

ASPECTOS FONÉTICO-ACÚSTICOS DA CANÇÃO ERUDITA BRASILEIRA *

Beatriz Raposo de MEDEIROS

SUMMARY *This work is a comparison of the formant patterns and the temporal structure between speech and singing. The focused song style was the Brazilian lied, often composed in middle frequency region, even for sopranos, as was the register of the experimental subjects. A song by Francisco Mignone, Cantiga de Ninar, was chosen from which a phrase was extract to be the carrier phrase in a low middle frequency region. The non-word /la'CV/ was introduced in the carrier phrase without altering its rhythmic and melodic structure. Measurements were made of the non-word segment duration and the first three formants of the vowel target in the CV syllable. In this syllable the consonant was [p,t,k] and vowels were [a, e, e, i, o, o, u], oral, in stressed position. The main finding was an acoustic effect that can be described as the matching of the vowel formants to the harmonics (partials) of the sung note.*

1. INTRODUÇÃO

É o canto com texto, ou o canto que acolhe a fala, fazendo dela uma fala cantada, que constitui uma contraparte do objeto mais amplo deste trabalho e da qual deriva o problema da inteligibilidade. A outra parte do objeto deste estudo é a fala falada, ou apenas fala.

O canto é um modo de expressão humana de origem já longínqua no tempo. Quando o homem começou a cantar? Não se sabe ao certo. A resposta é mais ou menos igual àquela que serve para responder a outras questões parecidas sobre origem, às quais responde-se com alguma especulação.

O texto no canto é perpetuado na nossa cultura ocidental, desde os hinos gregos até a canção de hoje, existindo tanto na cultura erudita, como na popular, sob as

* Texto resultante da Tese de Doutorado, apresentada ao Curso de Linguística, do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas (Unicamp), no dia 9 de agosto de 2002, sob a orientação da Profª. Drª. Eleonora Cavacante Albano.

mais variadas formas. No Brasil, além dos hinos cristãos largamente ensinados aos índios pelos jesuítas, canções trazidas d'além-mar começaram a criar cor local na colônia, fundindo-se com o lundu¹, o que acabou por criar a canção brasileira. É nesta fusão, em que não se vislumbra uma fronteira nítida entre o popular e o erudito, que nasceu o que podemos chamar de canção erudita brasileira, a partir do século XIX².

E a fala o que é? A fala é mesmo um objeto fascinante: é movimento mutante desde sua produção fisiológica, deixando os lábios, propagando-se no ar (movimento acústico) chegando ou não aos ouvidos (movimento perceptual); é forma de realização de uma das faculdades humanas mais difíceis de explicar: a linguagem natural. Por ser assim objeto de tantas faces, difícil apreendê-lo e resolver sua definição de uma hora para outra. É grande: o tipo da coisa que se pega de um lado e não se consegue pegar do outro. Daí então termos escolhido apenas dois aspectos fonético-acústicos da fala e do canto a serem estudados comparativamente: a duração dos segmentos e a característica das vogais.

1.1. O Mito da má-dicção

Mário de Andrade dedicava uma preocupação especial à questão da inteligibilidade do português brasileiro cantado, tendo realizado o Primeiro Congresso da Língua Nacional Cantada, em 1937, em São Paulo. No entanto a idéia de, não só cantar, como compor de modo a se valorizar o reconhecimento do texto, não se difundiu. Em vista da não continuidade do trabalho do Congresso de 1937 e da pouca divulgação das suas discussões que são muito mais ricas do que os resultados das Normas³, atribuíam-se e atribui-se até hoje, entre cantores, outros músicos e maestros, a falta de inteligibilidade do texto cantado à má dicção.

A questão que ficou latente no início deste trabalho foi justamente sobre a inteligibilidade do canto. Antes, porém, de tentar explicar como é percebida a fala cantada, propomo-nos a estudar como ela é produzida. Uma contribuição relevante da presente investigação é o esclarecimento, para a área a ela pertencente (onde se tocam a música e a lingüística) dos fenômenos fonético-acústicos observados na produção do texto cantado. A área da Lingüística que se mostrou pertinente para nosso trabalho foi a Fonética Acústica.

