

PARA A REPRESENTAÇÃO DINÂMICA DO TAP EM GRUPOS E CODAS: EVIDÊNCIAS ACÚSTICAS

Adelaide H. P. Silva¹

Felipe Costa Clemente²

Gustavo Nishida³

adelaide@ufpr.br

tela_letras@hotmail.com

gustavo_linguista@yahoo.com.br

RESUMO: Descrições acústicas do *tap* em grupos e em codas revelam que esse segmento é sempre realizado apoiando-se em um evento acústico de natureza vocálica. Tal elemento, porém, tem natureza variável em função da posição silábica onde o *tap* ocorre: se em grupos, o elemento será a mesma vogal nuclear, entrecortada pelo *tap*; se em codas, o elemento vocálico assume características de vogal neutra. Considerando-se que a natureza do elemento vocálico adjacente ao *tap* é específica da língua, e não universal, é necessário representar tal fato na gramática da língua. Argumentamos que a melhor maneira de se fazer isso é adotando-se modelos dinâmicos de produção da fala, os quais prevêem uma “tradução” direta do nível simbólico no nível numérico devido ao primitivo de análise postulado. Fornecemos, então, uma representação para o *tap* nesses ambientes silábicos, embora observando que o condicionamento do lugar que o *tap* ocupa na sílaba sobre a natureza do elemento vocálico coloca um problema para os modelos dinâmicos, que não prevêem a noção de estrutura silábica.

PALAVRAS-CHAVE: *taps*; grupos; codas; representação dinâmica.

INTRODUÇÃO

Em seu trabalho de descrição acústica das líquidas do português brasileiro (PB), Silva (1996) observa que, diferentemente de vibrantes alveolares e laterais – alveolares

¹ Departamento de Linguística, Letras Clássicas e Vernáculas (DELIN), Universidade Federal do Paraná (UFPR).

² Graduação Universidade Federal do Paraná (UFPR).

³ Graduação Universidade Federal do Paraná (UFPR).

e palatais – o *tap* ocorre sempre intermediário a eventos de natureza vocálica, ainda que em grupos ou em codas finais de palavra. Silva relata que, nesses ambientes, diferentemente da posição intervocálica, o evento acústico de natureza vocálica à esquerda do *tap* – no caso de grupos – e à direita dele – no caso de codas – tem duração muito breve, variando de 10 a 20 milissegundos⁴.

Além disso, baseada em inspeções visuais do sinal acústico, Silva aponta, ali, para uma possível semelhança entre a estrutura formântica do evento à esquerda do *tap* em grupos e a estrutura formântica da vogal nuclear da sílaba. Entretanto, não apresenta uma descrição acústica sistemática do evento vocálico para verificar se a semelhança de fato ocorre.

Mais tarde, Carvalho (2004), em seu estudo comparativo entre vibrantes do PB e do espanhol colombiano, observa, como Silva (*op. Cit.*), a existência do evento acústico de natureza vocálica à esquerda do *tap* em grupos e uma aparente semelhança entre tal evento e a vogal nuclear do grupo. Também Carvalho, porém, baseia-se apenas numa análise visual do sinal acústico, sem realizar as medidas das estruturas formânticas do evento vocálico e da vogal nuclear, comparando-as. Por outro lado, o estudo de Carvalho aponta para fato análogo em espanhol, i.e., para uma semelhança visual entre a estrutura formântica do evento vocálico à esquerda do *tap* em grupos e a vogal nuclear.

Sua descrição, portanto, corrobora o estudo de Silva (1996), no caso do PB, e estudos como os de Quilis (1993), já que esse autor, remetendo a descrições impressionísticas conduzidas, no final do século XIX, por Lenz, nota que haveria uma semelhança entre o evento vocálico e a vogal nuclear de grupos com *tap*.

Por outro lado, ambas as descrições para o PB vão em sentido contrário àquelas disponíveis para o búlgaro ou o norueguês: tanto no caso da primeira língua, foco do estudo de Jetchev (1995), como no caso da segunda, enfocada por Kvale & Foldvik (1995), o *tap* em grupos mostra-se adjacente a uma vogal neutra, do tipo *schwa*.

Tais fatos apontam para preferências distintas, de línguas também distintas, na escolha da natureza do elemento vocálico existente à esquerda do *tap* em grupos. Assim sendo, tem-se, através desses estudos, evidências da característica gramatical dos fatos em questão, o que coloca a necessidade de fornecer-lhes uma representação fonológica.

Como, porém, os estudos relativos ao PB, citados acima, constituem-se de uma descrição visual, é preciso conduzir uma investigação sistemática da natureza do evento

⁴ Observe-se, para efeitos comparativos, que em posição intervocálica as vogais adjacentes ao *tap* duram em média, aproximadamente, 40 milissegundos (ms).

vocálico adjacente ao *tap* em grupos e também em codas⁵, para que se possa, através de uma análise quantitativa, observar se, de fato, tem-se nesses eventos a mesma natureza da vogal nuclear. Serão então apresentadas neste texto, nas duas seções seguintes, análises acústicas do evento vocálico nas duas posições silábicas mencionadas para, a partir delas, tratar de outra questão que apontam tanto as descrições acima mencionadas como as análises por nós conduzidas: como representar, na gramática do PB, o fato observado.

Tal questão, como se procurará mostrar adiante, não é banal, mas de grande importância para os estudos do nível fônico da língua, uma vez que, ao evidenciar o caráter gramatical do detalhe fonético, toca na necessidade de se abandonar as análises fonológicas herdeiras de SPE⁶ e de se adotar uma análise dinâmica para os eventos vocálicos e os *taps*, análise essa que prevê um primitivo distinto daqueles preconizados pela fonologia até então, como os traços distintivos ou os segmentos, para garantir a comensurabilidade entre os níveis simbólico e dinâmico do processamento da fala.

1. ANÁLISE ACÚSTICA DO *TAP* EM GRUPOS DO PB

Devido à falta de um estudo sistemático que realizasse quantificações sobre os elementos vocálicos que são produzidos entre a obstruinte o *tap* em grupos do PB, decidiu-se elaborar um experimento acústico para se averiguar qual é a natureza desse elemento vocálico. Assim, havendo duas possibilidades: o elemento vocálico ser um *schwa*; ou se tratar da própria vogal nuclear do grupo sendo entrecortada pelo *tap*, uma vez que devido ao seu aspecto descontínuo seja necessário um som contínuo para ele se apoiar, o que faz com que “o *tap* sempre ocupe, de qualquer forma, uma posição intervocálica” (Silva, op cit: 69).

5 especialmente neste caso, pois até onde se tem notícia, o *tap* em coda, no PB, só foi descrito, acusticamente, por Silva (1996).

6 *The Sound Pattern of English*, obra em que é exposto o modelo de fonologia gerativa de Chomsky & Halle. Tal modelo, não é exagerado afirmar, influencia todos os seus sucedâneos, mesmo os que, de alguma maneira, o criticam.

1.1 METODOLOGIA

1.1.1 DESIGN EXPERIMENTAL

Com o auxílio do *software* Listas⁷, decidiu-se utilizar como palavras-alvo dissílabos paroxítonos, tendo o grupo localizado na sílaba tônica. Os grupos continham as consoantes [p, b, t, d, k, g, f] e as vogais alternadas, sempre que possível, entre as 7 orais do PB.

Para não enviesar a coleta de dados, decidiu-se por inserir as palavras-alvo em sentenças de uso cotidiano, obedecendo a uma estrutura sintática direta, localizando os grupos em posição de objeto direto, não em posição de sujeito, a fim de se evitar efeitos de topicalização⁸. Não foram utilizados distratores, pois o *corpus* teria um grande número de sentenças, o que poderia cansar os informantes e enviesar a coleta de dados. Vale dizer que os informantes não notaram o que se estava investigando, dando maior naturalidade ao experimento.

As 37 sentenças foram lidas 5 vezes cada, em ordem aleatória, por dois informantes naturais de Curitiba (informante D e M), totalizando 370 sentenças. As gravações foram realizadas em sala com tratamento acústico, no LACOMUS/UFPR, com taxa de amostragem do sinal a 44100 Hz.

Como será visto adiante, um novo experimento teve que ser elaborado para dar conta do comportamento dos elementos vocálicos em grupos compostos por fricativas. Assim, um procedimento teve que ser alterado na metodologia. Decidiu-se inserir as palavras-alvo com fricativas em sentenças-veículo do tipo “*Digo ___ baixinho*”, para que se tivesse maior controle da estrutura prosódica das sentenças, isolando com mais rigor o segmento a ser analisado.

⁷ Programa desenvolvido no LAFAPE/UNICAMP que possibilita realizar quantificações das estruturas de sílabas e palavras da língua.

⁸ Algumas das sentenças utilizadas no experimento: “*Minha mãe guardou o prato no armário.*”; “*O brasileiro venceu o grego na corrida*”; “*A polícia renovou a frota de viaturas.*”; “*Coloquei a fronha na máquina de lavar.*”; “*Minha mãe coloca cravo no arroz doce.*”.

