

# UMA INTRODUÇÃO À ACÚSTICA DA VOZ CANTADA

**Maurílio Nunes Vieira**  
Departamento de Física/ICEx/UFMG  
e-mail: maurilio@fisica.ufmg.br

Resumo: Este artigo é dedicado àqueles não-familiarizados com as ciências exatas, mas que têm interesse na acústica da voz e do canto. Ele faz uma breve introdução ao modelo Fonte-Filtro da produção de vogais e o utiliza no estudo de alguns aspectos do canto lírico feminino, do canto lírico masculino, do canto harmônico e do vibrato vocal. O texto foi escrito de forma sucinta, mas valendo-se de detalhamento nas figuras e em suas legendas.

## 1. ACÚSTICA DA PRODUÇÃO DA FALA

A fala humana utiliza sons produzidos de diversas maneiras. As consoantes fricativas, como o [f] e o [s], são articuladas com um estreitamento na passagem do ar. Este estreitamento ocorre entre o lábio inferior e os dentes superiores, no [f], e entre a língua e o alvéolo, no [s]. Para gerar o ruído de fricção característico das fricativas, há um aumento na pressão expiratória, o que torna o fluxo turbulento. Já nas consoantes plosivas, ou oclusivas (e.g., [p] ou [t]), o som é criado, como os nomes sugerem, pela rápida liberação de uma massa de ar comprimida em uma cavidade intraoral. Em vogais, por sua vez, o som é produzido exclusivamente pela fonação, ou vozeamento, isto é, pela vibração das pregas vocais. Esta vibração, que pode ser percebida apalpando-se o pescoço, também ocorre em consoantes vozeadas ([v], [z], etc.) e é o traço distintivo entre pares como [s][z], [f][v], ou [p][b]. Há outras categorias de sons na fala (veja, e.g., Stevens, 1998; Kent & Read, 1992), mas o restante deste texto aborda apenas a produção das vogais orais e sua utilização na interpretação acústica de alguns fenômenos da voz cantada.

O painel superior da Fig. 1 destaca as características anatômicas mais relevantes na acústica das vogais. Sucintamente, há, durante a expiração, uma interação entre o ar proveniente da traquéia (T) e as pregas vocais, que vibram. Esta vibração é a *fonte* sonora que injeta pulsos de ar periodicamente na cavidade oral. Nas vogais orais, o trato nasal está acusticamente desacoplado devido à elevação do véu palatino (P) e a forma do trato vocal é determinada pela língua, lábios, mandíbula e posição vertical da laringe. As cavidades orais interagem sobre o som da vibração das pregas vocais, modificando-o de uma maneira específica para cada vogal. Por fim, o som resultante, filtrado pelo trato vocal, propaga-se dos lábios para o ambiente externo.

O processo descrito acima está esquematizado na parte inferior da Fig. 1, que apresenta o modelo Fonte-Filtro da produção de vogais (Fant, 1970). Este modelo separa os fenômenos acústicos em três partes supostamente independentes: a fonte sonora, o filtro acústico e a irradiação. A separação entre fonte e filtro é justificada porque a abertura (glote) por onde o ar flui entre as pregas vocais é muito menor que a seção transversal das cavidades supraglóticas.

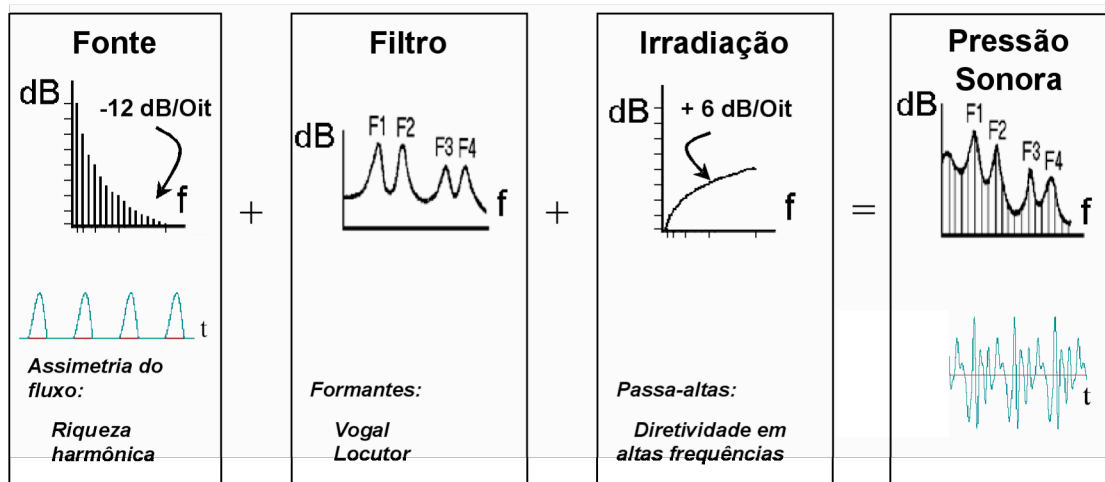
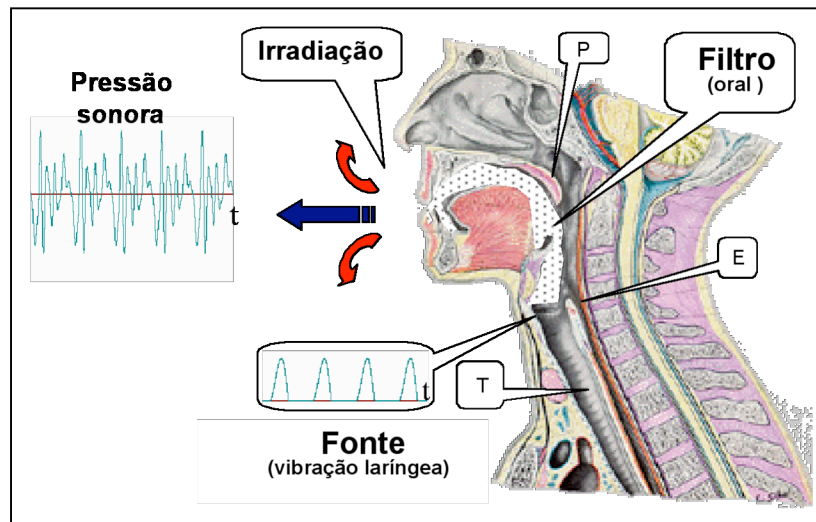


