

# O canto fala mais

*da Acústica à Cognição*



Beatriz Raposo de Medeiros



fflch

O Canto Fala Mais:  
da Acústica à Cognição

 **Universidade de São Paulo**

**Reitor**

*Carlos Gilberto Carlotti Junior*

**Vice-Reitora**

*Maria Arminda do Nascimento Arruda*



**Faculdade de Filosofia, Letras e Ciências Humanas**

**Diretor**

*Paulo Martins*

**Vice-Diretora**

*Ana Paula Torres Megiani*

DOI 10.11606/9788575064856

Beatriz Raposo de Medeiros

O Canto Fala Mais:  
da Acústica à Cognição



São Paulo, 2024

Catálogo na Publicação (CIP)  
Serviço de Biblioteca e Documentação  
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo  
Elizabeth Barbosa dos Santos – CRB-8/6638

---

M488 Medeiros, Beatriz Raposo de.  
O canto fala mais [recurso eletrônico] : da acústica à cognição /  
Beatriz Raposo de Medeiros. -- São Paulo: FFLCH/USP, 2024.  
12.300Kb ; PDF.

ISBN: 978-85-7506-485-6  
DOI 10.11606/9788575064856

1. Linguística aplicada. 2. Canto e fala – fonética. I. Título.

CDD 418

---



**Esta obra é de acesso aberto. É permitida a reprodução parcial ou total desta obra, desde que citada a fonte e autoria e respeitando a Licença *Creative Commons* indicada**

## **Dedicatória**

Dedico este livro às mulheres e à liberdade feminina. Portanto, é também dedicada a homens voltados à compreensão dos tempos novos e que conseguiram romper o ciclo de uma formação. Paradoxalmente, estamos, a todo momento, falando de um mundo de homens. Sobretudo na música e nos achados científicos. Não raro acontece de o trabalho ou a descoberta ser de uma mulher, mas quem leva o crédito ser um homem. Na nossa música, há sambas com letras machistas, de melodias memoráveis. Essas referências só deixarão de ser representativas quando findada a transição de um mundo de ideias para um em que incorporemos a vida do outro/da outra, cognitiva e compassivamente. Por enquanto há muito que fazer.

## SUMÁRIO

<b>POR QUE LER ESTE LIVRO?</b>	8
<b>PRINCÍPIO</b>	9
A Fala Contida no Canto e Travessias	9
<b>PARTE 1</b>	15
As Canções Foram Feitas para Serem Entendidas?	15
Dados Suplementares sobre a Descrição Fonética do Canto em Comparação à Fala	28
<b>PARTE 2</b>	37
A Fala, O Canto & A Linguística	37
A Fala e o Canto como Objetos de Estudos da Linguagem	49
Ideias Centrais à Ciência da Linguagem	50
A Importância da Fonética	51
Levantamento de Dissertações e Teses	52
Conclusão e Perspectivas Futuras	61
<b>PARTE 3</b>	63
Ritmo na Fala e na Música como	63
Organização da Temporalidade	63
<b>PARTE 4:</b>	82
A Questão da Estabilidade do Canto versus	82
A Instabilidade da Fala	82
Estabilidade em Termos Latos e em Termos Microscópicos também	89
Mas o que Seria a Instabilidade de Fonação na Fala?	94
Estudos sobre a Estabilidade de f0 no Canto e na Fala	96

<b>PARTE 5:</b>	99
Primitivos da Ciência:	99
O Primitivo Linguístico & O Primitivo Musical	99
<b>FIM</b>	116
O Sabor do Saber &	116
O Prazer de Cantar	116
<b>Agradecimentos</b>	127
<b>Referências</b>	128
<b>Sobre a Autora</b>	137



## **Por que ler este livro?**

Porque ele tem o objetivo de mostrar um percurso de estudos sobre fala e canto a partir da fonética, que é uma subárea da linguística, utilizando uma linguagem que, tanto leigos como acadêmicos de diversas áreas, compreendem.

Procuramos deixar mais acessíveis conhecimentos sobre os sons do canto, comparativamente à fala. Portanto, estamos tratando da acústica e da articulação desta capacidade tão admirável que é a produção de fala.

Aqui, o canto é considerado um tipo de fala. Falamos da inteligibilidade do canto e suas manobras que, por vezes, podem deixar o texto ininteligível. Tratamos do canto como um assunto que pode ser estudado pela ciência da linguagem, a linguística. Traçamos paralelos entre ritmo de fala e ritmo da música. Nossa intenção é sermos claros sem sermos simplistas, é falarmos da complexidade de coisas complexas, apontando para sua parte mais relevante relacionada às nossas discussões, sem devaneios.

Sobretudo, queremos atar pontas que as fronteiras entre áreas do saber parecem ter deixado eternamente soltas. Não deveria ser assim. Atar tais pontas, neste livro, significa tornar acessível muitas coisas que são estudadas na universidade e que devem ser amplamente divulgadas para toda e qualquer pessoa interessada.

## **PRINCÍPIO**

**A Fala Contida no Canto e Travessias**

Somos seres sonoros. Tão refinadamente sonoros, que desenvolvemos habilidades maravilhosas para produzir som, como a voz. É um pouco arriscado dizer isso aqui, pois corremos o risco de fixarmos a ideia de que a voz é apenas um produto acústico, ou sonoro. A voz é muito mais do que isso, já que para produzi-la temos de coordenar uma série de sistemas do nosso corpo. Simplificando: músculos respiratórios, laríngeos e faciais, bem como a língua e a mandíbula e demais articuladores são convocados para falarmos (e/ou cantarmos) as frases mais simples até as mais complexas. Então vejamos: a simples capacidade de cantarolar uma canção. Vamos lá: a sua preferida. Eu tenho muitas, mas só para te entusiasmar, ao escrever estas linhas, em silêncio canto: “fala baixinho para eu ouvir, porque ninguém vai mesmo compreender...” do Pixinguinha. Podemos cantar baixinho, ou subvocalizar, cada um para si. Aí, então, já desempenhamos uma tarefa belíssima: ajustamos a nossa laringe para emitir notas musicais. Este ajuste está principalmente movimentando nossas pregas vocais, de alguns centímetros (cm) de comprimento (em média, 1,7 cm). Outro grupo de músculos e de cartilagens age aqui para os movimentos articulatorios precisos que fazemos para entoar frequências fundamentais estáveis: sim, isso, as notas musicais, com as quais produzimos melodias. Vejam: somos todos seres sonoros e musicais.

Só por esta razão, pelo fato de cada um de nós cantarolar sua canção preferida (lembrando toda a letra ou não, isso não importa) já é de se confirmar que sabemos música, ainda que não formalmente. Se já sabemos isso, devemos confirmar e espalhar essa ideia por aí. Se não pensávamos nisso, devemos nos persuadir de uma vez por todas. Se pensávamos o contrário – que não temos ritmo, nem afinação –, devemos nos dissuadir disso imediatamente.

Tanto a vivência na música, como a convivência com músicos (mulheres e homens), desde pequena, ajudaram-me a construir a ideia de que somos todos seres musicais. O conhecimento formal e científico que tento desenvolver na idade adulta pode, em princípio, comprovar isso. Aprendi que não podemos confundir diferenças de desempenho, ou *performance*, com capacidades facilmente desenvolvidas pelo ser humano. A capacidade linguística do ser humano é uma dessas. E aqui trato capacidade de perceber e ouvir música como pertencente a todos nós, irmã, portanto, da linguística. O desempenho individual é que vai ser diferente.

Um das pessoas parecem ter mais talento do que outras. Falar de talento ou dom, nos domínios científicos, é um problema. Acadêmicos (como psicólogos, estudiosos da cognição, etc.) dirão que talento não existe e o que haveria seria a soma de fatores e de fatos que compõem o desenvolvimento de uma pessoa. Mas voltando ao senso comum sobre as diferentes performances de diferentes indivíduos a dar piruetas no ar, driblar o zagueiro e/ou cantar ópera, é possível emitir impressões sobre essas, sem medo de errar. Há *performances* estupefacentes, como as de ginastas olímpicas e craques do futebol, diante das quais só podemos dar um parecer: “Que talento!”. Em comparação a essas performances, podemos elencar atividades corriqueiras: pular poças de água em dias de chuva, conseguir sair do metrô lotado e cantar no banheiro. Isso todo mundo pode e sabe fazer. Tais tarefas oferecem seu grau de dificuldade também, que nós, sem percebermos, resolvemos com comportamentos mais complexos do que podemos imaginar que sejam, graças a nossas capacidades de movimentos e nosso desempenho cognitivo em música que revela faculdades como a de saber escalas musicais e entoá-las com a voz, como cantar afinado — mesmo que não nos demos conta disso.

Em suma: não importa se você canta desafinado. A afinação é importante para que o canto veicule traço elementar da música: é ela que dá a devida distância sonora entre as notas, a tal da melodia. Mas o fato é que todos nós cantamos, de um modo ou de outro. O mesmo pode-se dizer da capacidade de falar: alguns entre nós são mais talentosos linguisticamente, sob as mais diversas formas: os tagarelas, os que estão sempre fazendo trocadilhos, os que têm (e usam!) um repertório grande de palavras, os que têm o dom de convencer, de rimar, enfim... e de, até silentes, ouvir estrelas! “Ora direis... ouvir estrelas! Certo, perdeste o senso...!”<sup>1</sup> Então voltemos à razão. Mesmo aqueles quietos, quantas frases cabem em seu silêncio? Então, todos nós somos seres linguísticos, uns menos falantes do que outros. Assim, mantemos nossa proposta de defender que somos seres aptos, muito aptos a desenvolver a língua e a música.

A ideia de que assim somos levou-me a buscar respostas para fenômenos que ocorrem no território comum entre fala e canto. Como veremos, tudo partiu de uma inquietação a respeito de uma tarefa: cantar agudo, afinado e com volume e sem

---

<sup>1</sup> Olavo Bilac. Antologia: Poesias. São Paulo: Martin Claret, 2002. pp. 37-55: Via-Láctea.

microfone. Este cantar possui um texto e, portanto, pode ser definido simplesmente como canção. Espera-se então que quem canta torne o que canta musical e inteligível. No entanto, quando isso não ocorre, qual é o problema? “Qual a problemática?”, perguntam as agências de financiamento de pesquisa. Levantar esta problemática foi o olho d’água da pesquisa empreendida até o momento presente. Ao encontrar curso, mais ou menos livre, segue correndo a água do conhecimento. Vai correndo, para um mar de perguntas. Academicamente, tentei respondê-las. E nessa tentativa, teceu-se um conjunto de estudos aos quais subjazem a ideia (frequência ou pulso) fundamental:

1. Música e língua: são linguagens.
2. O canto e a fala devem ser estudados em conjunto.
3. O canto e a fala devem ser um objeto de estudo.

Nesta obra, retomo assuntos já tratados em trabalhos ao longo dos anos de estudo. Assim, apresentarei as ligações, as mais estreitas que eu conseguir apontar, entre as questões acadêmico-científicas e as questões de senso comum sobre a fala e o canto. Seguramente, aquelas primeiras originaram-se dessas segundas. Não há ciência sem intuição.

A ciência pergunta, inquire, fala por códigos, às vezes oculta-se: são segredos. A arte escancara nos sons, nas cores, no movimento — até à pedra a arte dá vida. A arte responde e revela. Assim, concluímos que ciência e arte se completam e estando latente esta propriedade no canto que eu cantava e nas indagações linguísticas que eu fazia, acabei, quase sem perceber, indo para o domínio da fonética.

Essa complementaridade entre ciência e arte é abordada em áreas fronteiriças como a música, o cinema (atualmente repleto de ciência em suas produções, ou seja, em sua feitura) e arquitetura e isso não é exatamente uma novidade. Talvez o novo seja entender de modo melhor e mais definitivo como se dá essa simbiose ou intercâmbio. Vinha sempre pensando nesse encontro, quando em certa feira do livro (aquela que acontece nos finais de anos no campus do Butantã da USP), topei com este título “A ciência que sonha e o verso que investiga”. Evando Mirra, o autor do livro, arrebatou o verso final de um poema de Jorge de Lima para nomear sua publicação que reúne textos sobre a construção da universidade, tecnologia e inovação, psicanálise, cinema, *jazz* e futebol. Esse livro é uma referência importante para aqueles pesquisadores que estão entre

o verso e a equação, entre o agudo da voz que canta e o *script* que mede a frequência fundamental, a *big data* e o infinito de cantares do Brasil e do mundo.

As primeiras perguntas que me fiz, de forma intuitiva, mas claro, com a bagagem de linguista, a qual me fez duvidar da simples afirmação que cantores não eram cuidadosos com a dicção para entoar seu canto. Data imprecisa: tudo começou em algum mês de 1994, quando meu grupo de estudo de canto, liderado pelo professor Francisco Frias<sup>2</sup>, discutia sobre a dificuldade de se entender o texto cantado por sopranos. Comecei então a ler sobre pronúncia no canto, mas nada havia que explicasse o fenômeno. E as pessoas em geral seguiam reclamando da falta de inteligibilidade, sobretudo do canto erudito, das vozes femininas e quando estavam cantando notas agudas. Então, resumindo, a questão foi tomando vulto e ficou importante em minha vida. Assim, escrevi um singelo projeto para iniciar o doutorado no IEL<sup>3</sup> da UNICAMP, buscando respostas na fonética.

Os desdobramentos, desde o doutorado (1998-2002) do questionamento ocorreram em momentos diversos e de formas diferentes no âmbito da pesquisa científica. A questão ou questões iniciais eram em torno da pronúncia do canto erudito e chegaram a elaborar hipóteses explicativas sobre a sua inteligibilidade. No entanto, este objeto de estudo, inicialmente célula única, passou a ser um tópico no interior de um projeto de investigação maior, que coloca perguntas mais abrangentes, sobre assuntos como a capacidade musical humana e suas relações com a capacidade linguística, incluída aí a capacidade de fala oral.

O fato de a linguística e mais precisamente a fonética terem me respondido questões sobre o canto e sua relação com a fala levou-me a um caminho epistemológico sobre a própria linguística e mais tarde sobre a cognição musical. Afinal, qual o lugar da fala cantada nos estudos da linguagem? Isso dará tecido para a Parte 2, que tem um caráter de defender os estudos da fala cantada no interior da linguística. Trataremos de iniciar uma discussão mais clara a respeito das diferenças entre fala e canto.

Seguiremos em mar aberto, no entanto mais tranquilo, para a Parte 3. O que veremos ter ficado para trás então, é um pedacinho de continente, ali onde aportaram as descobertas sobre as manobras articulatórias do canto e suas resultantes sonoras. Também

---

<sup>2</sup> Francisco Frias foi cantor de ópera e professor de canto na Escola de Música de Brasília (EMB).

<sup>3</sup> Instituto de Estudos da Linguagem da Universidade Estadual de Campinas - IEL/UNICAMP.

ali nesse horizonte de terra, situaram-se as questões de como o som está presente na linguística e pudemos falar de uma espécie de busca do som nos domínios da língua/linguagem que parece prescindir dele.

Definitivamente, na Parte 3, singramos pelos mares do ritmo.

E na Parte 4, enfrentaremos uma tormenta em alto mar, propondo uma hipótese desafiadora: a estabilidade do canto.

Esperamos que a Parte 5, não obstante sua natureza intuitiva e de exploração, consiga, trocando em miúdos, dizer o que é uma unidade primitiva da música, a exemplo do que a linguística já mostrou como primitivo da linguagem.

E por fim, o capítulo que deve ser o último, esforçar-se-á por um final feliz. Talvez concluir que os estudos sobre fala e canto não só devem constituir um objeto de estudo, mas que se dedicar a tal estudo significa fortalecer o elo com um fenômeno que nos constitui de forma mais plena como seres humanos.

**PARTE 1**  
**As Canções Foram Feitas para Serem Entendidas?**



A resposta é simples e é sim. Se foi simples chegar a ela, por via das comprovações experimentais, já não posso afirmar. O que farei é tentar mostrar aqui de forma a mais clara possível para leigos e não leigos os meios utilizados para iluminar as indagações em torno da inteligibilidade do canto.

Ao iniciar este capítulo também inicio um exercício de dizer as coisas a partir de vários pontos de vista. Uma das formas de pluralizar esses pontos de vista é empregar ora o “eu”, ora o “nós” ora o “você”. Nos momentos em que uso “eu”, estou desvendando uma iniciativa individual, sobre a qual assumo toda e qualquer responsabilidade. Quando uso o “nós”, o movimento é o de colocar-me como uma voz aliada a várias outras, seja para dizer coisas a partir do senso comum, ou a partir de descobertas acadêmicas.

Sendo brasileiros e vivendo no Brasil, estamos imersos todos os dias em canções. Se pensarmos num intervalo de tempo de pelo menos 60 anos (mais de meio século, então) já estamos, há muito, sintonizados no rádio, e há muito ligados na televisão, que no Brasil é e sempre foi veiculadora de canções em português brasileiro, malgrado o temor da invasão da música norte-americana – que na verdade sempre esteve aqui presente e sob a forma de canção – nos anos 70. E agora temos a internet. Sobre “A Sonoridade Brasileira”, capítulo do livro *O Século da Canção*, Luiz Tatit resume com maestria:

A atuação do corpo e da voz sempre balizou a produção musical brasileira. A dança, o ritmo e a melodia por eles produzidos deram calibre à música popular e serviram de âncora aos voos estéticos da música erudita. Em todos os períodos, desde o descobrimento, a percussão e a oralidade vêm engendrando a sonoridade do país<sup>4</sup>...

Subjaz a este trecho inicial do livro de Tatit todo um programa de estudo sobre a música e a canção como capacidades humanas em contexto do Brasil e da língua portuguesa brasileira – muito embora não seja esta a proposta do seu livro que é, isso, sim, um olhar semiótico e histórico sobre a canção. Vamos estender um pouco os significados aqui veiculados por algumas palavras isoladamente:

**1. O Corpo; 2. A Voz; 3. A Oralidade; 4. A Melodia e 5. A Percussão.**

---

<sup>4</sup> Luiz Tatit. *O século da canção*. Cotia, Ateliê Editorial, 2004, p. 19

O aspecto corporal a que Tatit se refere é o da dança. Mas aqui quero propor o desenvolvimento dessa ideia do corpo em relação à música: cantar já é uma ação corporal e talvez não atentemos facilmente para isso, pois, no senso comum, focamos mais no produto sonoro das capacidades de cantar e falar do que no seu movimento físico e fisiológico. Também, a capacidade de entoar a voz e produzir a fala é considerada a partir mais do produto e seus efeitos comunicacionais ou estéticos. E além de tudo, os movimentos dos articuladores da fala parecem de tão pouca magnitude (como a batida da ponta da língua nos dentes superiores) e tão “invisíveis” (o abaixamento da úvula, quando produzimos sons nasais), que a eles não damos o status de movimentos corporais ou gestos corporais. No entanto, todo o movimento corporal para se chegar à determinada sonoridade é importante para que entendamos todo o processo da fala e do canto.

A voz humana é um fenômeno absolutamente fantástico e imenso. No caso do estudo inicial sobre os aspectos acústicos de trecho cantado em português brasileiro, o aspecto vocal focado foi a capacidade da voz cantada de interagir com as unidades fônicas da fala (vogais e consoantes) e assim simultaneamente transmitir “duas linguagens” a da música e a da língua natural. Mas como eu ia dizendo, imenso é o fenômeno da voz. Muitas ocasiões, no senso comum, a palavra voz evoca o posicionamento das pessoas pelos seus direitos, opondo-se a um silêncio que seria a negação da existência e do clamor por esses direitos. Em abordagens mais específicas, como, por exemplo, na área da saúde, a voz é o objeto, por excelência dos fonoaudiólogos, cuja tarefa primordial é assegurar a saúde vocal dos indivíduos.

A oralidade é um termo bastante abrangente, que circula por várias áreas do saber. A palavra oral (e, portanto, oralidade) vem de *os*, do latim, que apresenta curiosamente, vários significados: tanto pode ser boca (o mais conotativo) como pode ser língua ou idioma. Em anatomia, os músculos da boca são adjetivados com o termo “*oris*”, como, por exemplo, o *levator angulis oris*, que eleva o canto da boca. Em fonética, para descrever processos tanto acústicos como articulatórios, o termo *cavidade oral* é responsável por designar o espaço interno e vazio da nossa boca. Nos estudos linguísticos, o termo oralidade pode se opor à escrita, para apontar para suas diferenças. Seria infundável a tarefa de descrever e explicar o emprego da palavra “oralidade”, além de desviar o assunto para outra dimensão dessa relação corpo-linguagem. Então o que me

interessa neste capítulo é estabelecer o elo dos meus pensamentos com o que o colega Luiz Tatit chamou de “oralidade” e que pode ser simplesmente a habilidade que temos de gesticular com partes do corpo que estão na boca e cercanias. Trato vocal ou trato oral podem ser usados como substitutos, embora nomeiem áreas diferentes do nosso corpo.

Acredito que nossa ideia mais corriqueira sobre o que seja melodia é aquela ligada à música. Mas podemos nos referir também a uma voz melodiosa, quando queremos apontar para sua qualidade musical. De novo, peço que se volte para sua verve musical cantante e que entoe, baixinho, aquela mesma canção preferida. Quando você faz isso, produz uma melodia. Essa melodia, embora possa ser comparada à melodia da voz falada, tem características próprias. A melodia da fala, tecnicamente estudada como a curva entoacional de alguma sentença que falamos, apresenta duas diferenças básicas da melodia musical: instabilidade e indefinição de alturas. Ou seja, ao falarmos, não afinamos nossas frequências fundamentais. Em outras palavras, nossa habilidade (ou capacidade) de fala não necessita que produzamos a voz em uma determinada altura do som, ou seja, em tons graves, médios e agudos definidos. Isso é uma obrigatoriedade na música/canto. No entanto, é preciso frisar e notar que ao falarmos, produzimos curvas entoacionais mais ou menos definidas, como, por exemplo, a entoação que produzimos ao proferir sentenças interrogativas. Na maioria das línguas, o contorno ou curva de entoação de uma pergunta é diferente daquele da afirmação. Vou tratar isso mais detalhadamente no capítulo 5.

E ao bater palmas no *show* da sua banda preferida, acompanhando seu pulso fundamental, você percutiu. Portanto realizou uma percussão. A nossa habilidade de percutir nas mais diferentes superfícies (a caixinha de fósforo) e na superfície do nosso próprio corpo está ligada intimamente com a ideia de ritmo. O conceito de ritmo é tão vasto quanto o de voz (ou mais) e perpassa todas as áreas do saber humano. Da medicina à física, da psicologia à música! Das batidas do coração aos gráficos que representam os pulsos de um fenômeno rítmico qualquer, passando pela ideia de como percebemos esses pulsos até chegar no tum-tum do tambor, podemos dizer que o ritmo é uma organização vital e necessária. No caso específico da fala, o que chamamos de ritmo deveria ser chamado de organização temporal. Dentre os muitos aspectos dessa organização temporal, foquei na relação de duração de vogais e consoantes no canto em comparação

com essa relação na fala. Por enquanto, podemos fazer uma comparação das consoantes com os instrumentos de percussão. Uma classe delas, chamada de oclusivas, seria assim a responsável pela batida surda de uma cadeia de melodiosas vogais. Mas disso falaremos mais no terceiro capítulo.

Os conceitos de corpo, voz, melodia e ritmo eram intrínsecos à pesquisa que propus para o doutorado. Tive, no entanto, que recortar e reduzir esses objetos maiores. Enfim, tive de ir para o detalhe. Foi o que disse na introdução, sobre a fala, apenas:

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 de coisa que se pega de um lado e não se consegue pegar do outro<sup>5</sup>.

A fala a que me referi nos primeiros estudos (e depois os que se seguiram) podia ser tanto a fala propriamente dita, como o canto. Ao que comumente chamamos de fala, em minha tese de doutorado, chamei de fala falada e ao que chamamos de canto, fala cantada. Ao propor um estudo sobre a produção da fala cantada, com base ao que já se sabia largamente sobre a fala falada (com tradição de pelos menos 50 anos de estudos acústicos<sup>6</sup>), estava desenvolvendo um conhecimento sobre uma ponta do processo e não sobre a percepção diretamente. Mas o intuito era, na prática, desafiar um mito. O da falta de dicção por parte das cantoras.

Então, voltando à pergunta que é título deste capítulo, quero enredar as ideias num tecido apenas: *As canções foram feitas para serem entendidas?* Tentei mostrar que sim. O problema é que uma tese de doutorado com um título como a da minha tem o dom de afastar até mesmo interessados. “Descrição comparativa de aspectos fonético-acústicos selecionado da fala e do canto em português brasileiro”. No entanto, hoje, com olhar

---

<sup>5</sup> Raposo de Medeiros. Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro. Tese de Doutorado, Universidade Estadual de Campinas. Campinas, Brasil. 2002, p. 26.

<sup>6</sup> A primeira edição de *Acoustic theory of speech production* de Gunnar Fant, é de 1960. Neste livro é apresentada a teoria fonte-filtro que é tida por especialistas como uma proposta de entendimento da acústica da fala que perdura até a atualidade.

distanciado, parece mais fácil ligar algumas pontas que estiveram sempre presentes de forma mais ou menos evidente no meu trabalho.

A questão mais abrangente e latente desde o início em meus estudos/pesquisas, porém que não fora evidenciada, é a questão da brasilidade. Trata-se de nossa exposição a canções populares e a disposição maior ou menor (deve variar de indivíduo para indivíduo) de cantar nossas canções. De forma espontânea. Uma das experiências mais marcantes da minha vida e que costumo recontar em palestras, sobretudo a estudantes, apontou-me para quão cancionais somos. (que somos musicais, isto é muito comum ouvir de estrangeiros). Esta experiência relato mais à frente, no destaque: Saber música de cor.

Além da nossa afeição às canções, outra marca inescapável de meu objeto de estudo era, e ainda é, a língua em que se canta, em que cantamos: a língua portuguesa brasileira. Outro traço importante foi escolher um compositor brasileiro. Escolhi uma canção erudita de Francisco Mignone. Eis o defeito da proposta: se ela queria ser mesmo brasileira e ter brasilidade, teria de ser popular. Será. Mas aqui argumento em meu favor. A canção foi “Cantiga de Ninar”<sup>7</sup> e representava um elo com a dissertação de mestrado, na qual as preocupações estavam voltadas para o que se chama de aprendizado da língua materna em idade do ensino fundamental em seus últimos anos. Em sua epígrafe, já aderida, sem saber, à ideia de entender os movimentos do corpo e sua relação com nossas capacidades mais gerais, citando Pablo Casals ao se referir às crianças:

[...] We should say to each of them: Do you know what you are? You are a marvel. You are unique. In the whole world there is no child like you. And look at your body – what a wonder it is!  
Yours legs, your arms, your cunning fingers, the way they move...<sup>8</sup>

Acalentar naquele momento – no momento de escrever a tese de doutorado – uma criança bem pequena era uma tarefa diária, mas era também o fortalecimento de uma convicção de que uma sociedade inteira deveria se sentir responsável pelo acalanto das crianças do país todo. A canção escolhida também era parte de um repertório brasileiro e de um nacionalista. Em termos musicais, apresentava uma tessitura de notas médias e

---

<sup>7</sup> Ver belíssima interpretação de Sandra Félix em *Cantiga de Ninar - Francisco Mignone (1897–1986)*, disponível em: <<https://www.youtube.com/watch?v=jOOLjufqq8Y>>. Acessado em 20 nov. 2023.

<sup>8</sup> Casals foi um grande violoncelista e maestro nascido na Espanha. Sugiro a leitura de *Joy and Sorrows: Reflections by Pablo Casals*, Albert E. Kahn, Simon & Schuster; First Edition, 1970.

meio-agudas e, portanto, mais estreita do que a da ópera. No entanto, pertencendo à classe de canção erudita (podendo ser classificada como *lied*), a técnica vocal empregada para se cantar este tipo de canto é a mesma ou aproximada daquela de cantar a ópera, conhecido como bel canto. Além disso, a questão que vinha me fazendo desde antes de iniciar o estudo fonético- acústico era sobre a inteligibilidade do canto sem o microfone, o canto exatamente da chamada música erudita ou música clássica<sup>9</sup>. Por fim, Francisco Mignone, malgrado a marca europeia de sua música, trazia, além da temática, o ritmo espontâneo brasileiro para a orquestra (ver Maracatu de Chico Rei) e isso é um tipo de brasilidade nas muitas facetas do Brasil.

Obviamente a escolha do canto em português e mais precisamente da variedade do português a que chamamos de português brasileiro, termo que está bastante difundido, circunscreve meu trabalho a uma cultura específica. Assim, só posso dizer que os padrões de canto que encontrei com base nos dados fonéticos acústicos de qualidade das vogais cantadas e a temporalidade dessas em relação às consoantes cantadas, apontam para que sim, cantamos para sermos entendidos, não só musicalmente, mas linguisticamente, em português brasileiro. Se quiser estender isso para outras línguas, a hipótese será de que outras culturas, cantantes em outras línguas também tem a intenção de fazer seus textos cantados serem entendidos. Mas já não posso afirmar, não cientificamente.

O que é importante frisar a respeito do que se disse da brasilidade é o seguinte: a música brasileira é, por excelência, a canção. Qualquer e toda canção: da mais urbana até a mais ribeirinha e aquela que ecoa nas chapadas, nos igarapés, nos pampas. Sairei ainda um dia a coletar cantos brasileiros (Mário de Andrade fez isso um dia) e tentar mapear suas tipologias. Por enquanto, salva-me o site Radio Garden, sobre a qual falo no destaque: Todos os cantares e todas as canções do Brasil.

Mas volto ao gênero ou tipo de canção que escolhi como o objeto a partir do qual estudar a inteligibilidade da canção. Ainda que seja o *lied*, no qual as temáticas e as exigências musicais são diferentes do canto operístico, suas particularidades são basicamente as mesmas do bel canto. Primeiramente, trata-se de um canto sem microfone,

---

<sup>9</sup> Os estudiosos/ acadêmicos da Música nomeiam música erudita – e não clássica – todo o tipo de música que tem formas musicais definidas, compostas a partir de um conhecimento sistemático e formal de tradição europeia. Aqui, não farei essa distinção. Salvo quando necessário: por exemplo, quando houver referência ao período clássico (Mozart) em comparação a outro período, por exemplo, o moderno (Villa Lobos).

portanto, sem amplificação mecânica. Em seguida, sua tessitura, no caso da voz feminina no registro mais aguda (sopranos). A tessitura, ou seja, a faixa das notas musicais em que uma peça é cantada ou tocada pode ir de muito grave a muito agudo, ou ter uma faixa mais estreita. Pode-se dizer que fala apresenta uma tessitura estreita, salvo em situações de gritos de medo ou de dor. No canto, podemos explorar mais o recurso de passear pela nossa extensão vocal, que é a capacidade de nossa laringe de emitir sons de graves a agudos, obedecendo-se a um sistema escalar musical, ou seja, cantando afinado. Pois então. O *lied* escolhido, Cantiga de Ninar, tem uma tessitura que vai de um Mi 3 A um Mi 4, que é uma tessitura média.

Tanto por trabalhos de outros pesquisadores, como o do sueco Johan Sundberg, como pelo meu estudo piloto que levou em consideração, entre outras canções, a Bachianas número 5, já sabíamos que as notas agudas, cantadas por vozes femininas comprometiam a inteligibilidade do texto cantado. Diante disso, propusemos uma investigação complementar às já existentes sobre o canto erudito. A motivação dessa proposta era a ideia de senso comum no meio dos cantores eruditos de que a má dicção é a responsável pela falta de inteligibilidade do texto cantado, sobretudo por sopranos.

Embora, inicialmente, eu não focasse as hipóteses que levantei para descrever o canto comparativamente à fala na questão da má dicção, diante dos resultados, pode-se apontar a ideia de má dicção como um mito que é derrubado, tão logo desvendamos as manobras articulatórias que as sopranos fazem para cantar afinado, com grande volume e, no caso da canção erudita, sem microfone.

### **Todos os Cantares e Todas as Canções do Brasil**

Se quisermos falar do canto como uma capacidade humana especial, ainda que partamos de pequenos exemplos, é preciso lidar com a ideia da variedade, da diversidade e do tamanho do fenômeno. A pergunta “Qual é a canção brasileira?”, é, pois, inócua, se feita assim com a palavra no singular. Ou, então, colocamos a palavra canção no centro e em torno dela todos os tipos possíveis. Uma coisa interessante na nossa cultura musical é que, na maioria das vezes, nomeamos o gênero/tipo de canções a partir do nome do ritmo e não de sua temática. Assim temos: samba, samba-

canção, baião, carimbó, valsa, etc e etc. O elemento rítmico é que define o que ouvimos, musicalmente. O que me interessa aqui é delimitar a canção pela língua em que é cantada. Uma ferramenta digital do tipo streaming me ajuda a verificar, passeando pelas rádios brasileiras desde o norte até o centro-oeste, que há sempre cantos ou “ritmos” em português. E português do Brasil. Essa ferramenta, ou site, é a Radio Garden que oferece milhares de estações de rádio do planeta, nas quais podemos nos ligar dando apenas um clique sobre um ponto no globo terrestre. Ver essa maravilha em: <https://radio.garden>.

Outro aspecto que emergiu da curiosidade de saber se é inteligível aquilo que cantamos e de descrever o canto como fala cantada e seus traços em detalhe fonético foi o da nossa capacidade cognitiva para cantar. Que recursos temos para isso? O que acontece com a fala quando alguém canta? Na tese de doutorado, propus que há, um *trade-off*, ou seja, uma negociação entre a língua (que falamos) e música (que cantamos). De forma bastante informal e descontraída o que quero iluminar é o seguinte: quando cantamos, utilizamos recursos tanto da nossa capacidade linguística (ou seja, para a língua natural) como da música.

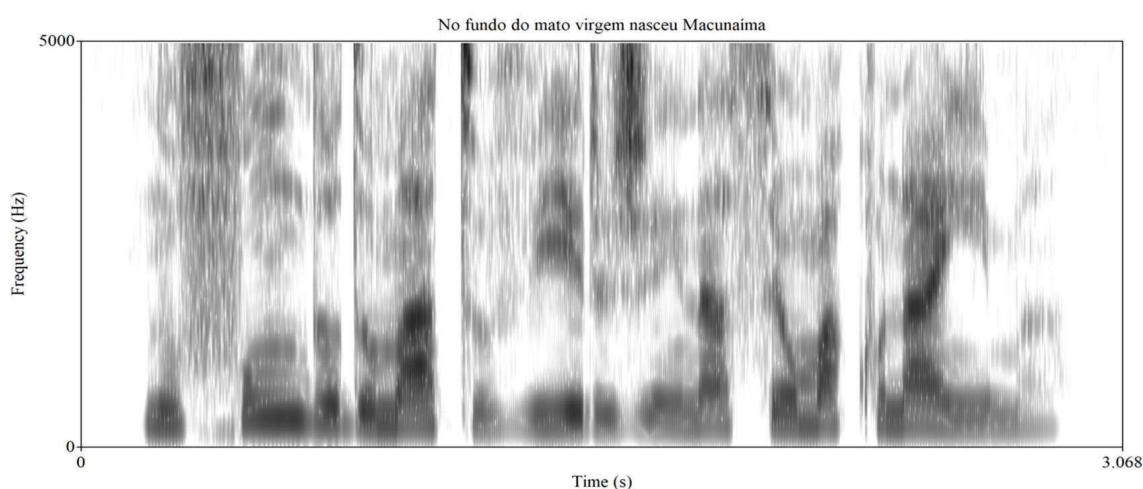
No entanto, antes de atingir essa explicação sobre o enlace fala e canto e sua doce negociação, quantifiquei as vozes cheias de luz das cantoras que vocalizaram, não mais a cantiga original do Francisco Mignone, mas uma frase adaptada para poder ser cantada por várias sopranos, repetidas vezes no laboratório de fonética. Onde a arte canta com piano a música inteira, iniciada pelos versos “Canto baixinho, uma velha canção de ninar”, a ciência canta, à capela, um fraseado deselegante que chamamos de frase-veículo: “Canto lapá baixinho...” Nesta frase-veículo, a palavra sem sentido “lapá” vai se alterando para “lapé”, “lapi” e outras sílabas finais diferentes para podermos observar um grupo de vogais e consoantes do português brasileiro representativas de nossa língua falada.