	[i]	[e]	[E]	[a]	[O]	[o]	[u]
[p]	primo	preso	pressa	prato	prova	-	-
[b]	briga	-	breque	braço	broche	broto	bruxa
[t]	tribo	trecho	treco	trave	trote	troco	truque
[d]	-	-	-	drama	droga	-	-
[k]	crime	creme	crepe	cravo	-	-	-
[g]	grito	grego	greve	grade	-	grosso	grupo
[f]	fritas	frevo	frete	frase	frota	fronha	fruta

Tabela 1: Dissílabos paroxítonos utilizados nas sentenças do experimento⁹.

Os *taps* foram identificados nos espectrogramas auditiva e visualmente. Articulatoriamente, os *taps* são caracterizados como sendo um “som em que um breve contato entre os articuladores é feito pelo movimento de um articulador ativo diretamente ao céu da boca” (Ladefoged & Maddieson, 1996: 231), o que resulta em espaços em branco no espectrograma, já que há a total interrupção da passagem do ar pelo trato vocal, e na diminuição da energia de produção (caracterizada pela diminuição na amplitude da onda).

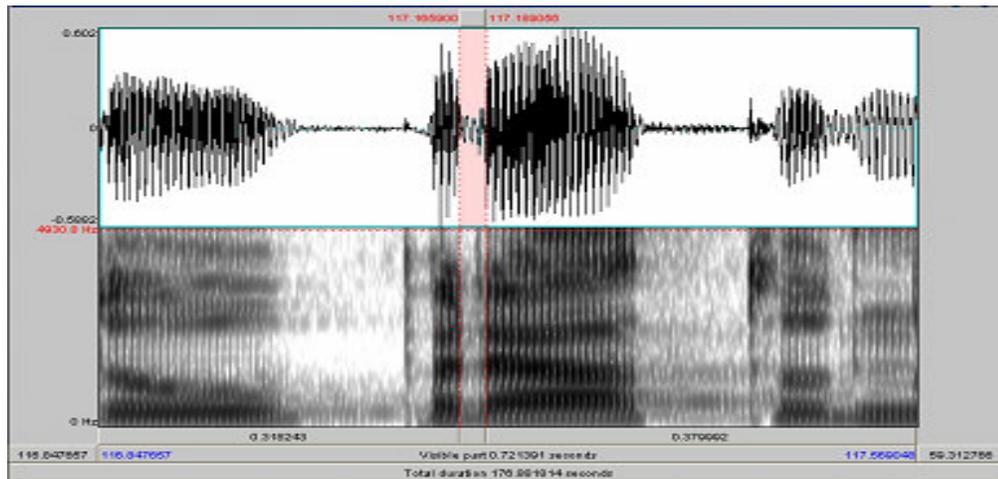


Figura 1: Espectrograma mostrando, entre linhas verticais, um *tap* na palavra-alvo “prato”. Pode-se notar, a sua esquerda, a produção de um segmento de aspecto contínuo. Esse segmento é o elemento vocálico.

No espectrograma da Figura 1, acima, é possível notar que entre a oclusiva [p] e o *tap* (entre linhas verticais pontilhadas) há a produção de um segmento de aspecto contínuo e com estrutura formântica que parece seguir ao da vogal nuclear [a]. Esse segmento é o elemento vocálico. Como se pode perceber visualmente, sua duração é inferior ao da vogal nuclear do grupo, o que talvez faça com que não seja percebido auditivamente.

⁹ Para as vogais anteriores e posteriores semi-abertas utilizaram-se, respectivamente, [E] e [O].

Os dados coletados foram analisados no software Praat¹⁰. Extraíram-se os três primeiros formantes do elemento vocálico e da vogal nuclear, em seus respectivos pontos médios, para se caracterizar a natureza do elemento vocálico. As medidas foram feitas com o auxílio da extração automática do Praat e conferidas manualmente para se evitar possíveis erros.

1.2 DISCUSSÃO DOS DADOS

Tendo coletado e analisado os dados, decidiu-se plotar os valores de F1 e F2 em gráficos, com a finalidade de se comparar o comportamento das vogais nucleares do grupo com a do elemento vocálico.

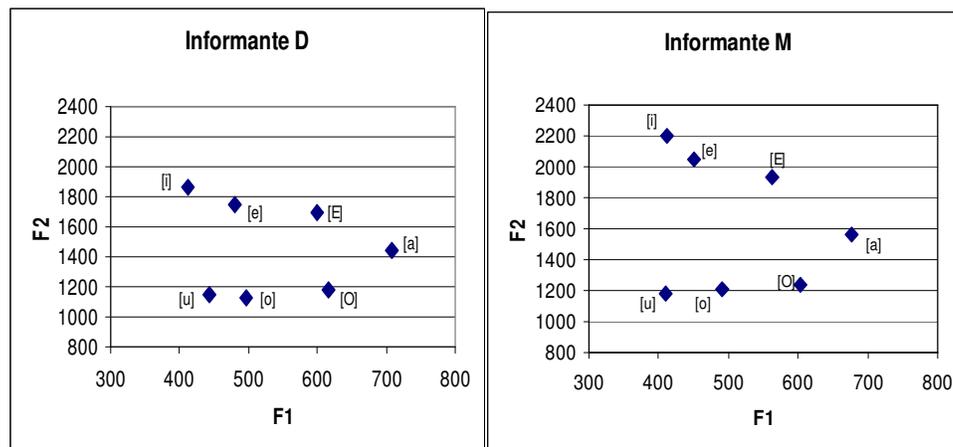


Gráfico 1: Médias de F1 e F2 das vogais nucleares dos grupos dos informantes D e M.

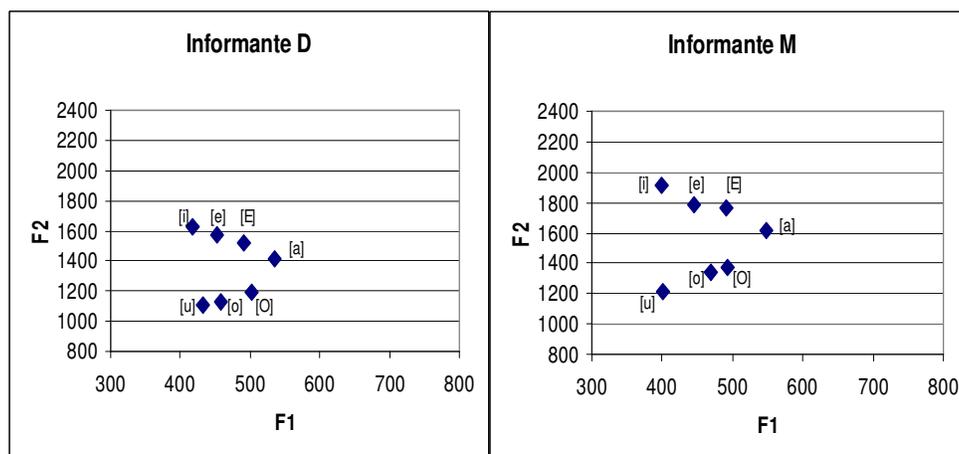


Gráfico 2: Média de F1 e F2 dos elementos vocálicos em grupos com oclusivas, informantes D e M.

¹⁰ Programa gratuito desenvolvido por Paul Boersma e David Weenik no Instituto de Ciências Fonéticas da Universidade de Amsterdã. Disponível em www.praat.org.

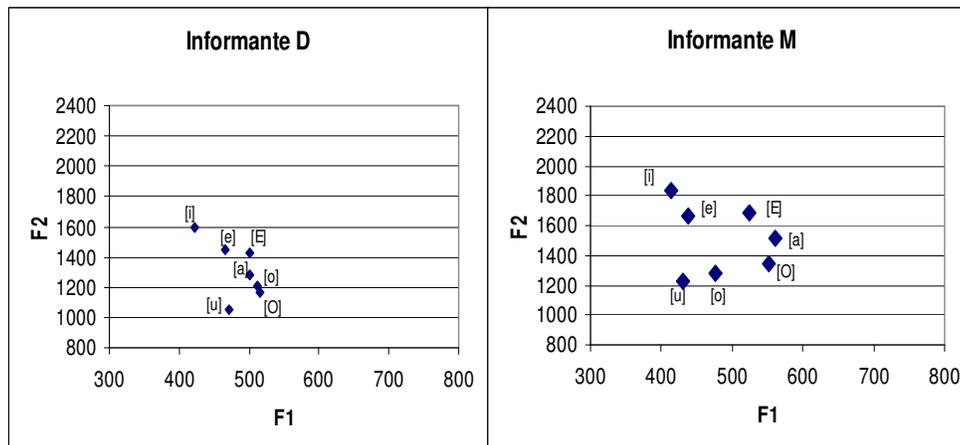


Gráfico 3: Média de F1 e F2 dos elementos vocálicos em grupos com fricativas, informantes D e M.