Figura 1: **Aparelho Fonador e o modelo Fonte-Filtro.** *Acima:* Estrutura anatômica do trato vocal. A elevação do pórtico velofaríngeo (P) desacopla as cavidades nasais. Note a proximidade da entrada do esôfago (E) com a região das pregas vocais, situadas acima da traquéia (T). Os pulsos gerados na fonação são filtrados pelas cavidades acústicas do trato vocal, e propagam-se pelo ar. Na irradiação, como sugere o desenho, parte do som (baixas frequências) espalha-se enquanto outra parte (altas frequências) propaga-se para frente. *Abaixo* (adaptado de Fant, 1970): modularização introduzida pelo modelo Fonte-Filtro (em alguns espectros há divisões em intervalos de 6 dB, no eixo vertical, e em oitavas, no horizontal). A fonte é representada por uma série harmônica, o filtro por formantes ( $F1, F2$ , etc.), e a irradiação pela sua característica passa-altas. Comparando-se o oscilograma do fluxo com o da pressão sonora vêm-se, de forma desordenada, as ondulações introduzidas pelos formantes, mas, no espectro da pressão irradiada, distingue-se cada um dos formantes.

O modelo Fonte-Filtro utiliza uma descrição dos fenômenos acústicos no domínio das frequências. O filtro oral é caracterizado por picos ( $F1, F2, F3$  e  $F4$ , na Fig. 1) no espectro de frequências. Estes picos, que correspondem aos modos normais dos tubos acústicos, são tradicionalmente chamados de *formantes*. Os três primeiros formantes ( $F1, F2$  e  $F3$ ) têm menor dependência com o locutor e prestam-se, principalmente, para diferenciar as vogais. Os formantes superiores ( $F4, F5$ , etc.), por outro lado, têm menor conteúdo lingüístico e maior variação com o locutor. Acusticamente, os formantes amplificam seletivamente os harmônicos gerados pela vibração laringea.

Para compreender as características do espectro da *fonte* sonora é importante discutir a vibração das pregas vocais. Cada prega (Fig. 2) é uma dobra de tecidos que, mecanicamente, comporta-se como um *corpo* (músculo e ligamento) relativamente rígido e pesado, revestido por uma *cobertura* (tecido conjuntivo e epitelial) flexível (Hirano, 1981). O fluxo aéreo é modulado à medida que o espaço entre as pregas vocais é aberto e fechado ciclicamente, de acordo com a deformação da camada de cobertura (Fig. 3). A vibração glótica dá-se de forma aproximadamente periódica, mas a velocidade do fechamento de cada ciclo é maior que a velocidade de abertura. Assim, o fluxo é inclinado para a direita. Se os pulsos glóticos fossem perfeitamente senoidais, haveria apenas uma

componente harmônica e não haveria excitação de múltiplos formantes, não sendo possível a produção de vogais diferentes. Devido à *assimetria* do pulso, o espectro passa a possuir não apenas a componente fundamental,  $F_0$ , mas uma *série harmônica* (painel inferior, à esquerda, da Fig. 1) onde a frequência de cada termo é um múltiplo inteiro de  $F_0$ . Para a voz modal, esta que a quase totalidade das pessoas utiliza na conversação normal, a amplitude dos termos da série harmônica cai, de forma aproximada, a uma proporção de 12 dB por oitava, isto é, a amplitude do harmônico de ordem  $n$  é apenas 25% da amplitude do harmônico de ordem  $n/2$ . Ajustando-se a musculatura laríngea é possível gerar-se outras vozes (e.g., falsete, fry) com características acústicas e perceptivas marcadamente diferentes da voz modal (Fig. 4).

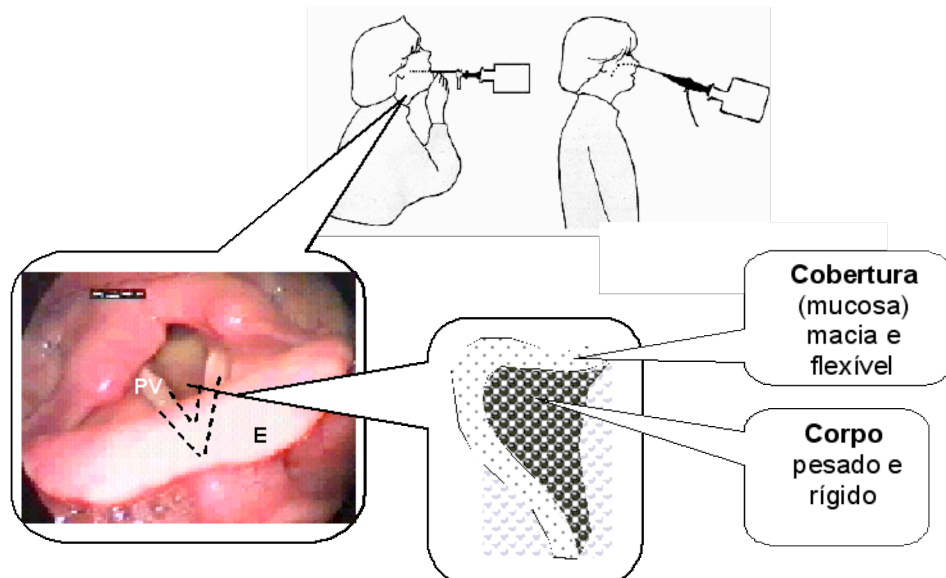


Figura 2: **Pregas vocais.** *Acima* (de Södersten & Lindestad, 1992): endoscopia oral (direita) e nasal (esquerda). *Abaixo*, à esquerda: visão endoscópica das pregas vocais (PV). As linhas tracejadas indicam a posição das pregas sob a epiglote (E). *Abaixo*, à direita: corte transversal de uma prega vocal, destacando as duas camadas mais relevantes para a fonação: o *corpo* e a *cobertura*.

