

Fonética Acústica

O processamento da linguagem envolve falante e ouvinte. O falante monitora a sua própria fala através do *feedback* (ouve o que ele próprio fala) fenômeno importante para a manutenção de um bom controle dos alvos (sons) a serem articulados. A cadeia da fala envolve três estágios: o linguístico que concerne à construção da mensagem; o articulatório que se refere à produção do sinal que carrega a mensagem; e por fim o acústico, que se diferencia dos demais por ser externo e comum tanto ao falante quanto ao ouvinte (Fig. 1). O nível fisiológico foi apresentado nos Capítulos 2 e 3 da Unidade A. O nível acústico será então aqui tratado.

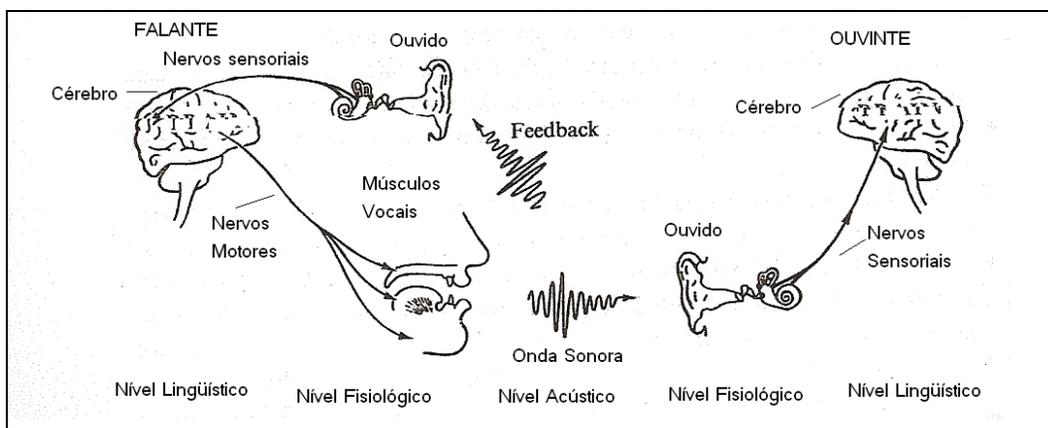


Figura 1. A cadeia da fala (GARMAN, 1990)

A Fonética Acústica estuda os sons da fala a partir de suas propriedades físicas. Esse estudo é processado sobre a cadeia da fala, ou seja, pela análise das ondas sonoras.

Relações acústico-articulatórias

Vejamos agora as relações dos movimentos no trato oral e os seus efeitos acústicos. As vogais geralmente apresentam 4 formantes (ressonâncias) - F1, F2, F3, F4. Mas, é possível caracterizar todas as vogais com os dois primeiros formantes, uma vez que o primeiro (F1) corresponde à altura da língua e o segundo (F2), ao movimento

horizontal da língua. Com relação à altura da língua, F1 é inversamente proporcional a ela, assim sendo, as vogais altas têm F1 baixo e as vogais baixas, F1 alto. Quanto ao eixo horizontal, as vogais mais anteriores apresentam F2 mais alto e as posteriores mais baixo. As médias e as centrais exibem valores intermediários. A Fig. 2 apresenta um diagrama esquemático do posicionamento das vogais com relação à altura e ao recuo/adiantamento da língua no trato oral.

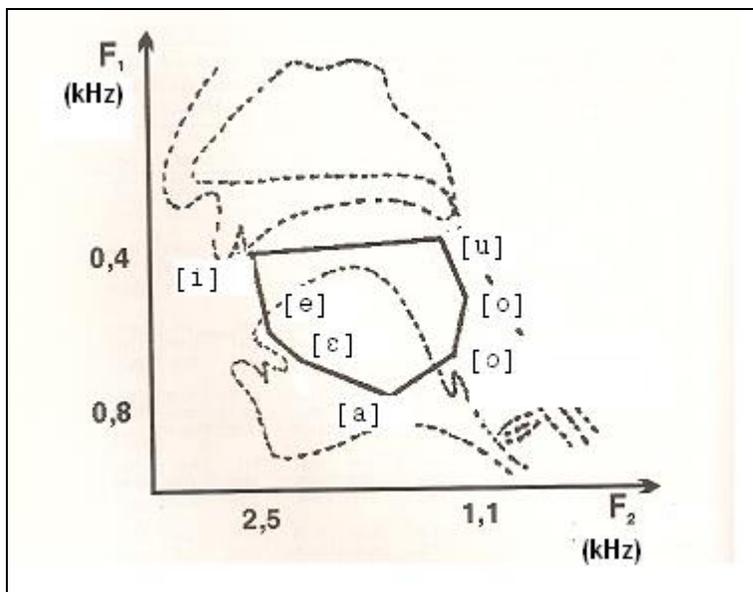


Figura 2. Representação acústico-articulatória das vogais com relação ao quadrilátero vocálico do PB (RUSSO & BEHLAU, 1993:36).

Observando isto, veja, na Tabela 1, a que vogais as frequências apresentadas correspondem. Complete a Tabela com as vogais [i e ɛ a ɔ o u]

Tabela 1. Correspondência de valores de F1 e F2 à altura e anterioridade/posterioridade da língua

Vogais	F1 (Hz)	F2 (Hz)
[i]	320	2110
[u]	320	850
[e]	460	1800
[o]	450	950
[ɛ]	570	1590
[ɔ]	560	1050
[a]	800	1200