No gráfico 1, acima, tem-se as médias dos valores de F1 e F2 das vogais nucleares dos grupos em contexto de oclusivas e de fricativas, produzidos pelos informantes **D** e **M**. Como se pode notar, os valores se dispõem de modo semelhante às sete vogais orais do PB no triângulo vocálico. Nota-se que os triângulos dos dois informantes não são idênticos, ou seja, não possuem os mesmos valores para cada vogal. Isso era de se esperar uma vez que, como uma mesma vogal não é produzida por um mesmo informante da mesma maneira, não se obtêm os mesmos valores na fala do próprio indivíduo, muito menos entre indivíduos. O que vai determinar as diferenças entre as vogais são as relações entre os formantes. Para vogais anteriores se tem F1 baixo e F2 alto, ou seja, distantes entre si, e para posteriores F1 e F2 baixos, mais próximos.

O gráfico 2, acima, corresponde às médias de F1 e F2 dos elementos vocálicos em grupos constituídos por “oclusiva + /r/ + vogal”. Como se pode notar, os pontos nos gráficos formam, como se tem para as vogais nucleares dos grupos, um triângulo que se assemelha ao triângulo vocálico das vogais do PB. É interessante notar que esses gráficos apresentam a mesma característica dos gráficos das vogais nucleares: o informante **M** utiliza maior ação de dorso de língua para fazer distinções entre os elementos vocálicos, o que resulta em um gráfico com maior área. Devido a essa semelhança entre os Gráficos 1 e 2 pode-se dizer que os elementos vocálicos dos grupos em que há oclusivas têm a estrutura formântica condicionada pela qualidade da vogal nuclear do grupo.

No entanto, os dados para as fricativas parecem apontar para um outro resultado. O Gráfico 3, acima, corresponde às médias de F1 e F2 dos elementos vocálicos em

grupos contendo “fricativa + /r/ + vogal”. Pode-se notar que, para o informante **M**, o triângulo dos elementos vocálicos parece assumir a mesma configuração das vogais nucleares e dos grupos compostos por oclusivas, o que levaria a pensar que o elemento vocálico tem sua configuração de formantes condicionada pela qualidade da vogal nuclear. O gráfico do informante **D** parece assumir configuração distinta. Embora seja possível perceber uma nítida diferença entre os elementos vocálicos que correspondem às vogais nucleares [i], [e] e [E], os dados para as vogais posteriores parecem se centralizar. Pode-se, então, pensar que em grupos contendo “fricativa + /r/ + vogal” produzem um elemento vocálico que se assemelha a uma vogal neutra do tipo *schwa* (como se tem para o norueguês, búlgaro e tcheco). Também pode-se levantar a hipótese de que esses gráficos traduzem idiossincrasias dos falantes (especialmente em ambiente de fricativa), já que apenas se analisou dois informantes até aqui. Essas diferenças, em um conjunto de dados maior, se dissolveriam. Um outro problema é o fato de que nem todos os pontos no Gráfico 3 correspondem a uma média aritmética; alguns desses valores são absolutos, pois os informantes não produziram o elemento vocálico em todas as repetições. Por essas razões se decidiu rever a metodologia do experimento e realizar uma nova análise dos grupos compostos por fricativas. Além da utilização de sentenças-veículo no novo experimento (ver seção 1.1.1), utilizou-se um grupo de controle com palavras que funcionassem como pares mínimos, ou seja, dissílabos paroxítonos que seguem seqüência idêntica ou análoga de segmentos. Os dados foram coletados junto a dois informantes (D e P) e analisados seguindo a metodologia utilizada anteriormente.

	[i]	[e]	[E]	[a]	[O]	[o]	[u]
grupo de controle	fita	-	feto	fato	foto	fone	fuga
palavras-alvo	frito	frevo	freto	fraco	frota	fronha	fruta

Tabela 2: Dissílabos paroxítonos utilizados como palavras-alvo e grupo de controle no novo experimento.

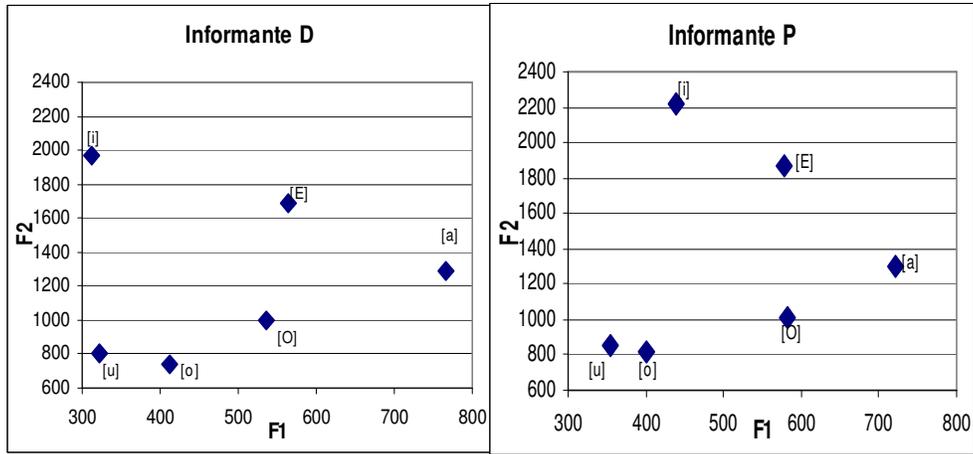


Gráfico 4: Médias dos valores de F1 X F2 das vogais nucleares das palavras do grupo de controle (informante D e P).

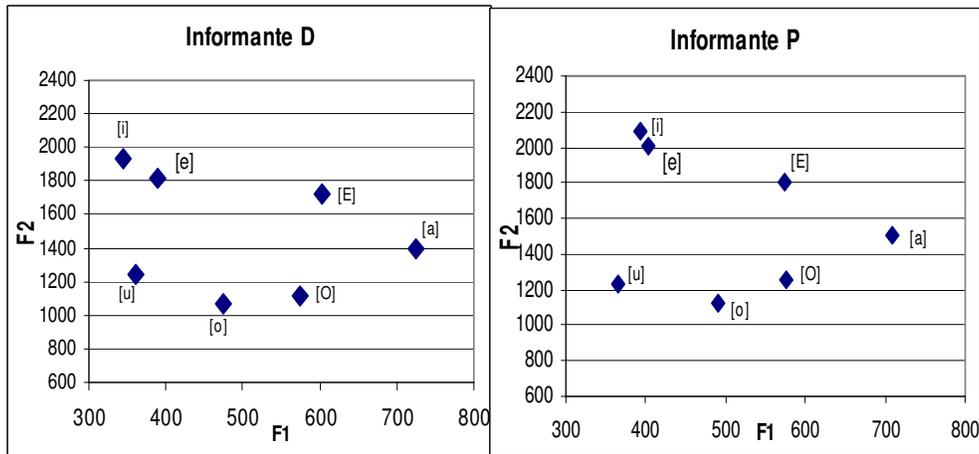


Gráfico 5: Médias dos valores de F1 X F2 das vogais nucleares dos grupos compostos por fricativas dos informantes D e P.

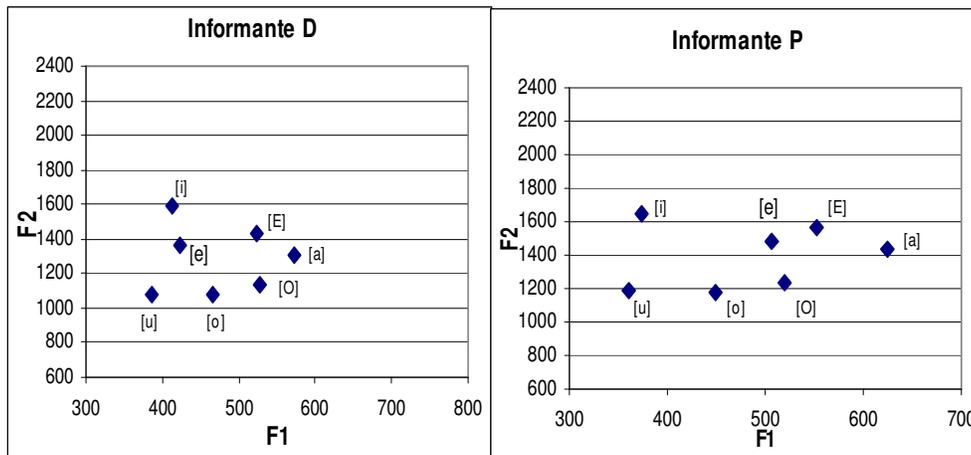


Gráfico 6: Médias dos valores de F1 X F2 dos elementos vocálicos em grupos compostos por fricativas dos informantes D e P.

No gráfico 4, acima, há as médias dos valores de F1 e F2 das vogais nucleares das palavras do grupo de controle. Pode-se notar que o gráfico forma o triângulo vocálico do PB, tal como se tinha para as vogais nucleares dos grupos do experimento anterior. É possível perceber que o ponto que corresponde à vogal [i] do informante **P** assume comportamento distinto do de **D**, sendo mais aberta que o [i] do informante **D** e que seu próprio [u]. **P** distingue [i] e [u] pela abertura de mandíbula (aumento no valor de F1), ou seja, [i] e [u] não se distinguem apenas por diferença no F2 como se tem para o informante **D**. Essas peculiaridades também poderão ser notada nos outros gráficos, sendo características específicas de cada informante, não se tratando de influências causadas pelas fricativas.