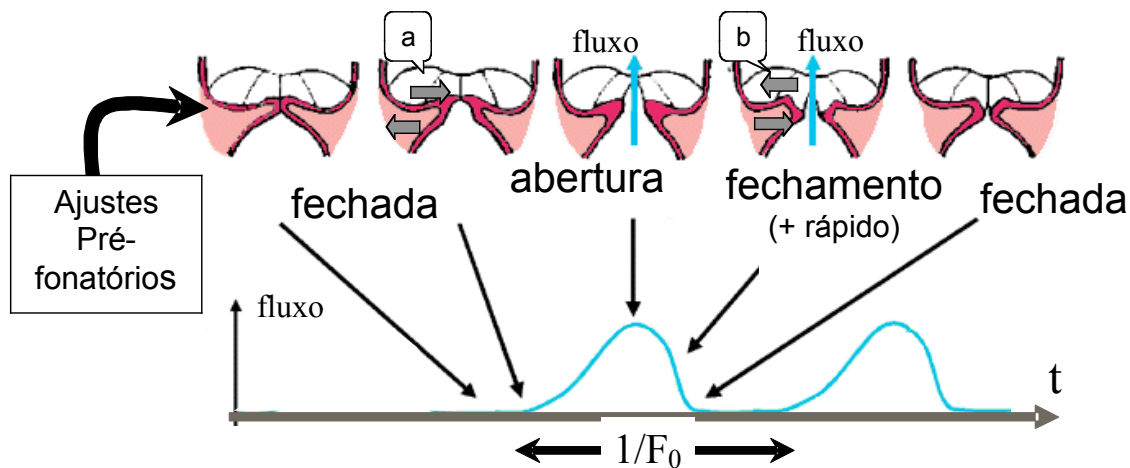


Figura 3: **Ciclo fonatório na voz modal** (adaptado de Hirano, 1981). A partir da esquerda: as pregas vocais estão inicialmente fechadas e, dependendo da frequência e intensidade do som a ser produzido, têm ajustes pré-fonatórios adequados de pressão subglótica, tensão longitudinal, aproximação da parte posterior das pregas vocais e força de compressão na parte medial. Com o esforço expiratório e com a glote ainda fechada, a pressão intraglótica aumenta enquanto as bordas inferiores afastam-se, acumulando energia potencial elástica na camada de cobertura. Com o aumento da pressão intraglótica, as bordas superiores finalmente separam e o ar flui pela glote. O fluxo de ar leva a uma queda na pressão intraglótica (Efeito Bernoulli) que ocorre num momento em que as bordas inferiores estão muito comprimidas. Isto resulta numa fase de fechamento mais rápida que a de abertura, causando uma assimetria na forma dos pulsos glóticos. O ciclo repete-se na frequência fundamental  $F_0$ . Os detalhes (a) e (b) mostram a oposição entre os sentidos dos movimentos das bordas superiores e inferiores. Esta diferença de fase está associada a uma onda (onda mucosa) que se propaga pela cobertura das pregas vocais desde sua parte inferior até sua parte dorsal. Doenças que dificultam ou interrompem o deslocamento deste movimento muco-ondulatório podem levar à rouquidão.

Com relação à irradiação, as baixas frequências tendem a espalhar-se por difração, enquanto as altas frequências propagam-se de forma mais direta. Em outras palavras, a pressão sonora numa região mais distante do locutor tende a ter uma maior intensidade relativa nas altas frequências. Pode-se dizer, assim, que a irradiação é equivalente a um filtro passa-altas. Este filtro pode ser representado, de forma simplificada (Fig. 1), por um ganho de 6 dB/oitava, isto é, a resposta dobra (+6 dB) cada vez que a frequência duplica (uma oitava). O estudo da irradiação sonora é relativamente complexo (e.g., Kinsler *et al.*, 1982; Fant, 1970) e a simplificação acima, que é adequada até cerca de 6 kHz, vem da analogia com um pistão (a massa de ar entre os lábios) vibrando à frente de uma esfera refletora (a face).

Tendo apresentado as várias facetas do modelo Fonte-Filtro para a produção de vogais, as discussões seguintes aplicarão estes conceitos a situações de interesse na voz cantada.

