusivas, ao passo que [crítico] caracteriza fricativas. [estreito], [médio] e [largo], por sua vez, combinados com descritores de local de constrição, caracterizam diferenças entre vogais, sendo a diferença entre os três descritores determinada pelo contraste articulatorio e perceptual num sistema vocálico. A diferença entre [largo] e [estreito] pode ser empregada ainda, de acordo com Browman & Goldstein, para distinguir o tamanho da abertura glotal, relacionada a consoantes aspiradas e não-aspiradas.

Os descritores gestuais de local de constrição, por sua vez, são atribuídos aos articuladores. Os lábios podem receber especificação [protruso], [labial], [dental]. A ponta da língua recebe especificação [labial], [dental], [alveolar], [pós-alveolar] ou [palatal]. A lâmina, ou corpo, da língua podem ser atribuídos [palatal], [velar], [uvular], [faríngeo]. O modelo prevê que os mesmos descritores de local de constrição podem ser atribuídos ao articulador raiz de língua que, porém, como já mencionei, não é incorporado ao modelo.

Haveria ainda um terceiro descritor gestual, a “forma de constrição”, responsável, hipoteticamente, pela distinção entre gestos envolvendo os mesmos conjuntos de articuladores e os mesmos descritores de grau e local de constrição. Essa variável não foi incorporada a nenhuma versão do modelo³⁵.

Definidos em termos articulatorios, os descritores gestuais são variáveis categóricas do modelo e visam, segundo Browman & Goldstein, a conferir ao gesto um caráter discreto³⁶. O caráter dinâmico do gesto lhe é conferido por descritores também dinâmicos, como o parâmetro “rigidez”, variável em função do grau de constrição, acento e taxa de elocução.

³⁵ Talvez a incorporação de uma variável dessa natureza, como “lateral”, possibilitasse um modelamento satisfatório das líquidas na FAR.

³⁶ Quanto à definição articulatoria dos descritores gestuais, cabe notar que ela é, muitas vezes, impressionística, conforme crítica de Albano (2001). Os próprios Browman & Goldstein parecem reconhecer isso, mas não vão além de observar que a caracterização dos descritores se inspira em autores como Jespersen (1941, *apud* Browman & Goldstein, 1989) ou Pike (1943). Além disso, cf. Albano (comunicação pessoal), a estabilização dessas categorias discretas é passível de modelamento na teoria dos sistemas dinâmicos, mas precisa, para tanto, de dados articulatorios colhidos longitudinalmente no período prelingüístico, o que, com os recursos tecnológicos disponíveis, é impossível obter.

No que concerne aos descritores gestuais, é pertinente frisar que assumem na FAR a função que os traços distintivos desempenham nos modelos fonológicos estáticos. De fato, os autores reconhecem que dois gestos são distintos entre si se diferem em, pelo menos, um descritor. A própria notação dos descritores, entre semi-colchetes, lembra a notação dos traços. A diferença é que os descritores não são binários: apenas estão ou não presentes num gesto, i.e., uma vez presentes, podem apresentar comportamento gradiente.

Sem dúvida que a presença da natureza discreta dos descritores gestuais coloca a FAR num paradoxo, pois, embora se pretenda um modelo fonológico distinto dos demais, com um primitivo de natureza dinâmica – portanto distinta da natureza de segmentos ou traços – a FAR preserva resquícios de modelos estáticos. Parece, no entanto, que Browman & Goldstein fazem questão de que o modelo tenha essa estrutura, nesse momento. Tanto é que, ao introduzirem variáveis do trato e descritores gestuais, os autores alteram a representação do gesto, i.e., a pauta gestual a qual, nesta segunda versão, passa a dispor de camadas articulatórias³⁷, na dimensão vertical, e informação temporal, na dimensão horizontal.

Além disso, conforme Browman & Goldstein, é possível capturar as relações de fase que organizam uma constelação gestual, já que se pode especificar a coordenação espaço-temporal entre os gestos³⁸. Na nova pauta, o gesto articulatorio é representado como um retângulo, cuja dimensão horizontal carrega informação sobre a sua duração, i.e., sobre o tempo durante o qual permanece ativado. Tal notação capta aspectos de sobreposição gestual, algo de grande importância na FAR, que prevê como conseqüências da sobreposição: a) efeitos de invariância acústica, como as transições formânticas que emergem, quando um gesto consonantal invariante se sobrepõe a diferentes gestos

³⁷ As camadas articulatórias, ou *tiers*, são definidas em função de uma “independência articulatoria”, como dizem Browman & Goldstein. Ainda segundo eles, as camadas corresponderiam a organizações propostas tanto por foneticistas como por fonólogos: as camadas de lábio, lâmina e ponta de língua corresponderiam aos agrupamentos de pontos de articulação em três conjuntos maiores: labial, lingual e dorsal. Além disso, cf. Browman & Goldstein (1990: 346), *os fonólogos* (como Clements, 1980, ou Goldsmith, 1976) *sempre propuseram camadas correspondentes a sistemas articulatorios independentes, e.g., a laringe, o véu e articuladores orais.* (A tradução é minha.)

³⁸ Isso pode ser feito sobrepondo-se a trajetória dos articuladores dos gestos. A pauta que transcrevo abaixo não traz, porém, essa informação. Para maiores detalhes, vide Browman & Goldstein (1989).

vocálicos; b) variações alofônicas; c) vários tipos de “coarticulação”, como as formas variadas do trato, dependentes de contexto. Frise-se que outra previsão do modelo diz respeito ao fato de que a sobreposição gestual ocorre, necessariamente, entre gestos de camadas diferentes³⁹.

A pauta gestual incorpora, ainda, informação sobre os descritores gestuais, tanto de grau como de local de constrição. Assim, temos:

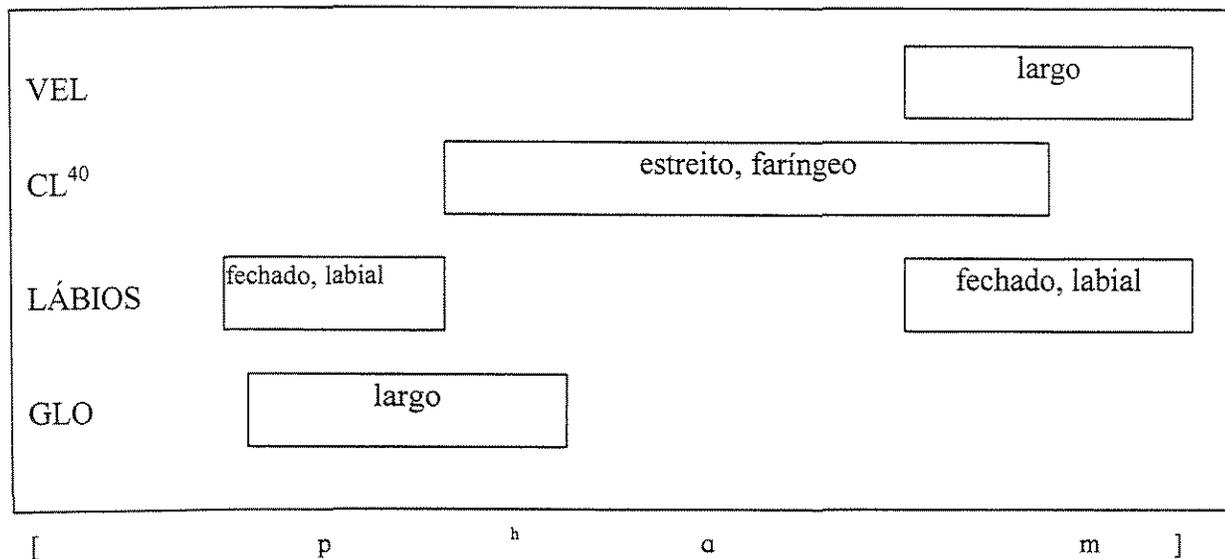


Figura 1.2 – Exemplo de pauta gestual para a palavra *palm* [pam], conforme sugestão de Browman & Goldstein (1989: 212).

Nessa pauta, diferentemente da pauta apresentada na primeira versão da FAR (vide item 1.3.1.1), não há o ordenamento “hierárquico” dos subsistemas do trato, de modo a posicionar o subsistema oral no alto da pauta, o subsistema vélico na porção inferior e o subsistema laríngeo intermediário aos dois primeiros. (Browman & Goldstein não explicam por que abandonaram essa disposição.)

Por outro lado, vê-se que o gesto de corpo de língua permanece ativado por grande parte da duração total de [pam], estando presente na realização da aspiração de [p], na

³⁹ Caso dois gestos da mesma camada se sobreponham, afirmam Browman & Goldstein que haverá, como consequência, perturbação do movimento das variáveis do trato de cada um desses gestos. Nesse caso, acrescentam, os dois gestos competem entre si, pois tentam realizar tarefas distintas com estruturas articulatórias idênticas.

realização da vogal e da consoante nasal. O gesto de oclusão labial é ativado duas vezes, primeiro para a produção da oclusiva bilabial e, em seguida, para a produção da nasal. Observe-se que a especificação dos descritores capta a identidade de classe entre as duas consoantes – ambas envolvem oclusão dos lábios. Observe-se também que não há sobreposição entre o primeiro gesto de oclusão labial e o gesto de corpo de língua; eles apenas se sucedem, coincidindo o *offset* do primeiro com o *onset* do segundo. Por fim, o gesto glotal se inicia logo após o primeiro gesto de oclusão labial e permanece ativado por algum tempo, além do tempo de ativação do gesto de oclusão labial. Esse gesto corresponde à aspiração da oclusiva surda inicial e se estende sobre o início do gesto de corpo de língua, sinalizando a coarticulação entre a aspiração da oclusiva e a vogal.

Browman & Goldstein apresentam, ainda, uma alternativa, arbórea, a essa notação da pauta gestual, tentando aproximá-la da notação da Geometria de Traços. Nessa representação, prescinde-se da informação sobre o tempo de ativação do gesto e, conseqüentemente, não se captam aspectos da sobreposição gestual. Em contrapartida, inserem-se linhas de associação, cuja função seria indicar as relações de fase intergestuais. É possível inferir, aí, uma relação de dominância entre os gestos e o papel de “traços distintivos” dos descritores gestuais⁴¹. Assim, a mesma palavra “*palm*”, pode ser representada da seguinte maneira:

⁴⁰ Abreviação para “corpo de língua”, tradução de *tongue blade*.

⁴¹ De fato, Browman & Goldstein chegam a prever uma organização hierárquica dos descritores gestuais, muito semelhante à estrutura de uma Geometria de Traços e que chegam a denominar “geometria articulatória”. Nessa organização, um “nó” Trato Vocal domina os “nós” Oral, VEL e GLO. O “nó” Oral, por sua vez, domina um “nó” Lábios e Língua, o qual domina um “nó” Ponta de Língua e um “nó” Corpo de Língua. Cada um dos “nós terminais” é especificado para os seguintes descritores: local e grau de constrição, rigidez e forma da constrição – esta última para os “nós terminais” de Língua. VEL, obviamente, não recebe especificação para local de constrição. (Fiz questão de utilizar uma nomenclatura próxima à da Geometria de Traços (GT) para descrever a organização proposta com o intuito de fazer notar a semelhança dela com a da GT.)

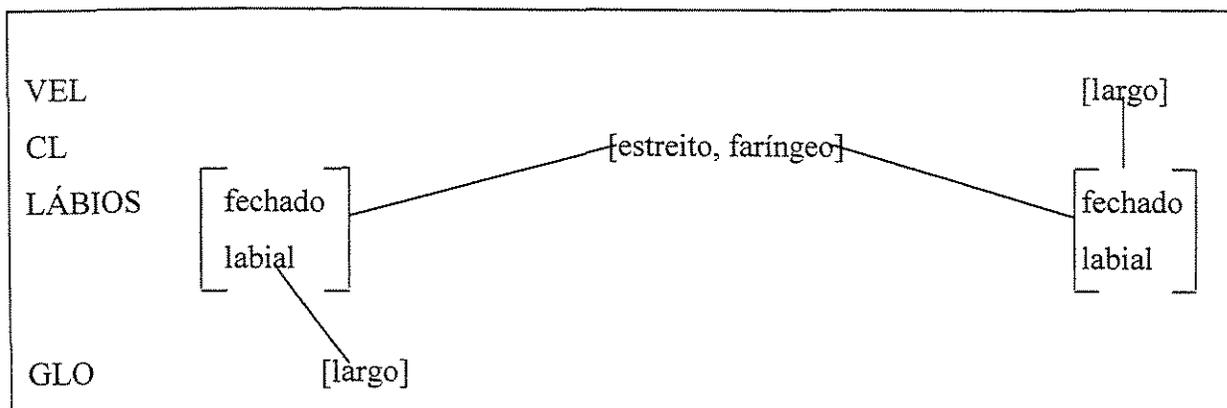


Figura 1.2 – Representação “arbórea” da pauta gestual de [pam], conforme sugestão de Browman & Goldstein (1989: 213).

Segundo Browman & Goldstein, a “vantagem” dessa representação arbórea é que ela permite comparar a organização gestual a uma organização nos moldes de uma Geometria de Traços. A tentativa de aproximar a FAR desse modelo chega ao ponto da proposta de uma “Geometria do Tubo”, que trabalha com a noção de “acoplamento hierárquico” de vários tubos uniformes constitutivos do trato vocal.

Assim, o nó mais alto da hierarquia, “Trato Vocal”, domina um nó “GLO” (atividade glotal) e um nó “Supralaríngeo”. Este, por sua vez, domina os nós “Nasal” e “Oral”. O nó “Nasal” domina um nó VEL (véu palatino); o nó “Oral” domina outros dois nós, “Língua” e “Lábios”. “Língua” domina os nós “Lateral” e “Central”. Por fim, “Lateral” domina um nó “Lâmina da Língua” e o nó “Central” domina os nós “Lâmina da Língua” e “Ponta da Língua”. Esquemáticamente, temos:

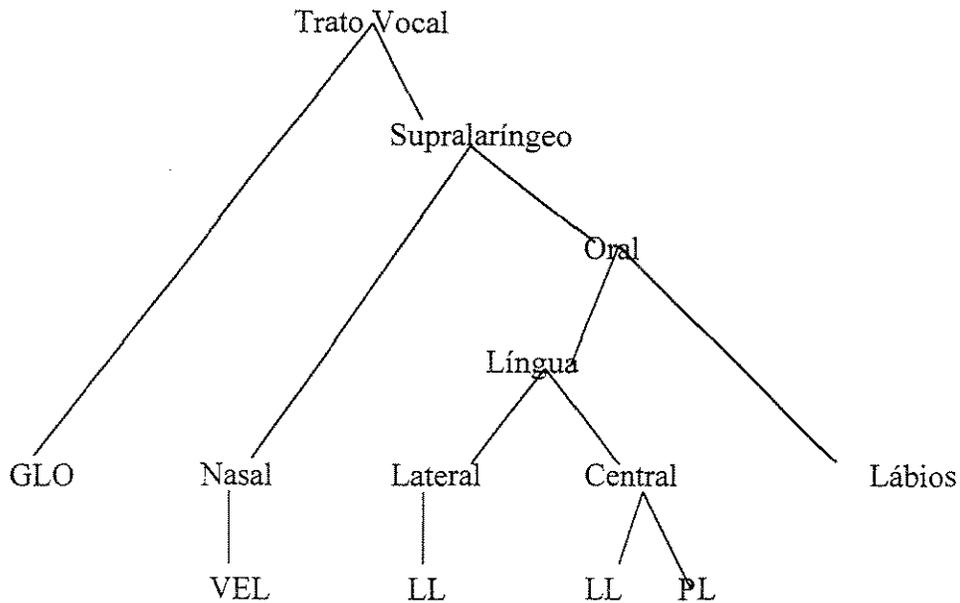


Figura 1.3 - Esquema de "Geometria do Tubo", conforme sugestão de Browman & Goldstein (1989: 236).

É evidente a semelhança estrutural entre a Geometria do Tubo e uma Geometria de Traços. Diferentemente desta, porém, a Geometria do Tubo assume que conjuntos de articuladores se movem dentro dos tubos – nasal, lateral e central – criando constrições. Além disso, prevê-se que o grau de constrição é caracterizado em cada nível e não apenas no nó mais alto da hierarquia. Dizem Browman & Goldstein a esse respeito:

Assim, cada um dos tubos e compostos de tubos da hierarquia terá seu grau de constrição completamente previsível, a partir do grau de constrição dos seus constituintes e, em última instância, do grau de constrição de cada gesto individual. (Browman & Goldstein, op.cit.: 235. A tradução é minha.)

Essa, sem dúvida, seria uma vantagem da Geometria do Tubo sobre uma Geometria de Traços que, como demonstrado por D'Angelis (1998) e como reitero no segundo capítulo, sempre teve dificuldade em lidar com os traços de modo, havendo sempre controvérsia sobre o lugar que tais traços ocupam na estrutura interna do segmento.

Um outro aspecto, que pode ser identificado como tentativa de aproximação entre a FAR e os modelos fonológicos que tomam primitivos de análise estáticos, é a tentativa de Browman & Goldstein de mostrar que seu modelo pode dar conta de fatos da fala corrente⁴², às vezes mal resolvidos por geometrias de traços ou pela Fonologia Autossegmental, como elisões, inserções ou assimilações. Ao invés de prever regras e ordená-las, tentando derivar esses processos, a FAR trabalharia com a variação da magnitude e da sobreposição gestuais.

Em *Tiers in Articulatory Phonology* (1990), Browman & Goldstein tratam dos fatos, que menciono acima, a partir de três exemplos:

/ˈmʌst bi/	—	[ˈmʌsbi]	(“ <i>must be</i> ”)
/hʌndrədˈpʌʊndz/	—	[hʌndrəbˈpʌʊndz]	(“ <i>hundred pounds</i> ”)
/ˈgrʌʊndˈpreʃə/	—	[ˈgrʌʊmˈpreʃə]	(“ <i>ground pressure</i> ”)

Quadro 2 – Transcrição fonológica (primeira coluna) e transcrição fonética larga (coluna do meio), para exemplos, respectivamente, de apagamento de consoante, assimilação de ponto de articulação de consoante e apagamento e assimilação simultâneos, cf. Browman & Goldstein, 1990: 359.

A argumentação dos autores se faz a partir da afirmação de que, enquanto modelos fonológicos tradicionais necessitam de um grande conjunto de regras, para dar conta dos fatos acima mencionados, a FAR pode resolvê-los mais parcimoniosamente, através apenas de mudanças nas pautas dos gestos, envolvidos nesses processos. Logo, a relação entre a representação e a realização de um item lexical é muito mais simples do que em geometrias de traços ou numa Fonologia Autossegmental.

Browman & Goldstein partem de dois pressupostos, para abordar os casos acima: 1) a cadeia da fala se organiza de modo que articulações consonantais se sobrepõem a articulações vocálicas contínuas, as quais, por sua vez, se sobrepõem minimamente; 2) os

⁴² Segundo os autores, a fala corrente é *subconjunto da fala rápida, no qual, tipicamente, ocorrem reduções*. (Browman & Goldstein, 1990: 359. Tradução minha.)

gestos de duas camadas diferentes podem se coordenar entre si, de modo a ficarem total ou parcialmente sobrepostos ou de modo a não se sobreporem.

No caso de “*must be*” ([ˈmʌsbɪ]), prevê-se que, numa situação de fala corrente, o gesto de oclusão labial “escorregue” sobre o gesto de ponta de língua, “escondendo” assim a oclusiva alveolar /t/. O gesto de oclusão alveolar continua presente na pauta, mas é totalmente sobreposto pelo gesto de oclusão bilabial, o que significa que o movimento de ponta de língua, em direção aos alvéolos, ocorre simultaneamente ao movimento de fechamento dos lábios. Numa situação de fala lenta, ter-se-ia uma sobreposição parcial dos gestos de oclusão alveolar e oclusão bilabial, sendo que o primeiro precederia o segundo.

Para “*hundred pounds*” ([hʌndrəbˈpaʊndz]), Browman & Goldstein propõem análise semelhante à anterior, embora a literatura fonológica tradicionalmente descreva este como um caso distinto do primeiro. Assim, o gesto de oclusão bilabial se sobrepõe ao gesto de oclusão alveolar, tornando o primeiro inaudível. Além disso, *a sobreposição do vozeamento sobre o início da oclusão bilabial resulta na transcrição [bp]* (Browman & Goldstein, 1990: 361).

Quanto a “*ground pressure*” ([ˈgraʊmˈpreʃə]), a análise é conduzida no sentido de argumentar que o gesto de oclusão bilabial de /p/ se sobrepõe ao gesto de oclusão alveolar, “escondendo-o”. Como dessa manobra resulta a sobreposição do gesto de abaixamento do véu de /n/ sobre a oclusão bilabial, tem-se, então, a realização de [m].

Vê-se, portanto, que a base para a explicação dos três fatos é a mesma e que, como resultado da sobreposição gestual, têm-se apenas gestos “escondidos”, mas não apagados. Ou seja, embora auditiva – e, às vezes, até acusticamente, segundo Browman & Goldstein – não se perceba um dado som, ele continua presente na cadeia da fala, em termos articulatórios. Os autores provam sua hipótese, através da observação da trajetória dos articuladores na seqüência *perfect memory*, onde, supostamente, é possível a “elisão” do /t/ final de *perfect*, em fala corrente. Observam que, neste caso, haveria sobreposição do gesto de oclusão bilabial de /m/ sobre o gesto de oclusão alveolar de /t/, de modo que este último não é audível e tampouco visível, na forma de onda da seqüência. Entretanto, a análise articulatória confirma sua presença.

Browman & Goldstein prevêem que a sobreposição entre gestos obedeça à organização gestual de uma língua específica, de modo que os gestos de *onsets* e codas silábicos estariam mais sujeitos à sobreposição parcial, o que preserva alguma informação desses gestos. Isso significa que numa seqüência como [pla] não ocorreria sobreposição total entre o gesto de oclusão bilabial e o gesto lateral. Gestos que não pertencem a *onsets* e codas, porém, podem se sobrepor em graus variáveis, inclusive ao ponto de um “esconder” o outro. Tais previsões se baseiam em Mattingly (1981, *apud* Browman & Goldstein, 1990), para quem a sílaba tem um papel não só fonológico ou prosódico, mas também de organizar eventos fonéticos, preservando as características dos seus componentes.

Apesar das evidências de que a FAR pode lidar, de modo mais elegante, com fatos que modelos fonológicos que assumem primitivos estáticos abordariam com um grande conjunto grande de regras derivacionais, Albano (2001) alerta para a ambigüidade da postura inicial de Browman & Goldstein. Tal postura, segundo Albano, mistura operações de natureza simbólica com fatos dinâmicos. Contra essa metodologia, declara:

Dois tipos de interdições tornadas clássicas pela Fonologia Gerativa parecem desautorizar essa mistura [entre simbólico e dinâmico]: a separação de componentes e a limitação ao poderio das regras. (Albano, op.cit: 65)

A maior coerência interna do modelo, para Albano (*op. cit*), é alcançada quando os autores passam a assumir que a FAR dá conta de fatos dinâmicos, como as alofonias. Isso acontece na apresentação mais recente da FAR, exposta em *Articulatory Phonology: an overview*, de 1992.– ou terceira “versão”, para utilizar a minha “classificação”.

1.3.1.3 – O gesto como unidade modeladora de alofonias gradientes

Frise-se que, embora deixando de lado a preocupação de demonstrar a possibilidade de emprego do gesto articulatório em modelos como a Geometria de Traços, Browman & Goldstein não abandonam, de todo, algumas noções, utilizadas em outras versões e que

assemelhavam a FAR a uma Geometria de Traços: é o caso da noção da “Geometria do Tubo”, a qual caracterizaria padrões resultantes de combinações gestuais sobrepostas e também da noção de uma geometria articulatória, organização hierárquica dos gestos, no interior do trato vocal. Browman & Goldstein, entretanto, tocam *en passant* nessas noções, atendo-se a argumentar que a Fonologia Articulatória pode dar conta de fatos gramaticais que outros modelos fonológicos não abordam. É o caso das alofonias gradientes ou de gradiência manifesta em lapsos.

Conforme Browman & Goldstein (1992), a razão por que o tratamento desses casos é possível na FAR está no fato de que esse modelo analisa *inputs*, enquanto modelos fonológicos estáticos analisam *outputs* no processamento da cadeia da fala. Dito de outro modo, a grande diferença entre FAR e modelos estáticos é que o primeiro modelo tem caráter representacional e implementacional, enquanto os do segundo grupo são derivacionais, i.e., a partir de uma forma acústica ou articulatória, eles procuram levantar regras que, ordenadas, expliquem o engendramento serial de cadeias de representações, a partir de uma estrutura abstrata⁴³. A FAR, em contrapartida, busca explicar como se dá a construção de uma dada estrutura fonológica, a partir de uma unidade básica, o gesto articulatório.

Nessa concepção do gesto articulatório como unidade organizacional da Fonologia de uma língua, dois pressupostos do modelo explicariam os casos de alofonias contínuas: a sobreposição e a magnitude dos gestos os quais, a depender de seu grau, podem resultar em diferentes efeitos acústico-articulatórios, e.g., a sensação de elisão de um segmento ou a percepção do enfraquecimento do mesmo, na cadeia da fala.

Assim, cito aqui dois – dentre vários – casos de alofonias gradientes, comentados por Browman & Goldstein como exemplo de fatos com os quais a FAR consegue lidar, diferentemente dos modelos fonológicos estáticos.

O primeiro deles concerne à “alofonia” das consoantes oclusivas aspiradas e não-aspiradas, em início de palavra no inglês. Estudos como o de Kahn (1976, *apud* Browman

⁴³ Trata-se, portanto, de uma metodologia que vem desde Chomsky & Halle (1995[1968]), como comento no item 1.2.1 deste capítulo.

& Goldstein, 1992), afirmam a ocorrência de um alofone aspirado, em início de palavra, ou em posição medial, antecedendo uma vogal acentuada⁴⁴. O alofone não-aspirado ocorreria nos encontros [s]-oclusiva ou em posição medial na palavra, antecedendo vogais átonas. Kahn (*op.cit.*) atribui um traço [+ spread glottis] para as aspiradas e um traço [- spread glottis] para as consoantes não-aspiradas. Trata-se portanto de uma descrição categórica para a aspiração das oclusivas, no inglês (cf. Browman & Goldstein, 1992).

Entretanto, Browman & Goldstein observam, com base em estudos articulatórios, conduzidos por Cooper (1991, *apud* Browman & Goldstein, 1992) haver atividade do gesto glotal, em todos os casos acima mencionados, e não apenas naqueles em que a análise categórica esperaria a realização de uma consoante aspirada. Cooper teria observado, ainda, a existência de efeito de acento e posição da consoante na palavra sobre a magnitude do gesto glotal, de modo que a posição inicial favorece gestos maiores. Na posição medial átona, ocorreriam os menores gestos de atividade glotal.

É possível afirmar, então, a existência de um contínuo físico pelo qual se estende a atividade do gesto glotal, de modo que magnitudes menores desse gesto seriam percebidas como ausência de aspiração e maiores magnitudes propiciariam a percepção de aspiração. Esse contínuo seria condicionado pela posição do acento e pela posição da oclusiva na palavra. Note-se que a FAR dá conta desse fato, de maneira parcimoniosa e elegante, trabalhando apenas com a noção da magnitude do gesto glotal, em função da posição da oclusiva na palavra e relativamente ao acento.

Uma análise categórica como a de Kahn (*op. cit.*) não conseguiria prever a existência do contínuo físico existente na “alofonia” da consoante aspirada, mesmo que considerasse a existência de uma grande interação entre fatores acentuais e posicionais. Está claro que, mesmo lançando mão de inúmeras regras, tal análise falharia em contemplar todos os fatos verificados.

Um outro caso de alofonia gradiente, comentado por Browman & Goldstein em *Articulatory Phonology*, concerne à variação de pronúncia da lateral /l/ em inglês. Tradicionalmente, reconhecem-se dois alofones posicionais, o de início de palavra/sílaba,

⁴⁴ Nesse caso, a análise de Kahn (*op. cit.*) considera a oclusiva como um segmento inicial de sílaba.

ou *light* /l/, e o de final de palavra/sílaba, ou *dark* /l/. Sproat & Fujimura (1993), porém, conduzem um estudo articulatório, no qual manipulam a força de fronteira adjacente à lateral não inicial. A força de fronteira obedece a uma “escala”, indo da maior – no caso em que ela se estabelece entre vocativo e sentença (e.g. *Beel, equate the actors*⁴⁵) – até a menor – no caso de fronteiras morfológicas, que se estabelecem entre um radical e um afixo (como em *feel-ing*⁴⁶). À guisa de parênteses, é interessante notar que Lehiste (1964) já observara uma variante “intermediária” à realização *light* e à *dark*, justamente nos casos de fronteira morfológica, a partir da descrição acústica da lateral.

Voltando ao estudo de Sproat & Fujimura: os autores observam inicialmente que a lateral se caracteriza por dois gestos, um de ponta e outro de dorso de língua. A diferença entre a variante *light* e a *dark* é que, na primeira, o *tip delay* – i.e., a diferença temporal entre a projeção anterior máxima da ponta da língua e a retração e abaixamento máximos do dorso da língua – é menor do que para a segunda.

Quando, porém, a lateral é inserida nas sentenças que manipulam a força de fronteira adjacente a ela, Sproat & Fujimura verificam que a lateral tende a *dark* quanto mais forte é a fronteira adjacente. Não há, portanto, uma variante *dark* e outra *light*, mas um contínuo físico que se estende entre essas duas categorias e que é condicionado pela força de fronteira.

Embora a interpretação de Sproat & Fujimura (*op.cit.*) não se dê em termos da Fonologia Articulatória⁴⁷, Browman & Goldstein comentam que é possível interpretar esses dados, em termos de mudanças na organização gestual, em função de sua posição na palavra. Dizem que, em posição não-final, as pautas dos gestos de ponta e dorso de língua estariam mais proximamente sincronizadas do que em posição final. Além disso, em

⁴⁵ Em Sproat & Fujimura (1993). Para outros exemplos de dados empregados na manipulação da força de fronteira adjacente à lateral, vide a mesma referência.

⁴⁶ Utilizo o traço para marcar a fronteira entre radical e afixo verbal.

⁴⁷ Na verdade, os autores defendem que o gesto dorsal tem característica vocálica – já que não produz obstrução extrema do trato – e que o gesto de ponta de língua tem característica consonantal – pois produz maior obstrução do trato. Assim, variantes tendendo a *dark* seriam vocálicas, pois a retração dorsal é mais significativa do que o gesto apical. Variantes tendendo a *light*, ao contrário, seriam consonantais, já que a atividade da ponta da língua é maior que a do dorso. Para uma resenha mais detalhada desse estudo, vide Silva (1996).

posição final, o gesto de constrição oral mais estreita – neste caso das laterais, o gesto de ponta de língua – se retarda significativamente.

Interessante é verificar o paralelismo que Browman & Goldstein identificam entre a lateral e a nasal: dados de Krakow (1989, *apud* Browman & Goldstein, 1992) apontam para diferentes coordenações dos gestos de oclusão bilabial e abaixamento do véu, durante a produção da nasal /m/, em função da posição que esse som ocupa na sílaba/palavra. Segundo Browman & Goldstein, nas nasais de início de palavra (como na seqüência *see more*) o final do gesto de abaixamento do véu é sincronizado com o final do gesto de oclusão bilabial. Em final de palavra (como em *seem ore*), o final do gesto de abaixamento do véu precede consideravelmente o final do gesto de oclusão bilabial (em 100 a 350 ms, de acordo com Browman & Goldstein). Sobre as semelhanças entre o comportamento da nasal e o comportamento da lateral no inglês, Browman & Goldstein afirmam que elas apontam para um comportamento mais geral da língua, o qual se relaciona à posição dos segmentos na palavra ou na sílaba. Acrescentam que:

*Visualizando esses comportamentos em termos gestuais, é possível sugerir uma eventual generalização mais ampla, especificamente a existência de um padrão organizacional único para final de sílaba, no qual as maiores contrações sempre precedem as menores contrações*⁴⁸. (Browman & Goldstein 1992: 167. Tradução minha.)

Parece-me que os exemplos relativos ao comportamento da nasal /m/ e da lateral /l/, adjacentes a fronteiras, envolvem um caso de diminuição gradual na magnitude de um gesto, seja o gesto de ponta de língua – para a produção da lateral – seja o gesto de oclusão bilabial – para a produção da nasal, e a conseqüente sobreposição dos gestos coocorrentes, i.e., o gesto de dorso de língua da lateral e o gesto de abaixamento do véu da nasal. Parece-me ainda interessante ressaltar que os gestos cuja magnitude é diminuída, em ambos os

segmentos, são gestos mais balísticos⁴⁹ do que os gestos cuja magnitude é preservada. Talvez esse fato permita ainda uma outra generalização: a de que gestos balísticos tendem a ser enfraquecidos, em posição de final de sílaba/palavra⁵⁰.

Há, no entanto, fatos dialetais que – de acordo com Browman & Goldstein – apontam para diferenças no grau de sobreposição gestual, as quais resultam na presença de um segmento para um dialeto e não para o outro. É o caso da oclusiva epentética que emerge em palavras inglesas como *tense* (realizada como [ten^hse], segundo descrições de Anderson, 1976 e Ohala, 1974, *apud* Browman & Goldstein, 1992), mas não emerge no inglês sul-africano. Os autores sugerem que a sobreposição gestual pode ser maior num dialeto do que no outro e que a coordenação entre os gestos pode ser maior num dialeto do que no outro. Essa questão, porém, é deixada em aberto e, tratada dessa maneira, poderia trazer o inconveniente de que a FAR tivesse de prever organizações gestuais distintas, dependentes de dialetos. Isso não é tão problemático para uma língua com poucos dialetos, mas para uma língua com vários deles, certamente é um empecilho para o modelo. A não ser que se prevejam – cf. Albano (comunicação pessoal) – ligeiras diferenças de sincronização entre os gestos.

1.4 – A Fonologia Acústico-Articulatória

Uma das maneiras de se avançar na solução do problema, suscitado pela FAR, acima referido, é lexicalizar as pautas gestuais similares, mas de resultados acústicos distintos, correspondentes, portanto, a variantes dialetais, como sugere Albano (2001), em sua Fonologia Acústico-Articulatória – doravante FAAR⁵¹.

⁴⁸ Seria interessante verificar se essa generalização é apenas específica do inglês ou ocorre também em outras línguas, como o português. Talvez se pudesse propor, a partir daí, uma explicação para a vocalização histórica da lateral de final de sílaba/palavra.

⁴⁹ Ou seja, gestos de curta duração temporal.

⁵⁰ Se assim for, poderemos, então, explicar a não realização do *tap*, em final de palavras do PB, já que esse som se caracteriza também por um gesto balístico de ponta de língua.

⁵¹ A proposta de lexicalização das variantes dialetais não se constitui um problema para a teoria, já que, tratando-se tanto a FAR como a FAAR de modelos implementacionais, não se coloca a questão do tamanho do léxico, como se colocaria num modelo derivacional, a exemplo da Fonologia Gerativa Padrão, para a qual o tamanho do léxico tem uma relação direta com a rapidez e a dificuldade do processamento.

O argumento de Albano (*op. cit.*) se baseia na observação de casos de “alofonias” do PB, como a africacão de [t,d]⁵² e tem desdobramentos sobre a própria organização do modelo, no sentido que a autora propõe: a) a dissociação entre as trajetórias dos descritores gestuais de ponto e de modo de articulação; b) a organização da pauta gestual, em termos de regiões acústico-articulatórias, nas quais se projeta o gesto articulatório e não em termos dos conjuntos de articuladores, preconizados pela FAR.

Albano examina, acusticamente, dados de um informante não-chiante (assim denominado por ela ao indivíduo que não produz a africada diante de [i], embora palatalize [t,d] nesse contexto) e, em seguida, compara-os a dados produzidos por um informante chiante.

A autora nota – a partir do exame de espectrogramas e de espectros FFT da seqüência V[t]V – que um indivíduo dito não-chiante pode “chiar” em contextos nos quais [t,d] ocorrem entre [i]s. O exemplo de Albano, concernente ao par “vítima/vitima”, mostra que o ruído fricativo, produzido em seguida à oclusão de [t], tem a mesma duração em “vitima” e em “vítima”, o que descarta um efeito da posição acentual sobre os padrões acústicos dos segmentos em questão. A preocupação de Albano, porém, não é explicar esse padrão fixo, mas dar uma interpretação gestual ao efeito acústico obtido e observar que tanto ouvintes treinados como os ouvintes não treinados em transcrição fonética percebem o [t] de “vítima” como uma africada e o [t] de “vitima” como uma oclusiva ou, no máximo, uma oclusiva palatalizada [t,].

Uma interpretação para a produção da africada por um indivíduo “não-chiante”, à luz da FAR de Browman & Goldstein (*op. cit.*), apontaria, conforme Albano, para a sobreposição dos gestos vocálico e consonantal⁵³, de modo que

(...) parece que o [i] “escorrega” de maneira fixa sob a consoante, embora, conforme esperado, só se prolongue muito além dela quando é acentuado. (Albano, 2001: 72)

⁵² Ressalte-se que essa alofonia tem uma interpretação gradiente: considera-se tradicionalmente que um dialeto exibe ou não as africadas [tʃ, dʒ].

Tal hipótese é corroborada pela análise espectral do sinal dos dois enunciados, a qual revela que o ruído fricativo concentra energia, nas mesmas faixas de frequência da vogal seguinte a ele.

Albano considera ainda que percepção da mesma sobreposição gestual como africacão em “vítima” e ausência dela ou, no máximo, palatalização, em “vitima”, poderia ser explicada, por sua vez, com base em fatores auditivos: ela cita experimentos de Stevens (1985, *apud* Albano, *op. cit.*), que observou em fala sintética uma fronteira categórica entre [s] e [ʃ] em torno da diferença zero, relativamente ao F3 da vogal. Ou seja, quando o F3 da vogal é mais intenso do que o ruído na mesma faixa de frequência, os estímulos sintéticos são ouvidos como [s]; quando, porém, o F3 é menos intenso do que o ruído, os estímulos são ouvidos como [ʃ]. É o que parece ocorrer nos casos descritos.

Em seguida, comparando os dados do informante não-chiante com os dados de um informante carioca, “chiante”, Albano nota primeiramente que, embora as palavras “vítima” e “vitima” tenham duração igual para os dois informantes, nos dados do falante não-chiante, o ruído da africada no par “vítima/vitima tem duração fixa, variando a duração do silêncio, conforme a tonicidade. Ao contrário, nos dados do falante chiante, o silêncio tem duração fixa, enquanto a duração do ruído se apresenta variável.

Some-se a isto a constatação de que o ruído fricativo, para o informante chiante, tem amplitude numa faixa de frequência superior a 3,5 kHz, enquanto o ruído fricativo para o informante não-chiante concentra amplitude abaixo dessa faixa de frequência. Citando mais uma vez o estudo de Stevens (*op. cit.*), Albano afirma que a implementação da fricacão depende de um aumento de estridência. Estes achados levam-na a apontar para a impossibilidade de propor uma única pauta gestual para os dois informantes – como previsto na FAR –, pois as diferentes estruturas acústicas devem corresponder a manobras articulatórias também distintas. Assim, uma das pautas coerentes com a FAR, para representar africadas, e que Albano toma como hipótese nula, em seu estudo, é descartada pela autora. A pauta em questão é:

⁵³ Correspondentes, respectivamente, aos gestos de dorso e de ponta de língua.

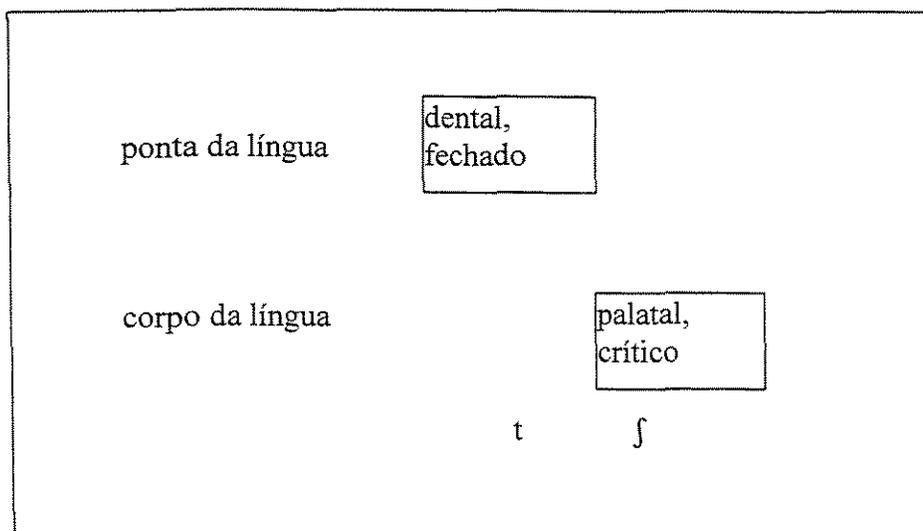


Figura 1.4 – Proposta de Albano (2001: 70) para pauta gestual de [tʃ]: a africada se caracteriza pela sobreposição dos gestos consonantal (ponta de língua) e vocálico (corpo da língua).⁵⁴

O abandono dessa pauta requer uma outra solução, para representar as africadas. Albano comenta, então, que uma saída simples seria prever dois graus distintos de sobreposição gestual, mas, empiricamente, essa solução não se verifica. Ao contrário, a sobreposição gestual prevê, falsamente, de acordo com Albano que, na versão chiante, haja um encurtamento da sílaba, como resultado de o gesto vocálico estar mais “escondido” sob o consonantal. Prevê também que, tanto nos dados do informante chiante como nos dados do não-chiante, *o grau de constrição seja uma média entre o da consoante e da vogal* (ibid: 80).

Tais constatações levam Albano a propor que a única maneira de diferenciar as duas versões da africada seria dissociar a duração das trajetórias das variáveis do trato para grau e local de constrição. Certamente, esta manobra diverge do arcabouço teórico da FAR, que considera indissociáveis as variáveis do trato que caracterizam o gesto articulatório. Albano reconhece essa divergência, mas a toma como uma premissa da Fonologia Acústico-Articulatória (FAAR).

⁵⁴ A Fonologia Articulatória permite ainda uma outra interpretação possível para a africada, que a toma como um segmento de contorno, no qual os gestos de ponta e dorso de língua têm *onsets* e *offsets* coincidentes. Essa pauta é descartada por Albano, por ser, segundo ela, contra-intuitiva. *No caso das africadas que emergem em*

Para Albano, dissociar a duração das variáveis do trato, relacionadas a grau e local de constrição, significa “relaxar” uma restrição da FAR. Segundo ela,

As durações das trajetórias das variáveis, ligadas ao grau e ao local de constrição, podem ser independentes, embora tendam, não obstante, no caso mais geral, a ser coincidentes. (Albano, 2001: 81)

Seguindo essa proposta, diz Albano que a versão chiante e a versão não-chiante para a palatalização de [t,d] diante de [i] seriam diferenciadas pelas durações de ativação dos graus de constrição da consoante e da vogal. Albano defende que a versão não-chiante seja representada pela pauta gestual, transcrita na Figura 1.4 *supra*. Para a versão chiante, porém, propõe uma pauta na qual o gesto de oclusão da consoante é encurtado e a constrição da vogal é adiada (não obstante, preserva-se a sobreposição entre os dois gestos). Assim, a sugestão de Albano é:

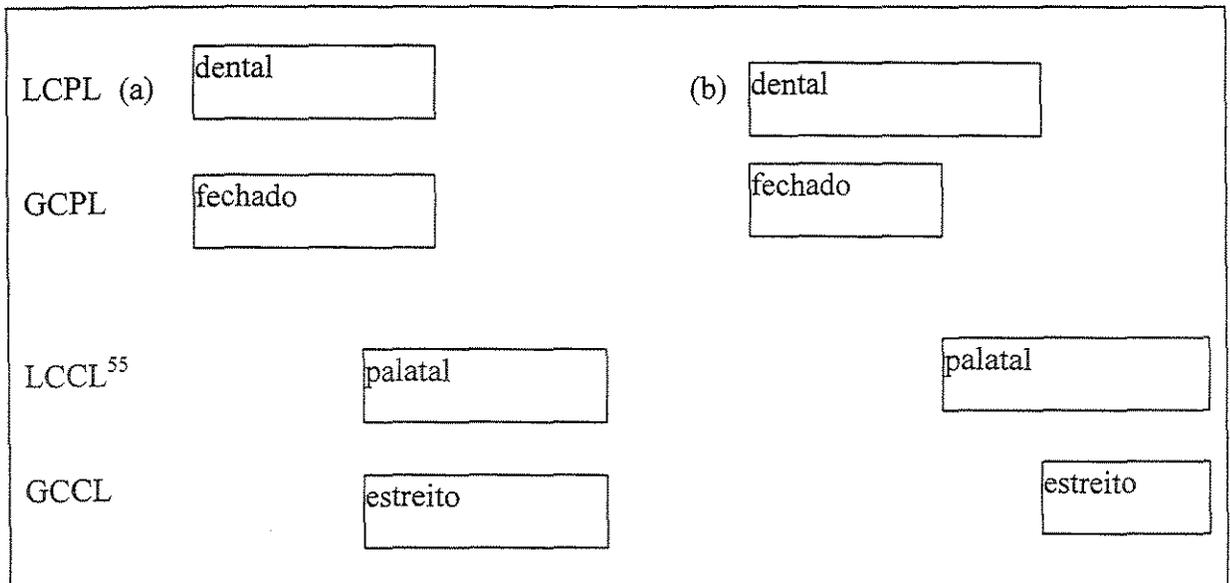


Figura 1.5 - Pautas gestuais sugeridas por Albano (*op. cit.*: 82) para a variante chiante e a variante não-chiante da oclusiva alveolar palatalizada, considerando-se a dissociação das durações das variáveis para grau e local da constrição.

contato com [i], seria muito mais intuitivo conceber a fricção como um efeito mecânico da sobreposição dos gestos consonantal e vocálico, diz ela (Albano, 2001: 69).