O gráfico 5, acima, contém as médias dos valores de F1 e F2 das vogais nucleares das palavras-alvo do novo experimento. Pode-se notar que há a produção de um triângulo vocálico como se tem para o grupo de controle.

O Gráfico 6, acima, apresenta as médias de F1 e F2 dos elementos vocálicos nos grupos com fricativas. Ao contrário do que havia no experimento anterior, todos os pontos desse gráfico são médias das medidas efetuadas, nenhum dado é um valor absoluto. Pode-se notar que há a produção de um triângulo vocálico para ambos os informantes. No experimento anterior, o gráfico de um dos informantes não chegou a formar um triângulo vocálico, levando a levantar a hipótese que em grupos compostos por fricativas se produzia um elemento vocálico com estrutura formântica semelhante ao de uma vogal neutra do tipo *schwa*, ou seja, uma vogal centralizada. No entanto, esses novos dados refutam essa hipótese, pois o Gráfico 6 ilustra, tanto para **D** como para **P**, um triângulo vocálico com as mesmas relações entre os pontos que se tinha para os dados do grupo de controle e das vogais nucleares dos grupos.

1.3 RESULTADOS

Conforme o que foi apresentado nas seções anteriores, pode-se dizer que os grupos com fricativas se comportam da mesma maneira que os grupos compostos por oclusivas, isto é, nos grupos do PB constituídos por “obstruinte + /r/ + vogal” há a produção de um elemento vocálico entre a obstruinte e o *tap* que tem sua estrutura formântica condicionada pela qualidade da vogal nuclear do grupo. Com isso, a hipótese que se levanta é a de que o *tap*, quando em grupos do PB, estaria entrecortando a vogal nuclear após alguns milissegundos e, assim, produzindo o elemento vocálico com a mesma estrutura formântica.

2. ANÁLISE ACÚSTICA DO TAP EM CODAS

Assim como já observado para o ambiente de grupos no PB (Nishida, 2004), o tap em posição de coda também é seguido de abertura oral, doravante chamado de elemento vocálico. Estudo anterior (Clemente, 2004), elaborado inicialmente para observar o fenômeno de apagamento do tap em posição de coda no PB, observa que, para fala de laboratório, tal apagamento não é recorrente, mas existe após o tap em posição de coda e antes da próxima consoante a presença de um elemento de natureza vocálica.

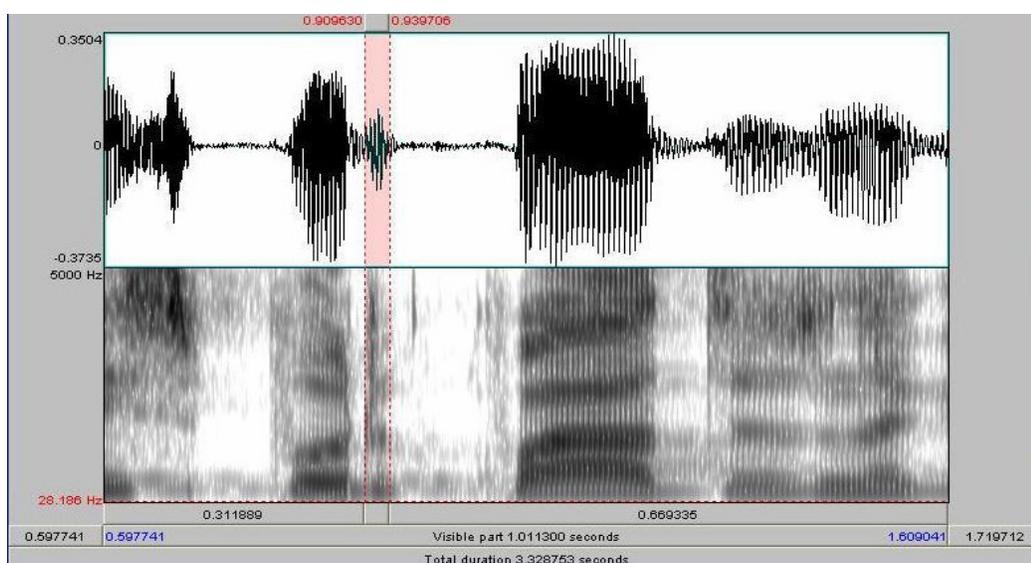


Figura 2: Espectrograma do tap seguido do elemento vocálico. A palavra-alvo é *expor* e o som em questão encontra-se entre as barras verticais.

Nesta seção procuramos averiguar se o elemento vocálico realmente ocorre em coda no PB, analisando um maior número de dados, visto que, para o PB, existe uma escassez de descrições acústicas sobre o comportamento do *tap* em coda¹¹.

Outro fator importante será caracterizar a natureza de tal elemento vocálico. Existem duas possibilidades: a primeira é que tal elemento trata-se de uma vogal neutra, apresentando comportamento diferente da vogal tônica que antecede o *tap*. A segunda hipótese é que o elemento vocálico apresenta as mesmas características da vogal tônica antecedente ao *tap*, que se entropia à vogal.

¹¹ Para uma caracterização inicial dos *taps* em codas, remetemos o leitor a Silva (1996).

2.1 METODOLOGIA

2.1.1 DESIGN EXPERIMENTAL

Um novo experimento foi elaborado para dar conta das questões elencadas acima, com o especial intuito de observar a natureza do elemento vocálico – se schwa ou a mesma vogal tônica – e verificar se a classe morfológica também é um fator que altera a ocorrência e a natureza do elemento vocálico.

Palavras-alvo foram selecionadas com o auxílio do programa Listas. Essas palavras eram oxítonas dissílabas, pois assim garantimos que o acento tônico recaísse sobre a sílaba onde estava inserido o *tap*, marcando a vogal antecedente a ele. Tais palavras pertenciam ou a classe dos substantivos ou a classe dos verbos.

A tabela 3 mostra as palavras-alvo do *corpus*.

Vogal	Verbos	Substantivos
[i]	PEDIR	FAQUIR
[e]	BATER	LAZER
[E]	*	TALHER
[a]	PEGAR	LUGAR
[O]	*	REDÔR
[o]	DEPOR	SABOR
[u]	*	FÊMUR

Tabela 3: Palavras-alvo analisadas no experimento.

Algumas observações a respeito do *corpus* são necessárias. Note-se que na tabela 3, alguns espaços estão em branco. Seriam os dados correspondentes aos contextos verbais das vogais [E], [O] e [u]. Como é sabido, no PB não há infinitivo verbal terminado com essas vogais. Dessa maneira, não haverá dados correspondentes a essas realizações. Outra observação relacionada aos dados diz respeito à palavra “fêmur”, paroxítona que foge do modelo de palavra utilizado no experimento. Isso se deve ao fato de que, após uma minuciosa busca, nenhum substantivo oxítono e de uso corriqueiro na língua foi encontrado contendo /u/ seguido de *tap*. Como não se optou pela utilização de logotomas, pois poderiam provocar estranhamentos nos informantes, resolveu-se utilizar uma palavra com acentuação diferente.

As palavras-alvo foram então inseridas em sentenças-veículo do tipo

“Digo.....pra ele”. Para a coleta dos dados foram utilizados quatro informantes: dois do sexo masculino (*F* e *R*) e dois do sexo feminino (*A* e *J*). Tanto os informantes do sexo masculino quanto os do sexo feminino são residentes na região do dialeto analisado (curitibano) por mais de oito anos. A faixa etária encontra-se entre os 18 e 35 anos. Dois destes informantes tinham o terceiro grau incompleto na época das gravações. Os outros dois apresentam terceiro grau completo.

É importante salientar que nenhum fator extralingüístico foi levado em consideração, como faixa etária ou grau de escolaridade. A escolha de informantes de sexo distintos se deu pelo fato de haver uma diferença quanto à qualidade de fala de homens e mulheres, de acordo com Fant (1960). O som produzido pelas mulheres durante o ato de fala possui frequências mais altas devido ao tamanho do trato vocal feminino que, geralmente, é menor do que o trato vocal dos homens.

As sentenças-veículo eram lidas de forma aleatória pelos informantes. Foram colhidas cinco repetições para cada informante, totalizando 55 dados por informante.

Como os dados do *tap* em grupos, estes também foram gravados digitalmente, no LACOMUS/DeArtes a uma taxa de amostragem de 44 kHz. Para a análise acústica, conduzida com o auxílio do programa Praat, a exemplo dos dados com *tap* em grupos, consideraram-se os valores de F1, F2 e F3 das vogais tônicas e dos elementos vocálicos, assim como a duração de ambos os segmentos.

2.2. DISCUSSÃO DOS DADOS

Os resultados obtidos através da inspeção visual dos espectrogramas relativos aos dados dos informantes *F* e *R* confirmam, nos dados onde há a produção de *tap* em posição de coda, a existência de um elemento de característica vocálica situado entre o segmento em questão e a consoante adjacente. Não houve sequer um caso onde ocorresse a produção do *tap* em coda sem o elemento vocálico. Esse fato foi recorrente para ambos os informantes.