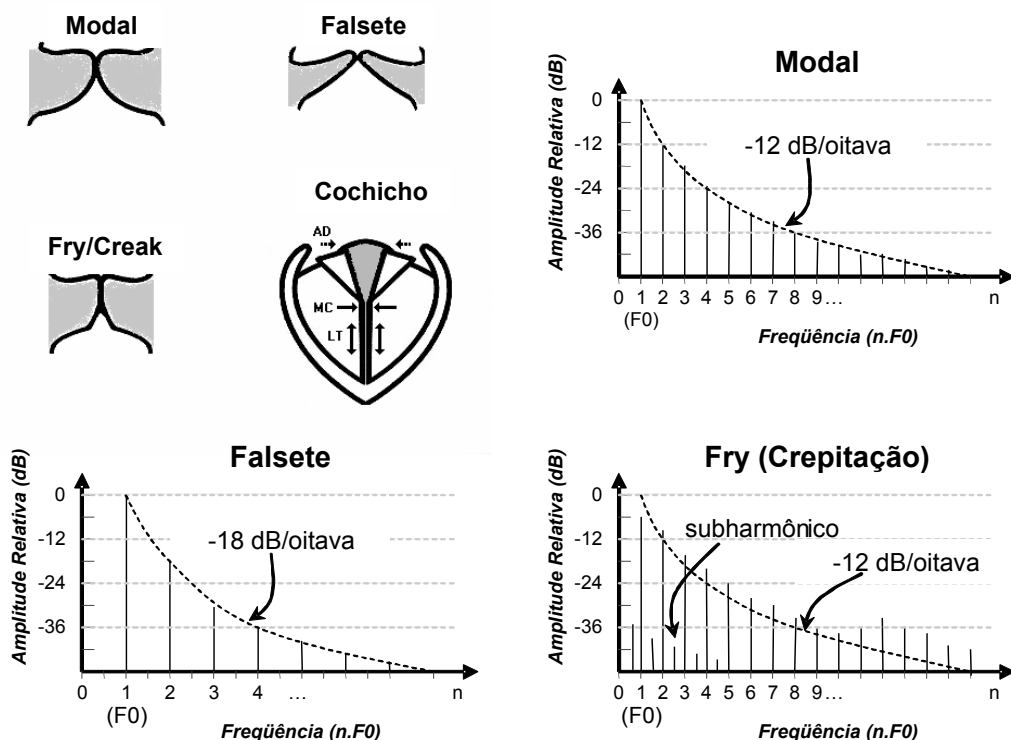


Figura 4: **Ajustes laríngeos e qualidade da voz.** *Acima, à esquerda* (baseado em Beck, 1988): formato da prega vocal em diversos ajustes laríngeos. Tomando-se a voz modal como referência e seguindo a descrição de Laver (1981), há, no falsete, grande aumento na tensão longitudinal (LT) e, conseqüentemente, em F0; uma fenda glótica é comum neste ajuste. No *fry* ou *creak* (denominado crepitação, pelos fonoaudiólogos, ou voz rangida, por alguns lingüistas), há grande redução na tensão longitudinal e na pressão subglótica, mas há aumento na compressão medial (MC), resultando num som mais grave, com aperiodicidade e subharmônicos. No *cochicho*, a falta de aproximação ou adução (AD) na parte posterior da glote resulta numa fenda onde é gerado um ruído de fricção; neste ajuste, não há vibração das pregas vocais. *Demais gráficos* (baseados em Klatt & Klatt, 1990): principais características acústicas de ajustes laríngeos, destacando as mudanças na frequência fundamental e na envoltória dos termos da série harmônica.

## 2. SINTONIA F1-F0 NO CANTO LÍRICO FEMININO

O canto lírico feminino requer sons bem mais agudos que os utilizados para a conversação. Enquanto o valor médio de F0 na fala feminina adulta está em torno de 220 Hz, as notas emitidas por uma soprano podem estar próximas a 1024 Hz (dó<sub>5</sub>). No canto, porém, à medida que a fundamental aumenta, aumenta também o espaçamento entre os termos da série harmônica gerada na laringe, diminuindo a concentração de harmônicos em torno dos formantes e perdendo-se os ganhos da ressonância com os formantes.

A Fig. 5 ilustra o problema. Hipoteticamente, uma soprano está emitindo uma nota em 500 Hz, mas o harmônico fundamental está entre o primeiro e o segundo formantes da vogal articulada ([u]). Desta forma, a amplitude do harmônico fundamental é relativamente pequena e a voz tem pouca audibilidade. A técnica utilizada para aumentar a amplitude do harmônico fundamental é ajustar o

primeiro formante (F1), fazendo-o coincidir com F0, através do abaixando da mandíbula, gesto que tem certa seletividade sobre F1 (Lindblom & Sundberg, 1971). Com isso, o harmônico fundamental aumenta pela ressonância com F1, resultando num ganho na audibilidade da voz sem que haja aumento na energia utilizada na fonação.

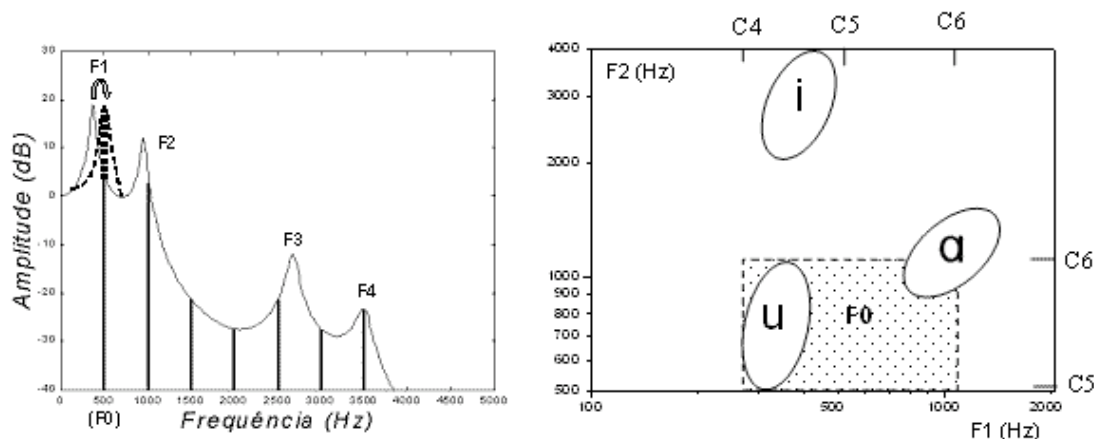


Figura 5: **Sintonia F1-F0.** *Esquerda:* aumento na amplitude do harmônico fundamental, F0, pelo deslocamento do primeiro formante, F1. *Direita* (adaptado de Rossing, 1990): representação de vogais no plano F1x F2; as elipses indicam valores médios de formantes e a área pontilhada é a região da frequência fundamental de sopranos:  $d\acute{o}_4$  (C4) a  $d\acute{o}_6$  (C6), aproximadamente). As demais vogais, não representadas, ocupam regiões intermediárias do plano, com superposições de áreas.

Como descrito por Fant (1986), outro benefício da sintonia F1-F0 é uma diminuição no consumo de ar devido a um grande aumento na pressão intraoral próximo ao instante de máxima abertura glótica. Este aumento de pressão supraglótica reduz o fluxo, mas não afeta a velocidade de fechamento.

O preço que se paga pela sintonia F1-F0 é uma perda na inteligibilidade das vogais. Segundo Di Carlo (1994), que estudou as vogais do Francês, somente as vogais nasais e as orais [e], [ɛ] e [o] são nitidamente individualizadas para notas emitidas entre o  $f\acute{a}_3$  e o  $l\acute{a}_3$  (350-440 Hz), que ela denomina *zona de tolerância*; entre o  $l\acute{a}_3$  e o  $mi_4$  (440-659 Hz), *zona de inteligibilidade eletiva*, somente [i] e [a] são diferenciados; acima do  $mi_4$ , *zona de não-inteligibilidade absoluta*, não se pode distinguir uma vogal das outras.

