UNIVERSIDADE DE SÃO PAULO

FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS

ANDRÉ RICARDO DE SOUZA

A intenção na melodia: estudo interdisciplinar sobre as relações entre entoação e gênero de discurso nas manifestações vocais (da fala ao canto)

São Paulo

2014

UNIVERSIDADE DE SÃO PAULO

ANDRÉ RICARDO DE SOUZA

A intenção na melodia: estudo interdisciplinar sobre as relações entre entoação e gênero de discurso nas manifestações vocais (da fala ao canto)

Tese apresentada à Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo para obtenção do título de Doutor em Letras

Área de Concentração: Filologia e Língua Portuguesa

Orientador: Prof. Dr. Waldemar Ferreira Netto

São Paulo

2014

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

S731i

SOUZA, André Ricardo de
A intenção na melodia: estudo interdisciplinar sobre as relações entre entoação e gênero de discurso nas manifestações vocais (da fala ao canto) / André Ricardo de SOUZA ; orientador Waldemar FERREIRA NETTO. - São Paulo, 2014.
116 f.

Tese (Doutorado)- Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. Departamento de Letras Clássicas e Vernáculas. Área de concentração: Filologia e Língua Portuguesa.

1. entoação. 2. intenção. 3. gênero de discurso. 4. melodia. 5. fala espontânea. I. FERREIRA NETTO, Waldemar, orient. II. Título.

Nome: SOUZA, André Ricardo de
Título: A intenção na melodia: estudo interdisciplinar sobre as relações entre entoação e gênero de discurso nas manifestações vocais (da fala ao canto)

Tese apresentada à Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo para obtenção do título de Doutor em Letras

Aprovado em:

Banca Examinadora

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

A toda minha família – pais, tios e avós
com quem aprendi a falar cantando.

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Waldemar Ferreira Netto, pelo constante incentivo ao desenvolvimento das minhas ideias, e pela palavra certa na hora certa, que evitou que eu me perdesse na rede das minha própria argumentação; sua preciosa orientação fez com que o resultado desse curso fosse muito mais que uma tese – foi uma verdadeira e profunda mudança na maneira de pensar e de me relacionar com o conhecimento.

À Profª. Drª. Beatriz Raposo de Medeiros, que primeiro acolheu meu projeto de pesquisa e me apresentou ao prof. Waldemar.

Aos membros da banca de qualificação, Prof. Dr. Antonio Vicente Seraphim Pietroforte e Prof. Dr. Manoel Mourivaldo Santiago Almeida, pelas observações, críticas e sugestões que foram de grande ajuda para direcionar a etapa final do trabalho.

Aos colegas de doutorado Daniel de Oliveira Peres e Marcus Vinícius Martins, com quem dividi dúvidas e especulações; e a todos os participantes do grupo ExProsodia que me ouviram e deram palpites.

A minha amiga Renata Pelloso Gelamo, que me chamou a atenção para o paradoxo da fala espontânea (que ela não acredita existir).

Aos meus amigos Dalton Martins e Pedro Beviláqua, que proveram uma interlocução musical que foi essencial para esta pesquisa.

Aos meus colegas da UNESPAR/FAP, que sempre me incentivaram na caminhada, trocando aulas e ideias e, especialmente, às coordenadoras do curso de Licenciatura em Música, professoras Solange Maranho Gomes e Simone do Rocio Cit, que me ajudaram a encontrar, em meio a aulas e compromissos acadêmicos, o tempo necessário para me dedicar ao doutorado.

Aos meus alunos na UNESPAR/FAP, que foram compreensíveis com as minhas ausências.

A Paula e Jorge pelo pouso seguro a qualquer hora ao longo destes quatro anos;
e a Ivani pelos almoços infalíveis.

E, finalmente, mas não por último, a Aninha, que me acompanhou em todo esse percurso e com seu amor me deu força para chegar até aqui.

Nós, os pesquisadores da área do conhecimento, nos desconhecemos mutuamente.
Isso tem seu motivo específico. Nunca nos procuramos,
como haveríamos de nos encontrar algum dia?

Friedrich Nietzsche

Resumo

Este trabalho apresenta uma reflexão acerca da natureza da entoação e de seu papel no funcionamento e no uso da linguagem, numa perspectiva interdisciplinar que traz para a discussão as contribuições de outras áreas do conhecimento que de uma maneira ou de outra se relacionam com a expressão na fala, seja por seus aspectos físicos, acústicos, seja por seus aspectos funcionais. Apresentamos evidências e argumentos que colocam a entoação como um fator primário na comunicação linguística que tem um papel fundamental na organização e transferência da informação de um sujeito a outro, tanto do ponto de vista da produção como da recepção. A tese aqui defendida é a de que a intenção comunicativa do falante – o modo como este usa a linguagem enquanto ação junto a outros seres humanos – é que determina a configuração melódica, precedendo a seleção e combinação dos constituintes do enunciado. Partimos da consideração da situação atual dos estudos prosódicos com respeito à entoação que encontramos descrita em Ladd (1996), Hirst e DiCristo (1998) e Fox (2000), entre outros, e apontamos alguns dos seus principais impasses e as questões epistemológicas relacionadas. Em seguida apresentamos as perspectivas de outras áreas de fora da linguística que direta ou indiretamente se reportam à entoação, seja como intenção, como é o caso da filosofia da linguagem, a sociologia, e a teoria literária, seja como melodia nas fronteiras do fazer teatral e da prática musical. Encontramos aí aproximações significativas entre as noções de estilo e gênero do discurso e a noção de intenção comunicativa que questionam a possibilidade de uma separação estanque entre o uso prático e o uso poético, musical, artístico da linguagem. Essa constatação nos levou a considerar uma grande categoria da atividade humana que denominamos práxis vocal que inclui, além da fala e do canto mais convencionais, formas intermediárias tais como narrativas orais, chamados, pregões, discursos políticos, declamações artísticas e religiosas como o recitativo e o salmodiar, entre outras. Estas formas empregam a modulação da frequência fundamental da voz de maneiras muito particulares e podem ser situadas na fronteira entre a linguagem e a música. Empreendemos, por isso, uma investigação acerca das relações, semelhanças e diferenças entre música e linguagem. Desta aproximação surgiu uma das ideias centrais desta tese, que é a relação entre estilo melódico e gênero de discurso nas manifestações orais. Percebemos que a realização de uma intenção comunicativa, seja numa fala espontânea, num discurso ou numa canção, começa com a escolha do gênero mais adequado à expressão do conteúdo desejado, e a partir dessa escolha é que fica determinado o estilo, que corresponde às regras de elaboração do discurso (linguístico, musical ou híbrido). A violação dessas regras que observamos quando há inconsistências melódicas na fala decorada, por exemplo, são decorrentes da necessidade de se criar uma melodia a partir de um texto já dado, fato que não acontece na fala espontânea.

Palavras-chave: entoação; intenção; gênero do discurso; melodia; fala espontânea

Abstract

This work presents a reflection about the nature of intonation, and about its role in how language works and it is used, within a interdisciplinary perspective that brings contributions from other areas that, in a way or another, relate to the issue of expression in speech, whether by its physical, acoustic features or by its functional aspects. We present evidences that show intonation as a primary factor in linguistic communication, having a fundamental role in organising and transferring information from one subject to another, from the point of view of production as well as perception. The main reasoning is that the communicative intention of the speaker - the way he uses languages as an action toward other human beings - determines the melodic configuration, that preceding selection and combination of the utterance contituents. Considering the recent development of investigation of prosody as it is described in Ladd (1996), Hirst e DiCristo (1998) and Fox (2000), among others, we show some of its hindrances and the epistemological issues related to them. Following that, we present perspectives brought from other areas outside linguistics that in a direct or indirect manner relate to the issues around intonation, whether as the speaker's intention, as we see in the field of philosophy of language, sociology and literary theory, or as a melodic elaboration in theater and musical practices. Within these new perspectives we found out striking similarities between the concepts of style and speech genre and the idea of communicative intention, that dispute a clear distinction of the practical use of language from its poetic, musical, artistic use. This remark have led us to consider a grand category of human activity that we named “vocal praxis”, including, besides ordinary speech and song, intermediate forms as those found in oral narratives, callings, auctioneering, political discourses and artistic declamation such as the recitative and chant. These forms employ the modulation of the fundamental frequency of the voice in very peculiar ways, and they may be situated close to the boundaries of speech and song. For that reason we endeavoured an inquiry on relationships, similarities and differences between music and language, This approximation produced one of the central ideas of this thesis, i.e. the connection of melodic style and speech genre along the continuum of vocal (oral) manifestations. We realized that the actualization of a communicative intention, whether in spontaneous speech, in a political discourse or in a song, begins with the choice of the adequate genre to the desired expression of the content; after this choice the style is determined, leading to the rules of elaboration of the particular discourse (linguistic, musical or hybrid). The violation of these rules, that we observe as melodic incoherences when someone speaks something by heart or reading aloud, are the consequence of an artificial situation that demands creating a new melody to a given text, which doesn't happen in spontaneous speech.

Key words: intonation; intention; speech genre; melody; spontaneous speech

LISTA DE SIGLAS

F0: frequência fundamental da voz

TMA: Teoria Métrica-Autossegmental

TLA: Teoria da Língua em Ato (Cresti)

ST: semitom (unidade de medida de intervalos musicais)

MIR: Music Information Retrieval (recuperação de informações a partir da música)

ASR: Automatic Speech Recognizing (reconhecimento automático de fala)

Sumário

Agradecimentos 4

Resumo 6

Abstract 7

Lista de siglas 8

Introdução 10

1 Questões epistemológicas e metodológicas no estudo da entoação
1.1 Conceitos fundamentais: prosódia, acento, entoação 18
1.2 Teoria métrica-autossegmental 31
1.3 Outros modelos 34
1.4 Impasses no estudo da entoação 35

2 Outros olhares sobre a entoação
2.1 Uma palavra sobre interdisciplinaridade 39
2.2 Os atos de fala 40
2.3 A economia das trocas linguísticas 45
2.4 Estética da criação verbal 47
2.5 Pragmática na criação literária 51
2.6 Gesto musical: ação e significação 53

3 Fala e canto
3.1 Representação musical da entoação 57
3.2 Música e linguagem 66
3.3 Origens comuns da fala e do canto....69
3.4 Práxis sonora 74
3.5 Aspectos cognitivos da percepção de linguagem e música 76
3.6 Investigação conjunta da fala e do canto 81

4 Contribuições da etnomusicologia
4.1 Etnomusicologia 83
4.2 Fronteiras entre a fala e o canto 84
4.3 Cantometria 87
4.4 Etnomusicologia computacional 92
4.5 Um problema persiste 96

5 Reflexões sobre a natureza da entoação
5.1 Ritual e protocolo 99
5.2 Intenção e gênero 102
5.3 O paradoxo da fala espontânea 103
5.4 A precedência da entoação 106

Conclusão 108

Referências 111

Introdução

O objeto da investigação que resultou nesta tese é a entoação, entendida como a variação da frequência fundamental (doravante F_0) de vibração da voz que acompanha as realizações linguísticas. Trata-se portanto de um trabalho que contempla a fala como campo privilegiado para compreensão do funcionamento e desenvolvimento da linguagem. Entretanto, a entoação ainda ocupa um lugar de certa forma marginalizado dentre os estudos linguísticos. Conforme aponta Anthony Fox (2000, p. 269), praticamente todos os autores que se propõem a tratar da entoação começam seus textos afirmando a grande dificuldade em se lidar com o tema, e na maioria das descrições fonológicas de línguas individuais o tema, quando mencionado, é relegado às seções finais ou apêndices.

Não é exagero afirmar que a entoação é o elemento mais evasivo da linguagem – e possivelmente por isso o menos estudado até hoje. Somente a partir do final do século XIX aparecem tentativas de sistematizar o conhecimento sobre a variação de altura, em parte incentivados pela possibilidade de se gravar o som, trazida pela invenção do fonógrafo. Mesmo assim, o tema continuou sendo periférico para a linguística, sendo considerado como algo fora da linguagem, "paralinguístico". O surgimento da fonologia nas primeiras décadas do século XX deu um grande impulso aos estudos linguísticos, e passou a ocupar o centro dos interesses na matéria. Tanto os pressupostos teóricos da fonologia, como o contexto em que esta se desenvolveu – a corrente estruturalista – apresentavam dificuldades para o estudo de formas dinâmicas e continuamente variadas, como é o caso dos contornos melódicos que acompanham os enunciados na língua falada, que resistem às classificações e relações distintivas entre signos pretendidas pela fonologia.

Desta forma, a entoação não foi de pronto abordada com modelos fonológicos. Em parte, segundo Ladd (1996, p. 20), isto também se deveu a uma concepção pré-estabelecida, implícita nos primeiros trabalhos desta época, de que a entoação carreava um significado completo, associado a emoções ou intenções, tais como alegria, medo, escárnio, queixa, e assim por diante. Assim, não haveria motivo para se investigar categorias fonológicas na entoação. Fox (op.cit., p. 275) também aponta que a redução de um conjunto de características

fonéticas para uma única categoria fonológica é uma tarefa árdua no caso da entoação, uma vez que há um gradiente de possibilidades, tanto para a altura como para a taxa de variação da mesma, que são significativas, pois criam nuances que podem ser interpretadas pelo ouvinte não como variantes alofônicas, mas como "significados" diferentes. Poderíamos acrescentar aqui também a constatação – que comentaremos mais adiante – de que o contorno melódico de um enunciado é o resultado da combinação de vários fatores linguísticos e extra-linguísticos, o que torna a realização fonética suscetível a influências não facilmente controláveis em um experimento.

Apesar destas dificuldades, constatamos que o número de pesquisadores e trabalhos na área dos estudos prosódicos nas últimas décadas se multiplicou, e avanços notáveis têm sido alcançados. Porém, ao estudarmos o desenvolvimento recente das principais linhas de investigação da entoação, desde os anos 70, foi possível perceber que há um certo descompasso entre avanço técnico e tecnológico nos instrumentos e metodologias para análise e descrição da curva de F_0, de um lado, e, de outro lado, a reflexão ainda superficial sobre a própria natureza da entoação, e sobre seu papel na comunicação lingüística. Há uma predominância da apresentação de resultados de métodos experimentais em detrimento de uma discussão mais ampla em torno da capacidade de descrição dos múltiplos aspectos do fenômeno que os modelos teóricos têm.

Este trabalho pretende contribuir com uma reflexão acerca da natureza da entoação e de seu papel no funcionamento e no uso da linguagem, numa perspectiva interdisciplinar, trazendo para a discussão as contribuições de outras áreas do conhecimento que de uma maneira ou de outra se relacionam com a expressão na fala, seja por seus aspectos físicos, acústicos, seja por seus aspectos funcionais. Pretende-se aqui fazer um contraponto ao grande número de trabalhos experimentais, não para contestá-los, mas para levantar pontos a serem investigados em experimentos futuros.

Ao longo dos cinco capítulos que compõem esta tese apresentamos evidências e argumentos que colocam a entoação como um fator primário na comunicação linguística, que tem um papel fundamental na organização e transferência da informação de um sujeito a outro, tanto do ponto de vista da produção como da recepção.

A tendência tradicional dos estudos linguísticos é pensar a entoação como algo acessório, que é aplicado *a posteriori* sobre um enunciado pré-concebido. Neste trabalho buscaremos demonstrar que não só a entoação transmite informações que organiza e direciona a interpretação do conteúdo linguístico, mas ela é também o ponto de partida para a construção dos enunciados que proferimos, pelo menos na situação de fala espontânea.

A tese aqui defendida é a de que a intenção comunicativa do falante – o modo como este usa a linguagem enquanto ação junto a outros seres humanos – é que determina a configuração melódica, o fio condutor ao qual serão conectados os fonemas que formarão o enunciado. Esta inversão de perspectiva faz frente à hegemonia da escrita no pensamento ocidental que envolveu a linguística desde o seu nascimento e praticamente estabeleceu a equação língua = escrita, considerando a fala como uma realização imperfeita da escrita. Nosso ponto de vista considera a escrita como uma representação idealizada da fala, que não corresponde à verdadeira natureza da linguagem, uma vez que esta nasceu falada e é por isso dinâmica, imprevisível, multifacetada.

O problema que suscitou esta pesquisa se apresentou na forma de um paradoxo. Quando o ator em cena fala o texto da personagem, ele precisa adequar as inflexões de voz à estrutura sintática do enunciado, ao contexto da cena, à relação com o interlocutor e, claro, às características particulares daquela personagem e a seu estado psicológico naquele momento. Tudo isto precisa ser elaborado para um texto que já está pré-determinado e, embora nem sempre haja a exigência de que a fala pareça natural, é necessária uma verossimilhança na maneira como os enunciados são proferidos. Se o ator em questão aplica uma entoação "errada", algo na cena perde o sentido, e muitas vezes até a assimilação do que está sendo dito fica prejudicada. Ora, se existe a possibilidade de se errar a entoação, então pode-se concluir que deve existir algum conjunto de regras que determinem o comportamento da altura da voz. Porém, até onde pude apurar, não existe nenhum método sistemático para o ator adquirir esta habilidade; tudo é feito empiricamente, por imitação de modelos conhecidos e também pela observação da fala no cotidiano. E mesmo no campo dos estudos prosódicos, não existe ainda uma prescrição de como deve ser a entoação para um determinado enunciado num determinado contexto, apenas diretrizes vagas, muitas ainda por se provar em definitivo. Ou seja, não é possível dizer qual é a entoação correta (na verdade sempre há múltiplas possibilidades), mas é possível dizer quando ela está errada.

Podemos então considerar que há regras implícitas, não enunciadas até o momento, que determinam a boa-formação (uma tradução e adaptação do conceito de "well-formedness" que encontramos na filosofia da linguagem e também na gramática gerativa) da entoação. Estas regras, assim como as regras fonológicas, são realizadas inconscientemente e com desenvoltura pelo falante na situação de fala espontânea mas, aparentemente, seu funcionamento entra em colapso quando a fala é decorada. Situação semelhante e até mais comumente observada acontece na leitura em voz alta. A observação deste fenômeno

paradoxal levou à elaboração do projeto de pesquisa que resultou no presente trabalho, trazendo, a princípio, as questões para o campo da fonologia prosódica; logo se percebeu, porém, que assim como a natureza do objeto é múltipla, os caminhos até alguma solução também seriam.

Em torno deste aparente paradoxo que motivou a investigação aqui apresentada reuniram-se outras observações do comportamento da entoação na fala cotidiana, no trabalho do ator e também na percepção do fenômeno no senso comum, o que permitiu reunir algumas constatações, listadas a seguir:

- distinguimos facilmente a fala espontânea da fala decorada e da leitura em voz alta; as duas últimas nos parecem artificiais e monótonas, e eventuais contradições entre a entoação e a estrutura sintática dificultam e às vezes até impedem a compreensão do enunciado;
- apesar de ser descrita usualmente como a melodia da fala, é difícil identificar com precisão as alturas correspondentes a cada sílaba, porque nos falta um sistema de referência (tonalidade).
- por outro lado, há ocorrência de certos intervalos ditos musicais nas falas individuais, e também a recorrência de alturas fixas, empregadas sistematicamente em ênfases ou finalizações
- um sotaque diferente do seu é habitualmente percebido como "cantado" por quem ouve mas não por quem fala, o que evidencia que as fórmulas melódicas características de cada falar pertencem ao sistema do dialeto particular;
- a entoação de um enunciado é influenciada por diversos fatores concorrentes, e por isso transporta inúmeras informações que são instantaneamente decodificadas pelo ouvinte;
- a "melodia" da fala também remete a situações específicas de comunicação ou expressão – rituais – que do mesmo modo são prontamente reconhecidas pelo ouvinte pertencente à mesma cultura , tais como salmodiar, o pregão, o discurso político, o canto, o recitativo, etc.
- há evidências de que a entoação desempenha papel importante na aquisição da linguagem;
- pessoas com ascendência estrangeira próxima, mesmo tendo o português como primeira língua, conservam melodias características do idioma de origem dos seus

pais, caso bastante perceptível no falar de descendentes de europeus no Vale do Itajaí (SC).

Se pensarmos no sentido comum, do dicionário, encontramos para o termo entoação duas acepções básicas: a entoação como modulação do tom de voz, controle da altura (F0), como melodia, enfim, e a entoação como intenção, como sinal de atitude do falante e modalidade do enunciado que a variação de altura confere a uma sentença.

Como a ideia de entoação como controle melódico depende de um conhecimento específico musical, é mais frequente, no senso comum, a acepção de intenção, como indica a expressão "não gostei de seu tom de voz". Pode-se perceber este fato também considerando a observação costumeira de que "tal pessoa fala sempre no mesmo tom", o que não pode corresponder obviamente, a uma F0 constante, que, sabemos, é característica de um outro tipo de elocução, o salmodiar. Assim podemos assumir que existem variações de altura que são devidas às restrições da produção da cadeia segmental e que são determinadas pela morfologia e pela sintaxe, mas que também podem existir variações devidas à singularidade do falante (sexo, cultura, etc) e ao estado emocional.

Estamos considerando, então, que ao decodificarmos a entoação em uma enunciação "eliminamos" o "ruído" da cadeia segmental e focamos nas variações que indicam foco, atitudes, modalidades e marcas da conversa, necessárias à condução do discurso. Infelizmente, por mais avançada que esteja a análise automática da prosódia, não existem algoritmos que executem esta tarefa. Este fato está perfeitamente de acordo com a questão teórica, pois essas componentes que denotam atitudes, modalidades, etc, deixam entender algo que ainda não sabemos definir: a intenção. A partir dessas inflexões do falante percebemos que há uma consciência ali tomando decisão, conduzindo o discurso.

A percepção de intenção é um dos quebra-cabeças de varias áreas do conhecimento semiótica, psicologia, filosofia, etc, e não nos atreveremos a se aprofundar nesse campo. No entanto, não podemos deixar de mencionar a proximidade entre intenção e sentido que se pode entrever na palavra "*meaning*" em inglês. "*To mean*" é "querer dizer" , ou, em última instância, "ter a intenção de", como em "*I didn't mean it!*". Pode-se notar que "intenção" e "significado" estão fundidos nessa palavra, que nada mais são que os dois aspectos do signo, o de quem o produz e o de quem o interpreta. Em português a palavra sentido tem aproximadamente o mesmo "duplo sentido" de referência e intenção, embora a etimologia aponte mais para o aspecto perceptual do que produtivo. A pergunta "Qual o sentido disso?"

pode ser interpretada como “Por que você fez isso?”, o que seria equivalente a “Como devo entender isso?”. Não devemos pretender identificar descobrir a intenção do falante, porém do ponto de vista da percepção existem algumas taxonomias que nos guiam ao acompanharmos discursos ou participarmos de conversas, que são as modalidades (curiosamente relacionados a “mood”, estado de ânimo) e as atitudes, como veremos no primeiro capítulo.

Estas componentes da entoação somadas às marcas de conversa, são fundamentais para que a linguagem cumpra a sua função primordial, que é agir sobre o mundo e interagir com os outros seres humanos. São elas que ditam o modo como o conteúdos das cadeias segmentais será interpretado, sem falar no poder de transformar a realidade que pode ser investido em um enunciado.