Então, a partir do fragmento de um exemplar em um mar de canções que é o cancionário brasileiro iniciei a análise acústica das vogais e consoantes definidas previamente. O fone, aquilo que parece uma unidade tão pequena e que proferimos tão rapidamente e vibra no ar por poucos centésimos de segundo, fica imenso. No pulsar e soar de um [i] temos uma infinidade de harmônicos da voz e podemos capturar sua



qualidade sonora, ou em outras palavras, seu timbre. É o que pergunto às alunas e alunos do curso de Fonética: qual é a cara acústica do [i]? Como soa esta vogal para você? O “i” soa agudo e isso tem uma explicação acústica. Não é mera impressão. Daí, então, posso explicar o que apresento como o primeiro achado da tese: o padrão articulatório específico do canto.

Inferimos o modo de articular os sons da fala a partir de dados acústicos expostos em uma espécie de fotografia da fala: o espectrograma (Figura 1). Podemos dizer que toda a imagem que se projeta ao longo do eixo vertical é informação sobre como ressoa nossa fala e nossa voz (ver também como essa imagem se estende ao longo do tempo). Essas informações são traços largos (ou faixas) mais escuros ou de cor mais intensa, mostrando que alguns harmônicos da nossa fala têm mais energia do que outros. Ora, essas faixas não são aleatórias e apresentam-se como um padrão, sobretudo por causa das distâncias entre si. Como todas as vogais, faladas e/ou cantadas apresentam tais harmônicos (ou mais tecnicamente, frequências de ressonância) é possível apresentar padrões quantitativos (medidas em Hertz) para falar da característica acústico-articulatória delas.



**Figura 1:** Espectrograma de enunciado falado (topo). No eixo x, temos o tempo (abaixo), de cerca de 3 segundos, com silêncios antes e depois do enunciado de cerca de 2.7 segundos. No eixo y, temos a frequência visível até 5 kHz (esquerda).

Por meio, então, de medições acústicas de cada vogal falada e cantada foi possível traçar um padrão acústico dessas vogais. Vimos que comparativamente à fala, o canto erudito tornava as vogais mais parecidas entre si. Isto quer dizer que manobras articulatórias extremas como abrir muito a boca para o [a] ou fechar demais para [i] eram atenuadas no canto. Chamamos a isso de manobra articulatória do canto e a relacionamos com a concepção do tubo retilíneo. A ideia de um tubo retilíneo para representar a cavidade oral não é uma invenção ou inovação. Ela pertence à fonética acústica, mas não entraremos em detalhes técnicos nesta parte do capítulo. Para o momento, basta pensarmos na nossa boca (ou cavidade oral) como um instrumento de sopro que pode tomar a forma de um tubo. Por exemplo: quando assobiamos, estaríamos alongando este tubo, quando sorrimos, estamos encurtando o tubo. O tamanho e o formato que esse tubo pode adquirir mudam o timbre, sobretudo, dos sons vocálicos. Graças à nossa capacidade de dar diferentes formas a este tubo, possuímos as unidades da fala: os fones!

Vimos então que o que as cantoras estavam fazendo era articular os fones de modo a saírem afinados e com volume, o que resultava em uma sobreposição das vogais, por exemplo, um [a] poderia soar como um [ó]. Isso, de fato, poderia prejudicar a pronúncia do texto cantado e, por consequência, comprometer a inteligibilidade. No entanto, não apresentamos um teste de percepção na tese, mas o fizemos logo em seguida e o resultado é de que quando mais aguda fica a nota musical, menor é a inteligibilidade da vogal.

Muito embora a manobra articulatória encontrada parecesse anular a distinção acústico-articulatória existente entre as vogais, havia outros aspectos a serem explorados para que pudéssemos hipotetizar explicações mais fidedignas e mesmo mais precisas sobre a questão da inteligibilidade da fala cantada.

Como linguista, ou ainda, como estudiosa da linguagem, estava apostando que, por mais que o canto erudito exigisse uma sonoridade tal que o afastasse da fala, os cantores cantam para que o texto seja entendido, de fato. Estava, pois, combatendo a ideia de que os cantores/cantoras cantavam tão somente para entoar sua bela voz, ou para cumprir uma tarefa musical. Ou ainda: soar como um instrumento – expressão muito utilizada por cantores quando querem explicar o que é cantar peças de Mozart, as quais exigem rigor técnico e contenção da expressão da emoção.

Trocando em miúdos, os aspectos que busquei olhar com detalhes no canto foram: como as vogais estavam estabelecendo distinções sonoras entre si e o que acontecia com a relação entre vogais e consoantes do ponto de vista da duração destes segmentos, ou fones. Para conseguir objetivar isso, usamos, além das frequências em Hertz, outra medida, chamada Bark<sup>10</sup>. Essa medida, diferentemente daquela em Hertz, traduz melhor o que o nosso ouvido ouve. Nosso ouvido funciona como um microfone e como um analisador de frequências. Capta todos os sons e pode discriminá-los de modo muito eficiente. As medidas em Hertz são extraídas de modo a espelharem a onda sonora captada (ou gravada), ao passo que as medidas em Bark, baseadas na onda sonora gravada, atribuem-lhe propriedades de vibração do nosso ouvido. Ao mapear os valores em Bark referentes a cada vogal cantada (a, e, é, i, ó, o, u), revelou-se uma espécie de afinação dos harmônicos da vogal com aqueles da nota musical entoada (um Lá bemol 3, correspondente a 392 Hertz). Além disso, as vogais na fala cantada (ou seja, quando são cantadas), não são deformadas de modo aleatório. Por exemplo: a vogal [a] pode soar como [ó], mas não como [i]. É possível que em tessituras agudas, o [i] seja entoado como [a]. Isso foi verificado em um estudo preliminar, em que observei o que acontecia com o [i] de “Cariri”, da bachiana número 5. Realmente, ali no agudo que tem de ser cantado com a última sílaba da palavra, portanto, [ri], se a soprano não tomar cuidado, acaba saindo uma vogal aberta, quase um [a].

Usamos metáforas para descrever o processo de ajuste das vogais faladas ao canto e, assim, dissemos que elas foram deformadas. Mas ainda deformadas, apresentavam distinções importantes e uma relação com as consoantes que entendemos como responsável pela inteligibilidade do texto cantado. Trata-se do aspecto da influência da consoante sobre a vogal, que, no canto, era bem menor do que na fala. Aliás, na fala, tal influência já é esperada e recebe o nome de coarticulação. Observamos que, além das consoantes cantadas manterem a mesma duração das faladas, apresentavam independência da “musicalização” das vogais. Neste momento, uso o termo “musicalização” das vogais para nomear, de modo mais intuitivo, o fenômeno de pareamento dos harmônicos da fala com os harmônicos da nota musical. O pareamento

---

<sup>10</sup> Traunmuller, Hartmut. Auditory scales of frequency representation, 1997. Ver mais em <https://search.r-project.org/CRAN/refmans/vowels/html/convert.bark.html>. Acessado em 20 nov. 2023.

foi um dos achados principais da tese: as vogais cantadas passam a soar como uma nota musical. Por seu lado, ao mesmo tempo, as consoantes pareciam estar blindadas em relação a isso, e não ofereciam nenhum vestígio de influência dessa musicalidade. Para aprofundar a exposição sobre estes detalhes acústico-articulatórios precisaríamos ser mais técnicos, mas não o seremos aqui. O fato é que só pudemos afirmar com certeza que o canto busca ser inteligível, mesmo com manobras articulatórias especiais, por causa dos detalhes fonéticos investigados. Muito provavelmente, se não tivéssemos perscrutado a vida acústica existente no vão vogal/consoante, estaríamos em um estágio muito inicial, concordando com as pessoas que dizem que as cantoras têm má dicção.

Vejamos quanta coisa, então, está em jogo, quando cantamos, não importa muito em que ritmo ou estilo. Em um sistema que podemos considerar complexo, a produção de um texto linguístico cantado requer nossos conhecimentos de música e língua natural. Atuamos ao mesmo tempo como falantes e musicistas e também como ouvintes da própria voz, que monitoramos, para manter a afinação e compreensibilidade linguística.

Como, então, podemos resumir as descobertas em relação à questão da inteligibilidade da fala cantada, que foi o aspecto motivador da pesquisa? Basicamente, dois achados devem ser apontados: o padrão acústico articulatório específico do canto e a capacidade cognitiva que chamei de negociação fala-canto. O primeiro tem a ver com o fato de fundirmos a entoação das vogais à de uma nota musical, de tal forma que uma é a outra e vice-versa. O segundo nos faz entender que para cantar, é preciso, às vezes, privarmo-nos do aspecto comunicacional da fala, portanto, embora cantemos para sermos entendidos musical e linguisticamente, em algum grau a inteligibilidade pode ficar comprometida.

No mais, se desafinamos ou se não nos fazemos entender, o que importa é conquistar o ato de cantar, tão maravilhosamente traduzido por Cecília Meireles no poema *Motivo*, que termina assim: “Sei que canto. E a canção é tudo. / Tem sangue eterno a asa ritmada. / Um dia sei que estarei mudo: / - mais nada.”

## Saber Músicas de Cor

Saber música de cor em português brasileiro quer dizer: saber cantar uma canção com toda a letra (ou quase toda a letra)! Certa vez estávamos em Belém do Pará, com dois convidados estrangeiros de um congresso. Um era alemão e outro irlandês. Conversávamos ali, em um bar-quiosque ao lado do Teatro da Paz. No violão, um músico local tocando um pouco de tudo em termos de MPB, entre outros. Lembro-me bem de tocar algo de Roberto Carlos. A cada canção que se iniciava, Graziela Bortz e eu cantávamos junto, lembrando quase toda a letra. Provavelmente esquecíamos alguns trechos. Mas isso não tirou o brilho da impressão que causamos aos nossos convidados. De certa forma, arrebatados pela nossa memória das letras, confessaram não ser nada comum saberem tantas letras de canções de cor, como nós sabíamos. Entendi aquela troca de impressões culturais como uma prova de nossa brasilidade.

### **Dados Suplementares sobre a Descrição Fonética do Canto em Comparação à Fala**

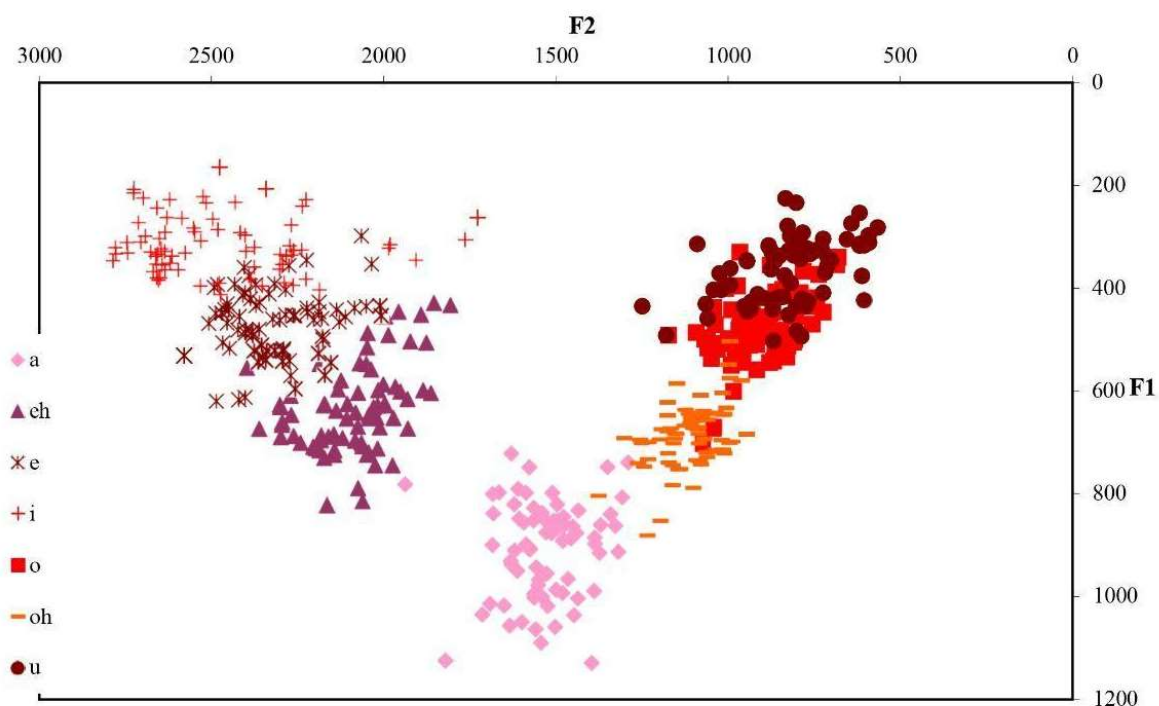
Queremos, nesta parte do livro, atar pontas. Em uma ponta, discorro de modo informal sobre a produção da fala cantada na canção erudita. Em outra ponta, apresentarei dados suplementares à tese de doutorado. Como vimos, essa tese levou em consideração a produção do canto, ou fala cantada, e comparou essa produção à fala falada, mostrando que as sonoridades diferentes têm sua razão de ser: a fala busca maior contraste entre as vogais; o canto, apesar de buscar os contrastes vocálicos presentes na fala, obedece à configuração do trato vocal que permite a nota musical soar como tal. Os gráficos, abaixo, contendo todas as vogais faladas e cantadas por todas as cantoras, demonstram essa diferença.

O desenho em forma de V (Gráfico 1), gerado pela posição que as vogais ocupam na área do gráfico x-y, mostra que essas soam tal como esperado para a fala: as vogais anteriores com valores de F2 mais alto [i, e, ε], a vogal baixa [a] com F1 elevado e as

arredondadas ([u, o, ɔ]), com F1 e F2 próximos<sup>11</sup>. Já o desenho em tons de azul do gráfico 2, apresenta as vogais mais sobrepostas, o que resulta em um desenho em forma de U achatado, com uma espécie de cauda à esquerda. Essa calda pertence à vogal [i] que tenta destacar-se das demais anteriores. Esses dois gráficos, contendo dados produzidos por todas as cantoras, resumem melhor o achado da tese sobre os padrões formânticos da fala (fala falada) e do canto (fala cantada).

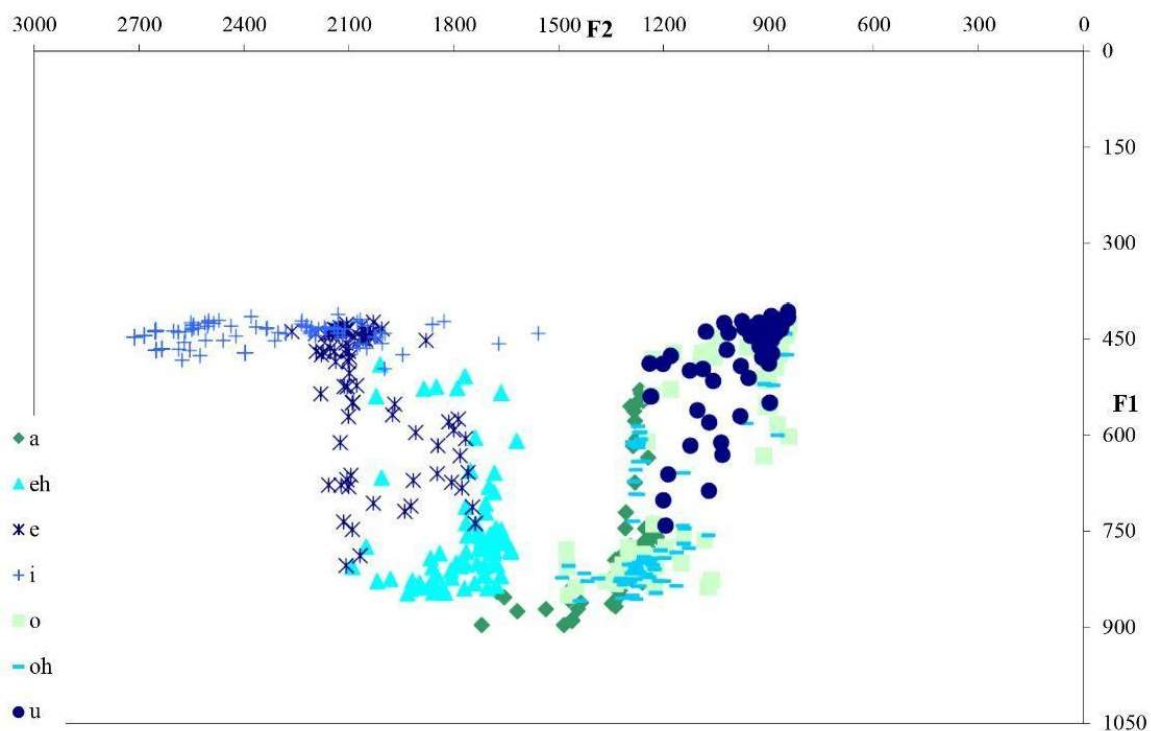
Em relação ao fenômeno articulatório que se infere a partir do comportamento acústico das vogais aqui apresentadas, houve uma tentativa de mostrar parcialmente o que estava ocorrendo, por meio de filmagens da face. Antes do término da tese, pude contar com a imagem e com a voz da soprano Marcia Guimarães (MG) capturadas em filmagem no NICS, Núcleo Interdisciplinar de Comunicação Sonora da Unicamp.

**Gráfico 1:** vogais faladas: F1 x F2



<sup>11</sup> F1 e F2 são abreviações para primeiro formante e segundo formante. A palavra formante nomeia uma propriedade de sons que possuem ressonância, como é o caso das vogais.

**Gráfico 2:** vogais cantadas: F1 x F2



Vale, assim, apresentar fotogramas oriundos das filmagens realizadas para a tese de doutorado (Figura 2) bem como medidas obtidas nos estudos apresentados nos eventos científicos.



/latu/ falado



/latu/ cantado



**Figura 2:** Conjunto de fotogramas obtidos da filmagem da cantora MG, sujeito da tese de doutorado. O momento da articulação é a porção medial da vogal-alvo: vogal da sílaba final do logatoma oxítono /laCV/

As imagens acima (Figura 2) mostram as diferenças de posicionamento dos lábios e mandíbula entre a fala e o canto e auxiliam em algumas explicações de achados com base nos dados acústicos. No entanto, não foram feitas medidas de distância entre articuladores.

Aprimorei as filmagens e medidas de abertura labial e abaixamento mandibular em trabalhos posteriores à tese de doutorado. As primeiras filmagens que fiz, nos primeiros anos de USP, geraram imagens como as duas abaixo (Figura 3), essas imagens desvendaram, na época, um trabalho em andamento: ainda usava fitas adesivas e o fundo de placa acústica de compensado e com furos, no Laboratório de Fonética antigo<sup>12</sup>. As fitas eram cortadas à mão. Havia também uma régua transparente fixada sobre a placa a

---

<sup>12</sup> Esta placa era muito comum nos anos 50 e 60 em estúdios. A cabine do Laboratório era bastante ampla e toda revestida com esse material. A cabine já não existe mais, tendo sido reformada e transformada em uma sala comum.



fim de servir de referência para os movimentos da cabeça (inevitáveis) e para a medição posterior.



**Figura 3:** Dois fotogramas obtidos da filmagem da face da pesquisadora (Beatriz Raposo). O momento da articulação é a porção medial da vogal-alvo: vogal da sílaba final do logatoma oxítono /laCV/. Esquerda: /laki/ falado. Direita: /laki/ cantado.

Para ambas as filmagens, seja “vintage”, seja a de “fundo infinito”, agradeço ao Didier Demolin pelo empréstimo da câmera de vídeo e doação de um notebook Toshiba para o Laboratório.

No estudo posterior ao de placa de furos, obtive imagens melhores (com “fundo infinito”, como na Figura 4) e refiz as medidas. A descrição técnica da coleta de dados é a seguinte, começando pelo design experimental:

Usamos o logatoma /la'CV/, tal qual na tese de doutorado, em que C (consoante) = [k,p,t] e V (vogal) = [a, eh, e, i oh, o, u]. A frase veículo era também a mesma já usada no doutorado: “Canto\_\_ uma velha canção de ninar”, extraída da *Cantiga de Ninar* de Mignone. Houve uma repetição para fala e uma para o canto. Os articuladores alvo eram aqueles visíveis a partir de uma filmagem simples da face: lábios e mandíbula. A filmagem foi realizada com uma câmera digital AG DV X100 Panasonic com microfone, a uma distância de 78 centímetros do sujeito (que era eu, de novo!).



/late/ falado

/late/cantado

**Figura 4:** Dois fotogramas obtidos da filmagem da face da pesquisadora (Beatriz Raposo). O momento da articulação é a porção medial da vogal- alvo: vogal da sílaba final do logatoma oxítono /laCV/.

A partir dos marcadores colados à face, puderam-se medir as distâncias, em centímetros, que assim foram designadas: ES – M: entre sobrancelhas e mandíbula; Ls – Li: lábio superior e lábio inferior; Ce – Cd: canto esquerdo e canto direito dos lábios. Usou-se o Adobe Photoshop7 que permite a visualização do fotograma e de uma régua na área de trabalho, nos eixos x e y, ou seja, do lado esquerdo e acima do fotograma. O marcador na ponta do nariz não foi alvo, mas funcionou como outro ponto fixo para se chegar a uma unidade de medida.

Vejamos, abaixo, a tabela com as medidas.

**Tabela 2**

	Fala			Canto		
	ES M	Ls Li	Ce Cd	ES M	Ls Li	Ce Cd
a	14,75	3,53	6,32	15,03	4,10	5,77
eh	14,70	3,37	6,47	14,75	3,75	5,93
e	14,53	3,07	6,32	14,23	3,40	6,08
i	14,3	3,07	6,32	14,1	3,38	5,98
oh	14,6	3,20	5,50	15,0	4,27	5,32
o	14,3	2,97	5,37	14,7	3,67	5,28
u	14,2	2,90	5,37	14,3	3,12	5,25

As medidas obtidas a partir do estudo relativo aos fotogramas da Figura 4, acima, mostram-nos que as medidas acústicas da tese de doutorado estavam coerentes com o padrão que se descrevia para o canto: a primeira ressonância do canto maior que a da fala. Lembrando que essa primeira ressonância (ou formante, que é o nome técnico) depende do abaixamento da mandíbula em conjunto com o abaixamento do lábio inferior. Na tabela acima vemos que a distância entre o ponto fixo (ES) e a mandíbula (M) é maior em todas as vogais do canto (valores em azul), exceto para [e] e [i], para as quais o aumento da distância se verifica nos valores relativos à distância entre Ls e Li (lábio superior e lábio inferior).

Alguns questionamentos persistiram a partir da discussão final desses estudos com imagens. A distância ES – M é bastante parecida entre fala e canto pois as diferenças são todas de menos de meio centímetro, a mais importante ficando para a vogal [ɔ] (ó), de 0,4 cm. Diante disso, indagamos se as diferenças de milímetros são importantes para a grandeza investigada. A distância Ls – Li parece mais robusta para explicar os aumentos de F1 no canto. As vogais [a, ε, e, o] apresentam maior abertura entre os articuladores labiais. A distância Ce – Cd indica que a aproximação dos cantos da boca no canto é maior no canto que na fala. É possível visualizar isso no fotograma da vogal [e] (Figura 4), o que é um dado qualitativo para atestarmos um arredondamento dos lábios no canto, mesmo para vogais não arredondadas.

Esses questionamentos surgidos a partir do que nos revelaram os estudos com filmagem da face apontam para detalhes da articulação na comparação da fala com o canto que a tese de doutorado não podia alcançar. Um último questionamento pareceu muito pertinente: deve-se partir para uma investigação de articuladores internos, como a língua? Ao obter imagens do movimento da língua, seria possível verificar se essa apresenta uma espécie de abaixamento compensatório na produção do [i] cantado. Em outras palavras, a mandíbula não abaixaria tanto, em compensação, a língua faria um abaixamento no espaço que lhe cabe na cavidade oral. O papel dos lábios também é o de uma manobra articulatória compensatória e vimos isso com as medidas das distâncias nesse último estudo das filmagens da face.

Vê-se então que vários estudos poderiam se desdobrar a partir do que fizemos filmando a face ao falar e cantar. Os desdobramentos exigiriam técnicas de imagem como a ressonância magnética e a ultrassonografia, por exemplo. Infelizmente para essas técnicas, fica impossível assegurar o canto em condições adequadas, já que para a ressonância magnética é preciso que o indivíduo fique deitado e para a ultrassonografia, há uma espécie de capacete que limita os movimentos da mandíbula. No entanto, cabe ressaltar que a simples necessidade de instrumental originário das ciências e práticas médicas para descrever o comportamento articulatório do canto já demonstra como a corporeidade presente no ato de cantar, assim como no ato de falar, pode se tornar um objeto de estudo efetivamente com detalhe e com relações objetivas com vários aspectos da voz cantada.

Neste meio tempo, ou seja, no qual se pensava com vários senões em como colocar um sujeito cantor/a dentro de uma máquina de ressonância magnética, surgiram outras indagações que pareciam mais importantes para as áreas da linguística e da música, as quais estavam sempre se entrelaçando nessas buscas. Tratava-se de questões em torno da capacidade de cantar e quais seriam os pontos de intersecção entre fala e canto. Na Parte 2 falaremos um pouco mais disso, focando alguns trabalhos interdisciplinares.

Para terminar de atar as pontas aqui, espero que tanto os dados acústicos como os dados articulatórios apresentados tenham trazido alguma luz sobre tópicos relacionados àqueles listados no início deste capítulo e que podem muito bem figurar em um programa de investigação científica. A manifestação da voz para a fala e para o canto pode ser

estudada a partir de diversos pontos de vista e, aqui, a proposta foi levantar dados mensuráveis e passíveis de serem observados da maneira mais objetiva possível. A contribuição, além de residir no âmbito de compreender mais o canto a partir do que já se sabe da acústica e da articulação da fala, está no fato de a proposta geral apresentada ser complementar ao que já se sabe em termos impressionísticos.

Veremos, a seguir, como a Linguística e, particularmente, a fonética são importantes para o desenvolvimento de estudos sobre o canto, e que este último deveria ser objeto de estudo da ciência da linguagem.

**PARTE 2**  
**A Fala, O Canto & A Linguística**

Por caminhos que mal conseguimos mapear, salvo uma investigação historiográfica detalhada, os estudos sobre as diferentes sonoridades humanas que necessariamente as relacionem à capacidade de linguagem, de alguma forma, se perderam. Os estudos sobre a língua natural buscaram entender sua estrutura e funcionamento, descartando a necessidade de se compreender sua substância sonora realizada através do que conhecemos por fala. É bastante disseminada a ideia de que a fala é simplesmente um canal do código linguístico e não tem relação essencial com o significado. Todos os aspectos físicos e fisiológicos da fala constituem conjunto de objetos não relevantes para a linguística e os fenômenos ligados a tais aspectos são considerados paralinguísticos.

Assim, não nos surpreende o fato de nossa cultura acadêmica, desde os anos escolares infantis, até os anos de ensino superior, estarmos expostos a sistemas altamente especializados que carecem de uma proposta de criar pensamentos de fato transdisciplinares. Há certa obediência cega a fronteiras entre assuntos e, assim, o professor/a de língua portuguesa não ousa explorar o capítulo dedicado à fonética, pois está mais preocupado em trabalhar a escrita. Este é, a meu ver, um assunto complicado porque sinaliza para as crianças que o estudo das línguas não deve abranger a sonoridade dessas línguas, ou seja, sua fonética. Vincula-se o aprendizado do funcionamento de uma língua à escrita. Ao contrário de uma abordagem que privilegie a compreensão da produção dos sons da fala em todos seus aspectos, a impressão que temos é que o aprendiz está sendo distanciado dos sons, para construir uma conexão nem sempre óbvia com os símbolos da escrita alfabética, que chamamos de grafemas. Ou seja, a velha letra. Somos letrados demais e sonorizados de menos.

Neste capítulo defenderei que o canto pode e deve ser objeto de estudo da linguística. Mais do que isso, deve ser considerado como um tipo ou modalidade de fala. Proponho a existência de um programa de investigação para o canto na linguística partindo das ideias fundamentais de que o canto emerge da fusão de duas capacidades humanas e, portanto, é um fenômeno único e para ele pode ser pensado uma unidade primitiva ou unidade a ser explicada e que virá a elucidar fatos sobre nossa capacidade de linguagem. Claro é que, como os estudos sobre a fala, sobretudo no domínio da fonética,

encontram-se bem avançados, os estudos da fala cantada (canto) só podem se beneficiar desse conhecimento acumulado.

Para argumentar em favor de darmos aos estudos do canto um *status* de objeto da linguística e da linguagem, iniciarei uma discussão sobre o elo comum entre língua e música e, na seção seguinte, apresentarei um breve levantamento de dissertações e teses que constituiu em um artigo intitulado “A fala e o canto como objetos de estudos da linguagem”.

Ainda nos anos em que desenvolvia o doutorado na Unicamp<sup>13</sup> — cujo campus era florido pelos ipês no inverno ameno e pelos flamboyants no chuvoso verão —, intrigava-me o fato de que música e língua natural fossem tratadas com tanta distância, a ponto de alguns acadêmicos dizerem que nada havia entre as duas. E eu me fazia algumas perguntas inocentes: bem, se as disciplinas não se encontram jamais, porque a fonologia, portanto uma área da linguística fala em “harmonia vocálica”, “contorno melódico”, “coda” e “andamento”? Todos esses termos são da música, sendo que o termo harmonia é normalmente utilizado na física. Encontrava-me em uma espécie de *carrefour* da terminologia científica e iniciava, assim, uma observação e uma reflexão sobre os termos comuns usados para descrever fenômenos linguísticos e musicais, o que deu origem a um artigo que veio a ser capítulo do livro “Em busca da mente musical”, organizado por Beatriz Senoi Ilari<sup>14</sup>.

O capítulo “Em busca do som perdido: O que há entre a Linguística e a Música” (o livro nomeou-se depois do capítulo, fazendo ressoar a ideia da procura) tentou mostrar que o objeto estudado possui a sonoridade como substância comum e o que muda é como a terminologia é empregada. Um dos termos muito interessantes, analisado no capítulo em questão, foi, justamente, o termo “harmonia”, argumentando-se que é usado na fonologia como metáfora. Harmonia é um termo emprestado da música pela linguística e resume um fenômeno sonoro mensurável que a física, digamos assim, conseguiu capturar.

Tentemos compreender a ideia mais geral de harmonia, a partir da ideia de harmônicos. Todas as notas musicais contêm harmônicos, pois a natureza física do que

---

<sup>13</sup> Como já dito, fiz meu doutorado de 1998 a 2002 no Laboratório de Fonética e Psicolinguística Experimental, nomeado pela simpática sigla Lafape. O Lafape fica no IEL, Instituto dos Estudos da Linguagem, no campus de Barão Geraldo em Campinas, São Paulo.

<sup>14</sup> Beatriz Ilari. Em busca da mente musical. Curitiba: Editora da UFPR, 2006



chamamos notas musicais são tons fundamentais que nosso ouvido capta como único e primeiro harmônico intenso ou de grande amplitude, mas que na verdade vem acompanhado de imperceptíveis múltiplos de si mesmo, sempre de ressonâncias mais altas. As notas ou tons fundamentais podem ser chamados de alturas, o que nos remete facilmente à ideia de sons graves e agudos. Tais tons ou alturas podem ser da voz, ou de um instrumento musical como a flauta, por exemplo. Esses tons combinam-se ou não combinam-se entre si: ou seja, podem estar em uma espécie de concordância ou afinação. Para que os tons fundamentais “se afinem” é preciso que seus harmônicos também coincidam. Em música, quando isso se dá, de maneira a se explorar as diferentes combinações para composição de uma peça musical, dizemos que o que está em jogo é a harmonia musical.

Para a explicação mais qualitativa e que apelou para a noção de elemento musical fundamental, que é a nota musical, há uma explicação correspondente no âmbito da física e da psicoacústica, que levará em conta necessariamente ideias básicas de acústica, como a natureza cíclica da onda sonora, as medidas em Hertz que caracterizam a onda como um tom fundamental e o entendimento mais matemático de como harmônicos de uma e outra onda estão em fase, ou seja, em maior sinergia que outros. Mas para saber detalhes disso, é só ir aos manuais de acústica que tratam de harmônicos e suas relações em termos de batimentos ou pulsos.

No entanto, a ideia mais básica de harmônicos e de harmonia é a seguinte: há eventos ocorrendo ao mesmo tempo e de forma a se ajustarem e combinarem. A partir dessa ideia fundamental é que a fonologia se apropria do termo “harmonia”. A harmonia vocálica, pois, tem a ver com uma especificidade da produção da fala que propicia a influência da característica acústico-articulatória de uma vogal sobre a outra. Em termos mais latos: "termo usado na fonologia para se referir a como a articulação de uma unidade fonológica é influenciada (está em harmonia com) outra unidade da mesma palavra ou frase"<sup>15</sup>. Em termos mais detalhados ou miúdos, o fenômeno em questão pode ser assim explicado: pensemos na vogal [e] tal como proferida na palavra “perigo”, que pode ser produzida mais como [i] do que como [e]. Isso acontece porque a vogal [i] da sílaba tônica

---

<sup>15</sup> Michael Kennedy. Dicionário Oxford de Música (tradução para o português Gabriela G. da Cruz e Rui V. Nery). Lisboa: Publicações Dom Quixote, 1994

parece exercer uma força sobre os movimentos articulatórios da vogal precedente, fazendo com que essa “combine” com aquela tanto em seus movimentos como em seu resultado acústico.

O exemplo do termo “harmonia” é apenas um dos casos de metáforas ou empréstimo na via da interdisciplinaridade entre linguística e música. Para não cair no simplismo de dizer que as duas diferentes áreas do conhecimento apenas tratam das mesmas coisas e podem ser aleatoriamente confundidas, o que não é verdade, chamei a atenção para o seguinte:

A música ressalta, reforça, exagera, condensa a paisagem sonora não musical; descobre, explora e elabora os harmônicos oferecidos pela natureza nos corpos vibrantes, extraindo alturas (*pitches*) que vão se combinar infinitamente gerando infinitas melodias. Neste arranjo e nesta condensação de sons, a música se instaura. E aí está a grande diferença entre o organizar da música e da fonologia. Ao organizar os sons a música faz música; e a fonologia não faz linguagem, antes a explica.<sup>16</sup>

É preciso ressaltar que a relação entre as duas disciplinas em tela não se dá só no domínio dos estudos fonéticos e fonológicos, domínio dedicado a explicar a gramática dos sons das línguas naturais. No tocante a unidades composicionais ou constituintes da música, há a teoria GTTM (do inglês *Generative Theory of Tonal Music* e em português *Gramática Gerativa da Música Tonal*) proposta por Fred Lerdahl e Ray Jackendoff, de visão gerativista sobre a capacidade musical. Ainda, na área das ciências cognitivas, incluída aí a neurociência, alguns estudiosos como Aniruddh Patel dedicam-se a mostrar que a sintaxe linguística e a sintaxe musical são compartilhadas. Patel faz isso com experimentos que observam o tempo de processamento dos sujeitos em tarefas que exigem, ao mesmo tempo, habilidades linguísticas e musicais, como preencher uma sentença com a palavra adequada e prever a nota de um acorde ou melodia para finalizá-los.

Assim, é possível dizer que, cada vez mais, os estudiosos de áreas como linguística, psicologia e música buscam a interação entre elas. A problemática ou desafio

---

<sup>16</sup> Raposo de Medeiros. Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro. Tese de Doutorado, Universidade Estadual de Campinas. Campinas, Brasil. 2002, p. 203

está no fato de que os estudos são diversificados e dispersos e não necessariamente dialogam entre si. Vejo que isto acontece por certo engessamento das partes envolvidas. Nem todos querem desenvolver suas pesquisas integrando conceitos que historicamente foram afastados uns dos outros. O conjunto de estudiosos que opta por circunscrever seu objeto de pesquisa a um dado terreno tem razão de fazê-lo, pois as especializações foram requeridas em dado momento para tornar possível a observação do objeto e sua análise. Por isso é que defendo que o programa que lance suas bases na via da interdisciplinaridade seja cada vez mais bem formulado e explicita seus fundamentos, sem prejuízo de uma convivência com outras propostas, e com o ganho de que as diferentes propostas sejam complementares, quando isso for possível.

No caso específico da minha proposta, a qual reúne trabalhos, em sua maioria, comparativos, os fenômenos estudados pertencem à fala e ao canto. Como vimos na Parte 1, a questão levantada requereu tanto a teoria como as ferramentas da fonética acústica no interior da linguística. Mostrarei, na seção seguinte, que outros estudos seguiram-se a esse, em sua maioria, no âmbito da linguística, o que é bastante curioso, já que a corrente mais conhecida e divulgada dessa ciência é aquela que observa, estuda e explica os fenômenos mais abstratos da língua: como a semântica e a sintaxe.