No canto lírico, não apenas de sopranos, outras distorções nos valores de formantes podem ocorrer. Os ajustes utilizados para alargar o trato vocal e diminuir a tensão da musculatura buco-facial não só alteram as ressonâncias como também limitam a amplitude dos gestos articulatórios, reduzindo as excursões temporais dos formantes, particularmente F1, que são elementos importantes da inteligibilidade (Di Carlo, 1994). Esta autora lembra ainda que consoantes mal articuladas pelo prolongamento excessivo das vogais também colaboram para a reduzir a inteligibilidade.

### 3. FORMANTE DO CANTOR CANTO LÍRICO MASCULINO

No canto lírico masculino as frequências mais agudas estão em torno do  $mi_3$  (330 Hz) para baixos,  $sol_3$  (392 Hz) para barítonos e  $d\acute{o}_4$  (523 Hz) para tenores, ao passo que o F0 médio na voz falada masculina é de 120 Hz. Assim, ao longo da tessitura dos cantores, há menor possibilidade de o harmônico fundamental localizar-se fora da região de ação do primeiro formante.

Os problemas no canto masculino surgem pelo ofuscamento da voz pelo som da orquestra. A técnica lírica vale-se, contudo, de ajustes há muito observados em cantores de ópera (e.g., Barhlomew, 1933), cujas vozes apresentam um pico espectral intenso e largo em torno de 3 kHz (Fig. 6). Este pico eleva a amplitude relativa dos harmônicos aí situados, destacando a voz sobre o som da orquestra. A produção deste pico, denominado *formante do cantor*, está relacionada com o abaixamento da laringe e o alargamento da cavidade faríngea, embora esta não seja o único mecanismo existente (Sundberg, 1979). Como descrito por este autor, o formante do cantor é uma aglutinação dos formantes superiores (F3, F4 e F5) devido às manobras citadas. No jargão dos cantores, esta técnica recebe nomes como *impostação* ou *projeção*.

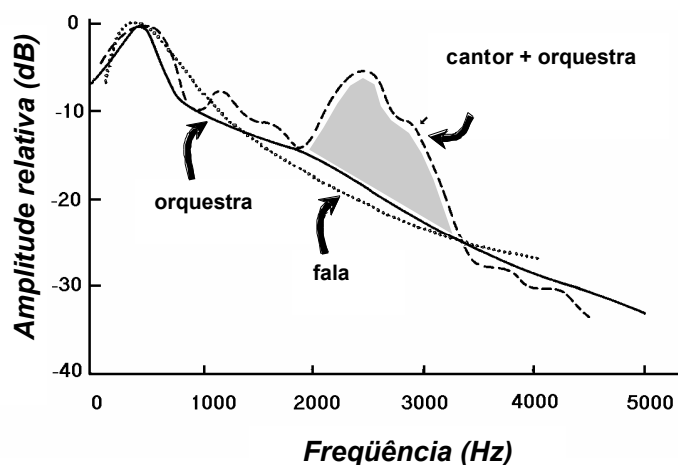


Figura 6: **Formante de cantor.** Espectros médios de longo prazo. Sem o recurso do formante do cantor (região escura), os sons da fala são mascarados pela orquestra. Com o abaixamento da laringe e o alargamento da região supraglótica, o formante do cantor é criado próximo a 3 kHz. Adaptado de Sundberg (1979).

Assim como na sintonia F1-F0 das sopranos, o formante do cantor propicia um aumento na audibilidade da voz sem aumento no esforço fonatório. Além disso, como lembra Sundberg (1979), há outras conseqüências favoráveis. Primeiro, a região de 3 kHz também coincide com a região mais sensível da percepção auditiva humana (onde ocorre uma ressonância com a cavidade do meato acústico do ouvido), sendo amplificada ainda mais pelo processamento auditivo periférico. Segundo, tendo-se em mente o comportamento passa-altas da irradiação sonora, os harmônicos de alta frequência propagam-se mais eficientemente em linha reta, o que pode auxiliar a platéia a discriminar melhor a voz do cantor no palco.

#### 4. CANTO HARMÔNICO

O recurso de destacar harmônicos é utilizado com maestria numa técnica de canto de regiões asiáticas, incluindo a Mongólia, o Tibet e a Sibéria (Levin & Edgerton, 1999). Chamada de “canto harmônico” (*harmonic chant/singing*), este estilo possui outras denominações que fazem alusão a sobretons (*overtone singing*) ou à garganta (*throat singing*). A primeira sensação que se tem ao ouvir esta manifestação artística é de certo espanto, por tratar-se de algo exótico, e vários ouvintes costumam admitir tratar-se de um som produzido por uma única pessoa.

Neste estilo (Fig. 7), o cantor emite um som laríngeo com uma frequência fundamental baixa (70-100 Hz) que permanece praticamente constante. Porém, refinando os ajustes utilizados na produção de vogais, ele consegue criar uma pequena cavidade intraoral que atua como um “superformante” de elevada amplitude e largura de faixa tão pequena que é capaz de destacar um único harmônico do espectro laríngeo. O cantor altera com precisão a forma da cavidade, mudando a frequência do formante e, literalmente, pinça harmônicos no espectro, criando uma seqüência melódica aguda que evolui sobre uma base grave. Detalhes sobre alguns ajustes articulatórios, arquivos de vídeo e áudio são disponíveis nos suplementos de Levin & Edgerton (1999).

A Fig. 7 mostra um exemplo de canto harmônico da região de Tuva, na Sibéria. No espectrograma de banda estreita, as linhas horizontais paralelas representam a evolução temporal da série harmônica; quanto mais escura a linha, mais intenso o harmônico. No início do espectrograma, o cantor utiliza uma voz sem os recursos do canto harmônico, como mostra a seção espectral “A”, onde vários formantes “convencionais” podem ser vistos. Assim que a técnica do canto harmônico é introduzida, um harmônico de alta frequência (seta) passa a destacar-se no espectro. Os harmônicos utilizados estão entre o 6<sup>o</sup> e o 12<sup>o</sup>, mas, como a sintaxe musical da região favorece melodias pentatônicas, o 7<sup>o</sup> e 11<sup>o</sup> harmônicos são evitados (Levin & Edgerton, 1999).