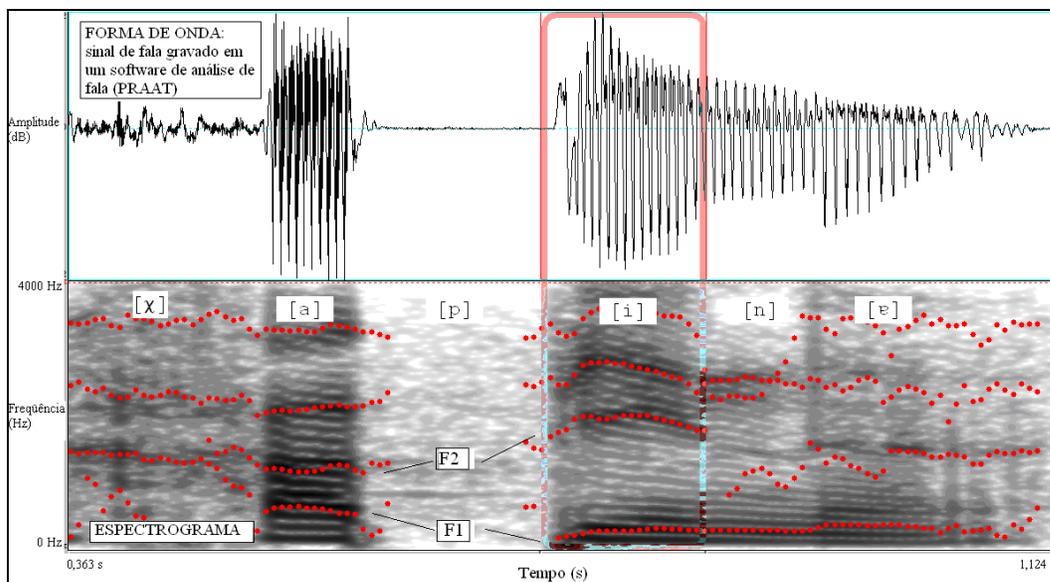


Figura 3. Forma de onda, espectrograma com indicação dos formantes da palavra *rapina* [xa'pinɐ]

A Fig.3 apresenta uma análise acústica da palavra *rapina*. Os parâmetros observados são descritos:

- pela forma de onda, representada no eixo Y pela amplitude (em decibéis) e no eixo X pelo tempo (em segundos);
- pelo espectrograma, representado no eixo Y pelas frequências (em Hertz) (as ressonâncias reforçadas (FORMANTES) estão marcadas em vermelho) e no eixo X pelo tempo (em segundos).

O programa de análise de fala utilizado para estes resultados denomina-se Praat e pode ser baixado livremente no endereço eletrônico: www.praat.org. No Manual para Uso de Praat, você encontra instruções para baixá-lo, gravar um informante e fazer análises com os parâmetros acústicos aqui mencionados. Faça esse exercício. Grave a sua própria fala ou a de um colega e tente analisá-la. Você terá uma visão mais concreta de como produz seus segmentos de fala.

A acústica das consoantes do PB

Com o já vimos, o modo de articulação está relacionado ao tipo de obstrução produzida no trato vocal. Na Figura 4, podemos verificar as constrictões realizadas na produção das consoantes oclusivas e fricativas.

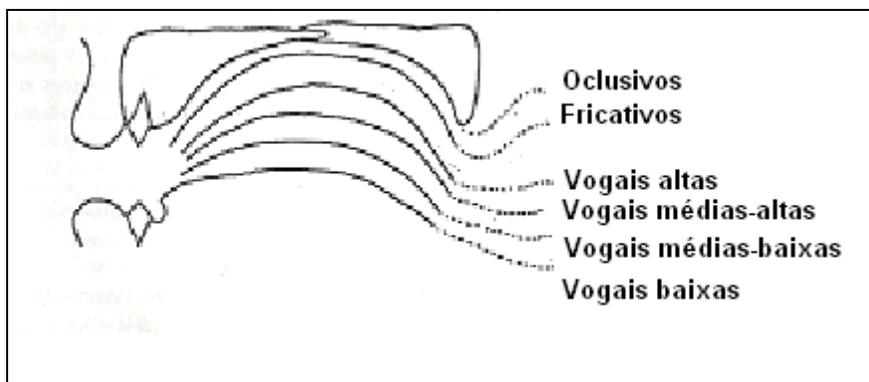


Figura 4. Constrictões do trato oral nas produções vocálicas e consonantais (MASSINI-CAGLIARI e CAGLIARI, 2001:128).

Segundo o **modo de articulação**, as consoantes classificam-se em:

<p>→ Oclusiva/plosiva:</p>	<p>produzida com uma obstrução total e momentânea do fluxo de ar nas cavidades supra-glóticas, realizada pelos articuladores (ativo e passivo), daí chamada de oclusiva. Quando a explosão acústica gerada na liberação da oclusão é considerada, esse segmento pode ser também chamado de plosivo. O véu do palato encontra-se levantado, sendo o fluxo de ar encaminhado apenas para a cavidade oral: paga, data, acaba.</p>
-----------------------------------	---

Acusticamente, estas consoantes podem ser observadas a partir de um período de silêncio que evidencia o tempo de oclusão do trato vocal, e por uma explosão que marca a abertura do trato vocal e a saída brusca de ar da boca. Veja exemplos dessas consoantes vozeadas e não vozeadas na Fig. 5.