No caso da acepção melódica da entoação, o sentido técnico musical do termo se refere a controlar com precisão as alturas que a voz percorre. Em outros idiomas, como o inglês e o francês, *intonation* pode ser traduzido para o português como “afinação” e vale não só para a voz como também para os instrumentos musicais.

Além de trazer o foco para a variação de F0, esta segunda acepção permite estabelecer ligação com outras manifestações vocais não consideradas como fala, apesar de terem conteúdo linguístico, e que serão importantes para o desenvolvimento da argumentação em torno de nossa hipótese principal. Estão incluídas aí as diferentes formas tradicionais de narração, chamados, pregões, discursos políticos, declamações artísticas e religiosas como o recitativo e o salmodiar, entre outras.

Estas manifestações vocais mais elaboradas melodicamente não são frequentemente estudadas, por se encontrarem na fronteira entre a linguagem e a música. A etnomusicologia tem procurado estudar estas manifestações no contexto de cada cultura, principalmente porque coloca em questão a concepção eurocêntrica do que é música. Neste ambiente encontramos a polêmica discussão acerca do que é gênero e estilo em música, que se assemelha em muitos pontos com a discussão sobre os mesmo termos na teoria literária. Desta aproximação surgiu uma das ideias centrais desta tese, que é a relação entre estilo melódico e gênero de discurso nas manifestações orais. Percebemos que a realização de uma intenção comunicativa, seja numa fala espontânea, num discurso ou numa canção, começa com a escolha/definição do gênero mais adequado à expressão do conteúdo desejado, e a partir dessa escolha é que fica determinado o estilo, que corresponde às regras de elaboração do discurso (linguístico, musical ou híbrido).

Esta homologia entre gênero musical e gênero de discurso permitiu esclarecer o papel da melodia no processo comunicativo das manifestações orais, que consideraremos na

perspectiva de uma categoria das ações humanas que chamamos de práxis vocal.

A tese está dividida em cinco capítulos. O primeiro expõe brevemente a situação atual dos estudos prosódicos com respeito à entoação e seus principais impasses e questões metodológicas. O segundo capítulo coloca a perspectiva interdisciplinar, apresentando outras áreas do conhecimento que se interessam pela entoação, como a filosofia da linguagem, a teoria literária, os estudos da performance teatral e a música. O terceiro capítulo se propõe a analisar as semelhanças entre a fala e o canto, e propõe um modelo abrangente de classificação para as manifestações vocais por meio da categoria da práxis vocal. O quarto capítulo desenvolve a classificação proposta tomando como modelo contribuições da etnomusicologia, e também propõe modelos de análise automática da variação de F0 inspirado nesses teóricos. Finalmente, o quinto capítulo revisita as questões e impasses levantados no primeiro, reinterpretando as fórmulas melódicas como protocolos de comunicação na fala e no canto.

Antes de prosseguir, uma palavra sobre a gênese deste trabalho. Esta pesquisa é fruto de uma inquietação que se tornou depois uma ideia fixa. A inquietação surgiu de um problema prático, qual seja, tornar verossímeis as falas decoradas dos atores. À época em que concluía a dissertação de mestrado (SOUZA, 2004), trabalhei no treinamento dos diálogos que os cantores de uma opereta tinham de fazer entre os números musicais; um ano depois, assumi a disciplina Expressão Vocal no curso de Bacharelado em Interpretação Teatral, no Departamento de Artes da Universidade Regional de Blumenau, agora trabalhando com atores, portanto. Contava com minha experiência como cantor e professor de canto, com meus conhecimentos teóricos de regente e alguma noção de linguística, área que sempre me interessou, desde os primeiros anos da graduação.

Minha dissertação, na área de teoria e análise musical, versava sobre o conceito de gesto musical, e chegava a tangenciar o problema da entoação, interpretando-a como um tipo de gesto musical pertencente à linguagem. Na conclusão do texto, indicava como uma das possíveis aplicações dos resultados da pesquisa o estudo sistemático da entoação com métodos e modelos da teoria musical contemporânea.

> "Uma primeira perspectiva seria o estudo da entonação (*sic*) lingüística interpretada como estrutura musical segundo o modelo schaefferiano, modelo, aliás, que já apresenta uma componente lingüística; este estudo pode partir de uma conciliação entre o modelo teórico da lingüística para o fenômeno da entonação e suas possíveis contrapartidas nas estruturas e objetos musicais, e posteriormente evoluir para análise da fala em gravações digitalizadas, valendo-se da parceria entre o know-how da música eletroacústica na investigação das estruturas internas do som e o conhecimento sistemático da fala já existente na lingüística." (SOUZA, 2004, p.

157)

Entretanto, esse conhecimento teórico não ajudou a resolver o problema prático. Além disso, a literatura específica sobre o treinamento do ator e sobre o fazer teatral raras vezes se detém no problema da entoação dramática uma vez que no teatro contemporâneo o maior interesse está no treinamento corporal. Alguns manuais práticos (NUNES, 1979; FO, 2000) chegam a dar algumas indicações e sugestões de exercícios, porém não aprofundam o assunto.

Assim, os cinco anos em que ministrei a disciplina foram um período de testes, experiências e exercícios, em que buscava desenvolver estratégias para tornar a fala decorada semelhante à fala espontânea. Os exercícios eram feitos individualmente pelos alunos e observados pela classe, que julgava se a fala do ator soara espontânea ou não, bem como se a intenção da personagem era verossímil com a entoação empregada. Foi possível observar que muitas vezes pequenas variações de altura – da ordem de um semitom – em algumas sílabas (em especial a última sílaba tônica) provocam significativas na percepção do que seria a intenção da personagem. Ao mesmo tempo, era impossível reconhecer algum padrão que permanecesse de aluno para aluno; parecia que as soluções eram muito individuais.

Estes experimentos informais, apesar de bastante limitados quanto à sua reprodutibilidade – pela própria natureza do exercício teatral – fizeram-me vislumbrar a possibilidade de um estudo sistemático a partir de amostras gravadas de fala espontânea, a ser desenvolvido no âmbito da fonologia, em especial a área dos estudos prosódicos.

A inquietação cresceu e se tornou assim uma ideia fixa, que me trouxe ao programa de pós-graduação em Filologia e Língua Portuguesa, para desenvolver esse estudo no curso de doutorado. Temos, então, um músico pesquisando na área de Letras para resolver um problema do fazer teatral. Os resultados, porém contribuem tanto para a compreensão da língua falada como de aspectos da música vocal, como se verá.

Agora, depois de quatro anos de estudos, leituras, reflexões e discussões com pesquisadores da área, tenho uma visão bem mais realista do estado atual dos estudos prosódicos, o que acabou resultando numa proposta de tese baseada menos em experimentos e mais na observação e reflexão; mais filosófica, portanto.

1. Questões epistemológicas acerca do estudo da entoação

1. 1. Conceitos fundamentais: prosódia, acento, entoação

Prosódia

Na sua acepção original, no grego clássico, prosódia (προσωδια) se referia à relação entre canto e fala, a fala cantada, ou o canto aplicado à fala, o canto (ode) da fala (prosa), a ode da prosa (FOX, 2000, p. 7). Ao que tudo indica o termo era empregado especificamente à situação da declamação poética, fosse lírica, épica ou dramática, cada uma com seu estilo específico, e não para a percepção que temos dos movimentos melódicos da fala cotidiana. Entretanto, é impossível saber (ou, pelo menos, temerário supor) como os gramáticos antigos percebiam aquilo que hoje chamamos de prosódia.

Ao nos transmitir a cultura clássica, a tradição escolástica nos legou a concepção de prosódia como algo ligado à composição poética, e até hoje as gramáticas escolares tratam do assunto, de passagem, nos capítulos dedicados à versificação, com um caráter normativo, próximo ao da ortoepia . Já para os músicos da tradição ocidental (europeia), "prosódia" sintetiza o conjunto de problemas e estratégias para sua solução relacionados à acomodação do texto poético à melodia, e é um dos requisitos para a composição da música vocal.

Na linguística contemporânea, que se desenvolveu notavelmente no século XX, o termo "prosódia" passou a reunir os chamados traços supra-segmentais, ou seja, que se estendem para além do domínio do segmento, relacionados com a sílaba, o morfema, ou até com o enunciado como um todo. Na primeira metade do século XX, o assunto despertou o interesse de uns poucos estudiosos, em parte pela limitação dos métodos de análise acústica, mas principalmente pelo predomínio da fonologia na vanguarda da disciplina a partir dos anos 30. A fonologia, como sabemos, revolucionou a linguística tratando os fonemas como conjuntos de traços distintivos, organizados sob a forma de oposições binárias entre pares mínimos. O foco era, portanto, o segmento e suas relações sintagmáticas e paradigmáticas, seguindo o caminho aberto por Saussure. As línguas passaram a ser descritas pelas regras fonológicas que preveem como a organização abstrata dos traços distintivos dá origem à superfície acústica dos segmentos postos em sequência. Os fatos acústicos – traços fonéticos – que não podiam ser explicados pelas regras fonológicas e ultrapassavam o domínio do

segmento eram então atribuídos à prosódia, esse campo ainda misterioso.

Este relativo ostracismo da prosódia começou a ser revertido a partir dos anos 60, com o desenvolvimento de diversas abordagens com diferentes pressupostos e métodos, em grande parte impulsionados pelo desenvolvimento da tecnologia de processamento e análise do sinal acústico. Porém, a grande dificuldade para o avanço do nosso conhecimento sobre a prosódia continuou sendo a própria definição do que seriam os traços prosódicos e a delimitação da sua manifestação acústica. Fox (idem, ibidem) questiona a associação entre prosódia e traços supra-segmentais, baseado na constatação de que existem características fonéticas de um segmento que "contaminam" os adjacentes, ultrapassando, portanto, o domínio do segmento, mas que nem por isso seriam traços prosódicos (o caso da nasalidade, por exemplo). Além disso os traços prosódicos são muito mais refratários à análise puramente fonética do que os traços distintivos dos segmentos, cuja descrição resultará sempre limitada1 (FRY, 1968, apud FOX, 2000, p. 8).

Por outro lado, a descrição fonológica da prosódia é limitada pela dificuldade em se determinar oposições binárias, já que os traços prosódicos têm uma natureza gradiente, frequentemente empregada para expressar nuances de significado ao mesmo tempo em que pode criar distinções entre palavras. Conforme aponta Crystal (1969, apud FOX, 2000, p. 8), "we may define prosodic systems as sets of mutually defining phonological features which have an essentially variable relationship to the words selected, as opposed to those features (…) which have a direct and identifying relationship to such words". Ou seja, ao contrário dos traços distintivos dos fonemas, os traços prosódicos podem variar para uma mesma palavra, "colorindo" seu significado conforme o contexto. E para tornar mais complexo ainda o problema, os correlatos físicos dos traços prosódicos não são univocamente determinados, sendo influenciados pela ordenação dos segmentos, pela interação dos parâmetros (frequência, duração e intensidade), e até por fatores externos, não-linguísticos, como as emoções e características individuais do falante. Apesar disso, a fonologia continua sendo uma referência teórica importante na maioria dos trabalhos significativos que trouxeram grandes avanços nas últimas décadas e fizeram dos estudos prosódicos uma das vanguardas da linguística contemporânea.

Para esclarecer como uma descrição fonológica da prosódia pode ser possível, vejamos o que diz Lehiste (1970, apud FOX, 2000). A autora aponta uma característica específica dos traços prosódicos: ao contrário dos traços distintivos dos fonemas, que se definem por relações paradigmáticas com outros elementos do léxico, os traços prosódicos "are established by a comparison of items in sequence" e são, portanto, sintagmáticos. Esta

visão é corroborada por Ladefoged (1975, apud FOX, 2000) que afirma que "all the suprasegmental features are characterized by the fact that they must be described in relation to other items in the same utterance". Trata-se, portanto, da busca de oposições entre itens sucessivos de um enunciado que explicará como funciona o sistema prosódico de uma certa língua. A descrição fonológica da prosódia tem a natureza de um conjunto de interações de relações sintagmáticas.

Segundo Fox (op. cit., p. 9), o predomínio da fonologia no desenvolvimento da linguística moderna trouxe uma supervalorização do segmento como objeto de estudo, e com isso relegou os traços supra-segmentais ao status de manifestações secundárias e episódicas da cadeia segmental. Nas últimas décadas, porém, as investigações conduzidas com relação à prosódia tem dedicado especial atenção à natureza dos traços e dos constituintes prosódicos. Hoje sabemos que os enunciados proferidos são uma trama de várias dimensões simultâneas, diferentes níveis de organização e uma interdependência estreita dos parâmetros acústicos. Além disso, acumulam-se evidências de que a estrutura prosódica do enunciado é que sustenta a cadeia segmental, ao contrário da visão tradicional de que a prosódia é aposta ao enunciado como uma característica secundária.

Não obstante as conquistas trazidas pela interpretação fonológica dos fenômenos prosódicos, as dificuldades ainda são muitas. Por um lado, os correlatos acústicos dos traços prosódicos sofrem enorme influência do contexto em que o enunciado é proferido, e estão longe de apresentar alguma regularidade. Este fato dificulta tanto o estabelecimento de corpora adequados como o desenvolvimento de métodos que possam isolar um traço em particular e associá-lo a alguma forma de significado. Além disso, é extremamente difícil isolar os correlatos acústicos que têm efetivamente um papel prosódico (fonológico, portanto) daqueles que se devem a efeitos paralinguísticos (emoções, gênero, hierarquia social, etc.).

Por outro lado, a interdependência dos parâmetros acústicos (altura, duração e intensidade) traz sérias limitações ao estudo isolado de cada um deles, no sentido de esclarecer o funcionamento do sistema prosódico de uma língua. Há a necessidade, portanto, de considerar a interação e a hierarquia entre eles. Demais, essa interação e essa hierarquia são características do sistema prosódico (que faz parte do sistema fonológico) de cada língua, e a compreensão da prosódia como um fenômeno linguístico universal requer necessariamente a comparação dos diferentes sistemas prosódicos, tópico que abordaremos mais adiante. No momento nos deteremos sobre a questão fundamental do acento, que subjaz a toda a organização prosódica dos enunciados e, em seguida, buscaremos estabelecer uma definição da entoação a partir da história do seu estudo e do estado atual do conhecimento que temos a

respeito. A seção sobre acento lida com os fenômenos que envolvem duração e intensidade (problemas rítmicos), enquanto a seção sobre entoação inclui comentários sobre o tom lexical (problemas melódicos). Uma definição mais precisa de prosódia, entretanto, deverá esperar o desenvolvimento destas seções.

Acento

O fenômeno do acento já era conhecido pelos gramáticos da Antiguidade e desde então é uma categoria prosódica conhecida (FOX, 2000, p. 115). Assim como o termo "prosódia", sua concepção transformou-se ao longo dos séculos, e ainda é motivo de intenso debate entre os linguistas. Ainda há considerável discordância quanto à natureza fonética do fenômeno e sua função fonológica, e sobre o modo adequado para descrevê-lo, e o termo "acento" é empregado por diferentes autores em acepções válidas, porém mutuamente incompatíveis, envolvendo fenômenos de duração, intensidade e até altura (op. cit., p. 114). Porém, diferentemente do estudo da entoação, a concepção contemporânea de acento se baseia na tradição descritiva dos gramáticos, motivo pelo qual começaremos por aqui.

Uma primeira constatação notável que Fox apresenta (p. 116) é que os gramáticos gregos empregavam termos musicais para caracterizar as sílabas, que poderiam ser agudas (oxys) ou graves (barys), donde se conclui que a altura desempenhava um papel central no acento do grego clássico, característica que se perdeu nos séculos seguintes. Os gramáticos latinos traduziram os termos gregos para os acentos das sílabas (acutum e grave), apesar de o latim não empregar a altura no seu sistema de acentuação, segundo a opinião da maioria dos estudiosos. Segundo Fox, esta confusão terminológica é uma das causas da controvérsia acerca da natureza do acento na tradição dos estudos da linguagem, que, apesar das evidências em contrário, durante bastante tempo considerou o acento como sendo um fenômeno de altura e não de intensidade. Basta considerarmos, por exemplo, a forma atual dos acentos gráficos empregados nas línguas românicas, que têm uma origem nitidamente musical, e que, inclusive, serviram de base para o desenvolvimento da notação musical dita neumática (SOUZA, 2004, p. 91 e ss.).

Na segunda metade do século XIX, os estudos de filologia comparada levam ao reconhecimento e à diferenciação entre acentos dinâmicos (que empregam a intensidade) e acentos musicais (que empregam a altura), bem como à hipótese de que as línguas indo-europeias primitivas tinham acento musical, que gradativamente foi sendo substituído pelo acento dinâmico. Porém, à medida que a fonética se desenvolvia, as características acústicas

dos fenômenos descritos vieram ao plano de frente e, como consequência disso, verificou-se que não havia de fato um correlato acústico consistente para o acento. Apesar de algumas tentativas de incluir alguma consideração física da parte dos filólogos europeus nas primeiras décadas do século XX, separando os fatores fonéticos por meio dos quais a "impressão mental de peso é produzida", a concepção de acento nestes autores está sempre mais próxima da ideia de maneira de falar, a anima vocis, que não é idêntica ao material fonético" (SARAN, 1907, p.19, apud FOX 2000, p. 118).

Outro autor do mesmo período, Schmitt (1924, apud FOX, 2000, p.118) reconhece sentidos ou concepções diferentes de acento: como propriedade de uma sílaba, como propriedade de uma sílaba com relação às outras no contexto da palavra, e como propriedade da relação entre as palavras em uma sentença. Segundo Schmitt, esta terceira concepção é a que mais se aproxima do sentido da tradição filológica para o acento, como maneira de falar, ou a "alma da fala", e que ele propõe que se empregue "acentuação". A primeira estaria relacionada com as línguas que possuem acento musical. Resta, então, como possibilidade de investigação sistemática a segunda concepção de acento, como uma propriedade relacional da sílaba com as outras. Schmitt também repara que, mesmo considerando este sentido "relacional", mais restrito, há ainda o problema em conciliar os aspectos psicológicos do fenômeno dos aspectos físicos.

Fox observa que apesar do desenvolvimento da linguística no século XX ter se distanciado da tradição filológica, estes autores apontam para questões centrais com respeito ao problema do acento que só serão percebidas nas últimas décadas do século passado. Podemos acrescentar que a distinção entre os aspectos físicos e psicológicos do acento ensejou a sua teorização no contexto da fonologia ao longo do século, culminando na grande dicotomia entre produção e percepção que perpassa toda teorização hodierna sobre os fenômenos prosódicos, e sobre os fenômenos da linguagem em geral.

Pudemos verificar que, no início do século XX, tanto a filologia como a recém-criada fonética se afastaram da visão do acento como maneira de falar, bem como da distinção entre acento musical e acento dinâmico. Porém, o impasse entre a descrição psicológica e física do acento só será resolvido com a inclusão da dimensão fonológica, capaz de unificar as diferentes realizações fonéticas sem se basear em uma vaga "impressão mental". A fonologia distingue o acento (abstrato) dos fatores acentuais (acústicos), que podem ser a qualidade (timbre), duração, intensidade, e altura. O aspecto abstrato do acento está relacionado com a função que este desempenha na estruturação dos enunciados, e muitas vezes esta função é realizada cumprida por diferentes correlatos acústicos, independentemente ou combinados.

Assim, a distinção entre acento dinâmico e acento musical não pode ser explicada pelo correlato acústico empregado (se a intensidade ou a altura), uma vez que ambos são combinados como fatores acentuais de acordo com o sistema particular da língua, resultando na proeminência relativa de uma sílaba ou mora em relação às adjacentes, ou de uma palavra em relação aos outros constituintes da mesma frase.

Para os teóricos do Círculo de Praga, esta distinção era irrelevante, uma vez que eram analisadas as funções do acento nos enunciados. Segundo Trubetzkoy, as funções fonológicas podem ser de três tipos: distintivas, culminativas e delimitativas. Considerando-se o acento como uma propriedade da relação de uma sílaba com a sua vizinhança, a ele não pode ser atribuída a função distintiva, devido a sua natureza intrinsecamente sintagmática. Conforme aponta Fox, para uma função distintiva a questão fundamental é "o traço está presente ou ausente?", enquanto que no caso do acento a questão fundamental é "em que posição da cadeia ele aparece"? (op. cit., p. 138).

Restam então, para o acento, as funções culminativas e delimitativas. Trubetzkoy propôs uma classificação das línguas conforme o acento seja fixo ou livre e a organização do acento seja pela sílaba ou pela mora. (resumo) Segundo Trubetzkoy (FOX, 2000, p. 138), todos os acentos têm função culminativa, uma vez que sua manifestação sempre é na forma de uma proeminência, porém só nas línguas de acento fixo este pode ter função delimitativa. Podemos concluir daí que nas línguas de acento livre a função delimitativa fica a cargo do contorno melódico. Jakobson, também observou estes dois aspectos do acento: por um lado aglutina as unidades próximas, unificando os constituintes (delimitativo), por outro estabelece contrastes e hierarquias entre os constituintes (*idem, ibidem*, p. 139). Seguindo a tradição de Praga, Martinet e Garde sustentam que mesmo quando o acento é livre, sua função é contrastiva (sintagmática) e não distintiva (paradigmática).

Há línguas em que a intensidade é o principal correlato do acento, acompanhado pela duração e a altura, enquanto há outras em que a altura é empregada para distinções lexicais, independentemente da intensidade e duração. Há, entretanto, o caso intermediário de línguas em que a altura é empregada para marcar a proeminência relativa de uma sílaba ou mora, sem ser apoiada pela intensidade. São as chamadas línguas de acento tonal, como o sueco, ou de tom acentual, como o japonês. O estudo do comportamento particular da altura nestas línguas abriu novos caminhos para a compreensão do acento.

No caso do japonês, que é uma língua moraica, o acento recai sobre uma mora a cada unidade acentual, sob a forma de uma variação de altura do agudo (na mora acentuada) para o grave (na mora seguinte, se houver), sem ter correlação com intensidade, e obedecendo a

regras de colocação conforme a palavra seja acentuada ou não, e conforme a sua combinação na sentença com outras palavras. As moras que precedem a que é acentuada também têm altura mais aguda, o que evidencia que o que marca o acento não é um pico de altura, mas uma variação descendente.

Beckman (1986, apud FOX, 2000), a partir de um estudo comparativo entre as características fonéticas do acento no japonês e no inglês, pode concluir que no japonês o acento é realizado primordialmente, se não exclusivamente, pela altura, enquanto que no inglês o acento é realizado primordialmente pela intensidade e duração, eventualmente acompanhados pela altura. Baseada nestas conclusões, Beckman mostra que do ponto de vista fonológico o acento nas duas línguas é o mesmo fenômeno, apesar da manifestação fonética ser diferente; o que os torna o mesmo fenômeno é o fato de compartilharem o papel de organização do enunciado. Este resultado resolve as controvérsias em torno da distinção entre acento musical e acento dinâmico, uma vez que, como a altura é um correlato do acento, independentemente de a intensidade e a duração participarem ou não, a diferença entre acento e tom, e entre acento e entoação não têm uma base fonética, mas se trata de uma diferença de função. Como veremos, o tom tem uma função preponderantemente lexical, e a entoação um papel preponderantemente pragmático, ambos submetidos à estrutura rítmica provida pelo acento. Este tema será retomado logo a seguir, quando tratarmos da entoação.