¹ Inicialmente dança trazida pelos negros africanos ao Brasil, acompanhada de canto e de caráter bastante sensual (Andrade, 1977). Segundo Aurélio Buarque de Holanda, *teve seu esplendor em Brasil em fins do século XVIII e começos do século XIX. Dos meados do século XIX em diante, canção solista, influenciados pelo lirismo da modinha e freqüentemente de caráter cômico.* (Caldas, 1985).

² As canções de Villa-Lobos, tido como um compositor erudito, são consideradas *cross-over*, na Europa: ou seja, não são nem populares nem eruditas.

³ Várias comunicações interessantes sobre a pronúncia das diversas regiões brasileiras e até peculiaridades do folclore encontram-se reunidas nos Anais do Primeiro Congresso da Língua Nacional Cantada.

1.2. Metodologia

Para comparar a fala e o canto em PB, escolhemos as sete vogais orais [i, e, ε, a, ɔ, o, u], da posição acentual tônica; e as consoantes oclusivas surdas [p,t,k], que foram realizadas no logatoma /la'CV/, contido na frase veículo: “Canto /la'CV/ baixinho numa velha canção de ninar”, na qual V era entoada numa frequência fundamental em torno de 420 Hz (lá bemol 3). Não foi difícil encontrarmos a frase veículo, que felizmente, está numa das canções, cujo compositor é um dos mais produtivos e reverenciados pela crítica musical brasileira⁴. Trata-se da Cantiga de Ninar de Francisco Mignone.

A frase veículo foi falada e cantada em cinco repetições por cinco cantoras (sopranos de câmara). Multiplicando o número de informantes, as cantoras, com o de repetições, com o de modalidades (canto e fala), com o de consoantes e, finalmente, com o de vogais, obtivemos um total de 1050 ocorrências ($5 \times 5 \times 2 \times 3 \times 7 = 1050$).

Uma vez terminada a gravação, realizada em cabine insonorizada do Lafape⁵, estava disponível, então, o conjunto de dados que se constituiu em nosso *corpus*.

Digitalizaram-se os dados gravados a 26500 Hz, utilizando-se o *software* de análise de fala CSL, modelo 4300b da *Kay Elemetrics*. A partir da forma, da onda gerava-se um espectrograma de um filtro de 375 Hz, este usado para medidas de duração.

Inicialmente mediamos o logatoma do *onset* do // até o *offset* da vogal-alvo. Em seguida medimos a consoante, do início do silêncio até o *onset* da vogal-alvo; depois a vogal-alvo até o *onset* da oclusão da palavra *baixinho*. Por último, medimos a vogal pré-tônica e o //, sempre com o auxílio visual da forma da onda e do espectrograma.

As medidas de frequência compreenderam as medidas dos três primeiros formantes, F_1 , F_2 e F_3 das vogais, em sua porção estacionária. O algoritmo utilizado foi o LPC (*Linear Predictive Code*), sobreposto ao FFT (*Fast Fourier Transform*)⁶.

O *software* de análise estatística utilizado foi o SAS. Tanto para os dados de duração como para aqueles de padrão formântico, o procedimento de análise utilizado foi o GLM (*General Linear Model*) para medidas repetidas e o Teste de Duncan, para comparação aos pares.

⁴ Para uma história bastante completa da canção brasileira, ver Mariz (1985).

⁵ Lafape: Laboratório de Fonética e Psicolinguística do IEL, UNICAMP.

⁶ A fixação de parâmetros para, respectivamente FFT e LPC, foi à seguinte: 1024 pontos para tamanho da amostra, 20 para tamanho do quadro e 24 a ordem de filtro.