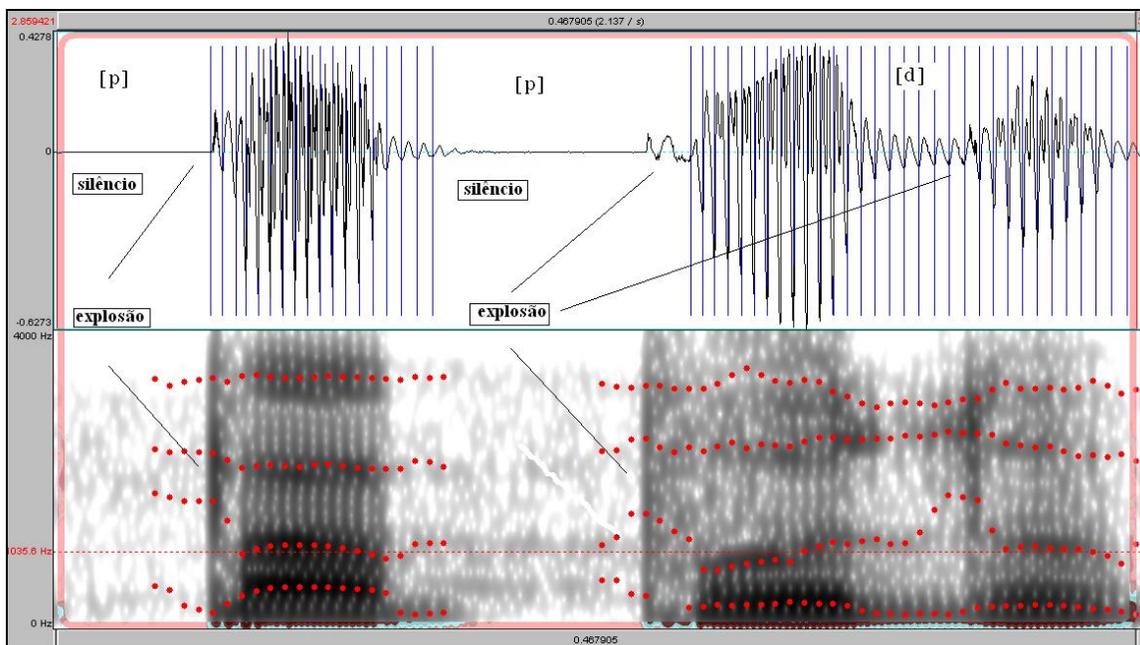


Figura 5. Forma de onda, espectrograma com indicação do silêncio e da explosão nas consoantes oclusivas/plosivas [p] e [d] na palavra *papudo* [pa'pudo]

➔ Nasal:

produzida com uma obstrução total e momentânea do fluxo de ar nas cavidades orais, e com um abaixamento simultâneo do véu do palato, permitindo a liberação do ar pelas cavidades nasais: **mano**, **banho**.

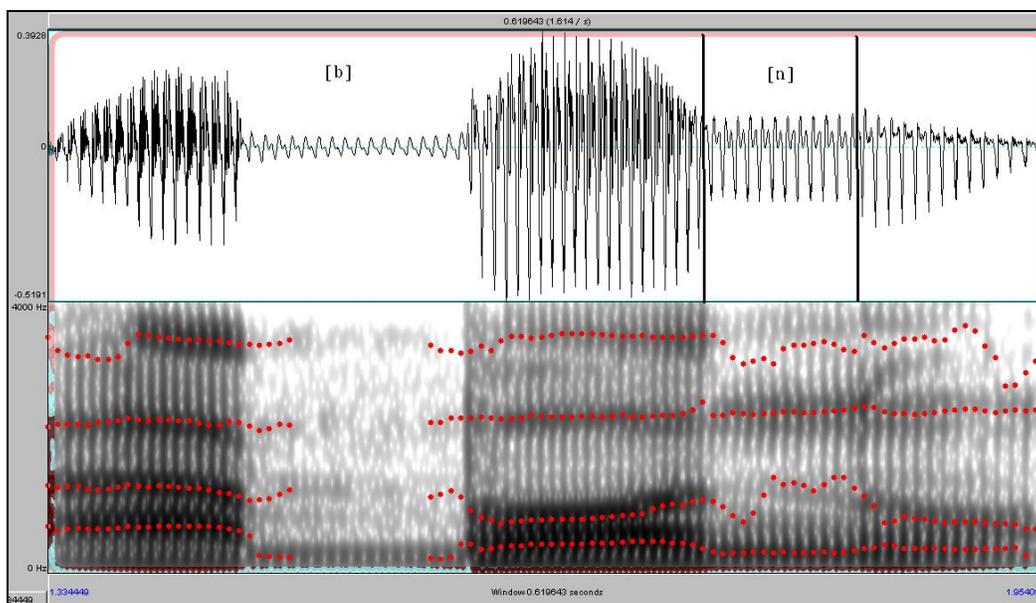


Figura 6. Forma de onda, espectrograma com indicação da região que corresponde a consoante nasal alveolar [n] no espectro acústico na palavra *abono*

[a'bono]. As regiões de grande amplitude do sinal são referentes a segmentos vocálicos.

Observe, na Fig. 6, que a consoante nasal [n] apresenta uma semelhança bastante grande com as vogais, ou seja, pulsos glotais e uma certa energia se comparada a outras consoantes como a plosiva bilabial presente também no espectro. A diferença de energia entre essas duas consoantes ocorre em função da consoante nasal, apesar de ter também uma oclusão no trato oral, permanecer com a cavidade nasal aberta à passagem de ar, gerando maior amplitude. Esse fato não ocorre com a plosiva bilabial que possui um bloqueio total à passagem de ar no trato oral, realizada nos lábios, mas é emitida com o véu do palato levantado não permitindo a passagem do ar pela cavidade nasal.

➔ **Fricativa:**

produzida com um estreitamento do canal bucal, ou seja, uma oclusão parcial, realizada pelos articuladores, fazendo com que o fluxo de ar nas cavidades supra-glóticas gere um ruído de fricção. O véu do palato encontra-se levantado, e o fluxo de ar é encaminhado apenas para a cavidade oral: **fava, saca, casa, chato, jato**