Antes, porém, vamos nos reportar à descrição feita por Ladd (1996, p. 46 e ss.) da natureza fonética do acento, do ponto de vista da teoria métrica autossegmental (TMA), que se estabeleceu como uma das principais tendências no estudo da prosódia nas últimas décadas do século XX, derivando da escola americana que inclui nomes como Pike, Bolinger e Pierrehumbert. Aqui enfrentamos uma certa confusão terminológica que, no entanto, não altera os resultados quanto ao aspecto fonológico do acento. Segundo Ladd, uma das premissas da TMA é a distinção entre pitch accent e stress. Vejamos no original:

> "Pitch accents, in languages that have them, serve as concrete perceptual cues to stress or prominence. However, they are in the first instance intonational features, which are associated with certain syllables in accordance with various principles of prosodic organisation. The perceived prominence of accented syllables is, at least in some languages, a matter of stress, which can be distinguished from pitch accent" (Ladd, 1996, p. 42, grifos do autor).

Podemos perceber que Ladd emprega a expressão pitch accent no sentido proposto por Bolinger a partir da constatação de que a altura era um correlato mais consistente para o acento do que a intensidade (FOX, 2000, p. 124), donde podemos concluir que quando usa o termo "stress", Ladd está considerando a organização rítmica que o acento provê e que permite a estruturação dos termos do enunciado proferido, que Beckman (op. cit.) chama de

“accent”, da qual a variação de frequência fundamental (no caso do pitch accent) é o correlato acústico principal. Traduzir “stress” por “acento” resolve uma parte da questão terminológica, mas resta ainda a expressão pitch accent, que seria talvez mais adequada a situações em que a altura é o único (e não somente o principal) correlato acústico do acento, que Beckman denomina non-stress-accent ao descrever as propriedades do acento no japonês. Cremos que traduzir por “acento de altura” não resolve o problema pois, no sentido usado por Ladd (d'après Bolinger, e, especialmente, Pierrehumbert), o fenômeno não está ligado ao acento e sim à entoação. Quando abordarmos as tendências contemporâneas no estudo da entoação, voltaremos a esta discussão.

Independentemente da discussão terminológica, a teorização apresentada por Ladd reconhece a natureza abstrata do acento. Entretanto, ela faz uma distinção entre acento de palavra (word-stress) e acento de sentença (sentence-stress), sendo o primeiro abstrato, enquanto que o segundo apresenta correlatos acústicos. Segundo Lehiste (1970 apud LADD, 1996, p. 48), “It appears that word-level stress is (…) a potential for being stressed, (…) the capacity of a syllable within a word to receive sentence-stress when the word is realized as part of the sentence”. Assim, o que estes autores ligados ao desenvolvimento da TMA entendem por word-stress é, podemos concluir, a estrutura abstrata sobre a qual as proeminências relativas se organizam na sentença, estas, sim, marcadas concretamente por correlatos acústicos (sentence-stress). Sintetizando esta concepção, Ladd afirma que “lexical stress is a phonological abstraction, and the 'stressed' syllable of a word is simply the place where actual sentence level prominence occurs, if the word is prominent in a sentence. (LADD, 1996, p. 48).

Mais útil para nossos propósitos aqui é a discussão em torno do que Liberman chama de “tune-text association” (LADD, 1996, p.52). Buscando uma analogia com a relação entre texto e melodia na canção, Liberman observou que as sílabas acentuadas do texto não coincidem necessariamente com notas mais longas, mais fortes ou mais agudas, o que seria de esperar se o acento estivesse diretamente relacionado com correlatos acústicos. Ao contrário, o texto da canção se adapta à melodia de maneira a fazer coincidir os acentos de palavra com as posições métricas fortes dentro do compasso. A respeito da noção de forte e fraco na posição métrica, Ladd comenta:

> “It is important to emphasize (...) that no absolute degree of prominence is implied by the labels 'strong' and 'weak'. There is no phonetic interpretation whatsoever on either label, but only of whole structures. (…) This abstract structural understanding of 'strength'(...) is crucial for the metrical interpretation of experimental studies of the perception of stress.” (idem, ibidem)

Vemos, portanto que na fala, assim como na música, a posição métrica é um dado

fundamental que orienta a percepção. De fato, como nos lembra Ladd (op. cit., p. 53), na música há uma distinção clara entre a posição métrica forte (o ictus2) e a real proeminência de uma nota, seja por intensidade, duração ou altura. O aparecimento de notas proeminentes em posição métrica fraca, como acontece nos contratempos e nas síncopes, não constitui má formação da frase musical, mas é usado para criar efeitos expressivos.

A esta altura podemos propor uma concepção de acento baseada nas propostas apresentadas acima. O acento será considerado neste trabalho como uma grade métrica abstrata sobre a qual se organizam as proeminências relativas das sílabas ou moras, conforme a língua, e cuja realização está sujeita tanto às regras do sistema prosódico da língua como ao contexto em que o enunciado é proferido. A interação dos correlatos acústicos de altura, duração, intensidade e qualidade (timbre) se dá sobre a estrutura temporal abstrata provida pelo acento, que ao mesmo tempo "amarra" os segmentos em ordem na cadeia, unindo léxico e prosódia. Vejamos agora as particularidades da relação especifica da altura com o o sistema prosódico como um todo. Esta grade se apoia nas recorrências periódicas (não necessariamente regulares) de proeminências em um ou mais níveis fonológicos, organizando o enunciado e orientando sua compreensão e interpretação.

Entoação

Ao contrário do acento, o fenômeno da entoação passou a ser investigado sistematicamente somente no século XX, embora sejam conhecidos registros anteriores que atestam a percepção das variações de altura na fala (FOX, 2000, p. 271). Mesmo durante o século XX, com os avanços trazidos pela fonologia, a entoação só passou a ser tratada como fenômeno genuinamente linguístico, passível de ter uma descrição fonológica, a partir do final da década de 1960, e a descrição e explicação dos fenômenos ainda é tema de debates e fonte de controvérsias.

Este cenário se deve a inúmeros motivos. Em primeiro lugar, como assinala Fox (2000, p. 269) vários contornos melódicos característicos que observamos na fala cotidiana são significativos per se, resistindo a uma decomposição em traços distintivos, essencial para uma descrição fonológica. Este fato pode explicar por que durante a primeira metade do século a entoação foi considerada por muitos um aspecto universal da linguagem, relacionada com funções comunicativas, contexto do discurso, emoções, atitudes, polidez, e assim por diante (FITZPATRICK, 2000, p. 3). A constatação da existência de aspectos universais, extra-linguísticos, e aspectos particulares de cada língua na entoação traz o problema da distinção

entre estes aspectos, que quase nunca é facilmente delineável.

Em segundo lugar, a complexa interação dos parâmetros acústicos dificulta a redução da profusão de realizações fonéticas possíveis para uma mesma função linguística. Embora a frequência fundamental desempenhe papel central na realização acústica, ela não é a única responsável pela construção do sentido da entoação, pois precisa ser considerada juntamente com a intensidade e a duração, além de ser interpretada tendo a grade métrica do acento como sistema de referência. Além desses motivos, podemos acrescentar as dificuldades na adaptação de métodos consagrados de investigação fonológica ao fenômeno, e, principalmente, a falta de uma definição clara e consensual do que seja a entoação, e que contemple as diferenças na relação da entoação com os sistemas prosódicos de cada língua (VAISSIÈRE, 2006, p. 238 e ss.).

Hirst e Di Cristo (1998, p. 3 e ss.) propõem que o conceito de entoação seja restrito aos traços prosódicos (supra-segmentais) com função supra-lexical, pós-lexical ou simplesmente não-lexical, ou seja, que não sejam parte essencial da identidade lexical das palavras. Desta forma a entoação pode ser considerada como um subsistema prosódico, que lida com os aspectos não-lexicais (sintáticos, pragmáticos, etc.) dos enunciados. Os aspectos lexicais ficam a cargo do tom, do acento (stress) e da quantidade. Esta restrição, entretanto, não permite uma definição física, acústica, do que seja a entoação, já que quase todos os traços prosódicos compartilham os parâmetros acústicos de altura, intensidade, duração e qualidade, apenas os combinam de formas diferentes (e simultaneamente).

A explicação de como estes parâmetros se combinam formando os diferentes subsistemas prosódicos se encontra na difícil e controversa área que é a interface entre a fonética e a fonologia. Na visão de Hirst e Di Cristo (op. cit., p. 6), a fonética corresponde à interface entre uma abordagem cognitiva, abstrata (a fonologia), e a descrição concreta, acústica dos fenômenos, e cada um destes "níveis" de análise (para usar o termo que eles empregam) requer ou resulta uma diferente definição de entoação. No nível abstrato, cognitivo, fonológico, a entoação é definida como o subsistema não-lexical do sistema prosódico da língua. No nível concreto, acústico, como vimos, não é possível definir com precisão a entoação, mas considerando a fonética como a interface entre o cognitivo e o acústico, os autores propõem uma segunda definição (que eles denominam fonética) de entoação como sendo a construção pela qual os traços prosódicos dos subsistemas lexicais e do não-lexical se relacionam com os parâmetros acústicos (op. cit. p. 7).

Repare-se que as duas definições estabelecem uma dicotomia entre forma e função da entoação. Na definição do ponto de vista cognitivo, o que distingue a entoação é a função

(não-lexical) desempenhada; na definição do ponto de vista fonético, o que está em jogo é a forma como os traços se combinam. Esta mesma dicotomia está presente quando descrevemos a entoação como "uma melodia que representa uma intenção". O primeiro termo desta fórmula (percepção de variação de altura) nos conduz a uma definição formal de entoação, baseada no comportamento físico do som da fala; já o segundo termo (percepção de intenção) nos levaria a uma definição funcional da entoação, ou seja, baseada na finalidade com que empregamos a entoação – definir a entoação a partir da função que ela desempenha na comunicação oral.

A noção trazida do senso comum de se considerar a entoação como melodia implica uma relação direta com a altura e consequentemente, com a frequência fundamental (F0) de vibração das pregas vocais durante a fala. A "forma" da entoação, portanto, se manifesta por meio das variações desta frequência ao longo do tempo. Esta forma pode, inclusive, ser visualizada num diagrama frequência versus tempo ou, como denominava Pierre Schaeffer (1966), no plano melódico (cujos eixos correspondem a altura e tempo).

Entretanto, se nos reportarmos à definição de melodia na teoria da música, veremos que a simples ordenação das alturas não identifica uma melodia. É necessário que sejam considerados os elementos rítmicos, não só de ordenação de durações e acentos, mas também (e, talvez, principalmente), o posicionamento das variações de altura com relação ao compasso e à subdivisão dos tempos, para que se reconheça a identidade de uma melodia (GROVE ONLINE, 2012). Na música, se um movimento melódico (qualquer que seja o intervalo, ascendente ou descendente) é deslocado com relação ao sistema de referência rítmico, baseado numa sucessão hierarquizada de posições fortes e fracas, ele muda sua identidade. Assim, um mesmo intervalo executado de um tempo fraco para um tempo forte ou de um tempo forte para um tempo fraco será percebido como melodicamente diferente.

Já é fato reconhecido que existem diferentes maneiras de a altura se relacionar com o sistema prosódico de uma língua. No caso das línguas tonais, por exemplo, variações de altura são empregadas para marcar distinções lexicais; já nas línguas chamadas entoacionais, como é o caso do português, isto não ocorre, fato que permite um uso mais livre da altura para outras finalidades, como expressar atitudes e modalidades. Existem também situações intermediárias, que são as línguas que marcam acentos de palavra com a altura, como é o caso do sueco e do japonês; essas duas línguas também se distinguem pela função do acento de palavra, resultando um uma distinção ulterior em língua de tom acentual (japonês) e língua de acento tonal (sueco) (HIRST; DI CRISTO, 1998).

Assim, para delimitarmos o que seria a "entoação propriamente dita", é necessário

excluir as variações de F0 que atendem a restrições lexicais e acentuais. Esta definição traz implícita a hipótese de que é possível, pelo menos teoricamente, separar os efeitos puramente lexicais sobre a curva de F0 dos efeitos intencionais, ligados a atitudes e modalidades. Abre-se, então, outro ponto de dúvida: essa distinção é viável a partir do sinal acústico, sem acessar informações no nível fonológico (acessível à análise automática, portanto)?

Acrescente-se a isso outras interferências não-linguísticas ou paralinguísticas sobre a curva de F0, como o estado psicológico e características idiossincráticas do falante, e veremos como é inviável estabelecer-se uma correspondência unívoca entre a forma da curva de F0 e a entoação propriamente dita.

Esta breve exposição demonstra, cremos, que uma definição em bases puramente formais/acústicas é insuficiente para apoiar um estudo bem sucedido da entoação. Demonstra, além disso, que a entoação propriamente dita, segundo Hirst e Di Cristo, precisa atender (não pode apagar, superpõe-se sem anular) as distinções lexicais, e por isso o emprego da altura com finalidades não-lexicais varia conforme o sistema prosódico da língua, tornando bem mais complexo o trabalho de comparar esses usos não-lexicais.

Esta constatação é a grande motivação para as abordagens que estamos chamando de funcionais. Uma descrição possível destas abordagens seria a identificação de usos equivalentes (funções), sua posterior classificação e isolamento experimental para daí chegar a alguma conclusão sobre seus correlatos acústicos. Vejamos a seguir os princípios e problemas de abordagens funcionais.

Fónagy (2003) apresenta um levantamento de inúmeras funções da entoação que ele encontrou em autores diversos. O autor fala em função demarcativa, de ênfase, gramatical, sintática, modal, imitativa, de apelo, lógica, preditiva, alusiva, identificativa, estética, expressiva, exploratória e preparatória, além de uma caracterologia vocal, esta ligada ao trabalho do ator. Apesar da grande abrangência do seu levantamento, o autor busca evidências em textos provenientes de quadros teóricos muito diferentes, e acaba incorrendo em certas redundâncias. Por exemplo, a função preditiva poderia ser considerada como parte da sintática, por estar relacionada com tipos de coordenação (aditiva e adversativa). Entretanto a exposição tem o mérito de se ater a usos deliberados e intencionais do falante.

Vaissière (2006), por outro lado, apresenta uma lista mais sucinta que a de Fónagy (que fora seu professor), porém introduz elementos que não podem ser considerados como usos intencionais do falante. Estes podem ser considerados funções somente do ponto de vista do ouvinte. Abaixo um resumo da lista de Vaissière.

Sintáticas	*Segmentação da fala contínua em unidades sintáticas de diferentes tamanhos*: palavras prosódicas, sintagmas, proposições, sentenças, parágrafos
Informacionais	*Segmentação da fala contínua em unidades informacionais*: tema/rema [tópico/comentário], dado/novo, foco/parêntese
Interativas	*Regulação da interação falante-ouvinte*: atração de atenção e interesse, turno, final/continuação do tópico
Modais	*Intenção comunicativa*: asserção/pergunta/ordem, etc.
Atitudinais	*Atitudes do falante com respeito ao que ele diz*: dúvida, descrença etc. *Atitudes do falante com respeito ao ouvinte*: polidez, ironia etc.
Emocionais	*Emoções do falante*: alegria, raiva etc.
Outras	*Características do falante*: identidade, gênero, idade, estado fisiológico, variantes regionais, variações estilísticas, aspectos sócio-culturais

Tabela 1 – Algumas das múltiplas funções da entoação (Vaissière, 2006, p. 237)

Se entendemos o termo função como "o uso com finalidade", somos obrigados a excluir as duas últimas linhas da tabela, que, no entanto, são informações importantes que também estão impressas na melodia da fala. Podemos considerar as informações que percebemos nas variações prosódicas não-lexicais divididas em dois grupos: aquelas que são causadas pela intenção do falante (no caso da tabela acima informacionais, interativas, modais e atitudinais) e aquelas que não são causadas pela intenção, que podem tanto ser restrições impostas pelo sistema fonológico da língua (as sintáticas) como contingências a que está sujeito o falante (emocionais e outras). Assim, sob uma perspectiva funcional, só seria considerada entoação aquela variação prosódica não-lexical que revelasse uma intenção comunicativa do falante: ênfases, focos, modalidades, atitudes, trocas.

Neste contexto, uma definição possível para a entoação seria o conjunto de alterações não-lexicais na estrutura de superfície de um enunciado que o falante emprega com o propósito de mostrar quais são as suas intenções (sinceras ou fingidas) ao proferir um enunciado, induzindo um comportamento responsivo ou então dando opções para o seu interlocutor. A entoação, vista sob esta perspectiva, mostra ao ouvinte como o enunciado que ele ouviu pode e deve ser interpretado, considerando, além disso, a possibilidade de se perceber a intenção do falante.

Assim como a concepção formal da entoação, a concepção funcional também tem

fragilidades, sendo a primeira delas o conceito de intencionalidade, que é um problema filosófico de primeira grandeza. Além disso, ela fica incompleta sem alguma consideração sobre a natureza física que transporta a informação interpretada.

Sobre a intencionalidade, podemos delimitá-la às escolhas possíveis, excluindo-se, portanto, as restrições da língua e as contigências a que o falante está sujeito no momento em que profere o enunciado, evitando investigações mais profundas no campo da filosofia. Quanto à natureza do suporte da informação, podemos recorrer à definição formal desenvolvida acima. Desta combinação, resulta a definição abaixo:

> **A entoação é o conjunto das realizações prosódicas não-lexicais que resultam da intenção comunicativa do falante, envolvendo a organização da informação e as diretrizes para a sua interpretação.**

Repare-se que o aspecto formal traduzido por "realizações prosódicas não-lexicais" acaba abrangendo todo o conjunto das línguas, independentemente de serem tonais, acento tonal, tom acentual ou entoacionais. Vale também a observação de que "organização da informação" envolve tanto os aspectos sintáticos de construção do enunciado como os aspectos pragmáticos de foco e ênfase, e as "diretrizes para a sua interpretação" correspondem a atitudes e modalidades. Esta definição ainda não resolve nossos problemas, mas ajuda a situá-los numa perspectiva mais ampla, ao mesmo tempo em que separa os aspectos linguísticos dos paralinguísticos.

1.2. Teoria métrica-autossegmental

A teoria autossegmental (TA) surgiu em meados dos anos 70 como um desenvolvimento da fonologia gerativa, tendo como principal novidade a distribuição da cadeia segmental em linhas ou fileiras (*tiers*) paralelas de traços distintivos, o que permite uma análise fonológica multi-linear em que os traços distintivos são associados transversalmente segundo regras fonológicas. A expansão de uma estrutura linear para uma multi-linear possibilitou não só uma visão mais refinada da estrutura interna dos segmentos, mas também a inclusão de linhas dedicadas à duração, intensidade e tom (GOLDSMITH, 1995). A teoria métrica (TM) também surgiu no mesmo período, paralelamente à TA e compartilhando com esta uma concepção não-linear da análise fonológica, dedicando-se, a

princípio, à questão do acento de palavra (KAGER, 2001, p. 360).

Uma das principais tendências do estudo da entoação, cuja origem pode ser atribuída aos trabalhos de Liberman, Bruce e Pierrehumbert, desenvolvidos no final dos anos 70, combinam elementos da TA e da TM, motivo pelo qual Ladd (1996, p. 42) denomina esta corrente de teoria métrica-autossegmental (autosegmental metrical theory), doravante TMA. Para a entoação, esta teoria

adota a meta fonológica de ser capaz de caracterizar adequadamente os contornos melódicos em termos de uma sequência de elementos categoricamente distintos, e a meta fonética de prover um mapeamento dos elementos fonológicos [discretos] para parâmetros acústicos contínuos" (idem, ibidem).

Para a TMA, o contorno melódico é constituído de uma sequência de eventos locais associados a certos pontos da cadeia segmental, entre os quais a variação da F0 não é fonologicamente especificada, podendo ser descrito em termos de transições de um evento ao próximo. Estes eventos são os pitch accents mencionados acima, que são formados por tons (level tones) ou alvos (pitch targets) que podem ser H (high, agudo) ou L (low, grave). A realização fonética destes eventos (a F0) está sujeita a inúmeros fatores, tais como ênfase, posição no enunciado, estado do falante, relação com o interlocutor, etc., que, no entanto, não influenciam a percepção do evento como H ou L (LADD, op. cit., p. 43).

O trabalho de Pierrehumbert, que propôs os conceitos de pitch accent e boundary tone (tom de fronteira), bem como uma notação sistemática para associar estes eventos com a cadeia segmental, é até hoje um dos mais influentes marcos conceituais no estudo da entoação (FOX, 2000, p. 287). O sistema ToBI (Tones and Break Indices) de transcrição da entoação, que deriva do trabalho de Pierrehumbert, foi aplicado para o estudo de inúmeras línguas, tornando-se possivelmente o sistema de notação mais conhecido, empregado até mesmo em estudos que não têm uma relação direta com a fonologia não-linear.

O ponto forte da TMA é o seu rigor fonológico, apoiado por uma teoria robusta que tem demonstrado resultados positivos para línguas dos mais diferentes tipos e origens. A concepção do acento como uma estrutura abstrata e a investigação de sua relação com a associação entre os eventos melódicos e a cadeia segmental pode esclarecer fenômenos como o acento tonal do sueco (LADD, 1996, p. 54). Ao abandonar a ideia de que os picos de F0 seriam correlatos acústicos do acento, a TMA pode explicar as variações significativas do alinhamento da curva melódica.

Segundo a TMA, "the invariant features of the pitch system appear to be the turning points in the contour rather than the transitions that connect them" (idem, ibidem, p. 63), ou

seja, o importante é atingir um determinado nível ou patamar de altura (pitch level) no momento certo, e não o movimento da F0 . Ladd afirma que "there is now abundant evidence that speakers control the level and temporal alignment of certain F0 target points with considerable precision, and that hearers can perceive the differences" (idem, p. 64)

A TMA, portanto, não pretende explicar as variações possíveis (realizações fonéticas) da curva de F0, apenas busca compreender como estes target points se relacionam com a cadeia segmental ritmicamente, sintaticamente, pragmaticamente. Há evidência experimental de alguma regularidade na frequência dos alvos (pitch targets) em situações emocionalmente neutras, porém ainda há dúvidas sobre as posições destes alvos no eixo das alturas.

Bruce sugere que a realização fonética dos tons H e L, abstrações fonológicas, se dê, de maneira idealizada, em uma escala de quatro níveis de F0, sendo 1 o mais grave e 4 o mais agudo. Vejamos a proposta de Bruce.