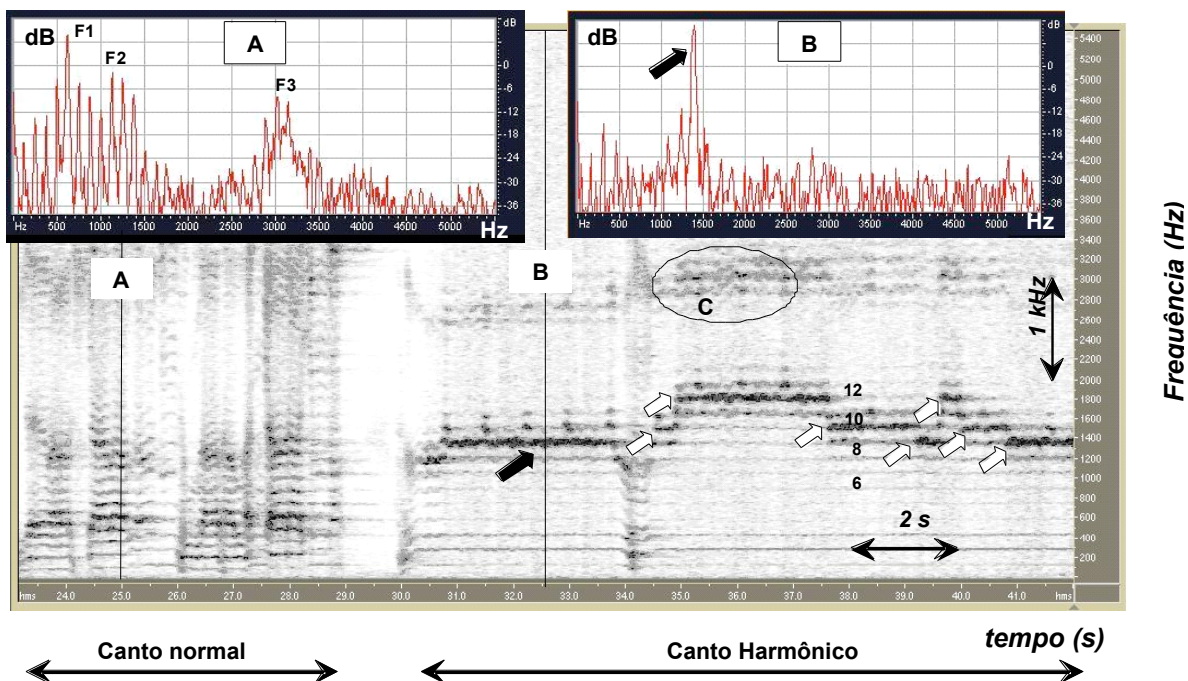


Figura 7: **Canto harmônico**. Na parte inicial, a técnica do canto harmônico não é utilizada. A seção espectral “A” mostra os formantes característicos da fala. Com o uso do canto harmônico, um harmônico (setas) destaca-se, como visto ao longo do espectrograma e na seção “B”. Os harmônicos selecionáveis estão entre o 6<sup>o</sup> e o 12<sup>o</sup>, mas o 7<sup>o</sup> e o 11<sup>o</sup> não são utilizados devido às escalas pentatônicas. O detalhe “C” indica uma ocorrência de vibrato. Elaborado com gravações de Levin & Edgerton (1999).

Este estilo também utiliza recursos fonatórios especiais e por isso é também denominada *khoomii*, palavra mongólica para “garganta”. Os ajustes laríngeos adotados têm semelhanças com o *fry* (Fig. 4). Com isso, além do baixo valor de F0, consegue-se uma maior velocidade de fechamento (que amplifica os harmônicos superiores) e uma diminuição da fase aberta de cada ciclo glótico (que diminui o consumo de ar e evita perda de energia pelo acoplamento com o trato vocal subglótico); as falsas pregas também podem ser utilizadas vibrando numa frequência sub-múltipla de F0 (Levin & Edgerton, 1999).

No espectrograma da Fig. 7 vê-se, no detalhe “C”, que as linhas horizontais não são retas, mas ondulam ao longo do tempo. Estas ondulações são conseqüências do vibrato, que será discutido a seguir.

## 5. VIBRATO VOCAL

O vibrato é uma variação regular na frequência de uma nota comumente utilizado na música (Fig. 8). Perceptivamente, sua qualidade depende de três parâmetros mensuráveis: a taxa, a profundidade e a regularidade. A taxa ( $F = 1/T$ ) é o número de ciclos por segundo; a profundidade, amplitude ou extensão ( $\pm\Delta F_0$ ), é o desvio da frequência em torno de ser valor médio, enquanto a regularidade carece de um parâmetro para medição, mas, desejavelmente, deve ter um aspecto senoidal. No *bel canto*, a taxa aceitável do vibrato varia de 5,5 Hz a 7,5 Hz e a profundidade vai de  $\pm 6\%$  a  $\pm 12\%$ , ou seja, de aproximadamente  $\pm 1$  a  $\pm 2$  semitons (Sundberg, 1987).