Mas antes vamos contextualizar a fala na linguística. De modo geral, a fala pode ser estudada como discurso na análise da conversação e na pragmática; e como evento físico produzido acústica e articulatoriamente. Em princípio, as abordagens das áreas que tratam do discurso linguístico não levam obrigatoriamente em conta a natureza física dos sons articulados da fala. O que está em jogo é a *práxis* do ato comunicacional que envolve a fala e isso tem de ser contextualizado a partir dos sujeitos falantes e do espaço ideológico que os circunda, ou que é produzido por suas falas. Algumas observações que levam em consideração o sinal sonoro podem auxiliar nas análises da conversação, como erigir explicações ligadas a pausas e contornos melódicos que compõem ou podem vir a compor o enunciado oral. No entanto, isso é feito de forma qualitativa. Já a fonética é a matéria responsável pela explicação da produção da fala em sua natureza física e fisiológica. Talvez os termos físico e fisiológico estejam um tanto distantes daqueles usados para explicar o que seja fonética, mas ao usá-los neste contexto, chamo atenção para o fato de

que a fala tem seu lado acústico, o qual não existiria sem os movimentos de partes do nosso corpo.

Anteriormente a este capítulo, foram levantados alguns exemplos de relações entre segmentos da fala e aspectos biológicos. Vimos que vogais seriam como que fisiologicamente diferentes das consoantes em termos articulatórios e aerodinâmicos. Essa já é uma visão diferente sobre os sons da fala na linguística. Mas há ainda uma proposta mais ousada, a de Robert Port, de que a classificação de vogais e consoantes como unidades da fala é um erro, já que memorizaríamos cadeias contínuas da fala e não segmentos. Esses segmentos seriam unidades artificialmente identificadas por uma contínua e maciça influência da escrita sobre a visão que temos da linguagem e por consequência, da fala.

As discussões sobre se há unidades da fala que podemos chamar de vogais e consoantes ou para usar termos mais genéricos, fones e/ou segmentos e qual seria sua natureza, levam-nos a questões mais profundas, no sentido de se colocarem em camadas iniciais para fundamentação de um saber sistemático sobre o conhecimento humano.

Vejamos como posso colocar o que quero dizer de uma maneira clara e organizada: partimos da ideia de que a fala é uma habilidade humana, assim como tocar um instrumento musical é uma habilidade humana. Então, uma questão que emerge é sobre quando e como começamos a falar em nossa história neste planeta. Até bem pouco tempo, na linguística, não era importante datar a origem da capacidade humana de linguagem. Hoje em dia sabe-se que a origem e a evolução da capacidade de fala dos humanos é um objeto relevante nos estudos da linguagem. Esses estudos foram rechaçados pela sociedade linguística francesa no final do século XIX e isso teve consequências, a meu ver, negativas para várias gerações de linguistas (entre essas, a minha).

Um exemplo disso é que perdemos visões como a de Ralph L. Holloway, na década de 70, sobre a origem da linguagem a partir de evidências neurológicas, estudadas a partir de métodos da paleontologia, ramo denominado, portanto, como paleoneurologia. Também parece ter passado despercebida para os linguistas a interessante relação do posicionamento da laringe e sua relação com a origem da fala articulada em humanos, em

um artigo 1988<sup>17</sup>, assinado por dois anatomistas (Jeffrey Laitman e Joy Reidenberg). Esses elos perdidos são retomados agora e até parecem novidade, mas não são. Ralph Holloway já falava de uma capacidade de linguagem a partir de evidências anatômicas relacionadas a habilidades manuais mais sofisticadas, como a fabricação de ferramentas e organização social para caça em grupo. Tais habilidades seriam ligadas ao desenvolvimento da linguagem que, por sua vez, viabilizava o que Holloway chama de controle social.

No entanto, o abismo aberto entre a linguística e outras ciências que desenvolveram ferramentas, teorias e conceitos para explicar as origens parece ter isolado a ciência de Saussure de achados no mínimo úteis para entendermos melhor o fenômeno da linguagem e mais especificamente a língua natural e um dos seus sistemas de realização: a fala. Como vimos, os estudos sobre a origem já relacionavam o fino controle dos movimentos das mãos com a capacidade de fala. Alguns estudiosos, como Maurizio Gentilucci e Michael C. Corballis aprofundaram-se mais na natureza gestual da fala para defender que os gestos manuais e faciais antecederam a fala do *homo sapiens*, espécie a qual pertencemos. Passamos de gestos mais amplos para gestos mais econômicos e que ainda possibilitavam a emissão de som. Emergiu, então, de forma gradual, a fala articulada e sonorizada. Os gestos mano-faciais, não só precederam, mas conviveram (e ainda convivem) com a produção da fala. A explicação para a evolução gradual se dá em duas frentes de conhecimento diferentes, mas que se complementam: (i) o da neurociência, que explica o desenvolvimento do córtex de homínídeos, estrutura do cérebro responsável pelas ações voluntárias e (ii) o da psicologia ligada à fonética, que, em um experimento relativamente simples, mostrou que os movimentos da mão estão ligados aos movimentos da boca, pelo menos na espécie humana. M. Gentilucci e colaboradores mostraram que, ao pegar uma maçã e dizer uma sílaba, o indivíduo abria mais a boca do que ao pegar uma cereja e dizer a mesma sílaba. As sílabas enunciadas foram analisadas acusticamente, e verificou-se que o parâmetro relativo à abertura da cavidade oral indicava maior abertura quando se tratava da fruta maior.

---

<sup>17</sup> O artigo é “Advances in understanding the relationship between the skull base and larynx with comments on the origins of speech”, publicado pela revista *Human Evolution*, volume 3, nas páginas de 99 a 109.

Desvenda-se, assim, um horizonte promissor para as pesquisas linguísticas e tudo parece apontar para a relação das capacidades consideradas historicamente como abstratas e as menos abstratas ou físicas e corporais. No entanto, a tarefa não é fácil, por várias razões, e uma delas é porque o pensamento humano é sempre estudado como algo fora do corpo e rotulado como mental. Mas não adentrarei este terreno, pois seria o caso de escrever um novo capítulo que fugiria aos objetivos deste trabalho.

Em alguns momentos, penso ter sido prudente separar a fala (corpo) da língua (mente) na linguística, pois explicar a relação do físico com o simbólico em nosso desenvolvimento é uma empreitada enorme. Um artigo que responde muito bem às minhas indagações em torno disso é o de Michael Studdert-Kennedy e Louis Goldstein, datado do início do século XXI.

Digamos que as ideias defendidas por esses autores ficam em um meio termo entre entender a fala como formada apenas de segmentos discretos (vogais e consoantes) e a ideia radical de Robert Port de que tais segmentos não existem. Para chegarem a uma explicação mais convincente sobre a natureza da língua natural, Studdert-Kennedy e Goldstein apresentaram uma visão evolucionista da linguagem, lançando, a meu ver, uma luz especial sobre a questão. Esses autores defendem a evolução de uma capacidade fonética originada em movimentos repetitivos como os de mastigar ou sugar, portanto uma capacidade baseada na produção, que evolui a partir de ajustes (*attunement*) entre falantes. Esta afinação ou ajuste entre falantes devem ser considerados como fatores importantes para entendermos a evolução das línguas e sua organização fonológica, a qual foi constituindo-se pelo uso. Trata-se, pois, de considerar a ideia do desenvolvimento de órgãos vocais cujos movimentos adquirem-se na base da imitação (pensemos nas crianças aprendendo a falar) para a expressão/realização da língua. Essa ideia já aproxima sobremaneira duas pontas que estavam/estão desatadas há muito tempo: de um lado a corporeidade da língua/fala e, de outro, o aspecto abstrato da língua mental. Assim, na base dos gestos fônicos das línguas estariam os gestos pré-linguísticos (mastigar, por exemplo) e posteriormente a este “treino” de nível biofísico, emergiria o léxico e a sintaxe.

O ganho de uma visão como estas – que a meu ver deveria ser mais divulgada na linguística – é de que se toma um caminho mais seguro e aclarador em torno de questões

que se tem arrastado na ciência, como aquela em que a capacidade linguística só pode ser explicada pelo seu núcleo ser a sintaxe. Em outras palavras, a capacidade de ordenar elementos em uma sentença seria fundadora de nossa capacidade de língua. Os outros níveis linguísticos seriam decorrências necessárias para veicular o que é pensado, ou seja, aquilo que tem uma organização abstrata.

Até agora espero ter persuadido quem lê a pensar na fala como uma habilidade que, dada à complexa evolução pela qual passou, presenteia-nos com uma formidável ligação entre o movimento e o som e que é parte integrante da nossa capacidade de linguagem, e mais especificamente da nossa capacidade de língua natural. Conquistou-se a fala há cerca de 170 mil anos. Mas como teria surgido o canto?

A ciência de hoje é vasta, o que se por um lado responde a muitas perguntas, por outro representa um desafio para se desenvolver a interdisciplinaridade. Parece que estamos numa teia em que todos os fios se conectam de maneira importante entre si. Assim, podemos também estudar a origem do canto. Essa visão de origem e evolução é importante por que situa melhor o fenômeno, lançando luz em dois aspectos que acho cruciais: (i) o da consciência de nossas capacidades como espécie e (ii) a preservação e valorização de uma expertise na canção, que é o caso da música brasileira. Estudos acadêmicos de história da música brasileira, só mais recentemente deixaram de focar a tradição europeia, passando a considerar o que é música brasileira de modo mais diversificado. Essa atitude é importante, pois pavimenta toda uma via na qual se erige conhecimento sobre a nossa música, desvenda os modos de fazer essa música e faz ressaltar os casos geniais de criação cancional e de ritmos realizados por instrumentos de percussão.

Mas voltemos à pergunta da origem do canto. Pesquisas referidas por Morley (2014) correlacionam a postura da laringe com a capacidade de vocalização de primatas<sup>18</sup>. Hoje em dia é bastante disseminado o achado de que a base do crânio flexionada (ou curva) posiciona a laringe de modo a proporcionar a vocalização e por consequência, a fala. É como se tivéssemos ganhado um espaço para mover o tubo da laringe, que ao se elevar, permite frequências fundamentais mais agudas e quando se abaixa, mais graves.

---

<sup>18</sup> Iain Morley. A multi-disciplinary approach to the origins of music: Perspectives from anthropology, archaeology, cognition and behavior. *Journal of Anthropological Sciences*, 92, 2014, pp. 147-177.

Além disso, a cavidade oral mais distante do tubo laríngeo parece ser totalmente relevante para o espaço articulatório dos sons da fala. Esse posicionamento da laringe, chamado de moderno, foi encontrado no antigo homo sapiens, (*Homo heidelbergensis*) que vivia há mais ou menos 350 mil anos; que não falava, mas já possuía, digamos assim, um protótipo da anatomia do homo sapiens.

Algumas ideias sobre a origem da música ou capacidade musical apontam para fatores neurobiológicos, para o inatismo ou para bases exclusivamente culturais. A problemática que Isabelle Peretz vê na explicação simplista de que o aparato neural proporciona a nossa musicalidade é que, se esse aparato fosse o único responsável por essa capacidade, os macacos também teriam desenvolvido música. Quanto à habilidade canora do ser humano, parece-me plausível apontar o achado sobre a fala, que é o abaixamento da laringe. Acredito que estudiosos concordariam com a afirmação de que o trato vocal como tubo de ressonância é um embelezador do som que se produz na laringe e, portanto, a “nova anatomia” deve ter um papel importante no desenvolvimento da voz cantada.

Mais ou menos recentemente, destacou-se na cena da linguística gerativa um pesquisador que publicou com Noam Chomsky sobre a origem da fala: o nome dele é W. T. Fitch e sua publicação com o nome mais importante da linguística mundial foi “the faculty of language: what is it, who has it and how did it evolve?”, que saiu em 2002, valendo também referenciar outra publicação de dois anos antes<sup>19</sup>. Entretanto, antes da entrada do novo milênio, já havia um grupo de estudiosos preocupados em desenvolver uma linha de pesquisa voltada para compreender a origem da língua. Trata-se da sociedade de Origens da Linguagem (*Language Origins Society*), criada nos anos 80 e que publicou três volumes intitulados “Studies in Language Origins”<sup>20</sup> que reúnem vários colaboradores, abordando temas tão diversos quanto o da anatomia do crânio, da fisiologia do ouvido, só para mencionar alguns tópicos que desvendam capacidades ditas de nível baixo, a tópicos que relacionam o controle motor como uma organização básica para a origem da capacidade de linguagem e — em um nível mais alto — a questão da

---

<sup>19</sup> William Fitch. The evolution of speech: a comparative review, Trends in Cognitive Science 4, 2000, pp. 258-267.

<sup>20</sup> Studies in Language Origins, Volume 3. Editores Jan Wind, Abraham Jonker Editora John Benjamins Publishing, 1994, só para dar um exemplo de referência.



relação da língua (linguagem) com a consciência, possibilitada pela atribuição de significado a enunciados compartilhados por indivíduos falantes.

Ainda que altamente atraente, o assunto da origem da linguagem e sua diversidade de tópicos e modos de abordagem não são objeto deste trabalho, o que faz com que seja aqui destacado mais por sua importância estratégica para argumentarmos em favor de uma visão mais integradora, do que ater-me aos detalhes. A visão integradora que proponho, de fato, e sobre a qual apresento argumentos e achados, é aquela sobre a fala e o canto. Ela está colocada em perspectiva a partir de uma visão de linguagem que conceba um pensamento mais holístico. Neste pensamento, as relações entre os aparatos neurológicos e fisiológicos e a capacidade simbólica que vem à tona pelo enunciado linguístico são compreendidas como constitutivas umas das outras e não em ordem de causa e consequência. Essa compreensão nos obriga a olhar para a relação do corpo com a linguagem.

Proponho que uma guinada epistemológica nos ajudaria a entender melhor a realização sonora da fala e também do canto. Há todo um universo e, por que não, universos paralelos de estudos que podem explicar minúcias de nossas capacidades falantes e cantantes. Mas o que temos feito em nosso quintal?

Durante os dois últimos anos do doutorado, assim como ao longo do tempo, construí uma sala de interdisciplinaridade sem janela. Sobretudo nos primeiros anos era difícil pensar epistemologicamente. A sala foi construída com experimentos do tipo apresentado na Parte 1 deste livro. O pensamento mais teórico tomou corpo com o capítulo de livro já mencionado, o “Em busca do som perdido”<sup>21</sup> e também no artigo “Ritmo na língua e na música” de 2009<sup>22</sup>. O tema ritmo mostrou-se promissor para falar de pontos comuns entre fala e canto e veremos mais detalhes sobre isso na Parte 3.

A seção a seguir apresenta o levantamento de estudos de que já falei anteriormente, através dos quais pude contextualizar uma defesa do canto como objeto de estudo da linguagem. Essa seção apresenta um estilo acadêmico de texto, com algumas citações, mas o que vale é ficarmos com o sentido geral e não nos preocuparmos em entender os detalhes das teorias linguísticas.

---

<sup>21</sup> Raposo de Medeiros. “Em busca do som perdido: o que há entre a Linguística e a Música”. In B. S. Ilari. (Org.). *Em busca da mente musical*. 1ed. Curitiba: Editora da UFPR, 2006

<sup>22</sup> *Idem*. *Ritmo na língua e na música: o elo possível*. *Música em Perspectiva*, v. 2, 2009, pp. 45-63.

## A Fala e o Canto como Objetos de Estudos da Linguagem

Como promover a inclusão dos estudos sobre o canto, ou fala cantada, na área da linguística e em suas áreas fronteiriças? Inicialmente buscamos situar brevemente a linguística como ciência formal e predominantemente mentalista, para, em seguida, apresentar pesquisas defendidas em nível de pós-graduação sobre o canto e a fala. Por fim, propomos a existência de um programa específico para estudos que elejam o canto como objeto de investigação.

Em seu livro *A Ciência e as Ciências*, Granger (1994) dedica um capítulo a discorrer sobre o que difere os objetos, verdades e validações das ciências formais e das ciências empíricas<sup>23</sup>. O distanciamento do pensamento matemático dos fenômenos sobre os quais versa conferiu status de ciência formal às matemáticas, em relação a outras ciências como a física, por exemplo. Em se tratando da ciência da linguagem, houve um esforço maior para se chegar a métodos formais, que encontraram e encontram até hoje, em algumas áreas, suas bases no gerativismo. Depois de longos anos dedicados a questões prescritivas, a linguística encontrou-se em um momento descritivo, quando da época do comparativismo. Para alçar-se à condição de ciência, pensamentos programáticos foram instaurados no estruturalismo e no gerativismo, como veremos a seguir, visando à formalização da linguística.

A fala é a capacidade humana que se caracteriza pelo seu aspecto físico, individual e existente no nível da realização e, portanto, apresenta variabilidade não “capturável”, o que dificulta o estabelecimento de previsões (hipóteses) e axiomas. A apreensão deste fato leva-nos a entender por que a fonética, a área da linguística que estuda os sons da fala, ficou relegada a segundo plano, quando esta última tornou-se uma ciência formal. Sendo assim, os estudos que vão ser apresentados na *Conclusão e perspectivas futuras* tratam de fenômenos que respondem mais a uma ciência aplicada do que a uma ciência de base, muito embora seja preciso ter cautela para aproveitar os achados que acenem para reflexões teóricas.

---

<sup>23</sup> Gilles Granger. *A ciência e as ciências*; tradução de Roberto Leal Ferreira. São Paulo: Fundação Editora da UNESP, 1994.

## Ideias Centrais à Ciência da Linguagem

Uma história das ideias linguísticas indicaria que o significado, ou aquilo que deriva do signo, é o objeto de estudo, por excelência, da linguística. A gramática gerativa desloca, de certa forma, o foco para a capacidade sintática dos seres humanos, ou seja, a capacidade de movimentar e ordenar constituintes, mas esta capacidade está a serviço do sentido, portanto do significado.

Embora não seja nosso objetivo explorar a história das ideias linguísticas, é importante notar que dada à necessidade de recortar e definir objetos de estudo, sobretudo a partir de uma ótica dualista (*langue/parole*, por exemplo), muitos aspectos da língua foram deixados de lado, ou apenas foram estudados de forma paralela à linguística. É o caso, por exemplo, do que se chamava, ou ainda hoje se chama, de paralinguístico e que envolve os aspectos de tom e ritmo da fala, ou o que é chamado não verbal.<sup>24</sup>

A partir da visão monolítica que se pode ter da linguística em seu desenvolvimento<sup>25</sup> é possível afirmar de maneira simples que a linguística ganha autonomia como ciência a partir da formalização do estruturalismo e passa a ser uma ciência explicativa a partir do gerativismo. À parte a crítica importante que Dascal tece sobre a visão monolítica e seu poder de homogeneizar os fatos do progresso da ciência, pode-se dizer que se consegue, com esta visão, abarcar as ideias centrais da linguística que se fortaleceram em inícios e meados do século anterior (século XX) e continuam sendo centrais até hoje. Um exemplo dentre essas ideias centrais é o conceito no nível abstrato e ideal de representação mental de unidades linguísticas. É sobre o nível ideal, insiste Chomsky (1962), que devemos nos basear para chegar a uma gramática formal da língua<sup>26</sup>.

As ideias até agora brevemente apresentadas: a de estrutura, a de gramática formal e o representacionalismo, se opõem à ideia de que os fatos da realização da língua, portanto os fatos da fala, sejam suficientes ou importantes para a explicação da capacidade linguística do ser humano. Neste sentido, os estudos da fala mantiveram-se mais como

---

<sup>24</sup> George Trager. *Paralanguage: a first approximation*. Studies in Linguistics, 13, 1958, pp 1-12. .

<sup>25</sup> Dascal, Marcelo. *Concepções gerais da teoria linguística*. Volume 1. São Paulo: Global Universitária, 1978.

<sup>26</sup> Noam Chomsky. “Modelos explanatórios em Linguística” [1962]. In: DASCAL, M. *Concepções gerais da teoria linguística*. Volume 1. São Paulo: Global Universitária, 1978, pp 61-95.

coadjuvantes, embora com força suficiente para grande desenvolvimento de métodos e instrumental. Além da evolução metodológica, influências da psicologia na compreensão da percepção de fala (por exemplo, a Teoria Motora de Percepção de fala)<sup>27</sup> proporcionaram, entre outras, uma autonomia da fonética que acabou por sua vez influenciando a fonologia e portanto a proposta de um paradigma fonológico em que o movimento do gesto (articulação por excelência) e o tempo de sua produção constam como primitivo linguístico, de caráter físico e abstrato. É o caso da fonologia articulatória.

Desta forma, entende-se que os estudos da produção e percepção de fala, ainda que se desenvolvam paralelamente a subáreas mais centrais da linguística, não deixam de ser metodologicamente e teoricamente elaborados e sofisticados. Epistemologicamente, tais estudos podem ser amparados pelo que Auroux (1998) propõe como nova versão do empirismo, o externalismo. Vale ressaltar um trecho do autor:

On a tort de trop concéder à l'intériorité. Il est vraisemblable que la capacité de calculer n'as pas commencé simplement dans la tête des gens et qu'elle ne serait pas née sans des manipulations d'objets externes (cailloux, abaques, etc.).<sup>28</sup>

### **A Importância da Fonética**

A razão de o canto ser estudado na linguística é porque esta é a ciência que desenvolveu melhor os estudos sobre a fala. Temos defendido desde Raposo de Medeiros (2002) que o canto é um tipo de fala<sup>29</sup>. Para diferenciar o canto da fala usamos os termos “fala cantada” e “fala falada”. É claro que esta denominação pode indicar um ponto de vista sobre o canto, possível, a partir da fonética. O estudo do canto também pode se realizar a partir de outros pontos de vista, como o musicológico e o da semiótica da canção, que em geral estão situados respectivamente em pesquisas de análise musical e do significado produzido no enlace entre sentido do texto e trajetórias melódicas. No

---

<sup>27</sup> Segundo avaliação crítica de S. Hawkins (1999), os teóricos da Motor Theory of Speech Perception não conseguiram provar que movimentos articulatórios são invariantes. Mas é inegável que a discussão em torno da percepção da fala é importante para a Fonologia na busca de estabelecer seu primitivo.

<sup>28</sup> Sylvain Auroux. *La raison, le langage et les normes*. Paris: Presses Universitaires de France, 1998, p. 7

<sup>29</sup> Raposo de Medeiros Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro. 2002. 177 f. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, 2002.

entanto, o que focamos neste artigo é a produção do canto e sua relação com a fala e seus aspectos fonéticos que podem ser acústicos, articulatórios e aerodinâmicos.

A fonética nunca deixou de pertencer aos estudos da linguagem, no entanto ocupou lugares de maior ou menor destaque dependendo das ideias centrais em um determinado período. É sabido que, no estruturalismo, teve grande repercussão a divisão entre fonética e fonologia em *Princípios de Fonologia* de N. Trubetzkoy. A ideia de uma fonologia que não precisava da fonética foi caindo por terra a partir de visões como as de J. Ohala, P. Ladefoged e J. Laver, a partir de fins dos anos 1960. Arriscamos dizer que devemos grande parte do progresso dos estudos experimentais em fonética e fonologia entre o final do século XX e início do século XXI a esses estudiosos.

A dedicação para colocar a Fonética à altura de uma disciplina linguística, tanto na organização de seus tópicos<sup>30</sup> como nas evidências de como dados fonéticos são imprescindíveis a teorias fonológicas<sup>31</sup> ou ainda na contribuição de Laver (1976) para uma melhor compreensão sobre qualidade de voz e como ela se relaciona às questões linguísticas<sup>32</sup>, serviu de base para o importante progresso da fonética e da fonologia mencionado na seção anterior. Também é possível depositar na conta da fonética a interdisciplinaridade hoje existente entre linguística e fonoaudiologia, só para dar um exemplo, sem mencionar as ciências da fala, de caráter multidisciplinar, que vêm unir ainda aos linguistas, engenheiros eletrônicos e da computação, entre outros.

É possível afirmar que, mesmo sem um programa pré-estabelecido, desenvolveu-se um território comum a várias disciplinas, as quais não concorrem entre si, mas complementam-se. Na seção que segue, ao apresentar os trabalhos elencados (ver tabelas de 1 a 5), teremos uma visão da complementaridade em tela.

### **Levantamento de Dissertações e Teses**

Foram encontrados 27 trabalhos que investigaram o canto, seja este comparado à fala, sejam seus aspectos acústicos, sejam seus aspectos fonológicos. A base de dados utilizada foi a BDTD (Biblioteca digital de Teses e Dissertações) acessando-se

---

<sup>30</sup> Peter Ladefoged. *Preliminaries to linguistics phonetics*, Chicago, 1971.

<sup>31</sup> John Ohala. "The contribution of acoustic phonetics to phonology", in: LINDBLOM, Bjorn, OHMAN, S. *Frontiers of speech communication research*. London: Academic Press, 1979, pp. 356-363

<sup>32</sup> John Laver. The semiotic nature of phonetic data, *York Papers in Linguistics* 6, pp. 55-62, 1976.

inicialmente a página: <http://bdt.d.ibict.br/vufind/>. As palavras de busca foram as seguintes: (1) acústica da fala e do canto; (2) linguística e canto; (3) fonologia e canto. Inicialmente priorizou-se buscar os trabalhos que teriam medidas acústicas, para em seguida verificar aqueles que apresentavam métodos relativos à fisiologia da voz e que, portanto, descreviam diretamente a articulação. Os primeiros foram mais numerosos que os últimos. Em seguida ponderou-se sobre a necessidade de verificar trabalhos versando sobre a fonologia e o canto, portanto de ordem mais qualitativa. Por fim, agruparam-se aqueles trabalhos que não se encaixavam em nenhuma das classificações: trabalhos sobre aspectos acústicos do canto em comparação com a fala, aspectos articulatórios do canto em comparação com a fala, aspectos acústicos do canto e aspectos fonológicos do canto.

A classificação privilegiou levantar trabalhos em que havia a comparação da fala com o canto. Parece-nos relevante, para defender estudos do canto no interior da linguística, que a fala seja uma referência importante. No entanto, isso não é obrigatório. Muitos trabalhos sobre acústica do canto, portanto, não comparativos à fala, são desenvolvidos no âmbito da linguística (ver Tabela 3) e isso não invalida nem a área do conhecimento nem o objeto estudado.

Outra questão a se destacar é a divisão entre acústico e articulatório, que jamais é total, pois um aspecto é relativo ao outro. No entanto, um dos trabalhos, isoladamente, classifica-se no quesito comparação canto e fala e predominantemente articulatório, é o caso de Nuñez, 2006 (Tabela 2).

Todas as tabelas são organizadas da seguinte maneira: a primeira coluna apresenta a área de conhecimento da dissertação ou tese; a segunda indica o sobrenome do autor ou autora, seguido do ano de defesa; a terceira apresenta o nível do trabalho e a quarta apresenta as medidas e/ou parâmetros utilizados para alcançarem-se os objetivos propostos. Por fim, a quinta coluna indica um aspecto específico do trabalho, como, por exemplo, o tipo de canção/canto observado, ou técnicas adicionais. Julgou-se mais apropriado deixar a informação sobre cada título no item das referências bibliográficas.

A Tabela 1 agrupa os trabalhos que se dedicaram a observar o que acontece entre a fala e o canto. O que há de convergente entre esses trabalhos é a explicação da interface fala/canto. Nesta questão, a linguística é importantíssima, pois é em seu âmbito, mais precisamente na fonética-fonologia (como disciplinas indissociáveis), que se desenvolveu

todo o conhecimento sobre a organização sonora das línguas. Portanto, nestas pesquisas, procura-se entender o canto, ou a fala cantada, em sua interação com a fala em seus aspectos segmentais e suprasegmentais. Trabalhos elencados na Tabela 4 também representam a compreensão do que acontece na fala cantada, a partir do que se sabe dos fenômenos fonológicos de uma língua, como acentuação, sândhi, paragoge, entre outros.

**Tabela 1:** Estudos Comparativos entre Fala e Canto

<b>Área</b>	<b>Autor/a e Ano</b>	<b>Nível</b>	<b>Medidas /Parâmetros</b>	<b>Especificidade</b>
<b>Linguística</b>	Medeiros, 2002	Tese	Duração; F0 –F3	Cantiga de Ninar, de Francisco Mignone; canto camerístico em português brasileiro
<b>Fonoaudiologia</b>	Rehder, 2002	Tese	F0, formantes; formante do cantor; vibrato	Avaliação auditiva qualitativa; vogal sustentada
<b>Linguística</b>	Pessotti, 2007	Diss.	Duração	Interação entre fala e canto de vários aspectos, sobretudo os prosódicos
<b>Linguística</b>	Pessotti, 2012	Tese	Formantes; intensidade; duração	Conselhos, de Carlos Gomes; escolas de canto
<b>Linguística</b>	Santos, 2012	Diss.	Duração de <i>lags</i> ; sincronização	Bossa-nova e rock
<b>Linguística Aplicada</b>	Lomba, 2017	Diss.	F0; intensidade; duração; decl. Espec.	Avaliação auditiva qualitativa
<b>Linguística</b>	Santos, 2017	Tese	F0; F1- F2; duração	Acento lexical na canção popular brasileira; a Banda de Chico Buarque

**Tabela 2:** Estudo Articulatório

Área	Autor/a e Ano	Nível	Medidas /Parâmetros	Especificidade
Fonoaudiologia	Nuñez, 2006	Diss.	Intensidade	Videofluoroscopia; avaliação perceptivo-auditiva; Bossa nova e samba

A Tabela 3 apresenta trabalhos predominantemente sobre a acústica da voz cantada e apenas da voz cantada. Em sua maioria, são trabalhos defendidos na área da fonoaudiologia. A tabela 5 agrupa trabalhos de aspectos variados, como a emoção e a pronúncia. Dentre eles, encontraram-se apenas dois defendidos na área da música, também os únicos nessa área no conjunto total das pesquisas.

**Tabela 3:** Acústica da Voz Cantada

Área	Autor/a e Ano	Nível	Medidas /Parâmetros	Especificidade
Fonoaudiologia	Oliveira, 2007	Diss.	Contorno de <i>pitch</i> ; F0; F1-F5	Avaliação auditiva qualitativa; voz de Roberto Carlos; Detalhes e Emoções
Linguística	Salomão, 2008	Tese	F0; H1-H2; quociente fechado, entre outros aspectos glóticos	Eletroglotografia; parâmetros articulatórios glotais relativos a registro modal e de falsete; canto coral
Linguística	Alves, 2013	Diss.	F0; F1-F3; intensidade, decl. espec. e LTAS	Videofluoroscopia
Linguística	Menegon, 2013	Diss.	Qualidade de voz	Utilização de Algoritmo: Evaluator Express aplicado ao Praat e VPAS
Fonoaudiologia	Loiola, 2013	Tese	F0; F1-F4; decl. espec.	Protocolos IDCM e IDCC



<b>Fonoaudiologia</b>	Silva, 2014	Diss.	Duração; F1-F2	Sujeitos: Cantoras líricas brasileiras. Peças de Schumann e Beethoven
<b>Fonoaudiologia</b>	Sousa, 2015	Tese	ELT; decl. espec.; F0 e formantes	Avaliação de auditiva qualitativa
<b>Fonoaudiologia</b>	Escamez, 2015	Diss.	ELT; decl. espec.; VOT	Melodia Sentimental, de H. Villa Lobos
<b>Linguística</b>	Waldow, 2015	Diss.	F0; intensidade; decl. espectral; ELT	Avaliação auditiva qualitativa
<b>Linguística Aplicada</b>	Coelho, 2017	Tese	F0; decl. espec.; LTAS	Dados dermatoglíficos; questionário de auto-avaliação; vogal sustentada e trechos cantados

**Tabela 4:** Fenômenos Fonológicos no Canto

<b>Área</b>	<b>Autor/a e Ano</b>	<b>Nível</b>	<b>Medidas /Parâmetros</b>	<b>Especificidade</b>
<b>Linguística</b>	Rostas, 2010	Tese	Não há	Avaliação de outiva de processos como monotongação, etc; reggae maranhense
<b>Linguística</b>	Migliorini, 2012	Tese	Não há	Análise de sândhi e paragoge; cantigas de Santa Maria
<b>Linguística</b>	Ricci, 2015	Diss.	Não há	Samba e bossa nova
<b>Língua Alemã</b>	Christmann, 2012	Diss.	Não há	Ária para voz feminina da ópera Der Freischütz

**Tabela 5:** Aspectos Vários do Canto

<b>Área</b>	<b>Autor/a e Ano</b>	<b>Nível</b>	<b>Medidas /Parâmetros</b>	<b>Especificidade</b>
<b>Linguística</b>	Rocha, 2013	Diss.	Não há	Pronúncia e IPA
<b>Música</b>	Beraldo, 2015	Diss.	Não há	Avaliação perceptiva-auditiva; emoções; Expanded Lens Model; Escala Emocional de Genebra
<b>Fonoaudiologia</b>	Tavano, 2011	Diss.	Posições dos articuladores laríngeos e supralaríngeos	Videofluoroscopia; produção de vogais cantadas do PB
<b>Linguística</b>	Machado, 2017	Diss.	Não há	Avaliação auditiva qualitativa de emoções na fala declamada e cantada
<b>Música</b>	Gusmão, 2014	Diss.	Harmônicos em 3-4 kHz	Fibronasolaringoscopia; formante do cantor

Do total de 27 trabalhos, 16 deles são da área de linguística. É interessante notar o aparente paradoxo que este número representa. De um lado, a linguística se mostra como a área do conhecimento mais fechada a eleger objetos como o canto, por outro, é a área em que se encontram as disciplinas com aparatos teóricos e metodológicos mais desenvolvidos, como a fonética e a fonologia, para tratar da produção de fala. A fonética acústica, por exemplo, conta com a teoria acústica da produção de fala<sup>33</sup> que tem como sucedâneo as correlações articulatórias. Soma-se a isso um conjunto de teorias fonológicas que hoje conta com um instrumental importante para desvendar manobras articulatórias da fala responsáveis por processos fonológicos explicativos da organização

---

<sup>33</sup> Gunnar Fant. Acoustic theory of speech production, The Hague: Mouton, 1960.

sonora das línguas. Tais teorias estão, podemos assim dizer, filiadas ao que se chama de Fonologia de Laboratório que é “uma posição metodológica dentro da fonologia que afirma que o estudo das representações fonológicas deve incluir o método experimental”<sup>34</sup>. Diante desta particularidade da Linguística, fenômenos de fala encontram excelente abrigo no interior da ciência da linguagem.

De um ponto de vista estrito, os trabalhos desenvolvidos na área da linguística não apresentam diretamente uma contribuição teórica. Então uma questão pode ser colocada: se os questionamentos em torno da fala cantada (canto) beneficiam-se dos métodos, materiais e teorias bem alocados na ciência da linguagem, não deveriam os estudos deles originados trazer uma contribuição importante para a linguística?

Uma resposta rápida e positiva seria: a contribuição, no conjunto destes estudos, é que complementam o que se sabe sobre o fenômeno da fala, já que esta é vista sob o ponto de vista da fala cantada, e iluminam discussões acerca de sua realização. Os fenômenos da fala são vistos como fenômenos a serem estudados empiricamente e, quando estudados como fala cantada, acabam por ter seu espectro alargado e, conseqüentemente, os achados tocam em questões tão instigantes como estas: deslocamento ou manutenção de acentos nas interações entre ritmos<sup>35,36</sup>, o aspecto opcional de fronteiras prosódicas<sup>37</sup>, configuração do trato vocal para diferentes estilos de cantar, comparados ao que se sabe sobre a configuração para a fala<sup>38,39</sup>.

Esses são achados das teses e dissertações em tela e que revelam que a linguística pode ter fronteiras mais fluidas, fato que proporciona uma postura crítica diante da relação entre as teorias e seus fatos, gerando novas perguntas diante de seu paradigma.

---

<sup>34</sup> Eleonora Albano. Fonologia de Laboratório. In: Hora, D., Matzenauer, C. L. Fonologia, fonologias. Uma introdução. São Paulo: Editora Contexto, 2017, p. 169.

<sup>35</sup> Cássio Santos. A sincronização em dois ritmos da canção: uma observação experimental acerca da fala cantada. Dissertação 104 f. (Mestrado em Filologia e Língua Portuguesa) - Universidade de São Paulo, 2012.

<sup>36</sup> Cássio Santos. Vogais cantadas e tonicidade: estudo experimental comparativo entre fala e canto com foco na duração. 2017. 129 f. Tese (Doutorado em Semiótica e Linguística Geral) – Universidade de São Paulo, 2017.

<sup>37</sup> Antonio Pessotti. O estilo na interpretação cantada e falada de uma canção de câmara brasileira: dados de cinco cantoras líricas brasileiras. 2007. 277p. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, 2007.