> "F0-level 1 is considered to the base level and is the true representative of the LOW pitch level [i.e. L tone]. (...) In certain contexts the LOW pitch level will also be specified as F0-level 2 (and occasionaly F0-level 3). The HIGH pitch level [i.e. H tone] can be specified as F0-level 2, 3 or 4, depending on the context. This means that F0 level2 can represent bot a HIGH and a LOW pitch level, which may seem paradoxical. But the pitch levels HIGH and LOW are to be conceived of as relative and contextually specified for each case as a particular F0 -level." (1977 apud LADD, op. cit., p. 71)

Estes níveis, no entanto, não são fixos, e sua realização está sujeita às circunstâncias da elocução. Vários autores, não só os ligados à TMA, reconhecem a existência de uma tendência geral de declínio na F0 ao longo da elocução, e este fenômeno é conhecido como declinação frasal, e está relacionado com as condições específicas da produção da voz enquanto sistema dinâmico. Entretanto, ainda há muito a explicar quanto à consistência ou não do falante em atingir estes alvos.

Nos estudos da entoação do PB, o modelo está presente em trabalhos das mais diferentes naturezas e alinhamentos teóricos, sejam trabalhos voltados para questões pragmáticas relacionadas com a organização da informação, ou nos estudos das modalidades e atitudes, ou até mesmo abordagens baseadas em modelos dinâmicos, para citar apenas algumas das tendências mais destacadas no país.

A TMA chegou ao século XXI com resultados importantes quanto à organização do discurso, identificando relações entre os pitch accents e aspectos pragmáticos como marcas de tópico e comentário, foco e ênfase, dado e novo (HIRSCHBERG, 2002; CRESTI, 2011), bem como aspectos sintáticos como delimitação, coordenação e subordinação de orações; também pode ser encontrada em estudos voltados à descrição dos contornos em termos de pitch levels, visando à identificação de contornos particulares relacionados a modalidades e atitudes

proposicionais (MORAES, 1998). Porém, a teoria ainda apresenta alguns inconvenientes.

Em primeiro lugar, a análise de amostras de fala em termos de pitch accents associados a pontos da cadeia segmental é feita manualmente e há considerável discordância entre os sujeitos que efetuam a transcrição (LUCENTE, 2012). Em segundo lugar, ainda está em aberto a explicação sobre a escolha dos pitch levels na realização fonética dos pitch accents, embora o comportamento da curva melódica tem recebido importantes contrimbuições dos chamados modelos dinâmicos (ver abaixo). Além disso, as variações na curva de F0 que são possíveis mantendo a estrutura dos tons parecem ter implicações atitudinais e de modalidades, e o modelo, ao desconsiderá-las como simples transições entre alvos, parece desistir de uma interpretação fonológica do fenômento (embora cf. CRESTI, 2011).

Fox (op. cit., p. 301) chega a duvidar de que os níveis de altura, mesmo que interpretados na configuração binária H e L, possam desempenhar o papel de traços distintivos (não podemos associar H e L à presença e ausência de um traço). Ele acredita que a busca destes traços deve se dar por meio da análise comparativa de padrões recorrentes da entoação numa determinada língua, considerando não só os alvos, mas também o movimento melódico e seu alinhamento com a cadeia segmental. Não obstante estas fragilidades, a TMA é certamente um dos pilares do conhecimento atual sobre a entoação, e seus pressupostos precisam ser considerados num estudo da entoação que pretenda incluir a conexão da prosódia com o sistema da língua como um todo.

1.3. Outros modelos

Os estudos que se alinham com título “modelos dinâmicos” adotam uma abordagem radicalmente diferente da TMA, buscando explicar o funcionamento da fala a partir dos mecanismos que a compõem, entendendo-a como um comportamento dinâmico auto-organizado.

A origem desta concepção pode ser atribuída ao modelo Fujisaki para a produção da fala, proposto no final da década de 60, voltado à princípio para a síntese de fala, e que foi aplicado como algoritmo de geração do contorno melódico por Gårding e colegas nas décadas seguintes (FOX, 2000, p. 283).

Na mesma época, estudos sobre o aspecto mecânico do comportamento humano levou ao desenvolvimento do conceito de motor phonetics, em trabalhos como os de Kelso,

Saltzmann, etc. Resultados experimentais levaram à proposta de caracterizar o mecanismo do comportamento físico da fala como uma combinação de diferentes osciladores que cuidam dos diferentes parâmetros. Esta nova concepção abriu caminho para explicações consistentes, especialmente para o ritmo e o acento (cf. BARBOSA, 2006).

Xu (2004, 2009) propõe um modelo para a entoação que resulta de estudos do inglês e do mandarim, que ele denomina modelo PENTA (Parallel Encoding and Target Approximation).

Lucente (2008, 2012) propõe para o PB um sistema de transcrição adaptado do ToBI que incorpora considerações dinâmicas na identificação e interpretação do alinhamento de contornos característicos da língua, o DaTo.

Os modelos dinâmicos têm ganhado terreno nos estudos da entoação, e têm a vantagem de se articular com outras áreas como a síntese e reconhecimento de fala. Porém, a ênfase na realização fonética deixa em segundo plano a questão da descrição fonológica do contorno melódico, o que inibe a explicação da entoação como um sistema linguisticamente organizado.

Uma alternativa a estas duas principais correntes (mais fonológica, como a TMA, ou mais fonética, como os modelos dinâmicos) pode ser encontrada no trabalho de Ferreira Netto (2006, 2008), que desenvolve métodos de análise automática da entoação diretamente da curva de F0, sem levar em consideração a cadeia segmental. Seu objetivo é encontrar padrões de comportamento global da fala, empregando métodos estatísticos.

Um resultado notável é a constatação da relação intervalar entre o tom médio e o tom de finalização, que tem se demonstrado consistentemente próxima do valor de 7 ST (para a leitura em voz alta), ou seja, uma quinta justa (a mesma relação entre dominante e tônica na música tonal). Este resultado propõe uma perspectiva para o estudo musical da entoação.

1.4. Impasses no estudo da entoação

Ao considerarmos a história do desenvolvimento dos estudos da entoação, podemos perceber, na segunda metade do século XX, um salto notável no número de trabalhos. Isso se deveu tanto ao crescente interesse pela prosódia como parte integrante do sistema linguístico e indispensável à sua descrição, como ao auxílio da tecnologia, que permitiu gravações acuradas e, na era digital, análises estatísticas e processamento por computador do sinal acústico. Entretanto, mesmo com todos os resultados e evidências crescendo em proporção

geométrica, algumas questões que já eram debatidas desde as primeiras décadas do século passado ainda desafiam os teóricos e são motivo de discussão. Fox (op. cit., p.) interpreta este fato como um descompasso entre o modelo teórico linguístico e a metodologia de investigação.

Uma das questões mais debatidas e conhecidas é a chamada controvérsia níveis versus configurações. Segundo Fox, a questão se o contorno melódico é composto de níveis de altura determinados fonologicamente (embora ainda não esteja totalmente claro como isto se dê) associados a pontos da cadeia segmental, ou se é composto da combinação e superposição de contornos específicos (que seriam, também, fonologicamente determinados). Ladd acha que a TMA resolveu o problema separando os eventos isolados de busca de um alvo (H ou L) como sendo os elementos significativos da entoação, ficando o comportamento da F0 entre estes eventos sujeito às circunstâncias da elocução. Porém, esta hipótese deixa de fora inúmeras características que percebemos na curva melódica e que não podem ser consideradas paralinguísticas, e certamente transportam informações relevantes à comunicação.

Os modelos dinâmicos ajudaram a esclarecer a formação dos contornos melódicos e sua relação com o acento. Mas por mais que o modelamento físico-matemático da entoação tenha se desenvolvido notavelmente nas duas últimas décadas, a investigação experimental ainda enfrenta um outro problema que é o estabelecimento do corpus, devido às peculiaridades da fala espontânea que são difíceis de serem reproduzidas em laboratório e, para serem extraídas de gravações de fala espontânea (normalmente com baixa qualidade de áudio) exigem um número muito grande de horas.

A comparação entre fala espontânea, decorada e leitura em voz alta mostra um padrão muito mais caótico para a primeira. Isso dificulta, entre outras coisas, a compreensão da questão da declinação, que ainda suscita debates.

Avaliando as três questões resistentes, acreditamos não ser exatamente um descompasso, mas talvez um distanciamento da teoria (linguística) da tecnologia. Os trabalhos com resultados mais consistentes tem sido aqueles em que o modelo teóricos incluem a descrição físico-matemática e a análise estatística viabilizadas pela gravação digital e pela facilidade e acesso a ambientes de programação.

Outra questão aparentemente pacificada mas que esconde dúvidas prementes é a da função ou funções da entoação. Em que pese a problemática natural ligada ao termo função, que carrega consigo uma miríade de significados oriundos das mais diversas áreas, vamos mantê-lo no sentido genérico de finalidade, propósito, o “para quê serve”, do ponto de vista da comunicação (produção e percepção). A ideia de função da entoação esconde uma

premissa da precedência da cadeia segmental, à qual o contorno melódico seria, num segundo momento cognitivo, acrescentado para especificar os significados pretendidos. Esta hipótese não pode ser testada, mas a tendência da compreensão da fala como comportamento dinâmico auto-organizado nos leva a questioná-la. De qualquer modo, explicitar esta premissa nos permite controlar o seu efeito sobre a argumentação.

Outro impasse está ligado ao debate histórico na linguística que é a disputa entre fonologia e fonética, que correspondem *grosso modo* às abordagens analíticas *top down* e *bottom up,* a primeira ligada à ideia de uma estrutura subjacente à superfície dos enunciados (caso, por exemplo, da gramática gerativa) e a segunda talvez mais próxima dos modelos dinâmicos.

Não temos a pretensão de sugerir alguma solução para este impasse, mas acreditamos que as regras fonológicas não se limitam a explicar a superfície da cadeia segmental, mas participam ativamente da estrutura do nível morfológico, por meio da prosódia. A morfologia prosódica é um dos grandes trunfos da fonologia.

Podemos supor, daí, que no horizonte da fala há um plano, um layout, um sketch, de enunciação que é acionado pela intenção comunicativa e empregado pelo mecanismo da fala (os osciladores acoplados, por exemplo).

O sistema fonológico é que executa a tarefa de especificar a formação dos morfemas a partir das escolhas semânticas; sua informação é lida pelo sistema que produz a fala, que é dinâmico e sujeito às circunstâncias. O léxico e as regras fonológicas compartilhadas é que permitem a comunicação linguística.

Acreditamos que a investigação experimental deve atender a um objetivo fonológico de descrever as regras e como o sistema dinâmico executa as instruções, e que ainda há um longo caminho a ser percorrido, mas que devemos nos animar com os avanços nas últimas décadas.

Não é, entretanto, o objetivo deste trabalho. O que importa aqui é reconhecer a possibilidade de descrição fonológica da entoação e daí derivar que o controle (inconsciente) das regras fonológicas denota a intenção comunicativa ou, propondo uma inversão da fórmula: o que distingue a intenção comunicativa dos acidentes na curva de F0 é o reconhecimento (inconsciente) da aplicação de uma regra fonológica, o que nos leva à última questão epistemológica, que diz respeito à intencionalidade na comunicação humana, tema que ultrapassa os limites da linguística, e mesmo dos estudos da linguagem, exigindo incursões em outros campos, como a filosofia e a as recém-criadas neurociências.

A avaliação desses impasses epistemológicos chama a atenção para certos aspectos da

entoação, que reunimos aqui à guisa de síntese. Em primeiro ligar, a entoação precisa ser definida no contexto mais amplo da prosódia, e não se limitar à variação da F0. A prosódia ganhou um estudo sistemático a partir da fonologia e esse trabalho, mesmo que não explicite, procura se valer da concepção de um comportamento (pelo menos dentro de certas condições) da entoação que segue regras de natureza fonológica. Neste sentido, a concepção de acento como grade métrica que engendra uma abstração rítmica é um dado fundamental a ser considerado, especialmente quando formos buscar analogias entre fala e canto, entre linguagem e música

A interface fonética é extremamente complexa devido à interação entre os parâmetros, e o problema da decomponibilidade (paradigmática ou sintagmática) da entoação leva à colocação de uma questão fundamental: "como distinguimos a intenção do falante em meio a todos os acidentes casuais (contigências) que "distorcem" a curva de F0?". A resposta que concluímos a partir da argumentação acima é "reconhecendo as regras fonológicas".

2. Outros olhares sobre a entoação

2.1. Uma palavra sobre interdisciplinaridade

A linguística é a área do conhecimento privilegiada para o estudo da entoação, mas o conhecimento sobre o comportamento melódico da fala interessa a diferentes campos de atuação humana, desde o aprendizado de uma língua estrangeira até o treinamento dos atores de teatro e locutores, passando pela perícia forense, reconhecimento e síntese de fala e até mesmo a inteligência artificial. Da mesma forma, simetricamente, todas estas áreas têm algo a acrescentar às pesquisas sobre entoação e por este motivo equipes inter- ou multi-disciplinares têm se dedicado à investigação da entoação nas últimas décadas.

Japiassu (1976) coloca a interdisciplinaridade como uma exigência atual e interna das ciências humanas, por fornecer um ponto de vista que permite uma tomada de consciência sobre o lugar real da posição e do tratamento dos seus principais problemas epistemológicos. O autor considera o problema da interdisciplinaridade como a questão do diálogo entre as disciplinas, que deve levar a uma "tentativa de formulação de uma interpretação global da experiência humana" (*op. cit.*, p. 29).

Segundo Japiassu (*ibidem*, p. 32), "a característica central da interdisciplinaridade consiste no fato de que ela incorpora os resultados de várias disciplinas, tomando-lhes de empréstimo esquemas conceituais de análise a fim de fazê-los integrar, depois de tê-los comparado e julgado". Entre várias razões que justificam a abordagem interdisciplinar, o autor cita o questionamento de uma "possível acomodação dos cientistas em seus pressupostos implícitos, em suas comunicações restritas que tornam difíceis as trocas e, com isso, favorece a explicitação de seus postulados epistemológicos, cada um deles fazendo dessa explicitação o acompanhamento necessário de sua prática e de suas eventuais descobertas" (*idem*, *ibidem*, p. 33). Por este motivo acreditamos que a incorporação de esquemas conceituais e resultados de diferentes disciplinas sobre um mesmo objeto (aqui, neste caso, a entoação) pode ajudar a superar os impasses que derivam de posicionamentos nem sempre implícitos, como os que foram apresentados no final do primeiro capítulo.

Japiassu alerta para o risco de superficialidade em virtude do "modismo" da interdisciplinaridade, importado das grandes universidades europeias e norte-americanas (*op. cit.*, p. 40 e ss.); avalia, porém, que o diálogo entre as disciplinas pode evidenciar as lacunas do próprio sistema universitário e das instituições de pesquisa, contribuindo para sua crítica e

seu desenvolvimento. Além disso, há outra razão prática para a aplicação de abordagens interdisciplinares, que é o que o autor chama de "curva exponencial seguida pela expansão científica", que exige uma organização interna da pesquisa, visando à sua economia e eficácia: "dado o elevado custo que ela implica, não é mais possível nem aceitável que cada disciplina se desenvolva independentemente das outras" (*ibidem*, p. 57).

Outra ressalva feita por Japiassu é o desafio de se encontrar uma linguagem comum, de haver concordância sobre as concepções iniciais das diferentes disciplinas malgrado as divergências terminológicas. O autor alerta que "precisamos estar conscientes das múltiplas dificuldades de levar a bom termo um vocabulário comparado das ciências humanas permitindo-nos esclarecer certo número de solidariedades e de oposições nos confins das diversas disciplinas" (*op. cit.*, p. 97). Para contornar estas dificuldades, Japiassu considera que o mais importante não é a busca de um vocabulário padrão pela formulação de novas definições, mas, sobretudo, "ressaltar todas vicissitudes dos conceitos através da diversidade do espaço epistemológico e segundo a renovação contínua das estruturas mentais" (*ibidem*, p. 98). Neste trabalho, e especialmente neste capítulo, tomamos o cuidado de evitar termos que possam representar concepções muito diferentes nas áreas contempladas, ou mesmo controversas dentro de cada disciplina, dando preferência aos conceitos já estabelecidos. Ao mesmo tempo procuramos explicitar as eventuais divergências que possam aparecer.

Asim como a entoação ocupa uma posição marginal nos estudos linguísticos, tendo sido inclusive considerada por muito tempo como um fenômeno paralinguístico assim como os gestos e as expressões faciais, este capítulo apresenta conceitos e propostas de disciplinas fronteiriças à linguística que lidam direta ou indiretamente com o conceito de entoação, especialmente na sua acepção de intenção expressiva. Cada seção é nomeada com o título de uma obra usada como referência principal, nas áreas da filosofia da linguagem, sociologia, teoria literária, estudos teatrais e da teoria e análise musical. Além destas disciplinas, é possível perceber ao longo do percurso conexões com as ciências do comportamento, a tecnologia da informação e as artes.

2.2. Os atos de fala

O filósofo britânico John Austin deu início à reflexão que levaria ao desenvolvimento da teoria dos atos de fala (speech acts), nos anos 50, analisando o que chamou de verbos performativos, tais como "batizar" ou "declarar", que modificam a realidade simplesmente

com a sua enunciação, distinguida dos verbos constativos, que apenas representam um estado do mundo (AUSTIN, 1962). Em textos posteriores, Austin argumenta que toda enunciação tem valor performativo, já que esta realiza uma ação sempre que houver um interlocutor, o qual será alterado ou influenciado em sua percepção do mundo, da conversa, do seu interlocutor (MAINGUENEAU, 1996, p. 7). Esta reflexão leva o filósofo a concluir que toda enunciação tem uma dimensão ilocutória, além do seu conteúdo proposicional. Austin denominou "atos ilocucionais" os atos de fala completos, ou seja, que envolvem referência, predicação e uma intenção, como os caracterizados pelos verbos afirmar, advertir, comentar, etc.; distinguia estes dos "atos locucionais", que só realizam referência e predicação.

John R. Searle, em continuidade ao trabalho de Austin, adota o conceito de ato de fala, porém recusa a distinção entre atos locucionais e ilocucionais, entendendo que tudo que é dito é dito com alguma intenção (SEARLE, 1981, p. 35). O autor propõe então que na enunciação de qualquer frase o falante executa necessariamente três tipos de atos distintos numa mesma ação: a) enunciar palavras (morfemas, frases); b) referir e predicar; c) afirmar, fazer uma pergunta, dar uma ordem, prometer, etc. A estes diferentes atos de fala, Searle dá o nome de atos de enunciação, atos proposicionais e atos ilocucionais, respectivamente. Também observa que os atos proposicionais não podem ocorrer sozinhos, pois necessariamente haverá uma componente ilocucional em toda enunciação (*idem*, *ibidem*, p.37).

O ato de enunciação está ligado à ação de proferir palavras articulando morfemas e frases; pode incluir também funções demarcativas e de associação (coordenação e subordinação).

O ato proposicional está ligado à referência e à predicação, ou seja, estabelecer a conexão do que se diz com o contexto e o mundo real; as marcas desta ação podem corresponder às marcas pragmáticas que apontam para termos do enunciado, indicando ênfases (foco) ou distinguindo o tópico de seu comentário, organizando assim a informação (o conteúdo, por assim dizer).

O ato ilocucional, por sua vez, é o que poderíamos chamar de vetor principal do ato de fala, seu impulso primordial; diz respeito às intenções do falante, sua atitude com relação ao que diz e a quem diz, e também envolve modos, graus de polidez, enfim, todas as pistas que o ouvinte utiliza para decidir como deve interpretar o enunciado. As marcas associadas ao ato ilocucional caracterizam todo o enunciado, mesmo estando localizadas fisicamente em um ponto específico da cadeia segmental.

Embora Searle mencione somente de passagem o papel da entoação no reconhecimento do tipo de ato de fala, em conjunto com outros fatores prosódicos e de

contexto, podemos concluir que é possível distinguir as diferentes funções da entoação de acordo com o tipo de ato de fala ao qual estão associadas. As marcas prosódicas associadas a aspectos emocionais e a contigências do falante são percebidas pelo ouvinte como circunstanciais e por isso não pertenceriam ao ato de fala propriamente dito; são assim "desconsideradas" no trabalho de contextualizar o enunciado de acordo com a intenção comunicativa do falante.

Em que pesem as possíveis ressalvas que podem ser levantadas na argumentação de Searle quanto a definição precisa de como tal ou qual aspecto do enunciado pertence à componente de enunciação, proposição e ilocução (e até mesmo o problema destes termos terem outros significados em disciplinas correlatas), a teoria dos atos de fala tem para nós dois resultados fundamentais que serão empregados na argumentação da tese que defendemos. O primeiro é a constatação de que a linguagem é um comportamento gerido por regras, e que nem todos os enunciados gramaticalmente possíveis são bem sucedidos, o que leva à noção de boa-formação (*well-formedness*). O segundo é a ideia de que o ato ilocucional comanda o o ato de fala como um todo, pois a organização das outras componentes depende da intenção que o falante tem com a realização do ato. Está, portanto, intimamente ligado ao contexto social em que a comunicação linguística se dá, seja pela fala ou pela escrita.

Antes de levar o nosso problema ao campo da sociologia, vejamos a influência da teoria dos atos de fala nos estudos prosódicos.

As relações entre os atos de fala e a prosódia já foram estudadas por inúmeros autores e representam uma tendência importante nos estudos prosódicos atualmente. Voltando ao texto de Daniel Hirst e Albert DiCristo (HIRST E DICRISTO, 1998) apresentado no primeiro capítulo, podemos perceber que, ao propor um esquema para descrição de padrões de entoação a ser adotado pelos demais autores da coletânea, este também é dividido em categorias semelhantes às dos atos de fala (enunciação, proposição e ilocução).

Apesar de reconhecerem a problemática inerente à postulação de um padrão de entoação básico neutro e não-marcado, Hirst e DiCristo (*ibidem*, p. 18 e ss.) fazem uma síntese dos resultados de diversos pesquisadores relacionando um padrão mais ou menos recorrente atrelado as características rítmicas de cada língua. Podemos imaginar que este padrão seria usado em situações de narração, descrição ou mesmo em um discurso em que o orador não esteja muito inflamado. Porém quando passamos ao uso da língua na conversa toda uma série de fatores que alterem o padrão de entoação devem ser considerados. A partir daí, os autores separam estes fatores em quatro grupos: modo e expressividade (que corresponderiam, segundo nossa proposta ao ato ilocucional), efeitos contextuais e

focalização (ato proposicional), fraseado e organização textual (ato de enunciação) e finalmente padrões estereotipados, como é o caso dos clichês melódicos associados a chamados, saudações, advertências, enumerações, etc. Estes padrões dão a impressão de estarem a meio caminho entre a fala e o canto, e revelam aspectos importantes tanto do uso da entoação na interação linguística como da própria natureza do fenômeno (cf. a análise do problema do alinhamento no contexto das provocações infantis em LADD, 1996, p. 52 e ss.). Dada a sua relação com as funções de interação da entoação, e tendem a caracterizar o enunciado como um todo, investindo-o de uma intenção particular associada a um uso habitual, poderíamos incluí-los no conjunto das marcas prosódicas características dos atos ilocucionais.