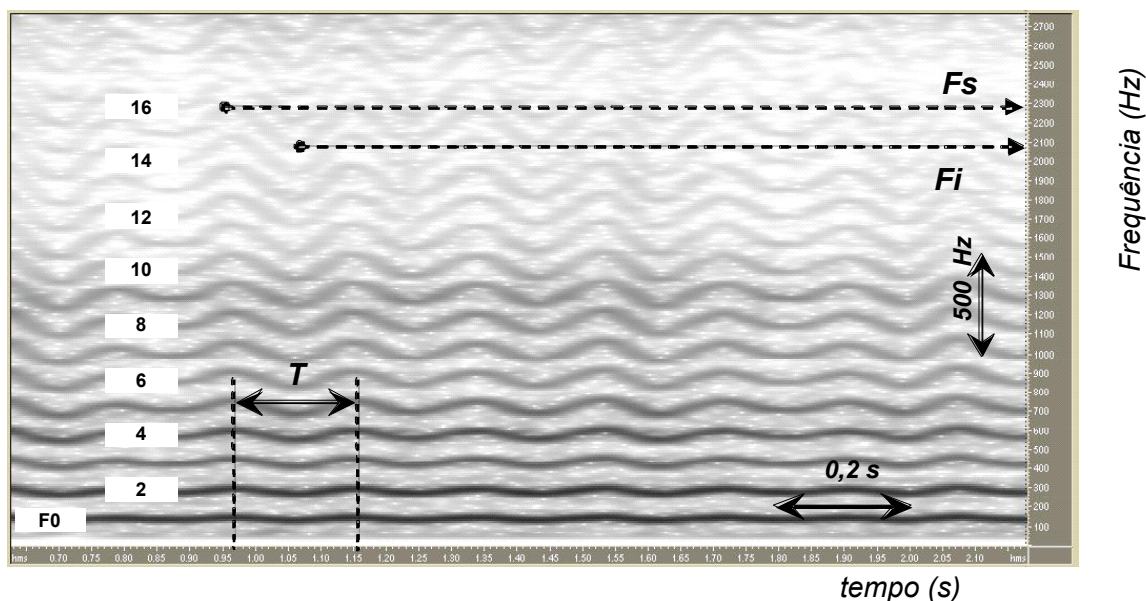


Figura 8: **Vibrato**. Espectrograma de banda estreita mostrando as ondulações introduzidas em toda a série harmônica devido ao vibrato. Os números crescentes à esquerda da figura (2, 4, 6, ...) são a ordem do harmônico. O período do vibrato é  $T \approx 0,19$  s e a taxa é  $F = 1/0,19$  s = 5,3 Hz. O 15<sup>o</sup> é utilizado para medir a amplitude do vibrato: sua frequência inferior é  $F_i = 2080$  Hz e a superior é  $F_s = 2280$  Hz. O valor médio de  $F_0$  é  $1/2 (2280 + 2080)/15 = 145,3$  Hz e a amplitude do vibrato é  $1/2 (2280 - 2080)/15 = \pm 6,67$  Hz, ou seja,  $\pm 4,6\%$  de  $F_0$ , que equivale a 0,8 semitons (vide texto principal).

A taxa e a profundidade do vibrato podem ser medidas em espectrogramas de banda estreita (Fig. 8) utilizando-se algum software que disponha de escalas e/ou cursores com resolução adequada. A leitura da duração do ciclo ( $T$ ) é imediata, mas, para estimar profundidade, é difícil obter-se resolução adequada medindo-se diretamente a frequência do harmônico fundamental. Porém, para um harmônico de ordem elevada, podem-se obter valores de frequência máxima e mínima com erros menores, e a profundidade do vibrato (valor de pico) pode ser expressa, em semitons, por:

$$\Delta F_0 = \pm 17,31 \times \ln \left( \frac{2F_s}{F_s + F_i} \right) \text{ [semitons]},$$

onde  $F_s$  e  $F_i$  são, respectivamente, os valores superior e inferior da frequência de um harmônico qualquer. No exemplo da Fig. 8,  $F_s \approx 2280$  Hz,  $F_i \approx 2080$ , donde  $\Delta F_0 \approx \pm 0,8$  semitons. Tendo-se em mente que o semitom da escala igualmente temperada equivale a uma relação entre frequências de  $\sqrt[12]{2} = 1,059... \approx 6\%$ , a amplitude do vibrato pode ser calculada sem o uso de logaritmos, obtendo-se  $\Delta F_0$  em valor percentual e obtendo-se a fração do semitom. No caso (vide legenda da Fig. 8),  $\Delta F_0 = 6,7/145,3 = 4,6\%$  e a amplitude do vibrato é  $\pm 4,6\%/6\% = \pm 0,8$  semitons.

No canto, o vibrato é produzido modulando-se a tensão da musculatura laríngea responsável pela tensão e enrijecimento das pregas vocais (músculos cricotireóideos e tiroaritenóideos, respectivamente); também é fisiologicamente possível produzir o vibrato variando-se a pressão subglótica (que resulta em maior deslocamento lateral das pregas vocais e, conseqüentemente, na elevação de sua tensão média), mas esta parece não ser a técnica natural de cantores experientes (Sundberg, 1979; Titze, 1994).

O vibrato produzido por variação da pressão pulmonar revela um mecanismo de interdependência entre a amplitude (causa) e a frequência (efeito). Por outros motivos, ocorre uma situação reversa no vibrato laríngeo, isto é, variações em  $F_0$  causam variações na amplitude da pressão sonora. Como ilustrado na Fig. 9, isto ocorre porque a amplitude dos harmônicos aumenta e diminui à medida que suas frequências oscilam em torno dos formantes, particularmente de  $F_1$ , que é o formante de maior amplitude (Sundberg, 1979).

A regularidade do vibrato é um aspecto que tem recebido atenção nos últimos anos (e.g., Diaz & Rothman, 2002). A regularidade pode ser analisada fazendo-se uma análise espectral da série temporal de valores de  $F_0$ , como ilustrado na Fig. 9. Este método pode ser adaptado para criar um espectrograma do vibrato e avaliar a dinâmica temporal dos ajustes.