<sup>38</sup> Maria Rehder. Análise perceptivo-auditiva e acústica da emissão de vogal sustentada falada e cantada de regentes de coral. 2002. 81 f. Tese – (Doutorado em Fonoaudiologia) Unifesp, São Paulo, 2002.

<sup>39</sup> Antonio Pessotti. Efeitos do treinamento e da prática vocal profissional sobre o canto e a fala, 2012, 175 p. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, 2012.

Se o canto pode deslocar acentos e fronteiras, deformar vogais, encurtar consoantes e mesmo assim ser compreendido do ponto de vista linguístico, isso leva-nos a ver que obviamente há uma relação necessária entre o que podemos chamar de restrições linguísticas e musicais. Este fato por si já justifica os estudos da relação (ou relações) canto e fala e, portanto, justifica-se dedicar atenção a seus achados.

Dos 27 trabalhos elencados, 8 são defendidos na área da fonoaudiologia que, como sabemos, está intimamente ligada aos estudos da linguagem. Os trabalhos nessa área (Tabelas 1 e 3) são eminentemente sobre a acústica do canto, aspecto estudado a partir da acústica da fala que prevê medidas de formantes. O formante do cantor<sup>40</sup>, que é um pico de energia existente na faixa de 3 a 4 kHz, sobretudo nas vozes de tenores do belcanto, é normalmente investigado em estudos sobre a acústica do canto e serve como um parâmetro para se distinguir estilos de canto, por exemplo, erudito versus popular, ou lírico versus musical.

O elo comum à maioria dos trabalhos que trata da acústica do canto, tanto os da área da fonoaudiologia como da linguística é a qualidade de voz, aspecto que abrange questões de registro vocal<sup>41</sup>; obtenção de uma determinada sonoridade a partir de metáforas sobre o trato vocal<sup>42,43</sup>, configuração do trato vocal específico do canto popular<sup>44</sup>.

Vê-se, portanto, que os estudos sobre qualidade de voz, introduzidos por John Laver com base na articulação da fala na área da linguística, criam uma área convergente entre dois campos do saber que se distinguem um do outro, em suas ideias centrais, da seguinte maneira: de um lado o compromisso linguístico de chegar à forma representacional e abstrata da língua (para se falar do paradigma gerativo) e, de outro, o

---

<sup>40</sup> Johan Sundberg. The acoustics of the singing voice, *Scientific American*, Vol 236, 3, 1977, pp. 82-91.

<sup>41</sup> Gláucia Salomão. Registros vocais no canto: aspectos perceptivos, acústicos, aerodinâmicos e fisiológicos da voz modal e da voz de falsete, 2008, 196 f. Tese (Doutorado em Linguística) - Pontifícia Universidade Católica de São Paulo, 2008.

<sup>42</sup> Cristina Alves. Effects of metaphors in the teaching of singing: acoustic and vocal tract image data, 2013, 166f. Dissertação (Mestrado em Linguística) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2013.

<sup>43</sup> Paulo Menegon. Metáforas no ensino de canto e seus efeitos na qualidade vocal: um estudo acústico-perceptivo, 2013, 91 f. Dissertação (Mestrado em Linguística) - Pontifícia Universidade Católica de São Paulo, 2013.

<sup>44</sup> Karissa Nuñez. Configurações do trato vocal na fonação habitual e no canto popular. 2006, 37 f. Dissertação (Mestrado em Linguagem e audição: modelos fonoaudiológicos) - Universidade Tuiuti do Paraná, 2006.

compromisso fonoaudiológico de se descrever os sistemas fisiológicos da voz e avançar na solução de suas patologias. No entanto, o que foi dito até aqui, assim como o conteúdo que constitui os trabalhos de pesquisa em foco, é possível afirmar que, na prática de investigação, a distinção entre os dois campos científicos enfraquece-se.

Enfraquece-se também a distinção entre linguística, fonoaudiologia e música. Considerem-se os trabalhos defendidos na área da música<sup>45,46</sup>. Sobretudo em Gusmão<sup>45</sup>, a ligação com a acústica e articulação da fala é direta, buscando-se entender a correlação acústico-articulatória do formante do cantor.

Dentro da proposta de levantar o estudo em áreas limdeiras e assumindo que o canto é um tipo de fala (a fala cantada), encontramos também trabalhos que tratam da fonologia do canto<sup>47,48,49,50</sup>, apresentados na Tabela 4. Tais estudos não apresentam medidas acústicas e/ou articulatórias ou ainda de outra natureza (por exemplo, aerodinâmica) relativas à fala. Inscrevem-se na linha mais tradicional da fonologia de análises de oitava.

O objetivo da tabulação dos trabalhos encontrados é o de dar um panorama geral, ainda que breve, das características de estudos que versam sobre o canto e seus aspectos linguísticos. Na tabela 5, encontra-se descrito o único estudo que não tem uma ligação direta com a linguística, mas sim com a psicologia e cognição, por tratar da comunicação de emoções na canção. Os demais trabalhos estão irmanados, ou aos métodos hoje normalmente usados na fonética/fonologia de laboratório – medidas acústicas e/ou articulatórias – ou mesmo a ideias de modelos fonológicos que buscam dar conta de fenômenos para além da palavra.

---

<sup>45</sup> Cristina Gusmão. O formante do cantor e os ajustes laríngeos e supralaríngeos em cantores barítonos: uma investigação acústica e fibronasolaringoscópica, 2014, 97 f. Dissertação (Mestrado) Departamento de Música, Universidade Federal de Minas Gerais, 2014.

<sup>46</sup> Doris Beraldo. Feedback cognitivo aplicado a voz cantada no repertório musical brasileiro, 2015, 111f. Dissertação (Mestrado) - Universidade Federal do Paraná, Programa de Pós-Graduação em Música, Curitiba, 2015.

<sup>47</sup> Márcia Rostas. Balizas suprasegmentais para a adaptação do reggae cantado em São Luís, 2010, 212 f. Tese (doutorado) - Universidade Estadual Paulista, Araraquara, 2010.

<sup>48</sup> Lívia Migliorini. De versos e trovas: análise de aspectos fonoestilísticos do português medieval por meio das “Cantigas de Santa Maria, 2012, 206 f. Tese (Doutorado) - Universidade Estadual Paulista, Araraquara, 2012.

<sup>49</sup> Juliana Christmann. Processos fonológicos em fronteiras de palavras no Canto Erudito em Alemão, 2012, 152 f. Dissertação (Mestrado em Língua e Literatura Alemã). Universidade de São Paulo, 2012.

<sup>50</sup> Gabriela Ricci. Samba e bossa nova: um estudo de aspectos da relação texto-música, 2015, 119 f. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, 2015.

## Conclusão e Perspectivas Futuras

Ao longo do texto buscamos entender como os estudos sobre fala e canto encontram-se situados na linguística ou mesmo em áreas fronteiriças. Buscamos compreender como disciplinas de caráter externalista, como a fonética, convivem com propostas formalistas no interior da linguística. Vimos como a fonética, a partir da metade do século XX, fortaleceu sua presença nos estudos da linguagem. Juntamente com a fonologia, foi a disciplina que melhor desenvolveu o conhecimento sobre a fala no interior da linguística e, assim, acabou sendo adotada por estudos de natureza interdisciplinar que versam sobre fenômenos fronteiriços à fala. Vimos que mais do que um fenômeno vizinho, o canto é um fenômeno produzido pela coexistência de duas capacidades: a linguística e a musical. Temos defendido sempre a existência de uma negociação entre fala e música responsável pela realização do canto. Neste sentido, quanto mais questões forem levantadas sobre a relação fala e canto, mais nos aproximamos de explicações objetivas para fatos mal entendidos e, portanto, mal discutidos e mal solucionados. É inegável, pois, a contribuição que a linguística, de um modo geral, e a fonética e a fonologia, particularmente, podem trazer para os estudos sobre a fala cantada.

Em um levantamento futuro, que dê continuidade a este, reconhecemos que é necessário buscar artigos de periódicos em inglês e aprofundar a discussão sobre o elo entre a linguística e a música, inclusive em seu viés formal que conta com uma gramática da música<sup>51</sup>. Este, no entanto, é um assunto que foge ao escopo do levantamento atual e por isso não fez parte deste artigo.

Uma visão panorâmica sobre os assuntos dos trabalhos levantados na seção 3 pode ser dividida em quatro grandes conjuntos:

1. Questões de inteligibilidade
2. Questões de pronúncia
3. Questões de qualidade de voz
4. Questões prosódicas

Propomos, então, um programa específico no âmbito de estudos sobre a fala, em que se deve haver uma limpeza de terreno e em que muitas questões ainda estão por ser

---

<sup>51</sup> Fred Lerdahl; Ray Jackendoff. A generative theory of tonal music. Cambridge: MIT Press. 1983.

respondidas: Todos os gêneros de canto se diferenciam da fala de uma forma básica? Todas as formas de cantar, através das diferentes, apresentam essa mesma forma? Sílabas corresponderiam a notas? Se sim, como se dá tal interação?

Essas questões somadas a todas aquelas possíveis pertencentes aos quatro grandes conjuntos de questões, acima expostos, podem servir de base para o programa proposto e ponto de partida para novas investigações em torno dos fenômenos da fala e do canto. Se este trabalho situou epistemologicamente tais estudos e justificou o estudo do canto na linguística, esperamos que já possa ser visto como uma contribuição. Parte da contribuição já existe, de fato, nos trabalhos de pesquisa aqui apresentados.

**PARTE 3**  
**Ritmo na Fala e na Música como**  
**Organização da Temporalidade**



A batida nossa de cada dia no violão brasileiro é prolongamento do nosso saber rítmico. Adquirimos sem querer, de ouvido, em algum processo muito parecido com aquele que aprendemos as canções (a que chamamos de música) de cor. Se formos falar da batida do samba, marchinha ou bossa nova – ritmos urbanos – não há ciência que explique. Já sai sincopado. Bem sincopado no formato do samba. E aqui deixarei uma pergunta suspensa: mas qual é, objetivamente, o ritmo do samba?

Outra pergunta deve fazer parte deste capítulo: qual é o ritmo do português brasileiro? Ou ainda: qual é o ritmo das línguas em geral?

Se focarmos na primeira pergunta, podemos mostrar cantando ou fazendo a batida no violão. Se fizermos uma busca na *internet*, veremos que vários músicos e professores de música brasileiros explicam a batida do samba com exemplos, tocando. Poucos são os que descrevem essa batida como “o deslocamento do tempo forte”, pois isso já soaria técnico demais, além de não ser uma explicação completa. Explicar o ritmo do samba foi um grande desafio quando eu iniciei a estudar um fenômeno que chamamos de sincronização da fala em estudos fonéticos voltados para a organização temporal de nossa fala. Sobre a minha “descoberta” do samba naquele momento, tratarei mais adiante. Sigo desenvolvendo algumas considerações sobre ritmo.

Uma das melhores definições de ritmo que já vi foi “o ritmo é o metro que flui”, do estudioso alemão Curt Sachs. É uma definição que se presta lindamente para a música, mas não sei se pode ser aplicada a todos os domínios do conhecimento. A Fred Cummins, colega, amigo e colaborador de pesquisa, devo também uma maior e melhor compreensão do que seja ritmo. Para ele, ritmo está em tudo. O núcleo da questão é deslindar como o ritmo está em tudo. Seguramente a intuição que temos sobre ritmo das línguas não é a mesma que temos para ritmo musical.

Já pude levantar uma pequena discussão em artigos acadêmicos sobre o uso do termo ritmo, no dia a dia das pessoas, que acaba ganhando significados os mais curiosos e interessantes. Por exemplo: as pessoas podem dizer que são ruins para música, não conseguem cantar, porque não têm ritmo. Muitas vezes, esta afirmação que as pessoas fazem sobre si mesmas é traída pelo fato de que ao cantarem, mostram ter ritmo e, no entanto, desafinam. Parece que a expressão “Não tenho ritmo nenhum” informa que a pessoa não consegue produzir as notas nos tons esperados, ou seja, nos tons da escala

musical em que se iniciou a música cantada. Outra expressão reveladora é a seguinte “Se continuar neste ritmo, acabo hoje”, ou “Se continuar nesse ritmo, vou me cansar”. Nessas afirmações, o significado pretendido parece estar de fato mais ligado à noção temporal. Mas não tem a ver com o ritmo em termos de organização de um pulso (binário, ternário ou outro), mas tem a ver com a frequência no tempo em que um evento acontece. Em outras palavras, pode-se dizer que tem a ver com a velocidade de um evento durante um determinado período de tempo.

Levando-se em conta esses dois exemplos, acho possível dizer que nossa ideia mais geral de ritmo está intimamente ligada à música e, especialmente, a um desempenho musical individual, bom ou ruim e com a velocidade das coisas. No entanto, vejamos o que seja ritmo para outros campos do saber.

A psicologia busca entender como é nossa percepção de ritmo, e a linguística já abrigou esta ideia, a seu modo, por exemplo, para entender como se dá a percepção do início de um evento, por exemplo, o da vogal – vogal falada, obviamente. Em linhas bem gerais, para todos os fenômenos temporais, temos alguma resposta perceptiva, que seja atribuir uma organização às batidas ou eventos percebidos. Os fenômenos podem ser os mais diversos. O motor de um carro, o zumbido do ar condicionado, os passos de alguém no corredor, ou a água em queda numa cachoeira. O que estudos científicos dizem nesta área e em outras áreas vizinhas (como a física, por exemplo), é que mesmo que a ocorrência de eventos sonoros não mostre nenhuma batida saliente, nosso cérebro tem a tendência de agrupar essas batidas, ou seja, de fazer salientar uma delas. A isso chamamos de atribuir algum acento, mesmo quando ele não existe. De preferência de duas a duas. Isso resultou em metros binários na música.

Começamos a necessitar de mais e mais termos para descrever os fenômenos rítmicos. Os mais comuns são: metro, pulso e batida. Esses termos também têm sido usados um pouco como sinônimos entre si, um pouco com certa confusão. Na linguagem do dia a dia não há nenhum problema em se substituir livremente ou distraidamente tais termos. Assim como não há problema de o indivíduo desafinado se autointitular “sem-ritmo”. A questão é quando acadêmicos querem tratar o fenômeno e, neste caso, algum rigor é sempre necessário, mesmo para que, depois, seja possível derivar outras explicações mais claras e mais intuitivas sobre os fenômenos ou processos. Podemos

lançar mão dos modelos matemáticos para explicar o ritmo de forma mais objetiva, sem necessariamente relacioná-los a uma área ou disciplina específica do conhecimento formal acadêmico. E ainda lembrar os comportamentos cíclicos do nosso corpo, como as batidas do coração. Assim, da nossa percepção auditiva dos eventos no tempo, passando pelos movimentos do corpo e chegando ao ritmo silencioso do nosso organismo (entendendo a circulação sanguínea como ciclo ou ciclos) podemos dizer que ritmo está em tudo. Mas na questão da música e sua relação com a língua, como explicar diferentes tipos rítmicos?

Isso me faz lembrar dois episódios que aconteceram em um espaço de tempo bastante grande, correspondente aos meses e depois anos em que fiquei matutando sobre questões de ritmo. O primeiro tem a ver com aquelas batidas quase natas do violão. Para mim, o ritmo de samba não tinha nada de difícil, fazia apenas parte de um todo: cantar, fazer as harmonias e tanger as cordas do instrumento. Mas ao desenvolver um trabalho com meu amigo irlandês, Fred Cummins, dei-me conta da nossa polirritmia, que ele considerava ritmo complexo. Foram suas palavras ao ouvir um samba a capela, entoado por um sujeito de pesquisa: “Como explicar este ritmo? A música de vocês é feita de polirritmo”. Para mim, inicialmente, aquilo foi muito complicado. Descobri, lendo etnomusicólogos fantásticos como Simha Arom<sup>52</sup>, que são considerados polirrítmicos vários ritmos subsaarianos e que são tidos como complexos para os europeus. No entanto, são o *default*, ou seja, o mais comum ou mais básico para os africanos, como aqueles que ele pesquisou e cujo ritmo registrou com escrita da música europeia. O segundo episódio com caráter anedótico, e que ilustra bem nossa intuição rítmica e nosso fazer, que se sobrepõe ao nosso saber formal, é a explicação do Zeca Pagodinho das diferenças entre samba, partido alto e pagode.

Transcrevo o trecho da conversa no qual Zeca Pagodinho explica a diferença. Trata-se de entrevista realizada no programa de Jô Soares, em 1995. Zeca Pagodinho inicia dizendo que pagode é uma coisa e partido alto é outra coisa. Para fazer partido alto é preciso versar. É preciso saber. E para o pagode é preciso versar também.

---

<sup>52</sup> Simha Arom é um musicólogo ou etnomusicólogo que desenvolveu estudos sobre polifonia e polirritmia africana e entre seus livros publicou *African Polyphony and Polyrythm: Musical Structure and Methodology*.

## Qual a diferença?

Ai, Jô, sabe, essa é uma coisa meio confusa da gente responder. Há oito anos que me fazem esta pergunta... eu não consigo... e eu vou levando assim... É, é não é...

Risadas do público. E Zeca continua:

Que é que eu vou falar? Qual é a diferença? Não tem diferença... é a mesma coisa.

É igual, né? – acrescenta Jô Soares

É igual, mas é diferente – responde o Pagodinho. A plateia irrompe em longas gargalhadas.

Zeca Pagodinho já tem, pode-se dizer, o samba incorporado a seus saberes e fazeres e, para ele, explicar de modo formal coisas que são distintas não é uma questão. Posto isso, sua hesitação não é uma hesitação de fato, mas antes certa impaciência diante de uma tarefa que não faz sentido para ele. Talvez seja isso que separa o artista do cientista. O artista faz. O cientista quer explicar: detém-se diante do indizível e dá um jeito de dizê-lo. É aqui que me intrometo um pouco para dizer o que há de diferente. No entanto, ressalto, antes, que a diferença mais evidente entre pagode e partido alto é que este último é realizado por dois cantores que se alternam em um duelo para ver qual faz os versos e as rimas mais desafiantes para a deixa que, a seu turno, o cantor adversário tem de tomar para desenvolver seu verso e rima. Exige, pois, um saber mais sofisticado de versos cantados em ritmo de samba do executor. Já o pagode tem uma letra fixa, previamente composta, além de outros traços específicos.

Ambos os ritmos, partido alto e pagode<sup>53</sup>, são tipos de samba e, portanto, são binários. A maioria dos ritmos brasileiros é binária. Dizer que uma estrutura temporal é binária é atribuir-lhe um metro. A fim de se capturar esse metro, entendemos duas batidas como pulso fundamental, como se nosso cérebro fizesse uma espécie de separação de várias batidas, de duas a duas. Em seguida, dado que conseguimos dividir ou salientar as

---

<sup>53</sup> Originalmente, o termo “pagode” servia para designar a reunião festiva ou festa organizada nas casas das tias baianas, como a famosa tia Ciata. O termo pagode também é usado no universo da viola caipira.

batidas duas a duas – um modo de realizar isso é contar mentalmente 1-2, e até mesmo batendo palmas alinhadas a esta contagem – criamos compassos de duas notas que correspondem a dois eventos temporais ou duas batidas. Nos ritmos brasileiros e em todos os outros ritmos musicais binários de todas as músicas do mundo é possível adicionar diversas batidas que acontecem em torno desses dois pulsos, como subdivisões dele de modo a soarem sincopadamente. Em outras palavras, basicamente, trata-se de não alinhar as notas musicais aos pulsos binários que podemos chamar de pulso fundamental. Este desalinhamento pode se dar de muitas formas, como aquela que é exemplificada no artigo que é apresentado na próxima seção. Mas isso não será nossa preocupação agora, já que teríamos que descrever vários detalhes da estrutura e representação rítmica musical.

Devemos concluir, para distinguir os dois tipos de samba, que a diferença mais sutil e menos obviamente notada entre pagode e partido alto é a distribuição das notas musicais da melodia em torno dos pulsos binários. Uma observação atenta levantaria padrões rítmicos diferentes para uma e outra. Além das outras diferenças salientes – o papel dos cantores, as temáticas, o conjunto de instrumentos, o andamento, os padrões melódicos – os padrões rítmicos que apresentam diferentes formas de síncope já justificam o fato das variantes serem nomeadas de formas diferentes. Vejamos o destaque<sup>54</sup>.

### **Quantos sambas!**

Samba de roda, samba de breque, samba canção e tantos outros! Vale vermos e valorizarmos documentários como o “Partido Alto”.

Cada tipo ou variante de samba do Brasil merece um estudo detalhado. Erigir um programa que estuda o samba nas suas mais diversas manifestações, observando diretamente e profundamente seus elementos musicais é crucial. Deveriam existir manuais de Samba, assim como existem manuais de linguística. É necessário manter as diferenças do objeto estudado e respeitar a tradição oral do samba e a ideia noeliana do verso “Ninguém aprende samba no colégio” (Feitio de oração, de Noel Rosa). Os objetivos devem ser sempre devidamente explicitados e diferenciados. Estudar o samba de modo formal

---

<sup>54</sup> Documentário de Leon Hirszman sobre samba de partido alto, de 1976, apresentando a arte de Candeia, Manacéia, Paulinho da Viola e de outros.

assegura, por um lado, o entendimento de conceitos já existentes e o registro escrito ou de outro tipo que represente sua divulgação e sua transmissão. “Estudar” o samba para tocá-lo ou cantá-lo é outra coisa. Os gestos de abordar o samba como objeto de estudo e fazer o samba em si deveriam ser vistos como gestos complementares e solidários que nos devolvem nossa imagem cultural e referenciam nossa identidade musical e como ser-humano.

Espero que a explicação acima, apesar de poder soar um tanto especializada, ainda que não exaustiva, venha a ser complementada por uma percepção que nós já temos do ritmo musical do samba. Mais e mais detalhes poderiam ser expostos e outros ritmos como a valsa (que é ternária, mas pode ser também binária como dizia o genial Eduardo Gramani em uma divisão 6 por 8), poderiam ser explicados. Não é, entretanto, objetivo deste livro expor exaustivamente questões musicais, mas antes lançar os conceitos básicos que investigo e estão presentes em uma disciplina e outra para fazer a relação da fala com o canto, da música com a linguística.

Mas será que uma compreensão de ritmos musicais como a que descrevemos sobre tipo de samba pode existir da mesma maneira sobre a fala? Recupero a pergunta perdida no início do capítulo, que foi suspensa para as descrições do samba: Qual é o ritmo das línguas em geral?

Uma das empreitadas da linguística, sob o selo da fonologia e também da fonética, é estudar o ritmo nas línguas. Em termos de tipologia rítmica das línguas, essas teriam três tipos de ritmo, explicados por questões de acento, duração silábica e de uma unidade um pouco diferente da sílaba, a mora. Mas se isso satisfaz a um grupo de pesquisadores, não é verdade para muitos outros. Em estudos mais recentes, a tipologia é considerada arriscada e não diferencia muito as línguas já que, na verdade, uma mesma língua pode vir a ser tipificada com dois rótulos diferentes.

Para ilustrar um pouco esta problemática no interior da linguística e mais precisamente que emerge de estudos em que fonólogos e foneticistas se veem na missão de explicar o lado rítmico das línguas, recorro a uma espécie de experimento informal e coletivo que fiz com alunos de duas disciplinas: a Fonética Experimental, da pós-

graduação e Elementos de Linguística I, do ciclo básico do curso de Letras, entre final de 2019 e início de 2020.

Pedindo a atenção dos alunos aos dois áudios, em um momento inicial de aula, meu objetivo era obter respostas sobre classificação de um dado objeto sonoro. Um dos áudios integrava o vídeo que tinha sido filmado por mim, na Praça do Relógio da USP e, apesar da imagem da relva e de árvores frondosas, essa não foi mostrada, pois o que interessava era o som captado. O segundo áudio era um trecho de 13 segundos e fora isolado da imagem de um filme falado em árabe<sup>55</sup>. A pergunta para as duas escutas foi a seguinte:

— Que ritmo vocês ouvem?

Ainda que titubeassem um pouco para responder ao ritmo da batucada do vídeo, as pessoas acabavam classificando-o como música. As respostas foram escritas em uma folha de papel. Dos 55 ouvintes, 43 disseram que era um ritmo musical ou fizeram uma descrição do som como sendo musical. Dentre esses 43, 21 responderam que o som do vídeo era “samba”, usando exatamente essa palavra. Nessa descrição, empregavam termos de metalinguagem musical e houve até quem empregasse os termos “semicolcheia e semínima”, além de “síncopa e harmonia”. Entre os 12 ouvintes que não chegaram a uma classificação efetiva do aspecto musical do som do vídeo, havia a utilização de termos genéricos demais, como “ritmo contínuo” ou “acelerado e depois pausado”.

Por outro lado, para o som da língua árabe a resposta jamais foi de que o ritmo era acentual, silábico ou moraico.

A questão que se instaura após esta averiguação informal a respeito de como identificamos ritmo de linguagens diferentes é que somos proficientes para detectar e nomear um ritmo musical, mesmo que não tenhamos formação formal em música, mas não temos o mesmo desempenho para identificar o ritmo das línguas. Talvez possamos lançar mão de uma explicação simplista e dizer que isso acontece justamente porque as línguas não têm ritmo.

Entretanto, diante de um conjunto grande de pesquisas feitas em torno do ritmo linguístico, não posso fechar os olhos (ou os ouvidos?) e dizer que está tudo resolvido. Lendo acadêmicos profundamente envolvidos na problemática de se as línguas têm ou

---

<sup>55</sup> Trata-se do filme palestino *Wajib* de 2017, dirigido por Annemarie Jacir.

não uma batida isocrônica entre os seus segmentos de fala, posso afirmar – convencida pelos achados mais recentes – que em termos de língua e de fala podemos tratar de estrutura temporal ou organização temporal, mais do que ritmo. Ou usar os dois termos como substitutos, sabendo, no entanto, que falar de ritmo na fala, em princípio, não é a mesma coisa que falar de ritmo na música.

Uma diferença que considero grande, comparando-se a estrutura temporal da música com a da fala é a não-isocronia da fala em concorrência com a regularidade prevista das batidas rítmicas da música. Muito já se falou na linguística que, por mais que se tenha buscado isocronia entre unidades da fala (entre sílabas, por exemplo), isto não se comprovou. Sabe-se que as sílabas têm durações diferentes e por um longo tempo tentou-se provar que, ainda que as sílabas não fossem isócronas, seria isócrono o intervalo entre as sílabas tônicas no discurso falado. O problema é que as durações dos segmentos da fala são variáveis. Uma palavra de duas sílabas, proferida isoladamente, pode ter uma duração maior do que quando é proferida em uma sentença longa com várias palavras. Trata-se de um fenômeno um tanto elástico e mesmo compensatório: quando temos muitas sílabas para pronunciar em uma só sentença ou enunciado mais longo, a tendência é fazer “caber” todas as sílabas nesta sentença encurtando suas durações em termos relativos. Soma-se a isso o fato de que, possivelmente, nessa sentença, a palavra ou palavras “reduzidas” não receba(m) um lugar de destaque, por exemplo, sendo uma informação nova ou informação focada.

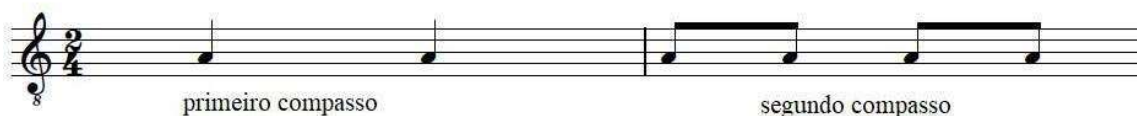
Na música, no entanto, uma vez definida uma série de notas em uma frase melódica, cada uma dessas notas pode ter seu tempo ideal pré-determinado. Posso ter unidades musicais – como o compasso – formando células binárias com duas, quatro ou oito batidas, igualmente subdivididas, como, por exemplo, no ritmo de marcha. A isocronia a que me refiro é o intervalo de tempo entre notas iguais: assim, se um compasso binário apresenta duas notas correspondentes a dois pulsos, o que se espera é que do início da primeira nota até o início da segunda nota haja um intervalo de valor temporal  $x$ , que deverá ser sempre o mesmo em compassos subsequentes iguais entre si.

É claro que as músicas são feitas de compassos variados e não apenas de marchas binárias de apenas dois pulsos. Mesmo que variadas em suas subdivisões e síncopas, e outros agrupamentos de notas que diversificam as figuras rítmicas, as unidades musicais



– ou frases musicais – sempre guardarão uma regularidade temporal entre as notas, de duração previsível. Assim, em uma velocidade considerada nem rápida nem lenta, que pode ser traduzida pelo acontecimento de um evento em intervalos de meio segundo (500 milissegundos), é possível pensar em um compasso de dois pulsos no espaço temporal de um segundo. O fenômeno de velocidade com que produzimos notas musicais é referido normalmente como andamento. É o andamento que define quanto dura uma nota musical. Se dentro de um segundo temos dois eventos igualmente espaçados no tempo, pode-se atribuir a cada um deles 500 milissegundos. E se forem subdivididos em mais duas batidas, teremos batidas de 250 milissegundos. Assim, num compasso como o da Figura 1, é possível dizer que o símbolo (uma semínima) que aparece duas vezes no primeiro compasso dura 500 milissegundos. O andamento atribuído a ela será o de 120 batidas por minuto, que poderia se chamar de 2 batidas por segundo ou 2 Hertz. Essa última nomeação nos dá a ideia de ciclos por segundo, portanto indica a quantidade de vezes no tempo que um evento ocorre. Salientando ou significando aspectos diferentes, esses conceitos referem-se ao mesmo fenômeno objetivo.

#### Exemplo de compasso binário com base na semínima em 120 BPM



**Figura 1:** O compasso binário notado nesta partitura-exemplo exemplifica duas notas mais longas no primeiro compasso e quatro notas menos longas no segundo compasso. No andamento estipulado, de 120 batidas por minuto, as notas mais longas terão 500 milissegundos ou 0,5 segundos de duração e as notas menos longas (colcheias) terão 250 milissegundos ou 0,25 segundos de duração.

O andamento, tal como estabelecido acima, pode se atribuir a outros eventos temporais humanos, por exemplo, passos da caminhada. Cientistas da área médica descobriram que o nosso ritmo mais comum para andar seria este de 2 Hertz, então é como se nossas passadas (cada uma, um evento no tempo) dessem-se numa organização temporal de duas a cada segundo<sup>56</sup>. Esta descoberta revela que nosso “estar no mundo” e

<sup>56</sup>“Marching to the beat of the same drummer: the spontaneous tempo of human locomotion”. Artigo escrito por Hamish G. MacDougall e Steven T. Moore e publicado no Journal of Applied Physiology em 2005.

uma grande quantidade de eventos temporais que produzimos gostam deste padrão de 500 milissegundos. Não vou adentrar a área da biologia, que é aquela que parece nos acenar agora. Assim, para fechar algo que iniciei dizendo, sigo à questão do ritmo cardíaco, ou melhor, o fenômeno da frequência cardíaca, já que esta está ligada a ciclos.

Ainda na escola aprendemos quão importante e vital é o bombeamento do sangue para oxigenar as células do corpo e sua característica de atividade involuntária ou visceral. Além dos ciclos sanguíneos, há outros que fazem parte do nosso organismo e que poderiam ser, por que não, estudados sob o tema do ritmo biológico do corpo humano. Poderíamos fazer ligações do ritmo biológico com a fala, como, por exemplo, a respiração e o tamanho das sentenças. No entanto, seria preciso outro fórum para discorrer com propriedade sobre as múltiplas relações entre as organizações temporais, ciclos e outros conceitos relativos ao funcionamento do organismo e movimentos corporais (dos mais simples aos mais sofisticados) que podemos colocar sob o guarda-chuva de ritmo.

Assim volto a uma comparação de dois universos bem mais próximos e que são os protagonistas aqui: o da língua e o da música. Seguramente, a estrutura temporal dessas duas linguagens é, em boa medida, diversa das estruturas temporais fisiológicas do corpo humano, uma vez que ambas são produtos de capacidades humanas simbólicas, ao mesmo tempo em que requerem gestos e movimentos precisos para se realizarem. Por exemplo, pensemos na diferença entre usar o trato oral para engolir um alimento e o mesmo trato oral, com a língua em ação, para proferir uma sílaba como “tlã” de *atlântico*. Para produzir uma sílaba, é preciso fina coordenação dos articuladores entre si. Em primeiro lugar, embora a língua seja protagonista nessa coordenação, a mandíbula deve estar ajustada em determinada abertura para que seja possível a língua, inicialmente, posicionar-se atrás dos dentes superiores, ficando ali por milissegundos para fazer o [t]. Em seguida, em vez de soltar-se desta posição, a língua mantém sua ponta no mesmo lugar para iniciar o [l], que só ficará completo quando as laterais da língua se abaixarem. Feito isso, a língua se soltará de sua posição e ao mesmo tempo o véu palatino se abaixará para desviar parte do ar para o nariz e assim conseguir o timbre nasal da vogal. Acontece, assim, não apenas uma sequência de movimentos, como também a concomitância destes movimentos, como é o caso dos gestos da abertura da boca, a soltura da ponta da língua a partir dos dentes

superiores e o abaixamento do véu! Já para engolir algo, também deverá haver aí uma coordenação de gestos, mas não será assim tão complexa, sobretudo porque não há intenção – salvo contexto específico – de se transmitir uma mensagem ao engolirmos.

Digamos então que esta capacidade refinada de colocar uma informação linguística em ação pela fala articulada e sonora desenvolve uma relação com o tempo que não é a mesma que se desenvolve para a respiração, ainda que desta dependa. Exposta a comparação entre fala articulada e a deglutição, também podemos acrescentar que a realização da língua falada se dá por esse corpo humano e, portanto, é dependente de seu funcionamento. Posto isso, podemos refletir sobre a fala e sua corporalidade ou corporeidade.

Já falei de quão diminutos são os sons da fala. Alguns deles podem ser alongados, como é o caso das vogais. Mas outros são fadados a serem eventos temporais brevíssimos, como é o caso da consoante chamada *tap* (“batida rápida” em inglês) e que apresenta, em média, 15 milissegundos de duração. No português brasileiro encontramos esse fone em palavras como “cara, para, sara, vara” e que comumente vemos descrito como “r” fraco. Embora os sons da fala, a que chamamos de fones, possam variar quanto à sua duração, essa variação não é *ad infinitum*, nem aleatória. Tanto para alongar vogais, quanto para alongar consoantes fricativas, como o “s”, o limite é nossa capacidade respiratória. Acabado o ar da expiração para produzir o fone, acabar-se-á o fone. É claro que o alongamento quase indefinido de tais sons num ato comunicacional do dia a dia é raro. Então, em bases corriqueiras, produzimos e utilizamos a fala obedecendo aos seus tempos intrínsecos, que por sua vez, resultam dos processos de articulação da fala. Então – e isso já é largamente sabido – os “is” que dizemos são em média mais curtos que os “as”, já que para articular esses últimos, a manobra articulatória exigida é o maior abaixamento da mandíbula, o que alonga a produção sonora. Por isso, é possível falar de tempos intrínsecos das vogais, assim também como é possível falar de tempos intrínsecos dos sons consonantais.

Não é só em termos temporais que os sons da fala se apresentam como intimamente relacionados à biologia do nosso corpo. Na base das ideias da fonologia articulatória – uma fonologia baseada na percepção dos movimentos dos articuladores da fala – está a de que as vogais distinguir-se-iam fundamentalmente das consoantes por

apresentarem uma espécie de fluxo contínuo que seria interrompido por uma produção de natureza oposta, ou seja, pelo traço de obstrução das consoantes<sup>57</sup>. A ideia é que a fala é um sistema capaz de ajustar dois movimentos articulatórios de naturezas diferentes e que, portanto, promove, de forma coordenada, sua coarticulação numa espécie de negociação entre imposições do aparato físico-biológico e requisições de um sistema mais abstrato. Além do tempo intrínseco das vogais, outro exemplo de que nosso organismo impõe limites à produção fonética é a questão da nossa tessitura vocal. Todos nós temos a intuição de que podemos controlar os tons graves, médios e agudos de nossa voz. Isso é possível por um ajuste de articuladores da laringe, que, basicamente, estende ou afrouxa nossas pregas vocais. Ao esticar e adelgaçar as pregas, podemos fazer sons mais agudos, e ao espessar as pregas, podemos produzir sons vocais mais graves. Ao som vocal que se produz por um determinado tempo e com uma altura determinada, chamamos de frequência fundamental ( $f_0$ ). Assim, variamos nossa frequência fundamental constantemente e tão naturalmente que, talvez, nem percebamos. O fato é que para conseguir uma voz muito aguda, por exemplo, que soe como uma voz de criança, uma pessoa com uma  $f_0$  normalmente grave irá até o limite de como estender suas pregas, que são, naturalmente mais densas e mais longas, mas nunca chegará a soar exatamente como a  $f_0$  do infante. A  $f_0$  da voz humana, pois, fica limitada a uma faixa ou tessitura.