A distinção entre as diferentes funções da entoação segundo características correspondentes aos tipos de atos de fala aparecem em outros autores (cf. FONÁGY, 2003; FOX, 2000, p. 324 e ss.), porém a classificação apresentada por Hirst e DiCristo parece ser a que tem uma relação mais explícita com a teoria dos atos de fala. Acreditamos que a distinção entre as componentes (marcas) características de cada tipo ato de fala pode facilitar o desenvolvimento de modelos que expliquem melhor as diversas funções que percebemos na entoação. Porém, neste trabalho não pretendemos desenvolver modelos específicos, mas sim refletir sobre o papel dos aspectos ilocucionais na formação da melodia da fala nas mais diversas situações de comunicação.

Dentre os diversos trabalhos que relacionam a teoria dos atos de fala com a prosódia, merece destaque aqui o trabalho de Cresti (2011). Trabalhando na interface entre a pragmática e os estudos prosódicos, Cresti propõe a Teoria da Língua em Ato para analisar os padrões de organização da informação na fala espontânea, a partir dos registros em italiano do *corpus* C-ORAL-ROM. A autora propõe uma distinção fundamental para esclarecer as confusões que rondam os pares de conceitos thema-rhema, tópico-comentário, foco-ênfase. Para Cresti, é necessário distinguir a natureza semântica do foco da natureza pragmática do comentário. Segundo a autora, a confusão se deve a falta da percepção da força ilocucionária do comentário (p. 61). Conforme a TLA, o tópico deve ser definido como o campo de ação da força ilocucionária do comentário (p. 49); tem uma dimensão pragmática, afetiva, portanto. Já o foco se relaciona com os aspectos cognitivos da interpretação do conteúdo linguístico, semânticos, portanto.

Esta distinção é a grande inovação no trabalho de Cresti: tanto o tópico como o comentário podem ter foco: “a Focus signals the apex of a semantic domain which develops a Topic or a Comment information function” (p. 71). Este foco é marcado prosodicamente por

meio de uma proeminência:

> "According to corpus data implying the consideration of sound counterpart, it seems useful to remember that a necessary feature of Focus is that it is marked by a prosodic prominence through different parameters. The most important are: a) pitch with a perceptually relevant F0 movement (rising-falling, or rising) or a strong modulation movement; b) duration with the lengthening of the syllables (plus a high intensity value). In all cases the seat of the prominence is the nucleus of the prefix PU or the root PU involved." (p. 72)

Ou seja, o foco é marcado por proeminência prosódica de maneira a indicar os elementos da sentença que exercerão o papel de tópico e comentário, mas não define qual será o papel do sintagma associado àquela unidade prosódica. Essa atribuição depende do contexto e da força ilocucionária. O foco pode ser marcado tanto por variações relevantes da F0 como por um alongamento da duração acompanhado de aumento de intensidade. Essa liberdade na maneira de marcar o foco chama a atenção, especialmente se considerarmos a expressão "strong modulation movement". Relembrando a ideia de que o comentário aplica a força ilocucionária sobre o tópico, podemos especular se estes movimentos melódicos que marcam o foco não poderiam qualificar, pela sua forma, diferentes atitudes e modalidades. O problema aqui é que falta à TLA uma descrição de como a força ilocucionária poderia depender da entoação. De qualquer forma, a proposta de Cresti parece sugerir que as tarefas de organizar a informação e indicar a força ilocucionária são simultâneas e independentes, o que nos levaria de volta à controvérsia de modelos por superposição e modelos lineares para a descrição fonológica da entoação. Voltaremos a este assunto mais adiante.

Para encerrar esta seção, destacamos a percepção de Cresti quanto aos limites de um modelo puramente linguístico para explicar o comportamento das relações entre pragmática e prosódia. Vejamos o trecho abaixo.

> "(...) the type of illocution depends on the affective disposition of the speaker toward the addressee; for instance, independently of what should be the content of an utterance, the same mental representation can be turned to the addressee as an order, a polite request, an instruction, a question, an invite, a suggestion, etc., following the kind of relationship occurring between the speakers. The type of the speaker's behavior depends directly on the affect motivating him. The psychic dynamics between speakers is the driving force of speech and it is continuously changing and becoming unpredictable." (p. 45)

Aqui é possível perceber a fronteira do campo de atuação da pragmática, que se limita com o estudo sociológico e psicológico dos usos da linguagem. Seguindo a ideia de que o que comanda a ilocução é a relação do falante com seu interlocutor, veremos a seguir um aprofundamento da noção de linguagem como ação, a partir da obra do sociólogo Pierre Bourdieu.

2.3. A economia das trocas linguísticas

Para que o ato de fala se realize plenamente é necessário que o interlocutor compreenda não só o conteúdo semântico e a estrutura sintática dos enunciados que ouve, mas também de que forma deve interpretá-los. Neste processo é fundamental perceber a intenção comunicativa de quem fala, associada logicamente ao contexto da conversa (ou do discurso). Apesar de haver algum grau de convenção neste jogo, não podemos falar de um código estabilizado; o que ocorre é um processo dinâmico, construído socialmente.

Bourdieu (1977) transpõe suas ideias acerca da troca de bens simbólicos – seu modelo para a compreensão das relações sociais – para o universo específico da linguagem, o que resulta numa crítica às tendências centrais da linguística que se balizam pela oposição entre *langue* e *parole* (Saussure) ou, mais recentemente, competência e performance (Chomsky). Para Bourdieu, a concepção usual de competência na linguística é por demais abstrata quando considera somente a capacidade de gerar infinitas combinações em um discurso gramaticalmente regular, independentemente do contexto, e não contempla a competência necessária que possibilita o uso adequado da competência puramente linguística. Desta forma, escapa ao discurso linguístico convencional a dimensão essencial, original, da linguagem que é, segundo o autor, ser um instrumento de ação (envolvendo necessariamente relações de poder) e não de compreensão. "Linguagem é uma práxis: é feita para dizer (…) [e] feita para dizer apropriadamente", ressalta Bourdieu, incluindo aí a decisão entre falar e calar, bem como escolha do estilo com que se fala (*op. cit.* p. 646).

O trecho a seguir ilustra bem a concepção da linguagem como instrumento de poder:

> "Assim, competência é também a capacidade de comandar o ouvinte. A linguagem não é somente um instrumento de comunicação, ou mesmo de conhecimento, mas também um instrumento de poder. Uma pessoa fala não somente para ser compreendida, mas também para ser crida, obedecida, respeitada, distinguida. Por isso [a necessidade de] uma definição mais abrangente de competência como sendo o direito de falar, i.e., a linguagem legitimada, a linguagem autorizada, que é, também, a linguagem da autoridade. <u>Competência implica o poder de impor recepção</u>. Aqui, mais uma vez, percebe-se a abstração da definição linguística de competência: o linguista considera as condições para o estabelecimento da comunicação como já asseguradas, quando, nas situações reais, esta é a questão essencial" (*idem*, *ibidem*, p. 648, grifo nosso).

Observamos aqui a crítica a uma concepção idealista da essência da linguagem, que seria uma gramática universal cuja estrutura captamos apesar das turbulências do uso prático da linguagem. Para Bourdieu é no seu uso que a linguagem se realiza plenamente e só aí pode ser plenamente compreendida, pois "a intenção expressiva, a maneira de realizá-la e as

condições para sua realização são indissociáveis" (*op. cit.*, p. 647). Daí se segue que "entender não é uma questão de reconhecer um significado invariável, mas de captar a singularidade de uma forma que só existe num contexto particular" (*loc. cit.*).

Bourdieu não pretende propor uma revisão da teoria linguística, mas antes chamar a atenção para equívocos que podem resultar da busca da compreensão de enunciados a partir da combinação de palavras, quando essas são compreendidas como uma forma invariante de inúmeras possibilidades de uso. Nesse sentido, o autor se aproxima de Bakhtin e também de Searle, na medida em que estes consideram o enunciado a unidade fundamental da língua, e não o morfema ou a palavra. Mesmo assim, Bourdieu considera que a explicação dos atos de fala em termos de uma "força ilocucionária" inerente ao enunciado também perde o dado fundamental que é o fato de que não é o enunciado que fala, mas um ator social no seu contexto particular. Se considerarmos, porém, a tendência contemporânea da aproximação entre pragmática e prosódia e a correspondente investigação de uma possível força ilocutória na realização acústica dos enunciados pelo falante, podemos dizer que este problema, pelo menos do ponto de vista teórico, está contornado.

Para nossa argumentação, duas ideias principais interessam aqui. A primeira está sintetizada na frase "competência implica o poder de impor recepção", e que, segundo nossa percepção, traz a realização acústica (prosódia e, principalmente, entoação) para o centro da atividade de linguagem, da língua em ação. Como consequência, a entoação não pode ser considerada como um acessório aplicado ao enunciado para "colori-lo", mas deve ser tomada como elemento essencial do uso da fala, localizado na origem da intenção comunicativa, tanto na fala do cotidiano como em situações formais, ritualizadas.

A segunda ideia é a noção do estilo, da maneira de falar, como a marca da autoridade que o falante se concede, ou se arroga, junto ao interlocutor. Vejamos o trecho abaixo.

> "Todo um aspecto da linguagem da autoridade não tem outra função senão enfatizar esta autoridade e dispor a audiência em acordo com a crença que é requerida. Neste caso, o estilo da linguagem é uma componente da parafernália imposta que serve para produzir ou manter a fé na linguagem. A linguagem da autoridade deve uma grande parte das suas propriedades para o fato de que tem de contribuir para sua própria credibilidade – p. ex. as elaborações estilísticas dos escritores literários (…) etc." (BOURDIEU, 1977, p. 649)

Apesar do termo "autoridade" remeter a uma relação unilateral, o autor está considerando aqui um acordo mútuo que legitima tanto o que está sendo dito quanto o como está sendo dito (se é que é possível separá-los, nessa perspectiva). Para nós, importa sobretudo reconhecer a função do estilo como elemento essencial, e não acessório, do enunciado. Reunindo as ideias que gravitam em torno dos conceitos de intenção, gênero e estilo dos

autores mencionados até agora neste capítulo, abriremos caminho para analisar o campo das manifestações vocais como o grande repositório de formas de contrato entre interlocutores que garantem a concordância quanto às trocas linguísticas.

Há, entretanto, uma outra ideia no texto de Bourdieu que é uma reflexão profunda sobre a natureza da linguagem; não tem relação direta com a questão dos gêneros e estilos, mas de alguma forma se relaciona com a hipótese central desta tese. Vejamos o trecho em que ela aparece pela primeira vez.

> "(...) os diferentes significados de uma mesma palavra não são percebidos como tais: somente a consciência educada que quebra a relação orgânica entre a competência e o campo [em que a palavra é empregada] expõe a pluralidade de sentidos, que não são possíveis de se captar na prática porque, na prática, a produção está sempre enraizada no campo de recepção." (*idem*, *ibidem*, p. 647)

Bourdieu chama a atenção para o fato de que é o uso de uma palavra (a relação da sua produção com o campo em que é empregada) é que produz o seu sentido, e não um "núcleo de significado" atrelado à sua forma, seu sinal acústico, o *signans*. A forma tem tão somente o papel de estabelecer conexões com outros usos na memória coletiva, e não pode ter um significado primordial do qual os demais derivam (vide a arbitrariedade do signo linguístico postulada por Saussure e que é um dos pilares da linguística moderna). A própria ideia de significado é um artifício que só cabe nos dicionários e nas gramáticas; o que existe é o sentido, dependente do uso e do contexto. Esta perspectiva se aproxima novamente da concepção do enunciado como sendo a unidade básica da língua; voltaremos a ela no quinto capítulo desta tese, quando reuniremos as evidências apresentadas nos quatro primeiros para argumentar em favor de nossa hipótese central.

2.4. Estética da criação verbal

Ao abordar a questão dos gêneros do discurso, Bakhtin (2003) propõe um modelo abrangente que pretende dar conta de todo tipo de manifestação da linguagem, desde o uso cotidiano, prático da fala, até o elaborado discurso científico, filosófico ou literário. Neste modelo, o enunciado é a unidade fundamental da comunicação discursiva.

> "O emprego da língua efetua-se na forma de enunciados (orais ou escritos) concretos e únicos, proferidos pelos integrantes desse ou daquele campo da atividade humana. Esses enunciados refletem as condições específicas e as finalidades de cada referido campo não só por seu

> conteúdo (temático) e pelo estilo da linguagem, ou seja, pela seleção dos recursos lexicais, fraseológicos e gramaticais da língua mas, acima de tudo, por sua construção composicional." (*idem*, *ibidem*, p. 261)

A partir desta concepção de enunciado, Bakhtin define os gêneros do discurso como sendo *tipos relativamente estáveis de enunciados* elaborados em cada campo de utilização da língua (*op. cit.*, p. 262). O autor também ressalta a extrema heterogeneidade dos gêneros do discurso, citando como exemplo as breves réplicas do diálogo, o relato, a carta, o comando militar, a ordem detalhada, os documentos oficiais, o texto publicitário, as publicações científicas e todos os gêneros literários. Cada um desses gêneros, dentre muitos outros que poderiam ser citados, abrangem uma infinidade de possíveis enunciados únicos e particulares, porém com características específicas que os ligam a cada gênero.

Bakhtin propõe uma primeira divisão dos gêneros em duas categorias: a dos gêneros primários (simples), que resultam da comunicação discursiva imediata, e a dos gêneros secundários (complexos) que surgem de um convívio cultural mais desenvolvido e organizado e aparecem predominantemente na forma escrita, como os romances, trabalhos científicos, peças jurídicas, etc. (*op. cit.*, p. 263). Essa divisão é necessária, segundo o autor, para que seja possível um aprofundamento da compreensão da natureza do enunciado, que é dificultada pela extrema heterogeneidade dos gêneros do discurso, já que "a própria relação mútua dos gêneros primários e secundários e o processo de formação histórica dos últimos lançam luz sobre a natureza do enunciado." (*op. cit.*, p. 264)

Desta forma, assim como se pode estudar os gêneros secundários do ponto de vista do estilo, o mesmo pode ser aplicado aos gêneros primários. Na fórmula sintética de Bakhtin, "onde há estilo há gênero." (*op. cit.*, p. 268). Esta associação entre estilo e gênero será útil para nossa argumentação, e será retomada adiante. Por ora passamos a analisar o que o autor fala a respeito da intenção do falante, que, veremos, se aproxima notavelmente do conceito de ilocução em Searle e de competência em Bourdieu.

Vejamos como a expressão aparece pela primeira vez no texto:

> "Em cada enunciado – da réplica monovocal do cotidiano às grandes e complexas obras de ciência ou de literatura – abrangemos, interpretamos, sentimos a *intenção discursiva* (…) ou a *vontade discursiva* do falante que determina o todo do enunciado, o seu volume e as suas fronteiras. Imaginamos o que o falante *quer* dizer, e com essa ideia verbalizada, essa vontade verbalizada (como a entendemos) é que medimos a conclusibilidade do enunciado." (*op. cit.*, p. 281, grifos do autor)

Essa conclusibilidade é condição necessária para que se possa responder ao enunciado, tomar uma atitude ou uma posição responsiva com relação a ele. Vemos aqui que nos mais diferentes usos da linguagem vigora uma lógica semelhante: a recepção do enunciado, a

resposta que o interlocutor dará, depende da percepção da intenção. Percebe-se portanto a semelhança com a dimensão ilocucional do ato de fala, que orienta o interlocutor na interpretação e no seu posicionamento com relação ao ato.

Bakhtin prossegue afirmando que "a vontade discursiva do falante se realiza antes de tudo na escolha de um certo gênero de discurso" (*op. cit.*, p. 282), ou, mais exatamente, como o autor colocará adiante, pela escolha de uma forma de gênero.

> "A intenção discursiva do falante, com toda sua individualidade e subjetividade, é em seguida aplicada e adaptada ao gênero escolhido, constitui-se e desenvolve-se em uma determinada forma de gênero. Tais formas[1] existem antes de tudo em todos os gêneros mais multiformes da comunicação oral cotidiana, inclusive do gênero mais familiar e do mais íntimo. Falamos apenas através de determinados gêneros do discurso, isto é, todos os nossos enunciados possuem *formas* relativamente estáveis e típicas de *construção do todo*" (*loc. cit.*, grifos do autor)

Nota-se aqui mais uma vez a relação entre a de intenção do falante e de conclusibilidade do enunciado, mediadas pelas formas dos gêneros do discurso. Para Bakhtin, essas formas são essenciais ao domínio da língua: "Se os gêneros do discurso não existissem (…), se tivéssemos que criá-los pela primeira vez no processo do discurso (…) a comunicação discursiva seria quase impossível." (*op. cit.*, p. 283) Esta posição levará o autor a declarar a primazia do enunciado sobre a oração e a palavra como unidade fundamental da comunicação linguística.

Assim como Bourdieu, Bakhtin considera que o conteúdo puramente linguístico (a oração, a frase) não tem um sentido concreto, por que é incapaz de, per se, "determinar imediata e ativamente a posição responsiva do falante" (BAKHTIN, 2003, p. 287). Para se tornar um enunciado, a oração ou a frase precisam incorporar elementos não-gramaticais que indiquem a intenção do falante com relação ao conteúdo e a seu interlocutor, e também a relação subjetiva emocionalmente valorativa do falante com o conteúdo do objeto e o sentido do enunciado. Um dos meios de que dispomos para expressar essa relação é o que o autor denomina "entoação expressiva"[2]. Bakhtin não chega a definir o que seja a entoação expressiva, mas podemos perceber que a origem do conceito é a entoação presente na fala que, segundo ele realizamos mentalmente na leitura muda de um discurso escrito (*op. cit.*, p. 290).

Aqui é necessário ter cautela com a homonímia entre dois conceitos diferentes. Para Bakhtin, a entoação expressiva não faz parte do sistema da língua; é um traço constitutivo (e,

---

1 Na tradução brasileira consultada, no lugar de "forma", está "gênero", o que criaria uma redundância; presumimos portanto um equívoco na revisão.

2 Entre os estudiosos de Bakhtin no Brasil é costume grafar-se "entonação"; para manter a coerência com o restante do trabalho, mantivemos nossa grafia.

acrescentamos, exclusivo) do enunciado. Pelos exemplos que emprega da expressão atribuída pela entoação, percebe-se que ele tem em mente basicamente as atitudes (seco, respeitoso, caloroso, sarcástico, etc.) e emoções (medo, alegria, raiva, e assim por diante). Esta concepção se deve provavelmente ao preconceito que tradicionalmente relegava a prosódia e a entoação ao domínio dos fenômenos paralinguísticos. Pelo que vimos no primeiro capítulo, já está pacificada a questão de que a prosódia tem implicações sintáticas e pragmáticas, e que estas se dão de maneira sistemática, mesmo que ainda não possamos descrevê-lo de maneira satisfatória. O próprio Bakhtin admite a existência de uma "entoação gramatical específica", que envolve finalizações, subordinações, enumerações, etc (p. 296); também se refere a modalidades (entoação narrativa, interrogativa, exclamativa e exortativa), situando-as no cruzamento entre a entoação gramatical e a entoação típica do gênero. Porém, seu interesse primordial no enunciado faz com que esses aspectos ligados ao sistema da língua não ganhem destaque.

Feita essa ressalva, apontamos a relação que Bakhtin estabelece da entoação expressiva com o estilo e o gênero. De acordo com o autor, a vontade discursiva individual se manifesta na escolha de um determinado gênero e de uma entoação expressiva, e certos gêneros requerem determinados "tons", isto é, incluem na sua estrutura uma determinada entoação expressiva (p. 284). Em outra passagem, Bakhtin afirma que "o estilo individual é determinado principalmente pelo seu aspecto expressivo" (p. 289). Percebemos aqui uma relação triádica entre gênero, estilo e entoação expressiva, os três se implicando e se influenciando mutuamente na concretização do enunciado segundo a intenção ou vontade discursiva do falante.

Um outro aspecto importante da análise do discurso empreendida por Bakhtin é o destaque que dá à relação de cada enunciado com os demais na cadeia da comunicação discursiva. Mesmo tendo limites determinados pela alternância dos sujeitos, cada enunciado reflete o processo do discurso e sobretudo os elos precedentes da cadeia (p. 299). Isso faz com que um outro traço essencial, constitutivo do enunciado seja o seu *direcionamento* a alguém, o seu *endereçamento* (p. 301). Todo enunciado tem um autor e um destinatário, que pode ser conhecido ou desconhecido, individual ou coletivo. Escreve Bakhtin:

> "A quem se destina o enunciado, como o falante (ou o que escreve) percebe e representa para si os destinatários, qual é a força e a influência deles no enunciado – disto dependem tanto a compisição quanto, particularmente, o estilo do enunciado. Cada gênero do discurso em cada campo da comunicação discursiva tem a sua concepção de destinatário que o determina como gênero."(*loc. cit.*)

Ou seja, a antecipação que o falante faz da reação do interlocutor ao seu enunciado

influencia o estilo do enunciado. Isso é particularmente notável em situações de fala que envolvem diferentes hierarquias e graus de polidez, mas também pode ser observado em gêneros do discurso escrito. Transpondo esta constatação para nossa percepção de uma situação de fala, é evidente, apesar de não demonstrado experimentalmente, a mudança de estilo melódico de acordo com o interlocutor na fala espontânea. Aliás, a própria presença de um interlocutor faz toda a diferença na realização prosódica – por mais paradoxal que pareça, é difícil ser espontâneo na fala quando se está sozinho.

Para encerrar esta seção, apresentamos um trecho que anuncia uma inversão de perspectiva, um radical posicionamento epistemológico de que nos aproximaremos gradualmente ao longo da tese.

> "Quando escolhemos as palavras, partimos do conjunto projetado no enunciado, e esse conjunto que projetamos e criamos é sempre expressivo, e é ele que irradia a sua expressão (ou melhor, a nossa expressão) a cada palavra que escolhemos; por assim dizer, contagia essa palavra com a expressão do conjunto." (p. 291)

Na nota de rodapé a este trecho citado, presente na edição russa, o autor voa mais alto:

> "Quando construímos o nosso discurso, sempre trazemos de antemão o todo da nossa enunciação, na forma tanto de um determinado esquema de gênero quanto de projeto de individual de discurso. Não enfileiramos as palavras, não vamos de uma palavra a outra, mas é como se completássemos com as devidas palavras a totalidade." (p. 292)

Antevemos aqui uma formulação possível para um dos aspectos da nossa hipótese central que é a precedência da prosódia na estruturação do enunciado. Apesar das incompatibilidades terminológicas e conceituais da obra de Bakhtin com o contexto atual dos estudos prosódicos, percebemos aqui uma visão profunda da natureza da linguagem, que questiona uma concepção combinatória da linguagem, trazendo a discussão para o campo da ação, da fala em ato. Por falta de conhecimentos mais precisos sobre a natureza dos fenômenos prosódicos, Bakhtin lançou-se intuitivamente na reflexão sobre as relações entre gênero, estilo e entoação, mas mesmo assim chegou a questionamentos que se situam nas fronteiras do conhecimento linguístico.

2.5. Pragmática da criação literária

Maingueneau (1996) faz uma interpretação da teoria dos atos de fala, aplicada à teoria literária, centrada na ideia de força ilocutória. Essa força ilocutória "indica que tipo de ato de linguagem é realizado quando se enuncia, como ele deve ser recebido pelo destinatário"

(MAINGUENEAU, 1996, p. 7).