⁵⁵ Albano denomina “corpo da língua” à mesma estrutura que eu denomino, em outros itens deste capítulo, “dorso da língua”

Apesar de reconhecer que a dissociação entre as variáveis promove enfraquecimento da FAR – embora essa manobra seja empiricamente sustentável, como evidenciam os dados expostos – Albano defende a idéia de que a lexicalização de diferenças entre pautas gestuais similares pode captar mudanças em curso, num determinado dialeto, como hipoteticamente a pronúncia de [t,d] não chiada, diante de [i], que passa a ser chiada, nesse mesmo contexto.

Vê-se, portanto, que Albano sustenta a idéia de um modelo não-derivacional⁵⁶, no qual inexistente uma forma básica, a partir da qual se engendram as demais, variantes. Ao contrário, todas as variantes devem ter as respectivas pautas gestuais lexicalizadas.

Essa medida resulta num outro ponto, fundamental, da FAAR, que é a tese da comensurabilidade, a qual defende a consideração de processos fonéticos no nível fonológico, fundindo-se os dois níveis.

Ressalte-se que Browman & Goldstein também concebiam um único módulo de processamento fônico, mas, para eles, a Fonologia era descartada desse módulo, sendo os processos fônicos considerados sempre numéricos, muito embora o gesto tenha, como já mencionei, no item 1.3.1, um caráter discreto que lhe é conferido, e.g., pelos descritores gestuais. É possível inferir essa concepção a partir de um fluxograma (Browman & Goldstein, 1990; 1992), que descreve o modelo computacional no qual o gesto articulatório é implementado. Aí, vê-se que a entrada do sistema, ou “enunciado visado” é enviada a um “modelo lingüístico gestual” de onde resulta uma pauta gestual. Esta, submetida a um Modelo Dinâmico de Tarefa, tem como resultado trajetórias articulatórias as quais, entrando num modelo de trato vocal resultam, finalmente, no sinal de fala.

Albano, por outro lado, reconhece ser necessário prever uma Fonologia, mesmo no arcabouço de um modelo dinâmico. Considerar apenas a Fonética não permitiria ao modelo capturar a relação entre processos fônicos e os fatos gramaticais que os condicionam – seja a posição acentual – como no caso da africacão de [t,d], descrito acima –, seja a força de fronteira – como no caso do estudo de Sproat & Fujimura (1993), mencionado no item

⁵⁶ A assunção de um modelo implementacional – e não derivacional – aproxima a Fonologia Acústico-Articulatória da Fonologia Articulatória. Uma diferença marcante entre as duas perspectivas, porém, está no fato de que a FAAR prevê que todas as variantes dialetais devam ter suas pautas gestuais estocadas no léxico

1.3.1.4. A Fonologia, por conseguinte, tem, na Fonologia Acústico-Articulatória, o papel de distinguir o simbólico do dinâmico sem, contudo, separá-los nos níveis propostos pela Fonologia Gerativa. Na verdade, o simbólico emerge, a partir da estabilização do dinâmico. Essa é a base do que Albano denomina “tese da comensurabilidade”.

Outros processos fônicos do PB, como a elisão e a alofonia de fricativa final de coda no PB, exibem um aspecto gradiente e um aspecto categórico, como mostra Albano. Segundo a FAAR, apenas a comensurabilidade do gesto articulatório é capaz de capturar essas duas facetas do processo.

Para garantir a comensurabilidade, Albano defende ser necessária uma especificação acústico-articulatória do gesto. Seu argumento parte de uma decisão de Browman & Goldstein, para ela, controversa. Tal decisão relaciona-se à atribuição da mesma duração a cada dupla de variáveis, e.g., GCPL (grau de constrição de ponta de língua); LCPL (local de constrição de ponta de língua). Como observa Albano, essa manobra se justifica teoricamente, mas é conflitiva com a possibilidade de definir uma equação dinâmica para as duas variáveis.

A justificativa teórica para a atribuição da mesma duração a duas variáveis é a seguinte: na FAR, os articuladores unem as variáveis relacionadas ao grau de constrição às variáveis, relacionadas ao local de constrição. As variáveis é que definem uma constrição e, conseqüentemente, um gesto articulatório. Dado que as variáveis se unem para formar a constrição, não haveria como dissociar os tempos de suas trajetórias. Além disso, conforme previsto pela Geometria Articulatória, um comando que ativa o domínio de um articulador domina, por conseguinte, tudo o que está sob esse articulador.

Para Albano, a atribuição de mesma duração a duas variáveis é elegante, porém insuficiente para lidar com o tipo de fenômeno que a teoria pretende tratar. Isso porque a FAR abre brecha para se atribuir uma equação dinâmica a cada variável. Como resultado, teríamos uma especificação do parâmetro rigidez por equação, o que significa que um mesmo gesto teria duas especificações de rigidez.

do falante, enquanto a FAR prevê que possíveis variações dialetais são o resultado de mecanismos de sobreposição gestual.

Por conta dessa inadequação empírica da FAR, Albano sugere que um espaço acústico-articulatório deva substituir o articulador, na tarefa de agrupar os comandos, para a execução de um mesmo gesto⁵⁷. O espaço acústico-articulatório, para Albano, é definido por regiões discretas do trato vocal, considerado um tubo acústico, tal como sustentado pelo Modelo de Regiões Distintivas (Carré & Mrayati, 1990; Carré, 1994, *apud* Albano, 2001). O MRD, segundo Albano, divide o trato vocal em oito regiões discretas com componentes acústicos distintos e equivalentes aos descritores gestuais.

Para Albano, a incorporação do espaço acústico-articulatório, na representação de um gesto, em substituição ao articulador,

(...) permite que as defasagens dos movimentos ao longo das coordenadas do plano sagital sejam melhor incorporadas ao modelo. Os móveis “local” e “grau de constrição” podem deslocar-se por esse plano em perfeita ou imperfeita sincronia. Nada há de contraditório numa assincronia parcial entre eles, já que um produz efeitos acústicos de ressonância e outro, efeitos acústicos de amplitude relativa e periodicidade (...), que nem sempre estão sincronizados no sinal acústico. (Albano, 2001: 99)

A FAAR prevê que duas variáveis que se deslocam pela mesma região do trato produzam o sinal acústico característico de um gesto, senão otimamente distintivo, pelo menos suficientemente distinto dos demais gestos articulatórios.

Representar um gesto, em termos de espaços acústico-articulatórios, pode, segundo Albano, resolver fatos com os quais a FAR tem dificuldade em lidar, como o [ʰ] intrusivo em palavras inglesas como *else* [ɛʰs]. Ressalte-se que Albano tentara representar sumariamente esse fato, antes, introduzindo na pauta gestual a variável forma

⁵⁷ É preciso esclarecer que a introdução do aspecto acústico, na Fonologia Acústico-Articulatória, tem por objetivo tornar mais precisa a constituição dos gestos articulatórios: nem sempre é possível inferir as características acústicas de um som, a partir de suas características articulatórias. Quer dizer, um mesmo efeito articulatório pode ter mais de uma realização acústica. Assim, definir os gestos que formam um determinado som, em termos de um espaço acústico-articulatório, possibilita contemplar todas as possibilidades de realização acústica de uma dada manobra articulatória.

de constrição, prevista, mas não implementada pela FAR. Sugere, entretanto, que a representação da pauta gestual desse item lexical, em termos de espaços acústico-articulatórios, capta um outro aspecto da questão, pois prevê que trajetórias co-ocorrentes de variáveis do trato podem produzir efeitos acústicos típicos da intrusiva. Eis, então, a pauta simplificada, proposta por Albano (note-se que ela não diferencia a fase [t̚] da fase [t̚] do gesto):

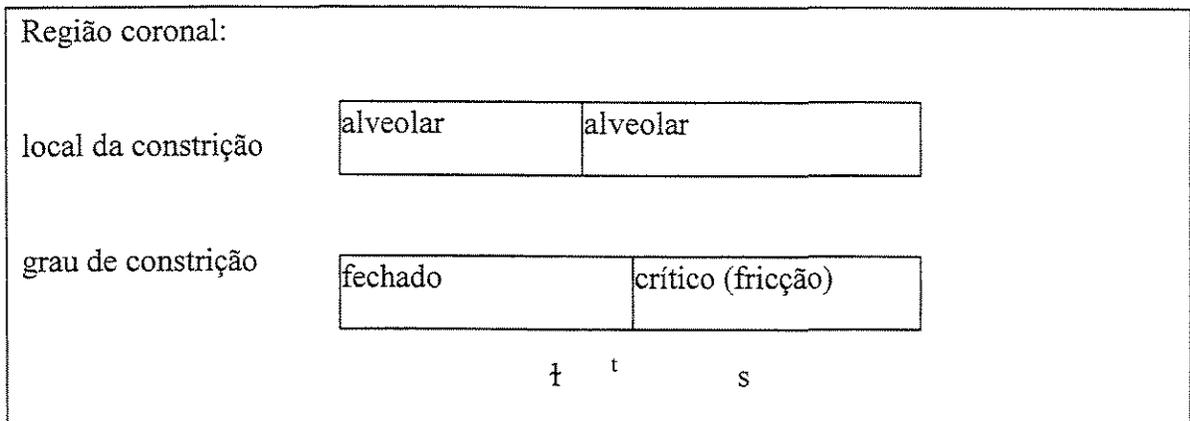


Figura 1.6 - Proposta de pauta gestual para *else* [ɛʔs], considerando especificações para variáveis de espaços acústico-articulatórios, conforme Albano (2001:100).

A interpretação para essa pauta seria a seguinte, conforme Albano: após a produção da lateral velar e enquanto o trato se vai obstruindo, para produzir uma fricativa estridente, a persistência do fechamento do trato introduz uma africação. Essa africação causa uma elevação abrupta da amplitude do ruído, no momento da soltura. Obviamente que essa interpretação só faz sentido dissociando-se a duração das variáveis do trato e substituindo-se os articuladores pelos espaços acústico-articulatórios.

Albano acrescenta que o gesto articulatório se instaura, nesse arcabouço teórico, a partir de um aprendizado do falante, para associar grau e local de constrição, visando a produzir um som com características acústicas estáveis. Esse aprendizado é seguido da experiência com situações propícias à antecipação da fricativa e, finalmente, o aprendizado termina com um processo que envolve toda a geração e grupo social do falante e que fez da oclusiva intrusiva uma marca estilística de alta probabilidade de ocorrência e, por isso, foi

lexicalizada. (Ressalte-se que a FAAR desenvolve uma tese que está implicada na escolha de unidades dinâmicas pela FAR, a saber: fatos fônicos com alta probabilidade de ocorrência são lexicalizados.)

1.5 – Considerações finais

A lexicalização das variantes dialetais e a caracterização dos gestos articulatórios, em termos de espaços acústico-articulatórios, como preconizado pela FAAR, pode dar uma grande contribuição ao estudo dos róticos iniciais em PB.

Como mostrarei, no terceiro capítulo, alguns dados revelam que a variabilidade de pronúncia dos róticos iniciais não é categórica, isto é, não se resume à presença de vibrantes alveolares e fricativas velares, na fala de um mesmo indivíduo, mas é gradiente, estendendo-se ao longo de um contínuo físico, tal que é possível verificar, na fala de um sujeito, variantes que se realizam como vibrantes, propriamente, outras que são fricativas típicas e, ainda, variantes intermediárias a esses dois segmentos, em diferentes graus.

Essa alofonia, que é prosodicamente condicionada, só pode ser tratada por modelos dinâmicos. Modelos fonológicos estáticos, como aqueles aos quais me refiro nos itens iniciais deste capítulo, a saber, a Fonologia Gerativa ou a Geometria de Traços não dão conta desse processo fônico por dois motivos: a) a natureza categórica do primitivo de análise adotado não permite que se contemplem fatos gradientes; b) a dissociação entre Fonética e Fonologia⁵⁸ é tal que fatos como esses são considerados como meros resultados implementacionais, apesar das evidências em contrário. (Disto tratarei mais detidamente no próximo capítulo.)

Uma abordagem dinâmica dessa alofonia contínua requer que se definam os gestos articulatórios que resultam nos efeitos acústicos observados. Como consequência, é preciso especificar as variáveis do trato que atuam na caracterização dos gestos. A FAR, entretanto, não implementou a variável forma de constrição, embora a mesma seja prevista. Por outro

⁵⁸ Como menciono no início deste capítulo, a dissociação entre Fonética e Fonologia decorre do fato de que se tomam como unidades de análise primitivos de natureza estática.

lado, considerar os espaços acústico-articulatórios, característicos dos gestos de cada um desses sons, como preconiza a FAAR, talvez possibilite não apenas a diferenciação entre eles, mas também a captura da sua identidade de classe.

2 - OS LIMITES DE UMA REPRESENTAÇÃO DAS LÍQUIDAS NA TEORIA DOS TRAÇOS DISTINTIVOS: DE *FUNDAMENTALS* À GEOMETRIA DE TRAÇOS

Ao mesmo tempo em que são concebidos como entidades psicológicas, os traços são definidos em termos de padrões específicos de realização acústica e articulatória que fornecem a ligação crucial entre a representação cognitiva da fala e a sua manifestação física. (Clements & Hume, 1995: 245. A tradução é minha.)

A afirmação de Clements & Hume, acima, transcrita do primeiro parágrafo de seu *The internal organization of speech sounds*, traduz uma preocupação - talvez não exatamente unânime, mas ao menos corrente - dos modelos de análise fonológica, que é a de se considerar a relação entre o fonético e o fonológico, tentando estabelecer uma conexão - ou “interface” - entre os dois aspectos.

Obviamente, tal preocupação é tão antiga quanto - ousaria mesmo dizer concomitante - a clássica dissociação entre Fonética e Fonologia, inaugurada pelo Estruturalismo.

Apesar de relegarem à Fonética o estudo do som em seu estado “bruto”, cabendo à Fonologia lapidá-lo e apontar sua função, dentro de um sistema lingüístico, Jakobson & Halle (1956) reconhecem a importância de considerar os aspectos fonéticos, na construção de seu modelo fonológico.

A maneira que encontraram de trazer a informação Fonética para dentro da Fonologia foi definir os traços distintivos - essencialmente os inerentes - pelas suas características acústicas e “genéticas”, isto é, articulatórias. No mais, como já mencionado no primeiro capítulo, a Fonética não tem participação qualquer no modelo já que, para Jakobson & Halle, os “feixes” de traços distintivos mantinham uma relação biunívoca com

os sons da fala, isto é, cada feixe de traços deveria corresponder a um e somente um fonema.

Com Chomsky & Halle (1995[1968]) e a Fonologia Gerativa, o modelamento fonológico ganha um caráter essencialmente abstrato. A preocupação com a representação mental não só dos segmentos, como também das operações que promovem a derivação de uma forma superficial, a partir de uma forma subjacente, abstrata, torna-se o cerne do modelo. E para tentar uma aproximação com a representação mental dos sons, Chomsky & Halle levam às últimas conseqüências o grau de abstração que subjaz ao modelo. Isto faz que a Fonética seja praticamente excluída da análise lingüística. Só há lugar para ela num dos últimos estágios da derivação da cadeia sonora, denominada “representação Fonética”. Essa representação, baseada numa transcrição Fonética da cadeia da fala, dá à transcrição um estatuto simbólico, cognitivo. Os autores deixam claro, aliás, que a transcrição deve ser entendida como uma representação dos sons da fala, e não como os sons da fala em si. Os sons da fala, *strictu sensu*, decorrem da implementação dessa representação por um sistema motor, portanto não lingüístico.

Tamanha abstração confere ao modelo um poder gerador incrível, ao ponto de que ele pode originar não só todas as cadeias sonoras universalmente possíveis e existentes, mas também cadeias sonoras inexistentes, em qualquer língua do mundo. Esse “excesso formal”, aliás, é um problema que os próprios autores acabam reconhecendo e que, no entanto, parece ter sido ignorado pelos modelos de análise fonológica subseqüentes, mesmo os chamados “não-lineares”⁵⁹. A preocupação principal destes últimos passa a ser a estruturação interna dos segmentos – não mais a formulação e o ordenamento das regras derivacionais que geram uma forma Y, a partir de uma forma X – buscando com isso dar conta de fenômenos não explicados pela Fonologia Gerativa Padrão, como o caso dos processos assimilatórios.

As observações acima são apenas um mote, para dizer que o fato de os modelos fonológicos em geral desconhecerem uma informação Fonética mais refinada – além

daquela essencialmente taxonômica, para usar um termo empregado por Ohala (1990) – traz como consequência a limitação em representar segmentos articulatoriamente mais complexos. Refiro-me particularmente, neste capítulo, às líquidas.

Maddieson (1980) relata que, numa amostra de 321 línguas constitutivas do UPSID (*UCLA Phonetic Inventory Database*), quase todas (95%) têm, pelo menos, uma líquida. Do total de línguas que exibem ao menos uma líquida, a maioria, isto é, 72% delas, têm mais de uma líquida. Maddieson observa também que, das 321 línguas do UPSID, 133 delas (41%) têm duas líquidas, sendo esse padrão distribucional o mais freqüente do inventário, pois línguas com mais que duas ou menos que duas líquidas não atingem sequer 25% do total.

Some-se a esse levantamento estatístico toda a descrição dos sistemas fonológicos de várias línguas do mundo, no *Handbook of the International Phonetics Association*, onde fica clara a distinção fonológica entre laterais e róticos e, em alguns casos, inclusive entre segmentos róticos.

Dadas todas essas evidências, era de se esperar que os modelos de análise fonológica fornecessem uma representação adequada para as oposições /l/ ~ /r/ ou /r/ ~ /l/, por exemplo, ao mesmo tempo captando a identidade de todos esses segmentos como uma classe. Não é, porém, o que ocorre.

Neste capítulo, busco evidenciar as falhas existentes na representação das líquidas, em alguns modelos fonológicos, em especial a Geometria de Traços, na configuração sugerida por Clements & Hume (1995).

A escolha por tal modelo se deve a duas razões principais: 1) parece haver uma certa hegemonia dele, na literatura fonológica – particularmente no Brasil –, o que faz crer que os fonólogos acreditam numa certa supremacia desta configuração de uma geometria de traços comparativamente às demais; 2) além de ter vantagens sobre as demais representações fonológicas – como a de que a estrutura interna dos segmentos é configurada por nós hierarquicamente ordenados – a Geometria de Traços tem paralelos com a Fonologia

⁵⁹ É preciso lembrar que tanto a Fonologia Natural como a Fonologia Gerativa Natural tentaram atacar o problema causado pelo imenso poder formal da Fonologia Gerativa Padrão. Curiosamente, no entanto, esses

Articulatória, ao menos nas versões iniciais deste modelo (Browman & Goldstein, 1986), conforme já abordado no capítulo anterior.

Neste capítulo, argumento a favor de considerarmos as líquidas uma classe de segmentos, a exemplo de fricativas ou oclusivas. Para isso, na seção 2.2 recorro a processos de alternância entre laterais e róticos, em línguas como o português, o espanhol, o catalão ou o basco. Tais alternâncias, que se verificam tanto na evolução histórica das línguas como em fatos sincrônicos – emergindo predominantemente em dialetos estigmatizados, mas também num dialeto padrão – são fortes pistas fonológicas de que as líquidas constituem uma classe natural.

Na seção 2.3, novamente trabalhando com dados, busco evidências da oposição fonológica possível entre laterais e róticos⁶⁰. São levantados pares mínimos especialmente do português e do espanhol, confirmando a ocorrência de duas “subclasses” – laterais e róticos – no interior da classe das líquidas. Ainda nesse item, e também com base no português e no espanhol, levanto dados mostrando que, na “subclasse” dos róticos podem opor-se entre si uma vibrante alveolar e um *tap*.

A partir das evidências empíricas trazidas em 2.2 e 2.3, se colocam as seguintes questões: 1) como representar fonologicamente a classe das líquidas, diferenciando-a das demais?; 2) como diferenciar laterais e róticos entre si e com relação a outros segmentos?

Os itens 2.4 e 2.5 se dedicam a discutir como a questão mencionada no parágrafo anterior é resolvida – ou melhor, não é resolvida! – dentro da Fonologia: se a representação da “classe líquidas” nunca foi um grande problema, o mesmo não acontece com a representação de laterais e róticos: há, desde propostas de traços *ad hoc* para representar a oposição entre laterais e róticos, passando por uma falta de consenso sobre qual traço captura as características fonológicas que unem esses segmentos em classes, até a solene negligência da existência dos róticos, em especial. Somem-se a todos esses problemas

modelos não alcançaram grande repercussão na teoria fonológica.

⁶⁰ A busca por essa oposição é fundamental para a discussão que instauo aqui, posto que há casos, como o dialeto paulista do kaingang (cf. D’Angelis, comunicação pessoal), em que não há oposição, mas alternância, do tipo “distribuição complementar” entre lateral alveolar e *tap*.

algumas “soluções” arbitrárias que tentam representar a oposição entre *taps* e vibrantes, através da estrutura silábica.

Depois da discussão desses problemas, sugiro, no item 2.6, uma representação das líquidas, dentro de uma geometria de traços. Para isso, tomo por base considerações, apresentadas nas duas seções anteriores, acerca de qual traço melhor captura as distinções que venho mencionando – proponho, inclusive, um traço [vibrante] para dar conta da oposição /r/ ~ /r/ – procurando uma representação segmental maximamente parcimoniosa, a partir da estrutura estabelecida por Clements & Hume (1995). Longe de querer ser a melhor, minha proposta pretende tão somente lançar alguma luz sobre esse tópico obscuro da teoria fonológica, que é a representação das líquidas.

Por fim, na seção 2.7, resumo as discussões prévias, apontando para a ineficiência da alternativa de representação, apresentada na seção anterior, argumentando que, mesmo no limite do modelo (a Geometria de Traços), a representação falha por limitações do próprio arcabouço teórico ao qual se filia. Esses problemas vão, desde uma certa arbitrariedade na escolha dos traços que pretendem dar conta das oposições mencionadas, até a impossibilidade de lidar com fatos como a gradiência existente na variabilidade de pronúncia dos róticos, o que se deve primordialmente à concepção, reinante nos modelos fonológicos, que é a de que Fonética e Fonologia constituem dois “módulos” distintos, sendo necessário um aparato – como os traços – que realize a tradução de um no outro.

2.1 - Questões inerentes à representação fonológica das líquidas

É provável que a falta de consenso no tratamento das líquidas e os problemas daí decorrentes, que aponto no parágrafo anterior, se devam às seguintes questões: a) é mesmo plausível propor a existência de uma classe de líquidas, que abarque laterais e róticos, ou é mais razoável propor que as laterais constituem uma classe fonológica distinta dos róticos?; b) qual traço distingue as líquidas das demais classes de sons?⁶¹; c) no interior da classe das

⁶¹ A princípio, esta questão poderia ser tomada como decorrência da primeira, se fosse respondida de maneira circular, ou seja, com uma afirmação do tipo: “sim, é possível propor uma classe líquidas, utilizando-se o

líquidas, qual traço distingue laterais de róticos?; d) no interior de cada classe – laterais e róticos – qual(is) traço(s) distingue(m) os segmentos entre si?

Tocar em tais questões exige, de início, a seguinte consideração, acerca da noção de “classe”: havendo boas razões para assumir que alguns segmentos constituem uma classe natural, é preciso que, na teoria, os mesmos compartilhem alguns traços comuns. Esse é o tratamento clássico de tal noção. Para caracterizar uma *classe natural de segmentos* diz-nos Hernandorena (1996: 29) que:

Hyman (1975, p. 139-140) estabelece critérios para a identificação de classes naturais. Diz que dois segmentos constituem uma classe natural quando um ou mais dos seguintes critérios são verificados em algumas línguas:

a – os dois segmentos sofrem regras fonológicas juntos;

b – os dois segmentos funcionam juntos nos ambientes de regras fonológicas;

c – um segmento é convertido no outro segmento através de uma regra fonológica;

d – um segmento é derivado no ambiente do outro segmento (como nos casos de assimilação).

2.2 - É sustentável a proposta de uma classe natural “líquidas”?

Existem dados, tanto de natureza diacrônica, como de natureza sincrônica, que sustentam ser mesmo possível propor a existência da classe das líquidas. Assim, nesta seção comento esses dados - principalmente relativos ao PB, o que não exclui dados de algumas outras línguas – com o intuito de mostrar que é plausível agrupar laterais e róticos numa mesma classe natural.

Num outro lugar (Silva, 1999), observo que, de acordo com Allen (1968), na Grécia antiga a denominação “líquidas” era usada para designar **rr** e **ll** como alusão ao caráter “fluido”, ou instável, que esses segmentos assumiam na métrica: quando a uma sílaba

traço ‘x’”. Se, em contrapartida, a questão for respondida *comme il faut*, i.e., apelando-se para fatos fonológicos que justifiquem uma classe líquidas, então teremos mesmo duas questões distintas.

contendo vogal breve se seguia um encontro consonantal do tipo plosiva + líquida, a quantidade da vogal permanecia duvidosa, i.e., podia variar entre breve e longa, devido justamente ao efeito da líquida⁶².

Considerando-se os critérios de Hyman *supra*, para determinar se alguns segmentos constituem ou não classes naturais, poderíamos usar o comportamento de róticos e laterais, na métrica grega, como um primeiro exemplo, já que todos “funcionam juntos, nos ambientes de regras fonológicas”, isto é, têm o mesmo efeito sobre vogais adjacentes, numa situação de métrica.

Tomemos, agora, os seguintes fatos diacrônicos presentes em línguas ibéricas, apresentados por Lloret (1997)⁶³:

Forma primitiva	Espanhol	Catalão	Italiano	Português	Basco
1. peregrinus	—	pelegri	pellegrino	—	—
2. armariu	—	—	—	almário	—
3. anima	alma	—	—	alma	*alima > arima ⁶⁴
4. *lusciniolu	—	rossinyol ⁶⁵	—	—	—
5. liliu	lirio	lliri	—	lírío	—

Quadro 2.1 - Dados diacrônicos evidenciando processos de lateralização de /r/ e de rotacismo da lateral, cf. Lloret (1997)

⁶² Devo esta referência ao prof. Jorge Piqué, do Departamento de Lingüística, Letras Clássicas e Vernáculas, da Universidade Federal do Paraná.

⁶³ Emprego a nomenclatura “línguas ibéricas” no mesmo sentido em que é empregada em Martínez-Gil & Morales-Front (1997), i.e., designando as diversas línguas faladas na Península Ibérica, o que inclui, num mesmo grupo, e.g., português, espanhol, catalão e basco. Não é meu intuito, aqui, tocar em questões concernentes aos grupos históricos de línguas, embora o tópico possa ser discutido.

⁶⁴ Lloret observa, na p. 144, que “*embora esteja claro o envolvimento da dissimilação nasal na evolução de ANIMA a arima ‘alma’, Michelena (1985) menciona um possível ‘estágio de lateral’ intermediário, i.e., *alima. Nesse caso, a mudança de *alima a arima deveria ser relacionada ao fato de que, em basco, l geralmente se torna r na posição intervocálica. (cf. *tselú > zeru ‘céu’, Alava > Arava ‘topônimo’).*”

⁶⁵ Traduz-se esse vocábulo como “rouxinol”. Logo, é possível pensar que, também na evolução do português, houve rotacismo da lateral, como é, de fato, observado por Huber (1986: 151), que afirma ocorrer aí um caso de “dissimilação a distância”, de modo que “rouxinol”, no português, resulta da evolução da forma “lusciniolu” a partir do provençal “rosinhol”, que passa a “roussinhol” no português antigo.

Nos exemplos acima ocorre a lateralização de /r/⁶⁶ - exemplos 1 e 2 – e o rotacismo da lateral – exemplos 3, 4 e 5, se aceitarmos que realmente tenha existido a forma intermediária *alima para o basco *arima*. Um outro exemplo de rotacismo da lateral pode ser encontrado na evolução de *animália*, do latim vulgar, a *alimária*, no português⁶⁷. Houve, portanto, a conversão de um segmento noutra através de uma regra fonológica, um dos “requisitos” necessários para considerarmos que /r/ e /l/ pertencem à mesma classe fonológica. Seguem, na mesma direção – rotacismo da lateral –, os dados do português, arrolados no quadro abaixo, a partir de Huber (1986):

Latim	Português atual
<i>regula</i> > <i>regla</i> (latim)	regra
<i>eclesia</i> > <i>egleia</i> (latim)	igreja
<i>nobile</i> (latim)	nobre
<i>cobla</i> (provençal)	cobra

Quadro 2.2 - Dados relativos à evolução histórica do português, a partir do latim e do provençal, ilustrando o rotacismo da lateral em grupos consonantais.

Os mesmos processos descritos, na evolução das línguas – Quadros 2.1 e 2.2 –, isto é, lateralização de /r/ e rotacismo da lateral, ocorrem sincronicamente, seja numa forma padrão das línguas, seja numa forma coloquial ou numa forma estigmatizada. Observe-se este novo quadro abaixo – adaptado de Lloret (1997) -, ilustrativo dos processos referidos em registros coloquiais do espanhol e do catalão:

⁶⁶ Esse processo também pode ser referido na literatura como *lambdacismo*, conforme, e.g., Obediente (1998).

⁶⁷ Este dado foi-me fornecido pelo prof. Décio Ranzani da Silva (comunicação pessoal).

Espanhol padrão	Espanhol coloquial	Catalão padrão	Catalão coloquial
1. glândula	grândula	glândula	grândula
2. julio	—	juliol	juriol
3. armari	almari	armari	almari ⁶⁸
4. cerebro	celebro ⁶⁹	—	—
5. fraile	—	frare	flare ⁷⁰

Quadro 2.3 - Dados sincrônicos evidenciando o rotacismo da lateral e a lateralização de /r/ em formas coloquiais do espanhol e do catalão (cf. dados de Lloret, 1997).

No caso do PB, abundam os exemplos de alternâncias entre lateral e rótico. Assim, em dialetos estigmatizados, tem-se, e.g., *craro*, *praca*, *broco*, *pranta*, *armoço*, *arface*, *arto*⁷¹. No entanto, também na forma padrão da língua, encontram-se dessas alternâncias, as quais estão inclusive dicionarizadas. Assim, o banco de dados de língua escrita do LAFAPE, constituído pelos itens lexicais do Minidicionário Aurélio, registra os seguintes exemplos:

1. aluguer *	aluguel
2. fruta	flauta
3. frecha	flecha
4. froco	floco
5. fror *	flor
6. grosa	glosa

Quadro 2.4 - Exemplos de alternância entre /r/ e /l/ registrados pelo Minidicionário Aurélio (as formas marcadas com * são registradas apenas na versão completa do Dicionário Aurélio).

É provável – digo isso intuitivamente, pois não disponho de dados estatísticos referentes à frequência de ocorrência desses vocábulos – que, na forma padrão da língua portuguesa, sejam mais frequentes as formas da coluna à direita, até mesmo porque o

⁶⁸ Tradução: “guarda-roupas”.

⁶⁹ Lloret observa (p.131) que, além de ocorrer num registro coloquial, a forma *celebro* era também a forma arcaica para “cerebro”. Ainda no espanhol arcaico, havia *miércoles*, forma que evoluiu para *miércoles* (quarta-feira).

⁷⁰ Tradução: “monge”.

rotacismo é um processo socialmente estigmatizado, o que, não raro, leva à hipercorreção. De qualquer maneira, continua se colocando o fato de que a alternância entre lateral e rótico não é exclusiva de dialetos estigmatizados. Fenômeno análogo é descrito por Obediente (1998) para o espanhol venezuelano:

Como em outras regiões de língua espanhola, ocorre na Venezuela (exceto nos Andes) o fenômeno da neutralização da oposição /-l/ ~ /-R/ em coda silábica. Esta neutralização obviamente elimina toda uma série de oposições semânticas, como “mal ~ mar”, “salta ~ sarta”, “alto ~ harto”, etc. A ausência de distinção entre l e r em posição final de sílaba (interna ou de palavra) se dá

** nas zonas rurais do país, exceto nos Andes*

** nos falantes da chamada “classe urbana marginal”*

** em setores da região dos Llanos e do Oriente, inclusive entre pessoas de um certo nível sócio-cultural.*

De maneira geral, é uma pronúncia estigmatizada, exceto no Oriente.

Deste modo, /l/ e /R/, em posição final de sílaba, perdem seu valor distintivo e se neutralizam em uma espécie de “superarquifonema líquido”, que pode ter como realizações fonéticas os seguintes fones: [l], [r] ou [ɹ]. (Obediente, 1998: 336, 337. A tradução e os grifos são meus.)

Também no tailandês, pode ocorrer alternância entre uma lateral e um rótico: de acordo com a descrição de Tingsabadh & Abramson (1999:147), a variação alofônica entre os dois sons ocorre principalmente em posição inicial de palavra, apesar de a distinção ser mantida por *cultivated speakers*, especialmente num discurso formal.

Além da alternância entre laterais e róticos, supra mencionada, há que se considerar “empréstimos” de laterais, no caso de línguas que possuem apenas róticos: ocorre aí uma espécie de “neutralização”, de onde resulta que a lateral da língua de origem resulta num rótico, na língua de chegada. Assim:

⁷¹ Formas correspondentes respectivamente a: claro, placa, bloco, planta, almoço, alface, alto.

1. bola	[ˈmbɔra] (kaingang)
2. galinha	[ŋgaˈrĩɲ] (kaingang)
3. Léry (nome próprio)	[reˈri] (tupinambá)
4. cavalo	[kɔwɔru] (kaingang)
5. mula	[muˈra]

Quadro 2.5 - Exemplos de neutralização da lateral em línguas indígenas brasileiras, a partir de "empréstimos" do português⁷²

Todos os dados apresentados nesta seção convergem para um mesmo ponto: a similaridade entre os dois sons – lateral e rótico – que faz haver freqüentemente a “troca” de um pelo outro, não só diacrônica mas também sincronicamente e, neste caso, em dialetos estigmatizados e, inclusive, no dialeto padrão.

A semelhança entre róticos e laterais, que promove a substituição de um segmento pelo outro, num mesmo ambiente fonológico, constitui-se num forte argumento para se sustentar a existência da classe “líquidas”. Resta a questão sobre qual traço distintivo diferencia as líquidas dos demais segmentos, nos sistemas fonológicos. Esta questão – a segunda que eu levantei, na primeira seção deste estudo - será abordada mais adiante.

2.3 - Como distinguir laterais de róticos numa mesma representação?

Obviamente, posta dessa maneira, tal questão pressupõe que laterais e róticos são funcionalmente distintos entre si, em alguns sistemas fonológicos, i.e., são fonemas, ou segmentos, fonologicamente distintos. O objetivo desta seção é, antes de mais nada, fornecer argumentos para esta afirmação. Então, num percurso semelhante ao já conhecido, vamos partir de dados do PB.

⁷² Dados fornecidos pelo prof. dr. Wilmar da Rocha D’Angelis (comunicação pessoal).

1. lata	rata
2. louco	rouco
3. fluir	fruir
4. planto (verbo)	pranto
5. falta	farta
6. alma	arma
7. calo	carro
8. calo	caro

Quadro 2.6 – Pares de vocábulos ilustrando distinções entre laterais e róticos no PB.

Observando-se os vocábulos nas duas colunas, é possível perceber que ao vocábulo da esquerda opõe-se o da direita, pelo seu valor semântico. Ora, esta é a prova clássica e cabal de que dois sons têm valor distintivo, num determinado sistema lingüístico, i.e., são fonemas. Ressalte-se que, no quadro acima, procurei captar diferenças que ocorrem em todas as posições silábicas possíveis, sem considerar, de início, a oposição /r/ ~ /r/ e a vocalização da lateral, em final de sílaba.

Também no espanhol é possível encontrar pares mínimos que evidenciam a oposição entre lateral e rótico, conforme ilustra o quadro abaixo:

1. mar	mal
2. salta	sarta
3. alto	harto

Quadro 2.7 - Pares mínimos evidenciando a oposição /r/ ~ /l/ em espanhol (dados de Obediente, 1998).

Pode-se ainda encontrar a oposição /r/ ~ /l/ nas seguintes línguas, segundo o *Handbook of the International Phonetic Association* (1999): amárico⁷³, árabe, búlgaro, catalão, croata, galego, hausa, hebraico, húngaro, persa, sindhi⁷⁴, taba⁷⁵, tcheco, tukang

⁷³ Língua semítica falada na Etiópia.

⁷⁴ Língua paquistanesa.

besi⁷⁶. Em esloveno, irlandês (ou gaélico), hindu e holandês a oposição existente dá-se entre /l/ e /r/.

É preciso notar que, dentro da classe dos róticos, se distinguem, no português, dois fonemas, /r/ e /ʀ/.⁷⁷ Prova disso são os dados 7 e 8 do Quadro 2.6, *supra*: neles, a *calo* se opõe, por um lado, *carro* e, por outro, *caro*, de onde se infere que *carro* e *caro* são, por sua vez, opostos entre si. Também no espanhol, a vibrante /r/ se opõe ao *tap* /r/:

1. cerro	cero
2. corral	coral
3. carro	caro
4. corro	coro
5. forro	foro
6. mirra	mira
7. morro	moro
8. parra	para

Quadro 2.8 - /r/ ~ /ʀ/ em espanhol, cf. dados de Obediente (1998).

De acordo com o *Handbook of the International Phonetic Association* (1999), no catalão, a vibrante se opõe ao *tap* em pares como *cera* x *serra*. Outra língua para a qual essa oposição é registrada é o galego. Para esta, entretanto, não são fornecidos exemplos de pares mínimos.

Se os dados do português e do Quadro 2.8 evidenciam a oposição entre a vibrante /r/ e o *tap*, é necessário, então, prever duas representações fonológicas distintas, cada uma correspondendo a um dos fonemas em questão. Contudo, este ponto é igualmente

⁷⁵ Língua austronésia, falada ao norte da província de Maluku, Indonésia.

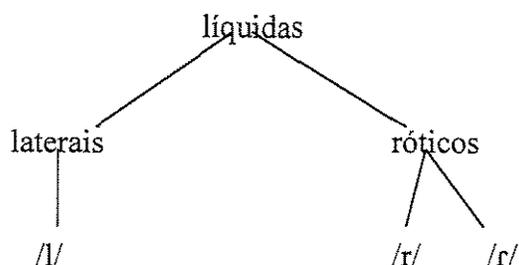
⁷⁶ Nome da língua falada no arquipélago homônimo, localizado ao sudeste da Indonésia.

⁷⁷ Estou tacitamente assumindo a existência de dois fonemas róticos, não apenas de um, que se realizaria diferentemente de acordo com a posição silábica, como assumem alguns autores. Essa discussão será posta adiante. Por enquanto, basta a observação que aí está.

problemático, na teoria fonológica, e, nas próximas seções, estarei remetendo às controvérsias existentes a esse respeito.

2.4 – Como representar a classe “líquidas”?

Os dados da seção anterior apontam para as oposições entre lateral e róticos e entre /ɾ/ e /r/, para o português e o espanhol, principalmente. Em resumo, podemos desenhar um esquema similar a este:



Ou seja, há uma classe de segmentos “líquidas”, diferente de outras classes de segmentos e no interior da qual se distinguem laterais de róticos. No interior da classe dos róticos, por sua vez, uma vibrante alveolar /ɾ/ distingue-se de um *tap* /r/. É preciso mencionar que não estou tocando na distinção entre a lateral alveolar /l/ e a lateral palatal /ʎ/ (evidenciada, e.g., em pares como *cala* x *calha*; *pala* x *palha*) porque essa é uma questão que parece mais tranqüila, uma vez que a distinção é de ponto de articulação.

Diante desse esquema, a primeira questão que se coloca é: havendo uma classe “líquidas”, distinta de outras classes de segmentos, na superfície, como representá-la adequadamente no nível fonológico?

Essa não parece uma questão crucial para a teoria fonológica, uma vez que os modelos sempre deram conta de opor – via combinações de traços distintivos – as chamadas “classes principais” de sons. Assim, Jakobson & Halle (1956) já caracterizavam a classe das líquidas pelos traços [+ consonantal, + vocálico], opondo-a às demais, caracterizadas como [+ consonantal, - vocálico], i.e., obstruintes e nasais; [- consonantal,

+ vocálico], vogais; [- consonantal, - vocálico], aproximantes. O argumento dos autores, para sustentarem tal caracterização, era de que a mesma capta a identidade vocálica e consonantal desses segmentos.

Chomsky & Halle (1995[1968]) fazem um percurso semelhante – considerando-se inclusive que o seu sistema de traços distintivos baseia-se, em grande parte, no sistema de Jakobson & Halle (*op. cit.*) – e caracterizam as líquidas como [+ silábico, + consonantal], também sob o argumento de que a especificação positiva para ambos os traços capta propriedades como o fato de poderem ocupar núcleo silábico, a exemplo das vogais, ao mesmo tempo em que travam sílaba ou oferecem obstáculo à passagem do ar no trato, características fonológica e fonética de consoantes. Analogamente ao sistema estruturalista, a Fonologia Gerativa Padrão, ao caracterizar as líquidas dessa maneira, distingue-as das demais classes, caracterizadas como [+ vocálico, - consonantal] (vogais); [- vocálico, + consonantal] (obstruintes e nasais); [- vocálico, - consonantal] (aproximantes). Com a substituição do traço [vocálico] por [silábico], no próprio SPE, as vogais passam a ser caracterizadas como [+ silábico, - consonantal]; nasais e obstruintes são [- silábico, + consonantal]; aproximantes são [-silábico, -consonantal]. As líquidas, por sua vez, são sempre [-consonantal], mas podem ter a especificação positiva ou negativa para [silábico], especialmente no caso de línguas, como o inglês, nas quais esses segmentos ocupam tanto margem quanto pico silábico. No primeiro caso, teriam especificação negativa para [silábico] e, conseqüentemente, integrariam uma classe natural maior com nasais e obstruintes, sendo necessários mais traços para distingui-las; no outro, a especificação seria positiva.

Nas chamadas fonologias não-lineares, em particular na Geometria de Traços, proposta por Clements & Hume (1995), as principais classes naturais são definidas pelos traços [soante], [aproximante], [vocóide], que constituem o nó Raiz. Esse nó, que domina todos os outros nós e traços que constituem internamente os segmentos, representa – de

acordo com os autores – o segmento como uma unidade fonológica⁷⁸. Além disso, tais traços identificam o grau de sonoridade das classes, estabelecendo uma escala de sonoridade⁷⁹.

Logo, a classe das obstruintes é definida como [- soante, - aproximante, - vocóide], sendo a menos sonora de todas as classes; a classe das nasais é definida como [+ soante, - aproximante, - vocóide], apresentando grau 1 na escala de sonoridade; as líquidas definem-se como [+ soante, + aproximante, - vocóide], apresentando grau 2, na escala; as vogais são [+ soante, + aproximante, + vocóide], apresentando grau 3, na escala de sonoridade. Esse tratamento – bem como os anteriores, de Chomsky & Halle e de Jakobson & Halle – traduz a possibilidade de as líquidas ocorrerem tanto em margem como em núcleo silábico⁸⁰. É preciso observar que, ao menos neste ponto, o uso classificatório dos traços distintivos dá conta de opor líquidas aos demais segmentos.

O esquema traçado, no início desta seção, porém, coloca duas outras questões, muito mais complexas: 1) como opor, numa representação formal, laterais e róticos?; 2) como opor, nessa mesma representação, os segmentos da classe dos róticos, i.e., vibrantes e *taps*?

Existindo as oposições mencionadas, uma gramática universal tem de dar conta delas – considerando-se inclusive que português e espanhol não são as únicas línguas em que se verificam e que essas oposições tampouco são restritas a uma família de línguas, pois o basco, por exemplo, também as exhibe. Contudo, não é isso que ocorre: os modelos correntes de análise fonológica – desde a Fonologia Gerativa Padrão até as geometrias de traços – assumem uma perspectiva classificatória, o que causa conflito com o uso fonético dos traços.

⁷⁸ Julgo desnecessário resenhar esse modelo, explicando a proposta de cada nó de classe e de cada traço, porque já há resenhas muito bem feitas a respeito. Remeto, em particular, a Hernandorena (1996) e D'Angelis (1998).

⁷⁹ Segundo essa escala, as classes de sons apresentariam uma “sonoridade inerente” – que regula sua distribuição na sílaba –, de modo que, quanto mais próximas de vogais, mais sonoras. Inversamente, quanto mais próximas de “consoantes típicas”, i.e., oclusivas, menos sonoras. Então, numa ordem crescente – dos menos para os mais sonoros –, os sons da fala se organizariam da seguinte maneira: oclusivas < fricativas < nasais < líquidas < aproximantes < vogais.

⁸⁰ Como é o caso do inglês, que apresenta, e.g., *light*, *novel*, *table*. Nos dois primeiros exemplos, /l/ ocorre, respectivamente em *onset* e rima; no terceiro, ocorre o chamado “l silábico”.

Tais modelos parecem ter atingido um grau “extremamente formal”, desconhecendo a ressalva que Chomsky & Halle fazem, no final do *The sound pattern of English*:

Toda a discussão de fonologia neste livro sofre de uma inadequação teórica fundamental. (...) O problema é que nosso tratamento dos traços, das regras e da avaliação foi extremamente formal. (...) Não há nada em nosso tratamento da teoria lingüística para indicar que o resultado seria a descrição de um sistema que viola certos princípios que governam as línguas humanas. Até onde isso é verdadeiro, nós falhamos em formular os princípios da teoria lingüística, da gramática universal, de um modo satisfatório. (Chomsky & Halle, 1995[1968]: 400. A tradução e os grifos são meus.)