2. VOGAIS LONGAS E CONSOANTES CURTAS

O fato de a vogal ser o som da fala eleito pela música não deve causar nenhuma surpresa. Os sons musicais, para serem considerados como tais, devem possuir uma estabilidade que não se encontra em sons não musicais. No capítulo 5 de Kent & Read (1992), *The Acoustic Characteristics of Vowels and Diphtongs*, temos a seguinte definição de vogais: (...) *vowels are associated with a steady-state articulatory configuration and a steady-state acoustic pattern. Supposedly, then, a vowel can be indefinitely prolonged as an articulatory or acoustic phenomenon.* (p.87) Ora, o prolongamento indefinido de que nos falam Kent e Read (op.cit.) é também o que acontece com uma nota musical. Assim, a vogal cantada segundo as partituras, tal como as conhecemos na música ocidental, deve obedecer à duração da nota, pois, no canto apenas vocalizado ou com texto, ela é a própria realização desta nota. Pela natureza mesma da vogal e pelas exigências temporais da realização de sons musicais, podemos supor que a vogal seja alongada no canto em relação à fala. O que acontece, então com o som da fala despojado de características musicais, a consoante?

Para poder responder a essa pergunta, apresentamos os resultados obtidos quanto à duração das oclusivas, bem como o resultado da análise estatística concernente, tanto às durações absolutas, como às relativas. Os gráficos 1, 2 e 3 ilustram bem as comparações entre fala e canto em termos absolutos.

Gráfico 1: Duração do logatoma: fala x canto

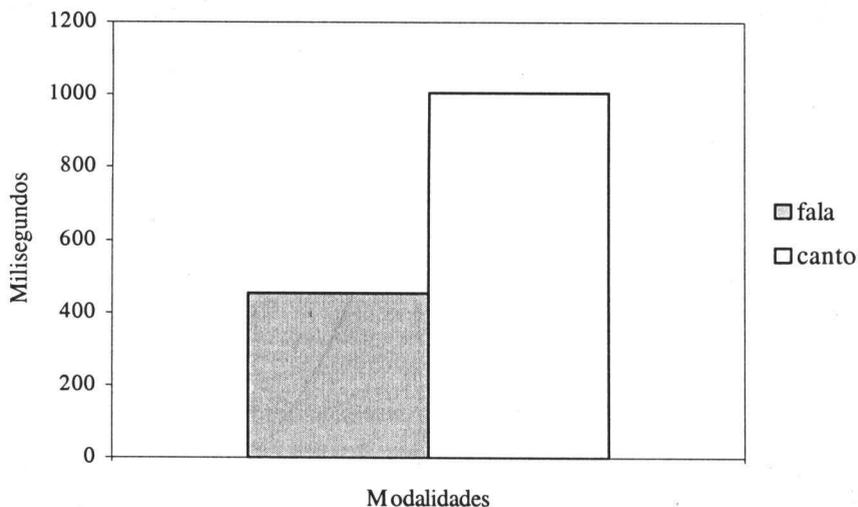
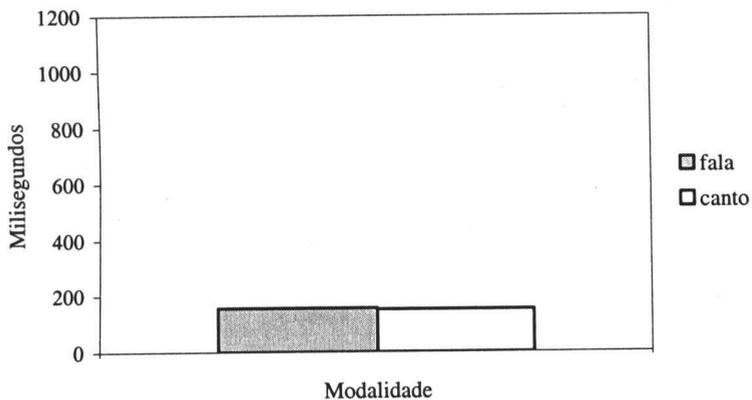


Gráfico 2: Duração da consoante oclusiva: fala x canto



As tabelas a seguir mostram as comparações feitas em termos relativos. P é o p-value, sendo que alfa é fixado a 0.05, e s indica que há diferença significativa.