A partir do que foi exposto sobre a perspectiva dos estudos prosódicos com respeito aos atos de fala, propomos o conceito de vetor ilocutório, constituído de marcas percebidas na curva de frequência, que indica como o falante pretende que o enunciado que ele profere seja interpretado. A proposta de substituir "força" por "vetor" deve-se tão somente a uma possível conotação metafísica da palavra "força"; o termo "vetor", por sua vez, está associado a uma descrição matemática, baseada em parâmetros ou coordenadas, mais próxima, portanto, da abordagem que encontramos na pesquisa contemporânea sobre a entoação. Este vetor ilocutório investe o enunciado de um valor performativo que orienta o ouvinte na interpretação.

Maingueneau também aponta para uma conexão entre a dimensão ilocutória e os gêneros do discurso, que ele denomina "macro-atos de linguagem": "Desse modo, interpretar corretamente um provérbio é ver nele não apenas uma asserção (…), mas também um gênero de discurso específico ao qual corresponde um macro-ato específico" (op.cit., p. 14). Mais adiante, encontramos: "a partir do momento em que identificou a que gênero pertence um texto, o receptor é capaz de interpretá-lo e comportar-se de modo adequado a seu respeito. Na falta disso, pode ocorrer uma verdadeira paralisia" (idem, p. 15).

Podemos deduzir, portanto, a existência de um vínculo direto entre a dimensão ilocucional da enunciação ao gênero de discurso. Expandindo o conceito de gênero de maneira a abranger todas as manifestações vocais, inclusive aquelas consideradas como pertencentes ao campo da música, podemos nos propor a investigar o papel das características melódicas na determinação do gênero, buscando assim abordar o problema da dimensão ilocucional do ponto de vista de uma práxis sonora, mais especificamente vocal, que visa a influenciar o ouvinte, induzindo nele ações desejadas, transformando assim a realidade. Numa aproximação com o campo da etnomusicologia, desenvolveremos no capítulo II, um estudo sobre gênero, estilo e melodia nas manifestações vocais.

Resumindo este primeiro capítulo, podemos afirmar que é possível perceber simultaneamente, numa situação de fala espontânea, a origem do falante (sotaque), o seu estado emocional, a estrutura sintática do enunciado (coordenações e subordinações as marcas de interação e a segmentação em unidades de informação, a intenção de comunicação (modo) e a atitude proposicional. Considerando cada um desses elementos como sendo transportado por uma sequência de alterações na curva melódica que se sobrepõe, podemos pensar na entoação como uma verdadeira polifonia em que cada voz desempenha o papel de canal de informação para cada uma das funções. Nesse contexto o ouvinte distingue as variações

prosódicas acidentais (inclusive as lexicais, no caso das línguas tonais) das inflexões características da curva melódica que denotam uma intenção comunicativa. Para essas últimas estamos propondo a expressão "vetor ilocutório". Esse vetor ilocutório determina como o ouvinte interpretará o enunciado – se é uma pergunta, uma asserção, uma ordem, uma sugestão, um pedido, uma declaração (performativa). Está relacionado, portanto, a uma intenção persuasiva com relação ao interlocutor. Independentemente da estrutura mais detalhada da classificação dos funções da entoação, acreditamos ser possível presumir uma distinção fundamental entre as componentes contingentes ou circunstanciais das condições da enunciação daquelas que resultam de uma intenção comunicativa determinada do falante, essas ligadas ao conceito de "tom" no senso comum.

2.6. Gesto musical: ação e significação

A gesticulação é um elemento fundamental da comunicação humana, embora muitas vezes seja difícil definir o que o gesto parece expressar ou comunicar. O fenômeno da gestualidade no contexto da comunicação verbal está associado ao campo da paralingüística. Já o estudo da gestualidade dentro de um horizonte mais amplo, que incluiria "textos gestuais" autônomos como, por exemplo, as saudações, a dança, a pantomima, a ginástica e a acrobacia, competiria à cinésica (ECO, 1987, p. 393). No entanto, estas áreas de investigação ainda não se encontram completamente sistematizadas como disciplinas autônomas, motivo pelo qual preferimos nos ater a trabalhos que estudem o gesto e a gestualidade dentro do quadro de uma semiologia geral, inserindo-os no fenômeno global da comunicação humana.

A questão da significação na comunicação gestual apresenta grandes desafios às teorias semiológicas, por várias razões. Primeiramente, porque se encontra na intersecção dos conjuntos dos fenômenos naturais (os movimentos inerentes ao corpo humano) e dos fenômenos culturais (os movimentos do corpo humano, aprendidos e transmitidos socialmente); em segundo lugar, porque, apesar da estreita ligação com a linguagem falada, não podemos falar de um código propriamente dito, uma vez que as variantes individuais apresentam uma enorme diversidade, e seu sentido, por sua vez, pode mudar radicalmente conforme o contexto; terceiro – que é, em parte, conseqüência do anterior – é que a gestualidade, enquanto fenômeno independente, transmite essencialmente a conotação (atributos, modalidades, estados psicológicos), ficando a denotação dependente de um código intermediário (linguagem verbal) que é acionado por gestos que apontam para unidades de

significados da língua.

Greimas (1968) empreendeu um importante estudos sobre o gesto, procurando situá-lo no contexto de uma divisão entre uma semiótica do mundo natural e outra do mundo cultural ou humano, e avaliando a possibilidade de uma comunicação por meio do gesto. "Conditions d'une sémiotique du monde naturel" é o título do artigo que aponta para um problema de grande abrangência, qual seja, a distinção entre os elementos naturais e culturais de um dado sistema semiótico, quando quase dois terços do texto se dedicam especificamente ao estudo da gestualidade. Na verdade Greimas analisa o problema utilizando o sistema da gestualidade como uma espécie de estudo de caso da relação entre cultura e natureza num sistema semiótico, e aí tece considerações de grande profundidade sobre a questão da significação em um contexto quase que "a-semântico".

Segundo Greimas, a "substância" da expressão gestual é o movimento do corpo humano, e este é um elemento do mundo natural, o que torna o gesto, a princípio, um signo natural. A transposição do gesto para o mundo significante exige a mediação de figuras – estáticas e dinâmicas – que são empregadas para enquadrar em um número limitado de categorias os infinitos objetos (incluindo processos) do mundo natural. Este conjunto catalogado de figuras "constitui o corpus a partir do qual um código semiótico de expressão visual pode ser construído" (GREIMAS, 1968, p. 8, tradução nossa). Podemos perceber aqui uma ligação com a psicologia da forma, segundo a qual um número finito de configurações ou formas básicas (Gestalten) é empregado na categorização dos infinitos objetos do mundo sensível. Estas configurações não são, portanto, entidades autônomas, mas sim atuam na mediação entre o significante e o significado gestual.

Greimas define a gesticulação como "uma empresa global do corpo humano na qual os gestos particulares dos agentes corporais são coordenados ou subordinados a um projeto de conjunto se desenvolvendo [*en déroulant*] em simultaneidade" (ibidem, p. 12, tradução nossa). Esta definição se aproxima de uma concepção mecanicista do gesto; note-se que não se faz menção a uma comunicação de um significado, mas aparece, em estado embrionário, a idéia de intenção (na expressão "projeto de conjunto"). Em conseqüência, a definição acaba por abranger desde os gestos que acompanham a fala e as tarefas manuais até os "textos gestuais" artísticos e religiosos. Entretanto, Greimas ressalta que uma abordagem puramente mecanicista não pode dar conta da interpretação do fenômeno da gestualidade como um todo. A gesticulação natural, ao participar da comunicação entre os indivíduos de uma sociedade, passa inevitavelmente a ser um fenômeno social, e podemos então falar da gestualidade como uma dimensão semiótica da cultura.

Para Greimas, a presença do sentido depende da percepção de uma direção na seqüência de movimentos daquele que realiza a gesticulação. A partir do momento em que exista um observador que interprete e desenvolva uma expectativa do que o outro faz ou parece querer fazer a partir de seus "enunciados gestuais", forma-se assim um par emissor-receptor e com ele as bases para que a gesticulação possa se tornar um código. Os enunciados gestuais, neste caso, compreendem desde uma complexa coreografia até seqüências de movimentos que contêm um programa (como vestir-se, por exemplo). O fato de que as seqüências de gestualidade prática sejam transmitidas pelos processos de aprendizagem demonstra a presença de significação nestas seqüências. Ao mesmo tempo, toda aprendizagem envolve alguma forma de metalinguagem (uma explicação verbal de partes do enunciado), o que confirma a possibilidade de segmentá-las. Entretanto, estas unidades mínimas portadoras de significação não podem ser distinguidas sem que se recorra à semântica das línguas naturais.

Greimas argumenta que esta dificuldade é insuperável, uma vez que nos enunciados gestuais, o homem é, ao mesmo tempo, sujeito do enunciado (aquele de quem se diz) e sujeito da enunciação (aquele que diz), o que impede o código gestual de atuar como metalinguagem, devido à confusão entre os níveis lingüísticos. Para superar este impasse, o autor recorre à sugestão de Julia Kristeva de substituir a categoria da comunicação por uma mais abrangente, que seria a da produtividade, se estamos interessados numa semiótica que possa dar conta da interpretação da totalidade dos comportamentos humanos. De fato, a comunicação é um comportamento humano que visa a transformar o mundo, assim como o trabalho; incluindo-se ambos na categoria mais geral da produtividade, podemos estudar num mesmo quadro teórico as práticas cinésicas (trabalho) e o discurso gestual (comunicação). Vista como produtividade, a gesticulação se organiza através de algoritmos e estratégias, que permitem segmentar o continuum gestual e de onde emerge a noção de intencionalidade, que dá o caráter significativo ao sintagma gestual. Nesse contexto, Greimas define então a "práxis gestual" como sendo "a utilização, pelo homem, de seu próprio corpo visando à produção de movimentos organizados em programas que têm um projeto, um sentido comum" (GREIMAS, 1968, p. 18, tradução nossa). Esta práxis gestual se divide em uma gestualidade prática, ligada ao fazer, e uma gestualidade mítica, ligada ao querer. Note-se que a intenção ou desejo de transformação do mundo é comum a ambos. Greimas assinala que a divisão da práxis entre os dois campos é desigual e variável, assim como sua fronteira, que varia de cultura para cultura, e admite, além disso, a existência de formas mistas, em que o mítico pode ser encontrado em meio ao prático e vice-versa. Segundo o autor, uma característica

fundamental da gesticulação mítica é, além da intenção de comunicar, a intenção de transformar o conteúdo que elas exprimem.

Podemos mais facilmente compreender a origem mítica em algumas formas da gesticulação (as que Greimas denomina lúdicas) se lembrarmos que os elementos constitutivos de seus enunciados gradativamente se dessemantizam pela repetição e transmissão de seus algoritmos. Assim, Greimas sugere que a gesticulação puramente estética, como o balé, é resultante de uma dessacralização do discurso mítico, que tem como conseqüência a dessemantizacão dos enunciados, restando-lhe somente as formas narrativas do discurso: "ao constituírem códigos de comunicação de conteúdos míticos, as formas gestuais se distanciam da comunicação lingüística e readquirem uma nova consistência, graças à aparição do princípio de organização funcional e narrativa que rege todos os discursos, sejam eles da ordem do fazer ou do dizer" (op. cit., p. 29, tradução nossa). Vê-se aqui uma espécie de equivalência entre discurso e ação, fazer e dizer pertencem ambos à categoria da práxis.

Podemos afirmar que a gesticulação que acompanha a fala constitui uma intersecção entre as áreas de atuação de ambas, uma vez que o gesto também atua como traço supra-segmental que participa da comunicação linguística, conferindo conotações aos enunciados. Por outro lado, podemos nos reportar a esta correspondência entre a gesticulação que acompanha, ou melhor, que integra o fenômeno da fala, e a entoação, passando então a considerar esta última como uma espécie de "gesto musical". Vista assim, a entoação pode ser percebida como uma espécie de equivalente da gesticulação que ocorre num campo perceptivo associado à emissão da fala. Certamente, se pensarmos na definição de Greimas para a gesticulação (uma empresa global do corpo humano que coordena e subordina ações simultâneas a um projeto de conjunto), não há por que não pensar na fala e na gesticulação como componentes de uma mesma "gestualidade" global, intrínseca à faculdade linguística, compartilhando, portanto, um mesmo programa e uma mesma intencionalidade que são revelados através de sua forma.

Esta homologia nos leva a considerar a entoação como parte de um conjunto maior de manifestações vocais-verbais, que empregam diferentes gestos musicais (variações ritmadas da F0) para transmitir diferentes intenções. Suas características melódicas são, portanto, determinadas pela intenção comunicativa de quem emite ou profere aquele enunciado, que estará associada ao gênero do discurso (aqui no sentido mais amplo possível) e o seu estilo correspondente. Nesta perspectiva, encontramos toda uma gama de entoações possíveis, desde as falas mais espontâneas até as falas mais ritualísticas, observadas na religião e na arte.

3. Fala e canto

3.1. Representação musical da entoação

No seu álbum “Festa dos Deuses” (PASCOAL, 1992), o músico brasileiro Hermeto Pascoal incluiu três faixas que correspondem a exemplos do que ele chama de “som da aura”, que seria “a vibração sonora da alma de cada um, refletida pela sua fala, que faz a ligação entre mente e corpo” (PASCOAL, 2011). A partir de três gravações de voz falada (um discurso do presidente Fernando Collor, a declamação de um poema pelo ator Mário Lago, e a fala de uma professora em uma aula de natação para crianças), Hermeto gravou, usando um sintetizador, uma melodia que corresponde, sílaba por sílaba, à entoação do enunciado, de acordo com a sua percepção melódica. Em seguida, gravou um acompanhamento com outro instrumento harmônico. Assim, as três faixas têm a mesma forma: na primeira metade aparece somente a voz, e na sequência a voz dobrada pelo teclado e com acompanhamento harmônico (no caso do poema, há também um acompanhamento de percussão em alguns momentos).

O primeiro registro dessa ideia aparece no álbum “Lagoa da Canoa, Município de Arapiraca” (1984), em que Hermeto empregou gravações dos locutores esportivos Osmar Santos e José Carlos Araújo; também foi realizada em outras situações tais como entrevistas em televisão e documentários, inclusive com enunciados em outras línguas. Conta Hermeto que

> “Aos 7 anos de idade descobri que a nossa fala é o nosso canto. O mais natural de todos, pois cada fala é uma melodia. Eu costumava dizer para minha mãe que ela e suas amigas estavam cantando quando conversavam, mas ela dizia:"Deixe disso, menino! Você está ficando louco?"” (PASCOAL, 2011)

Hermeto possui o que os músicos chamam de ouvido absoluto, que é a habilidade de reconhecer a classe de altura de um som (se é um dó, um ré, etc.) imediatamente, sem a ajuda de uma referência, como um instrumento ou um diapasão. Ainda não há uma explicação satisfatória sobre como esta habilidade é adquirida, mas sabe-se é bastante rara, aparecendo em menos de uma a cada dez mil pessoas (para uma descrição detalhada do fenômeno e possíveis explicações, cf. DEUTSCH, 2006).

Mais do que o fenômeno do ouvido absoluto, o que chama a atenção nestas gravações é o fato de que Hermeto resolveu de uma maneira peculiar, porém aparentemente sistemática, o problema da atribuição de uma altura a uma sílaba. Sabe-se que a F0 apresenta pequenas

flutuações ao longo da vogal central de uma sílaba, enquanto que um instrumento musical como o piano só pode produzir sons com frequências fixas, pertencentes à escala cromática. Aparentemente, Hermeto "arredonda" as alturas para a escala cromática temperada, qual a relação da frequência com o ritmo percebido pelo músico, entre outras.

Por meio de uma análise de um trecho de uma das faixas, encontramos evidências de que as alturas escolhidas de ouvido por Hermeto para cada sílaba se aproximam consistentemente dos valores correspondentes ao que Xu (2009) denomina alvos estáticos (SOUZA, 2011). Reproduzindo-se a parte gravada do teclado em velocidade reduzida, foi possível, com o auxílio de um piano, identificar com segurança as notas escolhidas por Hermeto para cada sílaba e, a partir disto, elaborar uma partitura da melodia executada pelo teclado (fig. 1).

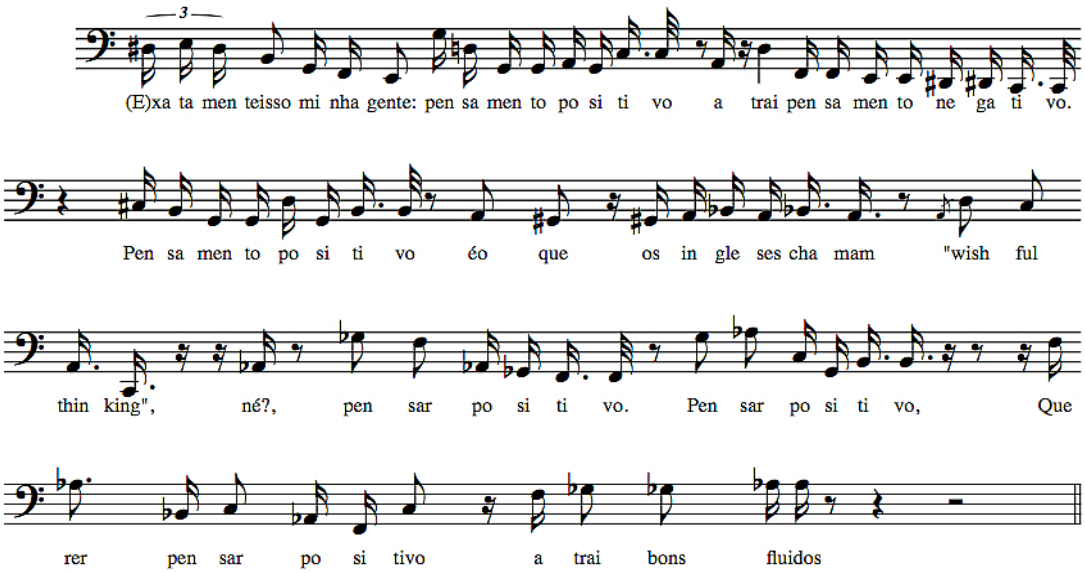


Figura 1 – Transcrição melódica de "Pensamento positivo" (PASCOAL, 1992)

O conceito de alvo estático se assemelha bastante ao que em música se chama portamento: uma nota que se inicia com uma altura ligeiramente mais baixa (ou, mais raramente, mais alta) que a esperada e aos poucos se aproxima da altura "real". O ouvinte, neste caso, não interpreta isso como uma desafinação, mas sim como um efeito expressivo, e entende como altura real aquela à qual o intérprete se dirige (o alvo). Tudo indica que Hermeto interpretou as variações intrassilábicas da F0 como portamentos.

Estas experimentações de Hermeto Pascoal remetem a questões importantes dos estudos prosódicos quanto ao comportamento da curva de F0. Por exemplo, concordam com

descobertas recentes de que os intervalos empregados na fala são múltiplos inteiros de um semitom (ROSS et al., 2007). Também levantam implicitamente a hipótese de que se pode associar uma altura definida, principal, representativa de uma sílaba, mesmo que a F0 varie ao longo desta. De qualquer forma, estas gravações oferecem uma situação privilegiada para que se observem as relações entre os modos de escuta de fala e música.

Outro músico que se notabilizou pelo interesse na melodia da fala foi o compositor checo Leoš Janáček (1854-1928). Janáček deixou aproximadamente 10.000 páginas de anotações em cadernos e, destas, cerca de metade contém o que ele que denominava *napěvky mluvy*, que poderíamos traduzir por "fragmentos melódicos da fala"[3] (PEARL, 2005, p. 37). Este interesse começou a partir do convite de seu amigo linguista František Bartoš para auxiliá-lo no registro de dialetos, e tornou-se uma obsessão. O principal objetivo era a princípio preservar a sonoridade da língua checa para as gerações futuras, e chegou a considerar um dicionário de melodias características (*idem*, *ibidem*, p. 74). Há evidências de que ele empregou as ideias melódicas anotadas em suas composições, especialmente nas óperas, mas ainda não há estudos suficientes para afirmá-lo com segurança.

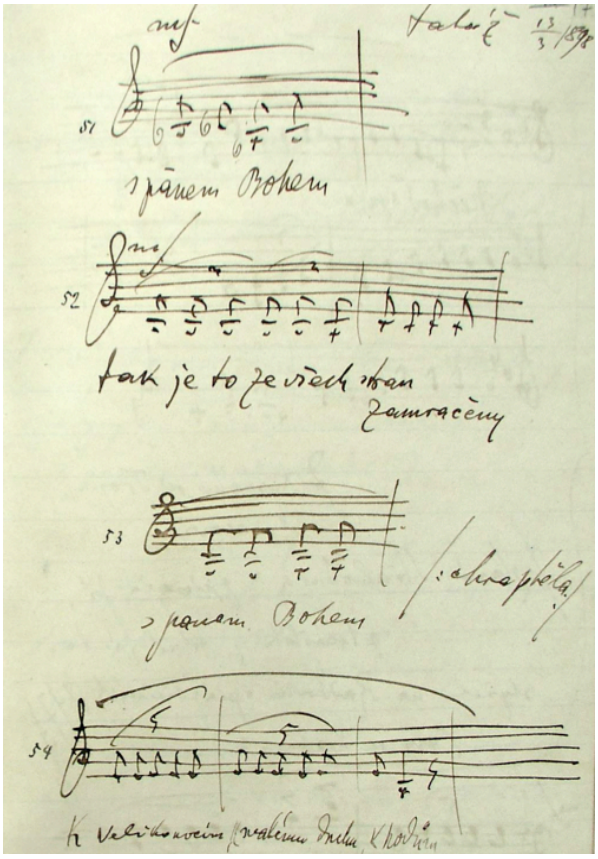


Figura 2 – Página de um caderno de anotações de Leoš Janáček (Fonte: PEARL, 2005)

3 Pearl traduz por "tunelets of speech"

Nas anotações de Janáček percebe-se que ele considerava a prosódia como sendo melodia, passível de ser registrada por meio da notação convencional. Nesse processo, o compositor simplificava os movimentos micro-melódicos (glides), que normalmente se observam no interior da sílaba, para alturas fixas. Segundo Pearl (*op. cit.*, p. 117) isso se devia a uma atitude perceptiva do compositor, decorrente do seu treinamento musical. Infelizmente, não há nenhum registro sonoro dos enunciados que Janáček anotou e não é possível estudar se houve algum método ou critério nesta transposição de alturas variáveis para alturas fixas em cada sílaba. Reproduzimos abaixo alguns exemplos editorados por Pearl (*op. cit.*), junto a fac-símiles das anotações originais (fig. 3); os códigos acima da partitura correspondem à catalogação dos manuscritos nos Arquivos de Música do Museu Morávio em Brno, na República Checa.