Como decorrência, por mais que reivindicuem para si o poder de *fazer uma ligação plausível entre a estrutura fonológica abstrata e a interpretação fonética* (Clements & Hume, 1995:298), nenhum modelo consegue uma representação satisfatória das líquidas, sequer num nível fonológico. Exigir, portanto, que eles reflitam a realidade fonética dos segmentos em questão é impossível.

Assim, os itens seguintes serão dedicados a apontar os problemas existentes na representação de laterais e róticos, explorando a Geometria de Traços de Clements & Hume (1995), nos seus limites, para propor uma representação fonológica e mostrar que, ainda no limite, devido à própria estrutura interna do modelo, qualquer representação, por mais adequada que se pretenda, será sempre arbitrária.

Para elaborar tal representação, recorrerei a modelos anteriores (Fonologia Gerativa Padrão, principalmente), na tentativa de encontrar suporte para o inventário de traços distintivos que constituem a estrutura interna dos segmentos laterais e róticos.

2.5 – Como distinguir laterais de róticos numa representação fonológica?

Neste item, aponto as controvérsias inerentes a essa tarefa, desde o modelo estruturalista de Jakobson & Halle (1956), até a Geometria de Traços de Clements & Hume (1995), propondo que laterais e róticos se diferenciem pelo traço [contínuo], porém não exatamente da maneira sugerida em SPE.

2.5.1 – A representação das laterais em *Fundamentals of Language* e SPE

Uma das primeiras tentativas de diferenciar laterais de róticos via traços distintivos remonta a Jakobson & Halle (1956) que propõem, em *Fundamentals of Language*, uma saída *ad hoc* para isso, empregando a oposição [contínuo/descontínuo]. O modelo prevê que laterais sejam [contínuo] e, róticos, [descontínuo] – como, aliás, farão, mais tarde, outros fonólogos, como os próprios Chomsky & Halle (abordo este ponto logo adiante). Para Jakobson & Halle, a oposição descontínuo/contínuo se define da seguinte maneira:

Acusticamente – silêncio (pelo menos na faixa de frequência acima daquela da vibração das pregas vocais) seguido e/ou precedido pelo espalhamento da energia sobre uma larga região de frequência (como um estouro ou como uma rápida transição aos formantes das vogais) versus a falta de transição abrupta entre o som e tal silêncio;

Geneticamente – rápida ativação ou desativação da fonte, tanto por meio de uma rápida oclusão e/ou abertura do trato vocal que distingue plosivas de constrictivas, como por meio de uma ou mais batidas que diferenciam as líquidas descontínuas, como um flap ou uma vibrante, de líquidas contínuas, como a lateral /l/. (Jakobson & Halle, 1956: 30. A tradução é minha.)

Na Fonologia Estruturalista, o emprego dessa distinção é a única maneira de diferenciar laterais de róticos.

Na Fonologia Gerativa Padrão, por sua vez, embora os autores abandonem essa saída *ad hoc*, há uma controvérsia explícita à diferenciação entre a vibrante alveolar /r/ e a

lateral alveolar /l/. Em princípio, Chomsky & Halle propõem, no SPE, que /r/ se diferencie de /l/ pelo traço [contínuo], sendo esta [- contínuo] e, aquela, [+ contínuo].

O [r] vibrante é mais difícil [relativamente à representação da variante fricativa, antes mencionada] , porque há interrupção da corrente de ar, pelo menos durante parte da duração do som. As vibrações da ponta da língua, contudo, são produzidas pela diminuição da pressão que ocorre dentro da passagem entre a ponta da língua e o palato, quando o ar se propaga rapidamente através dela (efeito de Bernoulli). A vibração é, pois, um efeito secundário do estreitamento da cavidade sem o efetivo bloqueio da passagem do ar. Conseqüentemente, há boas razões para considerar o [r] vibrante contínuo e, não, oclusivo. (...) A caracterização da líquida [l], em termos da escala contínuo/não-contínuo, é ainda mais complicada. Se a característica que define a oclusiva é tomada (como acima) como o bloqueio total da passagem do ar, então [l] deve ser visto como contínuo e deve ser distinguido de [r] pelo traço de “lateralidade”. Se, por outro lado, a característica que define oclusivas é considerada como sendo o bloqueio da corrente de ar, após a constrição primária, então [l] deve ser incluído entre as oclusivas. (SPE: 317, 318. A tradução é minha.)

Se o objetivo de qualquer modelo fonológico é propor para os sons da fala uma representação que tenha uma motivação fonética⁸¹, se não por outra via, pelo menos através da definição dos traços distintivos, então é no mínimo contra-intuitivo diferenciar /l/ de /r/ pelo traço [contínuo], de modo que a lateral seja [- contínuo] e a vibrante, [+ contínuo]. Afinal, para a produção da lateral, a língua se retém junto à região alveolar, resultando dessa configuração articulatória um espectro contínuo⁸², tal

⁸¹ O que varia entre os modelos é o grau de abstração com que tentam cumprir essa tarefa, uns levando fatos fonéticos em maior consideração que outros.

⁸² É preciso observar que, ao recorrer ao espectro, para apoiar minha hipótese, não pretendo privilegiar o acústico em detrimento do articulatório: na impossibilidade de realizar experimentos articulatórios, valho-me da informação acústica e de sua relação com a articulação dos sons.

qual o de vogais. Já a vibrante é que apresenta características temporais com “oclusões” (ou “fechamentos”, como chamo em outro lugar⁸³), entrecortando os períodos contínuos.

2.5.1.1 – Proposta para a representação da oposição /r/ ~ /l/ via [contínuo]

Pelo que exponho acima, pareceria mais razoável tratar a lateral como [+contínuo] e a vibrante como [-contínuo], como se faz em *Fundamentals*. Ou, então, seguir a outra proposta de SPE e especificar o traço [lateral] positivamente para /l/ e negativamente para /r/. Esta alternativa será refutada adiante, em favor do abandono do traço [lateral], pois procurarei mostrar nesta seção que prescindir desse traço, operando a distinção /r/ ~ /l/ pelo traço [contínuo] pode resultar numa representação mais parcimoniosa, principalmente devido à pouca operacionalidade de [lateral].

A princípio, porém, considerando-se as duas alternativas postas, teríamos duas matrizes de traços possíveis para /r/ e /l/⁸⁴:

	/r/	/l/
Silábico	+	+
Consonantal	+	+
Contínuo	-	+

Quadro 2.9 - Proposta de matriz de traços para /r/ e /l/, à luz de SPE, tratando a oposição pelo traço [contínuo].

	/r/	/l/
Silábico	+	+
Consonantal	+	+
Lateral	-	+

Quadro 2.10 - Proposta de matriz de traços para /r/ e /l/, à luz de SPE, tratando a oposição pelo traço [lateral].

⁸³ Silva (1996).

⁸⁴ Embora não esteja mencionada aí – por não envolver nenhum dos traços em questão – a diferença entre a lateral palatal /ʎ/ e outros segmentos líquidos do PB dá-se pelo traço [anterior], especificado positivamente para /r/, /r/ e /l/ e negativamente para /ʎ/.

Sobre os Quadros 2.9 e 2.10, acima, é necessário frisar que [lateral] não decorre de [contínuo], embora não haja segmentos [+lateral, - contínuo].⁸⁵

Além disso, faz-se pertinente uma questão colocada por D'Angelis (comunicação pessoal) sobre a pouca viabilidade de propormos um traço [lateral], porque ele deve ser menos operacional nas línguas do que o traço [contínuo]. Este, inclusive, é o ponto de partida da proposta de Spencer (1984), para a eliminação do traço [lateral]: seguindo ainda uma proposta linear, baseada na Fonologia Gerativa Padrão e no modelo de Halle & Stevens (1979)⁸⁶, o autor defende que o traço [lateral] seja eliminado das representações fonológicas e que a distinção da lateral, com relação aos demais segmentos, se dê através do traço [distribuído]⁸⁷.

Seu argumento é de que, embora haja uma definição articulatória clara para segmentos laterais, não há uma justificativa “puramente fonológica” para o traço [lateral], i.e.,

(...) ninguém mostrou que uma generalização lingüisticamente significativa sobre o padrão sonoro de uma língua ou sobre a estrutura sonora do seu léxico será perdida se esse traço ([lateral]) não estiver disponível. (Spencer, 1984:23. A tradução é minha.)

O autor ainda nota que, mesmo havendo evidências para a existência das laterais, nunca foi mostrado que essa classe não poderia ser caracterizada por outros traços, melhor motivados, o que poderia levar a uma análise mais parcimoniosa, no sentido de eliminar certas redundâncias na representação.

⁸⁵ O IPA não registra ocorrências desse tipo de segmento.

⁸⁶ Segundo Spencer (*op. cit.*), os autores propõem, nessa obra, a caracterização da lateral como [+ contínuo]. Não tive, entretanto, acesso a essa referência, para que possa fazer um comentário mais detalhado.

⁸⁷ Em comunicação pessoal, D'Angelis sugere que o traço [distribuído] seja utilizado para caracterizar a classe das líquidas como um todo. Nesse sentido, *tap* e lateral alveolar, além da aproximante alveolar, seriam [- distribuído], ao passo que a lateral palatal e a vibrante alveolar seriam [+ distribuído]. Essa poderia ser uma saída, mas, ainda assim, permanece a questão inerente ao tratamento das líquidas, via traços distintivos, que é a de como representar todos os alofones de /r/, por exemplo, já que, na Fonologia Gerativa, as alofonias são previstas na representação. Mais uma vez, dado o caráter “taxonômico” dos traços, também esta acaba sendo uma proposta arbitrária, tanto quanto as que proponho neste estudo.

Então, a proposta de Spencer consiste em caracterizar todas as laterais como [+ coronal, + contínuo]. As laterais palatais e as retroflexas são consideradas [- anterior] e, as demais, [+ anterior]. Além disso, as laterais seriam [- distribuído], pois sua produção envolve

(...) a propagação do ar entre o lado da língua e o lado das bochechas, sendo essa passagem uma das maiores constrições no espaço fonético, ao longo do qual o ar se propaga, durante a fala. (Spencer, 1984: 29. A tradução é minha.)

Tal característica articulatória vai de encontro à definição que Chomsky & Halle (1995[1968]) dão ao traço [distribuído]. Assim, a lateral seria caracterizada como [+ vocálico, + consonantal, + contínuo, + distribuído] e distinta de todos os demais segmentos, inclusive da vibrante, por essa matriz. É pertinente observar que Spencer toma a vibrante /r/ como [- contínuo], também com base em Halle & Stevens (*op. cit.*). Esta observação será retomada em seguida.

O problema que surge para esta proposta – segundo o próprio autor – é a distinção entre laterais apicais e laminais, existente, por exemplo, em línguas australianas, e que a caracterização por [+ distribuído] não capta. A saída de Spencer é desdobrar o traço [anterior] em [apical] e [dental]; conseqüentemente, as laterais apicais são [+ distribuído, + apical]. Pode ser objetado o recurso do desdobramento de [anterior], assim como o recurso de caracterizar as laterais como [+ distribuído]. De fato, Chomsky & Halle propõem [distribuído] para dar conta de sistemas lingüísticos nos quais a classe das obstruintes distingue mais do que os três pontos de articulação – posterior, alto e baixo – que caracterizam segmentos [- anterior, - coronal]. Esse é o caso de línguas com o seguinte sistema de obstruintes: /p, t̚, t, t̚, k₁/ ⁸⁸.

Surge novamente a questão concernente à operacionalidade do traço nos sistemas, ao menos no caso de [distribuído], para caracterizar as laterais. Não tenho elementos para

⁸⁸ Utilizei a mesma notação empregada por Chomsky & Halle (1995[1968]) para, respectivamente, a oclusiva dental, a oclusiva alveolar, a retroflexa e a oclusiva palato-alveolar.

levar adiante esta discussão, mas a contribuição que Spencer (*op. cit.*) traz à nossa argumentação é de que é possível – julgaria até mais pertinente – caracterizar os segmentos laterais como [+ contínuo]. Ao menos no caso do PB e de outras línguas como o espanhol, bastam os traços [+ silábico, + consonantal, + contínuo], para diferenciar /l/ dos demais segmentos no sistema. Qualquer outro traço seria redundante. Os róticos, por sua vez, como menciono acima, seriam caracterizados, dentro da perspectiva de Spencer, como [+ vocálico, + consonantal, - contínuo]. O autor não faz referência à caracterização desses segmentos pelo traço [distribuído] ou pelo traço [anterior]. É possível prever que aqueles realizados num ponto dental-alveolar sejam [+ anterior], enquanto retroflexos e demais róticos, produzidos na região posterior do trato, sejam [- anterior]. De qualquer modo, para os róticos do PB são necessários apenas os traços [+ silábico, + consonantal, - contínuo] para diferenciá-los dos demais segmentos, a exemplo do que propomos no Quadro 2.9.

2.5.2 – A representação de segmentos laterais, nas geometrias de traços

Nas Geometrias de Traços, tais como a elaborada por Clements & Hume (1995), a tarefa de caracterizar segmentos laterais é ainda mais complicada, pois, apesar de darem um tratamento extenso e exaustivo a classes como vogais e obstruintes – tentando propor um modelo unificado para representar consoantes e vogais –, têm problemas ao tratar das soantes, como um todo, e das líquidas, em especial. O problema, no caso das geometrias de traços, parece relacionado à dificuldade de se propor um tratamento satisfatório para o modo de articulação, como aponta D’Angelis (1998)⁸⁹. Ou seja, o lugar na estrutura interna do segmento ao qual atrelar traços de modo permanece obscuro. Assim, no caso do traço [nasal], persiste a questão sobre a pertinência de atrelá-lo diretamente ao nó Raiz ou a um outro nó, mais baixo, na hierarquia da organização dos segmentos, como, e.g., um nó SP⁹⁰.

Quanto ao tratamento das líquidas, é extremamente superficial, pois o modelo prevê apenas a existência de um traço [lateral] que distinguiria /l/ dos demais segmentos.

⁸⁹ Particularmente com relação a [nasal] e [soante], a partir de dados do kaingang, uma língua pertencente ao tronco Macro-Jê e falada hoje nos estados da região sul do Brasil, predominantemente.

Similarmente ao que ocorre com o traço [nasal], os autores têm problemas em definir exatamente o lugar da hierarquia onde colocar [lateral]. Essa limitação do modelo é explicitamente reconhecida pelos autores, quando afirmam:

(...) a discussão até agora não tocou em dois traços, cuja afiliação ainda não está clara, [lateral] e [estridente]. No caso de [lateral], as duas hipóteses que competem entre si são atrelar (o traço [lateral]) sob o nó coronal ou sob o nó Raiz. O maior argumento para atrelar sob coronal vem do critério de implicação do nó, (...); se atrelamos [lateral] sob o nó coronal, nós damos conta diretamente do fato de que todos os segmentos que possuem esse traço são fonologicamente [coronal], sem a necessidade de qualquer outra estipulação. Entretanto, há, pelo menos, quatro problemas nesta visão: (a) quando uma nasal assimila em ponto a um som lateral, ela normalmente não se torna lateral (...) ⁹¹; (b) quando uma lateral assimila em ponto a uma não-lateral, ela normalmente retém sua lateralidade (por exemplo, no espanhol e tamil); (c) quando o nó de cavidade oral se espalha de [l] a [s], na formação de oclusivas intrusivas (por exemplo, false [...l's]), a oclusiva intrusiva resultante é central, não lateral (...); (d) obstruintes laterais podem ser completamente transparentes a regras de assimilação a longa distância, envolvendo obstruintes coronais(...). Estes fatos argumentam fortemente que [lateral] ocorre acima do ponto, na hierarquia de traços. Se é assim, sons [lateral] são universalmente coronais apenas por conta da maneira como este traço é definido. (Clements & Hume, 1995: 293. A tradução e os grifos são meus).

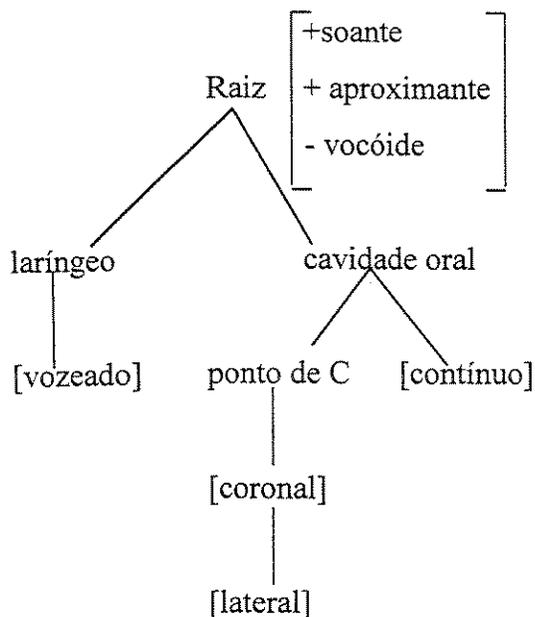
Clements & Hume, portanto, parecem preferir a alternativa que prevê o atrelamento de [lateral] diretamente sob o nó Raiz. Quanto aos róticos, o modelo desconhece

⁹⁰ Contribuições a essa discussão podem ser encontradas em D'Angelis (1998).

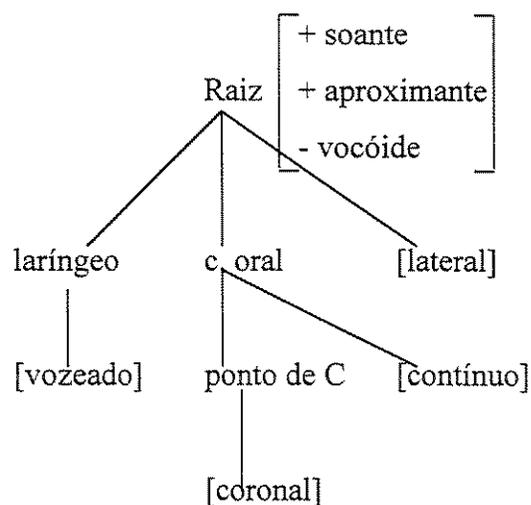
⁹¹ Este "problema" é discutível, já que tanto Lloret (1997) como Rice & Avery (1991) apresentam dados evidenciando o contrário (vide 2.6.1.2).

solenemente a sua existência, não propondo qualquer tratamento. Nesse modelo, as duas representações concorrentes para segmentos laterais seriam⁹²:

a) **[lateral] atrelado a [coronal]**



b) **[lateral] atrelado a Raiz**



Como menciono acima, a falta de consenso sobre o lugar ao qual atrelar o nó [lateral], na estrutura interna dos segmentos – sem falar na brevidade com que os autores propõem uma saída para [estridente], i.e., ligá-lo ao nó Raiz – relaciona-se a um problema inerente à Geometria de Traços, que é o de como dar conta dos traços de modo.

Um outro problema que se coloca para a teoria fonológica é a distinção dos róticos entre si. É a essa questão que me volto, a seguir.

⁹² Não julgo necessário comentar toda a estrutura proposta por Clements & Hume (1995), tocando na questão da motivação dos nós de classe – como denominam a cada um dos nós que dominam traços ou outros nós – nem tampouco a questão da mono ou bivalência dos traços. Boas resenhas desse modelo, que tocam inclusive nas questões que aponto, podem ser encontradas em Hernandorena (1996) e D'Angelis (1998).

2.5.3 – Como representar a oposição /r/ ~ /r̄/?

De acordo com o Quadro 2.8 da seção 2.3, o espanhol apresenta pares mínimos como “*cerro/cero*”; “*corral/coral*” ou “*mirra/mira*” que evidenciam a oposição fonológica entre a vibrante alveolar e o *tap*. Também no português, podem ser encontrados dados similares, como os pares “*forro/foro*”; “*erra/era*”; “*mirra/mira*”. Ainda nessa seção, observa-se – embora sem a disponibilidade de dados – que catalão e galego realizam a mesma oposição.

A oposição /r/ ~ /r̄/, em línguas como o português e o espanhol, sempre ofereceu problemas à teoria fonológica, especialmente porque ocorre apenas em posição intervocálica. Para tentar explicá-la, há duas alternativas na literatura: 1) afirmar que a oposição se estabelece entre um segmento simples e um segmento geminado; 2) propor um traço distintivo que dê conta de opor a vibrante alveolar ao *tap*.

A primeira alternativa é a mais recorrente, porém a menos sustentável. Vamos a ela.

2.5.3.1 – Oposição /r/ ~ /r̄/ estabelecida entre um segmento geminado e um não-geminado

Uma das primeiras propostas nesse sentido é elaborada por Câmara Jr., preocupado em precisar quantos são os fonemas róticos do PB. Na primeira edição de seu *Para o estudo da fonêmica portuguesa* (1953), o autor defende que o sistema consonântico do PB é composto de um único som rótico, o “r forte”⁹³. Para dar conta da oposição entre vibrante e *tap* em posição intervocálica, Câmara Jr. recorre a argumentos de natureza diacrônica. Assim, segundo ele, em latim havia um /r/, que podia ser geminado como qualquer outra consoante, e um outro /r/, simples, sendo a oposição simples/geminado distintiva.

Na evolução histórica do sistema consonantal, teriam ocorrido dois processos, a simplificação das geminadas e o enfraquecimento das consoantes simples. Nessa primeira versão de seu tratamento, Câmara Jr. defende que o “r brando” corresponderia a um enfraquecimento do /r/ simples latino, como consequência da posição intervocálica, sendo,

⁹³ Por “r forte” o autor denomina todos os róticos que se opõem àquele que ocorre exclusivamente em posição intervocálica ou grupos consonantais, chamado “r brando”, ou “r fraco”.

por isso, um alofone posicional do “r forte”. Este, por sua vez, decorreria de um processo de simplificação da geminada. Evidências favoráveis à tese de que esse segmento ainda preserva características de consoantes geminadas seriam dadas por processos de sândi externo, como em “ar roxo”.

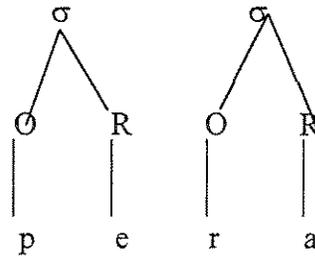
Na edição seguinte de *Para o estudo da fonêmica portuguesa* (1977), Câmara Jr. revê sua análise⁹⁴, argumentando a favor da existência de dois fonemas vibrantes, que se opõem unicamente em posição intervocálica, neutralizando-se nos demais ambientes, inclusive na posição inicial. A revisão de sua proposta deve-se, segundo ele, a que uma representação fonológica (fonêmica) não se pode afastar totalmente da “realidade fonética”. Além disso, tem de se calcar em fatos sincrônicos da língua e não diacrônicos, como fazia a proposta anterior.

Segundo o autor, a proposta de que o *tap* seria um enfraquecimento de outro fonema rótico se afasta de uma realidade fonética, porque /r/ não preserva uma identidade completa com a variante mais freqüente do /r/ forte, que é a posterior⁹⁵. Para Câmara Jr., isso seria um impedimento para se postular o /r/ forte como a contraparte geminada do /r/ brando. Daí o novo tratamento.

Uma outra proposta, que vai em sentido semelhante – pelo menos no que concerne ao argumento de que a vibrante é uma consoante geminada – é feita por Harris (1983) partindo de dados do espanhol. Para chegar a esse argumento, Harris se baseia nas seguintes observações e suposições: 1) o fonema rótico subjacente, em espanhol, é /r/; 2) em início de palavra, o único rótico possível é a vibrante; 3) apenas a vibrante ocorre, sucedendo /n, l, s/; 4) são proibidos encontros consonantais do tipo /rr/; 5) em encontros consonantais e final de palavra, somente /r/ é possível. As observações 2, 3 e 4 são formalizadas por regras, que transcrevo respectivamente como r2, r3 e r4, abaixo:

⁹⁴ De fato, este não é o primeiro lugar onde aparece a revisão de sua análise. Câmara Jr. já a havia revisto em *Problemas de lingüística descritiva* (1969) e *Estrutura da língua portuguesa* (1970).

Como, por outro lado, a regra r4 proíbe seqüências /r/ de ocorrer, o *tap* da primeira rima é apagado, resultando então:



De acordo com o autor, a derivação acima não demanda qualquer custo adicional para a gramática, mas traz reflexos na prosódia e na morfologia da língua.

Em termos prosódicos, seu tratamento explicaria a inexistência de proparoxítonas, cuja última sílaba se inicia por [r], e.g. * cáma[r]a; * víbo[r]a: a penúltima sílaba dessas palavras, de acordo com Harris, teria rima ramificada, que não acontece devido à regra de apagamento de /r/, em seqüências /r/. Morfologicamente, processos como amo[r]/* amo[r]es ou amo[r]/*amo[r]oso seriam agramaticais, porque pressuporiam uma forma amo/r/, a qual requereria um elemento terminal para ser uma palavra, visto que vibrantes são proibidas de ocorrerem em posição final.

Apesar de se assemelhar à de Câmara Jr., quanto à proposta de que a oposição /r/ ~ /r/ provém da oposição não-geminada/geminada, o tratamento acrescenta algo novo que é o apoio claro na estrutura silábica. Ora, temos aí um problema metodológico, que é o seguinte: embora, inegavelmente, os falantes reconheçam sílabas, até hoje a teoria fonológica não conseguiu formalizar a “unidade funcional sílaba”. Conseqüentemente, formalizar seus limites tampouco é possível. Logo, como dizer, com segurança, que nos casos da oposição intervocálica a vibrante decorre da geminção de *taps* heterossilábicos? Isto é tão arbitrário quanto admitir que os dois róticos de *cerro*, e.g., constituam o *onset* da segunda sílaba dessa palavra, como aliás é implicitamente admitido pela convenção ortográfica do espanhol.

Para o PB, há ainda um tratamento para os róticos, inspirado em Harris (1969, 1983), que é o de Monaretto (1997) e segundo o qual há, no PB falado no Rio Grande do

Sul, um único fonema subjacente, o *tap*, do qual se deriva a vibrante alveolar, a partir de um processo de gemação. É o mesmo argumento, portanto, de Harris, porém baseado – segundo a autora – na teoria da sílaba de Miliken (1988)⁹⁶.

De acordo com Monaretto (op.cit.) uma evidência favorável ao argumento de que existe uma única forma subjacente é fornecida por dados de aquisição da linguagem, já que a criança substitui tanto o r-forte como o fraco pela semivogal palatal [j]⁹⁷. O argumento de que essa forma é /r/ se basearia em processos fonológicos como o apagamento de /r/ pós-vocálico, em grupo consonantal, ou a sua substituição por lateral – e vice-versa. Tais processos, que segundo a autora são próprios de soantes, exigiriam que se postule o *tap* como forma subjacente. Tomar o r-forte como forma subjacente seria problemático, porque a sua variante mais freqüente é uma fricativa.

Com base nesses fatos, a autora postula que a vibrante seja uma consoante geminada e derivada de *taps* heterossilábicos⁹⁸. Assim, por exemplo, numa palavra como *carro*, [r] resulta da gemação do /r/ que trava a primeira sílaba e do /r/ que constitui o ataque da segunda.

A proposta, porém, fica apenas num nível descritivo: a tentativa de modelar os róticos, à luz do modelo mencionado (teoria da sílaba de Milliken, 1988), não vai além de uma representação estrutural de cada segmento, indicando que o *tap* se constitui de um único segmento, preso a um nó temporal, ao passo em que a vibrante se constitui de um segmento, preso a dois nós. [r], portanto, resulta do apagamento de um dos *taps* heterossilábicos, restando o nó temporal que se liga ao do *tap* remanescente. O suposto processo derivacional não é explicado ou sequer comentado.

Uma outra falha do tratamento de Monaretto é – também por conta do seu caráter exclusivamente descritivo – o fato de a autora não prever como derivar uma variante fricativa, a partir da vibrante: como já mencionado, a autora se restringe a uma representação da estrutura segmental de /r/ e /r/, atrelados ao esqueleto silábico. Outras

⁹⁶ Também a essa referência não tive acesso.

⁹⁷ Diga-se de passagem, aliás, que este não é um argumento muito convincente: a criança também pode substituir os róticos por uma lateral, como mostra Maia (1982), conforme menciono adiante, na nota 101.

etapas da derivação, como essa que aponto, são ignoradas. Ou simplesmente atribuídas a um nível pós-lexical, o que acaba sendo um subterfúgio para os problemas dos quais não se consegue dar conta, num nível lexical, – ao menos neste caso específico. Com isso, seu tratamento deixa de captar a identidade de classe, existente entre a vibrante e suas variantes fricativas.

Cabe também para o tratamento de Monaretto (*op.cit.*) a relativização que faço acima, a respeito da formalização de uma unidade fonológica “sílabas”. A mesma ressalva coloca-se, ainda, para a proposta de Mascaró & Bonet (1997).

Estes autores, contrariamente a Harris (*op. cit.*) ou Monaretto (*op.cit.*), não atribuem a oposição entre vibrante e *tap* a uma oposição geminada/não-geminada, atrelada à estrutura silábica. Antes, discordam de tratamentos desse tipo, observando numa nota que

Embora o tratamento de Harris funcione, alguém poderia se perguntar por que línguas que dificilmente têm geminadas na superfície deveriam ter flaps geminados subjacentes, os quais nunca se realizam como geminadas, enquanto outras línguas, como o italiano, que realmente têm geminadas na superfície, não exibem contraste entre róticos. O catalão tem apenas alguns casos (residuais) de geminadas (principalmente m, n, l e ʎ)⁹⁹. O espanhol não tem geminadas. No geral, nessas duas línguas as geminadas surgem apenas através de processos assimilatórios. (Mascaró & Bonet, 1997: 124. A tradução é minha.)

Assim, afastando a hipótese de que a oposição entre vibrantes e *taps* se dá pela oposição geminada/não-geminada, Mascaró & Bonet argumentam que há dois fonemas róticos distintos, /r/ e /r/, e que sua distribuição em línguas em que os mesmos se opõem é totalmente previsível e completamente determinada pela posição silábica, de modo que /r/

⁹⁸ Novamente, a mesma proposta de Harris (1983).

⁹⁹ Ressalte-se que Recasens (1991, 1999) argumenta fortemente contra os tratamentos fonológicos que tratam a vibrante como consoante geminada, fornecendo dados articulatórios do catalão que mostram a inexistência de róticos geminados, nessa língua.

ocorre em posição inicial de palavra ou sílaba e /r/, em segunda posição de *onset*, i.e., como segundo elemento de grupos consonantais.

A generalização da análise que liga fonema rótico a posição silábica, nos ambientes não-contrastantes, leva à previsão de que, no contexto de contraste V _ V deve ocorrer uma vibrante. A hipótese dos autores é que, em posição intervocálica, o segmento não-marcado é a vibrante, sendo /r/ o marcado. Em posição de *onset*, o rótico seria escolhido, a depender de sua posição interna nesse *onset*, uma vez que *taps* nunca ocorreriam em *onsets* simples, segundo os autores, e as vibrantes, por sua vez, nunca ocorreriam em *onsets* ramificados. De acordo com Mascaró & Bonet, a escolha do fonema se relaciona a diferenças entre vibrante e *tap*, na escala de sonoridade e no Princípio do Ciclo de Sonoridade, propostos por Clements (1990, apud Mascaró & Bonet, 1997). Esse Princípio orienta de tal modo a escolha do rótico, que estabelece que a especificação positiva para um traço [f]¹⁰⁰ – alusão a *flap* – tem como resultado, na superfície, um *flap* (ou *tap*). Por outro lado, se o traço é [-f], então o segmento resultante será uma vibrante.

Se o tap, como os glides, é mais sonoro que a vibrante, que pode ser considerada próxima às obstruintes em sonoridade, é natural que a vibrante ocorra como único elemento de um onset, enquanto o flap aparece, como os glides (ou as laterais), como segundo elemento de um onset. (Mascaró & Bonet, 1997: 108. A tradução é minha.)

A proposta da existência de um traço [f], especificado positivamente para [r], daria conta, como querem os autores, da ocorrência desse segmento em *onset* intervocálico.

Ao fim e ao cabo, esta proposta é tão arbitrária como as outras, mencionadas nesta seção. Além de não deixar clara a relação entre vibrantes e obstruintes, numa escala de sonoridade, o apelo a um traço [f] é claramente um artifício que tenta dar conta – de forma

¹⁰⁰ De acordo com Mascaró & Bonet, esse traço não é especificado subjacentemente, mas apenas durante o processo de silabificação.

totalmente desmotivada – de uma deficiência do modelo, que é a coocorrência de [r] e [r], em posição intervocálica. Além disso, poderíamos ainda argumentar que foneticamente [r] não está impedido de ocorrer, em início de palavras – isto é fato para alguns dialetos da região Sul do Brasil e dezenas de línguas indígenas¹⁰¹.

Na impossibilidade de explicar a oposição /r/ ~ /r/ pela oposição não-geminada/geminada (simplesmente porque as línguas em que a oposição acontece não têm quaisquer outras consoantes geminadas, o que faz de tal aproximação muito pouco natural) ou pela posição silábica, a alternativa é tentar dar conta dela, através de traços distintivos. É essa alternativa que persigo a seguir.

2.5.3.2 - Oposição /r/ ~ /r/ dada por traços distintivos

A tentativa de opor vibrantes e *taps*, por traços distintivos, existia já dentro de um arcabouço estruturalista. Tomando por base *Preliminaries to speech analysis*, autores como Llorach (1976) caracterizam a oposição pelo par [tenso]/[frouxo], de modo que a vibrante alveolar é [tenso] e o *tap*, [frouxo].

Seguindo o arcabouço da Fonologia Gerativa Padrão, autores como Harris (1983) ou Wheeler (1979, apud Mascaró & Bonet, 1997) também propõem que a oposição /r/ ~ /r/ do espanhol seja representada pela especificação do traço [tenso], tal que a vibrante alveolar é [+ tenso] e o *tap*, [- tenso].

Opor a vibrante alveolar e o *tap* pelo traço [tenso] não capta exatamente a diferença entre esses segmentos, pois, enquanto [tenso] pretende representar sons que requerem grande esforço articulatorio, [frouxo] representaria sons que requerem pouco esforço articulatorio, o que não parece ser o caso, já que o *tap* demanda um esforço articulatorio considerável. Tanto assim que é adquirido num estágio tardio do desenvolvimento do sistema fonológico das crianças.

¹⁰¹ Cf. D'Angelis em comunicação pessoal.

Na falta de um traço que consiga representar satisfatoriamente a oposição entre vibrante alveolar e *tap*, Ladefoged (1975) defende a existência de um traço [vibrante] que opere tal distinção. Para ele, esse traço deve representar a variação no grau de vibração de um segmento, havendo apenas duas possibilidades: ou um segmento requer vibração para ser produzido ou não. Nesse tratamento, a vibrante alveolar é [+ vibrante] e o *tap*, [- vibrante].

Surge, aí, uma nova questão – insistentemente perseguida por Maia (1982) –, que é como empregar um traço distintivo, para caracterizar diferenças entre uma vibrante alveolar e uma fricativa velar, por exemplo, ao mesmo tempo captando as relações que envolvem tais segmentos¹⁰². A autora aponta para a necessidade de se dispor de um mecanismo – ou representação – fonológico que *expresse a naturalidade das relações entre as variantes dos ‘erres’ nas línguas românicas* (p. 06)

A proposta de Maia é que as relações entre os róticos sejam dadas, num modelo gerativo, pelo traço [rotacizado]¹⁰³. Com isso, ela tenta estabelecer *classes de segmentos a partir do seu comportamento na língua e buscar propriedades fonéticas que as expliquem* (p. 01). Esse traço é concebido como escalar – portanto não mono ou bivalente – sob a alegação de que há uma variação individual, quanto à pronúncia dos róticos, e que uma escala que se estenda de vibrantes múltiplas às vogais e semivogais rotacizadas, passando por vibrantes simples e “fricativas rotacizadas”¹⁰⁴, pode dar conta satisfatoriamente da variabilidade de pronúncia. A opção por tratar [rotacizado] como um traço escalar, e não binário deve-se, segundo a autora, ao fato de que tal tratamento elimina a possibilidade de

¹⁰² As relações entre as vibrantes e as diversas variantes fricativas são notadas, cf. Maia (*op. cit.*) tanto em fatos diacrônicos da língua como em fatos de aquisição da linguagem: a autora menciona que existem substituições ao longo da aquisição que afetam toda a classe dos róticos e não apenas as variantes mais próximas de /l/, tal que *mo[l]eu substitui mo[x]eu ou mo[ɣ]eu aproximadamente na mesma época em que ca[l]a emerge em lugar de cará*. (p. 09)

¹⁰³ Cabe frisar que este traço já era proposto por Ladefoged (1975), porém com o único intuito de representar as vogais rotacizadas do inglês. Maia (1982) lança mão do mesmo traço propondo, porém, modificações, principalmente no que concerne aos sons que [rotacizado] representa – em seu tratamento, ele é estendido a todos os róticos.

¹⁰⁴ É interessante observar que Maia já menciona, nesse artigo, a existência de segmentos intermediários a vibrantes e fricativas, denominados, mais tarde, “vibrantes fricativizadas” (Silva, 1999).

uma associação mais natural das vibrantes às oclusivas e não a outros membros da escala mencionada, como as “fricativas rotacizadas”.

É preciso observar, entretanto, que o tratamento de Maia coloca um problema de ordem teórica para a discussão da representação das líquidas, pois geometrias de traços não concebem a existência de traços escalares, permitindo apenas traços mono ou bivalentes.

De qualquer modo, reconhecendo a existência de um traço específico para representar os róticos – quer o denominemos [rotacizado] ou [vibrante] – obtemos uma saída simples para um fato, tratado de forma tão complexa, na literatura fonológica, já que esse traço capta as relações entre todos os membros da classe dos róticos, inclusive as fricativas.

Um ponto adicional, digno de nota: outras tentativas de representar a oposição /r/ ~ /r/ foram feitas, empregando-se o traço [contínuo]. Essa solução, contudo, é completamente arbitrária, pois, digamos, pode-se alegar que a vibrante é [+ contínuo], em oposição ao *tap*, porque se caracteriza por uma continuidade de vibrações. Por outro lado, também é possível alegar que *tap* é [+ contínuo], porque não demanda seguidas interrupções da passagem do ar no trato. Tal é a arbitrariedade que Mascaró & Bonet (1997) observam, que, enquanto Mascaró (1976) assume que a vibrante é [+ contínuo], Bonet (1988) defende exatamente o oposto¹⁰⁵.

Em resumo, dado que a representação da oposição /r/ ~ /r/ via estrutura silábica se mostra problemática – já que não dispomos sequer de uma definição formal de sílaba – e dado que tratar essa oposição pelos traços [contínuo] ou [tenso] é igualmente problemática, pelas razões que exponho acima, parece mais razoável unificar a representação de todos os róticos por um traço [vibrante], bivalente – a exemplo de [contínuo], [anterior], [distribuído]. Nessa perspectiva, a especificação positiva para esse traço representa segmentos maximamente vibrados, ao passo que a representação negativa para ele concerne

¹⁰⁵ Não tive acesso a tais referências para que possa fazer um comentário mais aprofundado.

segmentos minimamente vibrados¹⁰⁶. Os segmentos maximamente vibrados correspondem a vibrantes, e segmentos minimamente vibrados, a *taps*.

No limite, é possível afirmar que também a representação dos róticos por um traço [vibrante] é arbitrária, pois é preciso lançar mão de um novo traço, o qual poderia receber qualquer outro nome e que se atém apenas a “rotular” os segmentos. O que quero dizer, em suma, é que o traço [vibrante] é, acima de tudo, uma unidade descritiva, e não formal, como é o próprio tratamento dado às líquidas, na teoria fonológica.

2.6 - Uma representação das líquidas dentro de uma geometria de traços

Como mencionado na Introdução, este capítulo privilegia as geometrias de traços, sobre quaisquer outros modelos de análise fonológica, porque as geometrias mantêm uma certa semelhança com a Fonologia Articulatória, que é o modelo pelo qual buscarei uma representação menos arbitrária para as líquidas, sem o conflito entre uso classificatório e uso fonético dos traços.

Para representar as líquidas dentro desse arcabouço teórico, baseio-me na Geometria de Traços de Clements & Hume (1995), por ser este um modelo mais largamente aceito pelos fonólogos. O problema deste modelo, contudo, é a sua dificuldade em tratar os traços de modo (cf. item 5.2). Tal problema poderia, aparentemente, ser resolvido, considerando-se a proposta de Rice & Avery (1991), que é a de unificar nasais e líquidas, numa única classe, a das soantes.

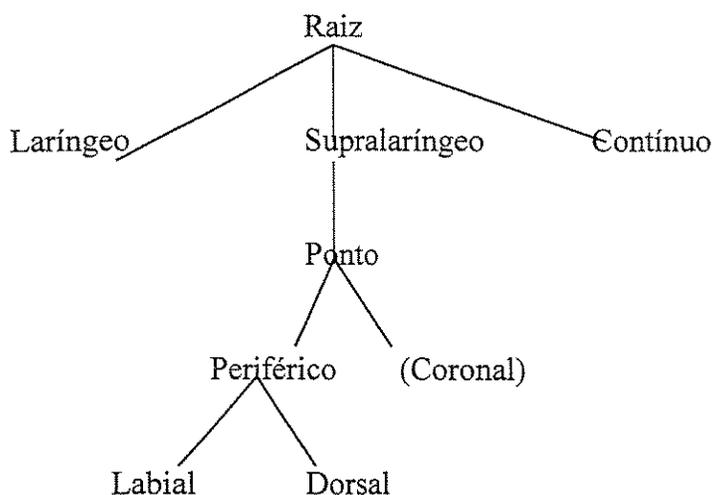
2.6.1 – Unificação de líquidas e nasais numa classe “soantes”

Argumentando que líquidas e nasais constituem uma classe natural, Rice & Avery (1991) defendem a filiação de [lateral] e [nasal] a um nó “vozeamento espontâneo” – ou *Spontaneous Voicing*, doravante SV – característico das soantes. Por conseguinte, esse

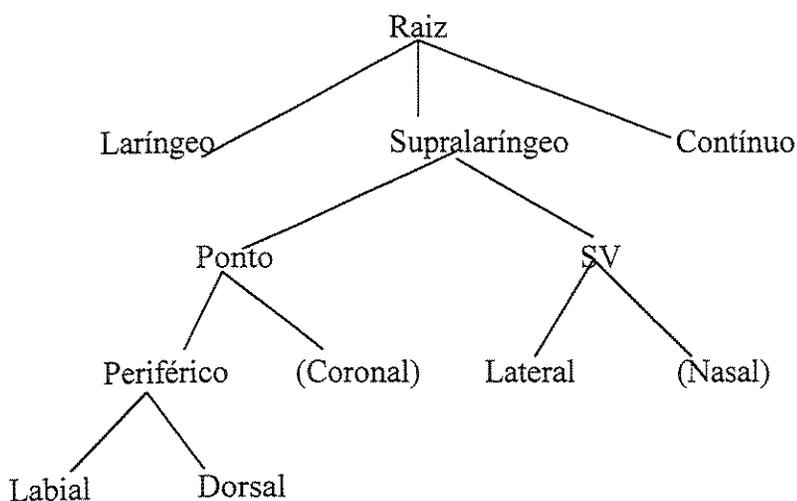
¹⁰⁶ Na verdade, embora propondo traços escalares, Ladefoged (1975) menciona que o traço vibrante só representa fonologicamente duas possibilidades: ou um segmento vibrado ou um segmento não-vibrado. Conseqüentemente, é também nesse modelo um traço bivalente.

modelo prevê representações distintas para obstruintes e soantes, as quais são transcritas abaixo¹⁰⁷:

(i) **Representação das obstruintes**



(ii) **Representação das soantes**



¹⁰⁷ Apesar de a representação proposta ter algumas semelhanças com outros modelos de geometria de traços, como o de Clements & Hume (1995), que seguimos mais de perto, nesta a noção de nós de classe ou nós terminais inexistente e, além disso, aqui todos os traços são monovalentes, o que significa que eles estão ou não presentes na estrutura. Cabe ainda uma última observação acerca da notação dos parênteses, que envolvem nasal e coronal: segundo Rice & Avery, eles seriam o *default* de Vozeamento Espontâneo e ponto, respectivamente, o que torna desnecessária – ou opcional – a sua especificação na representação.

As evidências empíricas que os autores afirmam haver para o nó SV são trazidas de processos de assimilação, dissimilação e assimilação da sonoridade – mas não do ponto – encontrados em algumas línguas. Assim, de acordo com Rice & Avery (*op. cit.*), dados em que há assimilação, como os do inglês, abaixo,

- a. *i[m]balance* *i[n]dentured* *i[ŋ]grown*
 i[m]possible *i[n]tangible* *i[ŋ]credible*
- b. *i[r]rational* *i[l]egible* *i[n]numerable* *i[m]measurable*

explicam-se – em especial os do item (b) – pelo espalhamento do nó SV de /r/, /l/, /n/ e /m/ para a consoante à sua esquerda, i.e., a nasal. Essa operação só é possível porque as nasais são tomadas como *default* de SV, de onde decorre que nasais são o próprio nó SV. As laterais, em contrapartida, são constituídas de um nó SV ao qual se atrela [lateral]. Portanto, há uma posição vazia sob SV das nasais, permitindo que algum traço (no caso [lateral]) se espalhe para elas. A recíproca, porém, não é verdadeira.

Nos casos em que a nasal é seguida por uma consoante obstruinte – como em todos os exemplos do item (a) – as condições da aplicação da regra de espalhamento de SV não seriam satisfeitas, porque não há um nó SV, a partir do qual os traços de soantes pudessem se espalhar – vide estrutura das obstruintes em (i) *supra*.