Gráfico 3: Duração da vogal alvo: fala x canto

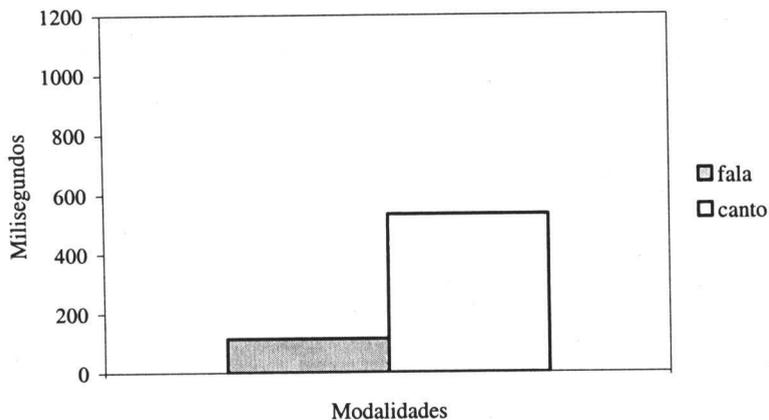


Tabela 1: Padrão geral da razão consoante/logatoma oclusiva. Todas as informantes, vogais e consoantes.

Razão consoante/logatoma			
Fala	Canto	P	Sig.
34%	15%	0.000	s

Tabela 2: Padrão geral da razão vogal alvo/logatoma. Todas as informantes, vogais e consoantes.

Razão vogal alvo/logatoma			
Fala	Canto	P	Sig.
25%	53%	0.000	S

Retomando à pergunta feita no antepenúltimo parágrafo, respondêmo-la, então, utilizando como exemplo as três oclusivas surdas que foram o foco da investigação nesta parte do estudo. Vimos que [p, t, k] ocupam uma porção importante do logatoma (34%) na fala e que essa importância cai para a metade no canto, muito embora, em termos absolutos, tenham durações semelhantes. A vogal, por sua vez, em termos relativos ou absolutos é sempre mais longa no canto que na fala, ocupando 53% do logatoma cantado e 25% do falado. Daí concluímos que as consoantes oclusivas surdas cantadas são encurtadas no canto (à exceção do [p] em termos absolutos), como pudemos aventar inicialmente, e têm uma importância relativa numa e noutra modalidade.

A partir dos resultados acima, podemos dizer que a fala, em comparação ao canto, dá importância às consoantes, enquanto este último destaca a importância das vogais. A importância atribuída a cada classe dos sons da fala explica-se pelas restrições próprias de cada modalidade.

A coordenação dos gestos cantados minimiza as características acústicas de silêncio e ruído que marcam a consoante, pelo fato de a voz cantada ser instrumento musical de alturas e não rítmico e assim necessitar das vogais, sons da fala que possuem uma estrutura harmônica. A coordenação dos gestos falados, comparados aos cantados, privilegia a consoante; Kessinger e Blumstein (1998), descobriram que quando a vogal se alonga, o VOT também se alonga; isto também deve ocorrer quando há foco ou acento frasal, provavelmente aumentando a duração da consoante. Por outro lado, Stevens (2000) verificou que a descontinuidade do sinal da fala marcado por uma mudança repentina cria pistas acústicas importantes em ambos os lados da consoante, que são fundamentais para sua identificação. Tais pistas parecem se manter no canto, dado que há um tempo mínimo de duração para as oclusivas investigadas.

Temos então, de um lado, a fala, em que as consoantes oclusivas são realizadas num tempo considerável, e, de outro, o canto, em que tais consoantes realizam-se

num tempo que poderíamos chamar de tempo suficiente, cuja função é ancorar a textualidade cantada, numa espécie de negociação entre as duas modalidades.