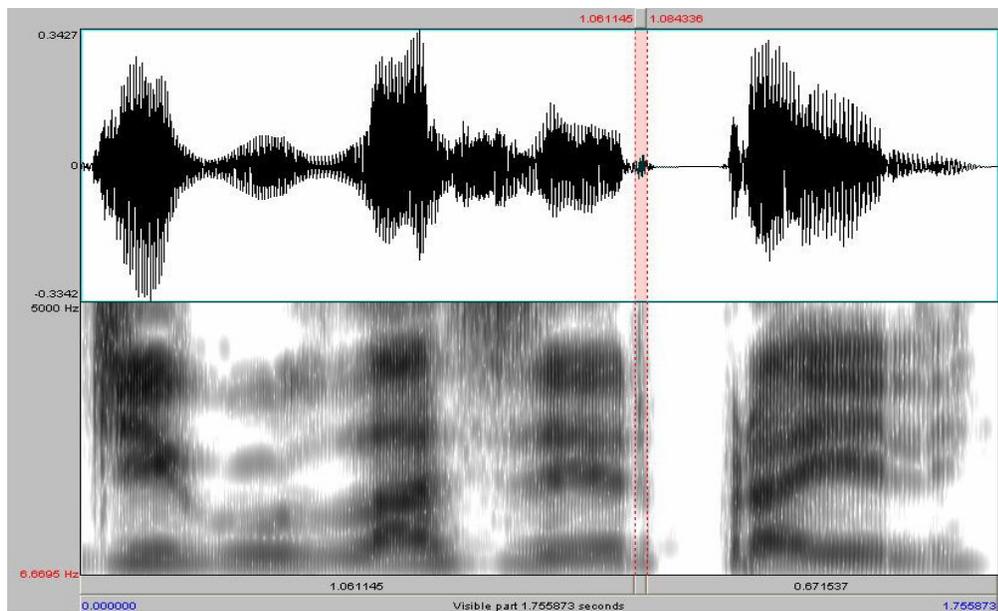


Figura 3: Espectrograma mostrando o elemento vocálico após o tap na palavra “lazer”, para dados do informante R. O som em questão encontra-se entre as barras verticais.

Certificada a existência do elemento vocálico após o *tap* em posição de coda, passou-se a observar o comportamento do elemento através da análise dos valores das frequências dos seus formantes. Tal análise foi de suma importância para identificar a natureza do elemento vocálico.

Um procedimento foi seguido para facilitar a comparação entre as vogais tônicas e o elemento vocálico. Das médias das frequências de cada vogal tônica, subtraíram-se as médias das frequências dos elementos vocálicos. O raciocínio é simples: comparando os resultados obtidos, é possível saber se o elemento vocálico se aproxima da vogal tônica ou se há um distanciamento entre os segmentos. Válido lembrar que existem correlatos entre as partes acústicas e articatórias. Logo, através dos resultados das frequências, é possível saber qual a natureza do elemento vocálico.

Os resultados dos valores das frequências foram plotados e gráficos do tipo F1 por F2. Esses gráficos auxiliam na compreensão dos resultados. Os resultados foram separados entre verbos e substantivos para os dois informantes.

Verbos	F1 v.t. ¹²	F1e.v. ¹³	Diferença		F2 v.t.	F2 e.v.	Diferença
[i]	425	333	92	[i]	2035	1712	325
[e]	X	X	X	[e]	X	X	X
[E]	397	323	74	[E]	1975	1678	298
[a]	779	341	438	[a]	1504	1474	30
[O]	X	X	X	[O]	X	X	X
[o]	381	344	37	[o]	826	1349	-523
[u]	X	X	X	[u]	X	X	X

Tabela 4: Diferenças dos valores médios das frequências entre vogais tônicas e elemento vocálicos nos verbos para o informante R.M.

Verbos	F1 v.t.	F1 e.v.	Diferença		F2 v.t.	F2e.v.	Diferença
[i]	350	422	-72	[i]	1741	1303	438
[e]	387	411	-24	[e]	1755	1408	347
[E]	X	X	X	[E]	X	X	X
[a]	696	503	193	[a]	1358	1327	31
[O]	X	X	X	[O]	X	X	X
[o]	478	445	33	[o]	942	1346	-404
[u]	X	X	X	[u]	X	X	X

Tabela 5: Tabela contendo as diferenças dos valores médios das frequências entre vogais tônicas e elemento vocálicos nos verbos F

Substantivos	F1v.t.	F1e.v.	Diferença		F2 v.t.	F2 e.v.	Diferença
[i]	381	302	79	[i]	2085	1746	339
[e]	611	336	275	[e]	1842	1518	324
[E]	386	355	31	[E]	1913	1583	330
[a]	780	284	496	[a]	1488	1507	-19
[O]	566	484	82	[O]	1208	1407	-199
[o]	473	446	27	[o]	1113	1355	-242
[u]	366	350	16	[u]	1043	1362	-319

Tabela 6: Tabela contendo as diferenças dos valores médios das frequências entre vogais tônicas e elemento vocálicos nos substantivos para o informante R.M.

Substantivos	F1v.t.	F1e.v.	Diferença		F2v.t.	F2e.v.	Diferença
[i]	324	332	-8	[i]	1894	1412	482
[e]	409	420	-11	[e]	1659	1391	268
[E]	606	492	114	[E]	1537	1438	99
[a]	711	525	186	[a]	1325	1331	-6
[O]	605	528	77	[O]	1045,4	1260	-215
[o]	531	459	72	[o]	1026	1366	-340
[u]	394	391	3	[u]	1057	1304	-274

Tabela 7: Tabela contendo as diferenças dos valores médios das frequências entre vogais tônicas e elemento vocálicos nos substantivos para o informante F.

¹² v.t. = vogal tônica.

¹³ e.v.= elemento vocálico.

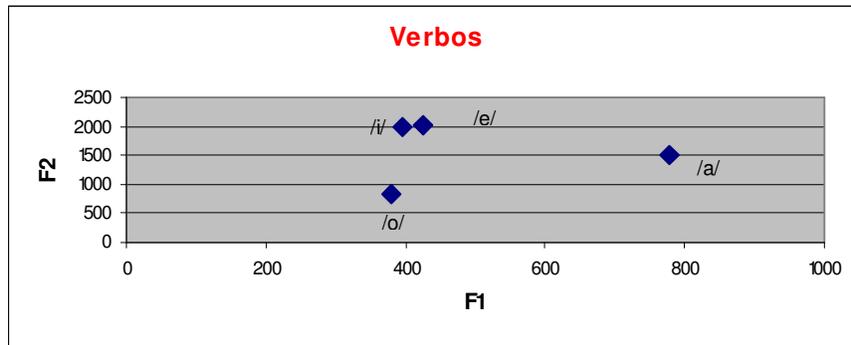


Gráfico 7: Valores de F1x F2 das vogais tônicas nas palavras-alvo da classe dos verbos para o informante R

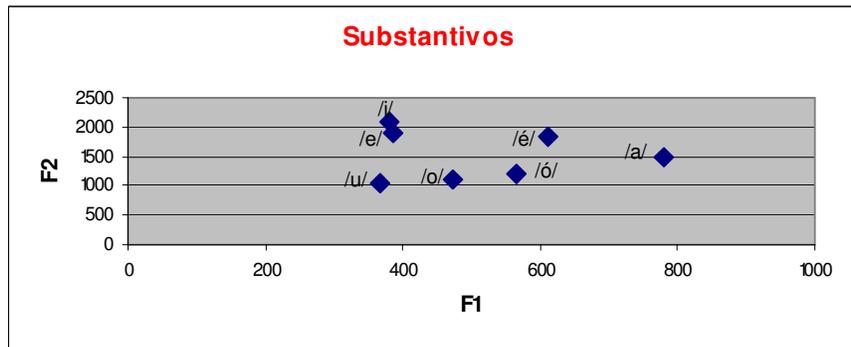


Gráfico 8: Valores de F1x F2 das vogais tônicas nas palavras-alvo da classe dos substantivos para o informante R

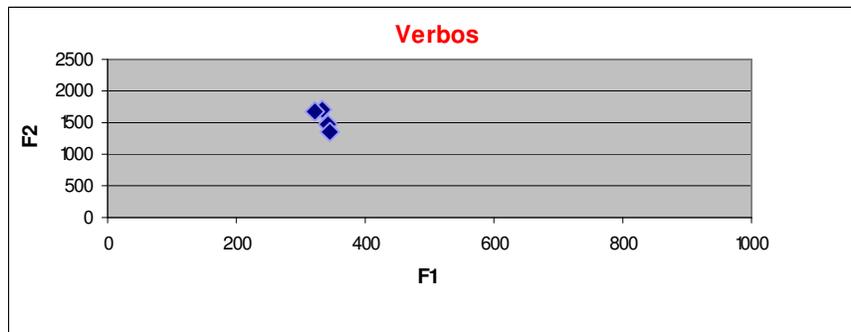


Gráfico 9: Valores de F1x F2 dos elementos vocálicos nas palavras-alvo da classe dos verbos para o informante

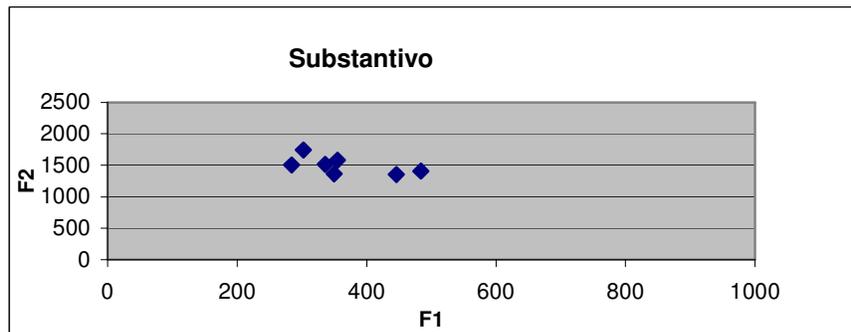


Gráfico 10: Valores de F1x F2 dos elementos vocálicos nas palavras-alvo da classe dos substantivos para o informante R

Para dados de informantes femininos, os resultados foram os mesmos, salvo que as medidas das frequências dos formantes foram mais altas, por motivos idiossincráticos.