A dissimilação é justificada por Rice & Avery (*op. cit.*) como um processo oposto: nos dados do kuman,

- yobul* + *na* → *yobutna* (meu osso)
yal + *nga* → *yatgna* (você planta)

haveria o desligamento do nó SV da lateral em seqüências /l/ + /n/, pois o mesmo é adjacente a outro nó SV. Como resultado do processo, emerge aquela que consideram a consoante coronal *default*, [t].

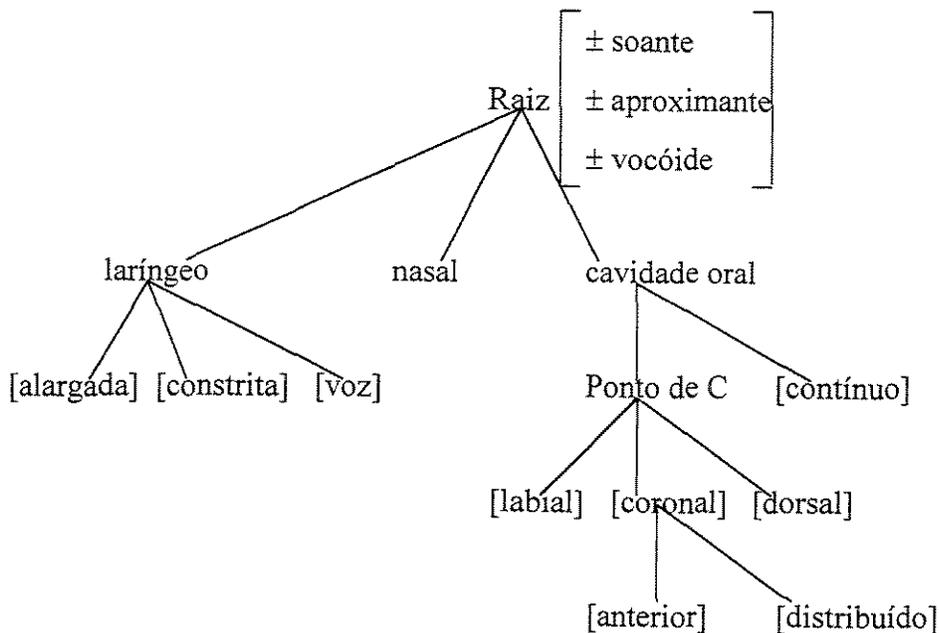
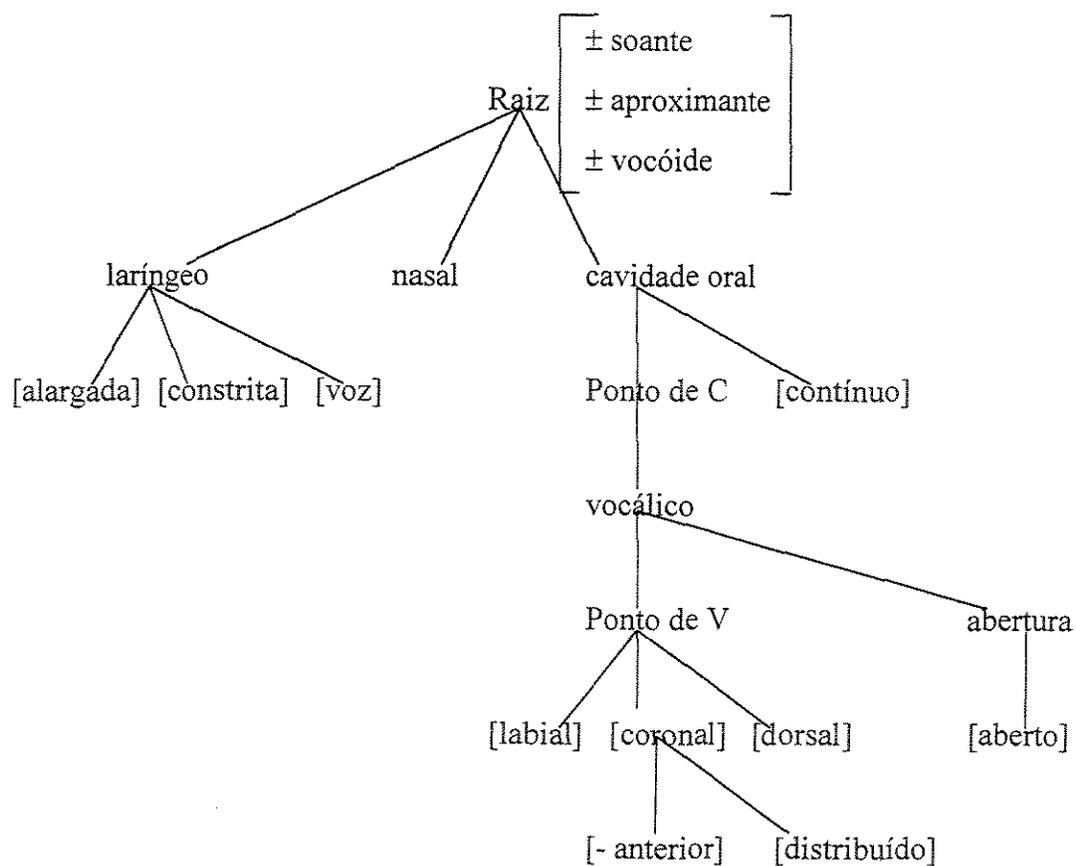
A terceira evidência empírica para o nó SV, i.e., a assimilação de sonoridade – ou cópia – é justificada, a partir de dados do coreano:

- a. *kukmul* → *kuɽmul* (sopa)
- b. *kakmok* → *kaɽmok* (madeira)
- c. *napnita* → *namnita* (germinar, brotar)
- d. *ka^hni* → *kanni* (ser o mesmo)
- e. *tikɽliɽ* → *tikɽliɽ* (as letras t e l)

A interpretação fornecida por Rice & Avery para esses dados é a que, em todos eles, estaria envolvido o espalhamento de SV da nasal ou da lateral para a obstruinte à esquerda. Como ilustrado pela estrutura (i) *supra*, as obstruintes não têm um nó SV, o que preenche uma condição estrutural, para que o nó SV da soante adjacente se espalhe. O sânscrito apresenta processo análogo. Para a análise de alguns dados dessa língua, vide Rice & Avery (1991).

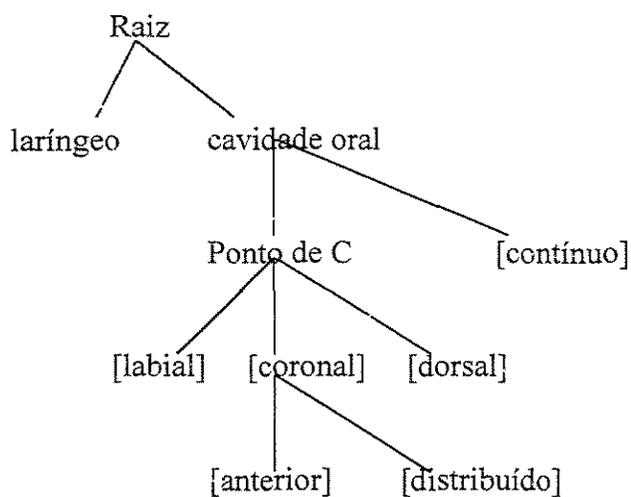
2.6.1.1 – O nó “vozeamento espontâneo” na estrutura interna dos segmentos proposta por Clements & Hume

Já que Clements & Hume (1995) não solucionam a questão dos traços de modo, uma possibilidade seria prever a existência do nó “vozeamento espontâneo” (SV) na estrutura interna dos segmentos, por eles proposta. A princípio, esses autores prevêm duas estruturas segmentais, conforme os segmentos sejam consoantes ou vocóides (vogais ou aproximantes). Transcrevo abaixo tais estruturas (Clements & Hume, 1995: 292):

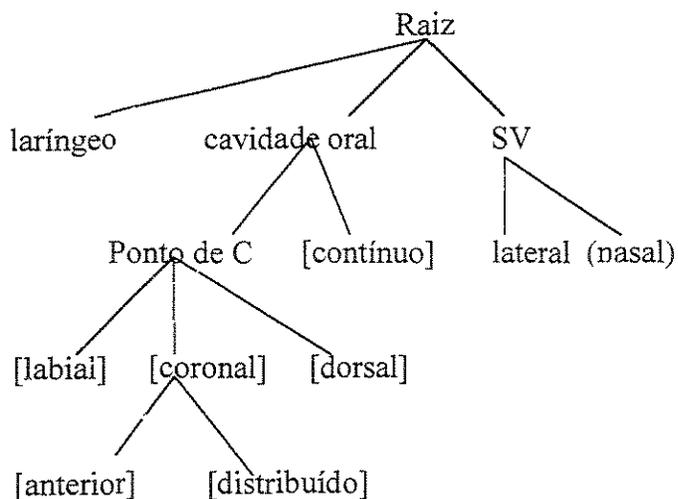
(iii) Consoantes:**(iv) Vocóides**

Inserindo o nó SV nas estruturas (iii) e (iv) *supra*, teríamos como resultado as seguintes representações:

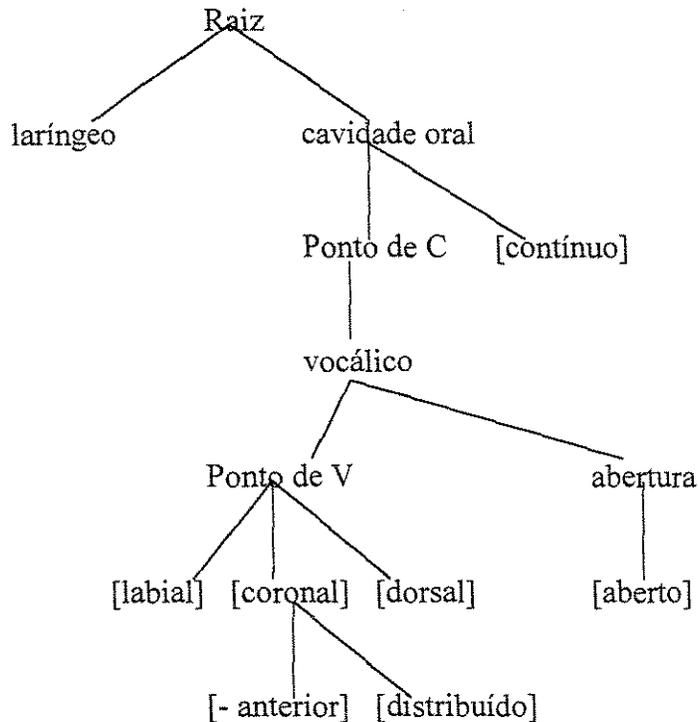
(v) Obstruintes:



(vi) Soantes



(vii) Vocóides



As principais diferenças entre as estruturas originalmente propostas por Clements & Hume e estas três últimas, além da evidente inserção do nó SV, são: 1) desaparecem os traços do nó Raiz; 2) não há unificação entre as representações das consoantes; ao contrário, distinguem-se agora obstruintes de soantes; 3) o nó nasal é tomado como *default* de SV, como apontado pela notação parentética, e nos mesmos moldes do que propõem Rice & Avery (vide nota 127); 4) o nó (nasal) desaparece da estrutura de vocóides, já que é previsto sob SV.

O desaparecimento dos traços do nó Raiz é decorrência direta da inserção do nó SV que, como se pode observar em (vi), é típico das soantes. Assim sendo, tornar-se-ia redundante caracterizar as soantes, ou quaisquer das outras classes, também pelos traços da Raiz. A consequência dessa manobra é que as classes não são definidas, na posição mais alta da hierarquia estrutural, mas abaixo dela: a classe das soantes é definida diretamente sob o nó Raiz, ao passo que obstruintes e vocóides se definem sob o nó cavidade oral. No

caso da distinção entre obstruintes e vocóides é ao nó “vocálico” – presente exclusivamente na estrutura interna de vocóides – que cabe estabelecer a diferenciação.

Quanto à estrutura interna do nó SV, ela segue à risca a proposta de Rice & Avery, segundo a qual (nasal) é o modo de articulação *default*, o que implica que o nó SV é inerentemente nasal, ou seja, a não especificação de qualquer traço sob SV significa que se trata da estrutura de uma consoante nasal. Caso a estrutura represente uma lateral, então (nasal) é suprimido.

2.6.1.2 – Problemas do nó “vozeamento espontâneo”

Se a unificação de nasais e laterais, pelo nó SV, aparentemente tem a vantagem de resolver o dilema sobre o que fazer com esses traços, ela tem uma série de desvantagens.

Em primeiro lugar, ela não prevê a caracterização dos róticos. Rice & Avery (*op.cit.*) chegam a sugerir, *en passant*, que isso se faça por um [*R-feature*], exclusivo para esses segmentos. No entanto, não mencionam nada, além de que esse traço, “caso exista”, deveria também ser previsto sob SV.

Como decorrência óbvia desse fato, as estruturas (v) e (vi) *supra* deixam de captar a identidade de classe entre vibrantes e fricativas, algo insistentemente perseguido por Maia (1982), conforme discutido em 2.5.3.2.

Desaparece também a identidade entre nasais e obstruintes, por exemplo, fato atestado por processos assimilatórios recorrentes nas línguas, como é o caso da assimilação do ponto da nasal, relativamente ao ponto de articulação da obstruinte adjacente.

Além disso, é impossível prever, em (vii), vogais nasais, uma vez que nasal desaparece da estrutura interna dos vocóides, por ser previsto sob SV.

Uma última ressalva a ser feita concerne à própria existência do nó SV: apesar das evidências arroladas por Rice & Avery para sustentá-la, parece problemático assumir esse como um nó universal.

Passo a expor as razões pelas quais não concordo com o nó SV, ao menos para línguas como o português e o espanhol. Em primeiro lugar, as descrições apresentadas no

Handbook of the International Phonetic Association (1999) não apontam as alternâncias entre róticos e nasais ou entre laterais e nasais, como fatos freqüentes. Em segundo lugar, embora haja tais alternâncias em línguas “ibéricas”, elas são bem menos freqüentes do que as alternâncias entre laterais e róticos, como afirma Lloret (1997), a partir, e.g., dos seguintes dados:

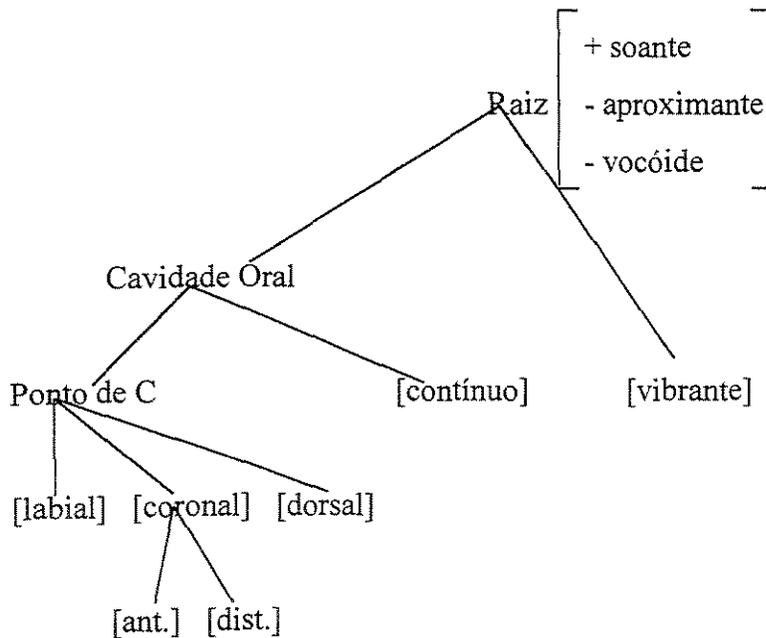
1. *bona nit* > *bora nit* (‘boa noite’ – registro coloquial do catalão)
2. *nonanta* > *noranta* (‘noventa’ – evolução histórica do catalão)
3. *Barcinone* > *Barcelona* (evolução histórica do catalão e do espanhol)
4. *monument* > *moliment* (‘monumento’ – dialeto do catalão, falado ao sul da Catalunha)
5. *ninho* > *linho* (português europeu coloquial)
6. *unicorne* > *licorne* (evolução histórica do português europeu)
7. *qualsevol* > *consevol* (‘qualquer um’ – registro coloquial do catalão)
8. *farigola* > *fanigola* (‘tomilho’ – registro coloquial do catalão)
9. *galeria* > *galenia* (‘galeria’ – registro coloquial do catalão e do espanhol)

Lloret afirma que, dentre as alternâncias acima, as mais comuns são as que provocam a mudança de nasais e laterais para um rótico, como nos exemplos 1 e 2. Casos como os dos exemplos 3, 4, 5 e 6, i.e., alternância entre nasal e lateral, são menos freqüentes. A alternância entre lateral e nasal, expressa no item 7, é tão rara quanto a alternância entre rótico e nasal, dos itens 8 e 9. (É preciso ressaltar que estas observações sobre a freqüência das alternâncias são comuns a outras línguas ibéricas, abordadas pela autora.)

Como essas alternâncias são menos produtivas do que as alternâncias entre laterais e róticos, é mais pertinente propor que nasais e líquidas constituam duas classes distintas.

2.6.1.3 – Outra representação das líquidas numa geometria de traços

Já que a tentativa de unificar nasais e líquidas numa classe soantes, se mostra ineficiente, dentro do modelo de Clements & Hume, pelas razões apontadas na seção anterior, e que dizem respeito não apenas às perdas que trazem para o próprio modelo – como a falta de unificação entre as classes consonantais ou a impossibilidade de representar a estrutura de vogais nasais – mas também à inexistência de dados suficientes para se sustentar essa unificação, uma outra alternativa é tentar representar as líquidas, no modelo de Clements & Hume, por traços distintivos, sem ter de se recorrer a um novo nó de classe. A partir das discussões, apresentadas desde a seção 2.4, essa alternativa, aparentemente mais razoável, é a seguinte:



As principais diferenças entre essa representação e a que Clements & Hume propõem para as consoantes (cf. (i), item 2.6.1.1) são: 1) descarta-se a existência de um traço [lateral]; 2) a oposição entre laterais e róticos é dada pelo traço [contínuo]; 3) prevê-se um traço [vibrante], que reúne róticos e os diferencia entre si (ao menos aqueles que demandam vibração de um articulador para serem produzidos). Este traço nem sequer era

cogitado em Clements & Hume (1995) e segue a concepção do traço [vibrante] de Ladefoged (1975).

Deve ficar claro que não estou discutindo toda a estrutura interna dos segmentos, estabelecida por Clements & Hume; o que faço apenas é tentar “enxugar” essa representação, recorrendo a um inventário mínimo de traços. O único traço, inserido na estrutura segmental, é [vibrante], o que não significa um “inchaço” do modelo ou tampouco a exigência de um processamento mais complexo.

O que deve ser dito sobre a estrutura proposta acima, primeiramente, é que a classe das líquidas é já caracterizada pelos traços que constituem o nó Raiz: [+ soante, + aproximante, - vocóide] isola essa classe das obstruintes – [- soante, - aproximante, - vocóide] -, das aproximantes - [- soante, + aproximante, - vocóide] – e dos vocóides [- soante, - aproximante, + vocóide]. Poderia ser argumentado que, dessa maneira, líquidas não se diferenciam de nasais, o que aparentemente recolocaria a questão sobre a unificação dessas classes numa única. Realmente, a oposição não se dá pelos traços que constituem o nó Raiz, mas pela presença ou ausência do traço [nasal], atrelado ao nó Raiz.¹⁰⁸

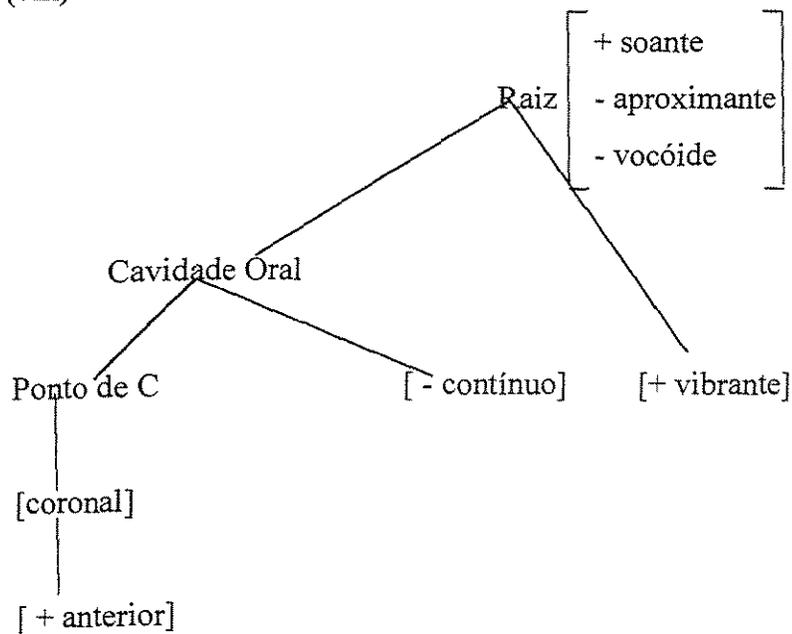
Quanto à diferenciação entre laterais e róticos, ela se dá pelo traço [contínuo] e não por [lateral]. Conforme já havia observado na seção 2.5.1.1, análises como a de Spencer (1984) suportam a visão de que [lateral] pode ser suprimido da estrutura segmental, sem prejuízo da representação dessa classe, já que o traço [contínuo], mais operacional nos processos fonológicos das línguas, dá conta dessa tarefa. Conseqüentemente, por esta proposta, segmentos especificados positivamente para [contínuo] são laterais; os especificados negativamente para esse traço são róticos.

A oposição entre as laterais, e.g., entre lateral alveolar e lateral palatal, deve ser dada pelo ponto de articulação: a lateral alveolar é [coronal, + anterior] e a palatal, [coronal, - anterior].

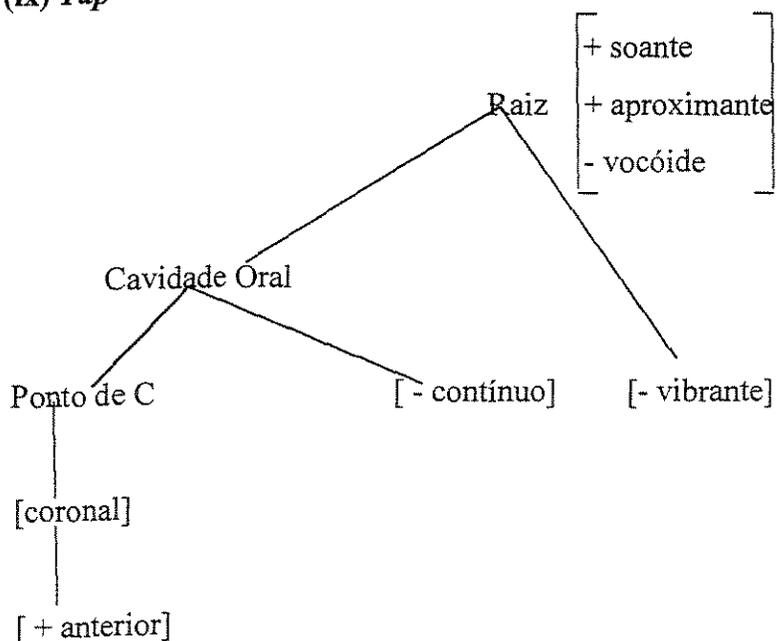
¹⁰⁸ Devo lembrar que existem controvérsias a respeito do local onde atrelar [nasal]. No entanto, essa discussão foge do escopo do meu trabalho.

Para representar a oposição /r/~r/, esta alternativa conta com duas estruturas segmentais possíveis, diferenciadas exclusivamente pelo valor do traço [vibrante] – este inserido na representação, pelas razões expostas em 2.5.3.2:

(viii) Vibrante alveolar:



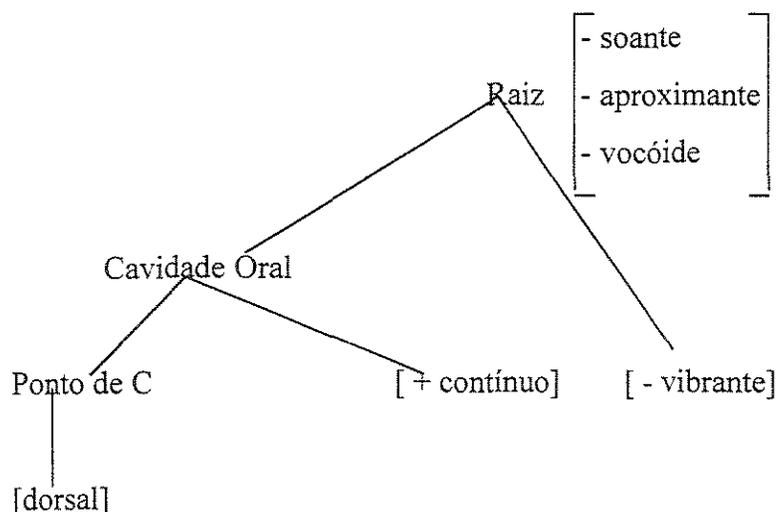
(ix) Tap¹⁰⁹



Na estrutura de uma vibrante uvular, ou de uma vibrante bilabial, o que muda, relativamente à representação (viii) é o ponto de articulação. Logo, uma vibrante uvular é também [+ vibrante] e, no entanto, [dorsal], enquanto uma vibrante bilabial é [labial]. A vantagem de uma tal representação, portanto, é captar as relações entre todas as vibrantes – inclusive as que não ocorrem em português.

Resta, ainda, verificar se a representação atende à exigência de Maia (1982) e capta também a semelhança entre vibrantes e suas variantes fricativas, como a fricativa velar [x], largamente disseminada como variante da vibrante alveolar, na maior parte do território brasileiro. Como eu já havia mencionado, no início desta seção, a classe das líquidas é caracterizada no nó Raiz, pela especificação dos traços que as constituem. Assim, fricativas diferem de vibrantes pelo traço soante, positivo para estas e negativo para aquelas. Nessa proposta que venho elaborando, a representação para a fricativa velar é:

(x) Fricativa velar



Assumo, portanto, que o traço vibrante não está restrito à classe das líquidas, podendo inclusive ocorrer na estrutura de obstruintes. Pressupor que as fricativas velares sejam especificadas em sua estrutura interna para o traço [vibrante] capta uma semelhança

¹⁰⁹ Pode haver também uma variante [- aproximante], além da que cito aqui, [+ aproximante].

histórica – além de uma identidade fonológica –, entre fricativas e vibrantes, semelhança atestada por estudos dialetológicos não só para o PB (e.g., Callou, 1987; Callou, Moraes & Leite, 1997) como também para diversos dialetos do espanhol (vide, e.g. Quilis; 1992; Obediente, 1998) e que envolve posteriorização e fricativização da vibrante alveolar, na direção da fricativa velar.

Uma questão inerente à representação (x) *supra* é: o traço [vibrante] deve permanecer com especificação positiva ou sua especificação deve ser negativa? Talvez seja mais plausível considerar que a fricativa velar é [- vibrante], pois, na definição do traço (vide 2.5.3.2 *supra*), mencionávamos que os segmentos maximamente vibrados, i.e., as vibrantes alveolares, são [+ vibrantes], ao passo em que os minimamente vibrados são [- vibrantes]. Quanto a este tópico, entretanto, ainda não estou certa, e ele seguramente mereceria discussão.

Restam ainda alguns problemas, na proposta apresentada: 1) não representa fricativas laterais; 2) não explica processos assimilatórios – ou trocas – envolvendo laterais e róticos; 3) não dá conta de representar segmentos intermediários a vibrantes e fricativas.

Quanto à representação das fricativas laterais, talvez pudesse ser argumentado que sua representação se faça analogamente à de fricativas velares, i.e., o caráter fricativo do segmento é definido pela especificação dos traços que constituem o nó Raiz. No entanto, surge o seguinte problema: as fricativas são especificadas positivamente para o traço [contínuo], em oposição às oclusivas. Como, então, expressar a lateralidade do segmento, se tal característica é prevista como decorrência de [contínuo]?

Voltando-nos agora para a explicação de processos assimilatórios, como os que ocorrem historicamente no português, espanhol, catalão ou italiano (vide Quadros 2.1 e 2.2 – item 2.2), ou nas alternâncias expostas nos Quadros 2.3 e 2.4 (item 2.2): a proposta sequer toca na maneira como se dariam tais processos. Certamente, seguindo o arcabouço teórico das geometrias de traços, isso aconteceria por operações de desligamento e/ou espalhamento de traços. Restaria ainda, contudo, investigar como são essas operações.

Por fim, há o problema de que a alternativa, proposta em 2.6.2, não dá conta de representar segmentos intermediários a vibrantes e fricativas. Este tópico será abordado mais detidamente, na última seção.

2.7 – Sumário e discussão

A despeito da abundância de evidências empíricas para a oposição entre laterais e róticos, ou ainda entre vibrantes e *taps*, apresentada nas seções 2.2 e 2.3, a Geometria de Traços sequer toca nessas oposições. O modelo tampouco se preocupa em tentar capturar a identidade de classe, existente entre laterais e róticos, ou ainda entre os segmentos róticos.

Ele se atém tão somente a identificar uma classe “líquidas”, diferenciada das demais pela combinatória dos três traços que constituem o nó Raiz, i.e., [soante, aproximante, vocóide]. Ora, fazer isto pouco mais é do que reeditar a velha caracterização das classes maiores de sons, como feito já por Jakobson & Halle (vide seção 2.4). O máximo previsto pelo modelo de Clements & Hume é a existência de um traço [lateral], cujo lugar na estrutura interna dos segmentos eles não conseguem definir, conforme mostrado em 2.5.2.

Essa limitação do modelo contradiz a proposta, anunciada por Clements & Hume, segundo a qual o modelo teria o poder de oferecer uma ligação plausível entre a representação fonológica e a realização fonética dos segmentos, e acaba criando um conflito entre o uso classificatório e o uso fonético dos traços distintivos. Ou seja, os traços não captam toda a realidade fonética das líquidas, atendo-se exclusivamente a “rotular” alguns segmentos como pertencentes à classe das líquidas.

Mesmo indo aos limites da Geometria de Traços, no sentido de tentar formalizar as oposições verificadas entre laterais e róticos e entre vibrantes e *taps*, a alternativa de representação, apresentada em 2.6.1.3, neste estudo, sofre igualmente de arbitrariedade, como conseqüência do conflito apontado acima. Dito de outro modo, propor que a oposição entre laterais e róticos se dê pelo traço [contínuo] – lembro que [lateral] é abandonado devido principalmente à sua pouca operacionalidade nas línguas do mundo (cf. 2.5.1.1) – e

que a oposição entre vibrantes e *taps* se dê por um traço [vibrante] – adotado com base nas razões levantadas em 2.5.3.2 – ainda parecem distinções construídas *ad hoc*.

Retomando a crítica que faço à distinção entre /l/ e /r/ pela oposição contínuo/descontínuo, no modelo de Jakobson & Halle, é como se, na alternativa, apresentada em 2.6.1.3, se “desse um jeito” de encaixar as oposições entre laterais e róticos e entre vibrante e *tap*, na estrutura interna dos segmentos, proposta por Clements & Hume (1995), devido à ausência de outros meios, para formalizar tais oposições. Isto torna a proposta essencialmente descritiva, mas pouco explicativa. Logo, perde-se a possibilidade de “traduzir” de forma eficaz a realização fonética na representação fonológica.

Por outro lado, a Geometria de Traços parece não deixar saída: qualquer proposta resultaria arbitrária, devido principalmente ao distanciamento desse modelo dos fatos fonéticos – por mais que Clements & Hume advoguem o contrário, isso é falso para as líquidas.

Um outro problema da alternativa, proposta em 2.6.1, é que ela não dá conta de representar a gradiência, verificada na produção dos róticos, como a que relato no capítulo seguinte. Embora a gradiência deva ter uma representação na gramática da língua, pois decorre da estrutura prosódica das sentenças nas quais os róticos ocorrem, não há como serem representados por uma estrutura segmental, como a que forneço. E a razão dessa impossibilidade está no fato de que, sendo os traços binários, têm-se apenas duas alternativas: uma vibrante ou uma fricativa, nunca os dois, ao mesmo tempo.

Poderia ser argumentado que uma maneira de resolver esse impasse seria conceber traços escalares, também na Geometria de Traços, analogamente ao que propunha Ladefoged (1975).

Sugerindo a incorporação do detalhe fonético à representação fonológica, esse autor defende a possibilidade de prover a representação fonológica de uma especificação fonética, em termos da porcentagem dos valores dos traços distintivos, de modo que cada traço seja concebido como uma escala física. Assim, no caso do traço [vibrante], seria teoricamente possível estabelecer uma escala que se estendesse de [0 vibrante] a [100 vibrante], devendo

os valores intermediários dessa escala refletir segmentos intermediários a um maximamente vibrado e outro minimamente vibrado.

Essa, contudo, não é uma solução. Primeiramente, porque a escala se referiria apenas ao efeito da vibração, e nas vibrantes espirantizadas (vide terceiro capítulo), há também efeito de fricção. A princípio, necessitaríamos de uma outra escala, estendendo-se de [0 fricativo] a [100 fricativo], que pudéssemos sobrepor à primeira. A sobreposição das duas escalas, por sua vez, é uma manobra inconcebível no modelo.

Outro problema de concebermos traços escalares é que não se verificam, nas alofonias de /r/, em início de palavra, porcentagens de fricção, isto é, não há um segmento [85 vibrante] ou [50 fricativo]. Verifica-se, ao contrário, uma gradiência entre vibrantes e fricativas, podendo ocorrer *n* vibrantes espirantizadas entre elas, a depender do grau de ruído fricativo sobreposto. Logo, não é difícil concluir que tal fato demandaria *n* categorias intermediárias a, por exemplo, [0 fricativa] e [1 fricativa]¹¹⁰. Assim, voltamos ao impasse decorrente, ao fim e ao cabo, de um problema inerente aos modelos de análise fonológica, em geral, e resultante do fato de

(...) a visão corrente das relações entre Fonética e Fonologia ser ainda inteiramente taxonômica: alguns fenômenos são classificados como fonéticos e outros como fonológicos, de tal forma que o mapeamento dos dois níveis torna-se em princípio arbitrário. Deste modo, cabe aos traços a tarefa ingrata de mediar entre dois níveis cuja relação é em si problemática. (Maia, 1982: 02)

Frise-se que, apesar dos insistentes argumentos da literatura fonética – em trabalhos como os de Keating (1985), Browman & Goldstein (1986, 1990, 1992), Sproat & Fujimura (1993), Gick (1999) – de que os traços distintivos não dão conta de mapear o físico no cognitivo, modelos como o de Clements & Hume (1995) insistem em advogar para si o

¹¹⁰ As restrições do contínuo observado na produção dos róticos de início de palavra, ao qual me refiro, não decorrem de fatores físicos, mas são induzidas por fatos lingüísticos. No caso do seu tratamento via traços, introduzir-se-ia circularidade, dado o caráter nominativo dos traços.

poder de estabelecer um elo entre a representação cognitiva da fala e sua manifestação física.

Os estudos, citados neste capítulo, e os dados que apresento, no capítulo seguinte, sugerem a necessidade de se abandonar os traços distintivos, em favor de outros primitivos teóricos, que realmente dêem conta de estabelecer esse elo. Para tanto, esses primitivos devem admitir a “dinâmica” inerente a segmentos intermediários a duas categorias, como é o caso das vibrantes espirantizadas, presentes na fala de um dos sujeitos do experimento fonético-acústico relatado no próximo capítulo. Adotando-se os “gestos articulatórios”, a mediação entre o nível fonético e o fonológico pode dar-se de maneira menos problemática, como mostra Albano (2001), a partir de diversos fatos do PB e análogos à gradiência, envolvida na variabilidade de pronúncia dos róticos, em posição inicial de palavra, a qual abordo no terceiro capítulo e para a qual esboço uma proposta de representação gestual, no quarto capítulo.

3 – A ALOFONIA CONTÍNUA DOS RÓTICOS INICIAIS EM PB: EVIDÊNCIAS PARA UM TRATAMENTO DINÂMICO DOS PROCESSOS FÔNICOS

Os róticos (sons da classe dos *rr*) apresentam uma grande variabilidade de pronúncia, seja entre línguas ou numa mesma língua. Trubetzkoy (1939)¹¹¹, e.g., já observava que, em alemão, antecedendo vogais /r/ podia ser uma vibrante dental, uma vibrante uvular ou, ainda, uma fricativa velar, sem que se estabelecesse oposição fonêmica entre essas variantes. Nas demais posições, podia ser pronunciado como uma velar “incompletamente articulada” ou como uma vogal não-silábica. Ainda segundo Trubetzkoy, em gilyak, língua falada na Sibéria Oriental e na parte norte da Ilha Sakhalin, /r/ podia ser pronunciado como fricativa surda e, no caso de ser articulado “energicamente”, diz Trubetzkoy que “*uma fricção tipo-3 pode claramente ser ouvida*”.

No PB também há grande variabilidade de pronúncia dos róticos, seja em posição inicial, seja em final de sílaba/palavra. Este fato tem sido alvo de estudos que abordam a distribuição das diversas variantes, como os de Callou *et alii* (1997) e Monaretto (1997).

No primeiro, os autores tratam da distribuição das variantes de /r/ em cinco capitais - Porto Alegre, São Paulo, Rio de Janeiro, Recife e Salvador -, observando um forte condicionamento geográfico sobre essa distribuição, tal que São Paulo e Porto Alegre tendem a privilegiar variantes vibrantes, enquanto as demais capitais privilegiam fricativas¹¹². Monaretto (1997) investiga a distribuição da vibrante nas capitais dos três estados da região sul do Brasil, comparando dados do Atlas Lingüístico-Etnográfico da

¹¹¹ Duas observações acerca dessa referência a Trubetzkoy: a) o emprego do tempo passado dos verbos é proposital, já que certamente houve mudanças nas línguas mencionadas e eu não disponho de informações a respeito da realização atual dos róticos nelas; b) a citação *supra* é extraída da tradução de D’Angelis (inérita), dos *Grundzüge der Phonologie*, a partir da tradução para o inglês de Christiane A.M. Baltaxe.

¹¹² É preciso lembrar que o estudo de Callou *et alii* (*op. cit.*) se baseia em dados do Projeto Norma Urbana Culta (NURC), colhidos na década de 70. Portanto, um estudo dialetológico mais atual possivelmente revelasse uma outra distribuição das variantes de /r/, de modo que apenas Porto Alegre privilegiasse ainda a vibrante alveolar e, em São Paulo, se observasse a disseminação da fricativa velar. Nesse sentido, Monaretto

Região Sul (ALERS) e do VARSUL. A autora observa, nesses *corpora*, a coocorrência do que denomina “vibrante anterior” e “vibrante posterior”. Pelas suas descrições¹¹³, pode-se inferir que se trata, respectivamente, da vibrante apical e de um som semelhante a uma fricativa.

Em comum entre esses estudos, há dois pontos: o primeiro deles diz respeito à evidência de que coexistem – num mesmo dialeto e entre dialetos do PB – diversas variantes de /r/, em posição inicial. O segundo ponto reforça uma observação de Callou (1987), que diz haver mudança em curso no PB, envolvendo a produção de /r/. De acordo com a autora, as variantes vibrantes – apicais ou uvulares¹¹⁴ – são variantes conservadoras que sofrem mudança, em seu ponto e modo de articulação, rumo a variantes fricativas, que são inovadoras.

Parece estar envolvido, nessa mudança, um enfraquecimento da vibrante. Digo enfraquecimento porque a vibrante perde seu caráter balístico – que lhe é conferido pelas sucessivas batidas da ponta da língua na região alveolar do trato –, em favor de um caráter contínuo, dado pela fricção.

Curioso é observar que essa variação não é atual e que a coexistência de variantes de /r/, num mesmo dialeto, era já observada por Gonçalves Viana, no final do século XIX. Relata ele:

Ele (/r/ múltiplo¹¹⁵) se pronuncia um pouco mais para trás que o r simples e é geralmente lingual. Podemos encontrar individualmente r vibrantes uvulares, mesmo entre aqueles que pronunciam o r simples como lingual [...] Algumas vezes

(*op. cit.*), já observa a disseminação da fricativa velar, também no dialeto de Porto Alegre. Este fato está de acordo com a mudança histórica, no sentido da espirantização da vibrante, relatada por Callou (1987).

¹¹³ Diz Monaretto (1997:17-18): *A pronúncia variável da vibrante será estudada em relação a quatro variantes, definidas por modo e zona de articulação. Sons apicais vibrantes, fricativos e aproximantes, que são articulados com a ponta da língua em direção aos dentes incisivos até o ponto palato-alveolar da boca, são enquadrados na categoria ‘vibrante anterior’. Sons produzidos com o dorso da língua levantado e a ponta da língua abaixada e pressionada contra os dentes incisivos inferiores, junto ao véu palatino ou à úvula, ou sons feitos na cavidade faringal, são classificados como ‘vibrante posterior’.* As duas outras variantes à quais a autora se refere são o *tap* e o “zero fonético”.

¹¹⁴ As vibrantes uvulares eram encontradas, na época do estudo de Callou (*op. cit.*), no dialeto carioca, mas já estavam caindo em desuso.

pronuncio o r inicial como uma fricativa sonora, uma espécie de rz [...] Encontrei raramente esta particularidade na pronúncia de outras pessoas portuguesas. Este r fricativo sonoro é entretanto bastante freqüente na pronúncia de brasileiros [...]; não saberia dizer, todavia, até que ponto esta pronúncia é individual ou dialetal: eu a notei entre naturais de Pernambuco e de São Paulo. (Essai de phonétique et de phonologie de la langue portugaise d'après de dialecte actuel de Lisbonne, 1883, apud Callou, 1987:10. Os grifos são meus.)

É interessante, na descrição de Gonçalves Viana, ele notar a existência de uma “variante intermediária”, digamos assim, a vibrantes e fricativas - a “fricativa sonora, uma espécie de rz” – que nem o estudo de Callou *et alii* (*op. cit.*) nem o de Monaretto (*op. cit.*) mencionam ao se reportarem a dialetos nos quais ocorrem a variante conservadora, vibrante, e a inovadora, fricativa. A ausência de menção a este fato se justifica, em certa medida, pela impressão auditiva que constitui a base dos estudos variacionistas¹¹⁶.

A existência de variantes intermediárias a vibrante e fricativa, em posição inicial de sílaba, é sugerida por dados acústicos de Silva (1996). Este achado foi, mais tarde, confirmado por outros dados (vide Albano *et alii*, 1998), colhidos junto a um novo informante. Os novos dados exibem, inclusive, um contínuo físico entre vibrante e fricativa, em início de palavra, e ao longo do qual se estende a variabilidade de pronúncia de /r/.

Finalmente, dados de um terceiro informante confirmam a existência do contínuo físico e, além disso, sinalizam que a emergência de variantes tendendo a vibrantes ou de variantes enfraquecidas, tendendo a fricativas, é condicionada pela estrutura prosódica do enunciado a qual, por sua vez, mapeia informações da estrutura sintática das sentenças contendo os segmentos em questão. Por isso, é possível postular que o contínuo físico ao longo do qual se estende a variabilidade de pronúncia de /r/, neste estudo, tem de ser representado no léxico, de modo que parâmetros fonéticos quantitativos sejam controlados

¹¹⁵ Esta é a denominação utilizada como referência à vibrante apical, em contraposição a “r fraco”, o *tap*.

¹¹⁶ Digo “em certa medida” porque a descrição de Gonçalves Viana era também impressionística. Com um detalhe: ele, obviamente, não dispunha de ferramentas robustas de análise, como os *softwares*, que se encontram hoje no mercado.

por parâmetros gramaticais qualitativos. Surge daí a questão de como, então, representar tal contínuo, na gramática de uma língua.

Neste capítulo apresento, primeiramente, os dados de Silva (1996) e Albano *et alii* (1998), os quais mostram que a variabilidade de pronúncia de /r/ se estende ao longo de um contínuo físico. Em seguida, apresento e discuto os novos dados que confirmam o contínuo, envolvido na variabilidade de pronúncia, e apontam para sua pertinência lingüística.

3.1 – Indícios favoráveis à natureza contínua da variabilidade de pronúncia dos róticos

A figura 3.1, abaixo, traz espectrogramas para três realizações do mesmo logatoma. Esses dados, que fazem parte do *corpus*, analisado em Silva (1996), foram colhidos junto ao informante ARM, paulistano, do sexo masculino e com 52 anos, na época da coleta de dados.

Cabe um esclarecimento: o experimento fonético-acústico desse estudo (Silva, *op. cit.*) não havia sido elaborado com o intuito de capturar a alofonia contínua dos róticos. A única pretensão, na época, era realizar uma descrição acústica das líquidas do PB, a partir da fala do informante ARM¹¹⁷. Para isso, foram analisadas cinco repetições de um *corpus*, contendo logatomas mono e dissílabos, com as líquidas em posição intervocálica e/ou em início de palavra (*taps* e laterais palatais ocorriam no *corpus*, apenas em posição intervocálica, para refletir a sua distribuição no PB¹¹⁸). A variabilidade de pronúncia da vibrante, atestada pelos espectrogramas da Figura 3.1, pôde ser verificada em três das cinco repetições do logatoma “varrá”.

¹¹⁷ O fato de o informante ter 52 anos garantia que ele produzisse a vibrante alveolar [r], presente majoritariamente na fala de paulistanos com mais de 50 anos, na época, e um dos alvos da análise, então realizada.

¹¹⁸ A análise dos dados foi realizada no LAFAPE, empregando-se o *Computerized Speech Laboratory* (CSL), modelo 4300B, da Kay Elemetrics.