3. VOGAIS DEFORMADAS

Falar de vogais no âmbito da fonética acústica parece fácil à primeira vista, devido aos bem sucedidos e propagados modelos da produção de vogais que conhecemos, o de Fant (1960) e de Chiba e Kajiyama (1941, apud Kent e Read, (1992)). Há uma literatura extensa referente a estudos de vogais a começar, por exemplo, pelo estudo da vogais do inglês americano de Peterson e Barney (1952). No que concerne as vogais cantadas, já não se pode dizer o mesmo. Os trabalhos mais divulgados no âmbito da fonética acústica, que tratam de segmentos, são de Sundberg (1969,1977, 1987) e Scotto di Carlo (1985, 1991).

Sundberg (1977) revela-nos um achado de grande relevância para repensar a questão da inteligibilidade no canto e atacar o mito da má dicção: ao investigar a acústica da voz cantada, o autor descobre duas manobras articulatórias, uma dos cantores⁷ e outra das cantoras, para conseguirem afinação e volume de voz, a fim de furarem a massa de som da orquestra – aqui, claro, ele está falando de cantores de ópera – sem realizar um esforço extenuante para as pregas vocais.

As cantoras realizam uma manobra bastante simples para obter maior energia nos agudos. Sundberg constata que tanto mais a nota deixa uma região grave (sol 3, por exemplo), para uma aguda (fá 4), maior é abertura da boca, ou seja, mais a mandíbula abaixa. Como F_1 é sensível ao abaixamento da mandíbula, o resultado acústico é de que, qualquer que seja a vogal, seu primeiro formante tende a parear com a frequência fundamental da nota, ou, como diz Sundberg, com a frequência de fonação. Isto acontece, em geral, a partir de um fá 4 que tem aproximadamente 700 Hz, o que vai resultar na alteração da cor da vogal cantada: por exemplo, se a vogal for um /i/ tenderá a ser ouvida como um /e/, devido à elevação de F_1 . As vogais, no canto, perdem em inteligibilidade, mas ganham em energia, pois, para ajustá-las às exigências musicais de afinação e volume, a manobra do abaixamento da mandíbula é uma solução do soprano para esforço vocal desnecessário ao cantar.

Quanto à questão da deformação das vogais, temos em Scotto di Carlo e Germain (1985) um estudo bastante ilustrativo. As autoras francesas investigam qual a influência do *pitch* sobre a inteligibilidade de vogais cantadas, utilizando 17 notas musicais em cinco regiões de frequência diferentes: grave, grave-média, média, aguda-média, aguda. Do lá 3 (220 Hz) ao dó 6 (1046 Hz), a soprano ligeiro M. Mesplé cantou 15 vogais ([i, e, ε, α, a, y, Ø, æ, u, o, ɔ, ê, œ, ð, ã] do francês em cada uma das 17 notas musicais, para serem identificadas por quatro ouvintes

⁷ Trata-se aqui de uma manobra que tem como consequência o “formante do cantor” (para detalhes, ver a tese na íntegra,).

ouvintes foneticistas, porém não músicos. A hipótese inicial era a de que tanto a natureza da vogal como a frequência em que era cantada teriam influência na inteligibilidade. Tal hipótese foi corroborada pelos resultados indicando 62% de identificação da vogal no registro baixo-médio, caindo para 38% no médio e para 9% no agudo. Isto quer dizer que quanto mais alta é a nota musical cantada, menos inteligível é a vogal, o que nos remete obrigatoriamente ao achado de Sundberg (1977): uma exigência articulatória do canto, o abaixamento da mandíbula, tem consequência acústica desejada que é o aumento de energia do sinal sonoro em detrimento da inteligibilidade do texto, uma vez que a vogal é deformada.

3.1 Resultados

Nossos resultados coadunam-se com os resultados dos autores acima resenhados, embora estejamos tratando da canção brasileira, ao passo que aqueles autores têm como fonte de dados o canto operístico.