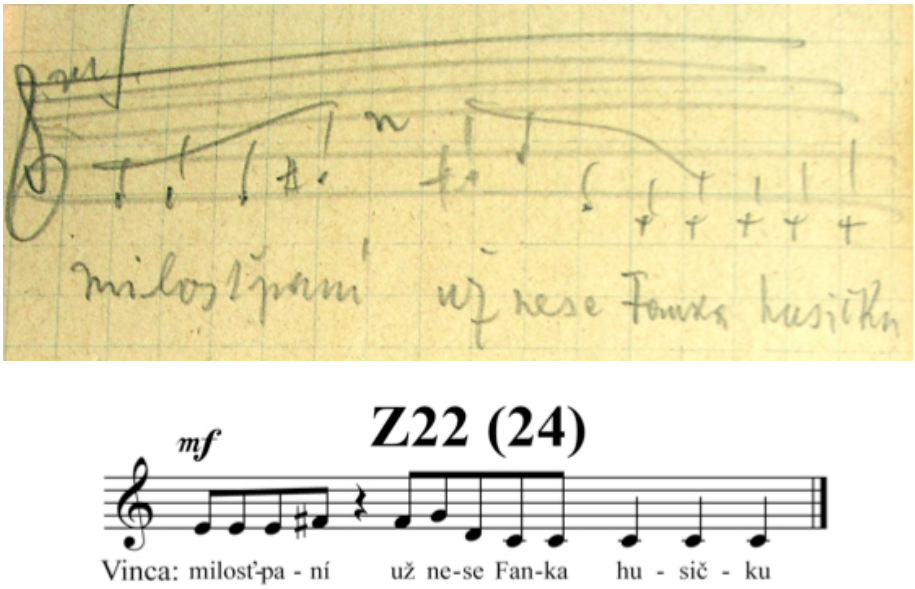


Figura 3 – Exemplo de anotação e correspondente editoração (Fonte: PEARL, 2005)

Janáček anotou inúmeras versões ouvidas de fórmulas típicas do cotidiano, como saudações e despedidas, registrando com detalhes as particularidades de cada falante (fig. 4).

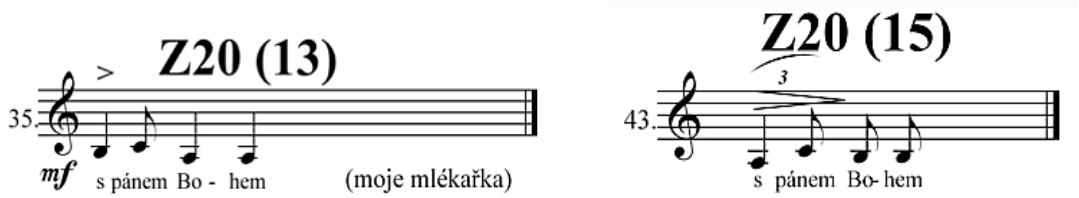


Figura 4 – Diferentes versões de uma fórmula de despedida (*s pánem Bohem*, em tradução livre "Fique com Deus!") anotada por Janáček (Fonte: PEARL, 2005)

O processo de transformação da entoação em uma melodia parece ser o mesmo tanto no caso de Hermeto Pascoal como no de Leoš Janáček. Ambos atribuem a cada sílaba uma altura do sistema fixo temperado de doze notas cromáticas, provavelmente a altura que consideram mais representativa daquela sílaba naquele contexto. A sua proximidade com a linguística (Janáček chegou a publicar um artigo intitulado "A fronteira entre fala e canção"[4]) faz pensar que o compositor checo estava em busca de alguma lógica na melodia da fala, e que ele considerava, mesmo que intuitivamente, a melodia da fala como sendo parte do sistema da língua; já Hermeto parece abordar a questão mais como um jogo, um exercício das suas notáveis habilidades musicais, a serviço de uma caracterização espiritual do falante (o conceito de "aura"). Se não trazem nenhuma contribuição direta para a compreensão dos fenômenos associados à entoação, o legado destes dois músicos é certamente um terreno fértil para estudos futuros, não só das relações entre linguagem e música, mas também da cognição humana.

Um outro caso interessante a ser citado é o estudo empreendido por Joshua Steele (1779), intitulado "*Prosodia Rationalis: or, an essay towards establishing the melody and measure of speech to be expressed and perpetuated by peculiar symbols*". Steele pretendia com seu ensaio contestar as ideias apresentadas num livro publicado anteriormente por James Burnet, que sustentava que a altura não varia durante a fala (HATFIELD, 2010). Para demonstrar sua tese, Steele desenvolveu um sistema de notação peculiar que descreve com detalhes a realização prosódica dos enunciados. Seu sistema é baseado na notação musical convencional, mas amplia as possibilidades de representação da altura, incluindo os quartos de tom (metade de um semitom). Este fato é notável para a época, e demonstra uma atitude perceptiva diferente dos dois músicos citados no início desta seção. Steele se preocupou em anotar nuances da curva de F0 no limite do perceptível, numa resolução mais fina do que a música da época empregava. Além do aspecto musical, também chama a atenção o fato de que o autor considerava essas nuances significativas para o efeito da entoação, pelo menos do ponto de vista expressivo. Vejamos os elementos de seu sistema de notação.

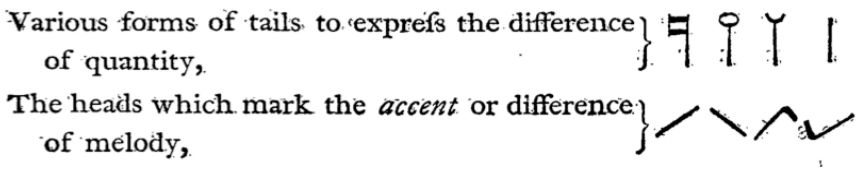


Figura 5 – Explicação de Steele para os símbolos empregados (STEELE, 1779)

4 Na tradução em inglês, "The borderline of speech and song" (PEARL, 2005, p. 266)

Baseando-se nas notas musicais Steele empregou símbolos para o movimento melódico no lugar das cabeças das notas, e símbolos para a duração no lugar onde normalmente ficam os colchetes (fig. 5). Note-se que ele usa o termo *accent* para as variações melódicas, o que evidencia o papel da altura na realização do acento (cf. supra, capítulo I, seção 1.2?). Estas "notas" eram então dispostas num pentagrama com uma subdivisão que permitia representar quartos de tom (fig. 6)

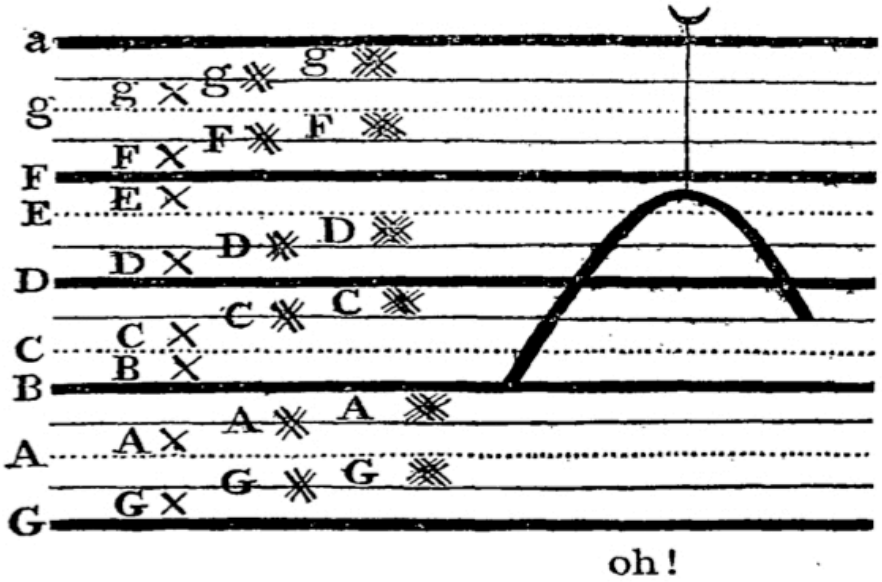


Figura 6 – Pauta empregada por Steele para registrar as alturas (1779)

Note-se que, por exemplo, entre as linhas que correspondem às notas sol (G) e sol# (G#), que estão à distância de um semitom, há um espaço que representa uma altura um quarto de tom acima de sol, que o autor representa com o símbolo Gx.

Steele empregou um método engenhoso para determinar com exatidão as trajetórias dos picos dos movimentos circunflexos (rise and fall): fez colar uma folha de papel sobre o braço de uma viola da gamba com uma escala graduada indicando os semitons e quartos de tom, de maneira que o dedo do instrumentista pudesse deslizar livremente entre os semitons. Com este artifício ele pode identificar o ponto de partida e chegada dos movimentos, bem como seus eventuais picos.

Por meio deste método ele pode demonstrar não só que a fala apresenta variações de altura, mas também que há variação tanto ao longo do enunciado como dentro da sílaba e que essas variações são consideráveis, da ordem de uma quinta (variação de cerca de 50% na F0 para cima ou 33% para baixo).

Posicionando as notas na pauta, com as palavra abaixo, do mesmo modo que é feito na

notação da música vocal, é possível grafar a notação de um enunciado (fig. 7).

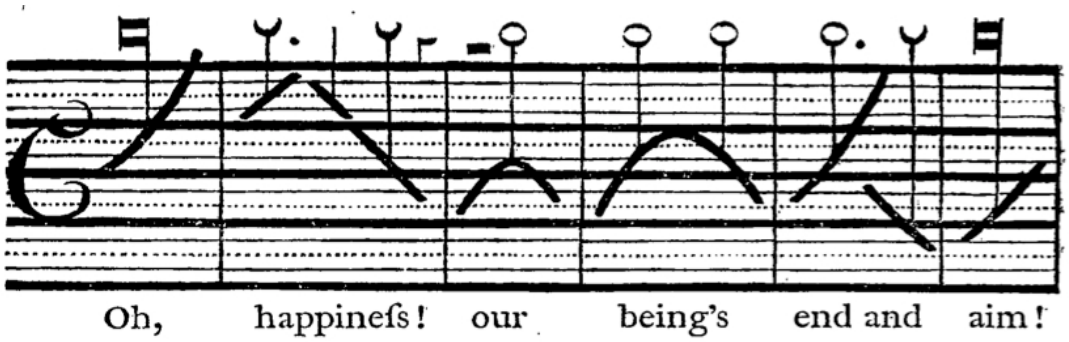


Figura 7 – Exemplo da notação de Steele para a entoação (STEELE, 1779)

Depois de apresentar sua proposta de notação, Steele passa a refletir sobre a natureza e a composição da melodia da fala, sempre de um ponto de vista musical. Steele resume:

> "(…) *there are five orders of accidents incident to melody and measure, essentially* different in their nature and effects *from each other, and* very material *to be attended to in the consideration of the melody and measure of speech.*"[5] (STEELE, p. 23, grifos nossos)

Logo abaixo, Steele lista os "acidentes":

> 1ſt, ACCENT. Acute ⟋ grave ⟍, or both combined ⋀⋁; in a variety of circumflexes.
> 2dly, QUANTITY. Longeſt, long, ſhort, ſhorteſt.
> 3dly, PAUSE or *ſilence*. Semibrief reſt, minim reſt, crotchet reſt, quaver reſt.
> 4thly, EMPHASIS or *cadence*. Heavy, light, lighteſt.
> 5thly, FORCE or *quality of ſound*. Loud, louder, ſoft, ſofter. Swelling or increaſing in loudneſs, decreaſing in loudneſs or dying away. Loudneſs uniformly continued.

Figura 7 – As cinco ordens de acidentes da prosódia, segundo Steele (1779)

A partir das características da sua notação, podemos inferir que Steele entende a prosódia como sendo constituída de cinco parâmetros ou propriedades (os acidentes). Chama a atenção a percepção da independência dos parâmetros prosódicos quanto à sua natureza e função e de sua materialidade acústica. É também notável a distinção que é feita entre acento, quantidade e ênfase (esta última não corresponde ao sentido atual que damos; está ligada, como veremos, ao conceito de posição métrica). E ficamos ainda mais impressionados ao

5 "existem cinco ordens de acidentes que ocorrem na melodia e no compasso, essencialmente diferentes entre si na sua natureza e efeitos, e verdadeiramente materiais para serem atentados em consideração à melodia e ao compasso da fala"

constatar que trata-se da primeira obra em língua inglesa (e possivelmente em qualquer outra) dedicada exclusivamente à entoação (HATFIELD, 2010), abordando questões atuais e ainda não resolvidas da linguística.

Além de contestar a tese de Burne, outro propósito de Steele com seu ensaio era propor um método de declamação com o apoio de um instrumento grave, a viola da gamba baixo[6], que serviria de referência para que os atores executassem a melodia de maneira mais segura e precisa. Uma boa parte do livro se dedica a explicar como isso se daria, com exemplos tirados de textos célebres, como ilustra o exemplo abaixo, extraído do célebre monólogo de Hamlet (fig. 8). Infelizmente não é possível saber como Steele pretendia que sua proposta soasse.

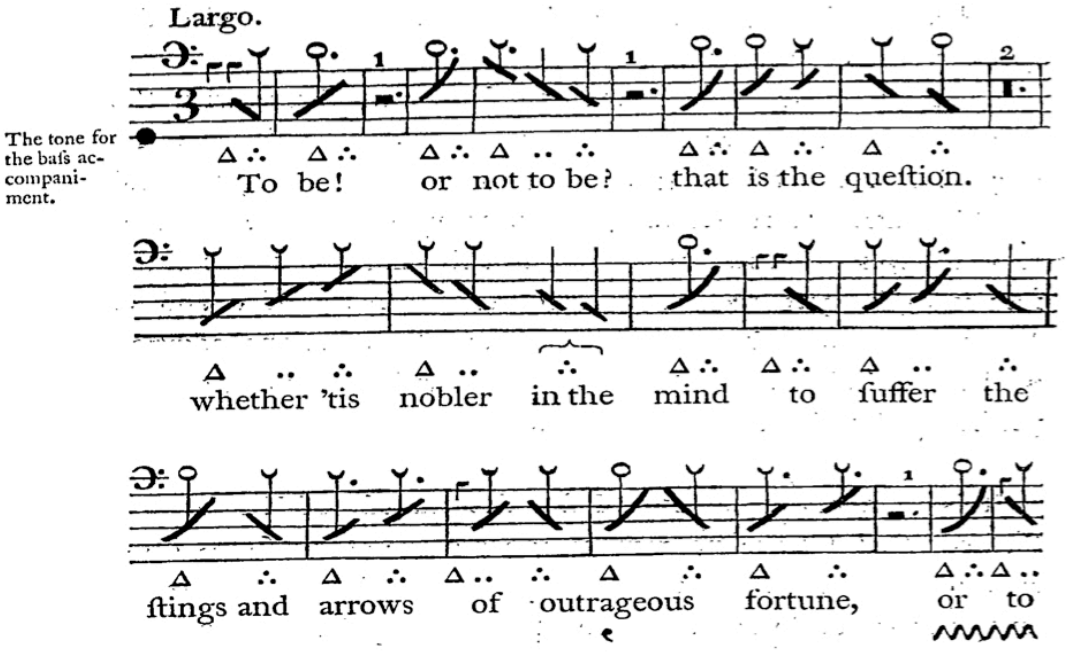


Figura 8 – Exemplo de indicação de interpretação teatral usando a viola da gamba baixo como referência (STEELE, 1779, p. 40)

A obra de Steele suscitou discussões quando foi publicada, mas em pouco tempo foi esquecida. Em perte, provavelmente, pelo fato de seu autor ter se mudado para a colônia inglesa de Barbados e nunca mais ter voltado à Inglaterra, mas também pelo pouco interesse que a filologia e a linguística demonstraram pela prosodia até meados dos século XX. Porém, acreditamos que um estudo mais aprofundado de sua obra pode revelar algum aspecto

6 Um instrumento semelhante ao violoncelo, com seis cordas, que caiu em desuso a partir do século XIX

significativo que nossas abordagens cientificistas podem estar deixando escapar. Steele escreve numa perspectiva universalista, unindo conhecimentos de música, teatro e, principalmente, de versificação clássica, em especial dos gregos. Esta perspectiva lhe permitiu perceber a independência das propriedades, que nas línguas atuais estão imbricadas de modo que é difícil separá-las.

Um estudo aprofundado desta obra foge dos objetivos da tese. Contribuem para este trabalho, no entanto, as ideias do autor a respeito da independência das propriedades de movimento melódico, quantidade e posição métrica, bem como a influência dos movimentos da curva de F0 no interior da sílaba para o sentido da entoação.

Talvez o aspecto mais importante da obra de Steele é levantar a questão a respeito da notação da entoação. O autor buscou registrar o mais fielmente possível as propriedades de altura e duração, mas isso revela algo sobre a sua estrutura fonológica? A julgar pela semelhança entre os desenhos de Steele e os gráficos produzidos por programas como o PRAAT, diríamos que não, pois estes diagramas precisam de uma interpretação ulterior baseada em um modelo teórico que, sabemos, está longe de ser consensual. Por outro lado, a preocupação em associar durações e posição métrica aos movimentos melódicos coloca a notação de Steele à frente dos gráficos F0 x t feitos em computador, malgrado sua imprecisão nas alturas. Hatfield (*op. cit.*) argumenta que o emprego de uma notação semelhante à notação musical para representar a entoação entra em choque com o fato de que a entoação não tem nenhuma relação com a tonalidade (combinação das notas em escalas e acordes), sendo significativos somente os movimentos melódicos; observa também que o tamanho dos intervalos na fala muda muito de falante para falante, e até para o mesmo falante, dependendo da situação, e portanto não haveria vantagem em se descrever estes intervalos com precisão. Discordamos de Hatfield no sentido em que a notação musical é descritiva, e não depende de uma tonalidade para ser interpretada, haja vista toda a música atonal escrita (aliás, cf. SOUZA, 2004, cap. 4, onde argumento que a crescente especialização e detalhamento da escrita musical foi um dos fatores que viabilizou o surgimento da atonalidade). Mas concordamos com a ideia de que o poder descritivo da notação musical não ajuda a revelar as estruturas internas da prosódia do enunciado, da mesma forma que acontece na empreitada de Steele. É necessária uma notação mais enxuta e que sinalize somente os movimentos significativos, sempre considerando a grade métrica. Até o momento, a notação derivada do trabalho pioneiro de Pierrehumbert, conhecida como ToBI (Tones and Break Indices), tem sido bem aceita por muitos estudiosos das mais diferentes tendências (LADD, 1996). Porém, ainda não está claro como a realização fonética dos picos e vales (H e L) influencia a

percepção do enunciado, não só pela caracterização de emoções e atitudes, mas também por uma possível interferência na sua estrutura fonológica.

Uma discussão mais aprofundada sobre notação foge ao escopo deste trabalho. As reflexões acima têm o objetivo de nos auxiliar no esclarecimento sobre as diferenças e semelhanças entre a fala e o canto. Consideremos agora a sua influência mútua no seu desenvolvimento histórico.

3.2. Música e Linguagem

O surgimento da ópera, nas últimas décadas do século XVI, representa um divisor de águas no desenvolvimento da maneira de se fazer e se pensar música na cultura ocidental. Pode-se dizer que um certo preconceito tardio (surgido no século XIX) com relação aos gêneros dramático-musicais, como a ópera e seus derivados (operetas, musicais, revistas, etc.) fez com que a história da música minimizasse o impacto deste acontecimento na cultura musical do ocidente. Estes gêneros, no mundo contemporâneo, acabaram se associando com a noção de entretenimento, em contraposição à música pura cultuada pela religião da arte pela arte. Entretanto, a própria concepção de música no ocidente é tributária da revolução provocada pela Camerata Fiorentina, que, buscando resgatar os princípios estéticos da tragédia grega (que eles supunham cantada), produziram as condições em que floresceu o que chamamos de música tonal, graças à necessidade de se coligar música e texto de uma maneira que ficasse convincente na voz de uma personagem no palco (CALENGANO, 2002).

Uma das características principais da música tonal é a organização do discurso em frases que se encerram com cadências (pontos de repouso rítmico-melódicos) de diferentes graus de conclusividade, demarcados pelas relações harmônicas em torno de um acorde principal ou tonalidade (daí a denominação "tonal"). Quanto mais distante do acorde de tônica, na hierarquia do campo harmônico, estiver o acorde final de uma cadência, menos conclusiva ela é. Este recurso técnico permitiu que a frase musical acompanhasse a dinâmica do texto, sua estrutura sintática, lógica e até retórica (hoje em dia diríamos "pragmática"), realizando o ideal de "*imitar col canto chi parla*" ("imitar com o canto quem fala"). Nessa perspectiva, podemos afirmar que o estabelecimento de uma noção comum (pelo menos no ocidente) do que seja música foi consequência de uma adaptação da composição musical às características estruturais da linguagem verbal.

Esta conclusão requer, antes de prosseguirmos, a análise de duas possíveis ressalvas. Primeiramente, música e linguagem estiveram ligadas desde tempos imemoriais, na forma de poesia cantada, ou simplesmente canto; assim, seria um exagero imaginar que a música só passou a assimilar a estrutura da linguagem a partir da ópera. Entretanto, o que merece destaque aqui é o fato de que a linguagem musical passou a ter uma lógica própria, aprendida com a linguagem. Por uma ironia curiosa, a música aprendeu com a linguagem a ter uma lógica, baseada nas relações harmônicas, e a partir daí rumou à sua independência e autonomia. Esta proposição é corroborada pelo fato de que a partir do Barroco (depois da ópera, portanto) é que a música instrumental independente ganhou autonomia e passou a ser apreciada independentemente de uma referência textual. Ou seja, por influência da necessidade de seguir, de acompanhar o texto, a música ocidental tornou-se a sua própria linguagem. Antes ela emprestava ao texto nuanças afetivas mas dependia deste para que a lógica do canto pudesse ser percebida.

Em segundo lugar, é preciso trazer ao nosso ponto a discussão infindável sobre o que é música. As pesquisas e experimentações das vanguardas musicais no século XX estenderam de tal maneira a concepção do que seria música (a ponto de considerar como sendo música a imobilidade de um pianista em frente ao piano por quatro minutos e trinta e três segundos) que somos levados a aceitar a definição de Luciano Berio: música é tudo aquilo que se ouve com a intenção de se ouvir música. Porém, se pensarmos no senso comum, no contexto da cultura ocidental, quando alguém (tanto leigos como a maioria dos músicos) emprega o termo música, está seguramente pensando em uma manifestação humana que emprega sons de altura definida organizados em melodias e acordes, com ritmo mais ou menos perceptível. E, na grande maioria dos casos, terá em mente algum tipo de música tonal. Por mais que a música de vanguarda tenha proposto novas maneiras de organizar os sons, o seu repertório atingem uma minoria das pessoas no ambiente urbano. Não só a música clássica (compreendida de Bach a Brahms), mas praticamente toda a música popular urbana, do sertanejo ao jazz, vem de uma matriz tonal. Podemos afirmar que o homem ocidental moderno (ou contemporâneo) é uma criatura tonal. É fundamental, portanto, que tenhamos em mente que quando falamos música aqui, estamos falando de música tonal e suas fronteiras, e não da concepção ampla das vanguardas.