Parece que houve uma digressão ao apresentar os exemplos de como a fala está conectada ao nosso corpo e, portanto, à sua anatomia, sua fisiologia, suas articulações e movimentos. Mas esse desvio foi importante para seguirmos entendendo como se estabelece a organização temporal da fala e do canto e, em termos gerais, da língua e da música.

Em trabalhos comparativos entre fala e canto, como os de Cássio Santos<sup>58</sup>, as indagações eram em torno do acento lexical da palavra falada em relação à palavra cantada. Sabemos que as palavras, em qualquer língua, têm uma sílaba acentuada. Mesmo unidades tais como os monossílabos átonos podem carregar acento, dependendo da sentença e de se essa é focada ou não. Uma questão pertinente a se levantar é qual, afinal,

---

<sup>57</sup> A ideia de Carol Fowler sobre as diferenças até fisiológicas entre vogais e consoantes é muito esclarecedora da organização temporal da língua. Vale a pena ver *Coarticulation and theories of extrinsic time*, no *Journal of Phonetics* de 1980.

<sup>58</sup> Cássio Santos A sincronização em dois ritmos da canção: uma observação experimental acerca da fala cantada. Dissertação inédita, Universidade de São Paulo, 2012.

a função do acento da palavra, a que chamamos também de acento lexical. Tenho a hipótese de que a palavra é um organizador temporal, assim como o compasso musical. Esta organização pode ser espacial também se pensarmos em articulação. Atendo-se, no entanto, só à questão acústica do acento, proponho a ideia de que este funciona como um atrator, ou seja, pertence a uma configuração de sistema tal capaz de atrair outros eventos sonoros, no entanto diferente dele, mas dele dependentes para formar uma unidade. Seria o caso das palavras a partir de duas ou mais sílabas. Também arrisco dizer que se não fosse esta organização, ou seja, se toda e cada sílaba que proferimos fosse acentuada, praticamente a fala soaria como que composta de monossílabos acentuados e levaríamos mais tempo para articular cada uma dessas unidades. Ou seja, seria uma cadência pouco econômica da corrente ou fluxo contínuo da fala.

Pensemos em um exemplo concreto para entender a organização das sílabas na palavra, em que apenas uma é a tônica:

Se repetirmos a palavra “batata” diversas vezes e formos alinhando uma batida do dedo com a sílaba tônica, ou seja, a segunda sílaba da palavra, perceberemos: (i) que essas batidas se dão ao mesmo tempo, portanto se sincronizam; (ii) e que as outras duas sílabas acontecem fora desse alinhamento, portanto não se sincronizam com a batida do dedo. Isso quer dizer que estão fora de fase em relação ao pulso substanciado pelo movimento do dedo.

Já falamos desse fenômeno, embora apontando para outro aspecto, a organização mais musical de eventos no tempo, referindo-nos a notas musicais correspondentes a batidas que se subdividem em torno de um pulso fundamental. Ora, a explicação sobre o comportamento temporal da palavra “batata” a partir da explicação de como funciona seu acento, está ligada à explicação do agrupamento de eventos para dar ritmo à música. A diferença é que o tamanho – e a imprevisibilidade – do conjunto de palavras a ser empregado para realizar a fala é muito maior do que o conjunto de unidades a serem empregadas em uma peça musical. É preciso lembrar, que a música – mesmo aquela improvisada – requer uma estrutura regular para seus elementos, a que pode ser considerada previsível. Embora as palavras propiciem uma estrutura que possa tornar-se regular, como é o caso do exemplo prosaico de “batata”, em sua produção comum em uma realização da fala, a intenção não é cadenciar o ritmo, nem fazer contornos melódicos

musicais. Quase ao contrário: os diversos tamanhos de sílabas, de palavras, de sentenças, bem como as hesitações para a escolha das palavras ou para formular as ideias, agem como forças opostas a uma organização temporal baseada em uma regularidade com fins de expressar justamente arranjos regulares e cadenciados: o ritmo musical.

Assim podemos dizer que o ritmo na fala e ritmo na música são e não são iguais. Ficamos no dilema do diferente que é igual e do igual que é diferente. Se são ritmos diferentes, como uma palavra se presta a ser representada por um compasso musical? E se são iguais, por que temos dificuldade de definir tipos rítmicos da fala? Proponho que indagações como essas deveriam nortear as pesquisas em torno dos aspectos temporais da fala/língua. Em termos mais efetivos, um programa de pesquisa sobre o tema deveria ter como perspectiva desenvolver estudos experimentais que espelhassem os fenômenos temporais da fala e da música. Um estudo possível, baseado na especulação com a palavra “batata” é observar o comportamento das palavras isoladas e levantar quais os padrões rítmicos em forma de compasso se nos apresentam. A partir disso, verificar, em modelos apropriados, se as métricas encontradas seriam mantidas em uma sentença, ou seja, no fluxo contínuo da fala em unidades maiores que a palavra. Em termos gerais, isso buscaria entender melhor a organização temporal de diferentes unidades da fala. Também poderíamos lançar mão de uma espécie de espelho ou experimento “ao contrário”. Frases melódicas de peças musicais apenas instrumentais poderiam ser subvocalizadas. De posse dessas sub vocalizações gravadas, teríamos como analisar a melodia e o ritmo, para em seguida, pouco a pouco ir transformando esta melodia em uma curva entoacional mais parecida com a da fala (manipulações do sinal sonoro são facilmente implementáveis em *softwares* de análise acústica da fala). Dessa maneira, seriam criadas condições similares à da melodia inicial, mas diferentes o bastante para supormos que os trechos seriam ouvidos como produção da fala. É claro que seria um experimento de percepção, já que a hipótese ou hipóteses dependeriam da escuta de indivíduos de cada trecho diferenciado. Sugeri os estudos acima, pois acredito que viriam a aclarar assuntos pouco explorados no espaço interdisciplinar existente entre música e linguística. Para mim esse espaço foi sempre muito evidente, ainda que, como vimos no capítulo anterior, tenha sido negligenciado e questionado por estudiosos, sobretudo aqueles do domínio da linguística,

já que esta é uma disciplina cujos objetivos, nos lindes científicos, estão mais bem definidos do que os da música.

No entanto, outro aspecto pouco explorado no reino do ritmo, em relação à fala, é o da sincronização da fala. Inspirada nos artigos de Fred Cummins sobre *entrainment* e nas inquietações de Cassio Santos sobre o desalinhamento do acento linguístico em alguns trechos do cancionário brasileiro, levantei algumas questões sobre o que haveria de comparável, em termos de sincronização, entre canto e fala. Seria a tarefa de sincronizar no canto mais fácil do que na fala? Os ritmos brasileiros, como samba e bossa nova já ofereceriam certa dificuldade para se sincronizar, comparados a outros ritmos?

Afinal o que é *entrainment*? E como se relaciona com os dois fenômenos: a fala sincronizada e deslocamento de acentos? A palavra *entrainment* e seu conceito encantaram-me por um longo período. Em uma de suas muitas explicações sobre *entrainment*, Fred Cummins diz o seguinte:

Where two processes are entrained, that entrainment must necessarily be based on some exchange among the processes, allowing the dynamics of one to influence the other. The first recorded example of entrainment was noted by Christian Huygens in the phase locked oscillation of the pendula of two clocks hung on the same wall (Spoor and Swift, 2000). In this case, the basis for entrainment was clearly a mechanical linkage between the two systems, as the coordination went away when the clocks were hung in different walls<sup>59</sup>

A fala sincronizada, assim como um conjunto de instrumentos tocando música juntos, é um tipo de *entrainment*, mas não mecânico. Para teorias fonológicas tradicionais, a capacidade de se realizar a fala sincronizada poderia ser explicada pela alternância rítmica de sílabas fracas e fortes ou por previsões de acento lexical. No entanto, como já dito, não há regularidade temporal, ou em outras palavras, não há periodicidade na fala. O achado de Cummins sobre a fala sincronizada é o seguinte: para estarem interagindo em tempos coordenados, os movimentos da fala estão também coordenados. Não se trata de termos um oscilador abstrato para a fala distribuindo mentalmente batidas fortes e fracas que depois, não sabemos muito bem de que maneira, produzem-se fisicamente.

---

<sup>59</sup> Fred Cummins. Rhythm as entrainment. The case of synchronous speech. *Journal of Phonetics*, 37, 2009, p. 6

Trata-se de uma capacidade de realizarmos movimentos corporais em conjunto, ao mesmo tempo, sendo esses movimentos corporais, os articuladores da fala.

Essa explicação vinha juntar-se à visão da fonologia dinamicista, da qual já falamos um pouco e que é conhecida entre as teorias fonológicas como fonologia articulatória. A visão de *entrainment* ligada à nossa capacidade de sincronizar a fala por meio do enlace dos gestos fônicos é também de base dinamicista que, mais popularmente falando, pode ser entendida como uma proposta que requer que abordemos e entendamos como se dá a relação entre a parte física e a parte abstrata de sistemas linguísticos, incluindo a fala.

Tentar compreender como se alinham as sílabas tônicas e átonas da fala – sobretudo as tônicas – com as batidas do compasso musical, é um modo de deslindar a integração de dois sistemas, quanto a suas organizações temporais, em termos de realizações de fato. Pensemos numa canção em que há o deslocamento acentual. Por exemplo, “O rio de Piracicaba”. O deslocamento acontece da sílaba “gar” para “jo” na frase “...vai jogar água pra fora”. Dentre as dezenas de sílabas cantadas nessa canção, apenas essa é deslocada, obedecendo a acento musical. Este é um exemplo do que chamo uma negociação entre língua natural e música: quando uma restrição musical (nesse caso o acento forte do compasso) requer da restrição linguística (o acento lexical) que esta se altere para que um terceiro sistema/fenômeno emergja, sendo esse fenômeno a canção ou o canto com palavras. Tal deslocamento, embora possa descaracterizar o signo linguístico falado no canto, não compromete a compreensão do texto. Podemos dizer, então, que o sistema linguístico de ritmo e o sistema musical de ritmo buscam alinhar-se na maior parte das vezes. Mas quando o desalinhamento acontece é um elemento bem-vindo na canção. E é um exemplo de dois sistemas dinâmicos influenciando-se mutuamente.

Diante de casos como esse, que são bastante intuitivos e chegam a chamar a atenção de amantes da linguagem e da canção, ouvido como deslize do cancionista ou como estilo, animo-me a propor outro estudo. A ideia é criar ou adaptar uma partitura musical, a partir de uma canção já conhecida, em que todos os acentos estejam deslocados. As questões mais gerais seriam: Como seria a inteligibilidade de uma canção dessas? A canção original ficaria reconhecível? Ela seria “cantável”?



E já que estamos falando de temporalidades, aproveito para dizer que, no mundo das pesquisas acadêmicas e científicas, a quantidade de ideias é inversamente proporcional à quantidade de tempo que temos para implementá-las. Arrebatada pelas ideias de *entrainment* na fala e pelo ainda não totalmente (naquela época) compreendido fenômeno da sincronização, resolvi enfrentar a questão do ritmo na língua em comparações com o canto, a partir, justamente, de um estudo sobre a fala conjunta e canto conjunto.

Tinha a meu favor o fato de falar junto, ou sincronizar falas, é uma ação bastante comum entre indivíduos. Em diversas ocasiões, como em atividades esportivas e em atos de protesto coletivos, lançamos mão de falar juntos. Torcidas de times esportivos têm seus “cânticos” que são entoados coletivamente e de forma sincronizada. Em relação à música cantada, diversificadas são as atividades que resultam em canto coletivo ou que já são previamente definidas como, que é o canto coral. Temos, respectivamente, exemplo de grupos de amigos que entoam canções conhecidas em uma reunião social e uma infinidade de peças compostas para serem cantadas por conjunto vocal desde, pelo menos, a Renascença<sup>60</sup>.

Mais uma vez estava diante do infinito. O dilema, então, que se coloca é aquele de tentar coletar milhares de dados ou decidir por observar o que acontece com alguns exemplares de fala e canto. Se optasse pela primeira proposta, teria que trabalhar com o que hoje se chama *big data*. Em palavras bem breves, levantar a maior quantidade de dados possível para uma pesquisa qualquer é viável, se esses dados estão disponíveis na rede mundial de computadores. Podemos, por exemplo, baixar o máximo de áudios (provavelmente serão vídeos, dos quais isolaremos a faixa de áudio) de torcidas entoando seu cântico e tratar esses áudios via algoritmos, para, daí extrair medidas e comprovar padrões, por exemplo: quais os intervalos musicais mais empregados em tais cânticos? No entanto, esse tipo de consecução de dados, conhecida como mineração de dados, pode tornar-se uma caixa preta para o pesquisador, já que não se tem controle das condições de gravação e de edição de sua fonte de dados. Ademais, seria necessário haver-se com vários aprendizados e procedimentos antes de partir para uma coleta como essa.

---

<sup>60</sup> Charles Nef. *Histoires de la musique*. Paris: Payot, 1948.

Assim, para responder a uma questão mais geral a respeito de se diferentes tipos de fala seriam mais facilmente ou mais dificilmente sincronizados, resolvi realizar um estudo experimental que contivesse tipos representativos de fala e canto. A ideia era observar desde um texto em prosa, falado, a um canto cujo ritmo espera-se muito periódico, que é o caso do *rock*. Meu parceiro nesta empreitada foi Fred Cummins da UCD School of Computer Science and Informatics, University College Dublin. Juntos, vimos que uma parlenda como "Hoje é domingo, pede cachimbo..." se comparada a um texto em prosa como "O relógio", de Clarice Lispector, não é melhor sincronizada por pares de falantes cuja tarefa é dizer o mesmo texto conjuntamente. Por outro lado, quando o samba "Preciso me encontrar" (Candeia) era comparado com o *rock* de Raul Seixas, *Aluga-se*, então verificamos que a batida mais regular deste último possibilita melhor sincronização entre cantores/cantoras.

Enquanto, humanamente, tento entender a temporalidade da fala e do canto, nas explorações da canção, o cancionista Caetano Veloso, divinamente, já traduziu nosso alinhamento com o tempo, entre o mistério e o deleite: “peço-te o prazer legítimo/e o movimento preciso/ tempo, tempo, tempo, tempo/ quando o tempo for propício/ tempo, tempo, tempo, tempo”.

**PARTE 4:**  
**A Questão da Estabilidade do Canto *versus***  
**A Instabilidade da Fala**

A capacidade de manter uma vogal alongada por tempo indeterminado ou produzir uma nota musical em uma vogal qualquer, pelo máximo tempo possível, é um recurso vocal fantástico. Estão acionadas para tal movimento a cavidade oral e a laringe, que podemos descrever como sendo dois tubos que se ajustam de modo a criar um instrumento vocal destinado a produzir efeitos estilísticos de fala (como o enunciado “Gol”, emitido pelo narrador de futebol) e notas musicais cantadas! Para ambos os fenômenos é necessário manter, digamos, assim, um modo estável do funcionamento da laringe. A isso chamamos de fonação sustentada.

Diante disso, e de tentar traçar uma linha que pudesse distinguir a fala do canto, propus, em um projeto de pesquisa, que o canto seria mais estável que a fala em termos de fonação, por tanto, acusticamente, produz e pode ser caracterizado por uma continuidade, ou estabilidade, ou constância. A fala, por sua natureza intrínseca, não oferece essa constância.

Tal característica, a da estabilidade, já estava se pronunciando nos trabalhos e estudos iniciais no meu percurso como foneticista. A maior duração das vogais cantadas e a necessidade de se manter uma cavidade oral sem grandes constrições já indicava isso. No primeiro capítulo, descrevemos essa cavidade oral como tubo mais uniforme para o canto, comparado ao da fala, que necessita constrições precisas para articular os fones. Depois constatamos uma sincronização maior para se cantar um tipo de ritmo, o *rock*. A organização temporal da música, por mais complexos que sejam suas frases ou células rítmicas, detém uma constância, diferentemente da produção da fala.

Se tomarmos a fala como um ponto de referência, ou *default*, pensemos da seguinte maneira: da fala para o canto há ajustes articulatórios necessários para que se chegue a uma sonoridade musical. Mas afinal como distinguimos o que é a música cantada da fala, a tal fala falada?

Não é preciso um lugar ou contexto específico, nem um posicionamento específico dos indivíduos ou instrumentos musicais, ou mesmo um ritual em ação para que diferenciemos ou percebamos a diferença entre uma voz falada e uma voz cantada. Estamos a todo tempo ajustando nossas escutas para diferentes tipos de som. Intuitivamente, também, sabemos que não se trata de um tudo ou nada entre a fala e o canto. Uma fala pode parecer cantada, assim como um canto pode parecer falado. Mas afinal de contas, que fator nos dá essa impressão?

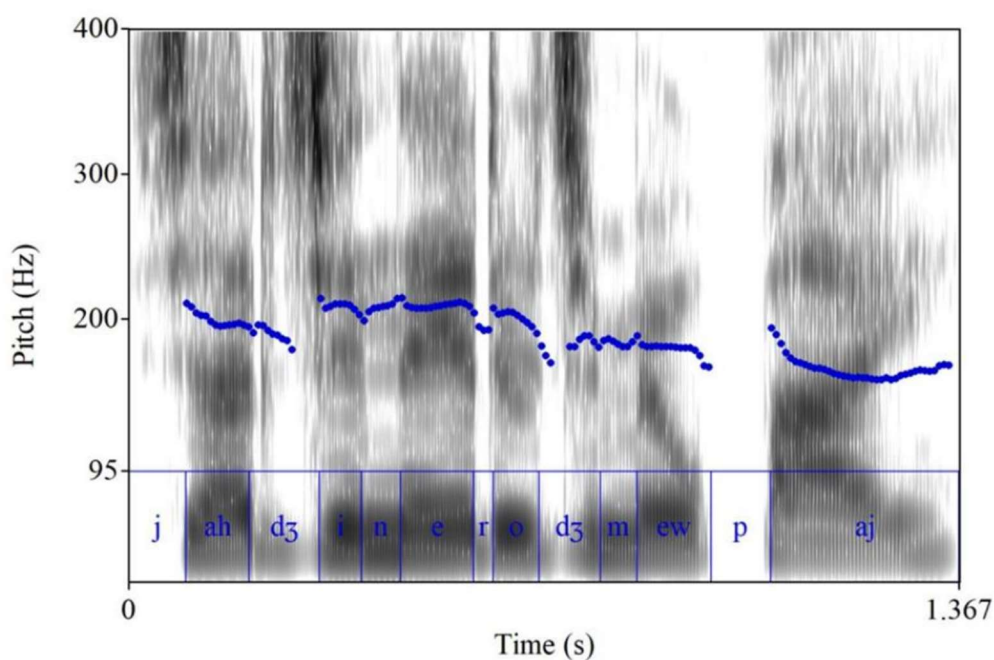
Estava, de novo, diante da complexidade existente no território de áreas comuns entre a produção e a percepção humana. Essas envolvem, ao mesmo tempo, a percepção auditiva (portanto, aí, estamos no mundo dos sons), e a percepção indireta dos movimentos do corpo, mais especificamente dos movimentos articulatórios que fazemos para realizar a fala. Nossa impressão ou percepção mais intuitiva é aquilo que nos possibilita expressarmo-nos sobre o fenômeno percebido. No entanto, essa impressão, expressa em palavras, costuma ser genérica e imprecisa, pois geralmente não temos quase nenhum conhecimento específico dos fenômenos. Assim, experimentos baseados em testes de percepção devem ser feitos com muito cuidado e buscar o máximo de controle das coisas que estão ocorrendo no momento da apresentação do dado ao ouvinte. Por isso, parece-me sempre inevitável, antes de fazer uma questão eminentemente na área do “perceptual”, abordar o fenômeno que (supomos) é o sinal produzido para tal percepção.

De novo, como no início dos estudos sobre canto, escolhi observar o fenômeno que queria circunscrever a partir da análise de fenômenos da produção, descobrindo o que os dados acústicos nos mostravam sobre a articulação da fala cantada (canto). Só que agora, estava voltada para o tom fundamental que emitimos quando cantamos, e que é produzido pela articulação da laringe. Diferentemente dos estudos iniciais, que eram mais descritivos, propus uma hipótese.

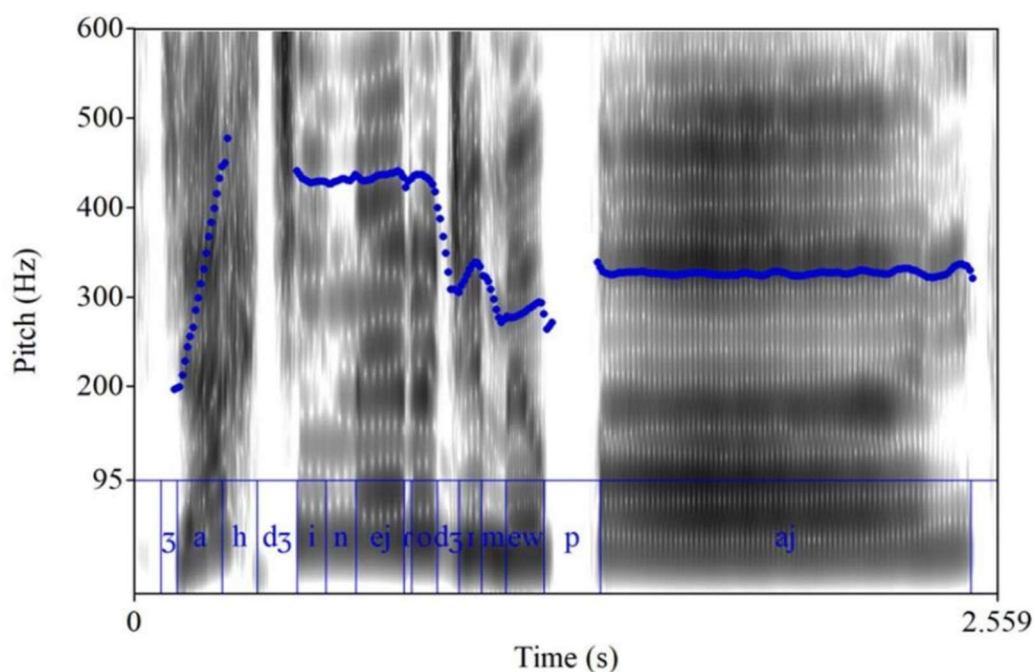
A hipótese é de que o sinal da voz emitido no canto é mais estável do que aquele emitido na fala. É algo não trivial de se provar. Estava partindo de alguns pressupostos, entre eles, o de que o termo “estabilidade de frequência fundamental” ou “estabilidade de  $f_0$ ” eram termos já previamente definidos.

Tratou-se, para mim, de uma espécie de questão de honra. Já que eu estava, há quase duas décadas, comparando aspectos da fala com aspectos do canto, deveria saber apontar qual o traço que os distinguia.

Ao tratar separadamente de fala e canto, usamos termos comuns, como, por exemplo, o termo melodia. É possível falar de melodia da fala e melodia do canto. O termo “melodia” da fala é menos metafórico do que pensamos. De fato entoamos as sílabas enunciadas em nossa fala do dia a dia, quase cada qual com uma frequência fundamental ( $f_0$ ), a depender de fatores linguísticos, como o tipo de sentença (tipos sintáticos), o tamanho da sentença (longa ou curta). A fala mais *default* ou canônica caracteriza-se por produzir uma sentença afirmativa simples – digamos assim com 1 ou 2 segundos de duração, constituída de categorias gramaticais como nome e verbo ou apenas nome – que tem um contorno melódico ligeiramente ascendente e depois descendente (ver Figuras 1 e 2 para comparação).



**Figura 1:** Espectrograma com contorno de *pitch*: exemplo de entoação da fala no enunciado “Jardineiro de meu pai”, produzido por uma atriz para o conjunto de dados do Projeto Macunaíma.



**Figura 2:** Espectrograma com contorno de *pitch* : exemplo de entoação do canto no enunciado “Jardineiro de meu pai”, produzido por uma cantora para o conjunto de dados do Projeto “Macunaíma”.<sup>61</sup>

<sup>61</sup> Projeto “Macunaíma” é o apelido dado ao projeto *Aspectos de produção, percepção e cognição na intersecção fala e canto*, financiado pela FAPESP (2015/06283-0). A canção da qual foi extraída o enunciado cantado é “Jardineiro” do cd de Iara Rennó, Macunaíma Ópera Tupi.

Digamos, então, que o contorno melódico da Figura 1 seria o padrão da fala, a partir do qual é possível fazer diferentes e novos contornos. É infinita a possibilidade de diferentes realizações de entoação da fala, portanto vamos nos ater ao que é mais *default*, que é como temos chamado padrões mais frequentes ou canônicos do ato de falar. Nessa realização canônica de uma frase curta e corriqueira, podemos assim descrever as constrações mais gerais da ação de falar:

1. As pregas vocais, localizadas na laringe, iniciam tal vibração de modo não instantâneo.
2. Ao mesmo tempo, articuladores como lábios e língua estão fazendo ajustes na cavidade oral.
3. Está em jogo, também, obviamente, a respiração, cujo controle parece ser o responsável pela inspiração necessária a ser feita para dar conta de acomodar poucos ou vários segundos de expiração (ou seja, o tempo em que o falante perderá ar).

Estamos diante de uma orquestração de movimentos e de passagem de ar nos tubos da fala: a laringe e a cavidade oral. Parece tudo muito fácil e natural, mas se formos analisar todos esses movimentos, à exaustão, veremos que se trata de sistemas bastante complexos. É a língua sonora em ação, cuja finalidade é emitir, na atmosfera, partículas sonoras com significado. A tais partículas chamamos mais comumente (e também formalmente) de fones. Essas partículas têm traços articulatorios bem característicos e, portanto, podemos classificá-las em vogais e consoantes. Nesse trabalho imenso de coordenar bem tal produção, as vogais e as consoantes já são emitidas com sua temporalidade própria. Ainda que seja possível alongar as vogais e algumas consoantes, elas têm sua duração mais ou menos esperada em uma fala canônica.

Vê-se, então, que um sistema articulatório está interagindo com um sistema de organização temporal, a fim de que se acomodem gestos fônicos e respiratórios de tal modo que aqueles primeiros sejam capturados e inteligíveis naquilo que querem ou precisam informar. Em geral, é possível prever a duração média dos fones, sobretudo se sabemos se são átonos ou tônicos.

A partir de toda essa explicação, espero que tenhamos uma noção já de que a produção de fala parece ao mesmo tempo atingir grandes ganhos sem requerer muito esforço. O ganho está em conseguir, por exemplo, com uma duração tão curta de 250 milissegundos, toda uma sílaba. Ou, ainda, o ganho reside no fato de que um fone tão

curto como um “r” ([r] = *tap*) de ponta língua, que pode ter diminutos 15 milissegundos, possa fazer toda a diferença ao ser percebido. Em outras palavras, a duração quase única e típica do *tap* faz com que o falante do português consiga distinguir “cara” de “cata”, por exemplo. A duração esperada de uma consoante como “t” é de mais de 100 milissegundos.

O *tap* é um bom exemplo de como as durações próprias de cada fone ou segmento de fala são inerentes a eles e os caracterizam sonoramente. As fricativas (sons como “s”, por exemplo) são normalmente longas, pois requerem um determinado tempo para soarem como tais, ou seja, para se chegar à fricção necessária das partículas de ar dentro da boca. Assim, dado que cada fone tem suas características temporais intrínsecas – e, portanto, esperadas – é mais do que plausível pensar que variem de fone para fone e que essa variação seja importante para a fala se constituir como fala: sinal sonoro com traços articulatórios específicos trocado entre falantes para fins de comunicação.

Diante dessas descrições sobre a temporalidade dos fones da fala, que devem encadear-se para a realização de uma frase simples, espera-se, pois, que a organização temporal constitua-se de unidades de durações diversas, portanto sem apresentar uma constância de ou isocronia entre estas durações. E como relacionar isso à questão da estabilidade ou instabilidade da frequência fundamental (doravante  $f_0$ )?

Sempre tendo em mente que se trata da fala canônica, podemos então pensar que, se no canto alongamos necessariamente a vogal, em geral obedecendo a uma métrica pré-determinada, é claro que há uma relação entre a produção da frequência de fonação e articulação do fone: há que se estabelecer uma espécie de continuidade entre dois sistemas, a saber, o laríngeo e o supralaríngeo (cavidade oral). Temos aí a coincidência da nota musical e da vogal: de um lado a música e de outro a língua. Ou ainda: de um lado o canto e de outro a fala. Ora, se a fala possui fones de diversas durações dentro de uma gama de durações curtas (abaixo de 200 milissegundos) é esperado que, normalmente, a sua melodia acompanhe estes fones e varie também, mesmo porque sua sustentação não é prevista para fala<sup>62</sup> e há, como já começamos a ver, razões fisiológicas para isso.

---

<sup>62</sup> Hiroya Fujisaki. “Dynamic characteristics of voice fundamental frequency in speech and singing – Acoustical analysis and physiological interpretations”, Proceedings of the Fourth F.A.S.E. Symposium on Acoustics and Speech, 2: pp. 57–70, 1981.



Vamos reunir alguns pontos importantes, bastante elucidativos da produção de fala, que, a meu ver, são quase nada ou pouco explorados para explicar fenômenos estudados no interior da Linguística:

1. O contorno melódico da fala e sua relação direta com nossas capacidades e limitações fisiológicas.
2. A compreensão mais completa e melhor disseminada (mesmo entre leigos) de que o contorno melódico da fala pertence ao conjunto de fenômenos da entoação da fala e é produzido em um tubo muito especial do nosso corpo: a laringe.
3. Na laringe estão abrigadas as pregas (ou cordas) vocais, que são articuladores não visíveis (internos), mas tão importantes quanto os mais visíveis, como lábios e língua, sem falar da mandíbula.

Toda a ideia sonora que temos das vozes faladas e cantadas parece estar desconectada da existência de uma fonte produtora. Em geral, nos estudos linguísticos, isso fica em segundo plano. Propiciar a reunião dos diferentes sistemas de produção da fala, tradicionalmente divididos e descritos como diferentes para fins didáticos e de estudos, seria uma boa iniciativa para promover uma apreensão mais significativa da forma como produzimos fala.

E poderia ser aplicada ao canto, com alguns ajustes. O primeiro ajuste seria aclarar que a voz cantada, com ou sem texto, realiza um sistema eminentemente sonoro, que tem “regras” próprias de organização de unidades que o constituem. Podemos dizer que a duração assim como a frequência fundamental de unidades como notas musicais são pré-determinadas e estão menos sujeitas a um sistema de produção mais espontâneo em relação a esses aspectos, como é o caso da fala.

Ao cantar melodias musicais e isso parece redundante, pois sempre cantamos melodias, fazemos isso entoando frequências fundamentais previsíveis, pois pertencentes a algum sistema escalar musical. Não é o momento de averiguar se todas as músicas do mundo obedecem a uma escala de tons, mas tudo leva a crer que sim. Vejamos o que nos diz Antonio Carrasqueira, ao tratar a gênese das escalas musicais:

É interessante notar que o intervalo de quinta, que é o segundo intervalo da série harmônica, é a base para a construção das escalas mais utilizadas no mundo todo:

- as escalas pentatônicas (escalas de cinco notas), presentes em culturas de todos os continentes;

- as escalas diatônicas (escalas de sete notas) modelo escalar da tradição europeia ocidental desde os gregos.

Conscientes disso, estamos também cientes de que os sons musicais pertencem a um sistema fechado ou mais ou menos fechado e previsível de alturas sonoras que se relacionam entre si. O que quero tentar mostrar é que, diferente da fala, que não tem uma escala de alturas para ser realizada, o canto é produzido com tais sons. Podemos dizer, então, que a musicalidade ou traço musical dos sons produzidos pelo sistema fonatório humano requer, basicamente, uma articulação específica das pregas vocais, ou em termos mais latos, da laringe.

Minha proposta é de que a laringe é o único órgão de nosso corpo a ser capaz de produzir um som estável e representante de um sistema de escala musical<sup>63</sup>, por sua capacidade de fazê-lo de forma contínua. Aliás, o traço contínuo na fonologia já é um traço de fones e/ou fonemas descritos formalmente nas teorias fonológicas no interior da linguística. A questão do gesto musical ligado à laringe como articulador será aprofundada no capítulo seguinte.

Agora trato de salientar a relação da laringe com o traço da estabilidade. Por sua capacidade articulatória de sustentar um tom, a laringe confere mais ou menos estabilidade ao som vozeado. Na música, espera-se que este som vozeado veicule tons diferentes que se organizam como uma melodia. A duração destes tons também é pré-determinada. Assim falamos aqui de estabilidade. Mas o que é estabilidade?

### **Estabilidade em Termos Latos e em Termos Microscópicos também**

O termo estabilidade parece estar presente em várias áreas da ciência, mas seguramente não as atravessa de maneira igual. Ou seja, para cada área, sua descrição mais ou menos objetiva deve diferir. O seu significado mais corriqueiro deve ser o de que nada muda ou pouco muda. Como esse sentido se relaciona com os demais? Ou seja, com aqueles mais científicos e que por isso devem ser mais objetivos?

Como ponto de partida para essa apresentação do termo estabilidade, lanço mão do verbete do dicionário, que transcrevo parcialmente:

---

<sup>63</sup> Há também a produção sonora do assobio, mas por falta de conhecimento, não vou levantar a questão acerca desse fenômeno. Sob a forma de projeto, há o estudo do assobio e suas características acústicas, assim como a abordagem de línguas assobiadas pelo aluno Gabriel Brasileiro.

**estabilidade.** [Do lat. *Stabilitate*] S.f. **1.** Qualidade de estável; firmeza, solidez, segurança. **2.** Fís. Propriedade geral dos sistemas mecânicos, elétricos e aerodinâmicos, pela qual o sistema retorna ao estado de equilíbrio após sofrer uma perturbação. **3.** Jur. Garantia que tem o funcionário público efetivo, depois de certo tempo de exercício, de não ser demitido senão por sentença judicial ou mediante processo administrativo.<sup>64</sup>

O aspecto estável de coisas ou eventos é abordado em alguns domínios da ciência a fim de definir um fenômeno ou sistema que mantém energia ou equilíbrio. Por exemplo, na química, a estabilidade termodinâmica se refere à menor energia em um sistema. Na física, o termo estabilidade está relacionado à capacidade do sistema de recuperar seu equilíbrio, bem como à localização do seu centro de gravidade. Interessante notar que a física está contemplada no verbete do dicionário acima citado, juntamente com a área jurídica. Então é claro que em diferentes dicionários, de diferentes épocas, será selecionada uma ou outra área do conhecimento para se definir o termo estabilidade.

A estabilidade que queremos abordar tem a ver com a capacidade de manter um movimento articulatório corporal sem grandes perturbações. Parece que essa visão tem mais a ver com a repetição de um movimento do que equilíbrio ou manutenção de energia. Mas foi buscando um título sugestivo em uma obra de divulgação científica<sup>38</sup>, que contivesse as palavras “estabilidade” e/ou “equilíbrio” ou a elas ligadas, como “energia”, tal como a vimos no contexto da termodinâmica, que encontrei o texto “Freud e os caminhos da energia”. Desse texto extrairéi a parte da explicação de Evando Mirra para o equilíbrio. Não resenharei a ideia da relação entre termodinâmica e energia psíquica postulada pelo genial Sigmund Freud, pois não sei se a capturei como deveria. O instigante é trazer a esta nossa discussão a descrição de Mirra. Vejamos como ele descreve o termo equilíbrio do ponto de vista da física:

Equilíbrio” vem da junção de *aequus* com libra: a igualdade de pesos entre os dois pratos da balança. O conceito foi primeiro utilizado pela Mecânica. E reinventado pela Termodinâmica para falar da situação em que não há diferenças, em que tudo é plano, igual. E imóvel: o equilíbrio é o último estado, depois do qual nada pode acontecer. O equilíbrio é a morte.

Para que haja mudanças é necessário haver diferenças: quando qualquer coisa se passa, o ponto de partida é o desequilíbrio.<sup>65</sup>

---

<sup>64</sup> Novo dicionário Aurélio da Língua Portuguesa. 2a. Edição. Editora Nova Fronteira, 1986.

<sup>65</sup> Evando Mirra. A Ciência que sonha e o verso que investiga. São Paulo. Editora Papagaio, 2009, p. 288.