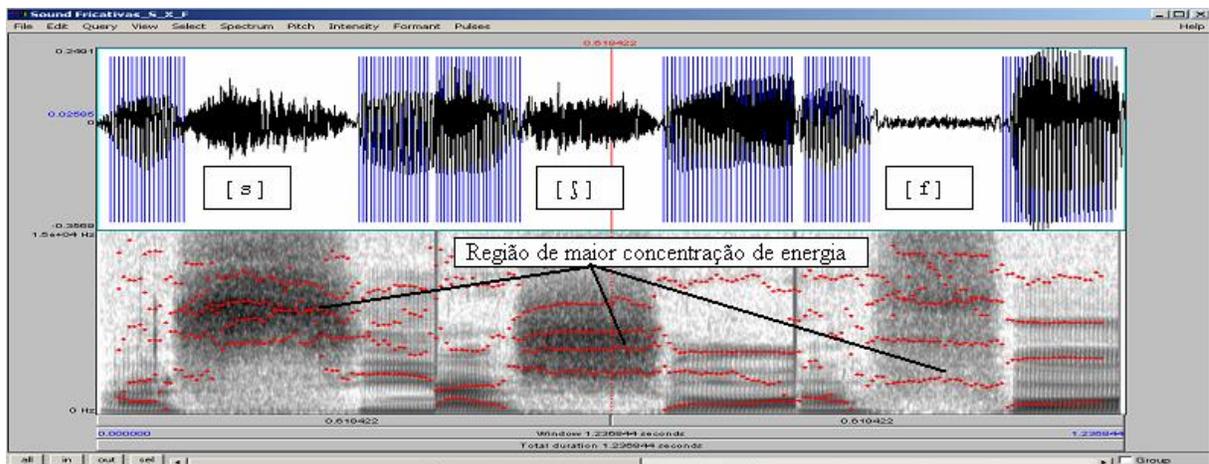


Figura 7. Forma de onda, espectrograma com indicação das regiões de maior concentração de frequências que correspondem a consoantes fricativas alveolar [s], alvéolo-palatal [ʃ] e labiodental [f].

Na Figura 7, observa-se o ruído de turbulência próprio das fricativas e que as três consoantes se diferenciam pelas regiões de maior concentração de energia. Há uma

relação decrescente em região de frequência da alveolar para a alvéolo-palatal e para a labiodental.

→ **Africada:** produzida com uma oclusão total e momentânea do fluxo de ar, seguida de um estreitamento do canal bucal, gerando fricção, no momento da soltura da oclusão. Aqui também o véu do palato encontra-se levantado, e o fluxo de ar passa apenas pela cavidade oral: **tchau**, **tia** e **dia** (no dialeto carioca)

Pela Figura 8, podemos observar que as africadas são compostas por dois momentos acústicos bastante distintos: um período de silêncio que corresponde ao bloqueio do trato oral na região alveolar e um ruído de fricção que ocorre imediatamente após ao momento de relaxamento da oclusão.

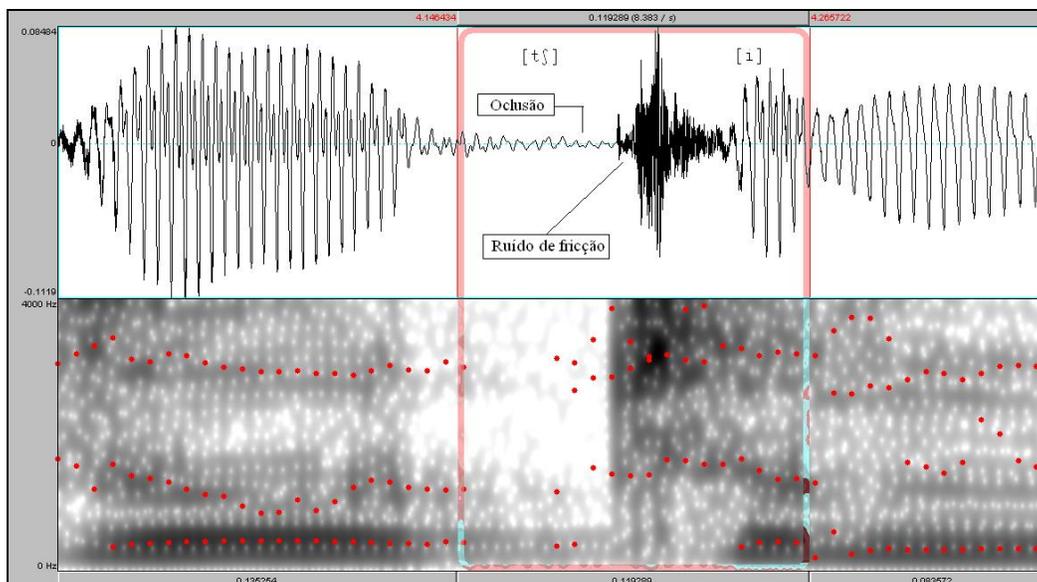


Figura 8. Forma de onda, espectrograma com indicação das regiões que correspondem ao bloqueio do trato oral e da região de ruído de fricção. A região marcada corresponde a sílaba [tʃi] na palavra *último* ['utʃimʊ].

→ **Tepe:** produzida com uma oclusão rápida do fluxo de ar nas cavidades orais. O véu do palato está levantado impedindo a passagem do ar pelas cavidades nasais: **caro**, **prato**.

Pela Figura 9, observamos que a consoante *tap* ([r]) é produzida com apoio vocálico tanto à sua esquerda quanto à sua direita. A pequena vogal que é inserida antes do *tap* em encontros consonantais e depois do *tap* quando ele é produzido em coda silábica (como na palavra *corte*), é chamada de vogal de apoio (SILVEIRA, 2007).