Quanto à duração do elemento vocálico, foi notado que tal segmento tem uma duração muito menor se comparado à vogal tônica. As tabelas abaixo contêm os valores de duração da vogal tônica e do elemento vocálico:

Informante	[i]	[e]	[E]	[a]	[O]	[o]	[u]
R	0,17	0,16	0,17	0,18	0,18	0,16	0,08
F	0,07	0,07	0,09	0,12	0,11	0,11	0,09

Tabela 8: Médias das durações das vogais tônicas, em s.

Informante	[i]	[e]	[E]	[a]	[O]	[o]	[u]
R	0,022	0,020	0,016	0,019	0,017	0,020	0,023
F	0,024	0,020	0,020	0,017	0,016	0,018	0,020

Tabela 9: Médias das durações dos elementos vocálicos, em s.

2.3. RESULTADOS

A inspeção visual dos dados evidencia a existência de uma abertura oral após a realização do *tap* em posição de coda. No que diz respeito à sua natureza, fica claro que o elemento vocálico não é a vogal tônica entrecortada pelo *tap*. Esta era uma das hipóteses levantadas no início da pesquisa. Os dados do primeiro experimento não se apresentaram claros o suficiente para elucidar esta questão, porém os resultados obtidos a partir destes mesmos dados nos conduziram a elaborar um novo experimento que desse conta de fornecer informações sobre a natureza do elemento vocálico. Tem-se, então, que o elemento vocálico se assemelha a uma vogal neutra centralizada.

Nesse sentido, o elemento vocálico adjacente ao *tap* exibe um comportamento distinto, a depender da sua posição no interior da sílaba: como observado na seção 1, *supra*, em grupos o elemento vocálico é a mesma vogal nuclear, que o *tap* entrecorta. Em codas, por outro lado, o elemento vocálico exibe, como mencionado, natureza distinta daquela da vogal nuclear.

3. PROPOSTA DE REPRESENTAÇÃO DINÂMICA PARA O TAP INTERVOCÁLICO

A representação que se segue abaixo foi proposta em Silva (2002) para o *tap* intervocálico. Baseada na Fonologia Articulatória (Browman & Goldsten, 1986, 1989, 1990, 1992) e na Fonologia Acústico-Articulatória (Albano, 2001) tal representação assume os gestos articulatórios como unidades de análise fonológica e fonética, a um só tempo, os quais têm suas dimensões formalizadas em pautas gestuais, como na Figura 4¹⁴. Nas pautas, cada gesto é representado por uma caixa, sendo o tempo de ativação gestual representado no eixo horizontal de cada caixa e a magnitude do gesto no eixo vertical. Variações nas dimensões horizontal e vertical das caixas, portanto, traduzem variações nesses parâmetros.

É possível, ainda, manipular a sincronia temporal entre os gestos, de modo que o *onset* de um gesto coincida com o *onset* de outro – como no caso dos gestos das regiões dorsal, faríngea e glotal, da figura abaixo – ou que o *onset* de um gesto coincida com o *offset* de outro¹⁵, ou ainda que o *onset* de um gesto ocorra durante o tempo de ativação de outro gesto como, e.g., o gesto da região coronal, comparativamente com o da região dorsal.

A representação prevê também – seguindo Albano (*op. cit.*) que os gestos se organizem regiões articulatórias, o que advém do fato de tal modelo incorporar informação acústica a um modelo dinâmico de produção da fala. Temos, então:

¹⁴ Para maiores detalhes sobre a natureza do gesto articulatório e sobre a arquitetura dos modelos dinâmicos citados, vide Silva (2002).

¹⁵ A Figura 4 não prevê esse tipo de relação entre os gestos. Para uma explicação mais detalhada sobre as relações de fase entre os gestos, vide, e.g., Browman & Goldstein, 1990.

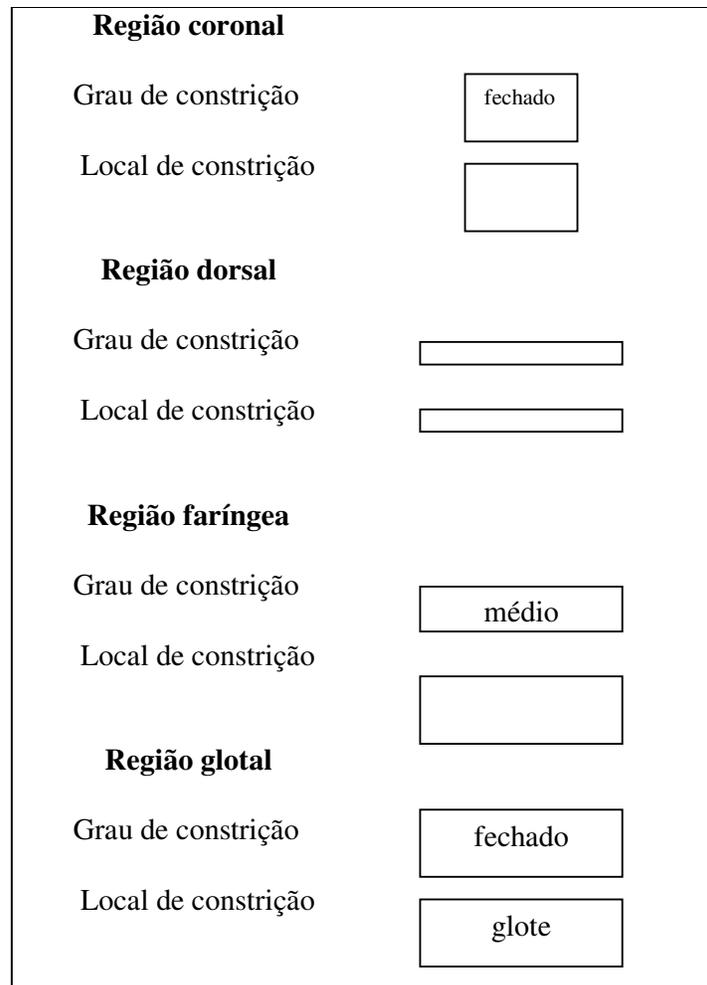


Figura 4: Pauta gestual para *tap* intervocálico, cf. Silva (2003)

Ainda é preciso observar, relativamente à Figura 4, que ela admite estarem envolvidos na produção do *tap* não só ponta e dorso de língua – em decorrência da ativação dos gestos das regiões coronal e dorsal, respectivamente – mas também raiz de língua e glote – esta especialmente no caso dos *taps* sonoros.

Observe-se que o tempo de ativação do gesto de região coronal (disposto no eixo horizontal das caixas), é bem menor que o dos demais gestos. Tal previsão implicará num gesto balístico de ponta de língua. O gesto da região dorsal, por sua vez, tem magnitude diminuída, em comparação com os demais (examine-se a dimensão vertical das caixas). Isso implica que o gesto da região coronal será preponderante na produção do *tap*, relativamente ao gesto de região dorsal.

A proposta apresentada na Figura 4 admite ainda, à luz da FAR e da Fonologia Acústico-Articulatória, que os gestos tenham grau e local de constrição caracterizados por descritores gestuais, como “fechado”, “médio” ou “glote”. Assim, especificamente

para a proposta do *tap* intervocálico, a previsão de que o grau de constrição da região coronal é caracterizado pelo descritor “fechado”, por exemplo, implica na implementação de movimento de ponta de língua que realiza obstrução no trato. A previsão de que o grau de constrição do gesto da região faríngea se caracteriza pelo descritor “médio” implica na implementação de uma constrição faríngea. Além disso, os descritores da região glotal, como previsto, devem implicar em aproximação das pregas.

A representação apresentada acima prevê que ocorram vogais, à esquerda e à direita do *tap*. Quer se admita que elas possam ser representadas por um único gesto – o de região dorsal –, quer se admita que elas devem também ser representadas por mais de um gesto – o de região faríngea, como no caso de [a], ou o de uma região labial, como no caso de [u] –, pode-se prever que o *offset* da vogal à esquerda coincida com o *onset* do *tap* e que o *offset* deste coincida com o *onset* da vogal à sua direita, fato que sinaliza alguma coarticulação entre os três segmentos.