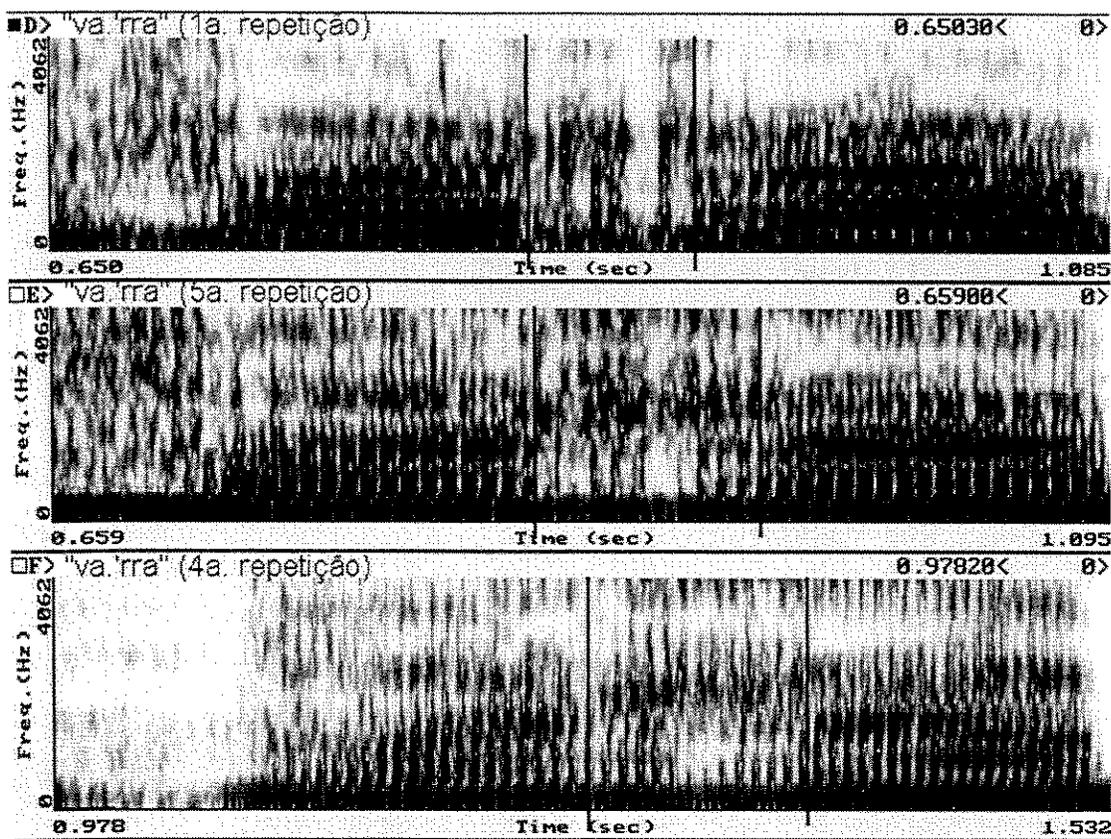


Figura 3.1 - Espectrogramas de três repetições do logatoma "varra", produzidas pelo informante ARM. Nota-se sobreposição do ruído fricativo à estrutura acústica da vibrante alveolar [r].

Observe-se que, neste caso – diferentemente de todos os demais que serão apresentados, neste capítulo –, a vibrante alveolar [r] se encontra no meio da palavra, em início de sílaba. De qualquer modo, é possível verificar, no espectrograma no topo da figura (janela D), entre barras verticais contínuas, uma “vibrante canônica”, com aberturas orais e fechamentos bem definidos¹¹⁹.

No espectrograma da janela E – no meio da figura – vê-se um segmento no qual ainda é possível identificar resquícios dos fechamentos e as aberturas orais da vibrante, inclusive com uma estrutura formântica. Porém, já é possível notar energia de produção numa região de alta frequência, por volta dos 4 kHz. Essa energia se realiza de modo muito

¹¹⁹ Seguindo Recasens (1991), denomino fechamento ao momento em que a voz quase desaparece, devido ao aumento da pressão subglótica, necessária para a realização da vibrante. No espectrograma, esse evento pode

semelhante ao ruído fricativo. O espectrograma da janela F traz, finalmente, uma estrutura na qual as aberturas orais são menos nítidas do que nas janelas anteriores e na qual há maior sobreposição de ruído fricativo. Há, portanto, uma gradação, na sobreposição do ruído fricativo à vibrante alveolar.

Dadas essas pistas da existência de contínuo físico, na espirantização da vibrante, era preciso verificar se o mesmo fenômeno ocorria na fala de mais informantes. Para tanto, colhi dados junto a um novo informante, OGF, também do sexo masculino, mas natural de Bofete (região centro-oeste do estado de São Paulo). O informante, um padre, tinha 41 anos, na época da coleta de dados, e foi escolhido porque, numa primeira impressão auditiva, parecia produzir vibrantes alveolares, em alternância com fricativas velares.

Houve dois procedimentos distintos para a coleta dos dados: fala espontânea e leitura de texto¹²⁰. Durante a fala espontânea, fiz perguntas ao informante sobre assuntos vários, nos moldes dos questionários sócio-lingüísticos labovianos (para detalhes acerca dessa metodologia, vide Tarallo, 1986). Para a leitura, selecionei um texto,¹²¹ relacionado à atividade do informante e que continha diversas ocorrências de /r/, em início de sílaba/palavra. Solicitei ao informante que lesse o texto três vezes. Entre cada leitura, eu fazia perguntas, versando sobre o próprio texto, e que tinham o intuito de evitar um possível viés, decorrente do emprego de uma mesma estratégia de leitura.

Os dados foram colhidos numa cabine com tratamento acústico¹²² e, em seguida, foram digitalizados no LAFAPE, utilizando-se para isso o *Computerized Speech Laboratory* (CSL), modelo 4300B, da Kay Elemetrics. A taxa de amostragem dos dados foi fixada em 10 kHz.

Análises visuais – tanto dos dados de fala espontânea, como dos dados de leitura – revelam que a vibrante alveolar [r] é a variante mais recorrente, na fala de OGF, em posição inicial de sílaba/palavra, tal como verificado nos dados de ARM. Também a exemplo dos

ser identificado como um espaço quase que em branco. Aberturas orais são eventos nos quais a voz é retomada e onde é possível, inclusive, reconhecer uma estrutura formântica definida.

¹²⁰ A adoção dessa metodologia tinha o único objetivo de inspecionar a fala do informante, buscando verificar se ele alternava variantes de /r/, em posição inicial de sílaba/palavra, e se essa alternância se estendia ao longo de um contínuo físico.

¹²¹ Trata-se do Evangelho de São João, capítulo 20, versículos de 1 a 9.

dados de ARM, os de OGF exibem alternância entre vibrante e fricativa e, mais ainda, segmentos intermediários a ambas, com diferentes graus de sobreposição de ruído fricativo à estrutura acústica da vibrante. Isso configura um contínuo físico, ao longo do qual se estende a variabilidade de pronúncia de /r/.

Os fatos, descritos no parágrafo anterior, foram constatados nas duas etapas de coleta dos dados. Os espectrogramas da Figura 3.2, abaixo, retratam o contínuo, presente na situação de fala espontânea:

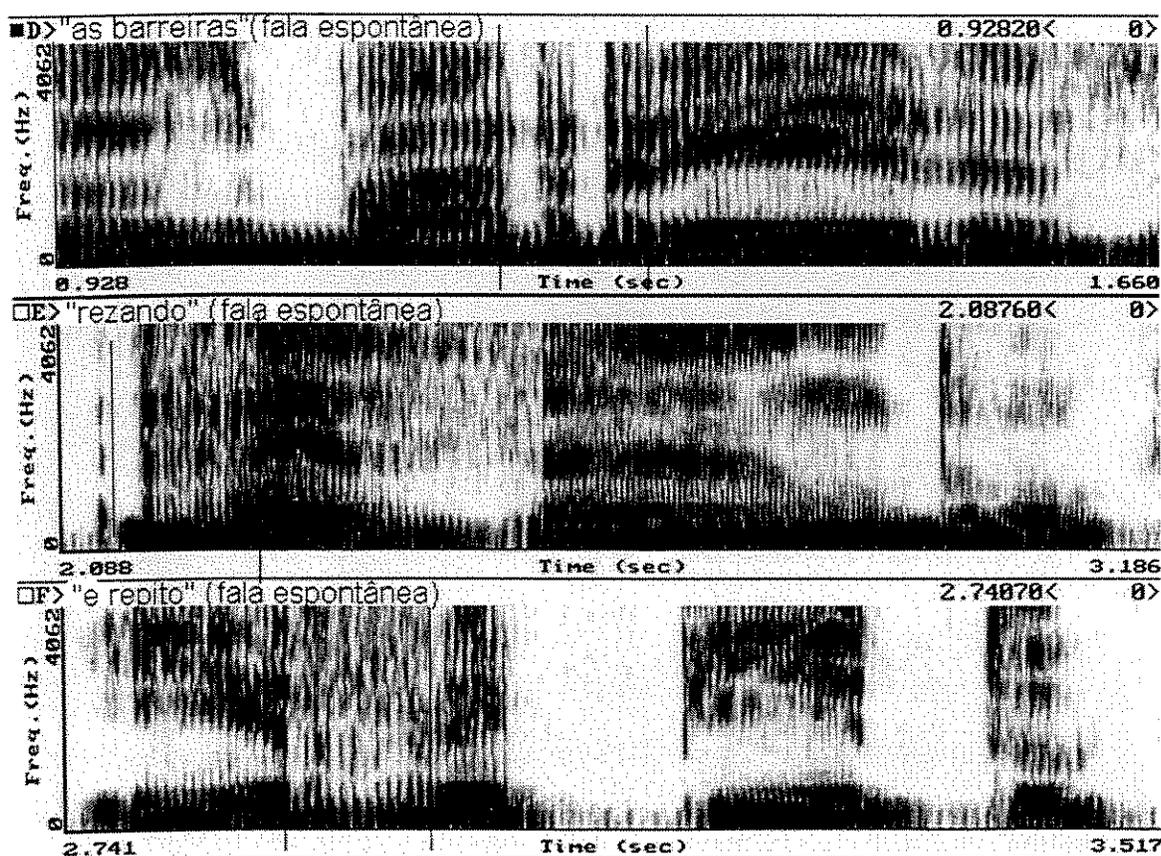


Figura 3.2 - Contínuo físico, na espirantização da vibrante, para os dados de OGF em fala espontânea: vibrante canônica (janela D), vibrante espirantizada (janela E) e fricativa (janela F).

No topo da figura, na janela D, tem-se, entre linhas verticais contínuas, uma vibrante apical canônica: é possível reconhecer, nesse segmento, um primeiro fechamento que

¹²² No estúdio da Rádio PRF8, de Botucatu (SP).

antecede uma abertura oral e, em seguida, um novo fechamento e uma nova abertura oral. A segunda abertura oral se une ao ditongo [ej], imediatamente subsequente. Isto pode ser inferido pela observação da trajetória dos formantes, no momento seguinte ao segundo fechamento.

A janela E, no meio da figura, traz, também entre linhas verticais contínuas, a vibrante espirantizada, na qual é possível identificar momentos semelhantes às aberturas orais – note-se o segundo e terceiro formantes bem definidos, neste segmento¹²³ – aos quais se sobrepõe ruído fricativo – atente-se para a energia de produção difusa e com grande amplitude, estendendo-se até 4 kHz.

A janela F, ao pé da Figura 3.2, traz, por fim, uma fricativa: observe-se o ruído contínuo, numa região de alta frequência do espectro. Diferentemente da vibrante espirantizada da janela E, não é possível reconhecer aqui uma estrutura formântica. A alternância entre aberturas e fechamentos, que se vê enfraquecida, mas ainda presente na janela E, é completamente ausente na janela F.

A tarefa de leitura, por sua vez, acrescenta novas informações sobre o contínuo físico sobre o qual se estende a variabilidade de pronúncia dos róticos.

¹²³ Cf. Silva (1999), para o PB, a frequência média dos formantes da vibrante alveolar, em início de palavra, fica em torno de 290 Hz (F1), 1290 Hz (F2) e 2000 Hz (F3), nas duas primeiras aberturas orais. Na terceira abertura oral, F1 sobe para uma média de 330 Hz, ao mesmo tempo que o valor de F2 decresce para 1150 Hz e o valor de F3, para 1900 Hz. Esses formantes não assumem, necessariamente, os mesmos valores dos formantes da vogal adjacente. Pode haver, entretanto, coarticulação entre a vibrante e a vogal adjacente (cf. Silva, 1996), donde resulta uma variação das médias dos valores da frequência dos formantes de /r/, em função dos valores da frequência dos formantes da vogal adjacente. Vale ainda notar que o F2 da vibrante é, em média, mais baixo que o F2 do *tap*, o que conferiria a esse segmento uma característica mais “centralizada” do que o *tap*. Essa mesma diferença é notada por Recasens para o catalão (1991).

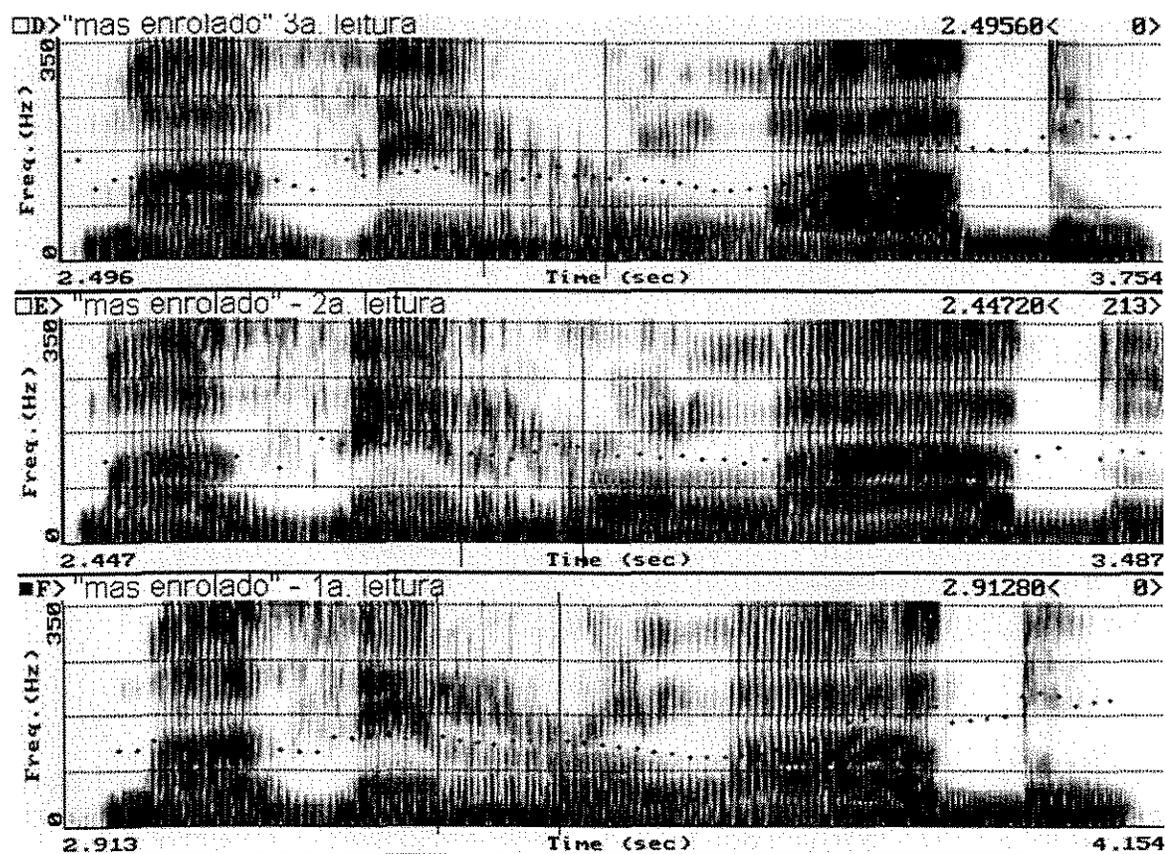


Figura 3.3 - Dados de OGF, exibindo o contínuo físico que envolve a variabilidade de pronúncia de /r/ em início de sílaba: vibrante canônica (janela D) e vibrantes espirantizadas (janelas E, F), com diferentes graus de sobreposição do ruído fricativo.

A janela D, no topo da Figura 3.3, traz a vibrante canônica: como na janela D das Figuras 3.1 e 3.2, é possível reconhecer, entre linhas verticais, uma vibrante, constituída por fechamentos e aberturas orais que se repetem.

A janela E, no meio da figura, traz, por sua vez, uma vibrante espirantizada: como na vibrante espirantizada da Figura 3.2, é possível identificar aqui momentos semelhantes às aberturas orais, aos quais se sobrepõe ruído fricativo, que se estende numa faixa de alta frequência no espectrograma. Observe-se que, no segmento em questão, nessa janela, diferentemente da janela D, há energia de produção difusa e na faixa dos 4kHz.

A janela F, na base da Figura 3.3, traz um segmento no qual há ainda maior sobreposição de ruído fricativo à vibrante: vê-se uma estrutura mais contínua que as duas

anteriores e, à semelhança do espectrograma da janela E, ruído fricativo, em torno dos 4kHz.

Até aí, nada de novo, com relação aos dados de fala espontânea. No entanto, o exame da forma de onda de cada uma das três repetições sinaliza o emprego de diferentes estratégias, em cada leitura do texto. Observe-se:

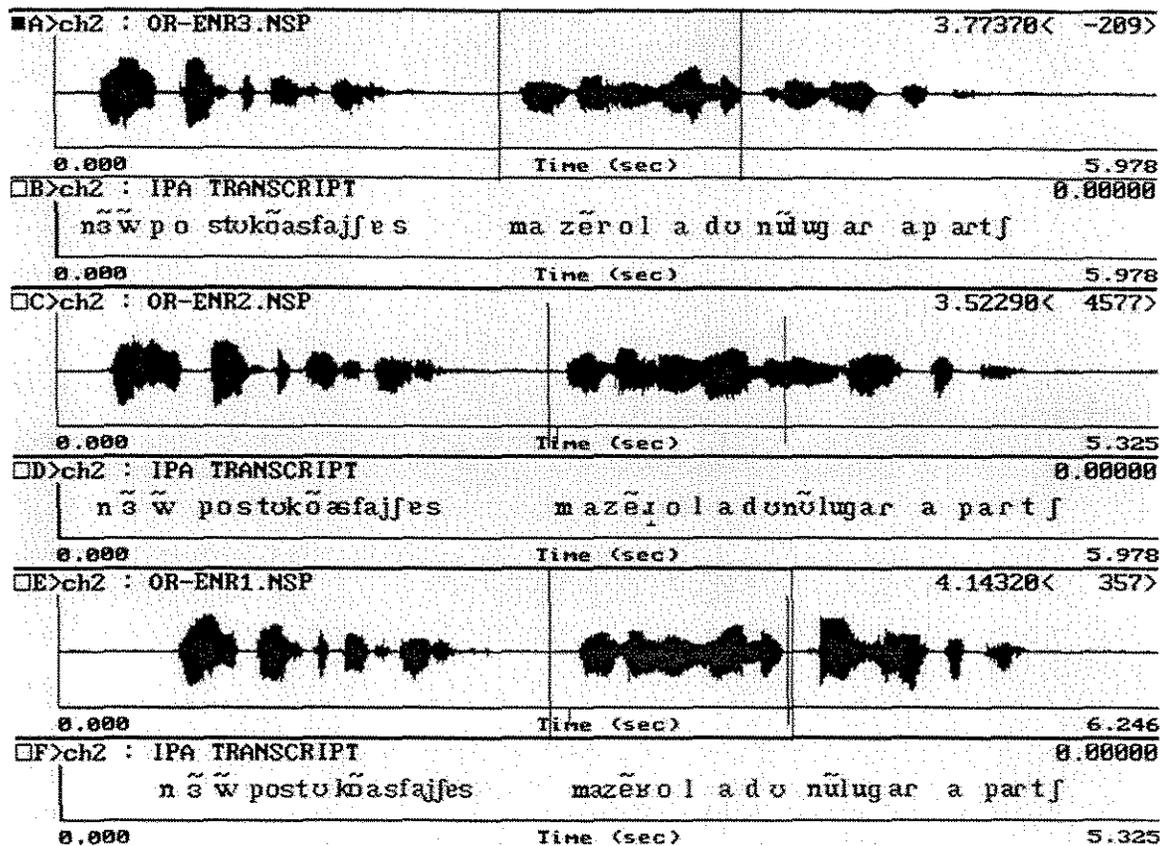


Figura 3.4 - Forma de onda de cada uma das três leituras do sintagma "mas enrolado". A ordem de apresentação das formas de onda obedece à ordem de apresentação dos espectrogramas da Figura 3.3. As transcrições para cada enunciado encontram-se nas janelas, abaixo das janelas das respectivas formas de onda.

Na figura acima, "mas enrolado" é destacado por linhas verticais contínuas, o que torna fácil perceber que, enquanto na terceira e na primeira leituras, OGF faz pausa entre "mas enrolado" e "num lugar à parte"¹²⁴; na segunda leitura, ele não realiza essa pausa.

¹²⁴ Esses sintagmas integram a sentença "viu as faixas de linho deitadas no chão e o pano que tinha estado sobre a cabeça de Jesus, não posto com as faixas, mas enrolado num lugar à parte." (in Jo 20, 7).

Podemos sinalizar as diferentes leituras – na ordem em que ocorrem nos espectrogramas e respectivas formas de onda – da seguinte maneira:

- 1º. espectrograma: mas enrolado # num lugar # à parte
- 2º. espectrograma: mas enrolado num lugar # à parte
- 3º. espectrograma: mas enrolado # num lugar # à parte (onde # sinaliza pausa)

Na terceira leitura, /r/ é realizado como vibrante apical, diferentemente das duas outras leituras, como é evidenciado pelos espectrogramas da Figura 3.3. Considere-se também que uma análise de outiva sugere que OGF faz ênfase sobre “enrolado”, apenas nessa leitura¹²⁵.

Há, então, dois fatos que parecem operar conjuntamente, na realização da vibrante apical: a colocação de pausa e a ênfase sobre a palavra contendo o segmento em questão. Pode-se dizer que a ênfase faz o segmento ser fortalecido. Nas outras duas leituras, a ausência de ênfase faz o rótico ser produzido de modo enfraquecido, o que traz como resultado a sobreposição de ruído fricativo à estrutura acústica da vibrante, em diferentes graus.

Apesar de pouco consistentes quanto à influência de colocação de pausa sobre a realização dos segmentos, os dados de OGF são claros na confirmação do contínuo físico sobre o qual se estende a variabilidade de pronúncia de /r/.

A literatura fonética, porém, evidencia condicionamento de fronteira prosódica sobre a natureza dos segmentos adjacentes a ela: Keating *et alii* (1999) mostram efeitos prosódicos agindo sobre as características articulatórias dos segmentos, de modo que há fortalecimento dos mesmos, quando ocorrem no início do que os autores denominam “domínios prosódicos altos”, i.e., enunciado, frase entoacional e frase acentual. Palavra e sílaba são considerados, nessa perspectiva, “domínios prosódicos baixos”.

Os resultados de Sproat & Fujimura (1993) vão na mesma direção, embora não se faça referência, nesse estudo, a “domínios prosódicos”, mas a “força de fronteira”. Os

¹²⁵ Esses dados, bem como todos os outros, tomados como exemplos nesta tese, encontram-se no CD anexo.

autores reportam que, quando adjacente a fronteiras mais fortes – como a que se estabelece entre vocativo e a sentença imperativa, e.g. –, a lateral de final de palavra, no inglês, tende à variante *dark*. Inversamente, fronteiras prosódicas fracas – e.g. a fronteira entre radical e afixo, no interior de palavra – fazem emergir variantes de /l/ tendendo a *light*¹²⁶.

Dadas as pistas do condicionamento da estrutura prosódica sobre a natureza dos segmentos, elaborei, em seguida, um experimento acústico com o intuito de verificar se esse condicionamento se mostra recorrente.

3.2 – Contínuo fonético na variabilidade de pronúncia dos róticos: prosodicamente condicionado?

Para confirmar a existência do contínuo físico na produção de /r/ e investigar um possível condicionamento prosódico sobre ele, foram colhidos dados junto a dois novos informantes. Um deles, a exemplo de OGF, alterna em sua fala vibrantes e fricativas¹²⁷. Esse informante, também do sexo masculino, é natural de Tibagi (PR), tendo residido a maior parte da vida em Curitiba. Tinha 49 anos, na época da coleta dos dados, o que garante que ele ainda produza vibrantes, já que, no dialeto falado no Sul do país, a vibrante ainda é encontrada, na fala de pessoas dessa faixa etária. JLM é professor universitário e morou na França entre 1974 e 1976 e, em seguida, entre 198 e 1979.

O outro informante do experimento fonético-acústico, AJS, alterna vibrantes e *taps*, talvez como decorrência de proceder de uma região de colonização ítalo-germânica. Ele é do sexo masculino, natural de Erechim (RS) e tinha 66 anos, na época em que os dados

¹²⁶ Como menciono no primeiro capítulo – item 1.3.1.3 –, considera-se, tradicionalmente, que a variante *dark* seja uma realização velarizada de /l/, ocorrendo em final de sílaba/palavra. A variante *light*, em contrapartida, ocorreria em início de sílaba/palavra e não é velarizada. Sproat & Fujimura (1993) mostram, porém, haver um contínuo físico envolvendo essas categorias de sons, tal que quanto maior a força de fronteira adjacente a /l/, mais esse segmento tende a *dark*. Quanto menor a força de fronteira adjacente a /l/, porém, mais o segmento tende a *light*. Os autores mostram, ainda, através de estudos articulatórios, que não há velarização de /l/ final, mas um gesto de dorso de língua com maior magnitude que o gesto de ponta de língua. (A produção de /l/ *light* envolve uma relação inversa entre os gestos de ponta e dorso de língua, responsáveis pela produção da lateral.)

foram colhidos. AJS, como JLM, é professor universitário e viveu na Alemanha, durante três anos.

Conforme argumento, no item 3.3.3 adiante, o contínuo, verificado na fala de ambos os informantes, envolve enfraquecimento dos segmentos, embora não na mesma direção: para JLM, o enfraquecimento se dá via abandono do gesto de ponta de língua e concomitante espirantização, conforme será definido abaixo. Para AJS, em contrapartida, o gesto de ponta de língua é preservado, mas vibra menos vezes.

Procurando investigar o possível condicionamento prosódico sobre a variabilidade de pronúncia de /r/ e com base em Sproat & Fujimura (*op. cit.*), o experimento acústico, relatado neste item, manipula a força da fronteira, adjacente ao rótico. Levanta-se a hipótese de que fronteiras prosódicas fortes favoreçam a ocorrência de segmentos também fortalecidos, ao passo que fronteiras prosódicas fracas devem favorecer a realização de segmentos enfraquecidos¹²⁸.

3.2.1 – *Design experimental*

O *corpus* do experimento acústico manipula a força da fronteira, numa cadeia de palavras, através do emprego de oito “pares mínimos prosódicos”, i.e., enunciados idênticos no nível segmental, porém distintos no nível prosódico. Eis um exemplo¹²⁹:

¹²⁷ Em um único dado, dos 64 disponíveis para esse informante, ocorre um *tap*, que poderia ser interpretado como um enfraquecimento da vibrante e que, contrariamente à fricativa, preserva o gesto balístico de ponta de língua.

¹²⁸ Note-se que não emprego a terminologia “hiperarticulado” ou “hipoarticulado”, cunhada por Lindblom (1990): evito tal terminologia porque não me parece que seja o caso, nos dados verificados nesta tese, de que um segmento não atinja seu alvo, ou o ultrapasse. Caso se considerasse a vibrante um segmento hiperarticulado, tanto para JLM como para AJS, e em oposição às demais variantes, todas hipoarticuladas, teríamos de resolver a seguinte questão: qual versão, dentro do contínuo, presente na fala de cada informante, seria o segmento *default*, quer dizer, aquele cujos articuladores não ultrapassam o alvo nem deixam de alcançá-lo?

A.1 - Aquela funcionária da biblioteca te atendeu mal?

B.1 - **Eu não achei a funcionária ríspida**. Ela até que foi bem delicada comigo.

A.2 - Você achou as duas funcionárias que você tava procurando?

B.2 - **Eu não achei a funcionária ríspida**. Só a funcionária delicada fica lá na hora do almoço.

Os “pares mínimos prosódicos” aos quais me refiro são as sentenças grafadas em negrito, nos mini-diálogos¹³⁰ acima; as cadeias de palavras examinadas (“seqüências-alvo”) estão grifadas.

A opção pelos pares mínimos prosódicos segue proposta de Guimarães (1998), segundo a qual as estruturas sintáticas são, de alguma forma, mapeadas na estrutura prosódica das sentenças e esse mapeamento é determinado por diferentes relações de c-comando. Assim, num enunciado, o *default* é que várias palavras estejam associadas por relações de c-comando, constituindo o que Guimarães (*op.cit.*) denomina “oração fonológica”. As relações de c-comando, porém, podem-se quebrar, em algum ponto dessa oração, resultando daí duas orações fonológicas, separadas por uma fronteira, a qual o autor denomina “fronteira de π ”.¹³¹

Nos pares mínimos prosódicos, a presença ou não da fronteira de π decorre das diferentes relações de c-comando, que resultam no contraste do escopo do adjetivo. Então, em B.1, “ríspida” é predicação secundária do NP “funcionária”, não havendo entre esses

¹²⁹ O *corpus* todo, que inclui oito pares mínimos prosódicos e dezenove pares de distratores, encontra-se, em anexo, no final desta tese.

¹³⁰ Os mini-diálogos foram utilizados como meio de direcionar a interpretação que se pretende que o falante faça do adjetivo nas sentenças-alvo: ora o adjetivo assume a função de predicativo do objeto, ora de adjunto adnominal, dependendo do contexto semântico-pragmático. A inexistência desse contexto impossibilitaria a distinção.

¹³¹ A proposta de Guimarães (*op. cit.*) dá um meio mecânico para se determinarem as unidades prosódicas de um enunciado, i.e., existem regras para fazer o mapeamento sintático na estrutura prosódica. Por isso, essa proposta é mais afeita a um estudo experimental do que as propostas – consagradas – de autores como Marina Nespor e Irene Vogel (vide, e.g., Nespor & Vogel, 1986) ou Elisabeth Selkirk (vide, e.g., Selkirk, 1984), nas quais os domínios prosódicos são determinados de modo apriorístico. Ressalto que a adoção da proposta de Guimarães não significa assumir a sua discussão acerca do mapeamento sintático na estrutura prosódica. Isto foge ao escopo desta tese.

dois termos relação de c-comando, como mostra a árvore da Figura 3.5a, abaixo. Em decorrência, emerge uma fronteira de π , nessa *small clause*.

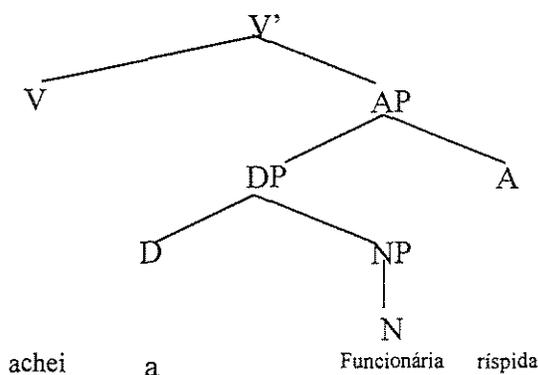


Figura 3.5a - "Rispida" como predicação secundária do NP "funcionária: não há relação de c-comando entre esses dois termos. Espera-se, portanto, uma fronteira forte entre eles.

Por outro lado, em B.2, "funcionária" e "rispida" constituem um único NP, estabelecendo-se relação de c-comando entre os termos, como se vê nas árvores das Figuras 3.5b, abaixo. Em decorrência, não se estabelece fronteira de π entre "funcionária" e "rispida". Espera-se, conseqüentemente, ausência de pausa entre os termos.

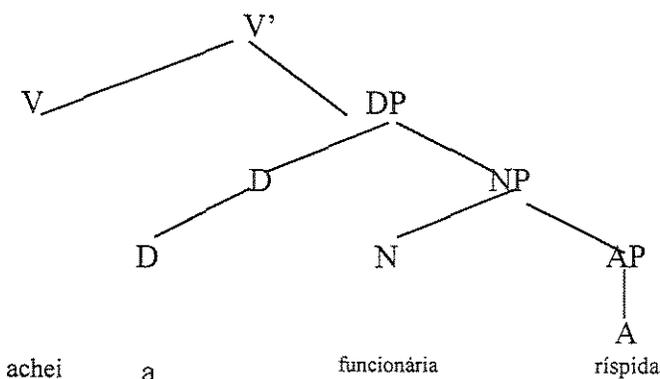


Figura 3.5b - "Funcionária" e "rispida" constituem um mesmo NP, estabelecendo-se entre eles relação de c-comando. Conseqüentemente, não há fronteira de π entre os termos. Nesta representação, o determinante fica fora do NP.

Fazendo um paralelo entre a tese de Guimarães e a "hierarquia de força de fronteira" de Sproat & Fujimura, assumo, que nos casos em que se estabelece uma fronteira de π – ou

onde se tem um adjetivo predicativo do objeto –, há uma fronteira prosódica forte, em oposição a uma fronteira prosódica fraca, nos enunciados onde a fronteira de π não emerge – i.e., onde o adjetivo é adjunto adnominal. Uma vez que fronteira prosódica forte favorece a produção de segmentos relativamente fortalecidos e fronteira fraca, segmentos relativamente enfraquecidos – ainda com base em Sproat & Fujimura (*op.cit.*) –, nossa hipótese é de que, nos enunciados onde o adjetivo é predicativo do objeto, devem ser produzidas variantes tendendo a vibrantes – para ambos os informantes – e, nos enunciados com adjetivo funcionando como adjunto adnominal, devem ser produzidas variantes tendendo a fricativas, nos dados de JLM e *taps*, nos dados de AJS.

Além do controle sobre a estrutura sintático-prosódica dos pares mínimos, foi controlada a estrutura segmental e acentual das seqüências-alvo, como garantia da ausência de qualquer viés, decorrente de diferenças acentuais e/ou segmentais, que impossibilitasse comparação entre os pares. Assim, quanto ao ambiente acentual, em três pares, as seqüências-alvo contêm substantivos e adjetivos paroxítonos e, em outros quatro pares, os substantivos e adjetivos são proparoxítonos¹³². Quanto à estrutura segmental, a vogal átona final dos substantivos é sempre /a/, e a vogal seguinte a /r/ se alterna entre /i/, /e/, /a/, /o/ e /u/. As seqüências-alvo dos pares mínimos prosódicos, empregadas neste experimento, são:

substantivos e adjetivos paroxítonos	substantivos e adjetivos proparoxítonos
criança risonha	funcionária ríspida ¹³³
menina rebelde	técnica rígida
camisa rasgada	música rápida
	cômoda rústica

Quadro 3.1 - Seqüências-alvo contidas nos pares mínimos prosódicos do experimento fonético-acústico.

¹³² No par “novela romântica”, que, embora não conste do Quadro 3.1, constitui também o conjunto dos pares mínimos prosódicos, tem-se, excepcionalmente, substantivo paroxítono e adjetivo proparoxítono.

¹³³ Observe-se que a afirmação de que o substantivo “funcionária” é proparoxítono se baseia numa transcrição fonética larga em que um dos critérios é considerar hiatos os ditongos crescentes. Também de acordo com os critérios dessa transcrição, desconsideram-se as vogais epentéticas que emergem em encontros “oclusiva-nasal”, como o que se tem no substantivo da linha de baixo.

Além dos 8 pares mínimos prosódicos, o *corpus* tem 19 pares de distratores, elaborados de maneira análoga, mas sem controle da estrutura segmental e acentual, na seqüência-alvo. Para a coleta de dados, os pares foram desmembrados e os mini-diálogos foram apresentados aos informantes, em ordem aleatória, tendo-se porém o cuidado de não permitir que os mini-diálogos de um mesmo par ocorressem em seqüência. Era solicitado aos informantes que lessem perguntas e respostas dos mini-diálogos. Essa tarefa foi repetida quatro vezes, por cada informante.

Os dados foram colhidos numa cabine acusticamente tratada do UNIVIDEO da Universidade Federal do Paraná (UFPR)¹³⁴. Em seguida à coleta, os dados foram digitalizados a uma taxa de 20 kHz, no CSL modelo 4300B, do LAFAPE/IEL/UNICAMP. Observe-se a mudança na taxa de amostragem para a digitalização do sinal (os dados de OGF tinham sido digitalizados a uma taxa de 10 kHz): isso se justifica pelo fato de a maior taxa de amostragem permitir uma análise mais acurada das fricativas (cf. Shaddell & Mair, 1996), que são sons cuja energia de produção se concentra em faixas de alta freqüência no espectro.

3.3 – Análise dos dados

Os dados de AJS e JLM foram analisados no CLS4300B, a Kay Elemetrics, no LAFAPE. Utilizei também, eventualmente, o Multi-speech, modelo 3700, também fabricado pela Kay¹³⁵.

Por motivos expositivos, passo a comentar separadamente a análise dos dados de cada um dos sujeitos. A discussão dos dados, no final do capítulo, retoma o conjunto.

¹³⁴ Agradeço ao pessoal do UNIVIDEO (Setor de Educação da UFPR) a presteza e eficiência com que realizaram a gravação dos dados, utilizados neste estudo.

¹³⁵ O Multi-speech realiza as mesmas tarefas do CSL. A diferença entre ambos está no fato de que o CSL possui um *hardware* externo para captura de dados, ao passo que o Multispeech utiliza uma placa de som qualquer para o mesmo fim. Portanto, a qualidade de digitalização dos dados no Multi-speech é proporcional à qualidade da placa de som. Todas as outras tarefas são realizadas com a mesma qualidade do CSL.

3.3.1 – Contínuo físico na variabilidade de pronúncia dos róticos: análises visuais e espectros de FFT¹³⁶ nos dados de JLM

A análise dos espectrogramas dos dados de JLM permite constatar a existência de variantes intermediárias à vibrante e à fricativa, às quais denomino, doravante, “vibrantes espirantizadas”. A figura abaixo ilustra essa variabilidade e traz três ocorrências do sintagma “funcionária ríspida”, em diferentes repetições:

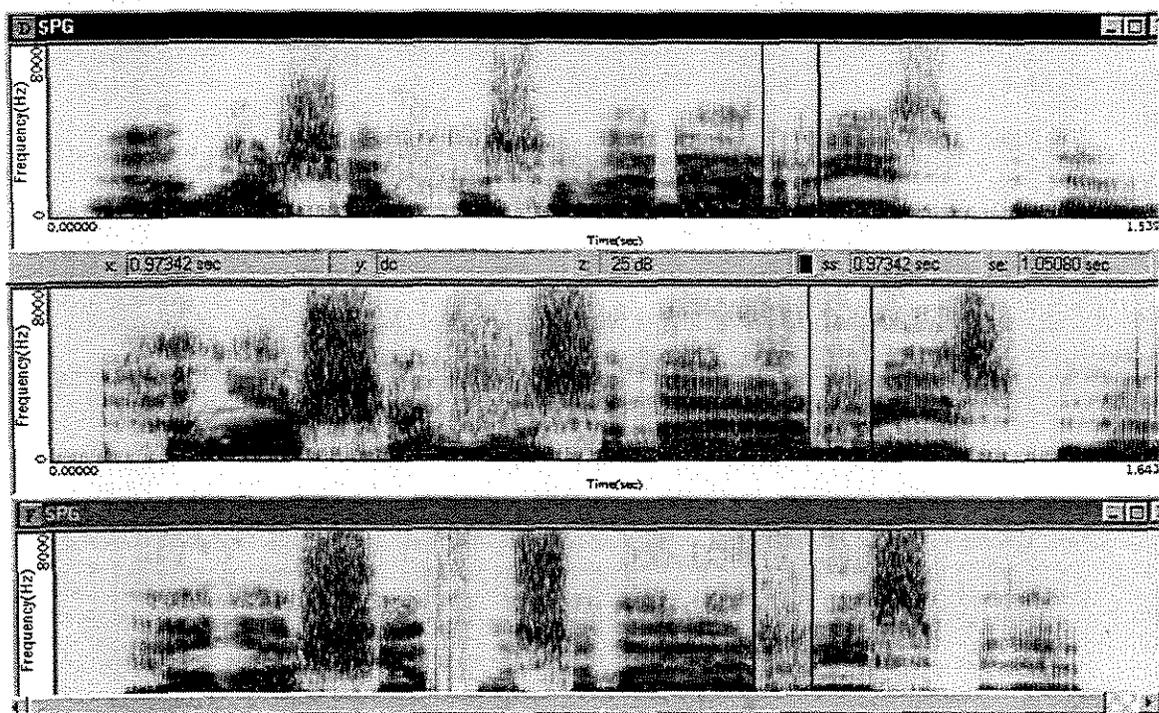


Figura 3.6 - Entre barras verticais contínuas, vibrante (janela superior), vibrante espirantizada (janela do meio) e fricativa (janela inferior), realizadas na produção do sintagma "funcionária ríspida". Na janela do meio, o adjetivo é predicativo do objeto; nas janelas superior e inferior, “ríspida” assume a função de adjunto adnominal.

¹³⁶ A *Fast Fourier Transform*, ou “transformada rápida de Fourier”, é um algoritmo computacional que, baseado no teorema do matemático Fourier, faz uma decomposição da forma de onda do sinal acústico digitalizado e, portanto, discretizado, em pontos no tempo. Como resultado, tem-se um espectro mostrando a amplitude de cada harmônico da frequência fundamental, em função do tempo. Cabe ressaltar que o número de pontos transformados, i.e., os pontos em que a FFT amostra o sinal acústico, é sempre potência de 2 e expressa a largura de banda da análise: quanto maior o número de pontos, mais estreita a largura da banda. Para as análises dos dados dos dois informantes, a FFT recorta o sinal acústico em 1024 pontos.

Observe-se, na janela superior, que o segmento entre barras verticais contínuas exibe um fechamento, seguido de abertura oral. A esta se segue um novo fechamento e uma nova abertura oral, adjacente ao [i] subsequente. Tem-se, portanto, uma vibrante típica, tal qual descrito por Recasens (1991b), Almeida & Dorta, (1993), Silva (1996), Obediente (1998).

Na janela do meio, o segmento selecionado exibe um fechamento, seguido de uma estrutura contínua, na qual, ao mesmo tempo em que é possível reconhecer formantes, vê-se também a sobreposição de ruído fricativo. Este, por sua vez, se espalha por uma grande região de frequência – aproximadamente de 1 kHz a 8 kHz.

Diferentemente das duas outras janelas, a janela inferior traz uma fricativa típica: veja-se que o segmento selecionado exibe apenas ruído, que se espalha desde uma região de frequências baixas, até os 8kHz. Neste segmento, ao contrário dos dois primeiros, é impossível reconhecer uma estrutura formântica.

Uma análise FFT, realizada no ponto médio de cada um dos segmentos, confirma a presença de ruído fricativo, em dois deles – os das duas últimas janelas –, em contraposição ao primeiro segmento. A figura 3.7, abaixo, traz as análises FFT, no ponto médio da vibrante, vibrante espirantizada e fricativa, ilustradas na figura anterior:

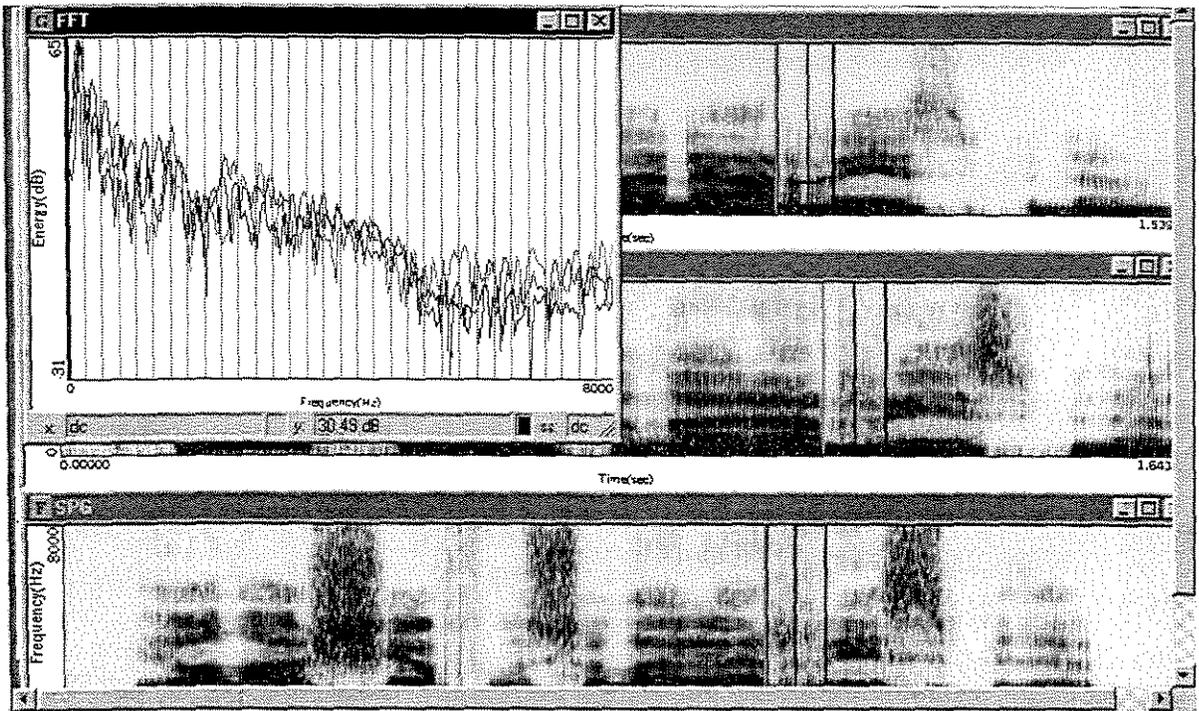


Figura 3.7 - Análise FFT, no ponto médio da vibrante, da vibrante espirantizada e da fricativa, na seqüência “funcionária rispida”. O ruído se sobrepõe gradativamente à estrutura formântica da vibrante.