Tenho me interessado não apenas por livros em certo formato e natureza de divulgação dos saberes científicos. Além disso, tenho tido a ajuda luxuosa dos colegas. Assim é que me escreveu gentilmente a física Adriana Tufaile, professora da Escola de Artes e Ciências Humanas (EACH) da USP, quando lhe perguntei o que era estabilidade:

### **Um livro deitado é um sistema estável**

Um jeito de mostrar estabilidade em física é estudar o comportamento do sistema quando ele sofre perturbações. Se o sistema volta ao mesmo ponto, ele é estável. Se ele vai para outro ponto, por causa da perturbação, ele é instável. O exemplo mais simples é de um objeto em pé no campo gravitacional, por exemplo, um livro. Imagine um livro fino (de capa dura) que você conseguiu colocar de pé numa mesa. Qualquer pequeno toque (perturbação) fará o livro tombar. Mas se você fizer isso com um livro muito grosso, como um dicionário de português, você não conseguirá tombá-lo com um toquinho, terá que empurrá-lo com mais força. Dizemos que o livrinho em pé é mais instável que o dicionário. Agora, se esses dois livros estiverem deitados, eles estão numa configuração muito estável, pois se você empurrá-los, eles não tombarão.

Ora, essa excelente explicação de Adriana Tufaile, trazendo-nos objetos e situações do dia a dia para entender a estabilidade, tem conexões, sim, com a estabilidade e a instabilidade da fala. A analogia a ser feita em primeiro lugar, se pensarmos nas vogais como sempre mais estáveis que as consoantes – tratando-se de fala ou canto, tanto faz – as consoantes são sons com grande potencial para inserir perturbações no sistema. Vejamos as obstruintes, por exemplo, consoantes como [p] e [f]. Elas podem ser chamadas de labiais, pois produzem um ruído utilizando os lábios como articuladores: a primeira produz uma explosão e a segunda um chiado contínuo. Para que isso ocorra, toda a configuração existente para as vogais tem de cessar, a passagem livre de ar pela cavidade oral se transforma em um canal ou de represamento de ar silencioso ou de ruído, ainda que as pregas vocais vibrem, o que fazem em baixíssima energia, no caso das consoantes vozeadas. É como se houvesse uma pequenina batalha entre sons melódiosos *versus* sons ruidosos. Felizmente uma batalha *win-win*. Todas as partes ganham. Na verdade os contrastes são desejáveis e já fazem parte do que os linguistas chamam de gramática da

língua. Em outras palavras, os fones têm de ser contrastantes para que, no ato comunicacional, tenhamos eficiência para veicular significados diferentes.

Se para a fala falada as alternâncias entre sons mais ruidosos e mais melódicos é bem-vinda e necessária, para o canto, esse fenômeno já é mais discutível. Não é rara, entre músicos, a queixa de que cantores e cantoras desafinam demais. Há até a seguinte anedota sobre a desafinação de alguns: — Não é o caso que cantores e trompistas **sempre** desafinam, mas que **às vezes** eles conseguem afinar. Não sei quanto aos trompistas, pois pouco sei de trompa (a não ser que tem um timbre divino), mas isso é muito injusto quanto aos cantores. Não importa qual seja o estilo de canto. Vou propor que abordemos a questão da afinação do cantor como o mito da desafinação.

Quando se fala que o cantor desafina, simplesmente, parece estar se falando de incompetência musical. Muitas das vezes o cantor afina ao longo de toda a peça cantada e desliza em alguns poucos momentos. Isso já é suficiente para dizer que desafina. Vou advogar que tais deslizos não representam incompetência musical do sujeito cantante, mas fazem parte mesmo do ato de cantar, dada sua complexidade. As consoantes podem ser verdadeiras vilãs nesta hora e é sua produção que leva ou pode levar à desafinação.

Pensemos assim: aquela perturbação da passagem de ar que cria um ruído característico de uma consoante, ou seja, que obstrui a passagem de ar – por isso se chama obstruente – pode ocasionar também a perturbação da vibração das pregas vocais. Imaginemos o seguinte: que o movimento articulatório realizado em um ponto x do nosso aparato fonador perturba ou outro ponto x'. Em uma analogia um tanto improvisada, tomemos o exemplo de um canudo maleável<sup>66</sup> em seu estado retilíneo. Ao assoprarmos em uma extremidade desse canudinho, o ar passará por aí livremente. Se dobrarmos o canudo em sua porção medial, e assoprarmos de novo, já o ar não passará da mesma maneira. Qualquer ponto demarcado em cada uma das metades do canudo terá sua passagem de ar diferente por causa do ponto medial.

Podemos dizer que é isso que acontece com o “tubo cantante” mais uniforme que se forma para cantarmos. O alvo articulatório, aí, seria manter uma configuração de estabilidades: enquanto a cavidade oral se encarrega de manter-se aberta, com articuladores sem se movimentarem, a laringe estaria também posicionada como uma continuação desse tubo, sem uma constrição na sua parte superior. Essa constrição pode

---

<sup>66</sup> O canudinho de plástico, que em breve deverá ser extinto de nosso dia a dia, por ser ambientalmente nocivo.

ocorrer dependendo do movimento da língua ou da pressão e volume de ar nos pontos na região da faringe.

Isso tudo para explicar que a produção de uma consoante barulhenta pode desestabilizar essa configuração bem articulada e provocar o efeito indesejável, musicalmente, que é a desafinação.

Ora, a desafinação é um fator de desestabilização para um sistema musical que busca relações fixas entre diferentes tons. Em música ocidental, são os graus da escala que se combinam mais ou menos entre si. E para a fala? Podemos atribuir estabilidade? Sim, mas em outra dimensão: a partir dos movimentos dos articuladores visíveis da fala. Por ser em pequena escala, essa dimensão requer uma explicação mais técnica. Um bom exemplo com o qual podemos começar a explicação é o da sílaba, cujo representante mais comum é aquele formado por consoante e vogal (CV). A maneira como se articulam em belíssima orquestração, consoante e vogal, é um exemplo de estabilidade. Acho que já falei disso em algum capítulo desta obra, mas não custa voltar aqui, adicionando o mais importante: a questão da fase.

Vários fenômenos de movimento envolvem estar em fase ou fora de fase. Ao bater palmas junto com outras pessoas – em uma plateia, depois de um *show* – podemos fazer isso em fase ou fora de fase. Em algumas situações dessas palmas, pode ocorrer de todos emitirem o som ao mesmo tempo, como se fosse um relógio ou marcação temporal. Quando isso acontece, dizemos que as palmas estão em fase. O mesmo pode acontecer com os gestos da fala, ou seja, com os movimentos para se fazer a consoante e a vogal. Diferente do que podemos pensar, os fones não são articulados um de cada vez, mas orquestrados de tal modo que podem estar em fase. A sílaba mais fácil para se colocar em fase é a sílaba CV. Não vou entrar em detalhes da teoria que explica isso, mas apelar para a intuição. Parece mais fácil a gente “arrumar” a cavidade oral para produzir uma consoante seguida de vogal, como que embutindo a articulação da segunda na primeira, do que articular a sequência vogal-consoante (VC). Podemos, então, atribuir à fala essa característica estável, concluindo que as sílabas CV têm mais estabilidade articulatória, digamos assim, do que outros tipos de sílabas (pensemos em “transparente”, em que só na primeira sílaba há apenas uma vogal emboscada em duas sequências de consoantes!).

Deve ficar claro, também, que a fala pode requerer sons contínuos e com estabilidade de  $f_0$  (manter ou sustentar sempre o mesmo tom), como já mencionamos quando falamos do grito de gol. A fala pode lançar mão de inúmeros recursos acústico-

articulatórios e rítmicos para marcar o discurso ou para destacar seu conteúdo, ao que, em geral, chamamos de ênfase entoacional. Ao tratar disso, adentramos o reino da prosódia da fala, encantador (espero que não encantado!) e que merece ser percorrido em outro momento da vida acadêmica.

Seria interessante continuar a buscar exemplos de estabilidade pela ciência afora, ou por outros reinos, mas isso desvirtuaria o propósito desta parte do livro.

### **Mas o que Seria a Instabilidade de Fonação na Fala?**

A instabilidade da fala é abordada comparativamente à música, ou mais especificamente ao canto. Esta abordagem pode parecer bem pouco objetiva, mas o fato que vou descrever ainda não foi objetivado, sendo que isso obriga-me a desbravar o terreno. Em vez de partirmos da fala, partamos do canto da música ocidental. Para cantarmos, entoamos notas musicais de uma escala de sete tons ou alturas, falando em termos bem gerais. Essas notas têm uma relação entre si que aqui podemos descrever como distâncias tonais esperadas, já que constituem peças integrantes de uma escala. Em geral sabemos isso, pois memorizamos uma melodia de forma fácil e justamente porque as notas guardam uma relação tonal entre si que o nosso sistema auditivo e nosso cérebro tratam de identificar.

O mesmo não é verdade para a fala, para a qual não fixamos uma determinada melodia correlata a uma determinada sentença ou qualquer outra unidade linguística. Em outras palavras: não predefinimos a frequência fundamental ( $f_0$ ) de uma unidade linguística falada. Isso não significa que não exista uma  $f_0$  na fala, e é óbvio que ela existe e que sem ela toda a nossa voz soaria desvozeada, o que parece um contrassenso.

A  $f_0$  da fala é responsável pelos fones vozeados e por sua melodia. Essa última, no entanto, não obedece a uma tonalidade (por exemplo, dó maior) e daí dizer que não sabemos nunca em que frequência de fonação estamos falando. Tampouco, ao falar, obedecemos à escala de uma tonalidade. Se pensarmos nos tons da escala musical como segmentos, podemos contrapô-los ao funcionamento da entoação de  $f_0$  na fala que não é segmental, pelo contrário, é uma linha contínua, aliás, trata-se de um fenômeno suprasegmental.

Essa ideia é fundamental para entendermos os funcionamentos diferentes dos eventos melódicos de fala e canto. Mas vejamos: se é verdade que não prevemos um valor de frequência fundamental para falar, também é verdade que entoamos a fala em uma

determinada faixa de frequência e sabemos produzir variações de  $f_0$  que nos ajudam a organizar o discurso e têm significado. Essa faixa de frequência está diretamente relacionada com traços fisiológicos da laringe do indivíduo, pode ser chamada também de tessitura da fala e tem a extensão de uma oitava, tanto para homens como para mulheres. É, pois, dentro dos máximos e mínimos dessa tessitura que falamos.

Mas, então, qual nome dar à falta de uma previsão de  $f_0$  na fala? Porque é disso que estamos falando: da ausência de um fato. Sugiro o termo “entoação livre” ou “não tonalidade”, sempre com ressalvas, pois um termo nunca abarca ou descreve tudo que queremos, com a objetividade que sonhamos. O termo “entoação livre” me parece mais adequado, já que o segundo pode trazer confusão sobre a ideia de tom. Há línguas tonais, para as quais as vogais em tons mais agudos ou graves fazem toda a diferença de significado, que é o caso do mandarim.

Farei mais um ajuste para dizer que a fala, em oposição ao canto, tem uma entoação livre relativa, já que essa liberdade dependerá, como já disse, de limites do organismo.

O conceito de entoação livre relativa implica, basicamente, os seguintes fenômenos, enumerando:

1. A produção contínua de um tom laríngeo ao longo da cadeia da fala.
2. A produção desse tom laríngeo nos limites de uma faixa de frequência da voz (tessitura da fala).
3. A variação contínua desse tom ( $f_0$ ), sem que se estabeleça, necessariamente, uma relação tonal entre essas variações.

Ou seja, a gente não “afina” para falar.

O que foi descrito acima de maneira qualitativa pode ser entendido como a instabilidade de  $f_0$  na fala, contraposta à maior estabilidade de  $f_0$  no canto. A estabilidade do canto pode ser melhor compreendida ao considerarmos que entoar melodias musicais é atingir alvos de entoação. Este alvo tem que se manter estável ao longo do tempo de uma nota. Tanto o alvo (nota) e o tempo de sua duração são predeterminados, seguem uma “partitura”, seja ela escrita, decorada ou apenas virtual.

O conceito de entoação cantada poderia ser o seguinte:

1. A produção contínua de um tom laríngeo ao longo do canto (ou fala cantada).
2. A produção desse tom laríngeo em limites estendidos da faixa de frequência da voz.



### 3. A sustentação do f0 com relações tonais, qualquer que seja a escala.

A estabilidade é um efeito desejado na produção do canto. Tanto é assim, que um estudo bastante conhecido de Diana Deutsch sobre o efeito ilusório que transforma fala em canto toma o fator estabilidade de f0 no canto como traço que se opõe à fala instável. Esse fenômeno, batizado por Deutsch como “illusory transformation from speech to song”, é o seguinte: indivíduos ouvintes expostos a trechos repetidos da fala, passam a ouvir aí uma melodia musical. É isso mesmo: mesmo os estímulos sendo de fala, acabamos ouvindo uma canção! Diana Deutsch e colegas explicam que a repetição do mesmo trecho leva o cérebro a atribuir-lhe uma estabilidade tonal<sup>67</sup>.

Essa descoberta superinteressante explica, juntamente com outros trabalhos que se sucederam ao de Deutsch, porque muitas vezes percebemos uma melodia cantada na fala. É como se o indivíduo falante estivesse diante de uma bifurcação e, de um momento a outro, sai do caminho da fala para ir pelo caminho do canto.

A questão da instabilidade de f0 na fala claramente ganhou um *status* nesta discussão em que explicá-la e nomeá-la revelou-se necessário. Mas nos estudos da fala (sejam eles abrangidos na fonética, na fonologia ou nas ciências da fala) trata-se apenas de estudos de entoação da fala, os quais se desenvolvem na subárea prosódia.

Em perspectiva, partindo-se da ideia de estabilidade do canto, foi preciso descrever a instabilidade na fala. A problemática que nos acompanhará para sempre, ao que me parece até o momento das reflexões e estudos, é que a instabilidade de que tratamos aqui não é categórica e pode bifurcar, desviando para uma estabilidade relativa a qualquer momento.

Daí podermos ouvir canto na fala e fala no canto.

#### **Estudos sobre a Estabilidade de f0 no Canto e na Fala**

A questão dos bastidores da ciência, então, é: como comprovar que a fala apresenta instabilidade de f0 comparada à estabilidade de f0 no canto?

Qual o plano para um roteiro que nos levasse à resposta? Trata-se do *Aspectos de produção, percepção e cognição na intersecção fala e canto*, projeto elaborado em 2015 por mim e que contou com a colaboração de João Paulo Cabral, Aleksandro Meireles e André Baceti. Isso me faz pensar o que seria de toda a arte e ofício não fossem os

---

<sup>67</sup> ver Diana Deutsch, Trevor Henthorn & Rachael Lapidis. Illusory transformation from speech to song. *Journal of the Acoustic Society of America*, 129, no. 4, 2011, pp. 2245-2252.

bastidores, *ateliers* e oficinas. Para os acadêmicos e cientistas, o projeto escrito é o verdadeiro repositório de ideias, e não só. É, também, o lugar em que estas se assentam, ainda que tenham que ser modificadas depois.

Do projeto escrito saltamos para a realização dos passos experimentais previstos no método: atrair os indivíduos falantes e cantantes para o laboratório (*atelier*) e gravar. Daí é fazer a escuta dos dados, organizá-los e fazer as análises.

Verificamos que, apesar dos melismas e das improvisações das cantoras, existe maior estabilidade no canto do que na fala. É claro que para além do laboratório, há todo um mundo de cantares, interminável, no qual outras instabilidades estáveis nos aguardam.

O recorte que fizemos para verificar a estabilidade maior no canto do que na fala envolveu quatro canções do álbum *Macunaíma Ópera Tupi* de Iara Rennó. O tipo de canto que as cantoras produziram é tal que a letra é facilmente entendida pelo ouvinte, o que faz esse canto parecido com a fala. Por exemplo, em geral, a tessitura não excedia uma oitava, o que é similar ao que ocorre com a tessitura da fala (196 to 392 Hz). Outra coisa importante de se ajustar nos bastidores experimentais é a uniformidade dos dados. As canções eram similares entre si, também em termos de andamento. As cantoras seguiram os andamentos da gravação de referência que iam de 96 a 120 batidas por minuto (bpm).

Poderia descrever outros detalhes do estudo, mas seria enfadonho. Em 2017, baseados nos primeiros resultados, conseguimos respostas iniciais e apresentamos, no congresso *Speech Prosody*, o trabalho *Acoustic distinctions between speech and singing: is singing acoustically more stable than speech?* A confirmação total da hipótese lançada foi conseguida depois e publicada, em 2021, no artigo *A comparative study of fundamental frequency stability*<sup>68</sup>.

E já que falei de bastidores, aproveito para frisar que os pensamentos e ideias gerados ao longo da feitura (*making off*) do estudo levantaram muitas indagações, tais como: e se gravássemos vozes masculinas, e se fizéssemos mais comparações com outros gêneros musicais e estilos de canto? E se a estabilidade de que falamos só pode ser comprovada com um *design* feito especialmente para capturá-la?

---

<sup>68</sup> Beatriz Raposo de Medeiros, João Paulo Cabral, Alexsandro R. Meireles & Andre A. Baceti. A comparative study of fundamental frequency stability between speech and singing. *Speech Communication*, Vol. 128, 2021, pp. 15-23

Vê-se que o trabalho de pesquisa não tem fim: responde algumas perguntas e gera outras. Então repito: temos mais assuntos a desenvolver do que tempo, mãos e braços.

**PARTE 5:**  
**Primitivos da Ciência:**  
**O Primitivo Linguístico & O Primitivo Musical**

Mais de uma vez, nesta, expressei o aspecto infinito que têm meus objetos de estudo: algumas falas e canções. Digo “algumas” porque o caráter experimental das investigações não abarca a totalidade do fenômeno. Passear pelas rádios de uma Radio Garden ou usar plataformas de *streaming* de música dá-nos uma ideia da quantidade e diversidade de música cantada que o ser humano produz.<sup>69</sup>

Ainda que eu conseguisse, neste momento em que escrevo, aqui e agora, acessar todas as canções disponíveis em gravações pela *internet* nesta “corredeira”, ficariam de fora os cantos espontâneos das crianças pequenas e todas as novas expressões cantadas que são criadas a cada instante.

De outro lado, em apenas alguns poucos minutos de gravação de um só banco de dados, posso extrair grande quantidade de informação da fala e do canto em vários aspectos: durações temporais dos segmentos falados e cantos; contorno melódico das sentenças; melodia musical; qualidade vocálica (extraindo formantes); inferências articulatórias das consoantes a partir de seu comportamento acústico; coarticulação consoante-vogal, vogal-consoante ou qualquer outra estrutura silábica ou para além da sílaba; alinhamento de acento fala-canto (ou não alinhamento) e qualidade de voz. Esses aspectos da fala que podem ser levantados também em estudos fonético-acústicos sobre o canto são os mais comuns para se chegar a um número vasto de explicações da produção da fala.

Por isso estou entre dois mundos: um de dimensão inalcançável e outro, que embora limitado, oferece a multiplicação de dados sonoros, a partir de contornos melódicos de sentenças, sílabas, vogais e porções de vogais. Só para se ter uma ideia, apenas o conjunto de dados de uma cantora, gerou 73.654 valores de  $f_0$  (frequência fundamental da voz). Esse conjunto de dados foi formado de quatro canções com uma média de 6 sentenças por canção. As figuras abaixo mostram o início e o fim da planilha que foi obtida como resultado da medida de vários parâmetros acústicos através de um programa de análise acústica da voz, o VoiceSauce<sup>70</sup>. A obtenção de tantos valores a cada vogal deve-se ao fato de que esse programa extrai um valor a cada milésimo de segundo

---

<sup>69</sup> Para um conjunto de dados específico e elaborado para estudo sobre universalidade da canção, vale a pena ver aquele utilizado nas pesquisas de Samuel A. Mehr e seus colegas. São aproximadamente 500 canções descritas do ponto de vista etnográfico e de seus elementos musicais.

<sup>70</sup> O VoiceSauce é um programa de análise de voz desenvolvido e continuamente aperfeiçoado por equipe da UCLA, University of California (<http://www.phonetics.ucla.edu/voicesauce/>). Além de extrair frequência fundamental e formantes, extrai uma série de outros parâmetros acústicos como amplitude de harmônicos ou mesmo a diferença entre amplitude de harmônicos para fins de análise de qualidade de voz.

da duração total de apenas uma vogal. Gera-se assim uma quantidade enorme de dados para cada parâmetro acústico (Figuras 1 e 2).

a)

The screenshot shows a large Excel spreadsheet with 51 rows of data. The columns are labeled with various acoustic parameters. The first few columns are: 'Filename', 'Label', 'seg', 'Start', 'seg\_End', 'sF0', 'sF1', 'sF2', 'sF3', 'sF4', 'sF5', 'sF6', 'sF7', 'sF8', 'sF9', 'sF10', 'sF11', 'sF12', 'sF13', 'sF14', 'sF15', 'sF16', 'sF17', 'sF18', 'sF19', 'sF20', 'sF21', 'sF22', 'sF23', 'sF24', 'sF25', 'sF26', 'sF27', 'sF28', 'sF29', 'sF30', 'sF31', 'sF32', 'sF33', 'sF34', 'sF35', 'sF36', 'sF37', 'sF38', 'sF39', 'sF40', 'sF41', 'sF42', 'sF43', 'sF44', 'sF45', 'sF46', 'sF47', 'sF48', 'sF49', 'sF50', 'sF51', 'sF52', 'sF53', 'sF54', 'sF55', 'sF56', 'sF57', 'sF58', 'sF59', 'sF60', 'sF61', 'sF62', 'sF63', 'sF64', 'sF65', 'sF66', 'sF67', 'sF68', 'sF69', 'sF70', 'sF71', 'sF72', 'sF73', 'sF74', 'sF75', 'sF76', 'sF77', 'sF78', 'sF79', 'sF80', 'sF81', 'sF82', 'sF83', 'sF84', 'sF85', 'sF86', 'sF87', 'sF88', 'sF89', 'sF90', 'sF91', 'sF92', 'sF93', 'sF94', 'sF95', 'sF96', 'sF97', 'sF98', 'sF99', 'sF100'. The data is organized in a grid format, with each row representing a different audio segment.

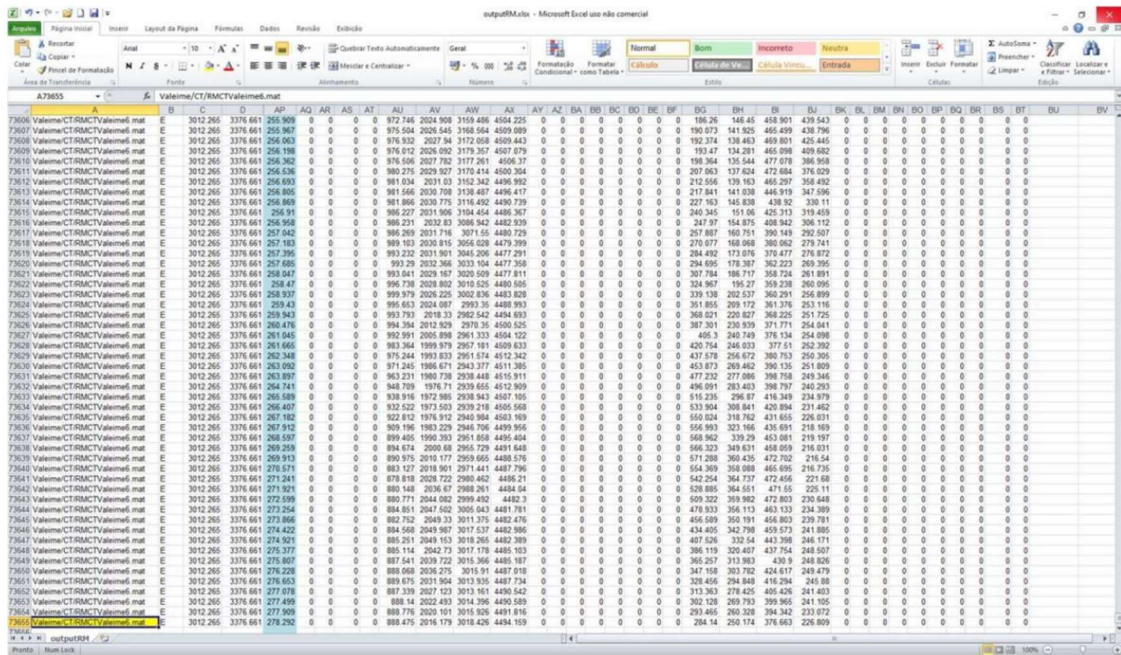
b)

The screenshot shows a detailed view of the first 14 rows of the Excel spreadsheet. The columns are labeled: 'Filename', 'Label', 'seg', 'Start', 'seg\_End', 'sF0', 'sF1', 'sF2', 'sF3', 'sF4', 'sF5', 'sF6', 'sF7', 'sF8', 'sF9', 'sF10', 'sF11', 'sF12', 'sF13', 'sF14', 'sF15', 'sF16', 'sF17', 'sF18', 'sF19', 'sF20', 'sF21', 'sF22', 'sF23', 'sF24', 'sF25', 'sF26', 'sF27', 'sF28', 'sF29', 'sF30', 'sF31', 'sF32', 'sF33', 'sF34', 'sF35', 'sF36', 'sF37', 'sF38', 'sF39', 'sF40', 'sF41', 'sF42', 'sF43', 'sF44', 'sF45', 'sF46', 'sF47', 'sF48', 'sF49', 'sF50', 'sF51', 'sF52', 'sF53', 'sF54', 'sF55', 'sF56', 'sF57', 'sF58', 'sF59', 'sF60', 'sF61', 'sF62', 'sF63', 'sF64', 'sF65', 'sF66', 'sF67', 'sF68', 'sF69', 'sF70', 'sF71', 'sF72', 'sF73', 'sF74', 'sF75', 'sF76', 'sF77', 'sF78', 'sF79', 'sF80', 'sF81', 'sF82', 'sF83', 'sF84', 'sF85', 'sF86', 'sF87', 'sF88', 'sF89', 'sF90', 'sF91', 'sF92', 'sF93', 'sF94', 'sF95', 'sF96', 'sF97', 'sF98', 'sF99', 'sF100'. The first row is highlighted in yellow, and the 'AP' column is highlighted in blue. The data is organized in a grid format, with each row representing a different audio segment.

**Figura 1:** (a) Imagem parcial da planilha gerada pelo programa VoiceSauce: linhas iniciais de 1 a 51. A coluna A apresenta o título da canção, produzida por uma determinada cantora em uma determinada sentença. A coluna B apresenta a vogal isolada. Da coluna C em diante, temos parâmetros acústicos da voz com destaque para a coluna em azul (AP) que foi o parâmetro alvo do estudo sobre a estabilidade da frequência fundamental na fala e no canto. Em (b) temos o detalhe de (A): a linha 1, marcada em amarelo, traz as seguintes informações: Conversa/CT/RMCT/conversa1.mat a 126.75. “Conversa” é a canção; “CT” é a sigla para identificar a condição do canto, “RMCT” são as siglas que nomeiam o sujeito e a condição; “conversa 1” identifica a sentença; “mat” é a extensão do arquivo; “a” é a vogal da qual se extrai a medida; 126.75 é o ponto inicial dessa vogal, em milissegundos.



a)



b)

73641	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	271.24	0	0	0	0	878.818	2028.72	2980.46	4486.21	0
73642	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	271.92	0	0	0	0	880.148	2036.67	2988.26	4484.04	0
73643	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	272.6	0	0	0	0	880.771	2044.08	2999.49	4482.3	0
73644	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	273.25	0	0	0	0	884.851	2047.5	3005.04	4481.78	0
73645	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	273.87	0	0	0	0	882.752	2049.33	3011.38	4482.48	0
73646	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	274.42	0	0	0	0	884.568	2049.99	3017.54	4482.99	0
73647	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	274.92	0	0	0	0	885.251	2049.15	3018.27	4482.39	0
73648	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	275.38	0	0	0	0	885.114	2042.73	3017.18	4485.1	0
73649	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	275.81	0	0	0	0	887.541	2039.72	3015.37	4485.19	0
73650	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	276.23	0	0	0	0	888.068	2036.28	3015.91	4487.02	0
73651	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	276.65	0	0	0	0	889.675	2031.9	3013.94	4487.73	0
73652	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	277.08	0	0	0	0	887.339	2027.12	3013.16	4490.54	0
73653	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	277.5	0	0	0	0	888.14	2022.49	3014.4	4490.59	0
73654	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	277.91	0	0	0	0	888.776	2020.1	3015.93	4491.82	0
73655	Valeime/CT/RMCTValeime6.mat	E	3012.27	3376.661	278.29	0	0	0	0	888.475	2016.18	3018.43	4494.16	0
73656														

Figura 2: (a) Imagem parcial da planilha gerada pelo VoiceSauce: últimas linhas das 73.654 linhas. Em (b), detalhe de (A), marcada em amarelo, está a última linha da Coluna A, cuja descrição se encontra na legenda da Figura 1.

Ora essa posição em que me encontro – entre dois mundos – é muito desconfortável para a ciência que precisa comprovar tudo, ou quase tudo. O aspecto diverso, criativo e também dinâmico das coisas do mundo impede que nossas categorizações ou definições sobre elas sejam peremptórias. No entanto, postular hipóteses teóricas vem a ser uma ferramenta de pensamento interessante e que parece ser um lugar mais tranquilo, ainda que, precisem ser comprovadas. Quero estabelecer uma pausa, abrindo um espaço para a *slow science*, com todo cuidado, para não parecer que assumo uma postura anti-ciência.

O sentimento anticientífico que perigosamente parece se espalhar nos dias que correm, ficou de certo modo profetizado na canção de Tom Zé “Ogodô, ano 2000”<sup>71</sup>. Vejamos a letra: “A ciência excitada/ fará o sinal da cruz/ e acenderemos fogueiras/ para apreciar a lâmpada elétrica.” O simples fato de pensarmos que há disseminação do conceito de terraplanismo gera uma espécie de mal-estar. Essa espécie de “contra-ideia” e outras tais refletem a existência de um sistema de comunicação caótico, sempre apressado e que desafortunadamente é terreno para notícias falsas, as *fakenews*.

Encontramo-nos, assim, em um lugar arriscadíssimo ao criticar ou elogiar a ciência. Ela pode estar sendo vítima de sua própria pressa e discurso altamente especializado, cuja linguagem deixa de ser acessível ao grande público. Suas verdades comprovadas por testes e experimentos sempre descritos como sofisticados e complexos investem-se de autoridade pela simples aparência e não exatamente pelo bem-estar que possam vir a promover, que deveria ser seu objetivo. Nesse território em que se interceptam discursos científicos e pseudocientíficos, pode-se criar um pântano em que ficam indistintos as invenções e os conhecimentos desenvolvidos pelas boas práticas científicas. E temos de partir em defesa do território livre da ciência, agindo em diversas frentes, entre as quais a reflexão sobre a linguagem científica e sua divulgação. Assim, a ideia de tentar dar um lugar mais claro aos limites da ciência insere-se nessa frente reflexiva a fim de propiciar uma autocrítica ao meu próprio trabalho científico. As questões, as hipóteses e cada passo de procedimento metodológico parecem estar sempre a poucos metros do erro.

Mas digamos que teorizar seja uma espécie de antessala das comprovações, sempre confortável. Ou uma sala VIP de um aeroporto, antes do embarque na nave ruidosa e gelada por horas a fio. Então, foi um tanto isso que me propus: entre a montanha de dados e o infinito, surge um recôncavo no qual podemos pensar no seguinte:

- Mas qual é a unidade musical?
- Em que a linguística pode nos ajudar?

Essas ideias só surgiram porque, ao longo dos meus estudos e pesquisas, estive sempre, de forma mais ou menos latente, a ideia de explicar as relações entre canto e fala, em vários níveis. Tanto nos níveis mais concretos, com dados sonoros, como foi o caso do primeiro estudo comparativo envolvendo os aspectos fonético-acústicos de fala e

---

<sup>71</sup> Do CD “O pirulito da ciência”. Tom Zé & Banda, com o selo da Biscoito Fino, 2009.



canto, como a explicação em termos mais abstratos, os quais, na área da cognição, entendemos como simbólicos.

Como evoluiu o pensamento desde o doutorado até recentemente? A narrativa da vida das ideias fica então assim:

No início, ouvi, com atenção, a crítica informal, advinda de cantores e ouvintes, de que a má dicção dos cantores atrapalhava o entendimento do texto cantado. Percebi que essa má dicção estava sendo atribuída à inabilidade do cantor ou cantora. Isso foi a mola propulsora para que eu tentasse entender melhor o fenômeno que se desenhava: a questão da inteligibilidade do texto cantado. Daí, então, deu-se a elaboração de um projeto de pesquisa que analisasse acusticamente os segmentos da fala quando cantados.

Essa análise conseguiu ver indícios de uma negociação entre dois sistemas ou duas estruturas de organização dos sons: de um lado, a fala e de outro, o canto. O tipo de canto investigado, tal como apresentado no primeiro capítulo, tem suas exigências específicas: sendo da linhagem erudita, como o belcanto, requer basicamente volume, afinação e a produção acompanhada de instrumentos não elétricos, ou mesmo à capela. Isso quer dizer que, tradicionalmente, esse tipo de canto não é produzido com amplificação através de microfone e alto-falante. As duas exigências, de afinação e volume são conseguidas, como vimos na parte 1, por uma configuração articulatória do trato vocal.

A partir daí, muitas outras questões pareciam estar submersas:

1. Como se dá exatamente este “ajuste” do tubo da fala com o tubo da laringe para que saia tudo afinado e bem audível?
2. Seria possível postular um nível cognitivo mais abstrato (nível alto) que planejasse o “ajuste”?
3. Ou o fenômeno de ajuste do formato do tubo cantante se explica por um nível mais baixo? Mais corporal?
4. Além de corporal seria no nível acústico e não necessariamente simbólico/abstrato?
5. Algumas vogais seriam mais fáceis de produzir do que outras no canto?
6. Como seria a relação da produção de vogais com sua percepção?
7. E o caso das consoantes: uma vez relativamente mais curtas, no canto, como seria sua percepção?

As questões que implicam uma relação entre ajustes articulatórios (portanto de nível mais mecânico e motor) e a natureza acústica dos sons produzidos nos coloca um

problema pelo seguinte: os tons da música ocidental estão organizados de modo a terem relação de graus de uma escala, ou escalas, que se relacionam. Este aspecto acena para o fato de que o nosso saber musical seja também simbólico. Ora, se as melodias que cantamos, ou mesmo aquelas que apenas cantarolamos, resultam afinadas é porque nosso aparato cognitivo já aprendeu e memorizou as distâncias entre tons que se combinam mais ou menos. Essa é, no entanto, uma questão que exige longa e profunda discussão sobre cognição musical e que envolve, antes de mais nada, aspectos físicos do som e a percepção desses. Vê-se então que apresentar rapidamente essas ideias seria tentar passar por cima de estudos introdutórios e mais qualificados. Sendo assim, para manter o espírito do texto, invoco alguma intuição nossa sobre o assunto do conhecimento musical.

### **Tons pulsando no ar e ouvidos que já sabem música**

Em linguística ou ciências da cognição, quando tratamos do conhecimento estamos falando de explicar a capacidade humana para alguma atividade que se manifesta em nossas ações. Um exemplo simples de que inferimos que há um pensamento que ocorre antes ou mesmo durante a ação, é o da devolução de um troco em dinheiro vivo. Mesmo que a máquina tenha dado o resultado final do cálculo, foi o indivíduo que informou à máquina o cálculo a ser feito. Então vejamos o que acontece com a música. Melodias advindas de diversos instrumentos musicais ou das diferentes vozes humanas, proporcionam-nos, como já vimos tons fundamentais. Esses tons são os que nossos ouvidos ouvem mais, pois possuem grande energia. Mas não vivem jamais sozinhos. Eles precisam de seus sobretons ou harmônicos, pulsando no ar, juntamente com eles. A pulsação de tons e sobretons (fundamentais e harmônicos) é de tal natureza que alguns se sobrepõem e se combinam e outros são detectados por nossos ouvidos como díspares, criando um desconforto auditivo. Ao descrever isso, explica-se que os sons da escala ocidental não são escolhas arbitrárias, mas que sua organização reside em algo que é natural de nossa percepção auditiva. E muito provavelmente é isso que explica nossa facilidade para aprender e decorar melodias.

Em relação às características mais individuais das vogais, uma sequência possível e que para mim soou natural do estudo de produção, foi fazer um estudo de percepção de

cada vogal cantada e suas contrapartes faladas. Elaborei um projeto, que, no entanto, não se aprofundou tanto quanto eu queria. Um caso que me intrigava e instigava era o comportamento da vogal “e” cantada. Parecia ser mais difícil articulá-la comparativamente a todas as outras. Na fala, essa vogal parece também ter o que chamarei intuitivamente de instabilidade articulatória, com a cautela de não relacionar esta impressão com o estudo sobre estabilidade que foi mostrado no capítulo anterior.