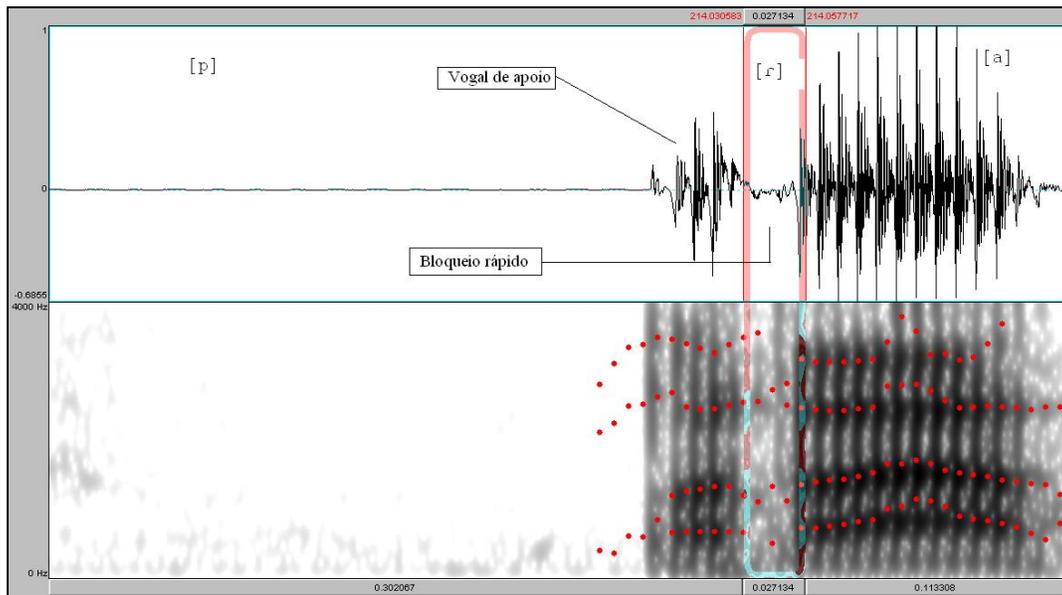


Figura 9. Forma de onda, espectrograma com indicação da região que corresponde ao bloqueio rápido executado na produção do tap alveolar. Esta consoante é produzida sempre com o apoio vocálico (vogal de apoio). A tela mostra a sílaba [pra] da palavra *prato* ['prato].

➔ **Vibrante:**

a ponta da língua ou a úvula provocam uma série de oclusões muito breves, seguidas por segmentos vocálicos extremamente curtos. A passagem do ar pelas cavidades nasais também está bloqueada: *roda*, *carro*.

Em PB, os dois sons de “r” que a língua distingue são muitas vezes chamados de **vibrante simples** (o tap) que aparece na palavra *caro* e **vibrante múltipla** que aparece na palavra *carro*. Pela Fig. 10, verificamos de maneira mais clara a razão desta colocação. Enquanto o tap (Fig. 9) constitui-se de um apoio vocálico e um bloqueio (daí a expressão simples), a vibrante apresenta mais de um bloqueio e mais de um apoio (daí a referência múltiplo), conforme se pode observar na tela da Fig. 10. Ou seja, a vibrante apresenta múltiplos contatos (aberturas e fechamentos) seguidos por um ou mais pulsos glotais nos

quais o fechamento do articulador não se completa, porém produz uma diminuição de amplitude.

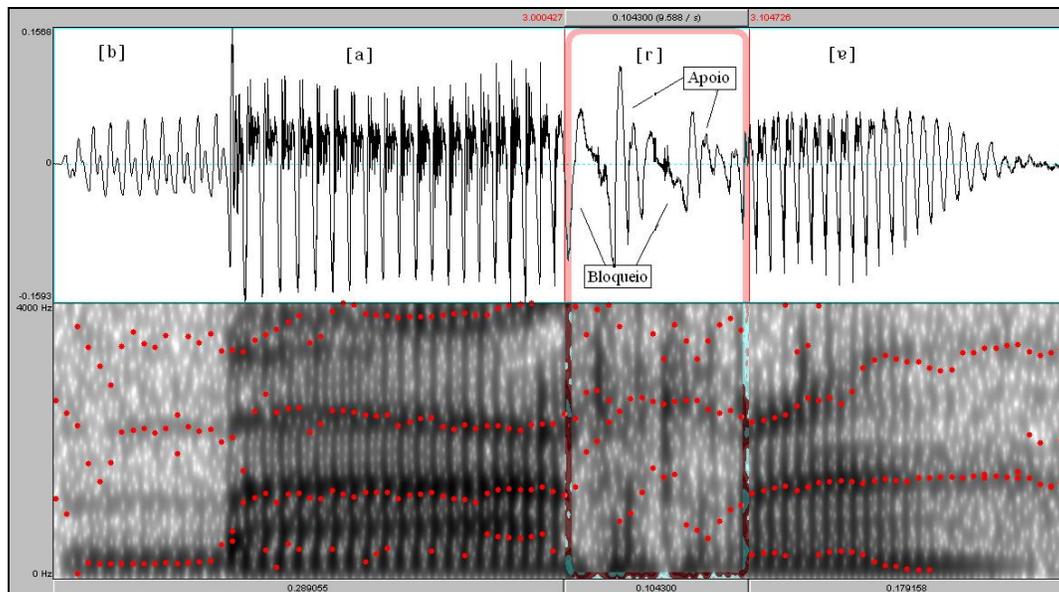


Figura 10. Forma de onda, espectrograma com indicação da região que corresponde à consoante vibrante que é produzida na sílaba [r] da palavra *barra* [ˈbarrɐ].

➔ **Retroflexa:** produzida com o levantamento e encurvamento do articulador ativo (ponta da língua) em direção do articulador passivo (palato duro). Cavidades nasais obstruídas pelo levantamento do véu palatino: marca (pronúncia de “r” no dialeto caipira)

Pela Fig. 11, podemos dizer que acusticamente a produção da variante retroflexa no PB, como no inglês, ocorre devido a uma fusão da vogal com a consoante retroflexa [ɽ] ou [ɽ̃]. Pode-se perceber também um abaixamento abrupto ou um enfraquecimento de algumas frequências mais altas. Este fenômeno caracteriza o traço distintivo denominado bemolizado. Esta variante é encontrada no PB em regiões do interior, cujo falar é caracterizado como “dialeto caipira”.