Para ilustrar os achados da tese no tocante ao padrão formântico, apresentamos dois gráficos de uma das informantes, bastante representativos do padrão geral⁸ da produção de vogais a que chegamos. O gráfico 1 indica-nos o padrão formântico da fala e o Gráfico 2 o do canto. Atente-se para o fato de que neste último gráfico há uma escala diferente do primeiro. Fizemos isso pois verificamos que no novo espaço acústico, as vogais tendem a se sobrepor consideravelmente, e não de modo altamente variável, mas acabam por reforçar os formantes sobre os harmônicos da frequência fundamental, que era sempre em torno de 420 Hz. (lá bemol 3). Daí termos traçado linhas sobre a área de plotagem, a fim de visualizarmos melhor o que estava acontecendo entre valores de formantes e harmônicos da nota, sendo estes : H_1 (420 Hz), H_2 (840 Hz), H_3 (1260 Hz), H_4 (1680 Hz), H_5 (2100 Hz) e H_6 (2520 Hz).

⁸ Chamamos de *padrão geral* os valores médios obtidos a partir da somatória das ocorrências de todas as informantes.

Gráfico 1: informante MG, vogais faladas

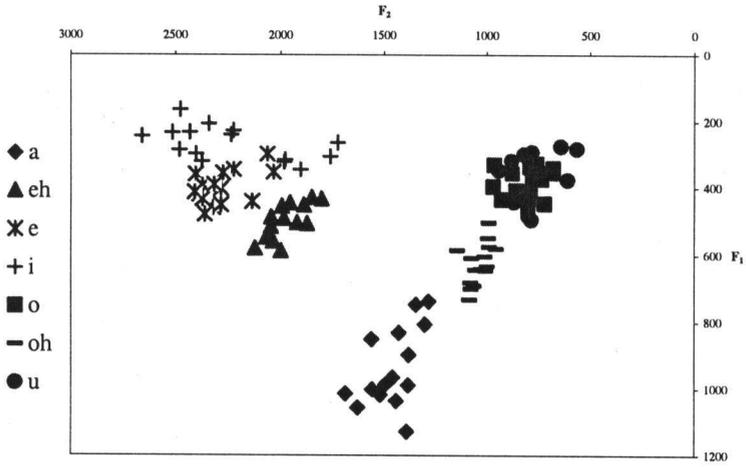
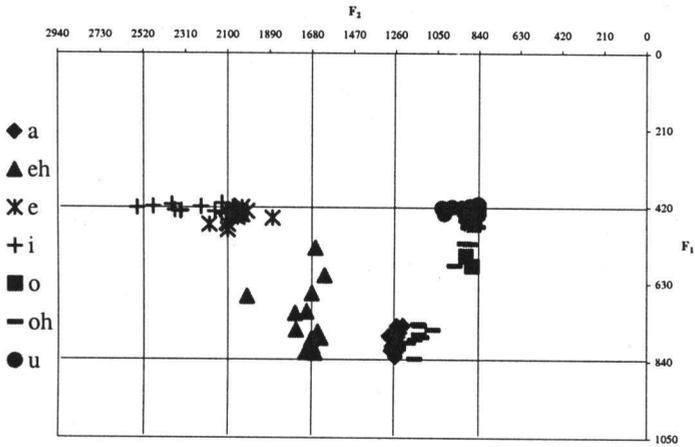


Gráfico 2: informante MG, vogais cantadas



Analisando os gráficos de $F_1 \times F_2$ das vogais cantadas, seja no padrão global⁹, seja por informante, foi-nos possível verificar pontos de atração para os dois primeiros formantes das vogais cantadas. Ao fenômeno de atração demos o nome de pareamento, pois o formante tende a parrear-se aos harmônicos da nota.

Mas qual a manobra articulatória que propicia o pareamento? Para responder a esta pergunta, propomos recuperar os movimentos articulatórios através das medidas acústicas dos três primeiros formantes, cujos valores constam da tabela abaixo:

Tabela 1: Valores médios da vogais faladas e cantadas por todas as informantes

Vogais	F ₁			F ₂			F ₃		
	fala	canto	Sig	fala	Canto	Sig	fala	canto	Sig
i	314	440	s	2476	2314	s	3078	3004	s
e	465	540	s	2310	2039	s	2952	2991	ns
ε	634	755	s	2087	1779	s	2878	2984	s
a	906	767	s	1529	1317	s	2742	3018	s
ɔ	682	752	s	1108	1234	s	2749	3009	s
o	475	614	s	918	1077	s	2864	3052	s
u	377	471	s	855	958	s	2937	3060	s