Poderíamos lançar uma hipótese assim um tanto iconoclasta e dizer o seguinte: já que a vogal “e” é relativamente ineficaz na língua, pois acaba sendo produzida sob outras várias formas, porque ela não é descartada da língua? E a resposta mais singela é: por que ela é um recurso distintivo como todas as outras vogais. Ainda que ela tenha uma articulação mais frágil – o que tem que ser comprovado – ainda assim, é um recurso importante ou uma riqueza da língua, útil para a criação e manutenção do signo linguístico. A natureza desse signo reside na diferença. Pensando em termos mais abrangentes, os gestos que compõem o signo, sejam eles sonoros ou visuais (libras, por exemplo), têm de contrastar entre si para que possam cumprir o papel de significar as diferentes coisas no mundo.

Vemos que a discussão sobre uma pequena unidade linguística invoca questões fundamentais de uma ciência. Ao falarmos de signo, distintividade e contraste, tocamos nas ideias fundadoras da linguística moderna. Assim, o fato de aprofundarmos nosso conhecimento sobre as vogais das línguas, em vários aspectos e níveis, já conduz a outras questões não só relativas à linguística, mas a todos os tipos de linguagem. Vejamos quais sejam:

1. Se as linguagens têm unidades maiores e menores, quais seriam as menores? Tanto na música como na língua?
2. Essas unidades são universais? Ou seja: a percepção de “partes menores” a que classificamos como notas musicais ou vogais é verdade para todos os seres humanos e culturas?
3. É possível falar de hierarquia entre as unidades: algumas mais importantes que outras?
4. É possível falar que as partes menores coincidem com as unidades primitivas, entendendo “primitivas” como aquelas unidades básicas ou fundamentais para a existência da capacidade humana, aqui em tela a língua e a música?

Perguntas como essas nos colocam na posição de explicar, nada mais nada menos, a capacidade fundadora da linguagem. Já de muito tempo essa é uma pergunta da linguística que tomou forma através do pensamento de Noam Chomsky na década de 1950. Para esse grande linguista e grande pensador, a capacidade de sintaxe é aquela responsável central por nossa capacidade linguística. Explica-se que o “saber sintaxe” dá-se em um nível abstrato e, portanto, a fala fica fora dessa explicação.

Mas é a fala que sempre me interessa. E para piorar a minha situação, é a fala cantada que me interessa mais. Vê-se então que uma porção de ingenuidade e um bocado de insistência levaram-me aos lindes dos estudos da linguagem. Ali, nesses limites encontrei a fonética refrescando-se em uma onda um tanto inusitada, a da fala sintetizada. Isso mesmo: máquinas falantes são uma espécie de sonho de consumo de vários setores da comunicação. A facilidade de trabalhar com processamento de sinal no computador e criar programas para análise do sinal sonoro ajudou a fonética a recuperar um pouco do *glamour* que tinha antes do início do estruturalismo. Desde o início do milênio, a fonologia, disciplina que está em interação e se confunde com a fonética, passou a ser fonologia de laboratório. Para toda a hipótese fonológica, teria que haver uma comprovação com dados acústicos, articulatórios e aerodinâmicos. Ou seja, comprovados com os parâmetros de muito já estudados na fonética. Já que a fala é feita de movimento, som e ar. Tudo junto.

O fato é que a fala se desenrola no tempo e, portanto, dados de duração são sempre muito elucidativos de questões atraentes como ritmo de fala, sobre o que tratei na parte 3.

A pergunta mais fundamental seria:

— É possível falar de uma organização temporal entre as unidades?

Respostas para isso a fonética já vem fornecendo há mais de meio século, por exemplo, com o caso das consoantes aspiradas do inglês (e de outras línguas). Elas acontecem em palavras como “pie” (torta) e “tie” (gravata) e são o pivô do nosso incômodo sotaque quando falamos inglês. Nossa língua brasileira não tem essas consoantes aspiradas, cujo detalhe acústico-articulatório está em atrasarem o vozeamento das vogais e serem percebidas com uma explosão mais ruidosa. É claro, pois, que essa distribuição de tempo entre os fones, que apenas parece imperceptível, é parte integrante das unidades da fala.

Vê-se, com esse exemplo, que o que chamamos de unidades da fala, não se caracteriza apenas por suas propriedades acústico-articulatórias, mas também pela

coordenação de tempo entre os fones em combinação, coordenação essa que bem podemos chamar de *timing*. Essa visão mais detalhada da produção do fone é possível pela fonética e aprofunda-se no quadro teórico de uma fonologia dinamicista, da qual já falamos anteriormente na parte 3, e para a qual o tempo é fator levado em conta na concepção de uma unidade abstrata da fala.

Penso que fica fácil, assim, ver que os fones são pequenas unidades linguísticas e possuem um lado físico inescapável. Obviamente são unidades de natureza muito diferente das unidades sintáticas. Além disso, não há evidências de que sejam primitivos no sentido de originarem toda nossa capacidade linguística. Aliás, sobre isso, vimos uma proposta de separar a origem das capacidades sintáticas da origem daquelas fonológicas na parte 2. Também tivemos a oportunidade de citar autores como T. Fitch, que juntamente com Noam Chomsky, há não muito tempo, mantém a proposta de que a origem da linguagem repousa sobre a faculdade estreita da linguagem e que deve ser entendida basicamente como capacidade cognitiva abstrata. O campo da discussão da origem e evolução da linguagem já conta com ideias importantes e autores consagrados, sendo assim não há muito que discutir sobre as capacidades de língua natural, uma vez que sequer este é o nosso objetivo.

A ideia central, agora, é propor um primitivo musical, contextualizando a proposta a partir da trajetória dos estudos comparativos entre fala e canto que desenvolvi a partir da tese de doutorado. Como vimos anteriormente, aqui mesmo nesta parte do livro, muitas questões já estavam latentes diante do achado da uniformização do tubo canoro e de um ajuste de vogais e consoantes para não perderem de todo sua inteligibilidade.

A questão de natureza epistemológica foi tratada na parte 2, na qual discuto os elos que atam a música à linguística. Podemos pensar que há vários elos em camadas diferentes, essas últimas sendo compreendidas como conexões entre as disciplinas, que muitas vezes são unilaterais. É o caso de empréstimos de termos como harmonia, do qual falamos bastante, esclarecendo a relação que existe entre o que se considera harmonia na fonologia e harmonia na música. Um exemplo interessante, também, na utilização de termos, é a palavra “coda”, bastante difundida entre os especialistas de fonética-fonologia para designar o final de uma sílaba, sobretudo do tipo CVC (consoante-vogal-consoante). Em música, a coda é a porção final de uma peça musical. Mesmo que outro trecho se repita depois dela, deve-se voltar a ela para terminar a peça.

Mas não só de empréstimos e trocas de termos vivem as linhas de conexão entre as áreas do saber e as disciplinas. Há relações mais diretas como aquelas de aspectos acústicos e fonéticos, desvendados em uma série de estudos específicos. Não deveria surpreender que tais estudos sejam empreendidos nas áreas da linguística e fonoaudiologia, já que a propriedade sonora das duas linguagens (a língua natural e a música) é obviamente seu elo por excelência.

O estudo de percepção fortaleceu as hipóteses que iam surgindo a partir das conclusões do doutorado de que há uma negociação entre linguagens diferentes, no nível da ação – ou seja, quando se produz a fala cantada (o canto). Haveria nas estruturas que se entrelaçam um ajuste, de modo a garantir que alvos fossem atingidos: afinação e compreensibilidade do texto cantado. De que forma, ou como explicar o sucesso da obtenção desses alvos? Que estrutura ou processo cognitivo estaria por trás disso?

Ainda não foi, porém, nos primeiros contatos com ideias direta ou indiretamente ligadas ao aspecto cognitivo que desenvolvemos estudos para responder perguntas sobre as unidades, sua estrutura hierárquica, e sequer de uma faculdade de fala cantada. Em geral, estudos não evoluem linearmente. Pouco tempo depois do término do doutorado, pude ter tempo para dedicar-me ao estudo mais aprofundado da fonologia dinamicista, estreitamente ligada à fonética. Nos domínios da linguística, foi essa teoria – e ainda é – que parecia responder melhor às questões da organização fônica e sonora da língua. O tópico escolhido para desenvolver uma explicação no seio da fonologia de abordagem dinâmica foi a nasalidade vocálica, para o qual escrevi mais de um projeto, inclusive o de pós-doutorado. Nasalizar as vogais no português brasileiro é ter capacidade de realizar um ajuste refinado de articuladores da fala. É fenômeno tão brasileiro como o ritmo do samba!

Findado o pós-doutorado, sem perder de vista a interdisciplinaridade – aspecto importantíssimo da minha pesquisa – elaborei um curso de pós-graduação que contemplava um diálogo entre várias áreas do conhecimento, tendo a fonética, portanto os estudos dos sons da fala, como ponto convergente entre elas. Em meio a discussões desse curso<sup>72</sup>, surgiu a questão do deslocamento do acento no canto. Essa questão foi

---

<sup>72</sup> O curso foi assim identificado: FLC6058 - Fonética e Fonologia: Integração e Interdisciplinaridade. Sua ementa trazia os seguintes objetivos: esta disciplina tem o objetivo de traçar um panorama no qual as fonologias de tradição gerativista e aquelas oriundas de uma teoria dinâmica (a fonologia articulatória) apontam para um percurso dos estudos dos sons dentro da linguística que leva à integração entre a fonética e a fonologia.

trazida por Cássio Santos<sup>73</sup>. Cantar uma palavra deslocando o seu acento é um fenômeno interessantíssimo. Para mim, é um recurso estilístico entre a fala e o canto, proposital, ou seja, faz parte da composição e não pode ser empregado a todo momento, sob o risco de perder a graça. Então, este novo tópico no terreno das comparações fala/canto trouxe-me aos mares do ritmo, navegados a bordo da sincronização entre dois cantores/falantes.

Claramente não foi um trabalho em um quadro teórico tradicional de ritmo. Por isso mesmo em vários momentos empreguei o termo “estrutura temporal” para me referir ao fenômeno estudado. O importante a destacar aqui nesta exposição dos estudos é o achado sobre a sincronização maior no *rock* do que em outros tipos de canto e de fala. A partir daí é possível afirmar que a estrutura temporal musical é tal que permite maior integração (*entrainment*), ainda que saibamos que os indivíduos, de forma geral, sincronizam bem suas ações. À música, e conseqüentemente ao canto, atribui-se uma estrutura temporal diferenciada. E por que isso seria assim? Por que isso aconteceria?

O grande risco é levantar muitas questões e não ter meios de respondê-las. No entanto, as descobertas pareciam cada vez mais apontar para o fato de que nossa linha de pensamento ia no rumo certo, ou, mais ambiciosamente, posso colocar assim: nosso programa de pesquisa estava conseguindo tecer uma trama consistente para explicar os aspectos da relação fala-canto. Em outras palavras: o fio das relações entre fala e canto parecia firme o suficiente para que atingíssemos compreensão razoável dessa relação. Essa compreensão, se não respondia categoricamente as perguntas, mantinha-as procedentes.

Dois fenômenos mais localizados lançavam luz sobre indagações mais amplas como as já listadas anteriormente. O primeiro fenômeno é aquele da configuração de nossa cavidade oral para o canto em comparação ao modo como configuramos essa cavidade para a fala. Ao cantar, sobretudo o canto sem microfone, fazemos outras escolhas articulatórias. No caso do canto erudito, vimos que buscamos um tubo mais uniforme para pronunciar as vogais sem ter de desafinar o que a laringe já afinou. Há outros tipos de canto que tradicionalmente não são amplificados eletricamente e que lançam mão de outras manobras articulatórias para garantir o volume de sua emissão. Cantos tradicionais religiosos e cantos de trabalho parecem ter suas próprias manobras ou

---

<sup>73</sup> Ver Cássio Santos. Vogais cantadas e tonicidade: estudo experimental comparativo entre fala e canto com foco na duração. 2017. 129 f. Tese (Doutorado em Semiótica e Linguística Geral) – Universidade de São Paulo, 2017.

posicionamentos articulatórios para proporcionar que uma voz alcance a intensidade desejada. Só essa constatação já justifica uma pesquisa para buscar aprofundar a teoria do tubo, essencial à fonética acústica, pensando, agora, em um tubo do canto. Ao mesmo tempo, é preciso abordar a contraparte abstrata destes gestos canoros; ou seja, relacionar a faculdade de música com as habilidades de movimento coordenado do trato vocal, aí incluída a laringe. O segundo fenômeno diz respeito à facilidade maior de se sincronizar a fala cantada no ritmo do rock, em relação à fala falada mais ritmada, como o caso da parlenda. Isso implica que além do casamento entre nota musical e sílaba, há uma “otimização” da coordenação desse canto quando tem de ser sincronizado entre dois indivíduos cantantes. Há, aí, uma influência do canto sobre a fala e, por conseguinte, podemos dizer que há um predomínio das restrições musicais temporais sobre uma ação linguística. Neste ponto, não podemos desprezar a ideia de unidades diferentes das duas linguagens buscando uma produção eficaz ou ótima. Assim, clarificar quais são essas unidades e como essas convivem, segue sendo importante para erigir conceitos capazes de designar bem os fenômenos em tela e conseqüentemente desenvolver cada vez mais estudos que configurem uma grande rede compreensiva dos cantares humanos.

A partir dos dois fenômenos constatados ao longo de alguns anos de investigação, parecia estar se delineando a ideia da estabilidade do canto *versus* a instabilidade da fala. Então a questão que emergiu foi: afinal qual é a grande diferença entre a fala e o canto?

A música é produto de uma organização sonora em que as grandes bases são a relação precisa entre tons diferentes e o ritmo – ou estrutura temporal – ser formado de eventos regulares. Os tons diferentes são chamados de alturas, pois distribuem-se em um espectro de frequências fundamentais. Organizados em escalas, tais tons são as notas musicais. Essas são organizadas temporalmente criando células rítmicas que podem e devem se repetir. Dito isso, afirmamos que a música representa a faculdade de produzir sinais sonoros estáveis, tanto no eixo temporal quanto no eixo das frequências. Esse foi o ponto de partida para desenvolvermos os estudos apresentados no capítulo 4.

Os resultados da estabilidade de  $f_0$  (frequência fundamental) trouxeram ânimo para desenvolver a ideia do gesto musical que eu vinha gestando desde 2012.

Esse gesto estaria para a música como o gesto fonológico está para a linguística. O problema é que eles não são conceitos muito intuitivos. É preciso, então, lançar mão de comparações e também de contextualizações a partir de conceitos mais divulgados na ciência. Parece estar razoavelmente bem divulgada a ideia de que o conhecimento é



constituído em nossas experiências diretas com objetos e estados das coisas e que nessas experiências as atividades sensório-motoras desempenham papel importantíssimo. A partir dessa visão, a ideia de corporificação ou corpóreo parece invadir todas as áreas da ciência que querem explicar o comportamento humano. Não deve ser diferente para a Linguística e a Música, ainda que para a primeira, desvendar os mistérios do movimento corporal tenha sido muito trabalhoso.

Já tivemos a oportunidade de dizer anteriormente que a cisão entre o mental e o físico foi necessária para que se pudesse entender a língua como objeto de estudo. Só na linguística moderna já temos mais de um século em que se propagam ideias mentalistas sobre a linguagem. Como se a língua estivesse fora do corpo. O mesmo acontece com a explicação da capacidade musical dos seres humanos. É como se a música existisse antes de qualquer manifestação, isso é: a priori. O apriorismo, ou seja, a ideia de que nosso conhecimento abstrato já está pronto é tão forte que se manifesta mesmo quando lembramos, de maneira enfática, que a primeira expressão musical do ser humano foi o canto. Ou seja, de um lado reconhecemos que o canto é primordial e por outro mantemos empregando a terminologia que sedimenta e difunde nossas ideias sobre faculdades cognitivas abstratas.

Pensar no canto como ação é um gancho para falar de um gesto musical como uma unidade que implica as noções de saber algo e fazer esse algo. Ao cantar acionamos o conhecimento que temos sobre os tons musicais e sobre como podem estruturar-se temporalmente. A parte do corpo que possibilita a produção musical nesses dois eixos é a laringe. Em princípio, podemos entoar qualquer melodia já criada pela invenção humana ou criar novas melodias com a voz. Proponho, então, pensar em dois conceitos.

O conceito de gesto musical, que sintetiza a nossa capacidade musical e o conceito de laringe como articulador desse gesto. Temos aí já a proposta de um primitivo musical que pode ser irmanado ao primitivo linguístico. Para este último já há muita discussão e artigos escritos. O mesmo não é verdade para o primitivo musical. A devida ênfase é dizer que só chegamos a tais ideias por que entrevimos nos eventos acústico-articulatórios do canto fatos potencialmente explicativos do entrelaçamento entre fala e canto, implicando vislumbrar também a relação entre música e língua. No entanto, não temos encontrado ideias que ressoam com as nossas, se não um mosaico de estudos com diferentes pontos de partida, para os quais buscamos revelar intersecções ou convergências.

Um exemplo rico desse mosaico é a publicação *Language and Music as Cognitive Systems*<sup>74</sup> que encerra trinta e três capítulos, incluindo a introdução e a conclusão. A maioria deles é escrita por acadêmicos cujas pesquisas baseiam-se em teorias representacionistas para explicar as capacidades de linguagem. Palavras e acordes – para dar exemplos de uma unidade linguística e de uma unidade musical – são “representados” a partir de um repositório desses itens mentais devidamente memorizados. Tais ideias ficam bastante evidentes na seção inicial do livro que trata de estruturas.

Embora as seções tentem agrupar títulos diferentes, várias delas apontam para o que chamamos de biolinguística e biomusicologia ou relacionam seus achados ao funcionamento do cérebro, como é o caso das seções *Evolution* e *Learning and processing*. Obviamente, a seção intitulada *Neuroscience* é aquela diretamente dedicada ao cérebro e em especial à questão da modularidade. A versatilidade que se exige do estudioso hoje em dia toma proporções inatingíveis. É preciso conhecer 6 milhões de anos de evolução das espécies que nos antecederam, além de compreendermos o cérebro em múltiplos aspectos, como o neuroanatômico e as relações entre suas regiões. Ultrapassamo-nos quando entendemos a ativação dos neurônios, as sinapses. Assim, louvamos a “floresta” bastante variada de ideias sobre a cognição linguística e musical em que estruturas como a rítmica e níveis como o sintático sejam comparáveis. Mas o que eu penso ser mais completo em termos de um programa de pesquisa é uma proposta seguida de uma discussão de um primitivo musical no campo das abordagens dinamicistas.

Em termos simples, colocamos assim tudo que foi dito sobre o gesto musical:

A produção do canto, assim como a produção da fala, é uma atividade ou ação cuja função é expressar uma linguagem. Essa produção é efetuada coordenando-se movimentos precisos de articuladores do aparato vocal de que dispomos, com ajustes próprios. A coordenação desses movimentos, que podemos entender como articulatorios, possui uma contraparte abstrata que ainda precisamos compreender e traduzir em algum tipo de modelo explicativo, tomando como base e exemplo, o que já se fez na fonologia de base dinamicista. No entanto, mais do que entender teorias que se desenvolvem no interior da Linguística, é preciso abraçar a ideia mais fundamental de um gesto musical e

---

<sup>74</sup> Patrick Rebuschat, Martin Rohrmeier, John Hawkins, and Ian Cross. *Language and Music as Cognitive Systems*. Oxford: Oxford University Press, 2012.

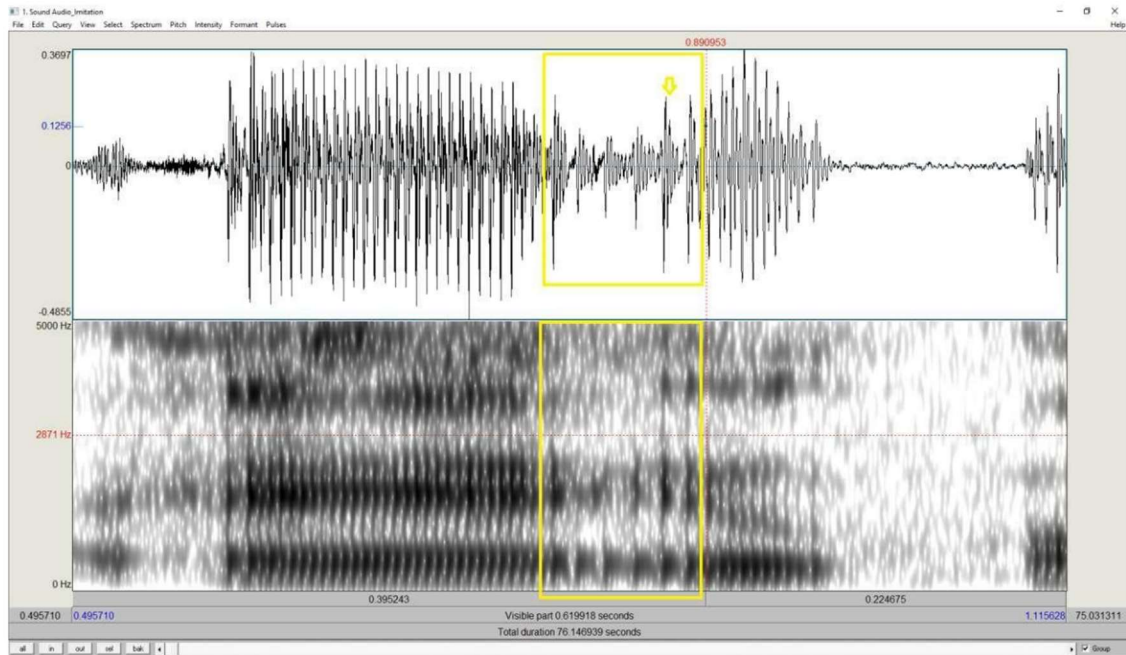
mesmo de uma laringe musical que produz esse gesto. Isso é possível ao encararmos o gesto musical como aquele no qual se encerram propriedades físicas e abstratas ao mesmo tempo. Uma questão pode resumir melhor:

— Ao cantar, quais conhecimentos mais corporais e mais abstratos estão sendo requisitados? E como eles estão integrados?

Veremos que é viável pensar na laringe como articulador<sup>75</sup>, o que é amplamente esclarecedor para qualquer estudioso da produção de fala. Assim como as unidades fônicas estão devidamente descritas, na fonética e fonologia, como produtos de posição articulatória e resultado acústico e aerodinâmico (por exemplo, a consoante [s] é uma fricativa dental surda), poderíamos descrever propriedades, tanto segmentais como suprasegmentais, que são produtos da articulação da laringe. Em relação a segmentos, sons uvulares, faríngeos e glotais (assim chamados tradicionalmente) têm suas constrições realizadas na região da garganta, para usar um termo mais popular. Para alguns estudiosos de como se produzem os sons da fala, tais sons, sobretudo os faríngeos e glotais deveriam ter sua articulação melhor explicada. No português brasileiro podemos falar dos “erres”: haveria diferença entre os sons do “erre” carioca e o mineiro (por exemplo, na palavra “porta”? Em relação a fenômenos suprasegmentais, podemos falar da voz rangida (*creaky voice* em inglês) e que é estudada como um aspecto acústico de qualidade de voz. A voz rangida soa, como o termo sugere, como se houvesse uma espécie de atrito entre duas superfícies em movimento (duas correntes por exemplo). Na forma de onda e no espectrograma (Figura 3), é possível ver a parte destacada, que representa uma maior separação entre os pulsos e uma energia maior do primeiro harmônico no início da palavra “eu”.

---

<sup>75</sup> Para linguistas e demais interessados nos sons laríngeos, indico a leitura do artigo “Articulatory Function of the Larynx and the Origins of Speech” de John H. Esling. Foi publicado em Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society (2014), pp. 121-150. Esses Proceedings encontram-se na plataforma digital eLanguage da Linguistic Society of America.



**Figura 3:** Forma de onda (acima) e espectrograma (abaixo) do enunciado “É eu”. Os retângulos amarelos enquadram a porção em que a voz rangida apresenta-se com maior evidência. No retângulo sobre a forma de onda vê-se que os pulsos (ciclos) estão mais espaçados em comparação com uma grande porção da vogal “é”, além de apresentarem picos iniciais maiores em relação aos demais picos. No espectrograma, a porção do retângulo destaca a trajetória de formantes interrompida pela sucessão de pulsos espaçados, espelhando o que acontece na forma de onda.

Este único exemplo já demonstra a riqueza de aspectos e detalhes que é preciso evidenciar para que compreendamos melhor a qualidade de voz que é, em grande parte, explicada por comportamentos diversificados da laringe. Por outro lado, se quisermos uma explicação apenas do seu funcionamento como aparato capaz de sustentar um som contínuo, delimitamos o objeto de estudo para a questão da estabilidade e seu papel no canto.

**FIM**  
**O Sabor do Saber &**  
**O Prazer de Cantar**

Eu me comprometi, na introdução, a chegar a um final feliz. Se por final feliz, ou qualquer momento feliz, possamos entender estar respirando bem e com planos futuros, cumpri o prometido ou cumpri-me. Falta cumprir-se o Brasil.

Mas acredito que o leitor e a leitora querem uma felicidade menos subjetiva e da qual possam compartilhar – exceto, claro, a ideia de que trabalhar pelo bem-estar da sociedade, que é missão da universidade, ajuda a tecer um país justo e igualitário. E há duas coisas que posso dizer sobre um percurso acadêmico como o que eu percorri: a primeira coisa é que os significados de “saber” e “sabor” entrelaçam-se, não apenas etimologicamente, mas de fato. “*Sapere*”, verbo no latim, é perceber pelo cheiro ou gosto, ou seja, é uma experiência de percepção. O significado de “*sapere*” é “gosto das coisas”. Para sentir o cheiro e o gosto, temos mais do que nos aproximar de alguma coisa. Temos que respirar o odor que ela exala e senti-la com o palato e a língua. Essa aproximação cria enorme intimidade com o objeto. No caso dos foneticistas, o objeto é a fala. No meu caso específico, é a relação entre a fala e o canto, cujo enlace (vejamos aí outro entrelaçamento) é o que me interessa estudar. Posso dizer que quanto mais sei sobre fala, mais fico maravilhada e penso que a produção de conhecimento deveria ser, para todos que vivem da pesquisa, um maravilhamento. A segunda coisa é que cantar é uma ação que envolve ajustes mais ou menos precisos na dimensão da fonação. Os processos — e não o produto — pode até não agradar os ouvintes, mas agrada a quem canta. Ou seja, cantar pode ser fonte de muito bem-estar e prazer para quem canta, principalmente para quem canta em grupo. Os benefícios que aparecem em efeitos fisiológicos como a respiração e a circulação já são conhecidos nas áreas médica e de musicoterapia e somam-se a outras consequências benéficas como a criação de elos sociais e sensações prazerosas<sup>76</sup>. Cantarolar para si mesmo, cantar acompanhado de instrumentos, cantar em coro, cantar com a multidão no show do artista preferido. Tudo isso dá prazer.

No entanto, difícil é aliar os dois lados desse canto. Há vários tipos e graus de envolvimento na relação que se pode estabelecer entre saber sobre canto e saber cantar. Ao longo da vida encontrei todo o tipo de saberes e cantares. Desde aquele cantor mais intuitivo que canta de forma avassaladora e bela até um cantor que tem conhecimento de toda a fisiologia, acústica e aerodinâmica da fonação e do canto com palavras (Johan Sundberg) e domina a técnica do belcanto. Só para dar um exemplo de terras brasileiras,

---

<sup>76</sup> ver levantamento de Jing Kang, Austin Scholp & Jack J Jiang. A Review of the Physiological Effects and Mechanisms of Singing. *J Voice*, 2018, pp. 390-395.

temos nomes importantes como o de Fabiana Cozza, que é excelente cantora popular e mestre em fonoaudiologia<sup>77</sup>, e de Ariel Coelho, estudioso da voz, especializado nas vozes do *rock*<sup>78</sup>. Menos conhecida, mas não menos talentosa é a Mariana Baltar<sup>79</sup>, também cantora da música popular brasileira e fonoaudióloga estudiosa da voz. Tenho certeza que se fôssemos fazer um levantamento dos cantores estudiosos da voz ou estudiosos cantores teríamos no nosso país um quadro enorme de vozes belíssimas de diferentes estilos de canto fazendo uma ponte do estudo e formação acadêmicos e o fazer artístico.

Um quadro menor, mas talvez não menos importante, é aquele no qual me incluo: o de linguistas que estudam os aspectos mais específicos do canto, obviamente no tocante à linguagem, portanto, no tocante à significação. Além do pequeno levantamento encontrado na parte 2, podemos destacar a faceta semiótica do estudo, não propriamente do canto, mas da canção. Aspectos prosódicos da fala podem ser vistos pelas lentes da semiótica tensiva, quando, por exemplo, a unidade analisada por Luiz Tatit é a sílaba tônica cantada e é vista como lugar de oscilações de sentido<sup>80</sup>.

O elo que liga as duas áreas maiores que, em termos de categorização, abarcam hierarquicamente a fala e o canto – linguística e música – é tão óbvio e mesmo assim é preciso resgatá-lo a todo tempo no oceano de estudos, pesquisas e ideias.

O elo é a fala. Esta afirmação deve vir com a compreensão e clarificação de que o canto é uma fala. Assim foi que minha orientadora de doutorado, Eleonora Albano, elegeu o termo “fala cantada” para diferenciá-la de “fala falada”, ou seja, para nomear os fenômenos de fala que estávamos a ver no canto. Pareceu-nos, à época, que apenas fala *versus* canto não nomeava exatamente o que estávamos observando: as intersecções entre os aspectos acústicos da fala e seus correspondentes no canto. Foi o caso, por exemplo,

---

<sup>77</sup> A sua dissertação pode ser encontrada na TEDE, Biblioteca Digital de teses e dissertações da PUC de São Paulo. A referência completa é: Santos, Fabiana Cozza dos. O canto em Linda Wise: ação imaginativa e interpretação. 2019. 110 f. Dissertação (Mestrado em Fonoaudiologia) - Programa de Estudos Pós-Graduados em Fonoaudiologia, Pontifícia Universidade Católica de São Paulo, São Paulo, 2019.

<sup>78</sup> Há vários tipos de canto no *rock*, desde o mais melodioso (*rock* melódico) até o mais ruidoso (por exemplo, o *heavy metal*). Para os tipos mais ruidosos, utiliza-se o que se chama de drive vocal e que designa uma qualidade de voz rouca, com alguma sopro e ao mesmo tempo mais intensidade do que, por exemplo, *flow*, que é um tipo vocal do canto em que a intensidade dos harmônicos dá-nos a percepção de uma voz não ruidosa.

<sup>79</sup> Sobre esta cantora, vale reproduzir o que o crítico Mauro Ferreira escreveu no seu blog do portal G1, em 13/10/2019: “Quão grande é o canto de Mariana Baltar! Quão pequeno pode ser o Brasil ao relegar uma cantora do porte dessa carioca à margem de mercado... Voz da resistência, Baltar arreventa o cordão de isolamento com álbum, **Os arcos – Paixão e morte**, gravado com o conjunto de choro Água de Moringa e dedicado ao cancionista de Aldir Blanc, bamba das letras.”

<sup>80</sup> Projeto de pesquisa de Luiz Tatit, como pesquisador do CNPq: *O papel da prosódia na epistemologia semiótica*.

da duração das consoantes: embora produzidas com a mesma duração média nas modalidades falada e cantada (cerca de 150 milissegundos), as consoantes oclusivas desempenhavam papéis temporais diferentes em cada uma dessas modalidades. Isso acontecia porque, relativamente à duração da palavra, a duração da consoante era menor no canto do que na fala. A consoante cantada ocupava apenas 15% da duração total da palavra, ao passo que quando falada, representava 34% dessa duração.

Não importa se usamos os termos “fala *versus* canto” ou “fala falada *versus* canto falado”, a questão é que tais fenômenos são objetos de estudo que, em sua natureza, já estão ligados. Aliás, é a ligação mais evidente e óbvia e, assim, penso que tardamos em erigir questões, estudos e hipóteses sobre eles. Felizmente, neste livro, pude tratar dessas questões. Espero que tenha feito isso de modo a não distanciar o saber do sabor.

Mas vamos imprimir um andamento mais rápido a isso aqui, já que não faremos ciência ou não teremos vida acadêmica se não pensarmos na contribuição a uma teoria ou a uma aplicação dos achados que, em última instância, são produtos da busca de conhecimento. Aliás, cabe frisar que nos últimos anos da vida na universidade, temos sido praticamente compelidos a mostrar o lado aplicado de nossas pesquisas. Então começo por ele.

Esse é um lado dos estudos aqui apresentados, que fica à sombra dos estudos linguísticos teóricos. No entanto, não é menos importante e sequer deve ser visto como menos relevante do que complexos estudos de comprovação de uma teoria explicativa.

Os resultados da minha tese de doutorado derrubam o mito da má dicção. Há dois ganhos iniciais na queda desse mito: a supressão da responsabilidade de um cantor/a considerados incompetentes em articular os sons da fala cantada e a possibilidade de se transformar o ensino do canto a partir do conhecimento acústico-articulatório da produção de notas agudas. Em larga escala, então, espera-se que, o ato de cantar tão natural e universal para nós humanos, desenvolva-se e faça florescer a habilidade de forma justa com os limites do nosso corpo e de seus movimentos.

Jamais teria chegado a este resultado se não fosse minha aproximação em relação à fonética. Este movimento em direção aos estudos dos sons da fala só pôde ter sido feito no âmbito da linguística. Vemos que há grande quantidade de trabalhos sobre a voz cantada no campo da fonoaudiologia. Só no intervalo dos anos de 2008 a 2012, houve 180 trabalhos acadêmicos relacionados à voz cantada, segundo levantamento de Marta



Silva e colegas, em 2013<sup>81</sup>. Outro levantamento, o de Lorena Drumond e colegas<sup>82</sup>, no âmbito das pesquisas fonoaudiológicas, volta-se para a voz no canto popular, e tendo se utilizado do banco de teses da CAPES, apresenta alguns títulos de pesquisas que figuram também no levantamento feito por mim, apresentado no capítulo 2 deste livro. Meu trabalho de doutorado tem ligação estreita com esses estudos na área da fonoaudiologia, mas jamais poderia ter se desenvolvido plenamente nessa área. A razão para isso é que o foco era a produção fônica do canto em comparação direta com essa produção na fala. Tratou-se, o tempo todo, de uma correspondência entre fala e canto, baseada no detalhe fonético, portanto estando em jogo a fala e sua faceta mais linguística: a da distintividade. Por isso a resposta à questão do primeiro capítulo é que cantamos para sermos entendidos.

Então, a canção e o canto são o lugar do encontro da emoção e do prazer com a razoabilidade: a comunicação em estado perfeito. Isso dá muito prazer, sem dúvida.

Mas não só de aplicações vive o espírito humano questionador. No capítulo 2 pudemos levantar várias questões no “leva e traz” da linguística e da música que podem ser erigidas a partir do fenômeno canto. Essas questões alimentam os estudos e descobertas em torno da fala e do canto, tocando aspectos vastamente investigados pelo mundo afora, como os de prosódia. Retomando as questões do capítulo 2 e mantendo a firme proposta de que constituam pontos de partida para um programa de pesquisa cujo foco seja desvendar como se fundem elementos da fala e elementos do canto:

— Todos os gêneros de canto se diferenciam da fala de uma forma básica? Todas as formas de cantar, considerando várias culturas, apresentam essa mesma forma? Sílabas corresponderiam a notas? Se sim, como se dá tal interação?

Quais as hipóteses em torno dessas ideias? Proponho que a música possa ser estudada (também) do ponto de vista biológico, ou seja, do ponto de vista de que é uma capacidade que se desenvolve por qualquer indivíduo, primordialmente na produção da voz cantada. Não se trata, neste momento, de falar da função do canto, de avaliar e/ou classificar tipos de canto. É preciso propor uma investigação que aborde a percepção das pessoas sobre música/canto e tratar de detectar qual é o traço que as distingue. A tímida questão de sílabas-notas musicais pode parecer de pouca envergadura, mas há, aí, um

---

<sup>81</sup> Marta Assumpção de Andrada e Silva, Ana Carolina de Assis Moura Ghirardi, Maria Fernanda Prado Bittencourt & Luciana Assanti. *A voz cantada*. São Paulo, 2013.