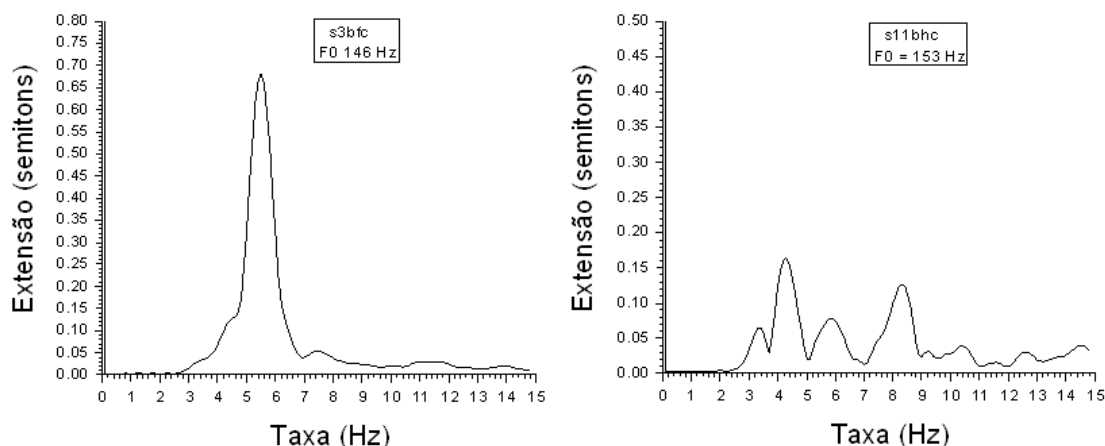


Figura 9: **Espectro da regularidade do vibrato.** *Esquerda:* voz com vibrato regular, taxa de aproximadamente 5,5 Hz e amplitude de 0,70 semitons. *Direita:* vibrato irregular e de baixa amplitude, apresentando dois picos maiores em 4,2 e 8,4 Hz. Os gráficos foram obtidos rastreando-se F0 a cada 20 ms, extraíndo-se a média da série temporal, passando por um filtro passa-altas com corte em 3 Hz e calculando-se a Transformada Discreta de Fourier em 100 pontos.

## 6. CONCLUSÃO

Este artigo abordou, de forma introdutória, o modelo Fonte-Filtro da produção da voz e sua utilização no estudo de alguns aspectos da voz cantada. A arte do canto tem estado à frente da ciência, mas aos poucos, a acústica vem explicando fenômenos que professores de canto conhecem há séculos. Com o avanço da tecnologia, o uso não só de registros acústicos, mas de sinais aerodinâmicos (fluxo aéreo e pressão sonora) e eletroglotográficos, tende a tornar-se um forte aliado à compreensão e evolução das técnicas de canto, como tem ocorrido na fonoaudiologia e na otorrinolaringologia (e.g., Hirano, 1981; Vieira, 1997; Vieira *et al.*, 2002). Estas especialidades médicas, aliás, têm tido crescente interesse no estudo da voz cantada visando o aprimoramento do tratamento de problemas vocais em cantores profissionais. Cabe aqui uma palavra de precaução. A facilidade atual de acesso a equipamentos de gravação e a *softwares* para a análise automática da voz (com resultados não-necessariamente confiáveis) não deve ocultar a necessidade de conhecimentos sobre os fenômenos básicos da acústica, as técnicas de gravação e os métodos de análise de sinais.

## 7. REFERÊNCIAS

- Bartholomew, W. T. (1933). A physical definition of "good voice-quality" in the male voice. *Journal of the Acoustical Society of America* 6, pp. 25-33.
- Beck, J. M. (1988). *Organic Variation and Voice Quality*. Ph.D. Thesis (University of Edinburgh).
- Diaz, J. A. & Rothman, H. B. (2002). Acoustical comparison between samples of good and poor vibrato in singers. *Journal of voice* 17, pp. 179-184.
- Di Carlo, N. S. (1994). L'intelligibilité de la voix chantée. *Medicine des Art* 10, pp. 2-15.
- Fant, G. (1970). *Acoustic Theory of Speech Production*. Mouton, The Hague.
- Fant, G. (1986). Glottal flow: models and interaction. *Journal of Phonetics* 14, pp. 393-399.
- Hirano, M. (1981). *Clinical Examination of Voice*. Springer-Verlag, New York.
- Kent, R. D. & Read, C. (1992). *The Acoustic Analysis of Speech*. Singular, San Diego.
- Kinsler, L. E., Frey, A. R., Coppers, A. B. & Sanders, J. V. (1982). *Fundamentals of Acoustics (Third Edition)*. John Wiley & Sons, New York.
- Klatt, D. H. & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87, pp. 820-857.
- Laver, J. (1981). *The Phonetic Description of Voice Quality*. Cambridge University Press.
- Levin, T. C & Edgerton, M. E. (1999). The throat singers of Tuva. *Scientific American* 281(3), pp. 80-82 + s.
- Lindblom, B. E. F. & Sundberg, J. E. F. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movements. *Journal of the Acoustical Society of America* 50, pp. 1166-1179.
- Rossing, T. (1990). *The Science of Sound, 2<sup>nd</sup> Edition*. Addison Wesley.
- Södersten, M. & Lindestad, P. A. (1992). A comparison of vocal fold closure in rigid telescopic and flexible fiberoptic laryngostroboscopy. *Acta Otolaryngologica (Stockholm)* 112, pp. 144-150.
- Stevens, K. N. (1998). *Acoustic Phonetics*. The MIT Press, Cambridge.

- Sundberg, J. (1979). Perception of singing. *Speech Transmission Laboratory Quarterly Progress and Status Report 1* (Royal Institute of Technology, Stockholm) pp. 1-48.
- Sundberg, J. (1987). *The Science of the Singing Voice*. Dekalb, Illinois.
- Södersten, M. & Lindestad, P. A. (1992). A comparison of vocal fold closure in rigid telescopic and flexible fiberoptic laryngostroboscopy. *Acta Otolaryngologica (Stockholm)* 112, pp. 144-150.
- Titze, I. R. (1994). *Principles of Voice Production*. Prentice-Hall, Englewood Cliffs.
- Vieira, M. N. (1997). *Automated Measures of Dysphonias and the Phonatory Effects of Asymmetries in the Posterior Larynx*, Ph.D Thesis, University of Edinburgh, UK.
- Vieira, M. N., McInnes, F. R., & Jack, M. (2002). On the influence of laryngeal pathologies on acoustic and electroglottographic jitter measures. *Journal of the Acoustical Society of America* 111, pp. 1045-1055.