Atente-se para o aumento do ruído fricativo: nulo para a vibrante, cujo espectro é traçado em preto (na janela G, à esquerda da figura), ele surge no espectro da vibrante espirantizada (traçado em cinza escuro) e se torna maior no espectro da fricativa, traçado em cinza claro. A FFT atesta o fato, através da amplitude dos picos, a qual se mantém alta em regiões de freqüência também alta. (O esperado, caso a vibrante fosse produzida sem ruído fricativo, é que seu espectro tivesse comportamento semelhante ao espectro de outros sons, como as vogais, ou seja, que a sua amplitude diminuísse concomitantemente ao aumento da freqüência.)

É mister frisar que os espectrogramas e as FFTs das figuras 3.6 e 3.7 não sinalizam a existência de uma categoria intermediária à vibrante e à fricativa. Ao contrário, apontam a existência de um contínuo físico, entre vibrantes e fricativas, na fala de JLM. Esse achado se reforça pela observação de que pode haver diferentes graus de sobreposição de ruído fricativo, nas vibrantes espirantizadas. É isto o que mostra a figura 3.8, abaixo, através dos espectrogramas de duas vibrantes espirantizadas e da FFT, realizada no ponto médio de cada uma delas, pertencentes à seqüência “menina rebelde”. A figura sobrepõe uma janela

com a análise FFT às janelas dos espectrogramas, para que se visualizem as duas vibrantes espirantizadas.

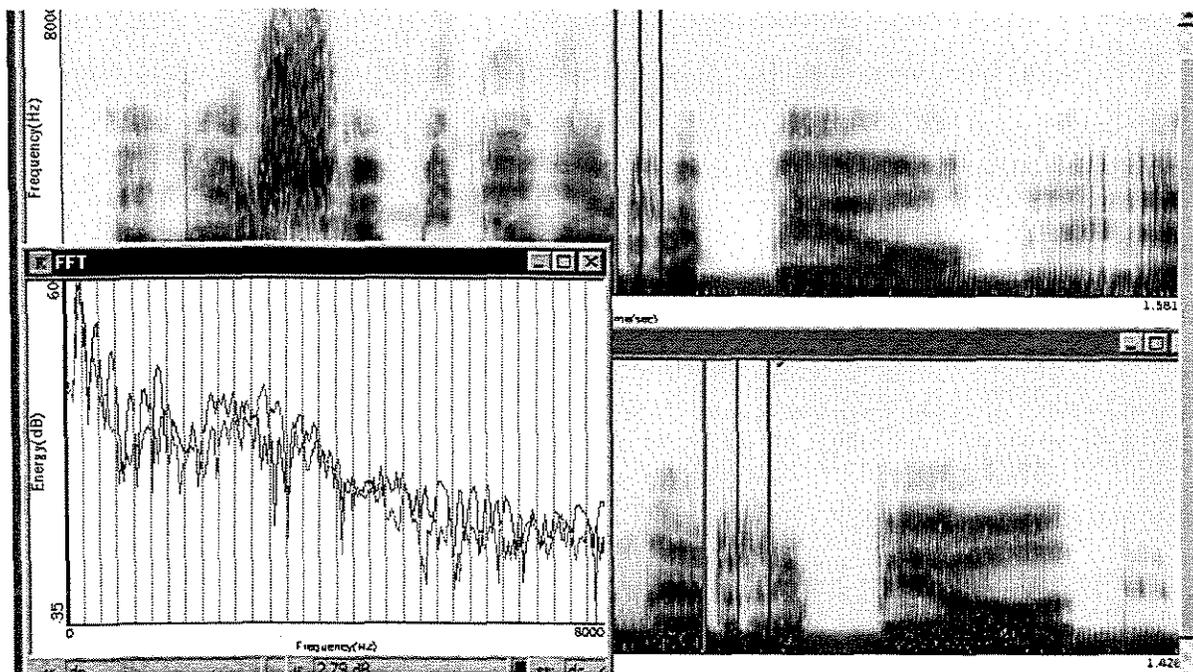


Figura 3.8 - FFT do ponto médio das duas vibrantes espirantizadas, marcadas por linhas verticais contínuas nos espectrogramas: notem-se os diferentes graus de sobreposição do ruído fricativo à estrutura da vibrante.

Observe-se que, em altas faixas de frequência – a partir de 4 kHz – o espectro FFT escuro, relativo à vibrante espirantizada da janela superior da figura, exibe amplitude maior do que a amplitude da FFT clara, relativa à vibrante espirantizada da janela inferior.

Para uma análise mais precisa do que esta, visual, que apresento acima, é necessário quantificar o ruído fricativo, presente em cada um dos segmentos espirantizados. Embora tal quantificação não tenha sido realizada para todos os róticos, produzidos por JLM, apresento na seção 3.3.1.2, adiante, as medidas do ruído fricativo, em alguns dados. Antes, é preciso lembrar que, no caso da Figura 3.8, acima, há apenas uma descrição da presença de mais ruído fricativo, numa vibrante espirantizada, do que noutra, através da comparação das amplitudes dos picos de uma e outra.

Também é necessário verificar, nos dados do informante JLM, se há condicionamento prosódico sobre a realização de /r/, tal qual hipotetizado inicialmente. É este tópico que abordo, a seguir.

3.3.1.1 – Alguns indícios de condicionamento prosódico sobre a realização do contínuo físico na variabilidade de pronúncia de /r/ para o informante JLM

Dois procedimentos foram adotados na investigação do condicionamento prosódico sobre o contínuo fonético sobre o qual se estende a variabilidade de pronúncia de /r/ inicial, na fala de JLM: a) o cálculo da duração relativa de cada segmento, inclusive de /r/, no interior de cada sentença, que constitui os pares mínimos; b) o confronto entre os valores de duração relativa de /r/ e uma análise visual, para identificar qual variante ocorria, no início de cada adjetivo, em cada sentença de cada par mínimo.

O cálculo da duração relativa dos segmentos consiste na razão entre a duração absoluta de um dado segmento de um enunciado, e.g., uma variante de /r/, e a duração absoluta do enunciado todo, e.g., “eu não achei a funcionária ríspida”. Como resultado, obtém-se o percentual de duração do segmento, visado num enunciado. Considere-se, por exemplo, o enunciado *supra* referido: na primeira repetição, para o caso em que “ríspida” assume função de adjunto adnominal, o enunciado tem duração de 1,453s. A variante de /r/, neste caso realizada como vibrante, dura 68 ms. Logo, a duração relativa da vibrante é de aproximadamente 5%.

Os valores das durações relativas das variantes de /r/, por si só, são inconclusivos, porém. Observe-se a tabela 3.1, abaixo:

Sintagma	Predicativo do Objeto				Adjunto Adnominal			
	1 ^a . rep.	2 ^a . rep.	3 ^a . rep.	4 ^a . rep.	1 ^a . rep.	2 ^a . rep.	3 ^a . rep.	4 ^a . rep.
Música rápida	6	11	12	7.5	8	7	9	7
Funcionária ríspida	7	5	5	5	5	4	5.5	6
Cômoda rústica	6	9.5	5.5	5	4	6	7	5
Técnica rígida	8	6	7	6	6	7	5.5	6
Criança risonha	3	4	5	4	4	5	4	5.5
Camisa rasgada	4	3	3	3	3	3	3	3
Novela romântica	5	4	5	4	6	5	5	5
Menina rebelde	5	3.5	5	4	5	6	6	5

Tabela 3.1 – Duração relativa (em %) dos róticos nas quatro repetições dos dados de JLM. As células sombreadas marcam os pares em que /r/ é mais longo iniciando adjetivos predicativos do objeto.

A tabela apenas aponta para uma aleatoriedade na duração de /r/, contrariamente à hipótese inicial, e segundo a qual as variantes de /r/ deveriam ser mais longas, adjacentes a fronteiras prosódicas fortes, ou seja, iniciando adjetivos com função de predicativo do objeto. Isto, de fato, ocorre em 11, dos 32 pares mínimos, no total. Esses pares estão nas células sombreadas da tabela.

Em outros 14 pares, contudo, as variantes de /r/ têm maior duração relativa, quando adjacentes à fronteira prosódica fraca, i.e., em início de adjetivos adjuntos adnominais. Em 7 pares, a duração relativa das variantes de /r/ é a mesma, independentemente da força de fronteira.

Um confronto desses valores com uma análise visual que identifica a variante de /r/, presente em cada enunciado, visa a uma maior clareza dos resultados. Isto porque a espirantização pode alongar o segmento – uma vez que fricativas são intrinsecamente mais longas que vibrantes, por exemplo –, fato que poderia estar obscurecendo os dados e impedindo que se relacionem efetivamente as versões da vibrante à força de fronteira adjacente.

Assim, a tabela abaixo, baseada numa análise visual de espectrogramas, traz as ocorrências das variantes de /r/, em todas as repetições de todos os enunciados do *corpus*:

Sintagma	Predicativo do Objeto				Adjunto Adnominal			
	1 ^a . rep.	2 ^a . rep.	3 ^a . rep.	4 ^a . rep.	1 ^a . rep.	2 ^a . rep.	3 ^a . rep.	4 ^a . rep.
Música rápida	V	VE	V	V	V	VE	V	V
Funcionária ríspida	VE	VE	VE	VE	V	V	F	F
Cômoda rústica	V	V	VE	V	V	V	V	VE
Técnica rígida	V	V	VE	VE	V	VE	VE	VE
Criança risonha	VE	VE	F	F	VE	V	VE	F
Camisa rasgada	VE	VE	VE	VE	VE	T	VE	VE
Novela romântica	V	V	V	VE	V	V	V	V
Menina rebelde	VE	VE	F	VE	VE	VE	V	VE

Tabela 3.2 - Ocorrências das variantes de /r/, nas repetições dos enunciados. V é "vibrante"; VE, "vibrante espirantizada"; F, "fricativa"; T, "tap".

Isolados, estes dados também não são esclarecedores. De fato, em apenas 5 pares – aqueles das células com sombreado mais intenso – a versão enfraquecida de /r/ ocorre seguinte a fronteira prosódica fraca: terceira e quarta repetições do enunciado contendo a seqüência “funcionária ríspida”; quarta repetição do enunciado contendo “cômoda rústica”; segunda repetição do enunciado contendo “técnica rígida” e segunda repetição do enunciado contendo “camisa rasgada”.

Comparando-se as Tabelas 3.1 e 3.2, chama a atenção o fato de que há nove pares – aqueles contidos nas células com sombreado mais claro – nos quais ocorre a mesma variante, tanto adjacente a fronteira prosódica forte como a fronteira fraca, assumindo, porém, maior duração relativa, quando adjacente à fronteira prosódica forte, ou no início de adjetivos com função de predicativo do objeto. Esse é um dado significativo, na medida que possibilita afirmar a influência da estrutura prosódica do enunciado sobre a estrutura acústica de um segmento. Conseqüentemente, é viável continuar perseguindo minha hipótese inicial, mesmo nos dados deste informante. Talvez a quantificação do ruído fricativo, que abordo em seguida, possa trazer dados mais esclarecedores a esse respeito.

3.3.1.2 – Quantificação do ruído fricativo em alguns dados de JLM

A quantificação do ruído fricativo foi realizada nos róticos dos seguintes pares mínimos prosódicos: “menina rebelde”, na terceira repetição; “camisa rasgada”, na quarta repetição; “funcionária ríspida”, na primeira e terceira repetições; “criança risonha”, na quarta repetição.

Para tanto, foram comparadas as amplitudes do /r/ de um mesmo adjetivo no início do fechamento – ou no vestígio de início de fechamento – desse segmento, através do espectro FFT. O número de pontos de amostragem do sinal acústico, para o cálculo da FFT, foi fixado em 1024 pontos, pois um maior número de pontos pode captar mudanças mais abruptas do sinal acústico, como parece ser o caso do início da espirantização de /r/, quando o ruído fricativo se sobrepõe a um evento acústico típico de vibrante, i.e., o fechamento. A tabela abaixo traz os valores da somatória do quadrado das amplitudes, para cada rótico, em cada par:

	Predicativo do objeto	Adjunto adnominal
Menina rebelde (3 ^a . repetição)	46785	28818
Camisa rasgada (4 ^a . repetição)	33653	46416
Funcionária ríspida (1 ^a . repetição)	29832	18483
Funcionária ríspida (3 ^a . repetição)	5302	10327
Criança risonha (4 ^a . repetição)	40622	17468

Tabela 3.3 - Somatória dos quadrados das amplitudes de /r/, em cada ocorrência de cinco pares mínimos prosódicos, produzidos por JLM.

Os valores da amplitude, para as duas ocorrências de /r/, em cada par selecionado, foram, então, submetidos a análise estatística, tendo sido considerada a região de frequência que, numa análise visual, apresentava maior diferença de amplitude. A Figura 3.9, abaixo, ilustra a região de frequência, tomada para a quantificação do ruído fricativo, no rótico de “funcionária ríspida”, na primeira repetição.

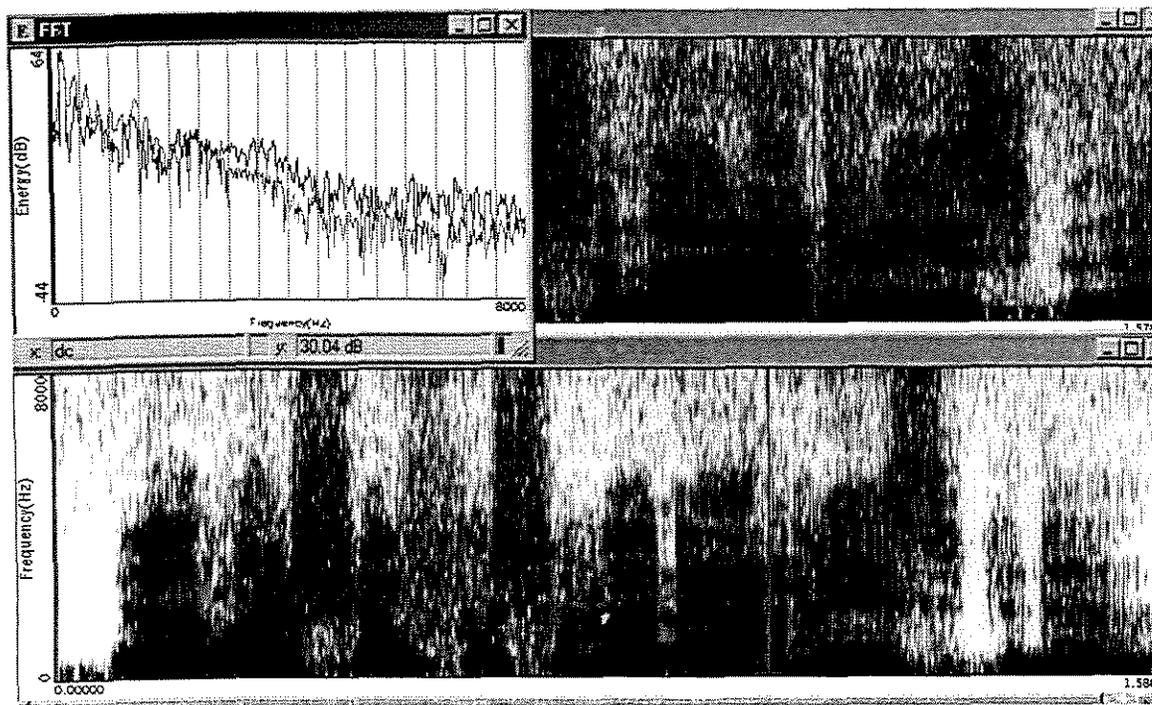


Figura 3.9 - Comparação dos espectros FFT, no par "funcionária ríspida" (1^a. repetição), para fins de quantificação do ruído fricativo dos róticos.

O teste estatístico realizado, o T de Student, visava a confirmar a existência de diferença de amplitude entre duas realizações de /r/. Ressalte-se que assumo que maior amplitude do sinal se relaciona diretamente com maior presença de ruído fricativo num segmento; inversamente, menor amplitude do sinal acústico aponta para pouca presença de ruído no segmento, ou seja, para presença de maior vibração nele. Obtiveram-se resultados significativos para todas as ocorrências de /r/, como se vê na Tabela 3.4, abaixo:

	“menina rebelde” (3 ^a . rep.)	“camisa rasgada” (4 ^a . rep.)	“funcionária ríspida” (1 ^a .rep.)	“funcionária ríspida” (3 ^a . rep.)	“criança risonha” (4 ^a . rep.)
P	0.000527	0.00036	2.78E-09	0.000145	1.04E-12
T crítico	1.969311	1.978819	1.975695	1.968442	1.975591
T	3.511457	3.666696	6.318427	3.851718	7.774111
	S	S	S	S	S

Tabela 3.4 - Resultados do teste-T para duas amostras em par para médias, tomando-se /r/ predicativo do objeto e /r/ adjunto adnominal, em cada sintagma observado.

Esses resultados mostram, portanto, que um dos róticos, em cada par tomado, é mais espirantizado do que o outro. Contudo, não apontam para um possível condicionamento da estrutura prosódica do enunciado sobre a realização de /r/, contrariando a hipótese expressa em 3.3.1.1 e segundo a qual a quantificação do ruído fricativo poderia fornecer pistas mais incisivas sobre a natureza acústica do segmento relativamente à força de fronteira adjacente a ele. Como se vê pela Tabela 3.3, *supra*, não é possível dizer que variantes mais espirantizadas do rótico ocorrem adjacentes a fronteira prosódica fraca.

É interessante notar, por outro lado, que os falantes percebem a espirantização: um experimento perceptual realizado em caráter preliminar, permite verificar que os sujeitos, no geral, identificam como mais espirantizado o segmento que, de fato, apresenta valores maiores para suas amplitudes. Observe-se o quadro 3.2, abaixo:

		<u>Menina</u> <u>rebelde</u> (3 ^a . rep.)	<u>Camisa</u> <u>rasgada</u> (4 ^a . rep.)	<u>Funcionária</u> <u>ríspida</u> (1 ^a . rep.)	<u>Funcionária</u> <u>ríspida</u> (3 ^a . rep.)	<u>Criança</u> <u>risonha</u> (4 ^a . rep.)
R	PO	*	*	*		*
	AA				*	
L	PO	*		*	*	*
	AA		*			
P	PO	*	*	*		*
	AA				*	
A	PO	*				
	AA		*	*	*	*

Quadro 3.2 - Resultados do teste perceptual: * indica /r/ mais espirantizado, conforme o julgamento de quatro informantes

O experimento perceptual consistiu em submeter a julgamento de quatro informantes – alunos de graduação ou pós-graduação do Instituto de Estudos da Linguagem da UNICAMP – os mesmos pares que foram tomados, para a quantificação do ruído fricativo de /r/. Eram apresentados aos sujeitos os dois róticos de um mesmo par mínimo prosódico. Só então se passava aos róticos de um outro par. A ordem de apresentação dos pares foi a mesma para os quatro informantes. Ao ouvirem os dois róticos de um mesmo par, os sujeitos eram solicitados apenas a dizer qual /r/ apresentava mais ruído fricativo.

Pelo Quadro 3.2, vê-se que os informantes foram unânimes em reconhecer que, no par “menina rebelde”, da terceira repetição, o /r/ iniciando o adjetivo com função de predicativo do objeto é o mais espirantizado. De fato, o julgamento coincide com a quantificação do ruído: a soma do quadrado das amplitudes é maior para esse /r/ do que para o /r/, com função de adjunto adnominal.

Embora o julgamento não seja unânime para os pares “funcionária ríspida” (primeira e terceira repetições) e “criança risonha” (quarta repetição), nota-se que três dos quatro informantes fazem o mesmo julgamento, acerca da espirantização do rótico. E, novamente, como no caso de “menina rebelde”, o julgamento coincide com a quantificação do ruído: /r/ que ocorre no adjetivo predicativo do objeto é mais espirantizado em “funcionária ríspida” (primeira repetição) e “criança risonha” (4^a. repetição), como se vê pelo valor da soma dos quadrados das amplitudes na Tabela 3.3. No sintagma “funcionária ríspida” (terceira repetição), por outro lado, o /r/ que ocorre no adjetivo adjunto adnominal é julgado mais espirantizado, como de fato é, segundo a Tabela 3.3. Apenas para o par

“camisa rasgada” (quarta repetição), o julgamento dos informantes torna-se aleatório, apesar de também aí haver diferença de amplitude, conforme se nota, na Tabela 3.3.

O experimento perceptual, portanto, sugere que os falantes percebem a espirantização dos róticos. Não há, porém, qualquer pista de que os falantes percebem a gradiência do processo. Por isso, talvez seja necessário estender o experimento perceptual a um maior número de sujeitos.

Outro fato que precisa ser melhor investigado é a razão do julgamento aleatório para os róticos do par “camisa rasgada” (quarta repetição), uma vez que a quantificação do ruído fricativo aponta para a diferença de amplitude entre os dois róticos, de modo que o /r/ que inicia o adjetivo adjunto adnominal tem maior somatória dos quadrados das amplitudes (cf. Tabela 3.3, *supra*) do que /r/ que inicia o adjetivo predicativo do objeto. Talvez esse par envolva uma espirantização mais intensa que os demais e, por isso, talvez seja preciso desenvolver um algoritmo mais poderoso – inclusive com o auxílio de um especialista em modelamento não-acústico – para quantificar o ruído fricativo, em casos como esse.

3.3.2 – Contínuo físico na variabilidade de pronúncia dos róticos: análises visuais dos dados de AJS

O segundo informante do experimento acústico, AJS, aparentemente alterna vibrantes e *taps* em sua fala. É isso o que sugeriria uma análise de outiva e que permitiria tomá-lo como mais um exemplo de falantes da região Sul do país que produzem *taps* em posição inicial de palavra (cf. descrições de Monaretto, 1997, por exemplo).

Contudo, uma análise mais cuidada, baseada nos espectrogramas dos dados, revela que AJS produz não só *taps*, em início de palavra: ele produz também uma variante intermediária a *tap* e vibrante, que denomino doravante “vibrante encurtada”. Cabe mencionar que a constatação da existência dessa “vibrante encurtada” surgiu a partir da observação de que /r/ em palavras como “rípida” não parecia uma vibrante – perceptualmente falando – mas tampouco era um *tap* de fato. Essa impressão se fortalece, quando se compara /r/, em início de palavra, com um *tap* prototípico, como o que ocorre em posição intervocálica na palavra “funcionária”.

Os espectrogramas dos dados, bem como as respectivas formas de onda, confirmam a existência do contínuo físico, na fala de AJS, não só na seqüência-alvo “funcionária

ríspida”, mas também em outras seqüências do *corpus*. A figura 3.10a, abaixo, traz os espectrogramas de três dados que permitem visualizar o contínuo: na janela da porção superior da figura, tem-se a seqüência-alvo “música rápida”. Nas outras duas janelas da figura, ocorre a seqüência “novela romântica”. Para uma observação mais detalhada do contínuo, as formas de onda dos espectrogramas são ilustradas pela figura 3.10b.

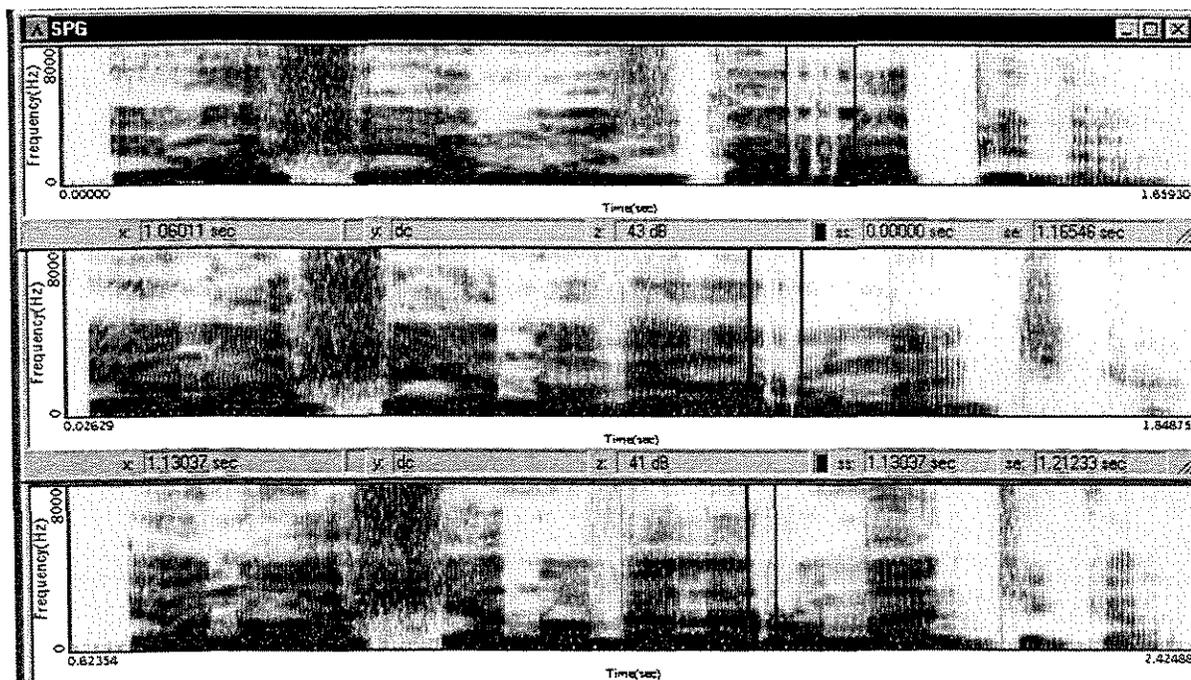


Figura 3.10a - Entre barras verticais contínuas, nos espectrogramas: vibrante (janela superior), "vibrante encurtada" (janela do meio) e tap (janela inferior).

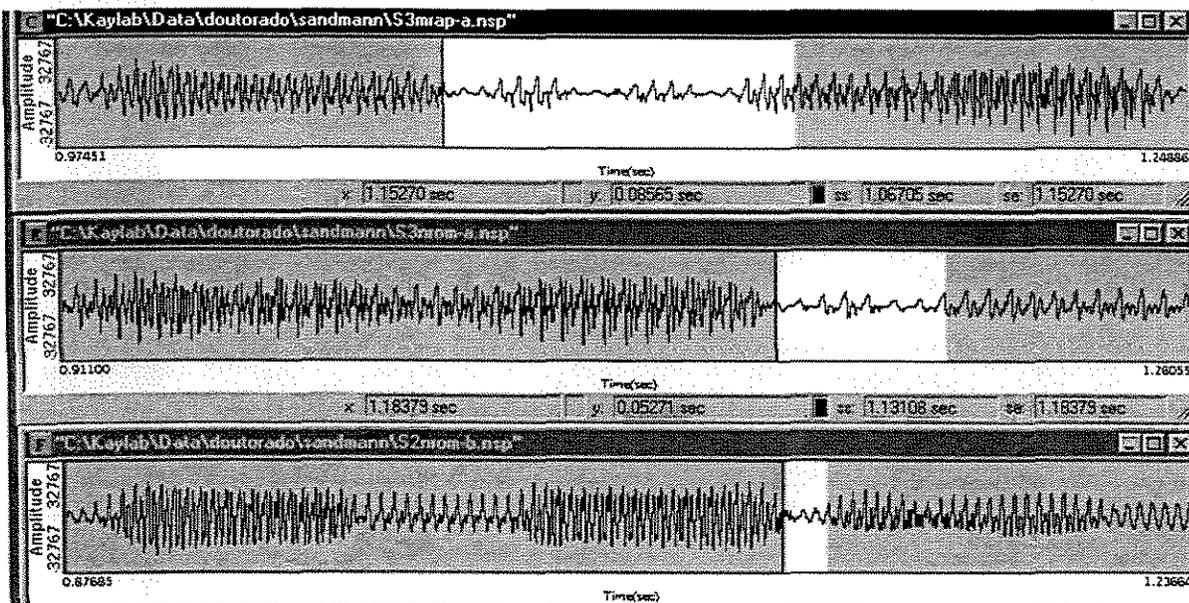


Figura 3.10b – Nas porções não sombreadas da figura, formas de onda da vibrante (janela superior), da "vibrante encurtada" (janela do meio) e *tap* (janela inferior).

Na janela superior da Figura 3.10a, tem-se o espectrograma de uma vibrante canônica: é possível reconhecer nele um fechamento, seguido de uma abertura oral, um novo fechamento e uma nova abertura oral e, finalmente, um terceiro fechamento e uma terceira abertura oral, já adjacente à vogal.

Como menciono no item 3.3.1, relativamente à vibrante, presente nos dados do informante JLM, essa estrutura acústica típica de uma vibrante, é descrita por autores como Recasens (1991b), Almeida & Dorta (1993) e Obediente (1998) – para o espanhol – e Silva (1996) para PB¹³⁷. O fechamento corresponde a um momento de diminuição da pressão supraglótica em que a voz quase desaparece. Podem-se identificar os fechamentos no espectrograma da porção superior da Figura 3.10a como espaços quase que em branco. Na forma de onda correspondente (janela superior da Figura 3.10b), esses momentos têm uma amplitude quase nula. Aos fechamentos se intercalam as aberturas orais, identificáveis, no espectrograma da janela superior da Figura 3.10a, como breves momentos nos quais o

¹³⁷ É digna de nota também a descrição dos róticos, apresentada por Ladefoged & Maddieson (1996), no sétimo capítulo de seu *The sound of the world's languages*. A descrição é bastante extensa e conta com exemplos de róticos, em várias línguas. Peca, no entanto, por considerar que, em línguas como o italiano e o espanhol, em que há contraste entre vibrante e *tap*, aquela seja uma versão geminada dessa. O motivo desta minha crítica deve ter ficado claro, no segundo capítulo, item 2.5.3.1.

vozeamento é retomado e nos quais se pode, inclusive, reconhecer uma estrutura formântica bem definida. Na forma de onda (janela superior da Figura 3.10b), esses momentos exibem pulsos regulares, semelhantes aos de uma vogal, e com amplitude semelhante à de uma vogal. Reconhece-se a terceira abertura oral dessa vibrante, adjacente à vogal /a/, especialmente pela trajetória dos formantes, similar à trajetória dos formantes, nas duas outras aberturas orais.

A “vibrante encurtada”, na janela do meio da Figura 3.10a, difere da vibrante, por apresentar uma única abertura oral, às vezes seguida de novo fechamento, como é o caso ilustrado na figura, às vezes adjacente à vogal seguinte. Novamente, os dois fechamentos que constituem o segmento na figura podem ser reconhecidos como espaços quase que em branco, no espectrograma. A sua forma de onda – como se vê na janela do meio na Figura 3.10b – apresenta amplitude muito baixa. A abertura oral, como na vibrante, é identificada no espectrograma como um breve evento acústico no qual se nota uma estrutura formântica bem definida. A forma de onda da abertura oral é constituída de pulsos regulares com amplitude razoável. Fato que chama a atenção na *vibrante encurtada* – novamente friso – é que perceptualmente se identifica algo que não é uma vibrante, mas tampouco um *tap* propriamente. A sua comparação auditiva com *taps* e vibrantes canônicos reforça essa impressão.

A janela inferior da Figura 3.10a, por fim, traz o espectrograma de um *tap*: vê-se aí um fechamento, seguido de uma barra vertical muito tênue e semelhante ao estouro das consoantes oclusivas. Na Figura 3.10b, o *tap*, destacado pela ausência de sombreamento, é reconhecido pela forma de onda com amplitude baixa. O estouro, por sua vez, não é visível. Lembre-se, inclusive, que nem sempre esse evento acústico é visível em espectrogramas.

A inspeção visual permite ainda notar uma diferença duracional entre os três segmentos, de modo que vibrante e *tap* são os extremos, numa escala decrescente de duração, e a vibrante encurtada se mantém intermediária. Volto a este aspecto, com mais detalhes, adiante.

Novamente friso – como já fizera no item 3.3.1, ao apresentar os dados de JLM – que este estudo não evidencia a existência de uma outra categoria, intermediária a vibrante

e *tap*, mas chama a atenção para o fato de emergir, na fala de AJS, um contínuo físico que tem vibrante e *tap* como pontos extremos.

3.3.2.1 – Índícios de condicionamento prosódico sobre a realização do contínuo físico na variabilidade de pronúncia de /r/: dados de AJS

Ainda perseguindo a hipótese de que há condicionamento prosódico, na realização do contínuo físico sobre o qual se estende a variabilidade de pronúncia dos róticos iniciais, calculei a duração relativa de /r/ em cada um dos 64 enunciados, produzidos por AJS. A tabela 3.5, abaixo, traz os resultados obtidos.

Sintagma	Predicativo do Objeto				Adjunto Adnominal			
	1ª. rep.	2ª. rep.	3ª. rep.	4ª. rep.	1ª. rep.	2ª. rep.	3ª. rep.	4ª. rep.
Música rápida	6	5	5	6	5	5	4	5
Funcionária ríspida	3	3	2.5	1	2.5	3	4	0.7
Cômoda rústica	4	3	3	3.5	4	3	3	3
Técnica rígida	3	4	3	3	3	4	2	5
Criança risonha	3	2	3	1	2	1	3	1
Camisa rasgada	3	3	3	3	2	2.5	1	1.5
Novela romântica	3	1	5.5	6	3	1	4	3
Menina rebelde	2	1	1	2	1	1	1	1

Tabela 3.5 Durações relativas (em %) de vibrantes, vibrantes encurtadas e *taps*, nas quatro repetições dos dados de AJS.

Diferentemente dos resultados do cálculo da duração relativa de /r/, obtidos para o informante JLM, os resultados obtidos para AJS são fortes indícios de que há realmente, de acordo com a hipótese inicial, influência da força de fronteira sobre a natureza dos segmentos: note-se que, de um total de 32 pares, em 17 deles – os que estão nas células sombreadas da tabela *supra* – a duração relativa de /r/ é maior nos casos em que esse segmento ocorre adjacente a fronteira prosódica forte, i.e., iniciando adjetivo predicativo do

objeto. Em apenas dois pares,¹³⁸ a duração relativa de /r/ é maior nos casos em que ele segue fronteira prosódica fraca, ou seja, em que inicia adjetivo adjunto adnominal. Nos outros treze pares do *corpus*, /r/ tem a mesma duração relativa, independentemente de seguir fronteira prosódica forte ou fraca.

Os dados sinalizam ainda que a duração relativa de /r/ está intrinsecamente relacionada ao enfraquecimento ou ao fortalecimento dos segmentos na fala de AJS, pois, como já mencionado, a duração decresce, à medida que a vibrante percorre o contínuo rumo ao *tap*. Tal relacionamento se torna mais visível quando verificamos a ocorrência das variantes, em cada repetição dos dados, como expresso na Tabela 3.6, abaixo.

Sintagma	Predicativo do Objeto				Adjunto Adnominal			
	1 ^a . rep.	2 ^a . rep.	3 ^a . rep.	4 ^a . rep.	1 ^a . rep.	2 ^a . rep.	3 ^a . rep.	4 ^a . rep.
Música rápida	V	V	V	V	V	V	V	V
Funcionária ríspida	VEN							
Cômoda rústica	VEN							
Técnica rígida	VEN	VEN	VEN	VEN	V	T	VEN	VEN
Criança risonha	T	T	VEN	T	T	T	VEN	T
Camisa rasgada	VEN	VEN	VEN	VEN	VEN	VEN	T	T
Novela romântica	VEN	T	VEN	VEN	V	T	VEN	VEN
Menina rebelde	T	T	T	T	T	T	T	T

Tabela 3.6 - Ocorrências de vibrantes (V), vibrantes encurtadas (VEN) e *taps* (T), nas quatro repetições dos dados de AJS.

Cruzando-se os dados desta tabela com os da Tabela 3.5, vê-se que, nas quatro repetições da seqüência “música rápida”, em que é produzida uma versão fortalecida de /r/, i.e., uma vibrante, a duração relativa do segmento é a maior de toda a tabela. Coincidentemente, nas demais repetições das outras seqüências, têm-se versões enfraquecidas de /r/, ou seja, vibrantes encurtadas ou *taps*. Observe-se, inclusive, que os

¹³⁸ A terceira repetição da seqüência-alvo “funcionária ríspida” e a quarta repetição da seqüência “técnica rígida”.

menores valores de durações relativas da Tabela 3.5 coincidem majoritariamente com as ocorrências de *taps*, na Tabela 3.6.

É o que acontece com /r/, na seqüência-alvo “criança risonha”, na primeira, segunda e quarta repetições, tanto para o caso em que o rótico sucede fronteira forte (quatro primeiras colunas da tabela) como para o caso em que sucede fronteira prosódica fraca (quatro últimas colunas da tabela). Este mesmo fato pode ser verificado ainda: na terceira e quarta repetições da seqüência “camisa rasgada”, para o caso em que /r/ sucede fronteira prosódica fraca; na segunda repetição de “novela romântica”, em ambiente seguinte a fronteira prosódica forte e a fronteira prosódica fraca; em todas as repetições e todos os ambientes prosódicos da seqüência “menina rebelde”¹³⁹.

As vibrantes encurtadas, por sua vez, exibem duração relativa intermediária à de vibrantes e *taps*, ora coincidindo com a duração relativa de *taps*, ora com a de vibrantes. Isso reforça seu caráter intermediário com relação aos outros dois segmentos e reafirma a estreita relação entre a duração e a natureza do rótico, – ou seja, nos casos em que tende a vibrante, /r/ apresenta maior duração relativa do que nos casos em que tende a vibrante encurtada ou simplesmente *tap*.

Pode-se dizer, portanto, que os dados de AJS são mais consistentes do que os de JLM, no que diz respeito à sinalização de efeito prosódico, condicionando a natureza dos segmentos, e à relação entre natureza dos segmentos e medidas de duração relativa.

3.3.3 – Considerações finais sobre os dados

Comparando-se os resultados das análises dos dados de AJS e JLM, pode-se dizer que todos têm em comum o fato de evidenciarem o enfraquecimento de /r/. Acontece,

¹³⁹ Sobre este dado, cabe ainda uma observação: em todas as repetições, a seqüência “menina rebelde” parece sofrer um processo de enfraquecimento geral. Isto porque “menina” é produzida, invariavelmente, como algo similar a [mĩ.nɐ], ou seja, o gesto nasal da sílaba tônica “escorrega” sobre a vogal [i] seguinte a ele, “escondendo-a”, e a nasalização se estende sobre [i] pretônico.

porém, que esse fenômeno envolve estratégias de produção distintas em cada caso, o que resulta, em última instância, em efeitos acústicos também distintos¹⁴⁰.

Assim, os dados de JLM exibem enfraquecimento de /r/, de modo que vibrantes apicais podem ser realizadas como fricativas, havendo entre essas duas categorias um contínuo físico, gradiente, sobre o qual são produzidas as vibrantes espirantizadas, as quais ora tendem a vibrantes, apresentando – ao menos visualmente – menor sobreposição de ruído fricativo, ora tendem a fricativas e apresentam maior sobreposição de ruído fricativo.

Está envolvida neste processo de enfraquecimento uma manobra articulatória que consiste na perda do gesto balístico de ponta de língua: deixa de ocorrer a batida da língua, na região alveolar, e forma-se apenas um estreito canal, por onde se propaga o ar¹⁴¹. Esse canal é estreito o suficiente para que haja fricção das partículas de ar que passam ao longo dele e variações no tamanho desse canal talvez provoquem as diferenças de ruído fricativo, sobreposto à estrutura da vibrante. Pode ocorrer ainda uma primeira batida da ponta da língua nos alvéolos, seguida não de um novo toque da ponta da língua nos alvéolos, mas de uma aproximação tal entre os dois articuladores, que se faz a constrição típica de fricativa. Isso explicaria a existência de casos em que se vê, no espectrograma, um primeiro momento característico de vibrante e, em seguida, um momento semelhante à de uma fricativa. Esses casos ocorrem não apenas nos dados de JLM, mas também nos dados de OGF e ARM, comentados no item 3.1.

Acusticamente, como já se disse, tem-se, como resultado da perda do movimento balístico de ponta de língua, concomitante à formação de constrição no trato, o aumento do

¹⁴⁰ Ressalte-se que não afirmo que haja sempre uma correspondência biunívoca entre uma manobra articulatória e uma manifestação acústica. Nos dados dos dois informantes deste estudo, coincidiu haver tal correspondência.

¹⁴¹ Sobre essa alteração na estratégia de produção da vibrante, que a torna um segmento “não-vibrante”, é elucidativa a descrição de Ladefoged & Maddieson (1996: 217): *a característica primária de uma vibrante é a vibração de um órgão de fonação contra outro, comandada por condições aerodinâmicas. Uma das partes moles móveis do trato vocal é posicionada suficientemente próxima de outra superfície, de modo que, quando uma corrente de ar de tamanho certo passa através da abertura criada por essa configuração, ocorre um padrão repetitivo de fechamento e abertura do canal da corrente de ar. Esse movimento(...) é muito similar à vibração das pregas vocais, mas uma abertura suficientemente estreita deve ser criada e uma corrente de ar adequada deve passar pela abertura. O tamanho da abertura e a corrente de ar devem cair em limites críticos para que a vibração ocorra, sendo que pequenos desvios acarretam a falha da vibração. Como resultado, vibrantes tendem a variar com pronúncias “não-vibrantes”.* (Tradução minha.)

ruído fricativo sobre o espectro da vibrante. (Curioso é notar que, nos dados de JLM, predominam as vibrantes espirantizadas, o que o poderia caracterizar como um daqueles sujeitos que produziam o “rz” de Gonçalves Viana.)

Variações no tamanho da constrição talvez sejam as responsáveis pela variação na sobreposição do ruído fricativo sobre a estrutura da vibrante. (Isto, porém, só ficaria evidente com um estudo articulatorio que medisse o tamanho da constrição, acompanhado pela quantificação do ruído fricativo.)

Ainda uma última observação: o fato de fricativas apresentarem duração intrínseca variável, mas maior que a duração intrínseca de vibrantes, pode ser uma das razões pelas quais as Tabelas 3.1 e 3.2 *supra*, relativas aos dados de JLM, não exibiam uma relação consistente entre força de fronteira e duração relativa de /r/. Ou seja, o caráter variável da duração da fricativa acaba por enviesar os dados, no sentido da confirmação de minha hipótese.

Os dados de AJS, por sua vez, exibem enfraquecimento de /r/, no sentido de perda de vibração: ao invés de o toque da língua nos alvéolos deixar de ser realizado – dando lugar à formação de uma estreita constrição que provoca a fricção das partículas de ar, no interior dessa constrição –, ele tende a ocorrer menos vezes – uma única, no caso do *tap*. No caso da vibrante encurtada, parece ocorrer um fechamento e uma abertura oral e a repetição de parte dessa estrutura, i.e., do fechamento, como ilustra a Figura 3.10a *supra*. Poder-se-ia argumentar que esse evento consiste, na verdade, em uma vibrante. Porém o confronto entre esse segmento, que chamo vibrante encurtada, e a vibrante *strictu sensu* evidencia diferença auditiva entre eles (vide item 3.3.2). Tampouco se trata de um *tap*, pois a comparação auditiva entre *taps* e vibrantes encurtadas evidencia diferença entre eles (também cf. item 3.3.2).

O enfraquecimento, nos dados de AJS, então, envolve o controle articulatorio necessário para a produção da vibrante: variações nesse controle fazem que haja menor quantidade de batidas da língua nos alvéolos, até acontecer apenas uma, na realização do *tap*.

Note-se também, ainda comparando os dados de JLM com os de AJS, que, enquanto naqueles dados o dorso da língua tende a se posteriorizar – concomitante à formação da

construção –, nesses dados o movimento dorsal parece permanecer inalterado. Daí ocorrer, nos dados de JLM, variação não só de modo, mas também de ponto de articulação.

Apesar de o enfraquecimento ocorrer, em direções distintas, para cada informante, há indícios de que ele seja prosodicamente condicionado: os dados de AJS são particularmente reveladores nesse sentido, já que exibem forte relação entre a duração relativa das variantes de /r/ e a força de fronteira adjacente a essa variante, como mencionado no item anterior. Os dados de JLM, embora não sejam tão consistentes, fornecem alguns indícios desse condicionamento.

O condicionamento prosódico sobre a alofonia contínua de /r/, i.e., sobre o contínuo físico, gradiente, no qual se estende a variabilidade de pronúncia de /r/, suscita uma questão-chave para a discussão acerca da fronteira entre Fonética e Fonologia: a alofonia é condicionada pela estrutura prosódica da língua – como mostram, especialmente, os dados de AJS, no item 3.3.2 – a qual, por sua vez, carrega diferentes mapeamentos da estrutura sintática dos enunciados tomados. Por isso, é necessário representar a alofonia dos róticos iniciais na gramática da língua. Entretanto, como mostrei no segundo capítulo, os modelos fonológicos que tomam primitivos de análise estáticos não dão conta de representar as alofonias contínuas. A alternativa que se coloca, então, é tratar desse fato à luz de modelos dinâmicos de produção da fala. É disso que trata o próximo capítulo.

4 – PROPOSTA DE REPRESENTAÇÃO DOS RÓTICOS À LUZ DE UM MODELO DINÂMICO DE PRODUÇÃO DA FALA

A proposta de representação para os róticos do PB, que apresento aqui, baseia-se na FAR, uma vez que considera uma organização gestual dos segmentos em questão. Uma das conseqüências de se filiar a um modelo dinâmico, para tentar representar esses segmentos – que, como menciono no segundo capítulo – até hoje as teorias fonológicas que tomam unidades estáticas como primitivos de análise não conseguem representar satisfatoriamente – é tornar possível uma “tradução” direta do dinâmico no simbólico, i.e., prever que toda a gradiência pela qual se estende a alofonia contínua dos róticos possa ser representada na gramática do PB.