3.1. TAP EM GRUPOS

Como relatado na seção 1, acima, em grupos o *tap* entrecorta a vogal nuclear sem que, no entanto, ele seja percebido como intervocálico. Uma aparente razão para o fato é a duração bastante reduzida da porção da vogal que fica à esquerda do *tap*.

É preciso frisar que, em ambiente de grupos, a representação necessita prever que a vogal à esquerda do *tap* deve ser igual àquela da direita.

Uma representação possível para tal fato é a que se apresenta na Figura 5, abaixo. Em tal representação, prevê-se que haja um único gesto dorsal, responsável pela produção da vogal¹⁶ e que, a ele, se sobreponham os gestos responsáveis pela produção do *tap*. Observe-se:

¹⁶ Ressalte-se que, aí, se admite que o gesto dorsal seja o principal responsável pela produção das vogais, embora seja necessário prever também a atividade de outros gestos, responsáveis por efeitos que se adicionam à produção das vogais, como arredondamento de lábios ou atividade faríngea.

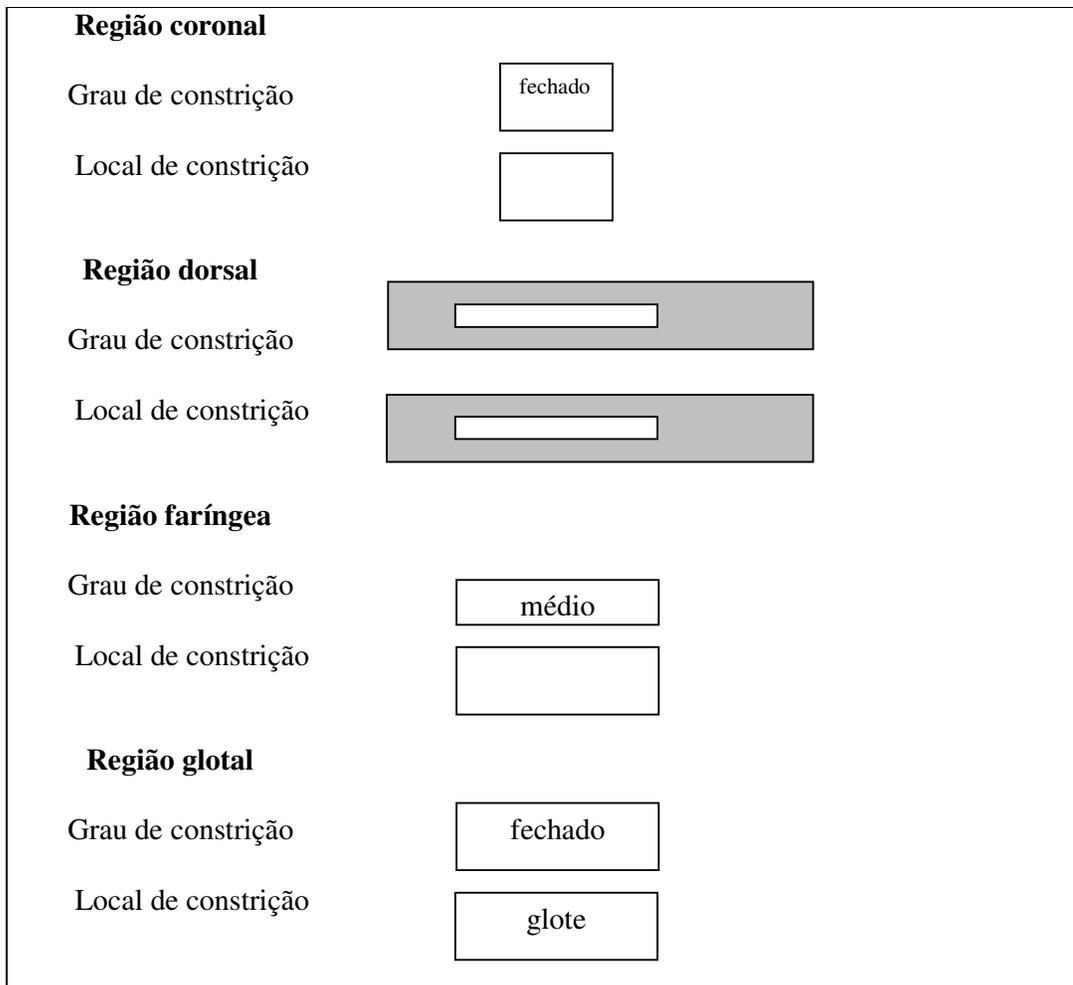


Figura 5: Pauta gestual para a seqüência vogal – tap – vogal, em grupos do tipo “obstruinte + tap + vogal”, prevendo-se um único gesto dorsal, com tempo de ativação mais longo.

É necessário esclarecer que, na pauta acima, o gesto de região dorsal representado pelas caixas maiores está sombreado apenas por fins expositivos: a intenção, aí, é sinalizar que o gesto responsável pela produção da vogal se realiza ao mesmo tempo que o gesto de região dorsal, presente na pauta do *tap* e representado, na Figura 5, pelas caixas menores. O gesto da região dorsal da vogal, representado nas caixas sombreadas, porém, se sobrepõe ao gesto de região dorsal do *tap* por ter magnitude e tempo de ativação maiores. Com isso, “esconde” o gesto de região dorsal do *tap*.

Ainda na pauta acima, nota-se que os gestos responsáveis pela produção do *tap* se realizam durante o tempo de ativação do gesto da vogal. O momento em que se inicia a ativação dos gestos do *tap* deve ser determinado pela relação de fase entre eles e os gestos da vogal.

3.2. PROPOSTA DE REPRESENTAÇÃO DINÂMICA PARA O *TAP* EM CODAS

Neste ambiente, pode-se adotar uma representação semelhante àquela do *tap* quando em grupos. Deve-se considerar, porém, que nesse contexto o *tap* exibe comportamento distinto daquele de grupos. Ou seja, independentemente do contexto vocálico precedente, o elemento vocálico que sucede o *tap* em codas de final de palavra parece tender à centralização, assemelhando-se a um *schwa*. Assim, propõe-se a representação disposta na Figura 6, a qual se assemelha sobremaneira à representação de um *tap* intervocálico, mas dela difere por considerar que à direita do *tap* deverá ocorrer sempre a mesma vogal a qual, por sua vez, difere da vogal à esquerda, não só pela qualidade, mas também pela duração, que é mais breve. Tem-se, então:

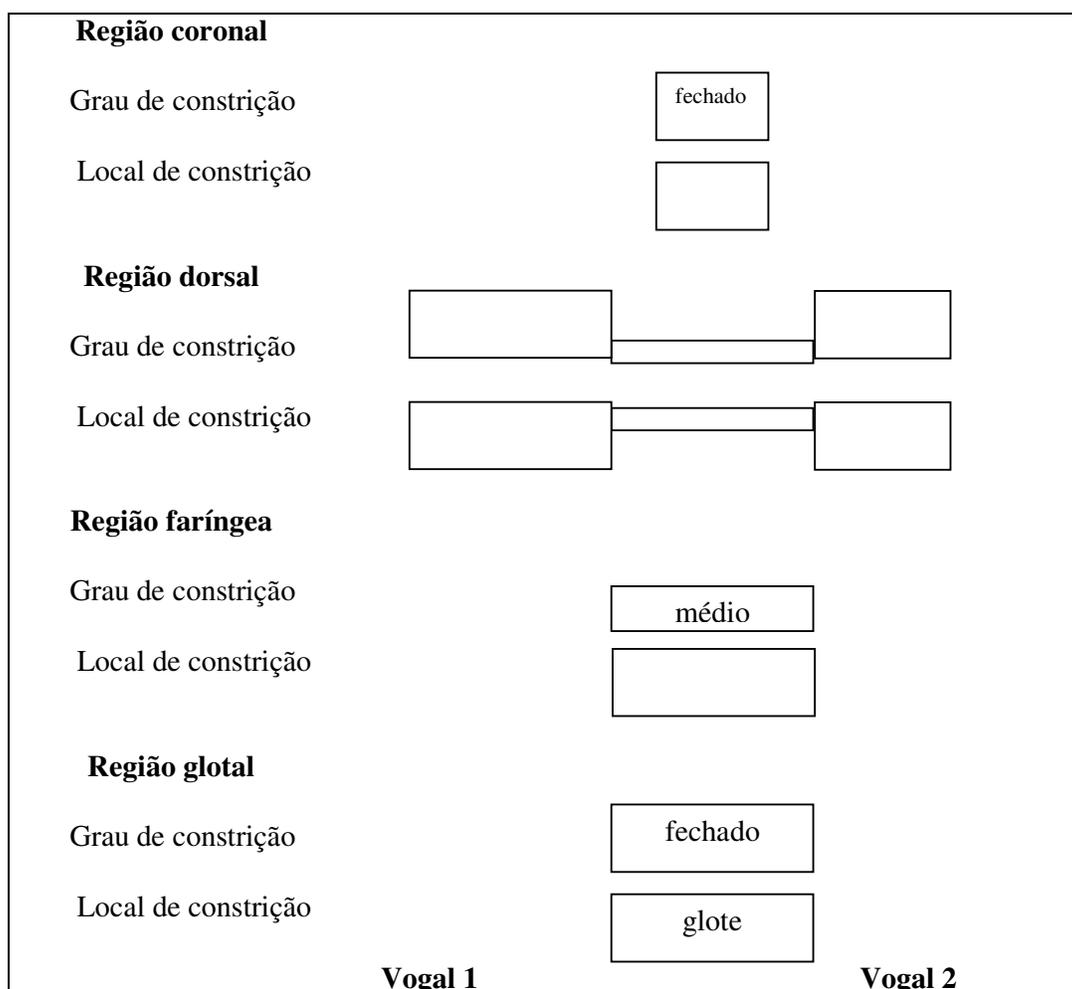


Figura 6: Pauta gestual para o *tap* em codas.