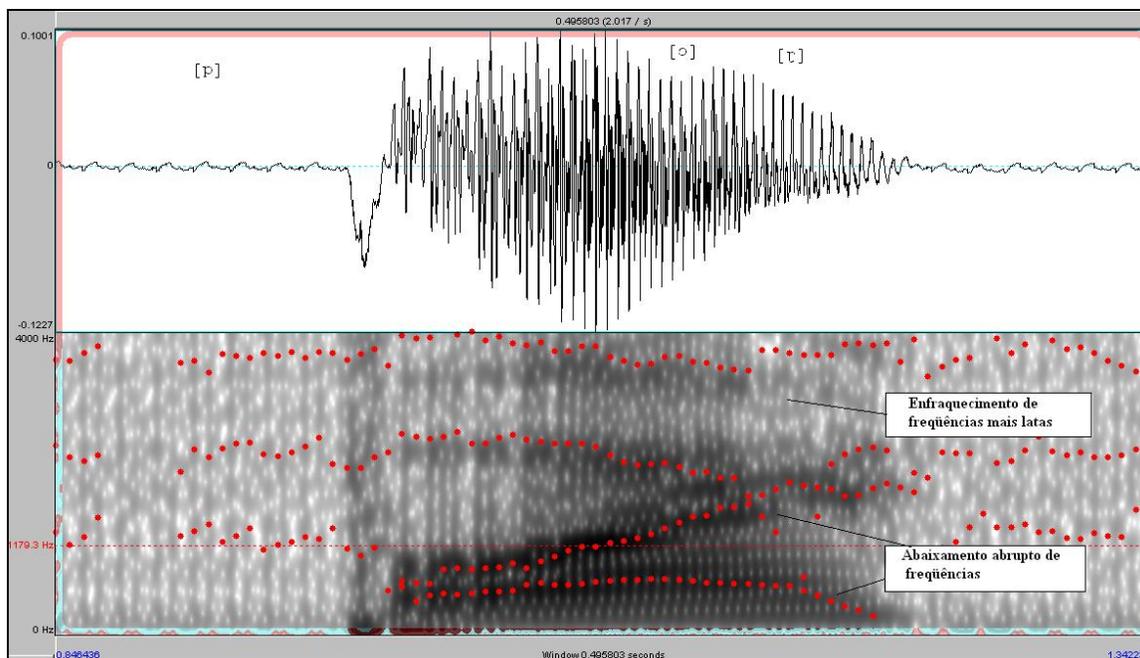


Figura 11. Forma de onda, espectrograma com indicação da região que corresponde à variante retroflexa na sílaba [pɣ] na palavra *porca* [ˈpɔɣkɐ].

➔ **Aproximante**

articulada com uma constrição que é maior do que a requerida para uma vogal, mas não radical o suficiente para produzir turbulência da corrente de ar.

➔ **Laterais:**

produzida com uma oclusão central, deixando que o ar escape pelas laterais do trato oral. O véu do palato encontra-se levantado, e o fluxo de ar passa apenas pela cavidade oral: **telha**, **lata**, **sal** (variante velarizada produzida em algumas regiões do Rio Grande do Sul)

Na Fig. 12, vemos as características acústicas de uma consoante lateral alveolar. Essas consoantes juntamente com os róticos (os sons de “r”) são chamadas de líquidas. Essa nomenclatura se deve a uma característica articulatória que combina uma oclusão e uma abertura do cavidade oral, de modo simultâneo como nas laterais e de modo sucessivo como nas vibrantes. As laterais caracterizam-se acusticamente por uma sonoridade próxima das vogais apresentando em seu espectro acústico uma estrutura de formantes bastante nítida, conforme se observa na tela da Fig. 12.

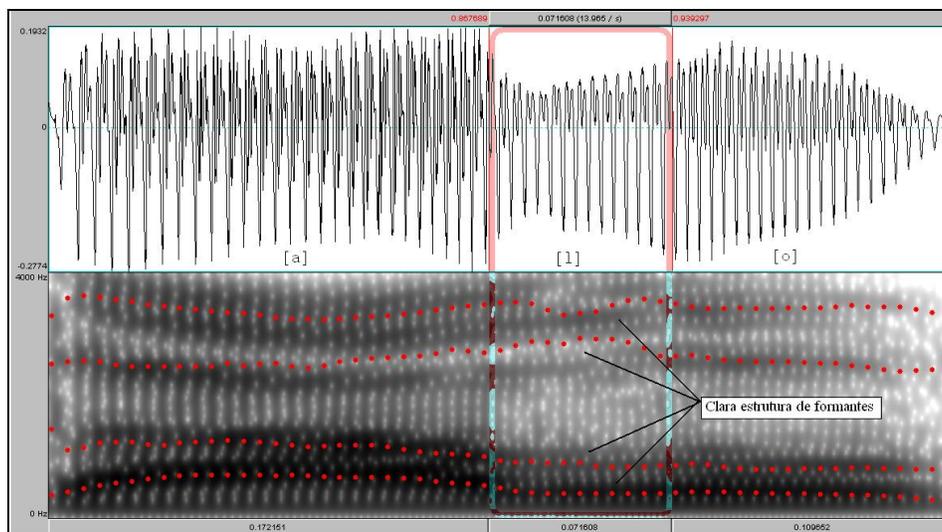


Figura 12. Forma de onda, espectrograma com indicação da região que corresponde à consoante lateral. Esta tela apresenta a seqüência de fonemas [alo]¹ na palavra *badalo* [ba'dalo]. A região marcada corresponde à consoante lateral alveolar.

Relações acústico-articulatórias relacionadas ao vozeamento

Observe a diferença articulatória entre sons consonantais surdos e sonoros na Figura 13. Agora, a partir de espectros acústicos do sinal de fala natural, vamos visualizar diferenças acústicas entre vogais e consoantes e também entre segmentos surdos e sonoros (Fig. 13).

O espectro de fala exibido na Fig. 13 foi concebido a partir do Programa Praat (já apresentado anteriormente). Nesta janela, podemos perceber a diferença entre vogais e consoantes. Nas vogais, dado que não há obstrução à passagem do fluxo de ar, observa-se uma grande quantidade de energia (maior amplitude dos pulsos glotais). Já, se observarmos as porções de tempo referentes às consoantes [k] e [b], veremos que há uma redução importante dessa energia (menor amplitude) nesses trechos. Isso se deve justamente às constrições necessárias a produção de consoantes. Outra diferença que se pode observar é concernente aos sons surdos e sonoros. Na região relativa à consoante [k], não há pulsos glotais, evidenciando a não vibração das pregas vocais. Já, na região referente à consoante [b], percebem-se os pulsos glotais (também presentes em todas as vogais pronunciadas), mostrando ser este um segmento consonantal sonoro.

¹ Nesta palavra, a vogal átona final foi produzida com média alta posterior e não com a alta. Isso deve ter acontecido em função do sinal de fala ser resultante de gravações de fala lida, uma vez que, no dialeto do sujeito gravado, essa posição não seria ocupada pela vogal média alta.