Embora tome por base a FAR, minha proposta incorpora sugestões da Fonologia Acústico-Articulatória (FAAR) de Albano (2001), também um modelo dinâmico de produção da fala e que traz contribuições importantes à FAR, as quais abordo em 4.2.3. Minha proposta, ainda, toma algumas conclusões do estudo de Sproat & Fujimura (1993) sobre a alofonia contínua da lateral em inglês, para o modelamento gestual dos róticos iniciais em PB.

Então, antes de expor tal proposta, retomo alguns pontos, abordados nos capítulos anteriores – inclusive relativos às propostas de Albano (*op. cit.*) e Sproat & Fujimura (*op. cit.*) – e que são fundamentais para a representação que apresento, no final deste capítulo.

No terceiro capítulo, os dados acústicos evidenciam a alofonia contínua de /r/ inicial, tanto na fala de AJS como na de JLM. As Figuras 3.6 e 3.10a tornam clara a gradiência, presente na variabilidade de pronúncia dos róticos para os dois informantes, e sinalizam o enfraquecimento de /r/¹⁴². Esse enfraquecimento – como comento naquele capítulo– segue direções distintas para cada informante: nos dados de AJS, o gesto de ponta de língua (coronal) é preservado, havendo redução do número de batidas da língua nos alvéolos. Parece tratar-se, portanto, de uma diferença no parâmetro rigidez, especificado na equação dinâmica (vide primeiro capítulo) que define os

¹⁴² Ressalte-se que, como observo no capítulo anterior, considero a vibrante a versão fortalecida dos róticos na fala dos dois informantes e todas as demais variantes, versões enfraquecidas, não no sentido de tomar a vibrante como versão geminada do *tap*, como fazem fonólogos, a exemplo de Harris (1983), – cuja análise critico, no segundo capítulo, e volto a questionar mais adiante –, mas no sentido de Fougeron (1999), i.e., a vibrante requereria, para sua

gestos de cada variante que se realiza ao longo do contínuo físico observado¹⁴³. Como resultado, verificam-se na fala de AJS vibrantes – a versão fortalecida de /r/ - vibrantes encurtadas¹⁴⁴ e *taps* – a versão enfraquecida de /r/.

Nos dados de JLM, em contrapartida, tem-se diferença na magnitude do gesto de ponta de língua: ele se reduz, concomitantemente ao aumento da magnitude do gesto de dorso de língua, o que causa a constrição, responsável pela adição do ruído fricativo à estrutura da vibrante, donde resultam as “vibrantes espirantizadas”¹⁴⁵ as quais se alternam com vibrantes e fricativas, na fala de JLM. Uma pauta gestual à *la FAR* representaria tal diferença como na Figura 4.1:

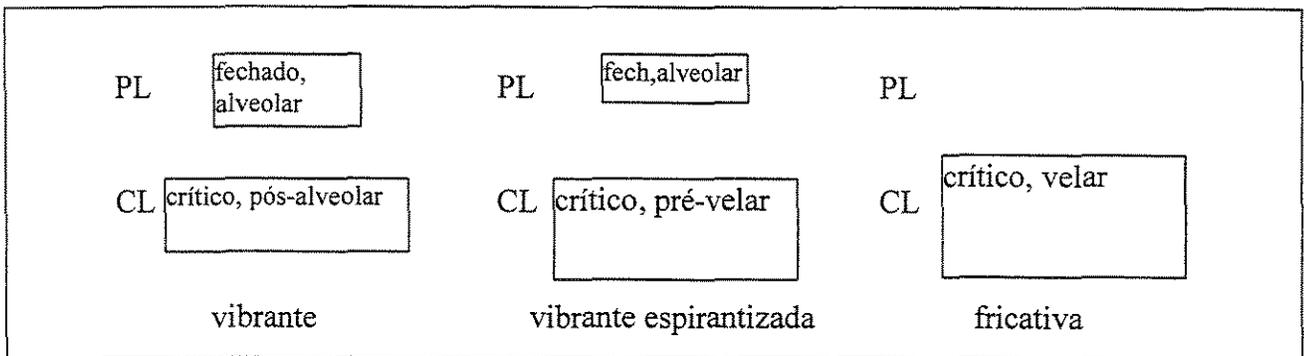


Figura 4.1 - Pautas gestuais para vibrante, vibrante espirantizada e fricativa, seguindo o arcabouço da FAR.

Essa pauta prevê que: 1) o gesto de ponta de língua tem menor duração que o de corpo da língua, o que lhe confere um caráter mais balístico; 2) à medida que a magnitude do gesto de ponta de língua diminui, aumenta a do gesto de corpo de língua, até um momento em que a magnitude do gesto de ponta de língua é nula e a magnitude do gesto de corpo de língua atinge seu máximo, o que caracteriza uma fricativa. Observe-se também que essas alterações na magnitude gestuais são acompanhadas por uma mudança no descritor do gesto de corpo de

produção, maior força articulatória, maior energia, maior tensão, ou articulação mais precisa que os demais segmentos.

¹⁴³ Como observo em 4.3, a propósito da pauta, esboçada na Figura 4.5, a diferença entre *taps* e vibrantes se dá pelo número de graus de amortecimento, maior para o primeiro segmento que para o segundo. Esses graus de amortecimento, cf. sugestão de Albano (*op. cit.*) são representados, na pauta gestual, através de uma senóide, que pode ter ciclo completo ou não.

¹⁴⁴ Na descrição dos dados da Figura 3.10a, do item 3.3.2, considero uma vibrante encurtada o segmento que apresenta um fechamento e uma única abertura oral, eventualmente seguida de novo fechamento.

¹⁴⁵ Este ponto também será abordado, mais detalhadamente, em 4.3, adiante.

língua: ele se posterioriza concomitantemente ao aumento da magnitude do gesto de corpo de língua¹⁴⁶.

Essa pauta, porém, prevê que os róticos devam ser caracterizados por dois gestos, algo que não é propriamente previsto na FAR: Browman & Goldstein defendem que cada segmento seja representado por um gesto. Nesse sentido, é preciso considerar que a FAR prevê muito bem a representação de vogais e obstruintes, mas as soantes trazem problema para o modelo. O tratamento dos róticos como segmentos constituídos de dois gestos segue sugestão de Sproat & Fujimura (1993), à qual retorno adiante, no item 4.2.1.

Outro fato para o qual apontam os dados de AJS e JLM é um possível condicionamento prosódico sobre a emergência do contínuo físico, ao longo do qual se dá a variabilidade de pronúncia de /r/ para os dois informantes. Isso é especialmente claro para AJS, cujos dados apresentam maior duração relativa de /r/, diante de fronteira forte – i.e., nos casos em que o adjetivo de uma seqüência-alvo tem a função de predicativo do objeto – do que diante de fronteira fraca – nos casos em que o adjetivo assume função de adjunto adnominal (vide Tabela 3.5). Há também evidência, nos dados de AJS, de que a duração relativa de /r/ se relaciona ao enfraquecimento ou ao fortalecimento dos segmentos, pois, como menciono no terceiro capítulo, à luz da Tabela 3.6, a duração de /r/ decresce, à medida que a vibrante percorre o contínuo, rumo ao *tap*.

Os dados de JLM não são tão consistentes, mas sugerem, igualmente, a influência da estrutura prosódica dos enunciados sobre a realização dos segmentos. Em 3.3.1.1 observo que, embora /r/ apresente maior duração relativa, quando adjacente à fronteira forte em apenas 11 dos 32 pares, há nove pares nos quais ocorre a mesma variante, tanto adjacente à fronteira prosódica forte como à fronteira fraca. Interessante é que, em todos esses pares, a mesma variante assume maior duração quando adjacente à fronteira forte. Como eu mencionava no referido item, este é um dado significativo, posto que nos pares mencionados se revela, nitidamente, a influência da estrutura prosódica sobre a natureza dos segmentos¹⁴⁷.

¹⁴⁶ O descritor “pré-velar” é empregado, nessa pauta, com o objetivo de captar a característica articulatória das vibrantes espirantizadas, característica essa que é intermediária a vibrante e fricativa.

¹⁴⁷ Considere-se que o fato de, em 14 pares, as variantes de /r/ apresentarem maior duração relativa, quando adjacentes à fronteira prosódica fraca (cf. Tabela 3.1) não significa necessariamente que esteja equivocada a hipótese

O condicionamento prosódico sobre a variabilidade da pronúncia de /r/ soma-se aos vários exemplos, já disponíveis na literatura, acerca do caráter contínuo de alofonias que modelos fonológicos tradicionais tratam como categóricas (vide, e.g., Albano *et alii*, 1998; Albano, 2001, para o PB; Browman & Goldstein, 1992; Sproat & Fujimura, 1993; Huffman, 1997; Gick, 1999, para o inglês). Ou seja, entre duas categorias de sons, existe um contínuo físico, gradiente. Como as alofonias contínuas não são apenas físicas, mas são gramaticais, é preciso representá-las, através de regras que incluam operações numéricas.

A questão é como fazê-lo, já que os modelos fonológicos que adotam primitivos de análise estáticos, conforme venho insistindo nesta tese, lidam apenas com regras de operações simbólicas. Como consequência, o tratamento dessa gradiência via traços distintivos é impossível. Ressalte-se, aliás, que a própria oposição /r/ ~ /r/ não é bem sucedida, num modelo de traços distintivos: como comento em 2.5.3.2, a caracterização desses segmentos por [tenso] e [frouxo], respectivamente, não capta a diferença exata entre eles, pois o *tap*, embora não seja tão “tenso” como a vibrante, demanda também grande esforço articulatório, o que o faz um dos últimos segmentos a serem adquiridos, durante o desenvolvimento do sistema fonológico. Além disso, esses traços não captam a identidade de classe entre as realizações vibrantes e as realizações fricativas dos róticos, como apontado por Maia (1982)¹⁴⁸.

A proposta de Maia (*op.cit.*), que prevê um traço [rotacizado] para caracterizar todos os róticos, vibrantes ou fricativos, capturando assim a identidade de classe entre as variantes, também não dá conta de representar a gradiência, observada nos dados de AJS e JLM, por dois motivos: o primeiro deles é que, embora esse traço seja escalar, e não bivalente, como os concebidos por Chomsky & Halle (1968), por exemplo, não é possível que se proponha um traço [1,5 rotacizado], para representar uma variante, e outro [1,6 rotacizado], para representar outra variante, porque a gradiência verificada implica na existência de variantes entre esses dois pontos. Além disso, há que se considerar que as escalas são muito aleatórias, no tratamento da gradiência, pois se pode postular escalas *ad hoc*. O segundo motivo é que a maioria dos modelos fonológicos,

de condicionamento prosódico sobre as variantes: como as fricativas têm maior duração intrínseca que outros sons, os resultados obtidos podem sinalizar mesmo o enfraquecimento da vibrante, nos dados de JLM.

¹⁴⁸ O tratamento da oposição entre as variantes de /r/ via [contínuo] também é problemático e controverso, conforme aponto no mesmo item 3.5.3.2.

desde SPE até as Geometrias de Traços, não aceita traços de natureza escalar em seu inventário. Adotá-lo para a representação dos róticos acarretaria, conseqüente e obviamente, uma incompatibilidade com os arcabouços teóricos disponíveis.

Quanto à representação dos róticos, numa Geometria de Traços, – da qual trato em 2.6, ao abordar a representação das líquidas – é igualmente problemática: mesmo que se proponha um traço [vibrante], seguindo sugestão de Ladefoged (1975) e que esse traço não seja restrito à estrutura interna de líquidas, mas possa ocorrer também na estrutura de oclusivas, sempre atrelado ao nó Raiz, ele deve ser, *by default*, binário ou monovalente. Portanto, não pode representar a alofonia contínua dos róticos¹⁴⁹.

Essa insuficiência dos modelos derivacionais – como a Geometria de Traços – advém do fato de que o máximo que esses modelos permitem é obter uma categoria “y”, a partir de uma categoria “x”. Modelar a gradiência verificada só seria possível se diversas formas “superficiais” fossem previstas. Isso, num modelo derivacional, é simplesmente inviável, porque significa prever uma representação fonética absurdamente imensa, com tantos símbolos quantos forem os graus de variação, observados para um dado segmento.

4.1 - Retomando alguns pressupostos da FAR

Dadas as limitações dos modelos estáticos é que proponho o tratamento dos róticos do PB, num modelo dinâmico de produção da fala. Para essa proposta, baseio-me em dois modelos dessa natureza: a Fonologia Articulatória, ou FAR (Browman & Goldstein, 1986, 1989, 1990, 1992) e a Fonologia Acústico-Articulatória, FAAR (Albano, 2001). Eles não são excludentes, mas complementares: a FAAR parte dos pressupostos da FAR, revendo alguns deles e trazendo ao modelo algumas contribuições, das quais as mais significativas são a incorporação de informação acústica ao modelo e a lexicalização de variantes dialetais. Estes pontos eu comento no primeiro capítulo e voltarei a eles, brevemente, no decorrer do presente capítulo.

¹⁴⁹ Além disso, algo que poderia ser questionado, relativamente a esse traço, é sua subordinação ao nó Raiz: poder-se-ia pensar na possibilidade de atrelá-lo ao nó “Cavidade Oral”.

Antes, é preciso lembrar que, ao contrário dos modelos fonológicos tradicionais, que prevêm a organização do processamento de fala sobre uma base física e outra cognitiva – sem qualquer relação intrínseca –, a FAR assume que as duas bases constituem descrições de duas dimensões de um sistema complexo de processamento da fala. Isto garante a possibilidade de “tradução” da base cognitiva na base física, intrinsecamente incomensuráveis.

A “tradução” de uma base noutra é possível, graças aos primitivos teóricos tomados, os “gestos articulatórios”. Eles são unidades fonológicas (cognitivas, portanto), dotados de unidades de ação articulatória (física), dinamicamente especificadas. Uma vez que cada gesto é definido como unidade dinâmica, não são necessárias regras de implementação para caracterizar as propriedades de cada nível (superior/cognitivo ou inferior/físico) do enunciado. Um padrão de movimento dos articuladores, variável no tempo, é “legalmente implementado” pelos sistemas dinâmicos, que são auto-implimentacionais.

Os gestos são dotados de tempo intrínseco¹⁵⁰, o qual é determinado pela relação entre a rigidez e o amortecimento dos articuladores, envolvidos na realização da “tarefa dinâmica”, que nada mais é do que atingir um alvo articulatório particular. A especificação de um gesto oral inclui, no arcabouço teórico da FAR, uma posição de equilíbrio, ou objetivo, para cada duas variáveis do trato, bem como uma rigidez. (Esses parâmetros são especificados por uma equação dinâmica, sendo sua variação responsável pelas diferenças intergestuais.) Cada objetivo funcional de um gesto é alcançado através da ação coordenada de um conjunto de articuladores.

As pautas gestuais – como já observado em 1.3.1, *supra* – formalizam a organização dos gestos que controlam diferentes conjuntos de articuladores, cinco, no total, conforme a previsão do modelo (vide Browman & Goldstein, 1989). Cada um desses conjuntos é descrito por variáveis relacionadas ao grau de amortecimento dos gestos, i.e., os descritores gestuais para local e grau de constrição.

Assim, na Figura 4.1, *supra*, o descritor “fechado” especifica o gesto de ponta de língua dos róticos, nas versões que tendem a vibrante, dada a característica oclusiva desse gesto, o qual tem

¹⁵⁰ Na verdade, o fato de serem dotados de tempo intrínseco é que confere aos gestos articulatórios seu caráter auto-implimentacional.

ponto de articulação alveolar¹⁵¹. O descritor “crítico” é atribuído ao gesto de corpo de língua, pois o articulador corpo da língua não promove obstrução do trato, mas se aproxima de um outro articulador, e.g., o véu palatino, fazendo uma constrição. O ponto de articulação, como noto acima, muda com o aumento da magnitude do gesto de corpo de língua, desde uma posição anteriorizada, até uma posição velar.

O problema de admitirmos, com a FAR, o conjunto de descritores articulatorios, é que não dispomos de informações precisas sobre a configuração dos articuladores, requerida pelos róticos do PB. Um outro fato a ser considerado, e a que concerne uma certa ressalva em adotar os descritores da FAR, é a crítica de Albano (2001) a eles: a autora observa que os descritores são estabelecidos com base em inventários como o IPA, o que lhes confere um caráter impressionístico.

Some-se a esses pontos um outro, no qual toco acima: a FAR dá conta especialmente bem de vogais e obstruintes, mas permanecem problemáticos nela os segmentos que os modelos fonológicos tradicionais reúnem na classe das soantes¹⁵².

4.2 – Para além da FAR

4.2.1 – O modelo “bigestual” de Sproat & Fujimura para as laterais do inglês norte-americano

Uma contribuição para a representação das soantes – em especial das líquidas – vem pelas mãos de Sproat & Fujimura (1993). Eles não modelam o trato tridimensionalmente, como seria

¹⁵¹ Frise-se que, embora eu diga que o ponto de articulação da vibrante é alveolar, trata-se de uma caracterização impressionística, posto que não se dispõe de dados articulatorios do PB que possibilitem precisar o ponto de articulação desses segmentos.

¹⁵² Cabe notar que estabelecer classes fonológicas só faz sentido no arcabouço de modelos fonológicos derivacionais, que necessitam do conceito de classes para lidar com os domínios de aplicação de regras. Em modelos implementacionais, como a FAR ou a FAAR, uma vez que não há regras que derivem uma estrutura Y, a partir de uma X, não há preocupação em estabelecer classes de segmentos. Uso, por vezes, nesta tese, o termo “classe de soantes”, ou “classe dos róticos”, para facilitar ao leitor a identificação, por exemplo, dos segmentos que, na escrita, são grafados com a letra r. Note-se, portanto, que, ao empregar o termo “classe”, no arcabouço da FAR, sigo apenas fins expositivos.

necessário no caso desses segmentos, mas prevêem, a partir de estudos articulatórios, que um mesmo segmento possa ser constituído de mais de um gesto.

Sproat & Fujimura (*op.cit.*) investigam a lateral do inglês norte-americano, procurando demonstrar, a partir de dados articulatórios, que não existe propriamente uma variante *dark* (velarizada, típica de final de sílaba) e outra *light* (típica de início de sílaba). O que existe é um contínuo físico pelo qual se estende a realização da lateral, que pode tender a *dark* ou *light*, conforme a força da fronteira, seguinte à lateral.

Os dados articulatórios levam Sproat & Fujimura a propor que a lateral é constituída de dois gestos, um de ponta e outro de dorso de língua – ou, como denominam, respectivamente, um gesto consonantal e outro vocálico. Esta é a grande contribuição dos autores à reflexão fonético-fonológica que assume a possibilidade de uma representação dinâmica: a previsão de que um único segmento pode ser constituído por mais de um gesto.

Para modelar a gradiência, envolvida na variabilidade de pronúncia da lateral, conforme a força da fronteira adjacente, Sproat & Fujimura prevêem a coordenação dos dois gestos, com base no parâmetro que denominam *tip delay*, isto é, o espaço temporal entre a realização do gesto de ponta e do gesto de dorso. Assim, se o *tip delay* é negativo, em decorrência de que o gesto de ponta da língua se inicia antes do gesto de dorso, tem-se uma variante anteriorizada da lateral, isto é, uma variante tendendo a *light*. Se, ao contrário, o *tip delay* é positivo, em decorrência de o início do gesto dorsal anteceder o de ponta de língua, então ocorre uma variante posterior, ou que tende a *dark*.

Com base neste estudo, a hipótese que orienta minha proposta de representação dos róticos é de que também essa classe de sons – como as laterais – envolva dois gestos articulatórios, em sua produção: um gesto de ponta de língua e outro de dorso, como já esboço na pauta gestual da Figura 4.1, *supra*. A hipótese de que róticos, como as laterais, são segmentos que envolvem mais de um gesto articulatório se calca na semelhança fonética entre as duas classes de sons e que permite processos como o rotacismo da lateral, o qual ocorre tanto diacrônica como sincronicamente. Assim é que se tem, no latim vulgar, a evolução de “*regula*” a “*regla*” e, no português, “*regra*”. Sincronicamente, tem-se no espanhol e no catalão padrões “*glândula*” e, na versão coloquial dessas línguas, “*grândula*”. (Vide segundo capítulo, item 2.2 para outros

exemplos de rotacismo em português, espanhol e catalão e para exemplos de alternâncias entre laterais e róticos em línguas pertencentes a outros troncos lingüísticos.)

Admitir de saída a pauta de 4.1 como representação dos róticos do PB, porém, traz alguns problemas: 1) como não há informações exatas sobre a configuração dos articuladores, requerida pelos róticos do PB¹⁵³, pode-se incorrer numa representação imprecisa, caso se tomem apenas parâmetros articulatórios para modelá-los; 2) a pauta não dá conta de representar aspiradas, por um lado, e vibrantes uvulares, por outro. Embora estas estejam em desuso, como menciono no terceiro capítulo, estudos variacionistas ainda as registram em dialetos como o carioca. Note-se: ainda que o *tap* não seja mencionado na Figura 4.1, sua pauta seria, em princípio, a mesma da vibrante alveolar. A distinção entre ambos se dá pelo parâmetro “rigidez”, especificado na equação dinâmica, que implementa os gestos articulatórios.

Para contornar os problemas que aponto acima, uma solução seria considerar aspectos acústicos, incorporando-os à definição dos gestos articulatórios¹⁵⁴. A Fonologia Acústico-Articulatória (FAAR), de Albano (2001), propõe algo nessa direção, postulando, ao invés de articuladores do trato, regiões acústico-articulatórias que definem os gestos. Nesse sentido, bem como em alguns outros que exponho adiante, a FAAR se constitui numa importante contribuição à FAR.

4.2.2 – O tratamento “bigestual” dos *glides* e líquidas do inglês norte-americano por Gick

Um outro estudo que reforça a hipótese acerca de se representarem os róticos do PB por mais de um gesto articulatório é o tratamento bigestual dos *glides* e das líquidas, proposto por Gick (1999).

A questão central desse estudo, na verdade, é buscar uma descrição da estrutura silábica, em termos acústicos e articulatórios consistentes, já que, apesar da sua importância para a teoria fonológica e apesar de falantes de qualquer língua aparentemente terem um conhecimento

¹⁵³ Refiro-me, obviamente, a estudos articulatórios e não às descrições impressionísticas disponíveis.

¹⁵⁴ Sobre os ganhos que a incorporação de informação acústica traz à FAR, vide nota 55, no item 1.4, do primeiro capítulo.

intuitivo sobre a delimitação de sílabas, a descrição e delimitação da estrutura silábica são um problema tanto para foneticistas quanto para fonólogos experimentais.

Mesmo que se assuma – como faz Turk, 1994, *apud* Gick, 1999 – que consoantes iniciais de palavra marcam o início de uma sílaba e consoantes finais de palavra marcam o final de uma sílaba, Gick observa que há evidências fonológicas abundantes de fatos que contestam essa assunção, como os casos de ressilabificação parcial. O *flapping*, verificado no inglês norte-americano, é um exemplo: o resultado da operação que converte uma oclusiva final de palavra diante de vogal em início de uma palavra subsequente, é um segmento ambissilábico, i.e., cujas características fonéticas se confundem com as de um segmento intervocálico.

Gick se ocupa, em sua investigação, de /j/, /w/, /r/ e /l/ ambissilábicos: ele observa que, embora haja estudos articulatórios que medem a magnitude dos gestos e sua sincronização para identificar os correlatos fonéticos de várias posições silábicas, como fazem Sproat & Fujimura (1993 – vide seção anterior) ou Krakow (1989, *apud* Gick, 1999), e que essa caracterização seja consensual, não há consenso acerca das propriedades articulatórias de alofones ambissilábicos (além de que são poucos os investigadores que se voltam para essa questão).

Tentando contribuir para o desenvolvimento de um modelo de estrutura silábica, baseado na magnitude e na organização temporal dos gestos articulatórios que constituem segmentos ambissilábicos, Gick conduz um experimento articulatório,¹⁵⁵ através do qual tenta identificar os gestos que compõem os alofones pré e pós-vocálicos de /w/, /j/ e /l/ do inglês norte-americano. A hipótese que orienta sua investigação é a que os gestos intervocálicos, que constituem os segmentos ambissilábicos, devem coordenar-se tanto com o gesto vocálico precedente quanto com o gesto vocálico seguinte. Observe-se que esta é uma perspectiva distinta daquela que tradicionalmente se toma para se referir à ambissilabicidade e segundo a qual uma consoante intervocálica se afiliaria, ao mesmo tempo, com dois picos vocálicos adjacentes (havendo ou não fronteiras).

¹⁵⁵ Os dados do experimento, segundo o autor, foram coletados utilizando-se o “articulômetro eletromagnético médio-sagital” (EMMA), desenvolvido nos Haskins Laboratories. Eletrodos foram fixados em quatro pontos médio-sagitais da língua dos sujeitos, um eletrodo foi fixado no lábio superior, outro no lábio inferior, mais um na mandíbula, nos maxilares e na cavidade nasal (estes dois últimos, cf. Gick, empregados com o objetivo de corrigir eventuais movimentos da cabeça).

Os estímulos do experimento foram elaborados de maneira a permitir uma comparação controlada de alofones de *glides* e líquidas, em final de palavra, silabificados com e sem ambigüidade. Para isso, no caso da comparação dos alofones de /l/ final e ambissilábico, foi utilizado o par *hall hotter* vs. *hall otter*, inserido em sentenças-veículo. Esse par foi repetido de 8 a 12 vezes por dois sujeitos.

Para um dos sujeitos, Gick mediu apenas a trajetória do gesto de ponta de língua, dada, segundo o autor, a impossibilidade de se medir as trajetórias do gesto dorsal. Para o outro sujeito, mediu-se a trajetória do gesto de ponta de língua; a trajetória do gesto de retração dorsal da língua; o *tip delay*¹⁵⁶ entre os dois gestos.

Note-se que, com base em Sproat & Fujimura (*op. cit.*), Gick reconhece a existência de dois gestos constitutivos da lateral¹⁵⁷, o de ponta de língua – ao qual ele inicialmente denomina, também a partir de Sproat & Fujimura, “consonantal” – e o gesto de dorso de língua, um gesto vocálico. De acordo com Gick, as medidas das trajetórias dos gestos de /l/ mostram que a previsão inicial, segundo a qual haveria um efeito intermediário,¹⁵⁸ na magnitude dos gestos e na coordenação entre eles, não se verifica totalmente: isso porque o segmento “ambíguo” se assemelha ao /l/ final, quanto à retração do dorso da língua, mas é intermediário a ele e ao inicial, quanto à projeção da ponta da língua. Essas afirmações advêm das observações de que o gesto de ponta de língua tem a maior projeção anterior para o caso em que /l/ ocupa início de palavra e menor projeção, no caso em que /l/ ocupa posição final de palavra. Esse mesmo gesto tem projeção anterior intermediária, no caso de um segmento “ambíguo”. Tal comportamento, em princípio, confirmaria a hipótese inicial.

O gesto de retração dorsal, porém, contrariando a previsão inicial, tem mesma magnitude em qualquer posição que ocupe, i.e., em posição inicial, final ou ambissilábica.

¹⁵⁶ Vide seção anterior deste capítulo, para detalhes acerca do que seja o *tip delay*, parâmetro inicialmente empregado por Sproat & Fujimura (1993).

¹⁵⁷ Na verdade, Gick vai além, trazendo evidências de que também os *glides* e /r/ são constituídos, cada um, por dois gestos articulatórios, um que se assemelha a uma vogal, o “gesto-V”, e um que se assemelha a uma consoante, o “gesto-C”.

¹⁵⁸ Intermediário ao que se verifica para /l/ inicial e /l/ final, segundo Sproat & Fujimura (*op. cit.*), i.e., para /l/ inicial de sílaba/palavra, espera-se maior ação temporal do gesto de ponta de língua. Contrariamente, para o /l/ final de sílaba/palavra, espera-se maior ação temporal do gesto de retração dorsal.

Vê-se, portanto, que o gesto consonantal – o de ponta de língua – é mais sensível ao contexto seguinte a /l/ do que o gesto de dorso de língua. Isto leva Gick a concluir que é exatamente esse o efeito relativo à distinção entre alofones ambissilábicos e de final de palavra, pois, no caso dos alofones finais, o gesto consonantal é menos sensível ao contexto seguinte.

Este estudo, portanto, mostra uma diferença de sincronização dos dois gestos que constituem *glides* e líquidas – em especial a lateral /l/, que é o segmento ao qual me atenho nesta seção, buscando fazer um paralelo com o estudo de Sproat & Fujimura – relativamente à sincronização desses mesmos gestos, nos alofones iniciais e finais de /l/ norte-americano. Temos, então, novas evidências de que diferentes organizações gestuais – seja essa diferença devida ao grau de sobreposição dos gestos, seja ela devida a diferentes intervalos de ativação – podem dar conta de representar diferentes alofones.

4.2.3 – O gesto articulatório definido em função de regiões acústico-articulatórias: a contribuição da FAAR

O ponto-chave da Fonologia Acústico-Articulatória (FAAR, Albano, 2001), é a tese da comensurabilidade. A propósito dela, constrói-se toda a argumentação da autora e, como consequência, surgem as contribuições à FAR. Tal tese é definida, pela própria autora, da seguinte maneira:

A tese da comensurabilidade afirma que a fonologia, qualquer que seja sua organização, e a fonética (isto é, o modelo lingüístico de produção da fala) têm muitos parâmetros em comum, que diferem apenas quanto à participação discreta ou contínua das dimensões envolvidas. O conjunto discreto e reduzido de pontos a serviço da primeira deve, em princípio, emergir a partir daquelas regiões dos contínuos

*mobilizados pela segunda nas quais atuam fatores que favoreçam a manifestação – via percepção*¹⁵⁹, por exemplo – de descontinuidades. (Albano, 2001: 104)

O excerto acima marca um distanciamento entre a FAAR e a FAR: Browman & Goldstein, tentando romper com a tradição que considera que o processamento da fala se dá em dois módulos distintos – o Fonético e o Fonológico – preconizam, de modo radical, que o processamento da fala se dê num único módulo, o Fonético. A Fonologia, portanto, seria “excluída” da gramática. Ora, excluir a Fonologia da gramática traz alguns problemas sérios, como impossibilidade de prever a ação prosódica sobre a natureza dos segmentos. Isto porque essa medida atribui à Fonética um poder enorme, já que esse módulo “enxerga” diretamente o módulo sintático, sem intermediações. Cria-se, então, um abismo entre os aspectos simbólico e físico de uma língua (cf. Albano *et alii*, 1998 ou Albano, 2001).

Albano propõe, por outro lado, a existência de um módulo “fonético-fonológico”, no qual os limites entre o categórico – como na Fonologia tradicional – e o gradiente – como na Fonética tradicional – são nebulosos. Decorre daí que uma mesma representação pode conter elementos categóricos e gradientes. Essa é a base da “tese da comensurabilidade”, que prevê, ainda, e como consequência do que foi dito anteriormente, que os extremos de um contínuo físico podem ser duas categorias, as quais se estabelecem nas bordas dos gestos articulatórios, através, e.g., de fatores como a percepção. Nesse sentido, as categorias se estabelecem a partir da “cristalização” de um padrão gestual, a qual se dá pela repetição desses padrões, tanto sincrônica como diacronicamente.

Eis aí mais um aspecto novo da FAAR com relação à FAR: a incorporação de informação diacrônica, que visa a explicitar a relação entre uma variante conservadora e outra, inovadora. Nesse sentido é que Albano (*op.cit.*) postula a fricativa como uma versão enfraquecida da vibrante uvular. Essa manobra teórica, para a autora, possibilitaria captar a relação entre versões fricativas e vibrantes dos róticos. Como consequência, poder-se-ia compreender, nos processos de *sandhi* externo, a escolha do *tap*, em final de palavra e diante de vogal, mesmo em dialetos nos

¹⁵⁹ Nota de Albano (*op. cit.*: 104): *Browman e Goldstein (1990b) mencionam também, sem aprofundar, a possibilidade de regiões discretas emergirem como atratores nas próprias funções contínuas que descrevem os*

quais ocorre uma fricativa em final de palavra, seguida de pausa ou de consoantes. Para Albano, o *sandhi* externo tornaria obrigatória a escolha do *tap*, sendo a variante, utilizada nos demais contextos, uma *fricativa faríngea acoplada a um gesto dorsal simultaneamente curto e reduzido* (Albano, *op.cit*: 133). As pautas gestuais, propostas por Albano para /r/, nesses contextos, são:

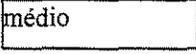
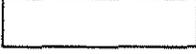
Final de sílabas		
	(a) outros contextos	(b) diante de vogal
Região coronal		
Grau de constrição		
Local de constrição		
Região dorsal		
Grau de constrição		
Local de constrição		
Região faríngea		
Grau de constrição		
Local de constrição		

Figura 4.2 - Proposta para pauta gestual de fricativas e vibrantes do rótico final de sílaba em PB, cf. sugestão de Albano (2001: 134).

Com relação à pauta, é preciso comentar, antes das implicações que ela traz para a representação dos róticos, a notação aí empregada: no primeiro capítulo e no item 4.1, *supra*, eu chamava a atenção para o fato de que a FAR postula que os gestos articulatórios são definidos por cinco conjuntos de articuladores, os quais, por sua vez, são especificados por descritores gestuais de local e modo de constrição. Assim é que a pauta da Figura 4.1 representa os gestos articulatórios, envolvidos na produção da vibrante e de suas versões enfraquecidas, até a fricativa.

Nas pautas de vibrantes e fricativas, que traz a Figura 4.2, entretanto, os gestos não são definidos em termos de conjuntos de articuladores, mas em termos de regiões acústico-articulatórias. A decisão de Albano em definir os gestos em termos de regiões acústico-articulatórias origina-se de um certo paradoxo que a autora identifica na FAR: Albano observa que, dado o importantíssimo papel conceitual que os articuladores assumem, na Fonologia Articulatória, – que é o de atribuir a um mesmo gesto as duas trajetórias que definem uma constrição – não é possível dissociar os tempos dessas trajetórias, as quais se deslocam na dimensão vertical e na dimensão horizontal de um plano sagital. Como – segundo frisa Albano – *um comando para ativar o domínio de um articulador deve atingir tudo o que está sob ele* (Albano, *op. cit.*: 97) – justifica-se, em termos teóricos, que a cada dupla de variáveis seja atribuída a mesma duração. Na prática, porém, é possível definir uma equação dinâmica para cada variável, o que resultaria numa especificação de rigidez em cada equação.

Para resolver esse paradoxo, substituindo o articulador no papel de unir as trajetórias de um gesto, Albano propõe que o gesto se defina a partir do espaço acústico-articulatório, estabelecido *por regiões discretas do trato vocal enquanto tubo acústico* (Albano, *id.*: 98)¹⁶⁰. Assim é que a pauta gestual, na FAAR, passa a considerar regiões articulatórias, como a coronal, a dorsal e a faríngea, da Figura 4.2, acima, ou ainda a labial e a região do “pórtico nasal”. As regiões acústico-articulatórias são definidas pelas mesmas variáveis do trato e descritores gestuais que definiam os conjuntos de articuladores¹⁶¹.

A incorporação de informação acústica ao modelo, mais uma contribuição da FAAR à FAR, é importante para a tese da comensurabilidade uma vez que, ao resolver o paradoxo da FAR, permite uma melhor “tradução” do dinâmico no simbólico¹⁶².

Um outro aspecto da tese da comensurabilidade é a previsão de que as variantes dialetais têm suas pautas gestuais lexicalizadas¹⁶³. Essas pautas são escolhidas pelo falante através de um

¹⁶⁰ Tal proposta se baseia, segundo a autora, nos trabalhos de Ohala, Lindblom e Stevens e encontra respaldo num modelo de síntese articulatória já implementado, o Modelo das Regiões Distintivas (Carré & Mrayati, 1990; Carré, 1994, *apud* Albano, *op. cit.*)

¹⁶¹ Ressalte-se que, como Browman & Goldstein, Albano não prevê uma variável do trato “forma da constrição”.

¹⁶² Este parece-me o ponto mais importante da incorporação de aspectos acústicos ao modelo. Poder-se-ia pensar ainda noutra, óbvia, que é a possibilidade de uma formalização mais completa do gesto articulatório através da consideração de mais um aspecto fonético, o aspecto acústico. O aspecto perceptual, tanto na FAR como na FAAR, é abordado rapidamente, não sendo considerado, como os demais, na formalização do gesto.

mecanismo de *look ahead*, à medida que ele planeja a cadeia da fala. Esse mecanismo é que permitiria a um falante de um dialeto em que o rótico final de palavra é realizado como fricativa, diante de pausa ou consoantes, “escolher” a variante [r], quando o rótico precede vogal. Essa previsão do modelo explica, portanto, mecanismos de *sandhi* externo, envolvendo os róticos. Ressalte-se que a proposta de lexicalização das variantes dialetais não se coloca como um problema para o modelo: dado seu caráter implementacional, o léxico pode ser constituído de um grande número de itens, sem que isso se torne um empecilho para o processamento da fala.

Ainda no que concerne à Figura 4.2, é preciso notar que a pauta gestual para fricativas considera que o grau de constrição da região dorsal tenha magnitude reduzida, relativamente ao local de constrição,¹⁶⁴ e que tanto o grau como o local de constrição da região dorsal sejam mais breves do que o grau e o local de constrição da região faríngea. O descritor “crítico” para o grau de constrição da região faríngea garante o caráter fricativo do segmento.

Albano prevê, para a variante vibrante de /r/, por seu turno, que os gestos da região coronal tenham menor duração¹⁶⁵ que os gestos da região faríngea. A menor duração dos gestos da região coronal, aliada à especificação “fechado” para o grau de constrição, nessa região, confere à vibrante seu caráter balístico. Note-se também que a senóide de meio ciclo, na caixa do grau de constrição da vibrante, indica o grau de amortecimento dessa variante. Essa notação é importante para distingui-la da vibrante múltipla, que talvez tivesse o grau de amortecimento sinalizado por uma senóide de ciclo completo. Albano não sinaliza essa possibilidade, ou tampouco qualquer outra. Observa, apenas, que são necessários dois graus de amortecimento não crítico – fechado, portanto, *by default* – para distinguir a vibrante simples da múltipla¹⁶⁶.

¹⁶³ A FAR não faz menção à variação dialetal, a não ser ao postular que esse fato pode resultar de diferentes graus de sobreposição gestual (vide Browman & Goldstein, 1986, 1992).

¹⁶⁴ Como se vê na dimensão vertical das caixas que representam os gestos, nessa pauta.

¹⁶⁵ Veja-se a dimensão horizontal das caixas que representam os gestos.

¹⁶⁶ Importante frisar que a representação gestual da lateral alveolar envolve as mesmas regiões articulatórias da representação das vibrantes. A diferença entre esses segmentos, porém, está no tempo de ativação dos gestos da região coronal, que é maior para a lateral, igualando-se à duração dos gestos da região faríngea. Esta previsão da FAAR capta a identidade entre os segmentos e explica os processos de rotacismo da lateral. Tal proposta encontra respaldo no fato de que a lateral, em grupos, tem menor duração do que nas demais posições em que pode ocorrer no PB (cf. Silva, 1996, *apud* Albano, 2001) e também no fato de que é possível “fabricar” um *tap*, a partir da lateral, cortando-se parte do sinal acústico desta.

Outro fato a se considerar, na Figura 4.2, é a ação da região faríngea na produção dos róticos. Justificando-a, a autora observa que

Uma constrição faríngea é obrigatória para todos os róticos, podendo ser acompanhada de uma constrição coronal, como nas vibrantes alveolares, ou de uma constrição dorsal, como nas vibrantes uvulares. (Albano, op. cit: 131)

As pautas gestuais, propostas por Albano para representar os róticos, não captam a relação entre vibrantes e fricativas velares com a fricativa glotal. Nesse sentido, proponho, na sessão seguinte, que a representação gestual dos róticos deva contemplar também a ação da região glotal.

Ainda um aspecto a ser observado, na proposta de Albano: a FAAR prevê que a fricativa velar seja uma redução da vibrante uvular, sob o argumento de que esse postulado resgata a relação da vibrante uvular com a vibrante alveolar, mesmo que em termos diacrônicos. Estudos variacionistas, como o de Callou (1987), entretanto, defendem que a fricativa velar seja um enfraquecimento da vibrante alveolar. Assim, embora ambas as propostas considerem a fricativa velar uma versão enfraquecida, as direções previstas para o enfraquecimento são distintas: no primeiro caso, prevê-se uma anteriorização dos segmentos, concomitante à mudança de modo de articulação. No segundo, postula-se posteriorização concomitante à mudança de modo.

Sigo, porém, em minha proposta (vide item 4.3, abaixo), o postulado da FAAR, pois ela parece mais plausível, diante da observação de Gonçalves Viana, que transcrevo no início do terceiro capítulo, e que relata, na fala de paulistas e pernambucanos do final do século XVIII, uma alternância entre vibrantes uvulares, vibrantes apicais e as variantes que denomina “fricativas sonoras” – algo semelhante às vibrantes espirantizadas, pela descrição disponível.

A descrição de Viana permite pensar que, no momento em que a língua se projeta para a região anterior do trato, sua ponta pode efetivamente tocar os alvéolos, realizando obstruções breves, ou pode não tocar os alvéolos. Nesse caso, o dorso da língua se aproximaria da região alveo-palatal, originando aí uma constrição, responsável pela fricção.

4.3 – Proposta de pauta gestual para os róticos do PB

Embora siga, para minha proposta, a sugestão da FAAR para a pauta gestual dos róticos no PB, tento contemplar também a fricativa glotal, que a FAAR não contempla, como observo no item anterior. Há que se ressaltar que minha proposta tenta contemplar os róticos de início de palavra, especialmente, embora sejam feitas algumas considerações acerca dos róticos, em posição final. Assim, temos:

Região coronal		Região coronal		Região coronal	
Grau de constrição		Grau de constrição		Grau de constrição	fech
Local de constrição		Local de constrição		Local de constrição	
Região dorsal		Região dorsal		Região dorsal	
Grau de constrição	crítico	Grau de constrição	crítico	Grau de constrição	
Local de constrição		Local de constrição		Local de constrição	
Região faríngea		Região faríngea		Região faríngea	
Grau de constrição		Grau de constrição	médio	Grau de constrição	médio
Local de constrição		Local de constrição		Local de constrição	
Região glotal		Região glotal		Região glotal	
Grau de constrição	ab. máxima	Grau de constrição	ab. máxima	Grau de constrição	fechado
Local de constrição	interariten.	Local de constrição	interariten.	Local de constrição	glote
	fricativa velar sonora		fricativa glotal sonora		<i>tap</i> ¹⁶⁷

Figura 4.3 - Pautas gestuais para versões fricativas sonoras e vibrantes do rótico, incorporando-se a ação dos gestos da região glotal.

A pauta da figura acima introduz um descritor gestual “interaritenóides”, para caracterizar o local de constrição do gesto da região glotal das variantes fricativas posteriores, e um descritor “abertura máxima”, para o grau de constrição dessa mesma região para as mesmas variantes. Isto

¹⁶⁷ Estou utilizando nesta Figura rótulos “categóricos” para as variantes apenas com fins expositivos. Leia-se, para “fricativa velar”, “variantes de /r/ que tendem a velar”, e assim por diante.

porque, para realizar as variantes fricativas posteriores sonoras, talvez seja necessária uma abertura máxima de glote, entre as aritenóides, para reduzir o fluxo com sopro.

Além disso, a pauta prevê que: 1) mais de um gesto caracterize as três variantes; 2) dois gestos – no caso o da região dorsal e o da região faríngea – são vocálicos; 3) o outro gesto, consonantal, é fechado, como no caso do *tap*, ou crítico, como no caso das variantes fricativas; 4) a gradiência ocorre, a partir da redução do fechamento e concomitante amplificação do gesto crítico, como já observado na página 138. Além disso, cabe notar que as três versões fricativas têm ativados os gestos da região faríngea¹⁶⁸. A pauta prevê também, e ainda seguindo Albano, que a versão vibrante tenha gestos de local e grau de constrição da região coronal de mesma magnitude e igualmente breves. O que difere nessa proposta, relativamente à de Albano, é que ela prevê que todos os róticos tenham, em sua pauta, especificação para a região glotal. Tal previsão, por sua vez, requer que se contemplem na proposta também as variantes surdas. Uma pauta possível para elas seria:

¹⁶⁸ Seguindo sugestão de Albano, exposta no item anterior, segundo a qual todos os róticos envolveriam atividade faríngea.

Região coronal		Região coronal	
Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição	
Região dorsal		Região dorsal	
Grau de constrição	crítico	Grau de constrição	
Local de constrição		Local de constrição	
Região faríngea		Região faríngea	
Grau de constrição		Grau de constrição	médio
Local de constrição		Local de constrição	
Região glotal		Região glotal	
Grau de constrição	Ab. máxima	Grau de constrição	Ab. máxima
Local de constrição	glote	Local de constrição	glote
	fricativa velar surda		fricativa glotal surda

Figura 4.4 - Pautas gestuais para versões fricativas surdas de /r/, incorporando-se a ação dos gestos da região glotal.

A pauta da Figura 4.4 prevê que as variantes surdas tenham a glote muito aberta, como as aspiradas. As sonoras, por sua vez, e como sugerem as pautas da Figura 4.3, seriam murmuradas.