Comparando fala e canto, tendo aquela primeira como ponto de partida, chegamos à seguinte movimentação dos formantes das vogais cantadas: F_1 se eleva em todas as vogais, exceto em [a], em que diminui. F_2 diminui em todas as anteriores, no [a] e eleva-se nas posteriores. F_3 diminui em [i], permanece igual em [e], eleva-se em [ε], [a] e nas posteriores. Com base na literatura clássica, (Sundberg e Lindblom, 1971) explicamos a movimentação de cada formante (diminuição e elevação) como efeito dos articuladores do trato vocal: lábios, língua, mandíbula e laringe. No entanto, embora tentemos entender os movimentos dos formantes como efeito da ação de diferentes articuladores, talvez um apenas seja responsável pela manobra que resulta no pareamento: a língua.

⁹ Ver a tese na íntegra.

Para sustentar o fato de que a língua é responsável pelo ajuste a que se submetem as vogais cantadas para criar um tubo mais livre de constrictões, apoiarme-ei, basicamente em dois trabalhos, o de Sundberg (1987) e o de Johnson, Ladefoged e Lindau (1993). O primeiro trata a língua como um articulador que pode tomar várias formas, dependendo da direção em que se avulta, seja em direção ao palato duro, ao véu ou em direção à parede posterior da faringe.

Digamos que, no canto, o avultamento da língua seria pequeno, a fim de permitir um trato vocal com as restrições apenas suficientes para criar uma cor vocálica. Tendemos a considerar a língua como um articulador importante do canto, sobretudo na região de frequência investigada, a grave-média, região em que a manobra de abaixamento da mandíbula é relativizada, e outra manobra é necessária para manter a energia do sinal. A língua é responsável por marcadas diferenças articulatórias, também na fala, segundo reportado em Johnson et al. (1993). Os autores descrevem a constrictão do corpo da língua de um dos sujeitos como mais estreita do que a de outro, para o qual a constrictão é mais distribuída.

Diferente do que acontece na fala, em que cada falante realiza um formato de língua diferente, no canto, as cantoras tenderiam a buscar sempre um mesmo formato, talvez menos avultado. Este formato de língua, por sua vez, seria responsável pelas manobras articulatórias que tornam o tubo mais livre de constrictões, o que propicia o pareamento dos dois primeiros formantes com os harmônicos da nota.

4. CONCLUSÃO

Ao comparar fala e canto sob dois aspectos fonético-acústicos diferentes, obtivemos evidências de que há restrições musicais se impondo à fala, no caso da deformação das vogais, e restrições da fala que se impõem às musicais, como, por exemplo, a duração apenas suficiente das consoantes. Isso nos levou a considerar que as restrições de fala cantada não têm caráter apenas físico, mas lingüístico, uma vez que a inteligibilidade se mantém em um nível considerável numa região de frequência média (Scotto di Carlo e Germain, 1985). Consideramos ainda que, ao se produzir a fala cantada, está-se coordenando planos e execuções lingüísticos e musicais, e não apenas reproduzindo um enunciado estereotipado com alturas determinadas por uma tonalidade musical. Para aprofundar a investigação nesse sentido, seriam necessários outros experimentos delineados em um estudo não apenas descritivo – como é o caso desta tese – baseado num modelo dinâmico de produção de fala, como por exemplo, o defendido por Albano (2001). Ao tratar das constrictões que definem gestos articulatórios, a Fonologia Acústico-Articulatoria (FAAR), defendida por Albano (op.cit) desloca a ênfase da Fonologia Articulatória

(FAR)¹⁰, do articulador para regiões articulatórias associadas a efeitos acústicos (cf. a Teoria Acústica de Fant). As manobras articulatórias hipotetizadas a partir dos resultados acústicos desta tese são consistentes com essa visão mais flexível da realização das constrictões que constituem os gestos articulatórios¹¹.