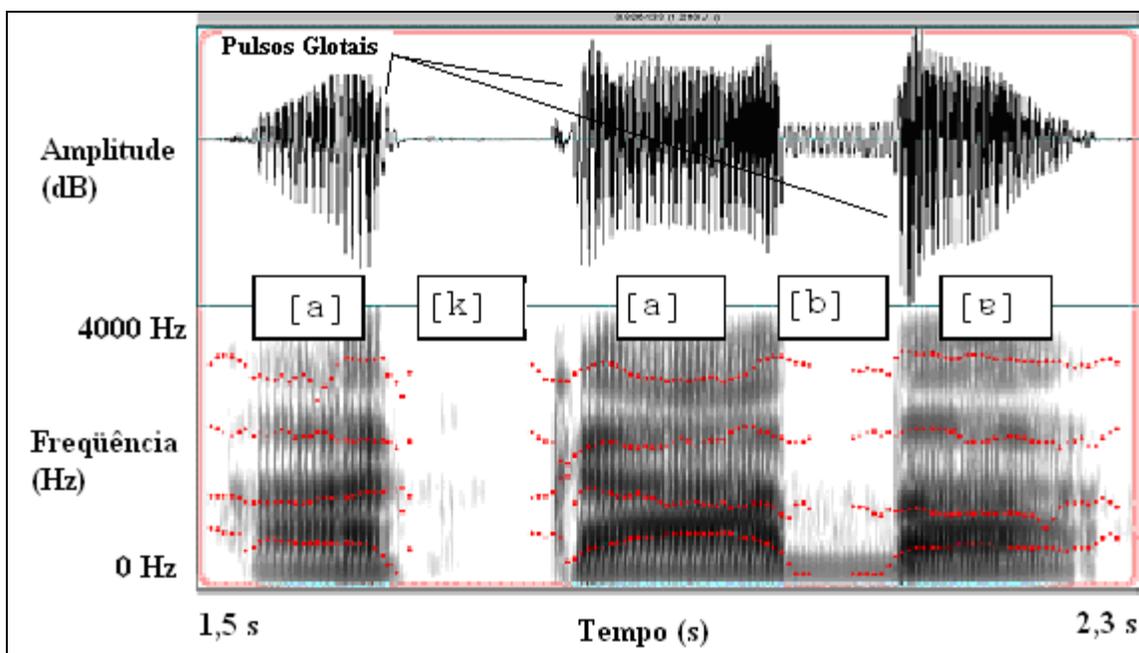


Figura 13. Forma de onda, espectrograma com indicação de formantes vocálicos da palavra *acaba* [a'kaba]

Prosódia e entoação

Apesar de muitos estudiosos não colocarem de maneira clara às áreas que cobrem os estudos da prosódia e da entoação e de muitas vezes chamarem os estudos sobre a entoação de prosódia, preferimos iniciar nossa conversa, deixando claro que, quando estivermos nos referindo a aspectos que dizem respeito a unidades maiores como as sentenças, estaremos tratando da entoação. Quando falarmos de prosódia, estaremos nos referindo a aspectos fônicos relativos ao acento de um dado segmento de fala dentro de uma sílaba ou vocábulo. Para a prosódia, três parâmetros acústicos são considerados: a intensidade, a curva de frequência fundamental (f_0) (*pitch*) e a duração. Na entoação, um importante parâmetro de análise é a curva de f_0 (contorno de *pitch*).

A definição de entoação apresentada em DUBOIS (1973, p. 217) parece deixar explícita esta distinção. Para ele, entoação são:

variações de altura do tom laríngeo que não incidem sobre um fonema ou sílaba, mas sobre uma seqüência mais longa (palavra, seqüência de palavras) e formam a curva melódica da frase. São utilizadas, na fonação, para veicular, fora da simples enunciação, informações complementares (...) reconhecidas pela gramática: a

interrogação (frase interrogativa), a cólera, a alegria (frase exclamativa), etc.

A prosódia é parte da fonologia que estuda os traços fônicos que se acrescentam aos sons da fala e que devem ser descritos com referência a um domínio maior do que um simples segmento. Dubois (1973) apresenta os três elementos estudados pela prosódia de uma maneira talvez mais clara: o acento dinâmico (de energia), relacionado à força com que o ar é expelido dos pulmões (intensidade), o acento de entoação (de altura), referente à frequência de f_0 (*pitch*) e a duração, relativa à sustentação sonora de um fonema.

Assim, pela prosódia veremos algumas diferenças observadas entre sílabas átonas e tônicas no PB. Parece ser consenso, se levarmos em conta somente a nossa experiência lingüística, que a sílaba tônica teria maior energia e maior duração do que as sílabas átonas. Então, vejamos tais parâmetros acústicos na elocução da palavra *alfabeto*.

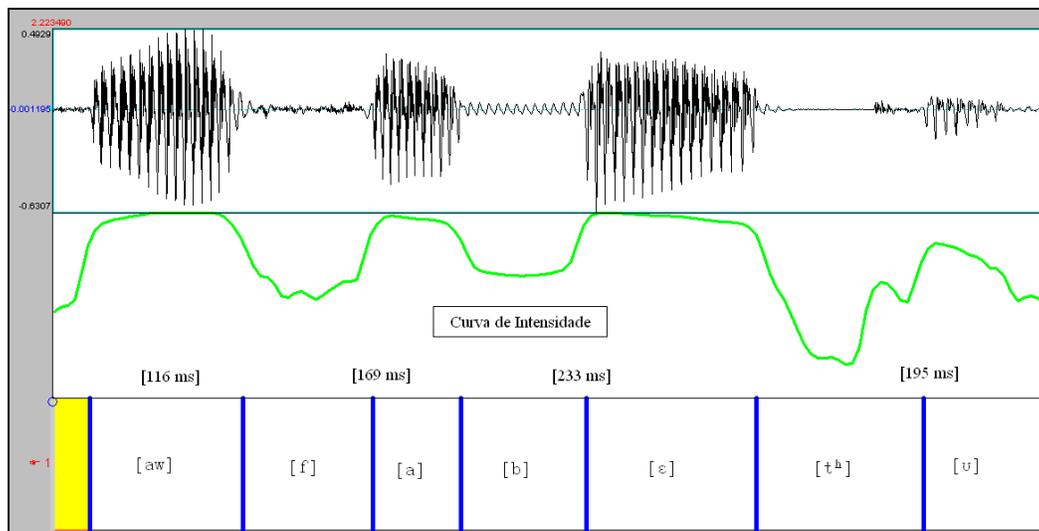


Figura 14. Curva de intensidade (energia) da palavra *alfabeto* [awfa'βetɔ] com a duração de cada sílaba elocucionada.