<sup>82</sup> Lorena Badaró Drumond, Naymme Barbosa Vieira & Domingos Sávio Ferreira de Oliveira. *Produção fonoaudiológica sobre voz no canto popular. Fonoaudiologia Baseada em Evidências*. J Soc Bras Fonoaudiol, 2011, pp. 390-7

mundo a ser buscado. No tempo de uma sílaba falada, cerca de 250 milissegundos (ou 0.25 segundos) podemos verificar e inferir comportamentos articulatórios e coarticulatórios (relação consoante-vogal), a frequência fundamental e até mesmo o timbre da vogal através do comportamento dos harmônicos dessa  $f_0$ . Sendo tônica, podemos observar, medir e chegar a explicações sobre sua relação temporal com as demais sílabas do enunciado, sobre a relação temporal e a vogal pronunciada (relação tempo-característica articulatória da vogal).

O conjunto de detalhes fonéticos contidos em uma sílaba, na qual podem ser inspecionados, desvenda as relações da fala e do canto e nos aponta para que alarguemos o escopo do estudo. Por isso passamos a falar da capacidade do canto como um objeto de estudo possível. Também propusemos a ideia que não se pode tratar da habilidade de cantar sem chegar à cognição musical.

Tendo aderido a uma visão de cognição em que o corpo e sua ação no tempo são elementos explicativos de como se aprende ou se sabe algo, a questão do ritmo logo se impôs. Fascinada pelas ideias de coordenação e de *entrainment* como relacionados ao fenômeno de ritmo na fala, apostei na observação da sincronização entre pares de falantes e cantores para diferentes tipos de texto. Neste momento viu-se a necessidade de aprofundar o conhecimento sobre o que chamamos de ritmo simples e complexo. Tal necessidade me fez aprender e entender, de forma objetiva, o ritmo do samba cantado. Aquilo que era tão naturalmente cantado desde criança pequena, apresentava-se, agora, como um mundo misterioso de combinações específicas de síncopas: padrões elegantes de tempos entoados das notas musicais ao longo de duas batidas (ou pulsos), agrupadas em uma janela de tempo que se chama compasso. Essa janela é constantemente atravessada por uma nota que começa na janela anterior e se alonga para a seguinte. Vê-se isso na partitura do samba de Adoniran Barbosa, na Figura 1. A célula rítmica apresentada nessa figura apresenta variações ao longo da canção e isso por si já representa um objeto interessante para mantermos o estudo sobre um padrão da síncopa do samba.

## Trem das Onze

Adoniran Barbosa

$\text{♩} = 90$

Não pos-so fi - car nem mais um mi - nu - to com vo - cê\_\_\_ sinto mui-

- to, A-mor,\_\_\_ mas não po - de ser\_\_\_\_\_ Mo - ro em Ja - ça - nã\_\_\_

se eu perder es - se trem\_\_\_ que sai a-go - ra às on - ze ho - ras\_\_\_ só a-ma-

**Figura 1:** 18 primeiros compassos do samba *Trem das Onze*, de Adoniran Barbosa. Os compassos destacados no balão (10º, 11º e 12º) apresentam o contraste entre um compasso sem síncopa como o 10º e os compassos subsequentes constituídos de uma célula rítmica típica do samba. Nesta célula, a síncopa se realiza do final do primeiro pulso (sílabas “Ja”), antecipando a batida do segundo pulso, representado pela ligadura. Também se realiza do final do segundo pulso (sílabas “nã”), que está ligado à nota do compasso seguinte.

A beleza do inusitado e o encantamento do ritmo do samba têm a ver com o paradoxo encontrado. Espera-se que a música seja mais regular. O *rock* foi o ritmo musical que melhor sincronizou. Em princípio, a facilidade de sincronização poderia ser aplicada ao samba, que é música também. Mas é mais difícil de sincronizar. Isso não é um problema artístico, mas apenas desvenda uma característica da organização temporal do samba, evidenciando uma característica singular deste tipo musical. É a sua graça: a síncopa que consegue fugir dos movimentos mais simples e comuns. E daí poder ser considerado ritmo complexo.

Novamente a ciência da linguagem e a da cognição tiveram, aí, um papel crucial. Para ir além da visão descritiva – necessária para a derrubada de mitos – e explorar as questões que advinham dos estudos de observação, era preciso manter-se conectada ao ideário das teorias sobre a linguagem humana e sua capacidade. Aqui neste livro tive a oportunidade de falar da maneira mais simples – apostando que seja também informativa ao leitor leigo – sobre algumas teorias, levantando alguns tópicos importantes:

- O detalhe fonético acústico e seu papel de elo entre disciplinas da linguística e outras disciplinas.

- A música e a linguística ligadas nas áreas da fonologia e da fonética.
- Uma teoria musical baseada na Gramática Gerativa.
- A fonologia dinamicista.
- A origem da capacidade da linguagem e também do canto.
- As teorias centrais da linguística e sua natureza mentalista.
- A visão de ritmo na linguística e uma visão alternativa, ligada à coordenação de gestos motores.
- A sustentação fônica como traço do canto: a estabilidade.
- O gesto musical: com base no gesto fonológico.

Foi uma missão possível tratar todos esses temas em um território fronteiriço que eu propus desde meu projeto de doutorado. Por isso ressalto a intersecção de áreas do saber como muito frutífera. A interdisciplinaridade instiga questionamentos inéditos e para os quais muito trabalho a muitas mãos ainda tem de ser feito. Aliás, para o empreendimento desta tarefa realizada por vários/as, estudiosos, há a chamada de Fred Cummins em seu livro *The ground from which we speak: joint speech and collective subject*. Para ele, que nos presenteou com uma palestra inigualável sobre “joint speech”<sup>83</sup> em 2020, em plena pandemia, há muito mais fenômenos inexplorados, como a fala conjunta, que ficam à margem da linguística tradicional e a consequência disso é que estamos perdendo este terreno. Entendo que não estamos erigindo as perguntas necessárias para abordá-los e, assim, ficando sem respostas, na maioria das vezes, mais do que necessárias para aclarar discussões intermináveis. Isso está alinhado ao que falei, na parte 2, sobre a distância entre falar da origem das línguas na paleontologia, por exemplo, e não se falar disso nos terrenos da linguística.

Trata-se de dedicarmos, de fato, tempo e atenção para as questões tidas como paralinguísticas ou extralinguísticas. Algumas delas podemos resumir aqui como alinhamentos de comportamentos que estudamos tradicionalmente como estranhos ao sistema linguístico e à realização da fala. É o caso das expressões da face e sua relação com aspectos fonéticos, da adaptação prosódica, dos gestos corporais alinhados à fala, e de qualidade de voz. Há estudos sobre esses temas, elaborados por grupos de pesquisadores nos quais muitos membros são linguistas. No entanto, a meu ver, não são

---

<sup>83</sup> *Joint Speech Language and Linguaging* foi proferida em 17 de julho de 2020, na Abralín ao Vivo (Linguists Online). Este evento contou com apresentações diárias de 4 de maio a 31 de julho em transmissão ao vivo, pela internet.

incorporados como deveriam a um programa de pesquisa cuja meta é explicar nossas capacidades cognitivas de linguagem. Na verdade, os estudos que podemos classificar como paralinguísticos têm se multiplicado, como é o caso desse estudo aqui: “Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction”<sup>84</sup> e são apresentados em fóruns pertencentes à Ciência da Fala. Cabe abrir agora um espaço para tecer considerações sobre o que representa um artigo como esse. Indubitavelmente ele é de interesse para os estudos da linguagem por levantar questões como a de adaptação prosódica que investiga detalhadamente comportamentos do seguinte tipo: ao concordar com uma pessoa em uma conversa, nós podemos entrar em adaptação sonora com ele, sim, produzindo, por exemplo, uma entoação semelhante, ou seja, ajustando nossa frequência fundamental aos contornos de *pitch* produzidos pelo nosso interlocutor. Vejamos alguns elementos de autoria e âmbito de publicação do artigo em tela. No tocante à afiliação dos autores e nome de seus grupos ou laboratórios, cabe notar que a palavra departamento parece não mais nomear as sub-unidades de uma instituição de pesquisa. Dois autores pertencem ao Laboratório de Comunicação da Fala (Trinity College Dublin) e os outros dois autores ao Instituto para Tecnologias Criativas da University Southern California. Parece que essas nomeações apontam certa autonomia dos grupos de pesquisa em relação à universidade a que pertencem. Talvez essa ideia de “pertencimento” ou parte integrante de uma universidade esteja já dando lugar para a ideia de uma rede de grupos financiados por seus próprios projetos em todos os aspectos, inclusive o da infraestrutura. O início do artigo já demonstra a necessidade dos autores de destacar o aspecto aplicado do estudo. Vejamos como fica com minha tradução do inglês para o português:

Sistemas de diálogo falado fazem uso de várias tecnologias de linguagem e são cada vez mais usados para facilitar e aumentar a comunicação humana, particularmente, através de seu uso na Human Computer Interfaces (Oviatt, 1996; Coulston et al., 2002). Essas tecnologias de fala são utilizadas em uma gama variada de áreas, incluindo a das comunicações móveis (Ward and Nakagawa, 2002; Lu et al., 2011; Agarwal et al. 2011) máquinas de busca da internet (Google, 2011, Apple Inc, 2011), jogos e tecnologia de assistência desenvolvidas para idosos (Kleinberg et al. 2007) ou para deficientes comunicacionais (Zhou et al., 2012, p. 11).

---

<sup>84</sup> Céline De Looze, Stefan Scherer, Brian Vaughan & Nick Campbell. Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Communication* Volume 58, 2014, pp. 11-34

A proposta do artigo é compreender os elementos de interação entre indivíduos em uma situação de comunicação linguística, elementos que são nomeados como afeto, emoção, atitudes e intenções do falante. Fica claro que o que está em jogo é sofisticar cada vez mais os estudos sobre o comportamento humano em contexto comunicacional e achar padrões que possam alimentar um aparato de inteligência artificial, para que, então, o computador saiba reconhecer e interagir com o ser humano. Nessa empreitada, unem-se cientistas das linguagens e cientistas da computação, mas o objetivo não é mais desenvolver uma gramática da língua, explicativa do que sabemos dela. O objetivo é criar jogos e aplicativos. Ficam então bem contempladas as áreas da indústria da diversão e a da saúde. Portanto, propostas mais teóricas da cognição e do comportamento humano ficam em segundo plano. Isso causa, ou pode causar um empobrecimento das teorias linguísticas e da filosofia da linguagem, que se encontram abrigadas nas ciências humanas.

Assim é que apesar de ter escrito dois artigos de maior alcance, por que em inglês como o é *Towards a musical larynx: preliminary Proposal e Singing and speech as comparable phenomena: a dynamical approach*<sup>85</sup>, em dado momento, tive de pensar qual era a contribuição aplicada de ter desvendado a estabilidade de  $f_0$  no canto comparado a uma instabilidade da fala. Estabilidade comprovada, pode-se colocá-la em algoritmos e aí são os especialistas computacionais que derivam a inteligência capaz de distinguir fala de canto e o que mais desejarem. Isso, para nós na fronteira das ciências e das artes, é a mesma coisa que fabricar tecido e aquecer o mercado mundial. E só.

Tecer o tecido que nos aquece de fato, porque aquece os sentimentos e nossa capacidade de expressão, é necessário, mas facilmente temos nos esquecido de fazê-lo. Quero ressaltar com veemência que não podemos estudar algo porque devemos aumentar o número de estudos científicos. Esse “algo”, na minha pesquisa, nos é dado pelo fazer de inúmeras vidas humanas que se expressam e se conectam pela língua que falam e pela música que cantam. O objeto que estudamos é o produto da criatividade e do estar no mundo do ser humano. Então é nosso dever explicá-lo, para que seja compreendido e torne-se sabedoria. É preciso enaltecê-lo para que seja devidamente acolhido por diversas gerações. Como vimos, podemos ter, no país, um quadro excepcional de estudiosos do canto com as mais diversas formações. Um programa dos

---

<sup>85</sup> Ambos artigos são capítulos de livro e defendem estudos comparativos entre canto e fala, respectivamente sobre o gesto musical e a laringe musical.

cantares e cantos brasileiros e suas relações com diversas áreas do saber só viria a fortalecer nossa identidade. Reverência e valorização de um fazer musical e linguístico que é único no mundo. Incide diretamente sobre a felicidade.

A contribuição mais profunda é conceber que a fala cantada (o canto) é objeto natural dos estudos da linguagem e merece um lugar bem definido na linguística, obviamente considerando suas relações com a música. Os estudos comparativos que desenvolvi ao longo de duas décadas versam sobre como são cantadas as vogais, a sincronização na fala e no canto, como forma de abordar o fenômeno do ritmo e a busca de entender a estabilidade no canto. Tudo isso tem como ponto de partida a fala e veio dar na praia em que o mister é explicar o gesto do canto. Vejo este gesto como um fenômeno importante para entender a capacidade musical e estou pronta para me lançar nas águas de mais reflexões e de estudos sobre essa capacidade.

## Agradecimentos

A todos os encontros. Inclusive àqueles do meio do caminho.

E àqueles que permanecem, pois por mais despedidas que tenha havido na curva extrema, não desaparecem no caminho extremo: Eleonora Albano, Beatriz Ilari, Didier Demolin, Marcos Nogueira, Márcia Duarte, Fred Cummins, João Paulo Cabral, Alex Meireles, Antonio Carrasqueira, Olga Coelho e Claudia Wanderley.

Aos companheiros de sempre, ombreados, nos debates, nas assembleias, nos atos, nas passeatas, nas greves, vivenciando os diversos espaços da USP e das ruas, a pé, de carro ou de “bicicreta”.

Dentre esses companheiros, alguns são das saídas de bicicleta pela cidade, desde 2014, no grupo Bicicreta: Alexandre Bebiano, Adrián Fanjul, Daniel Puglia, Elisabetta Santoro, Elisabeth Lima, Helder Garmes, Kimi Tomisaki, Marta Marques, Pierluigi Benevieri, Rosângela Sarteschi, Tércio Redondo, Valéria de Marcos e Vima Lia de Rossi.

Aos colegas que generosamente disponibilizaram seus memoriais para mim: Adrián Fanjul, Luciana Storto e Ana Lúcia Müller.

À FAPESP pelo apoio valioso em vários projetos.

A tanta gente agradeço, com quem convivi e trabalhei nas diversas instâncias da FFCLH. Pessoas admiráveis, cuja opção é a liberdade de pensamento, respeito ao outro e discussão incansável. De suas falas sempre aprendi lições.

Aprendi lições também com colegas de outras unidades, os quais conheci na maioria das vezes, nas assembleias da Adusp e aos quais externo minha gratidão.

Aos vizinhos do Maracá, edifício que é minha morada, no Butantã.

Às amigas afetuosas de Mariana Dordetti, Mayumi Ilari e Alessandro Penezzi.

E com afeto desmedido à Catarina Raposo, a filha linda, a melhor que uma mãe pode ter.



## Referências

- Albano, E. Fonologia de Laboratório. In: Hora, D., Matzenauer, C. L. Fonologia, fonologias. Uma introdução. São Paulo: Editora Contexto, 2017. pp. 169 – 181.
- Alves, C. Effects of metaphors in the teaching of singing: acoustic and vocal tract image data. 2013. 166f. Dissertação (Mestrado em Lingüística) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2013.
- Arom, S. *African Polyphony and Polyrhythm: Musical Structure and Methodology*. Cambridge: Cambridge University Press, 1994.
- Auroux, S. La raison, le langage et les normes. Paris: Presses Universitaires de France, 1998.
- Beraldo, D. Feedback cognitivo aplicado a voz cantada no repertório musical brasileiro. 2015. 111f. Dissertação (Mestrado) - Universidade Federal do Paraná, Programa de Pós-Graduação em Música. Curitiba, 2015.
- Bloomfield, L. Um conjunto de postulados para a ciência da linguagem (1926). In: DASCAL, M. Concepções gerais da teoria linguística. Volume 1. São Paulo: Global Universitária, 1978. pp 45-60.
- Botelho, T. e Botelho, E. *Metamorfoses*. Brasília, 2010.
- Carrasqueira, A. *Divertimentos-Descobertas: Estudos criativos para o desenvolvimento musical – sopros e cordas friccionadas*. São Paulo: Editora da Universidade de São Paul, 2017.
- Chomsky, N. Modelos explanatórios em Linguística (1962). In: DASCAL, M. Concepções gerais da teoria linguística. Volume 1. São Paulo: Global Universitária, 1978. p 61-95.
- Christmann, J. Processos fonológicos em fronteiras de palavras no Canto Erudito em Alemão. 2012. 152 f. Dissertação (Mestrado em Língua e Literatura Alemã). Universidade de São Paulo, 2012.
- Coelho, C. Cantores líricos e de musicais: dados dermatoglíficos e acústicos. 2017. 133 f. Tese (Doutorado em Linguística Aplicada e Estudos da Linguagem) - Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem, Pontifícia Universidade Católica de São Paulo, 2017.
- Cummins, F. Rhythm as entrainment. The case of synchronous speech. *Journal of Phonetics*, 37, 2009, pp. 16-28.

- Cummins, F. *The ground from which we speak: joint speech and collective subject*. Cambridge Scholars Publishing, 2018.
- Dascal, Marcelo. *Concepções gerais da teoria linguística*. Volume 1. São Paulo: Global Universitária, 1978.
- De Looze, C., Scherer, S., Vaughan B. & Campbell, N. Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Communication*, 58, 2014, pp. 1-34. doi.org/10.1016/j.specom.2013.10.002
- Deutsch, D., Henthorn, T. and Lapidis, R. Illusory transformation from speech to song. *Journal of the Acoustic Society of America*, 129, no. 4, 2011, pp. 2245-2252.
- Drumond, L. B., Vieira, N. B., Ferreira de Oliveira, D. S. Produção fonoaudiológica sobre voz no canto popular. *Jornal da Sociedade Brasileira de Fonoaudiologia*. 23(4), 2011, pp. 390-7.
- Escamez, Na. Cantoras eruditas e populares: comparação de características vocais na canção *Melodia Sentimental* de Villa-Lobos. 2015. 75 f. Dissertação (Mestrado em Fonoaudiologia) - Pontifícia Universidade Católica de São Paulo, 2015.
- Esling, J. H. "Articulatory Function of the Larynx and the Origins of Speech". Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society, 2014, pp. 121-150.
- Fant, G. *Acoustic theory of speech production*, The Hague: Mouton, 1960.
- Fitch, W. T. The evolution of speech: a comparative review. *Trends in Cognitive Science*, 4, 2000, pp. 258-267.
- Fowler, C. Coarticulation and theories of extrinsic time, *Journal of Phonetics*, 8, 1980, pp. 113-133.
- Fujisaki, H. Dynamic characteristics of voice fundamental frequency in speech and singing – Acoustical analysis and physiological interpretations. Proceedings of the Fourth F.A.S.E, Symposium on Acoustics and Speech, 2, 1981, pp. 57-70.
- Gentillucci, M. e Corballis, M. C. From manual gesture to speech: A gradual transition. *Neuroscience and Behavioral Reviews* 30(7), 2006, 949-960. doi: 10.1016/j.neubiorev.2006.02.004.
- Gentilucci, M., Santunione, P., Roy, A.C., Stefanini, S. Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *European Journal of Neuroscience* 19(1), 2004, pp. 190-202. doi: 10.1111/j.1460-9568.2004.03104.x.

- Gusmão, C. O formante do cantor e os ajustes laríngeos e supralaríngeos em cantores barítonos: uma investigação acústica e fibronasolaringoscópica. 2014. 97 f. Dissertação (Mestrado) Departamento de Música, Universidade Federal de Minas Gerais, 2014.
- Granger, G. A ciência e as ciências; tradução de Roberto Leal Ferreira. São Paulo Fundação Editora da UNESP, 1994.
- Hauser, M. D., Chomsky, N., Fitch, W.T. The faculty of language: what is it, who has it and how did it evolve. *Science* 22, vol. 298, issue 5598, 2002, pp. 1569-1579. doi: 10.1126/science.298.5598.1569
- Hawkins, S. Reevaluating assumptions about speech perception: interactive and integrative theories. In: PICKETT, J.M., The acoustics of speech communication. Boston: Allyn and Bacon, 1999. pp 232 - 288.
- Holloway, R. L. Human paleontological evidence relevant to language behavior. *Human Neurobiology*. 2 (3), 1983, pp. 105-14.
- Ilari, B. S. *Em busca da mente musical*. Curitiba: Editora da UFPR, 2006.
- Kennedy, M. *Dicionário Oxford de Música* (Gabriela G. da Cruz e Rui V. Nery, Trad). Lisboa: Publicações Dom Quixote, 1994.
- Kang, J., Scholp, A. Jiang, J.J. A Review of the Physiological Effects and Mechanisms of Singing. *Journal of Voice*, 2017. doi: 10.1016/j.jvoice.2017.07.008
- Ladefoged, P. Preliminaries to linguistics phonetics, Chicago, 1971.
- Laitman, J. e Reidenberg, J. Advances in understanding the relationship between the skull base and larynx with comments on the origins of speech. *Human Evolution*, v. 3, 1988, pp. 99-109.
- Laver, J. The semiotic nature of phonetic data, *York Papers in Linguistics* 6, p. 55- 62, 1976.
- Lerdahl, F. and Jackendoff, R. *Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press, 1983.
- Loiola, C. Canto popular e erudito: características vocais, ajustes do trato vocal e desempenho profissional. 2013. 108 f. Tese (Doutorado em Fonoaudiologia) - Pontifícia Universidade Católica de São Paulo, 2013.
- Lomba, J. Um estudo sobre as palavras cantada, falada e declamada e seus efeitos impressivos. 2017. 106 f. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) - Programa de Estudos Pós-Graduados em Linguística

- Aplicada e Estudos da Linguagem, Pontifícia Universidade Católica de São Paulo, 2017.
- Machado, A. Diferenças na percepção de emoções em enunciados verbais declamados e cantados. 2017. 109 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) Universidade Estadual Paulista, Araraquara, 2017.
- MacDougall, H. G. e Moore, S. T. Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *Journal of Applied Physiology*, 2005. doi: 101152/jappphysiol.00138.2005.
- Mehr, S. A., Singh, M., Knox, D., Ketter, D.M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A.A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., , Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. N., Glowacki, L. Universality and diversity in human song. *Science* 366, eaax0868. Downloaded from <http://science.sciencemag.org/> on November 28, 2019.
- Menegon, P. Metáforas no ensino de canto e seus efeitos na qualidade vocal: um estudo acústico-perceptivo. 2013. 91 f. Dissertação (Mestrado em Linguística) - Pontifícia Universidade Católica de São Paulo, 2013.
- Migliorini, L. De versos e trovas: análise de aspectos fonostilísticos do português medieval por meio das “Cantigas de Santa Maria. 2012. 206 f. Tese (Doutorado) - Universidade Estadual Paulista, Araraquara, 2012.
- Mirra, E. *A ciência que sonha e o verso que investiga: ensaios sobre inovação, poesia, tecnologia e futebol*. São Paulo: Papagaio, 2009.
- Morley, I. A multi-disciplinary approach to the origins of music: Perspectives from anthropology, archaeology, cognition and behavior. *Journal of Anthropological Sciences*, 92, 2014, pp. 147-177. doi: 10.4436/JASS.92008
- Nef, C. *Histoires de la musique*. Paris: Payot, 1948.
- Nuñez, K. Configurações do trato vocal na fonação habitual e no canto popular. 2006. 37 f. Dissertação (Mestrado em Linguagem e audição: modelos fonoaudiológicos) - Universidade Tuiuti do Paraná, 2006.
- Ohala, J. The contribution of acoustic phonetics to phonology in LINDBLOM, Bjorn, OHMAN, S. *Frontiers of speech communication research*. London: Academic Press, 1979, p. 356-363

- Oliveira, S. A voz de Roberto Carlos: avaliação perceptivo-auditiva, análise acústica e a opinião do público. 2007. 202 f. Dissertação (Mestrado em Fonoaudiologia) - Pontifícia Universidade Católica de São Paulo, 2007.
- Patel, A. *Music, language and the brain*. Oxford University Press, 2010.
- Pessotti, A. Efeitos do treinamento e da prática vocal profissional sobre o canto e a fala. 2012. 175 p. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, 2012.
- Peretz, I. The nature of music from a biological perspective. *Cognition*, 100, 2006, pp. 1-32. doi:10.1016/j.cognition.2005.11.004
- Port, R. All is prosody: Phones and phonemes are the ghosts of letters. Proceedings of the 4th International Conference on Speech Prosody, 2008. [https://www.isca-speech.org/archive/sp2008/papers/sp08\\_007.pdf](https://www.isca-speech.org/archive/sp2008/papers/sp08_007.pdf)
- Raposo de Medeiros, B. Estudo preliminar das vogais cantadas no português brasileiro. *Estudos Linguísticos XXIX*, 2000, pp. 657-662.
- Raposo de Medeiros, B. A questão da relação entre fala e canto e a coordenação dos gestos. Congresso Internacional da Abralín, Fortaleza. Boletim da Associação Brasileira de Linguística, v. 26, 2001, pp. 666-668.
- Raposo de Medeiros, B. *Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro*. Tese de Doutorado, Universidade Estadual de Campinas. Campinas, Brasil, 2002.
- Raposo de Medeiros, B. Estudo perceptual de vogais faladas e cantadas em português brasileiro. Intercâmbio de Pesquisas em Linguística Aplicada, São Paulo. 13º. InPLA, Caderno de Resumos, 2003, p. 209.
- Raposo de Medeiros, B. Aspectos fonético-acústicos da canção erudita brasileira. *Síntese: Revista dos cursos de Pós-Graduação*, 8, 2003, pp. 217-229.
- Raposo de Medeiros, B. O português brasileiro e a pronúncia do canto erudito: reflexões preliminares. *ARTEunesp*, v. 1, 16, 2003, pp. 11-240.
- Raposo de Medeiros, B. Formant and harmonic matching in Brazilian lied. 147th Meeting Acoustical Society of America. *The Journal of the Acoustical Society of America*. Danvers MA: Acoustical Society of America, v. 115, 2004, pp. 2633-2633. doi.org/10.1121/1.4784874

- Raposo de Medeiros, B. Estudo articulatório via filmagem dos movimentos da face. XIX Encontro Nacional da Anpoll, Maceió AL. Boletim Informativo da Anpoll, v. 32, 2004, p. 168.
- Raposo de Medeiros, B. Uma visão sobre cognição musical a partir de estudos interdisciplinares na linguística e na música. *Música em Contexto (UnB)*, v. 1, 2004, pp. 43-75.
- Raposo de Medeiros, B. Uma correlação entre padrões articulatórios e acústicos da fala e do canto. 1o. Simpósio Internacional de Cognição e Artes Musicais, Curitiba. Anais do 1o. Simpósio Internacional de Cognição e Artes Musicais, v. I, 2005, p. 287.
- Raposo de Medeiros, B. Um panorama de estudos comparativos entre fala e canto. I Encontro Nacional de Cognição e Artes Musicais, Curitiba. Anais do I Encontro Nacional de Cognição e Artes Musicais. Curitiba: Editora do Departamento de Artes da Universidade Federal do Paraná, v. 1, 2006, pp. 93-99.
- Raposo de Medeiros, B. Em busca do som perdido: o que há entre a Linguística e a Música. In B. S. Ilari. (Org.). *Em busca da mente musical*. 1ed. Curitiba: Editora da UFPR, 2006.
- Raposo de Medeiros, B. Ritmo na língua e na música: o elo possível. *Música em Perspectiva*, v. 2, 2009, pp. 45-63.
- Raposo de Medeiros, B. Os Rumos da Pesquisa em Cognição Musical - Ainda: o que há entre a Música e a Linguística? Conferência proferida, Simpósio Internacional de Cognição e Artes Musicais (V SIMCAM), Universidade Federal de Goiás, 2009.
- Raposo de Medeiros, B. Cognição musical e suas fronteiras: por uma fronteira livre entre Música e Linguística. Conferência proferida, VI SIMCAM, Universidade Federal do Rio de Janeiro, 2010.
- Raposo de Medeiros, B. Towards a musical gesture in the perspective of music as a dynamical system. Proceedings of the 12th International Conference on Music Perception and Cognition, 1, 2012, pp. 843-847. Thessaloniki.
- Raposo de Medeiros, B., Cummins, F.; Oliveira, L. F. Cantando junto: por uma unidade do conceito de ritmo para fala e canto. Mesa-redonda, IX SIMCAM, Universidade Federal do Pará, 2013.
- Raposo de Medeiros, B. & Cummins, F. Speech and song synchronization: a comparative study, Proceedings of Speech Prosody, 2014, pp. 748-75.

- Raposo de Medeiros, B. A organização temporal na fala e no canto: estudo com foco na sincronização. *Percepta*, 1, 2014, pp. 27-50.
- Raposo de Medeiros, B. Aspectos de produção, percepção e cognição na intersecção fala e canto. Relatório científico parcial à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), 2016. Processo 2015/06283-0.
- Raposo de Medeiros, B. A estabilidade instável da fala cantada e a integração fonética/fonologia. Comunicação, II MiniEnapol de Fonética e Fonologia, Universidade de São Paulo, 2017.
- Raposo de Medeiros, B. & Cabral, J. P. A sílaba na fala e no canto e a questão da estabilidade: um estudo preliminar. Simpósio de Cognição e Artes Musicais, 2017, Curitiba. Anais do XIII SIMCAM. Curitiba: ABCM, v. 1, 2017, pp. 69-77.
- Raposo de Medeiros, B. & Cabral, J.P. Acoustic distinctions between speech and singing: Is singing acoustically more stable than speech? *Proceedings of Speech Prosody*, 2018, pp. 542-546. At 10.21437/SpeechProsody.2018-110.
- Raposo de Medeiros, B. Towards a Musical Larynx: Preliminary Proposal. In Antenor Ferreira Corrêa. (Org.). *Music, Speech, and Mind*. 1ed. Curitiba: Association of Cognition and Musical Arts - ABCM, 2019, v. 1, pp. 171-185.
- Raposo de Medeiros, B. Estabilidade de f0: um estudo de percepção. Comunicação, Encontro Anual do Grupo de Pesquisa DINAFON, Universidade Estadual de Campinas, 2019.
- Raposo de Medeiros, B. Singing and Speech as Comparable Phenomena: A Dynamical Approach. In: Frank A. Russo, Beatriz Ilari, and Annabel J. Cohen. (Orgs.). *Singing and Speech as Comparable Phenomena: A Dynamical Approach*. 1ed. Nova Iorque: Routledge Taylor & Francis Group, v. 1, 2020, pp. 97-107.
- Raposo de Medeiros, B., Cabral, J.P., Meireles, A., Baceti, A. A comparative study of fundamental frequency stability between speech and singing. *Speech Communication*, 128, 2021, pp. 15-23. <https://doi.org/10.1016/j.specom.2021.02.003>
- Rebuschat, P.; Rohrmeier M.; Hawkins, J.; Cross, I. *Language and Music as Cognitive Systems*. Oxford: Oxford University Press, 2012.
- Rehder, M. Análise perceptivo-auditiva e acústica da emissão de vogal sustentada falada e cantada de regentes de coral. 2002. 81 f. Tese – (Doutorado em Fonoaudiologia) Unifesp, São Paulo, 2002.

- Ricci, G. Samba e bossa nova: um estudo de aspectos da relação texto-música. 2015. 119 f. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, 2015.
- Rocha, J. Contribuições da fonética no processo ensino aprendizagem da pronúncia de línguas no canto. 2013. 327 f. Dissertação (Mestrado em Linguística, Letras e Artes) - Universidade Federal de Uberlândia, 2013.
- Rostas, M. Balizas suprasegmentais para a adaptação do reggae cantado em São Luís. 2010. 212 f. Tese (doutorado) - Universidade Estadual Paulista, Araraquara, 2010.
- Sachs, C. *Rhythm and tempo: A study in music history*. New York: Columbia University Press, 1953.
- Salomão, G. Registros vocais no canto: aspectos perceptivos, acústicos, aerodinâmicos e fisiológicos da voz modal e da voz de falsete. 2008. 196 f. Tese (Doutorado em Linguística) - Pontifícia Universidade Católica de São Paulo, 2008.
- Santos, C. A. A sincronização em dois ritmos da canção: uma observação experimental acerca da fala cantada. Dissertação inédita, Universidade de São Paulo, 2012.
- Santos, C. A., Raposo de Medeiros, B., Pessotti, A. Para medir a sincronização de dois cantores: o caso da bossa-nova. VI Simpósio Internacional de Cognição e Artes Musicais, Brasília. Anais do VII SIMCAM, 2011, pp. 143-148.
- Santos, C. A., Raposo de Medeiros, B. Cantar junto em dos tipos de ritmo: um estudo preliminar. *Percepta*, 3(1), 2015, pp. 17-37.
- Santos, F. C. *O canto em Linda Wise: ação imaginativa e interpretação*. Dissertação de Mestrado Pontifícia Universidade Católica de São Paulo, São Paulo, Brasil, 2019.
- Silva, M. A. de A., Ghirardi, A. C. de A. M., Bittencourt, M. F. P., Assanti, L. A voz cantada, 2013. [http://www.sbfal.org.br/portal/voz\\_profissional2013/](http://www.sbfal.org.br/portal/voz_profissional2013/)
- Silva, V. Características acústicas do idioma alemão na interpretação de cantoras líricas brasileiras. 2014. 84 f. Dissertação. Mestrado em Fonoaudiologia. Pontifícia Universidade Católica de São Paulo, 2014.
- Siqueira, V. P. & Raposo de Medeiros, B. Por que falamos em conjunto? E o que falar em conjunto diz sobre o comportamento humano? Campinas: Blogs de Ciência, 2018.
- Siqueira, V. P. & Raposo de Medeiros, B. Proposta metodológica para o estudo da sincronização da fala em interação com a semântica. *GRADUS - Revista Brasileira de Fonologia de Laboratório*, v. 5.1, 2020, pp. 99-124.



- Studdert-Kennedy, M. & Goldstein, L. (2003). Launching Language: The Gestural Origin of Discrete Infinity. In Morten H. Christiansen and Simon Kirby (Orgs.), *Language evolution*. Oxford: Oxford University Press, 2003.
- Sousa, N. Escolas de canto italiana, alemã e francesa: características perceptivo-auditivas e acústicas na voz do soprano. 2015. 104 f. Tese (Doutorado em Fonoaudiologia) - Pontifícia Universidade Católica de São Paulo.
- Sundberg, J. The acoustics of the singing voice, *Scientific American*, Vol 236, 3, 1977. pp. 82-91.
- Tatit, L. *O século da canção*. Cotia: Ateliê Editorial, 2004.
- Tavano, K. Análise acústica e imagem da emissão de vogais em cantoras líricas e populares. 2011. 98 f. Dissertação (Mestrado em Linguagem e audição: modelos fonoaudiológicos) - Universidade Tuiuti do Paraná, Curitiba, 2011.
- Trager, G. Paralanguage: a first approximation. *Studies in Linguistics*, 13. p 1-12. 1958.
- Waldow, M. Respiratory strategies and their effects on voice quality in singing: an acoustic and perceptual study. 2015. 111 f. Dissertação (Mestrado em Linguística) - Pontifícia Universidade Católica de São Paulo, 2015.
- Wind, J., Jonker, A., Allott, R., Rolfe, L. *Studies in Language Origins*, volume 3. John Benjamins Publishing Company, 1994.

## Sobre a Autora

por Olga Ferreira Coelho Sansoni<sup>86</sup>



Beatriz Raposo de Medeiros estuda a fala humana já há algumas décadas. Tem se interessado pelas condições nada ordinárias de produção da fala, examinando-as em suas nuances físicas e cognitivas. Também se interessa pelas condições extraordinárias de emprego da voz, que desde sempre envolvem a arte do canto, e, mais recentemente, também a voz afetada por enfermidades. Beatriz, cantora e linguista, transita pelas intrigantes relações entre cantar e falar, e pelos saberes que permitem, por exemplo, colocar o conhecimento fonético a serviço de estudos na interface com a área médica,

buscando biomarcadores em aspectos da fala. Professora associada da USP, desenvolveu uma série de projetos, orientações e publicações cuja tônica é compreender, por um lado, como fatores linguísticos interagem com a música, e, por outro lado, como uma teoria fonológica dinâmica permite lidar com fenômenos fonéticos e fonológicos. Seu estudo de livre-docência juntou a esses persistentes interesses (que ela sintetizou como um caminho “da acústica à cognição”) o desejo de tornar mais palatáveis esses saberes acumulados. Esta versão em livro do trabalho cumpre essa missão.

São Paulo, março de 2024.

---

<sup>86</sup> Linguista, especialista em Historiografia Linguística, professora doutora da Universidade de São Paulo.