Por fim, sobre as diferenças entre fala e canto podemos concluir o seguinte:

Quanto à *Estrutura Temporal*, o canto alonga as vogais, responsáveis, nessa modalidade, pela entonação da nota musical, produzindo as consoantes num tempo mínimo. Ainda que encurtadas em termos absolutos e relativos, as consoantes são responsáveis pela manutenção do texto da canção.

Quanto ao *Padrão Formântico*, podemos dizer o seguinte: no longo tempo em que as vogais são produzidas no canto, a manutenção da frequência fundamental para garantir afinação é propiciada por duas manobras articulatórias: abaixamento do maxilar (Sundberg, 1977) e menor avultamento da língua (exceto para [a]).

As conseqüências acústicas destas duas manobras articulatórias são: a sobreposição desejável de [i,e]; [e, ε]; [a, ɔ] e [o,u] , o pareamento de F₁ com a frequência fundamental ou H₁(420 Hz) e com o segundo harmônico ou H₂ (840 Hz) e o pareamento de F₂ com os harmônicos de 2 a 6.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBANO, E. (2001). *O gesto e suas bordas*, Campinas, Mercado de Letras.
- ANAIS do Primeiro Congresso da Língua Nacional Cantada, (1938). São Paulo, Departamento de Cultura do Estado de São Paulo.
- ANDRADE, M. (1977). *Pequena história da música*, 8ª ed. São Paulo, Livraria Martins Editora.
- BROWMAN, C. & GOLDSTEIN, L. (1992). Articulatory Phonology: An Overview, in *Phonetica* 49: pp 155-180.
- FANT, G. (1960). *Acoustic Theory of Acoustic Perception*, 2a. edição, The Hague-Paris, Mouton, 1970.
- JOHNSON, K.; LADEFOGED, P. & LINDAU, M. (1993). Individual differences in vowel production, in *The Journal of the Acoustical Society of America*, vol. 94, no. 2, pp 701-714.
- KENT, R.D. & READ, C. (1992). *The acoustic analysis of speech*, San Diego, Singular Publishing Group.
- LINDBLOM, B. & SUNDBERG, J. (1971). Acoustical consequences of Lip, tongue, Jaw and Larynx movement, in *The Journal of the Acoustical Society of America*, Vol. 50, no. 4, pp 1166-1179.

¹⁰ A FAR tem como autores seminais Catherine Browman e Louis Goldstein. Uma visão geral da proposta pode ser encontrada em Browman e Goldstein (1992).

¹¹ Para mais detalhes, ver a tese de doutorado, sobretudo o Capítulo 3 (aqui não resenhado por falta de espaço) e a Conclusão.

- KESSINGER & BLUMSTEIN, S. (1998). Effects of speaking rate on voice-onset time and vowel production: some implications for perception studies, in *Journal of Phonetics*, 26, pp 117-128.
- MARIZ, V. (1985). *A canção brasileira: popular e erudita*, 5ª. Edição, Rio de Janeiro, Nova Fronteira.
- PETERSON, G.E. & BARNEY, H. E. (1952). Control methods used in a study of vowels, in *The Journal of the Acoustical Society of America*, 24, pp. 175-184.
- SCOTTO DI CARLO, N. & GERMAIN, A. (1985). A perceptual study of the influence of pitch in the intelligibility of sung vowels. *Phonetica*, 42, 4, pp.188-197.
- SCOTTO DI CARLO, N. (1991). L'organisation temporelle de la syllabe dans la parole e dans le chant: étude préliminaire, Communication au *XII Congrès des Sciences Phonétiques Aix-en-Provence*, Août.
- STEVENS, K. (2000). Diverse acoustic cues at consonantal landmarks in *Phonetica*, 57, pp139-151.
- SUNDBERG, J. (1/1969). Articulatory differences between spoken and sung vowels in singers, *STL-QPSR*, pp 33-46.
- _____. (1977). The acoustics of the singing voice, *Scientific American*, Vol 236, no. 3, pp 82-91.
- _____. (1987). *The science of the singing voice*, Northern Illinois University Press, Dekalb.