No geral, e considerando as duas pautas (Figura 4.3 e Figura 4.4), a fricativa glotal teria ativados os gestos de grau e local de constrição nessa região. Observe-se que, por essa proposta, o que diferencia uma variante que tende a velar de uma variante que tende a glotal é a região acústico-articulatória, onde os gestos são ativados: nessa, os gestos são ativados numa porção posterior do trato; naquela, numa porção um pouco mais anteriorizada. Também nas versões anteriorizadas, estaria especificada a presença de gestos da região glotal, mas com nenhuma magnitude.

Torna-se, então, possível captar as semelhanças entre essas variantes e explicar sua realização em posição final, nos dialetos que alternam versões vibrantes e fricativas – inclusive a velar – dos róticos. É também possível explicar a coocorrência dessas variantes, em posição inicial, no caso daqueles dialetos onde ainda se verificam versões vibrantes, em especial a

vibrante alveolar, mas também versões fricativas. É o caso do dialeto de OGF, cujos dados comento no item 3.1, do terceiro capítulo, e do dialeto de JLM, cujos dados são investigados no item 3.3.1, no mesmo capítulo.

Para explicar a coocorrência das variantes vibrantes e fricativas, na fala de OGF e JLM, a hipótese é de que os gestos da região coronal diminuem em magnitude, concomitantemente ao aumento da magnitude dos gestos da região dorsal. A vibrante espirantizada, presente nos dados de JLM, teria ativados gestos de três regiões: da coronal, da dorsal e da faríngea. Assim:

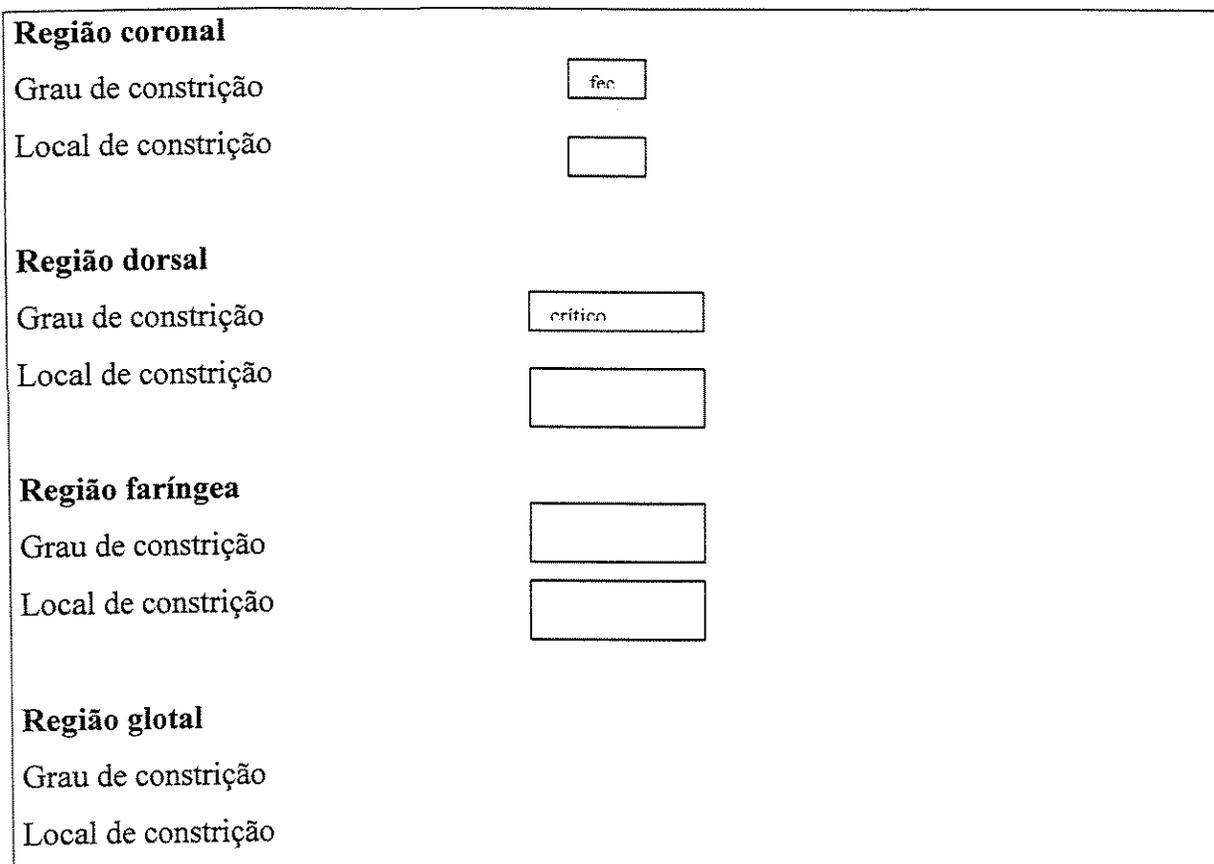


Figura 4.4 - Pauta gestual da vibrante espirantizada.

Essa pauta prevê que haja, na variante em questão, a atividade balística, oclusiva, da ponta da língua – fato característico das vibrantes -, em coocorrência com a atividade do dorso da

língua, que o faz aproximar-se do palato, constituindo uma constrição, sinalizada pelo descritor “crítico” para o grau de constrição da região dorsal¹⁶⁹.

Deve haver graus variáveis da magnitude dos gestos da região coronal e da região dorsal para que se module a gradiência do ruído fricativo, sobreposto à estrutura da vibrante, tal como observado no terceiro capítulo: gestos da região dorsal, apresentando magnitude maior que os gestos da região coronal, trariam como resultado maior sobreposição de ruído fricativo à estrutura da vibrante. Gestos da região dorsal de igual magnitude aos gestos da região coronal – note-se a referência ao parâmetro “magnitude”, e não à extensão temporal do gesto¹⁷⁰ – teriam como resultado maior sobreposição de ruído fricativo sobre aberturas orais e fechamentos. Quanto menor a magnitude do gesto da região coronal, mais o segmento tende a fricativa, até que o segmento se realize efetivamente, como tal, para a magnitude da região coronal igual a zero.

Ainda sobre a Figura 4.4 é preciso acrescentar que a pauta para a vibrante espirantizada, bem como para as demais versões da vibrante (Figuras 4.5 e 4.6, adiante), prevê que os gestos da região dorsal têm igual duração aos gestos da região faríngea. Essa foi uma decisão tomada *ad hoc*, porque a equivalência motora entre raiz e dorso de língua precisa de estudos articulatórios para ser confirmada. A proposta de equivalência entre os gestos das duas regiões pode suscitar um questionamento acerca de se admitir, com Albano (2001), que é possível reunir os gestos numa região “dorso-faríngea”. Prefiro, porém, deixar os gestos desmembrados, pela seguinte razão: Albano propõe uma região dorso-faríngea para tratar das vogais do PB, mas há língua, como o inglês, que fornecem evidências para se manter os gestos dissociados. No inglês, há diversos graus de abertura da vogal, relacionados ao seu abaixamento, o que quer dizer que o gesto de dorso não acompanha, necessariamente, o de raiz de língua. É importante que se diga que uma resposta definitiva às questões que pontuo neste parágrafo só poderá ser alcançada através de estudos articulatórios.

¹⁶⁹ Também para as vibrantes é possível tecer as mesmas generalizações, feitas a propósito da pauta da Figura 4.3, i.e., que as vibrantes se caracterizam pela ação de mais de um gesto, e que um desses gestos é vocálico, enquanto o outro, consonantal, é fechado ou crítico. Cabe observar que, embora não se contemple a vibrante uvular na pauta da Figura 4.3, ela teria uma pauta análoga à da vibrante alveolar, com a diferença de que os gestos da região glotal é que seriam ativados, e não da região faríngea.

¹⁷⁰ O gesto da região coronal é sempre mais breve, *by default*, o que lhe garante o caráter balístico.

Uma outra assunção de minha proposta – também à luz da FAAR – é que todas essas pautas estão presentes no léxico dos falantes do dialeto, no qual coexistem todas as variantes, i.e., as que tendem a vibrantes e as que tendem a fricativas. O falante “escolhe” a pauta gestual a ser empregada, a partir de informações como a força de fronteira anterior a /r/. Note-se que, como frisado no terceiro capítulo, o fato de haver nos dados de JLM a mesma variante de /r/, ocorrendo com maior duração depois de fronteira forte e com duração menor depois de fronteira fraca, é forte indício do condicionamento da força de fronteira sobre a variabilidade de pronúncia de /r/. Portanto, parece plausível postular que, seguindo fronteira forte, o falante se incline a escolher uma pauta de um segmento que tende a vibrante e que, seguindo fronteira fraca, o falante escolha uma versão enfraquecida – portanto tendendo a fricativa – do rótico.

Isto para posição inicial. Para posição final, embora eu não tenha realizado experimentos nesse sentido, uma análise de outiva dos dados de JLM sugere que o informante realize o *tap* aí, ou o suprima¹⁷¹. Então, a única pauta lexicalizada para este informante é a pauta do *tap*.

Para os dialetos em que não existe a alternância, verificada nos dados de JLM, as pautas lexicalizadas são, obviamente, outras. É o que ocorre, no dialeto de AJS. Como também observado no terceiro capítulo, coocorrem na fala deste informante, variantes vibrantes – simples e múltiplas – além de segmentos intermediários, as vibrantes encurtadas.

A FAAR prevê que a diferença entre vibrantes simples e múltiplas, i.e., entre *taps* e vibrantes alveolares, se dê pelo grau de amortecimento dessas variantes – seguindo, aliás, sugestão da FAR – e que, além disso, seriam necessários dois graus de amortecimento, para distinguir tais variantes. Assim:

Região coronal		Região coronal	
Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição	
Região dorsal		Região dorsal	
Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição	
Região faríngea		Região faríngea	
Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição	
Região glotal		Região glotal	
Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição	
	<i>tap</i>		vibrante alveolar

Figura 4.5 - Pauta gestual de *taps* e vibrantes, seguindo sugestão da FAAR.

A senóide de meio ciclo, na pauta do *tap*, indica, na notação da FAAR e como já observado nos comentários acerca da Figura 4.2, o grau de amortecimento dessa variante, em oposição à vibrante, cujo amortecimento pode ser sinalizado na pauta por uma senóide de ciclo completo ou pela ausência da senóide. A FAAR não faz qualquer sugestão a respeito, mas, como o amortecimento da vibrante simples é, por definição (cf. FAR), maior que o da vibrante múltipla, podemos prever que a pauta desta variante não contenha a senóide, na especificação do gesto de grau de constrição da região coronal.

A questão que os dados de AJS colocam para a FAAR concerne ao número de graus de amortecimento, necessários para prever também essa variante. Seriam necessários três? Neste

¹⁷¹ Cabe um estudo também sobre os róticos finais nesses dialetos em que o *tap* coocorre com um pretenso zero fonético, para se investigar se realmente os gestos do *tap* não se realizam ou se há algum vestígio deles, auditivamente imperceptível, mas visível acusticamente.

caso, pode-se prever que o amortecimento do *tap* seja sinalizado na pauta por uma senóide de ciclo completo para o grau de constrição da região coronal, que a vibrante encurtada seja sinalizada por uma senóide de meio ciclo e que a vibrante não tenha senóide, na especificação do grau de constrição da região coronal. Assim:

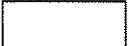
Região coronal		Região coronal		Região coronal	
Grau de constrição		Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição		Local de constrição	
Região dorsal		Região dorsal		Região dorsal	
Grau de constrição		Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição		Local de constrição	
Região faríngea		Região faríngea		Região faríngea	
Grau de constrição		Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição		Local de constrição	
Região glotal		Região glotal		Região glotal	
Grau de constrição		Grau de constrição		Grau de constrição	
Local de constrição		Local de constrição		Local de constrição	
<i>tap</i>		vibrante encurtada		vibrante alveolar	

Figura 4.6 – Diferença entre *tap*, vibrante encurtada e vibrante alveolar, através de graus distintos de amortecimento, a partir de sugestão da FAAR.

Há ainda um outro ponto a ser observado, relativamente às pautas das Figuras 4.5 e 4.6: nelas, assume-se que o gesto da região dorsal tem duração variável, sendo menor para as variantes tendendo a *tap* e maior para as que tendem a vibrantes alveolares. As vibrantes encurtadas teriam gesto de região dorsal com duração intermediária à de vibrantes e *taps*. O fato de se prever maior duração do gesto de região dorsal para a vibrante relaciona-se ao fato de que o gesto coronal repetitivo requer a ampliação do gesto dessa região. Outro fato a ser notado é que se assume que o gesto dorsal e o gesto faríngeo tenham mesma duração, para cada uma das variantes. Por fim, as pautas das Figuras 4.5 e 4.6 poderiam suscitar uma questão acerca da distinção entre vibrantes encurtadas e vibrantes espirantizadas (vide Figura 4.4), já que ambas têm especificados os gestos,

nas mesmas regiões articulatórias. É preciso lembrar que o parâmetro dinâmico que distingue essas duas variantes é a magnitude gestual: assume-se que, quanto maior o gesto da região dorsal e menor o da região coronal, mais espirantizado é o segmento. No caso da vibrante encurtada, não se prevê aumento da magnitude do gesto de região dorsal nem diminuição do gesto de região coronal. O parâmetro magnitude, neste caso, mantém-se inalterado.

Novamente, e a exemplo do que foi postulado sobre os dados de JLM, todas as pautas das variantes vibrantes são lexicalizadas no dialeto onde coocorrem. Uma pauta de uma variante específica é escolhida, durante o planejamento da fala, a partir da informação da força de fronteira que antecede o rótico: fronteiras prosódicas fortes orientariam a escolha da pauta para a vibrante alveolar. Em contrapartida, fronteiras prosódicas fracas requereriam pautas de versões enfraquecidas: a vibrante encurtada ou o *tap*.

Para final de palavra, coloca-se a mesma observação, feita para os dados de JLM, pois também AJS parece realizar o *tap* ou um zero fonético, nessa posição. É relevante observar que também a posição do segmento na palavra parece um fator que decide a pauta gestual a ser empregada, já que apenas uma variante, dentre todas as observadas, nos dados dos dois informantes, ocorre nessa posição.

Uma última consideração acerca da sugestão de representação, exposta neste capítulo: não se procura, em momento algum, sustentar a tese de que inexitem categorias, na fala dos dois informantes, cujos dados são analisados nesta tese (AJS e JLM), ou mesmo nos dados dos informantes que inicialmente levantaram a suspeita da existência de um contínuo físico, envolvendo a variabilidade de pronúncia de /r/, ARM e OGF. Claro está que existem variantes vibrantes, fricativas e/ou *taps*, na fala de todos eles. O fato, porém, é que, além daqueles cuja categorização é possível, existem “n” variantes entre uma categoria e outra de /r/, as quais se realizam ao longo de um contínuo físico. É isto que fonologias que adotam primitivos estáticos não conseguem prever e é isso, por outro lado, que modelos dinâmicos, como os expostos acima, permitem representar.

Ainda é preciso que se diga – à luz da tese da comensurabilidade – que as categorias vibrantes, fricativas e *taps* se estabelecem nas bordas dos gestos que representam os róticos, i.e., nos limites temporais desses gestos. Assim, vibrantes teriam bordas assíncronicas para os gestos

da região faríngea e da região glotal, enquanto fricativas têm bordas sincrônicas para esses mesmos gestos. Como sugere Albano, a percepção talvez tenha um papel importante, nesse estabelecimento das categorias. Vai ao encontro dessa hipótese o fato de que muitos falantes não conseguem perceber as variantes que se estabelecem, ao longo do contínuo, mas apenas aquelas que jazem em suas bordas, i.e., as categorias, propriamente. Um estudo que contemple o aspecto perceptual dessa variabilidade de pronúncia, contrapondo-o às informações acústicas, talvez traga alguma contribuição no sentido de prever o papel da percepção para o estabelecimento das categorias fônicas. Mas isto é matéria para pesquisas futuras.

Conclusão

Esta tese objetivou chamar a atenção para dois fatos, especialmente: 1) a alofonia dos róticos iniciais de palavra, em PB, não é categórica, mas gradiente; 2) modelos dinâmicos de produção da fala podem dar conta da representação desses fatos.

Os dados, apresentados no terceiro capítulo, sugerem que, intermediando duas categorias entre as quais varia a pronúncia de /r/ para os sujeitos do experimento fonético-acústico – i.e., vibrante apical e fricativa velar, para um deles, e vibrante apical e *tap*, para o outro – há um contínuo físico, ao longo do qual é possível encontrar “n” variantes que se caracterizam por carregarem, concomitantemente, características das duas categorias. É como se houvesse uma linha contínua, em cujas bordas ocorressem dois pontos discretos.

Pode-se também inferir dos dados um condicionamento da estrutura prosódica: isto fica claro para os dados do informante AJS, pois as versões fortalecidas do rótico tendem a ocorrer adjacentes a fronteiras prosódicas fortes, e suas versões enfraquecidas ocorrem adjacentes a fronteiras prosódicas fracas. Ora, as forças de fronteira são manipuladas pelo estabelecimento de diferentes relações sintáticas, no interior de um enunciado. Esse fato requer, em decorrência, que os parâmetros fonéticos quantitativos sejam representados no léxico da língua e controlados por parâmetros gramaticais qualitativos.

A questão, por conseguinte, é como representar os fatos fonéticos gradientes adequadamente ou, dito de outro modo, como “traduzir” o dinâmico no simbólico. Conforme observado no segundo capítulo, não é possível que essa representação seja feita por um modelo fonológico que toma como primitivos de análise traços distintivos ou segmentos. Tal impossibilidade decorre, como eu frisava nesse capítulo, da própria natureza dessas unidades: os traços ou os segmentos têm natureza simbólica e, portanto, não conseguem capturar a “linha que existe entre dois pontos”, digamos assim.

Qualquer que seja a perspectiva do modelo, i.e., considere ele que a vibrante é a versão geminada de dois *taps* ou considere ele a introdução de um traço específico para lidar com os róticos, na estrutura interna de um segmento, a “tradução” do numérico no simbólico não se faz. Isto porque, qualquer que seja o traço, [vibrante] ou [contínuo], e.g., ele só conseguirá representar os pontos discretos. Além disso, quer se considerem traços monovalentes, binários ou escalares, continua sendo impossível a um modelo que toma

primitivos estáticos captar a gradiência das alofonias contínuas. No caso dos traços escalares, particularmente, por mais que pareçam capazes de fazer essa ponte entre o dinâmico e o simbólico, eles também falham na tarefa, já que qualquer escala pode ser arbitrária.

Assim, a solução é adotar um modelo dinâmico de produção da fala, como a Fonologia Articulatória (Browman & Goldstein, 1986, 1989, 1990, 1992), para dar conta da representação dos fatos contínuos. Tais modelos são particularmente eficazes no tratamento das alofonias contínuas – quaisquer delas e não apenas as dos róticos –, pois tomam como primitivos de análise os gestos articulatórios. Sendo definidos a partir de uma equação dinâmica, os gestos podem representar as “n” variantes existentes entre os pontos discretos, a partir de mudanças na especificação dos parâmetros dessa equação, como, por exemplo, o parâmetro “rigidez”.

Embora se adotem aqui esses pressupostos da FAR, “comprá-la” como um todo não seria suficiente para modelar os róticos iniciais do PB. Primeiro porque a proposta de representação dos róticos de Browman & Goldstein não considera uma gama muito vasta de variantes possíveis para /r/. Como menciono no quarto capítulo, o modelo considera apenas as vibrantes, deixando de capturar a identidade de classe entre elas e outros róticos, como, por exemplo, as aspiradas, para as quais sequer há uma proposta de representação. Outro problema de se adotar exclusivamente a FAR para a representação dos róticos é que o modelo parte de dados articulatórios, e não dispomos de meios, no momento, para fazer estudos articulatórios.

Por isso é que a proposta de representação dos róticos do PB – apresentada no quarto capítulo – adota também os pressupostos da Fonologia Acústico-Articulatória (Albano, 2001). A definição dos gestos articulatórios, em termos de espaços acústicos (e não mais de regiões articulatórias), parece permitir uma representação mais acurada dos róticos e a proposta de lexicalização das variantes possibilita explicar a coocorrência de variantes vibrantes e fricativas, na fala de um mesmo indivíduo. É possível pensar que, diante de uma determinada estrutura sintático-prosódica, o falante “escolhe” a variante mais apropriada àquela estrutura gramatical. Note-se que essa proposta também permite uma “tradução” direta do dinâmico no simbólico, sem a necessidade de se postularem inúmeras regras transformacionais que derivem “n” variantes, a partir de um único segmento.

A proposta de representação gestual para os róticos do PB incorpora, ainda, achados de Sproat & Fujimura (1993), para as variantes de /l/ no inglês norte-americano, e de Gick (1999), para as variantes ambissilábicas de /j/, /w/, /r/ e /l/, também nessa língua. Os achados concernem à existência de dois gestos articulatórios que constituem tais segmentos. No caso de nossa proposta, preconiza-se a ativação dos gestos de mais de uma região articulatória, na tentativa de, com essa medida, conseguir capturar a identidade de classe entre as diversas variantes de /r/ no PB.

Embora a proposta – como mencionado – realize a tradução direta do numérico no simbólico e embora ela capte a identidade de classe entre as variantes de /r/ no PB de maneira mais parcimoniosa que modelos que adotam unidades de análise estáticas, restam ainda algumas questões que, na verdade, são questões subjacentes à própria Fonologia Articulatória.

Uma delas é a seguinte: eu afirmo que o estudo conduzido na presente tese aponta para a existência de um contínuo físico entre duas unidades discretas. Afirmo também que a equação dinâmica que define o gesto articulatório dá conta de engendrar esse contínuo, através da manipulação dos parâmetros dinâmicos. Como, porém, estabelecer os pontos discretos nas extremidades do contínuo? A FAR não o resolve. Uma solução plausível é assumir, como Albano (2001), que a percepção desempenhe papel fundamental no estabelecimento desses pontos. Contudo, a maneira como a percepção atua, especialmente no casos das variantes de /r/, precisa ser investigada.

A FAR também não dá conta de fazer a ponte entre o nível segmental e o nível prosódico. De fato, a preocupação do modelo, até sua apresentação mais recente, de 1992, volta-se ao nível segmental. Browman & Goldstein mencionam, *en passant*, a possível existência de uma camada rítmica, que organizaria os padrões gestuais. Essa camada, porém, funcionaria de modo muito semelhante à camada temporal de modelos como a Fonologia Autossegmental e a Geometria de Traços. Também para a Fonologia Acústico-Articulatória a interação entre o nível segmental e o prosódico não é muito clara. Talvez no momento em que se conseguir fazer essa ponte, fique mais claro o processo de “seleção” de variantes, como no caso do informante JLM.

Por fim, cabe mencionar que ainda há muito a ser investigado sobre o comportamento dos róticos no PB, não só em outras posições silábicas, mas também de

outras variantes. Como menciono na nota 1 da Introdução, privilégio, no presente estudo, os róticos iniciais, por considerar a alofonia contínua deles um fato mais intrigante do que a alofonia dos róticos finais, dado que se espera variação de pronúncia de /r/ em final de sílaba/palavra, por ser esta uma posição mais fraca na palavra, e, portanto, favorecedora de tais processos. A posição inicial, porém, por favorecer uma articulação mais cuidada dos segmentos, não seria, *a priori*, um lugar onde se esperasse encontrar enfraquecimento de segmentos. De qualquer modo, é preciso verificar também como se realiza o /r/ de final de sílaba, i.e., se, também nessa posição, há uma espécie de contínuo físico, tanto nos casos dos dialetos ao norte de São Paulo, que alternam variantes fricativas velares e glotais, como nos casos do dialeto do interior de São Paulo e norte do Paraná, que alternam *taps* com variantes retroflexas, em posição final. Cabe igualmente investigar ainda o chamado “zero fonético” de final de palavra, para observar se, de fato, se tem o apagamento de um segmento ou se há algum vestígio do gesto dele, talvez na vogal antecedente.

De forma análoga ao que se fez, cabe investigar se a “variação alofônica” de /r/, em final de sílaba, é sensível à estrutura prosódica do enunciado, e de que maneira. Caso haja um contínuo físico sobre o qual se estende a variabilidade de pronúncia de /r/ final e caso esse contínuo seja sensível à estrutura prosódica, teremos novos exemplos de alofonias contínuas, somando-se aos outros vários, já descritos pela literatura fonética, e que só podem ser representados num modelo dinâmico de produção de fala.

As laterais também merecem investigação: similarmente aos róticos, são segmentos muito pouco estudados – seja na literatura fonético-fonológica geral, seja na literatura fonético-fonológica do português. De fato, como laterais e róticos são segmentos peculiares, dado que não são nem consoantes nem vogais típicas, a maioria dos modelos teóricos acaba por não se deter neles, o que resulta ou em modelos que tratam de modo insuficiente desses segmentos, ou em modelos que simplesmente não tratam deles. Estudos de róticos e laterais, podem, então, trazer contribuições importantes a modelos teóricos, testando-os e, até, fornecendo informações para melhorá-los.

Referências bibliográficas

- ALBANO, E.C. *Da fala à linguagem: tocando de ouvido*. São Paulo: Martins Fontes, 1990.
- ALBANO, E.C. *O gesto e suas bordas: esboço de Fonologia Acústico-Articulatória do português brasileiro*. Campinas: Mercado de Letras/ALB/FAPESP, 2001.
- ALBANO, E.C.; BARBOSA, P.; GAMA-ROSSI, A.; MADUREIRA, S. & SILVA, A. A interface fonética-fonologia e a interação prosódia-segmentos, in *Estudos Lingüísticos XXVII*, UNICAMP: p. 135-143, 1998.
- ALLEN, W.S. *Vox Graeca - a guide to pronunciation of classical Greek*. Cambridge: The Cambridge University Press, 1968.
- ALMEIDA, M. & DORTA, J. Datos acústicos de las líquidas españolas, in *En homenaje a José Pérez Vidal*. Edición al cuidado de Carmen Díaz Alayón. La Laguna: Tenerife: 214-217, 1993.
- BROWMAN, C. & GOLDSTEIN, L. Towards an Articulatory Phonology, in *Phonology Yearbook*, 3:219-252, 1986.
- Articulatory gestures as phonological units, in *Phonology Yearbook*, 6: 201-251, 1989.
- Tiers in Articulatory Phonology, in J. Kingston & M. Beckman, *Papers in Laboratory Phonology I: between the grammar and physics of speech*. Cambridge: The Cambridge University Press, pp: 341-376, 1990.
- Articulatory Phonology: an overview, in *Phonetica*, 49: 155-180, 1992.
- CALLOU, D.M.I. *Variação e distribuição da vibrante na fala urbana culta do Rio de Janeiro*. Rio de Janeiro: PROED/UFRJ, 1987.
- CALLOU, D.M.I.; MORAES, J.A. & LEITE, Y. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: Koch, I.V. (org.). *Gramática do Português Falado, vol. VI*. Campinas: Editora da UNICAMP, 1997.
- CÂMARA JR., J.M. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Organização Simões, 1953.
- _____ *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Padrão – Livraria Editora, 1977.

- CHOMSKY, N. & HALLE, M. *The sound pattern of English*. Cambridge: The MIT Press, 1995 [1968].
- CLEMENTS, N. & HUME, E. The internal organization of speech sounds, in J. Goldsmith (org.) *The handbook of phonological theory*. Cambridge: Blackwell, 1995.
- D'ANGELIS, W. R. *Traços de modo e modos de traçar geometrias: línguas Macro-Jê e teoria fonológica*. Tese de doutorado, inédita. Campinas: LAFAPE/IEL/UNICAMP, 1998.
- FERREIRA, A.B.H. *Novo dicionário da língua portuguesa*. Rio de Janeiro: Editora Nova Fronteira. 2^a. edição, 39^a. impressão, 1986.
- *Minidicionário Aurélio*. Rio de Janeiro: Editora Nova Fronteira, 1977.
- FOUGERON, C. Prosodically conditioned articulatory variations: a review, in *UCLA Working Papers in Phonetics*, 97: 1 – 73, 1999.
- FOX, R. & TERBEEK, D. Dental flaps, vowel duration and the rule ordering in American English, in *Journal of Phonetics*, 5: 27-34, 1977.
- GAMA-ROSSI, A. *Relações entre o desenvolvimento lingüístico e neuromotor: a aquisição da duração no português brasileiro*. Tese de doutorado, inédita. Campinas: LAFAPE/IEL/UNICAMP, 1999.
- GICK, B. *The articulatory basis of syllable structure: a study of English glides and liquids*. Tese de doutorado, inédita. Yale University, 1999.
- GUIMARÃES, M. *Repensando a interface sintaxe-fonologia a partir do Axioma de Correspondência Linear*. Dissertação de Mestrado, inédita. Campinas: IEL-UNICAMP, 1998.
- Handbook of the International Phonetic Society*. Cambridge: The Cambridge University Press, 1999.
- HARRIS, J.W. *Syllable structure and stress in Spanish: a nonlinear analysis*. Cambridge: The MIT Press, 1983.
- HERNANDORENA, C.L.M. Introdução à teoria fonológica, in L. Bisol (org.) *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1996.
- HUBER, J. *Gramática do português antigo*. Lisboa: Fundação Calouste Gulbenkian, 1986.

- JAKOBSON, R.; FANT, G. & HALLE, M. *Preliminaries to speech analysis: the distinctive features and their correlates*. Cambridge: The MIT Press, 1976, 11^a. edição.
- JAKOBSON, R. & HALLE, M. *Fundamentals of Language*. The Hague: Mouton, 1956.
- KATAMBA, F. *An introduction to Phonology*. New York: Longman, 1989.
- KEATING, P. Universal phonetics and the organization of grammars. In: Fromkin, V. (ed.) *Phonetic Linguistics: essays in honor of Peter Ladefoged*. New York: Academic Press, 115-132, 1985.
- KEATING, P.; CHO, T.; FOUGERON, C. & HSU, C. Domain-initial articulatory strengthening in four languages, in *UCLA Working Papers in Phonetics*, 97: 139- 151, 1999.
- LADEFOGED, P. *A course in phonetics*. New York: Harcourt Brace Jovanovich, 1975.
- LADEFOGED, P. & MADDIESON, I. *The sounds of the world's languages*. Cambridge: Blackwell, 1986.
- LEHISTE, I. *Acoustical characteristics of selected English consonants*. The Hague: Mouton, 1964.
- LINDAU, M. The story of /r/, in V. Fromkin (ed.) *Phonetic Linguistics: essays in honor of Peter Ladefoged*. New York: Academic Press, 157-168, 1985.
- LINDBLUM, B. Spectrographic study of vowel reduction, in *Journal of the Acoustic Society of America*, 35: 1773-1781, 1963.
- LINDBLUM, B. Models of phonetic variation and selection, in *Phonetic Experimental Research in the Institute of Linguistics, University of Stockholm (PERILUS)*, 11: 65-100, 1990.
- LLORACH, E.M. *Fonología española*. Madrid: Editorial Gredos. 4^a. edição, 1976.
- LLORET, M.R. Sonorant dissimilation in the Iberian Languages, in F. Martínez-Gil & A. Morales-Font (eds.) *Issues in the phonology and morphology of the major Iberian languages*. Washington: Georgetown University Press, 1997.
- MADDIESON, I. A survey of liquids, in *UCLA Working Papers in Phonetics*, 50: 93-112, 1980.
- MAIA, E.A.M. Evidência portuguesa para uma nova versão do traço rotacizado. Trabalho apresentado no Encontro Nacional de Linguística, inédito, 1982.

- MASCARÓ, J. & BONET, E. On the representation of contrasting rhotics, in F. Martínez-Gil & A. Morales-Front (eds.) *Issues in the phonology and morphology of the major Iberian languages*. Washington: Georgetown University Press, 1997.
- MONARETTO, V.N.O. *Um reestudo da vibrante: análise variacionista e fonológica*. Tese de doutorado. PUC-RS, inédita, 1997.
- NESPOR, M & VOGEL, I. *Prosodic Phonology*. Foris: Dordrecht, 1986.
- OBEDIENTE, E. *Fonética y fonología*. Mérida: Facultad de Humanidades y Educación/ Consejo de Publicaciones de la Universidad de los Andes. 3^a. edição, 1998.
- OHALA, J. There is no interface between phonetics and phonology: a personal view, in *Journal of Phonetics*, 18: 153-171, 1990.
- PIKE, K.L. *Phonemics – a technique for reducing languages to writing*. Ann Harbor: The University of Michigan Press, 1947.
- QUILIS, A. *La lengua española en cuatro mundos*. Madrid: Mapfre, 1992.
- RECASENS, D. On the production characteristics of apicoalveolar taps and trills, in *Journal of Phonetics*, 19: 267-280, 1991.
- RECASENS, D. Predicting directionality trends in segmental assimilations and insertions, in J. Ohala *et alii* (eds.) *Proceedings of the XIVth. International Congress of Phonetic Sciences: 1847-1850*, 1999.
- RICE, K. & AVERY, P. On the relationship between laterality and coronality, in C. PARADIS & J.F. PRUNET (eds.) *Phonetics and phonology: the special status of coronals – internal and external evidence*. San Diego: Academic Press, 1991.
- SAGEY, E. *The representation of features and relations in non-linear phonology*. Tese de doutorado, inédita, MIT, 1996.
- SAPIR, E. A realidade psicológica do fonema, in Dascal, M. (org.) *Fundamentos metodológicos da lingüística – volume 2: Fonologia e Sintaxe*. Campinas: IEL, 1981 [1933].
- SAUSSURE, F. *Curso de Lingüística Geral*. São Paulo: Cultrix/EDUSP, 1969.
- SELKIRK, E. *Phonology and Syntax: the relation between sound and structure*. Cambridge, the MIT Press, 1984.

- SHADDLE, C. & MAIR, S.J. Quantifying spectral characteristics of fricatives, in *Proceedings of the International Conference on Spoken Language Processing*, vol. 3: 1521-1524, 1996.
- SILVA, A.H.P. *Para a descrição fonético-acústica das líquidas no português brasileiro: dados de um informante paulistano*. Dissertação de mestrado, inédita. LAFAPE-IEL-UNICAMP, 1996.
- ___ Pistas para o condicionamento prosódico sobre a variabilidade de produção de /r/, in *Estudos Lingüísticos*: 28: 682-688, 1999a.
- Caracterização acústica de [r], [r̥], [l] e [ʎ] nos dados de um informante paulistano, in *Cadernos de Estudos Lingüísticos*, 37: 51-68, 1999b.
- SILVA, A. & ALBANO, E.C. Brazilian Portuguese rhotics and the phonetics/phonology boundary, in J. Ohala *et alii* (eds.) *Proceedings of the XIVth. International Congress of Phonetic Sciences*, vol. 3: 2211- 2214, 1999.
- SILVA, T.C. *Fonética e Fonologia*. São Paulo: Contexto, 2000.
- SPENCER, A. Eliminating the feature [lateral], in *Journal of Linguistics*, 20: 23-43, 1984.
- SPROAT, R. & FUJIMURA, O. Allophonic variation in English /l/ and its implications for phonetic implementation, in *Journal of Phonetics*. 21:291-311, 1993.
- TARALLO, F. *A pesquisa sócio-lingüística*. São Paulo: Ática, 1986.
- TRUBETZKOY, N. *Principles of Phonology*. Trad.Christiane A. M. Baltaxe. Berkeley/Los Angeles: University of California Press, 1969.

ANEXO
CORPUS DO EXPERIMENTO FONÉTICO-ACÚSTICO

1. Por que você não quis fazer a propaganda da campanha de vacinação com a criança que tínhamos selecionado?

Eu não achei a criança risonha. Eu achei ela mal-humorada demais, por isso ela não ia ficar legal nos outdoors da campanha.

2. Você achou as duas crianças que você tava procurando pro comercial de leite?

Eu não achei a criança risonha. Eu só achei a mal-humorada até agora.

3. No ano passado eu tive muitos problemas com aquela menina da primeira série. Você também não achou ela uma criança super difícil de lidar?

Eu não achei a menina rebelde. Ela era um pouco agitada, mas fazia tudo o que eu pedia.

4. Você achou todas as crianças daquela classe? Precisamos reunir todas elas, logo depois do recreio, pra passar os avisos sobre a festa.

Eu não achei a menina rebelde. Parece que ela fugiu de novo da escola.

5. Como você sabe que a empregada consertou aquela sua camisa velha?

Eu não encontrei a camisa rasgada. Eu encontrei ela remendada quando fui procurar ela ontem.

6. Você encontrou as duas camisas velhas que você tava procurando pra usar como pano de chão?

Eu não encontrei a camisa rasgada. Só a remendada tava na gaveta.

7. Aquela funcionária da biblioteca te atendeu mal?

Eu não achei a funcionária ríspida. Ela até que foi bem delicada comigo.

8. Você achou as duas funcionárias que você tava procurando?

Eu não achei a funcionária ríspida. Só a funcionária delicada fica lá na hora do almoço.

9. O nosso time tá indo tão bem no campeonato! Por que você tá pensando em dispensar a técnica justo agora?

Eu não acho a técnica rígida. Ela deixa as jogadoras soltas demais e elas precisam de disciplina na fase final do campeonato.

10. Nosso time tá meio perdido. Por que você ainda não contratou uma das duas técnicas que entrevistamos?

Eu não acho a técnica rígida. Ela sumiu! Eu só acho a outra, que é meio molenga, e o nosso time precisa de alguém de fibra pra melhorar!

11. Por que você não quis dançar aquela música comigo?

Eu achei a música rápida. Pra dançar a dois, ela tinha de ser mais lenta.

12. E então, você não vai me mostrar aquela balada que você sugeriu pra trilha sonora do filme? Eu tô curioso.

Só achei a música rápida. Aquela música lenta, eu não tô achando a fita agora.

13. O Paulo achou a estória de amor da novela das oito super comovente. E você, o que você achou?

Eu não achei a novela romântica. Eu achei ela tão brega que parecia até novela mexicana.

14. E aí, você já achou as duas novelas do Cervantes que você tava procurando pra fazer o seu trabalho de literatura?

Eu não achei a novela romântica. Eu só achei a satírica até agora, mas eu ainda vou procurar na biblioteca pública.

15. Por que você não comprou aquela cômoda que a gente viu na loja? Ela ia combinar tão bem com os móveis do seu quarto!

Eu não achei a cômoda rústica. Eu achei ela sofisticada demais pra combinar com os outros móveis.

16. Por que nosso cliente ainda não recebeu a cômoda que ele comprou na semana passada?

Eu não achei a cômoda rústica. No estoque só tem uma outra, mais sofisticada. Mas essa não deve ser a que o cliente comprou.

DISTRADORES

17. Você acha que vale a pena indicar a leitura daquele romance aos alunos da sétima série? Ele parece tão infantil que a leitura pode se tornar monótona.

Eu não acho o romance infantil. Eu acho que a trama dele é até sofisticada para alunos dessa idade.

18. E então, a biblioteca tinha os dois romances que você queria?

Eu não achei o romance infantil. Do autor que eu procuro a biblioteca só tem o romance regional.

19. Por que você demorou pra reconhecer o filho da Maria?

Eu não encontrei o menino obeso. Eu encontrei ele tão magro que parecia até doente.

20. Os dois filhos da Maria estavam na festinha de aniversário?

Eu não encontrei o menino obeso. Eu só encontrei o outro, o magrinho.

21. E aí, conseguiu chegar rápido no Ibirapuera, apesar daquele caminho congestionado que te indicaram?

Eu não achei aquele caminho movimentado. Ele até que era calmo pra São Paulo.

22. Por que você não foi visitar a chácara da nossa amiga ontem, como tínhamos combinado?

Eu não achei aquele caminho movimentado. Eu só achei um outro, muito calmo, que pelo jeito não era o que você tinha me falado. Aí eu me perdi!

23. Por que o sindicalista voltou do Congresso sem fazer acordo com o deputado que o recebeu?

Ele não achou o deputado idôneo. Ele achou o deputado tão cara de pau que resolveu tentar outros contatos.

24. Os dois políticos que o relator da comissão de sindicância procurava para fazer a acareação estavam no Congresso?

Ele não achou o deputado idôneo. Ele só achou o corrupto, por isso a acareação teve de ficar para outro dia.

25. Por que você não se importa com os comentários que aquele aluno faz sobre seu curso?

Eu não considero o aluno desleixado. Eu só considero o aplicado.

26. Você não acha que aquele aluno do primeiro ano poderia se esforçar mais?

Eu não considero o aluno desleixado. Eu considero ele um dos mais aplicados da turma.

27. Você nem bem começou a jogar “Paciência” e já desistiu. Por quê?

Eu não achei o jogo fácil. Eu achei ele complicado demais.

28. O técnico em informática instalou no seu micro todos os jogos que você queria?

Eu não achei o jogo fácil. Eu só achei outros, muito complicados. Talvez o técnico não tenha instalado tudo o que eu pedi.

29. Não sei o que você viu naquele cara. Todo mundo achou ele tão arrogante aquela noite, no barzinho!

Eu não achei o cara arrogante. Se você quer saber, ele é um dos caras mais sensíveis que já conheci.

30. Por que você ainda não fez aquele trabalho de psicologia envolvendo a comparação da personalidade de dois sujeitos?

Eu não achei o cara arrogante. Eu só achei o humilde até agora, por isso ainda não posso comparar a personalidade de ambos.

31. Por que você não quis seguir a sugestão do vendedor de pagar a mercadoria em várias prestações?

Eu não achei o vendedor confiável. Eu achei ele metido a espertalhão, e aí eu podia me dar mal.

32. E então, você pediu ajuda pro vendedor pra conferir a contabilidade da loja, como o dono pediu?

Eu não achei o vendedor confiável. Eu só achei o desonesto, por isso resolvi conferir a contabilidade outro dia.

33. Parece que você não foi muito com a cara do antigo porteiro do seu prédio, por quê?

Eu não achei o porteiro eficiente. Eu achei ele um dos mais irresponsáveis que já trabalharam no condomínio.

34. Por que você não quis deixar a chave do apartamento na portaria?

Eu não achei o porteiro eficiente. O porteiro que tava na recepção é irresponsável, e por isso não acho muito seguro deixar a chave lá.

35. Como nosso candidato se saiu no debate?

Eu não achei o candidato convincente. Eu achei ele muito inseguro, não demonstrava convicção no que dizia.

36. Por que você não entrevistou os dois reitoráveis? O jornal precisa da entrevista pra edição de domingo.

Eu não achei o candidato convincente. Eu só achei o inseguro, por isso preferi deixar a entrevista pra outro dia.

37. Por que você não continua assinando aquele jornal?

Eu não acho aquele jornal imparcial. Eu acho ele um dos mais tendenciosos que existem por aí.

38. Por que você tá preocupado? O arquivo da biblioteca não tem os dois jornais que você procura?

Eu não acho aquele jornal imparcial. A biblioteca só tem o outro, que é extremamente tendencioso, e eu preciso dos dois pra minha pesquisa.

39. Todos os convidados repetiram o bolo. Por que você não quis repetir também?

Eu não achei aquele bolo gostoso. Ele tava doce demais!

40. Por que você não comprou o bolo pra festa desta noite, como tínhamos combinado?

Eu não achei aquele bolo gostoso. Na confeitaria só tinha um outro bolo, que não parecia muito bom. Por isso decidi conversar com você pra ver o que vamos fazer.

41. Por que você não gravou a fala do primeiro informante que entrevistamos?

Eu não achei o informante ideal. Eu achei ele muito conservador, por isso não iríamos conseguir observar na fala dele o fenômeno que estamos estudando.

42. E aí, por que você ainda não começou a rodar seu experimento?

Eu não achei o informante ideal. Todos os que apareceram até agora não têm o nível de escolaridade que queremos.

43. Por que você não consulta aquela fonte que o professor sugeriu?

Eu não acho aquela fonte atual. Eu acho que ela já está ultrapassada.

44. E então, a biblioteca tinha todas as fontes que você tava procurando?

Eu não achei aquela fonte atual. Eu só encontrei as outras, mais antigas.

45. Por que você não toma mais um pouco de vinho?

Eu não achei o vinho encorpado. Eu achei ele muito fraco e eu não gosto de vinhos fracos.

46. Por que você não pegou lá na adega o vinho que te pedi?

Eu não achei o vinho encorpado. Eu só achei o fraco. O vinho que você quer deve ter acabado.

47. Por que você não comprou o vestido que a gente viu na loja?

Eu não achei a estampa alegre. Eu achei ela muito séria e eu queria um vestido mais informal pra ir na festa.

48. Me disseram que você ainda nem comprou tecido pra fazer o traje dos bailarinos, e o espetáculo estréia no mês que vem. O que tá acontecendo?

Eu não achei a estampa alegre. Eu achei só uma outra, muito séria, e não vai ficar bem.

49. Por que você não contratou o garoto pra ser o dublê do filme?

Eu não achei o garoto ousado. Eu achei ele muito medroso e isso não é bom para quem quer trabalhar como dublê.

50. É verdade que você ainda não conseguiu achar os dois garotos pra fazer o comercial?

Eu não achei o garoto ousado. Eu só achei o medroso até agora e assim não dá pra gravarmos a cena da escalada.

51. Da última vez que conversei com aquele seu amigo ele tava muito contente com o novo emprego. Ele continua contente?

Eu não encontrei o meu amigo alegre. Eu encontrei ele triste, acho que as coisas não tão indo bem na empresa.

52. E aí, você saiu ontem com os dois amigos que queria rever?

Eu não encontrei o meu amigo alegre. Eu só encontrei o triste e aí a noitada foi uma droga.

53. Por que você preferiu comprar à vista ao invés de fazer um crediário?

Eu não considero os juros baixos. Eu considero eles muito altos, não compensa pagar a prestação.

54. Como você faz o cálculo do preço final dos seus produtos, se eles só estarão prontos pra venda dentro de alguns meses?

Eu não considero os juros baixos. Eu só considero os altos, aí garanto sempre algum lucro sobre o preço final.