É preciso observar, relativamente às vogais da pauta, que sua qualidade é dada

por descritores gestuais, espécies de índices que carregam informações relativas, e.g., à altura de mandíbula. Neste caso, como a vogal da esquerda pode ser qualquer uma, optou-se por não especificar o descritor gestual. No caso da vogal à direita, como até agora as descrições têm apontado para uma semelhança com uma vogal neutra, do tipo *schwa*, também não se especificam os descritores gestuais.

Considerando-se as representações propostas nas Figuras 5 e 6, acima, e considerando-se também a pauta para um *tap* intervocálico faria a previsão de uma vogal à direita do *tap* com qualquer qualidade e com duração maior do que a da vogal da Figura 6, resta ainda um problema, que aqui será apenas mencionado – até pela limitação do espaço – mas que deverá ser abordado em estudos posteriores.

O problema, considerando-se os modelos dinâmicos adotados, é prever que as pautas para os *taps* teriam de “enxergar” o ambiente silábico no qual se encontram, o que implicaria admitir a existência de um nível fonológico, um problema especialmente para a Fonologia Articulatória, que prescinde totalmente dele. A Fonologia Acústico-Articulatória, por sua vez, reconhece a necessidade de se prever um nível fonológico para se dar conta de alguns fatos do português brasileiro – em especial os casos de junção intervocabular. Talvez o comportamento distinto do *tap*, a depender de sua posição na sílaba, seja um fato que reforce o pressuposto da Fonologia Acústico-Articulatória. Mas, como já mencionado, esse é objeto de reflexão para um outro estudo.

4. À GUIA DE CONCLUSÃO

A necessidade de propor uma representação dinâmica para os *taps* decorre do fato de que tais segmentos “escolhem” a natureza do elemento vocálico que ocorre adjacente a eles. E isso varia entre as línguas; afinal, como mencionado no início deste estudo, em línguas como o búlgaro ou o norueguês os elementos vocálicos são, invariavelmente, vogais neutras (ou *schwas*) e noutras línguas, como no espanhol ou no português, o elemento vocálico é, especialmente em grupos, a vogal do núcleo cuja produção o *tap* interrompe para se realizar. O interessante, no caso do português brasileiro, é que em codas os segmentos em questão exibem comportamento distinto, de modo que os elementos vocálicos assemelham-se a vogais neutras.

Esse tipo de fato não é capturado por nenhum modelo de análise fonológica que se baseie na Fonologia Gerativa de Chomsky & Halle (1968) – ou seja, praticamente todos os disponíveis – devido à natureza categórica dos primitivos de análise adotados,

que não admitiriam que um segmento fosse “entrecortado” para a realização de um segundo.

Por outro lado, modelos dinâmicos de produção da fala, como a Fonologia Articulatória ou modelos nele inspirados conseguem fornecer uma representação parcimoniosa para o *tap* em quaisquer das situações verificadas. Ainda assim, resta o problema, não previsto por esses modelos, do condicionamento da estrutura silábica sobre a natureza dos elementos vocálicos.

Assim, ao mesmo tempo em que os *taps* fornecem fortes argumentos para a adoção de representações fônicas dinâmicas, eles colocam problema para esses modelos, que adotam primitivos de natureza gradiente, uma vez que os mesmos não prevêem noções como as de sílaba em seu arcabouço. Os *taps* colocam, então, a necessidade de um refinamento dessa classe de modelos.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ALBANO, E. C. *O gesto e suas bordas: esboço de Fonologia Acústico-Articulatória do Português Brasileiro*. Campinas: Mercado de Letras/ALB/FAPESP, 2001.
2. ALMEIDA, M. & DORTA, J. Datos acústicos de las líquidas españolas. In: C. D. Alayón (org.) *Homenaje a José Pérez Vidal*. La Laguna: Tenerife, 1993: 97-110.
3. BROWMAN, C. & GOLDSTEIN, L. Towards an Articulatory Phonology, in *Phonology Yearbook*, 3: 219-252, 1986.
4. BROWMAN, C. & GOLDSTEIN, L. Articulatory gestures as phonological units, in *Phonology Yearbook*, 6: 201-251, 1989.
5. BROWMAN, C. & GOLDSTEIN, L. Tiers in Articulatory Phonology, in J. Kingston & M. Beckman (orgs.) *Papers in Laboratory Phonology 1: between the grammar and physics of speech*. Cambridge: The Cambridge University Press, pp: 341-376, 1990.
6. BROWMAN, C. & GOLDSTEIN, L. Articulatory Phonology: an overview, in *Phonetica*, 49: 155-180, 1992.
7. CARVALHO, K.C.H.P. *Descrição fonético-acústica das vibrantes no português e no espanhol*. Tese de doutorado, inédita. UNESP/Assis, 2004.
8. CHOMSKY, N. & HALLE, M. *The sound pattern of English*. The MIT Press, 1995 [1968].
9. CLEMENTE, F. C. Características acústicas do tap em coda no PB. Trabalho

- apresentado no 6º encontro do CelSul, Florianópolis, 2004.
10. JETCHEV, G. Rhotics, jers and schwa in the history of Bulgarian. In: *Proceedings of the XIVth. International Congress of Phonetic Sciences*. Stockholm, 1995, (4): 662-665.
 11. KVALE, K. & FOLDVIK, A.K. An acoustic analysis of the retroflex flap. In: *Proceedings of the XIVth. International Congress of Phonetic Sciences*. Stockholm, 1995, (2): 454-457.
 12. LADEFOGED, P. & MADDIESON, I. *The sounds of the world's languages*. Cambridge: Blackwell, 1996.
 13. NISHIDA, G. Características acústicas do tap em grupos no PB. Trabalho apresentado no 6º encontro do CelSul, Florianópolis, 2004.
 14. QUILIS, A. *Tratado de fonética y fonología españolas*. Madrid: Editorial Gredos, 1993.
 15. 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.
 16. SILVA, A. H. P. As fronteiras entre Fonética e Fonologia e a alofonia dos róticos iniciais em PB: dados de dois informantes do sul do país. Tese de doutorado, inédita. LAFAPE/IEL/UNICAMP, 2002.
 17. SILVA, A. H. P. Pela incorporação de informação fonética aos modelos fonológicos, in *Revista Letras*, 60: 319-333, 2003.
 18. SPROAT, R. & FUJIMURA, O. Allophonic variation in English /l/ and its implications for phonetic implementation, in *Journal of Phonetics*, 21:291-311, 1993.

RESUMO: Descrições acústicas do *tap* em grupos e em codas revelam que esse segmento é sempre realizado apoiando-se em um evento acústico de natureza vocálica. Tal elemento, porém, tem natureza variável em função da posição silábica onde o *tap* ocorre: se em grupos, o elemento será a mesma vogal nuclear, entrecortada pelo *tap*; se em codas, o elemento vocálico assume características de vogal neutra. Considerando-se que a natureza do elemento vocálico adjacente ao *tap* é específica da língua, e não universal, é necessário representar tal fato na gramática da língua. Argumentamos que a melhor maneira de se fazer isso é adotando-se modelos dinâmicos de produção da fala, os quais prevêm uma “tradução” direta do nível simbólico no nível numérico devido ao primitivo de análise postulado. Fornecemos, então, uma representação para o *tap* nesses ambientes silábicos, embora observando que o condicionamento do lugar que o *tap* ocupa na sílaba sobre a natureza do elemento vocálico coloca um problema para os modelos dinâmicos, que não prevêm a noção de estrutura silábica.

PALAVRAS-CHAVE: *taps*; grupos; codas; representação dinâmica.

ABSTRACT: Acoustical descriptions of the tap in clusters and codas reveal that it is always produced adjacent to a vowel-like segment. Nevertheless, its nature varies according to the place in the syllable that the tap occupies: if it is in clusters, the vowel-like segment is the same vowel that occurs in the syllable nucleus; if, on the other hand, the tap lays in codas, the vowel-like segment is similar to a neutral vowel. Considering that the nature of the vowel-like segment is language specific – and not universal – it is necessary to provide a phonological representation for such facts. We argue, then, that the best way to do it is to adopt dynamical models of speech production, which postulate a direct “translation” of the symbolic level of the language into the numerical one because of their units of analysis. So, we provide a representation for the tap in clusters and in codas in the light of a dynamical model, although we observe that the fact that the nature of the vowel-like segment is conditioned to the place in the syllable that the tap occurs in constitutes a problem for such models, which do not consider the notion of syllable structure.

KEY-WORDS: taps; clusters; codas; dynamical representation.