Pela Fig. 14, notamos que a curva de intensidade não é capaz de distinguir a sílaba tônica ['βɛ] das átonas pré-tônicas. A única que se diferencia das demais é a sílaba pós-tônica final [tʰ] que perde intensidade. Se observarmos o parâmetro duração, veremos que, aí sim, a sílaba tônica se diferencia das demais. Pode-se verificar ainda que a sílaba pós-tônica final tem duração maior do que as pré-tônicas.

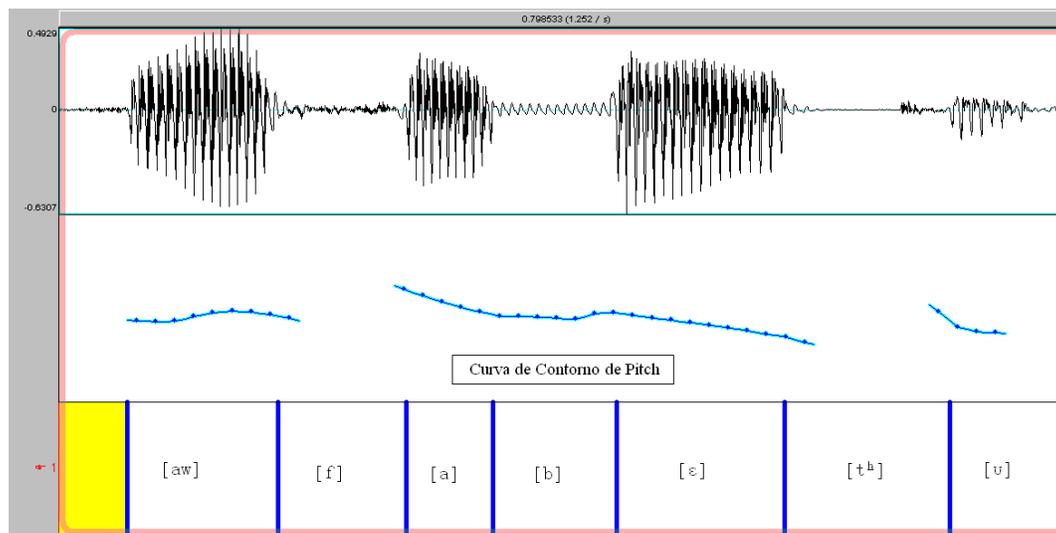


Figura 15. Curva de contorno de *pitch* da palavra *alfabeto*.

Pela Fig. 15, notamos que a curva de contorno de *pitch* sozinha também não seria um parâmetro seguro para a diferenciação entre sílabas acentuadas e não acentuadas. Parece-nos claro, todavia, que a duração é um parâmetro mais consistente nesta distinção. O que queremos mostrar com essas poucas e simples análises realizadas sobre a palavra *alfabeto* é que os parâmetros devem ser analisados em conjunto, levando em consideração o tipo de fala sobre o qual se faz a análise, visto que se tem observado uma grande diferença entre resultados obtidos sobre fala lida daqueles obtidos sobre a fala espontânea.

Agora, se o nosso objetivo é verificar as diferenças de entoação entre, por exemplo, uma frase declarativa e uma interrogativa, o contorno de *pitch* já é suficiente para tal distinção. Vejamos a Fig. 16. Na sentença declarativa, o contorno de *pitch* se mostra ascendente no início da sentença e cai ao final. Na interrogativa, ele se eleva no final da sentença, e cai abruptamente.

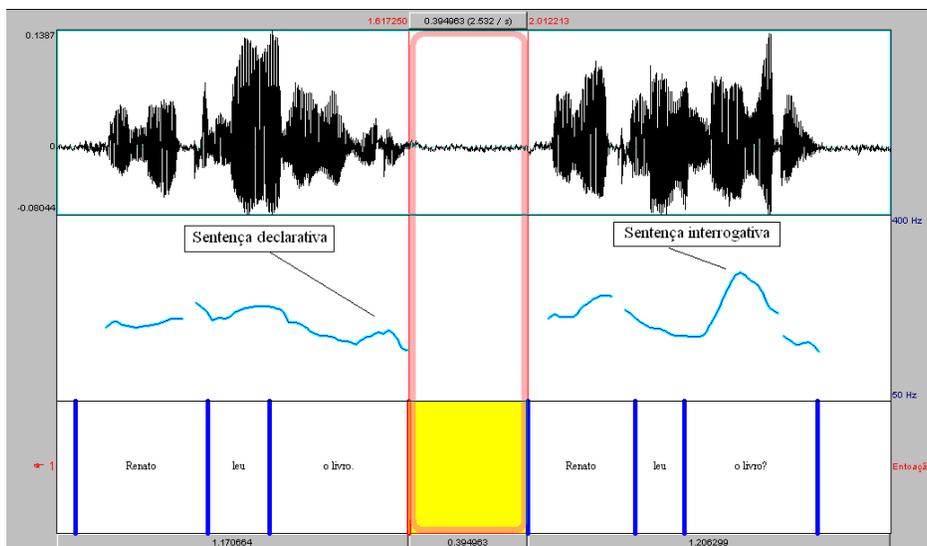


Figura 16. Curva de contorno de *pitch* da sentença *Renato leu o livro* na modalidade declarativa e na interrogativa .

Queremos deixar claro, porém, que nem todas as análises sobre a entoação são assim evidentes. Mas esses dados já nos dão idéia do que se pode tratar via entoação do PB.