Feitas essas duas ressalvas, podemos retomar o que foi dito acima numa fórmula sintética: a música tonal (equivalente de “música” para a grande maioria das pessoas) surgiu porque a música quis emular a estrutura da linguagem para, num primeiro momento, acompanhá-la com precisão e, depois, não mais dela depender. Sendo assim, quando

analisamos o desenvolvimento histórico da fala e do canto é preciso lembrar que o que entendemos hoje por música não existia antes do mundo moderno. Aliás, seria interessante aqui trazer alguns resultados das investigações sobre o conceito de mousiké.

Para os gregos, mousiké era uma espécie de "erudição estética", uma inspiração das musas em direção ao conhecimento por meio da apreciação artística e contemplação das estruturas subjacentes (a ideia platônica), e não uma habilidade com os sons. Estas habilidades, desprezadas pelos nobres por serem atributos de escravos (artes mecânicas indignas dos privilegiados), eram denominadas de acordo com o instrumento: aulética, citarística, e assim por diante. Durante a Idade Média, sob o domínio da igreja católica, a música é considerada já sob os dois aspectos: música prática e música especulativa, esta última puramente teórica, voltada para as relações matemáticas que regem os intervalos e sua relação com a cosmologia (que encontramos, por exemplo, na obra de Boécio). A música prática era vista como atividade inferior, mundana.

Curiosamente, foi no ambiente da música eclesiástica que se desenvolveu a técnica de composição pelas regras do contraponto, que permitiu a composição de novas melodias a partir de outras já existentes, servindo, portanto, a uma aplicação prática. Esta técnica de composição levou progressivamente ao conceito de acorde e assentou as bases sobre as quais a escrita para a ópera seria desenvolvida, levando à melodia acompanhada e à harmonia tonal.

Toda esta digressão a respeito da importância central da música tonal na concepção ocidental contemporânea do que seja música é necessária para que não tomemos o termo música na acepção mais ampla das vanguardas e da teoria musical contemporânea, e para que estejamos cientes do quanto esta música, suas formas e suas estruturas, é tributária da linguagem. Por isso, argumentos em favor de uma estrutura comum à música e à linguagem por meio de paralelos entre a linguagem e a música tonal (p. ex. Lerdahl e Jackendoff), além de terem um viés eurocêntrico, pecam por esconder uma tautologia, uma vez que a música tonal foi construída em cima de estruturas linguísticas.

Passaremos a seguir a um breve panorama do desenvolvimento histórico das manifestações sonoras humanas que podemos de alguma forma associar à música e à linguagem.

3.3. Origens comuns da fala e do canto

Muito já se especulou a respeito da origem da linguagem, bem como da origem da música. À medida que ciências como a arqueologia e a antropologia nos revelam fatos conclusivos dos primórdios da humanidade, podemos construir um quadro mais preciso de como estes dois elementos fundamentais da cultura e, portanto, da humanidade se entrelaçam e participam do desenvolvimento de estruturas sociais cada vez mais complexas.

Brown (2000) aborda o problema a partir da seguinte pergunta: música e linguagem têm uma origem comum ou uma evoluiu a partir da outra? O autor começa analisando as características comum aos dois domínios: um método de produção de enunciados por combinação (que ele denomina sintaxe combinatória) e agrupamento dos elementos destes enunciados por meio de propriedades acústicas unificadoras (que ele chama de "intonational phrasing", uma expressão difícil de traduzir, algo como fraseado entoacional). Esta semelhança estrutural leva o autor a enunciar a hipótese de uma origem comum da música e da linguagem, que teriam evoluido e se especializado a partir de um estágio anterior, uma espécie de "elo perdido", que ele chamou de "musilanguage", e que traduziremos aqui, por comodidade, usando também um neologismo: "musilíngua".

Uma das evidências que Brown apresenta em favor desta hipótese é o fato de que, apesar das diferenças notáveis que se desenvolveram durante a evolução humana, os dois domínios apresentam em maior ou menor grau aspectos referenciais e aspectos emotivos, e a diferença entre eles estaria na ênfase que é dada a um e outro aspecto. O autor sugere que ao invés de universos distintos com intersecção em alguns pontos, música e linguagem são melhor descritas como se estivessem dispostas ao longo de um espectro contínuo que iria da referencialidade absoluta à emotividade absoluta.

Outra evidência que Brown aponta é o grande número de línguas (a maioria das línguas existentes no mundo) que usam tons lexicais Estas línguas são mais comuns em regiões que demoraram mais a ter contato com a civilização ocidental. Este fato leva a crer que as línguas entoacionais são possivelmente estágios posteriores de desenvolvimento que aconteceu pela perda do tom lexical, e que, numa perspectiva evolucionária, a linguagem tem certamente uma das suas raízes num sistema de controle de alturas com valor semântico e sintático, não apenas expressivo (BROWN, 2000, p. 281). Vale a pena lembrar aqui as conjecturas apresentadas no primeiro capítulo a respeito do papel da altura no acento do grego antigo, bem como o sugestivo uso da palavra *accent* por Steele para se referir aos movimentos melódicos.

Brown também comenta as implicações das principais conclusões da TMA (de que tratamos no primeiro capítulo) com relação à sua concepção de musilíngua. A noção de que a entoação se desenvolve por meio de movimentos que buscam alvos, e que estes alvos se mantêm relativamente estáveis se for efetuada uma normalização dos resultados considerando o âmbito da fala de cada um (pitch range) aponta para a existência de esquemas de alturas semelhantes a escalas na fala, o que seria outro ponto em comum originário do estágio primordial proposto pelo autor. Desta forma, a vocalização de alturas específicas (discretas) dentro do âmbito do falante (a presença de uma escala) estaria presente na musilíngua desde o início.

A TMA ainda serve de base para um insight que vai ao encontro da nossa hipótese. Vejamos a argumentação de Brown (op. cit., p. 284).

> "The insight from autosegmental theory for the musilanguage model is that sequences of level tones can be the basis for semantic strings. The fact that intonation languages dissociate such strings of level tones from semantic strings emphasizes the earlier point that language's meaning level has no obligatory relationship to its phonological level or even to the acoustic modality. Intonation languages, like gesture languages, highlight teh primary importance of creating semantic meaning from meaningless components, whatever these components may be"

Somos levados a completar o que faltou falar: a melodia, ou seja, a sequência de *level tones* recebe a cadeia segmental, à maneira de um repentista que improvisa a letra sobre um padrão melódico pré-existente, e a principal justificativa para esta conclusão seria a hipótese da evolução das línguas a partir de uma matriz tonal, que é corroborada pela distribuição das línguas tonais no mundo e dos casos intermediários de línguas de acento tonal e de tom acentual, e também pelas evidências da relação entre altura e acento nas línguas indo-europeias antigas (como o grego), que se perdeu no desenvolvimento das línguas entoacionais. Brown também não fala, mas achamos importante explicitar que o fato de os estudos da entoação terem começado em línguas entoacionais criou um viés (eurocêntrico) que obscureceu por algum tempo estes aspectos da linguagem que, pouco a pouco, vão sendo revelados.

Quando, porém, Brown passa a analisar a formação de estruturas mais complexas, por combinação de frases, ele vê dificuldades em explicar a combinatória com a produção de padrões de entoação globais. Cremos que faltou só levar às últimas consequências a hipótese da precedência da melodia, que poderia organizar os enunciados em níveis sucessivos, num princípio semelhante ao da dupla articulação, relacionado com o duplo foco resolutivo do cérebro.

O autor deixa uma pista, entretanto: a presença de mecanismos de proeminência

possibilitaria a construção de hierarquia dentro dos primeiros enunciados (p. 285). Proeminências relativas podem criar relações de subordinação e coordenação, e estas relações podem ser especificadas pela configuração melódica associada a cada proeminência. A proeminência aglutina o que está à sua volta e se relaciona com outras proeminências.

No desenvolvimento das línguas entoacionais, os tons (originalmente com função lexical) vão aumentando o seu âmbito de ação para marcar sintagmas maiores e assim enriquecendo as possibilidades sintáticas e pragmáticas. Esta seria uma consequência do modelo musilíngua para a compreensão da natureza e do desenvolvimento da entoação. Brown também faz referênca às possibilidades expressivas que a entoação apresenta, mantendo o esquema geral de unidades discretas, cuja forma final depende de outros fatores como dinâmica (intensidade), registro (abertura do âmbito melódico) e andamento (correspondente na fala à taxa de elocução), ligados a aspectos emocionais e de atitude.

No fundo percebe-se uma base fonológica no modelo de Brown, e é curioso notar que ele hesita em admitir que o mesmo mecanismo de controle de altura que está na origem dos elementos fundamentais da fala (morfemas) pode estabelecer relações sintáticas. Parece que o autor não considera a proeminência relativa compatível com a ideia de uma escala para as alturas da fala (initializng vs. normalizing hypothesis; cf. LADD, 1996). Podemos objetar que esta questão ainda não está resolvida na TMA, e Brown tomou partido cedo demais.

De qualquer forma, é impossível resgatar os caminhos percorridos, mas as especulações com base na fonologia fizeram surgir ideias que encontram ressonância nas outras abordagens que vimos até aqui, notadamente a concepção de produção da fala proposta por Bakhtin (cf. cap. II, seção 4).

Além disso, uma deficiência que aparece ao considerarmos o desenvolvimento ulterior divergente de música e linguagem a partir da musilíngua é o fato de o autor não considerar aí a natureza particular das formas cantadas, que ocupam o centro do espectro que vai da fala ao canto. Brown não explicita o processo de descoberta dos intervalos fundamentais para o surgimento dos sistemas musicais (a quinta e a oitava), que muito provavelmente dependeu do desenvolvimento dos instrumentos. Resumidamente e hipoteticamente, podemos dizer que a construção de instrumentos aos poucos chamou a atenção para as relações harmônicas, e estas foram aos poucos infundidas nas manifestações vocais que já tinham alguma função ritualística, explicitada por notas longas e estabilidade das alturas.

Brown se perde ao não considerar a relação dialética histórica do desenvolvimento dos instrumentos musicais com a evolução dos sistemas de altura (ou ele não menciona por achar óbvio), e, por não explicitar como seria a dinâmica da incorporação dos sistemas de altura nas

manifestações vocais.

Independentemente de aceitarmos completamente a proposta de Brown, há que se considerar dois processos simultâneos que acompanham o desenvolvimento da humanidade nascente, a saber: 1) o controle de durações e alturas (usando instrumentos) formando padrões (que poderíamos chamar de protomúsica) e cujo uso coletivo leva ao estabelecimento de códigos de avisos e chamados, fenômeno que também acontece com os animais superiores (anfíbios, aves e mamíferos) [destaque para a característica exclusivamente humana que é a construção de instrumentos (ferramentas)], e 2) o emprego do aparelho vocal para expressar estados de ânimo, acompanhando gestos e expressões faciais na construção das relações entre os indivíduos de um grupo, e que aos poucos foi criando um repertório que permitia referências a objetos e processos do mundo real, servindo para comunicar experiências impulsionando o desenvolvimento do que chamamos cultura (que poderíamos chamar de protofala).

Estas noções de protofala e protomúsica não podem ser delimitadas ou mesmo comprovadas, pelo menos não no estágio atual do conhecimento do homem sobre suas origens, mas podem nos ajudar a considerar a influência mútua desses dois fatores que estão constantemente presentes no desenvolvimento da humanidade, que são o controle de durações e alturas e a expressão e referência por meio de sons vocais. Numa visão bastante simplista, porém útil, podemos dizer que é mais plausível que num primeiro momento a influência da organização da protomúsica sobre a protofala ampliou sua capacidade de estruturação e permitiu o surgimento de enunciados (cf. a dialética entre gesto e figura em SOUZA, 2004). Dito de outro modo, a organização de alturas e durações (que se desenvolve por meio da experimentação livre, até mesmo lúdica) insemina os gestos vocais primevos e os capacita de se conectarem por relações de contraste e repetição, ensejando o surgimento do que viria a ser as bases das regras fonológicas (que são feitas de relações de oposição). Em que ponto linguagem e música começam a divergir, é muito difícil, se não impossível de se afirmar.

Acreditamos que é necessário, para avançar nesse terreno movediço que é o conhecimento das origens da humanidade, considerar o papel dos rituais e manifestações religiosas no desenvolvimento da linguagem e da música. Benjamin nos chama a atenção para a relação intrínseca de arte e magia, e nos leva a especular sobre uma origem comum às diversas manifestações artísticas nos rituais de magia das primeiras religiões. Podemos distinguir os dois usos que vimos no capítulo II sendo aplicados na linguagem nascente: um prático, referencial, destinado a facilitar a organização do trabalho e a interação dos indivíduos, e outro mítico, mágico, necessário à manutenção do grupo pelo estabelecimento

de uma identidade e canalização da violência. Nesta esfera do uso mítico da linguagem podemos vislumbrar situações propícias para o desenvolvimento da poesia, campo por excelência para a criação de linguagem. Já é amplamente reconhecida a relação direta entre a poesia e as fórmulas mágicas e as orações das religiões. Assim, podemos concluir que este campo de criação de linguagem tem em sua base [como fator fundamental, propulsor] a influência dos padrões de duração e altura (o lógos da mousiké) sobre as manifestações vocais, não mais preocupadas com a referência e a expressão, mas buscando o encantamento, a conexão com um estado alterado de consciência e, principalmente, a preservação do conhecimento (não esqueçamos que as Musas são filhas de Mnemósine). Percebe-se, também, a dimensão política da linguagem poética enquanto mantenedora (direcionadora) das relações sociais, fator fundamental para o surgimento da civilização por meio da divisão do trabalho.

O encantamento da poesia deriva da inseminação do lógos na fala, continuamente produzindo linguagem nova. A música trouxe o encantamento para dentro da fala, a qual torna-se instrumento de poder e ao longo da história se impõe das classes dominantes (escribas, sacerdotes, e depois os políticos) para o resto da sociedade. Como resultado deste processo temos o estabelecimento das regras fonológicas que, num segundo momento, serão explicitadas (pelo menos em parte) no alfabeto fonético. A implementação tecnológica da representação visual dos fonemas teve um impacto de tal monta na organização da sociedade, e consequentemente no desenvolvimento ulterior da linguagem, que chegamos a esquecer que a linguagem nasceu oral. Tem-se facilmente a impressão de que a escrita representa a matriz do pensamento, da qual e fala é uma representação imperfeita, quando, na verdade, tanto ontogenética quanto filogeneticamente a linguagem nasce fala. É uma inversão de valores que podemos atribuir a 25 séculos de platonismo, mas que a filosofia do século XX já está francamente superando.

Sintetizando, as proporções de altura e duração (lógos) infundiram um poder encantatório na fala que, quando empregada ritualisticamente, adquire uma dimensão política de direcionamento e manutenção da ordem social – surge aí a poesia, campo de criação de linguagem e preservação do conhecimento. Este processo de constituição da linguagem como cadeias de segmentos acusticamente definidos, construídas segundo regras de combinação por contraste e repetição (herdadas da proto-música), depende portanto da delimitação dos usos da linguagem entre usos práticos e míticos; estes últimos, por sua estreita relação com o poder, direcionam o desenvolvimento da linguagem.

Chegamos assim ao problema dos usos da linguagem e sua relação direta com a utilização de elementos que hoje consideramos como musicais, ou seja, a questão central

desta tese que é a relação entre melodia e gênero de discurso. Passaremos, por isso, de uma abordagem diacrônica para uma abordagem sincrônica, apresentando um novo conceito, o de práxis vocal.

3.4. Práxis sonora

Araújo e Paz (2011), trabalhando no contexto da etnomusicologia, questionam as concepções tradicionais e eurocêntricas do que seja música e arte, aproximando-as de uma concepção mais atual e abrangente de linguagem, entendida por eles como "trabalho humano com extensão multifacetada em sonoridades, gestos e grafias" (*idem*, p. 212, grifo meu). Neste sentido, a linguagem é, segundo os autores, "o campo de expressão por excelência de conflitos e mediações entre interesses nem sempre conciliáveis, em outras palavras, um terreno de ação política – seja em sentido de controle do Estado ou das micropolíticas do cotidiano de grupos e indivíduos" (*idem*, *ibidem*). Para eles, o aprofundamento da compreensão dos fenômenos musicais pela pesquisa acadêmica requer que se reconheça a posição tanto da música como de outras formas sonoras neste universo da linguagem vista como trabalho envolvendo "disputas sociais à base das diferentes modalidades e concepções de linguagem".

Os autores alertam para "o emprego muitas vezes irrefletido da categoria música no debate acadêmico e a consequente assimilação inadvertida de hierarquias e esquemas de dominação entre visões hegemônicas e subalternas" (*idem*, p. 214). Observa-se aqui o movimento que tomou a etnomusicologia desde a década de 1960 e questionou definitiva-mente a delimitação do que seriam manifestações musicais nas diferentes culturas. Não é demais relembrar que a noção de música como atividade autônoma é uma característica da cultura ocidental moderna, e que para muitas culturas não há nem mesmo um termo equivalente, sendo usados termos que se referem às circunstâncias em que o fenômeno musical ocorre. Este novo posicionamento da disciplina frente ao sonoro implica a ampliação das áreas de interesse, buscando de maneira mais abrangente a "compreensão do fenômeno sonoro, do papel da sonoridade na linguagem, que irá se prolongar no gesto corporal, no discurso e na tomada do espaço de atuação" (*idem*, p. 219).

Para superar a limitação do termo música, os autores propõem a categoria "práxis sonora", entendendo "práxis" como a "contínua tensão e recíproca interferência crítica entre

reflexão e ação". A práxis sonora busca tratar de uma "totalidade que (...) enfoca estrategicamente a dimensão sonora da atividade prática humana, sem isolá-la de outros aspectos dessa mesma atividade geral, e, particularmente, sua dimensão política, i.e., de ação que propõe alianças, mediações e rupturas" (idem, p. 220). Mesmo não explicitando, podemos deduzir que os autores propõem que uma disciplina que trata do sonoro como trabalho humano tenha interesse também pela dimensão sonora da linguagem verbal (a linguagem *par excellence*), mesmo que delimitada ao seu uso não mundano, como chamados, pregões, parlendas, etc. Assim, decidimos nos apropriar da categoria práxis sonora e propor uma sub-categoria que nos interessa aqui, a saber a de "práxis vocal".

Para delimitar e mapear esta práxis vocal, traremos de volta a proposta de Greimas para a práxis gestual (vide cap. 2) e traçaremos suas diretrizes por meio de uma analogia. Assim como Greimas dividiu o campo da gestualidade humana em prática (expressão de um fazer) e mítica (expressão de um querer), já havíamos anunciado (SOUZA, 2004, p.) a possibilidade de se transpor para a entoação esta classificação, de maneira que as manifestações vocais, por mais distintas que fossem suas formas e seus objetivos, fossem compreendidas dentro de uma categoria do trabalho humano. Seguindo a sugestão de Araújo e Paz, propomos chamar o universo dos sons produzidos com a voz humana de práxis vocal, lembrando que esses sons não podem ser analisados em separado, mas dentro da cadeia de influências recíprocas nas relações humanas.

Esse universo da práxis vocal poderia, seguindo Greimas, ser dividido em vocalizações práticas (de uso cotidiano nas relações interpessoais) e míticas (de uso ritualístico, que ganha, portanto, uma dimensão necessariamente política). O primeiro caso compreenderia em grande medida a fala espontânea do cotidiano, em que pedimos e fornecemos informações necessárias à sobrevivência e preservacão; o segundo corresponde aos usos mais ou menos ritualísticos da voz, desde discursos, pregões, orações até recitativos, canções, árias e vocalizações puras. A vantagem que esta abordagem oferece é a de abrir caminho para um quadro teórico que possa tratar qualquer manifestação vocal por meio de um modelo consolidado, o que favorece uma abordagem sincrônica do uso da linguagem não limitado pelas fronteiras com usos específicos tais como a declamação e a canção.

Um problema que decorre, entretanto, é a eventual necessidade de incluirmos todo o estudo da linguagem nesse escopo, o que exigiria um trabalho gigantesco de adequação teórica. O modelo de Greimas não tinha este problema, pois o gesto não pode ser referencial, a não ser que se aplique uma convenção. Podemos, contudo, considerar a capacidade referencial da linguagem como uma necessidade da esfera prática, e não como algo intrínseco

à natureza da manifestação vocal. Aos poucos a capacidade referencial invade o campo expressivo, à medida em que a fala é inseminada pelo lógos, e este reflui para o campo referencial fornecendo nuances expressivas até para as manifestações vocais de ordem prática.

Mesmo reconhecendo as limitações e simplificações que um tal modelo apresenta, acreditamos que a diferenciação sistemática das manifestações vocais segundo sua utilização ritual ou não, referencial ou não, e assim por diante, ajuda a posicionar o problema da entoação num contexto mais amplo do que o puramente linguístico, criando assim um campo em que as diversas disciplinas que se interessam pela entoação possam se confrontar e, deste confronto, nasçam novas perspectivas para compreender o fenômeno.

É assim que a visão sociológica das trocas linguísticas pode dialogar com a teoria dos atos de fala e com a concepção de gênero de discurso num terreno comum, cuja predominância pode ser reivindicada, em última análise, pela filosofia da linguagem.

A aproximação com a etnomusicologia ainda renderá novas abordagens possíveis de análise, como veremos no próximo capítulo. Antes, porém, de proceder a uma síntese desta investigação sobre as relações entre fala e canto, permitiremo-nos um excurso a respeito da percepção auditiva e suas implicações na cognição de música e linguagem.

3.5. Aspectos cognitivos da percepção de fala e música

Aparentemente há uma mútua exclusividade entre a percepção de um dado enunciado como sendo “cantado” ou “falado”. Não distinguimos imediatamente a “melodia” de uma sentença proferida, a menos que deixemos em segundo plano a atenção ao conteúdo lingüístico do texto. Uma hipótese para explicar este fato seria que a percepção de intervalos está fundamentada num contexto de tonalidade musical, que seria “desligado” quando o foco está na comunicação lingüística. Mesmo assim, a percepção da fala é significativamente sensível às variações de altura. Sabemos que entoação de um enunciado é rica em informações sobre o falante, sua atitude e suas intenções, bem como as componentes sintáticas e pragmáticas do enunciado. Com o objetivo de esclarecer esta aparente contradição, apresentamos a seguir alguns resultados significativos das investigações das chamadas neuro-ciências sobre o processamento da percepção de altura no córtex cerebral.

Estudos recentes, empregando técnicas de visualização da atividade cerebral, vêm demonstrando que há uma especialização lateral de determinados componentes do

processamento da melodia. Mesmo assim, não é possível afirmar uma especialização lateral do processamento da altura, pois os resultados variam conforme o treinamento musical anterior do sujeito testado. Além disso, o processamento global tanto da linguagem como da música lança mão da ação coordenada de mecanismos neurais que ocorrem em diferentes localizações do cérebro. Ao invés de uma lateralização completa de um determinado processo, temos a localização específica e distinta dos mecanismos componentes do processo nos diferentes hemisférios.

Zatorre et alii (2002) reúnem uma série de evidências constatadas a partir de descobertas na área da neurologia, a respeito da lateralização dos processos envolvidos na percepção de fala e música. Os autores apontam que a fala emprega sons complexos que para serem identificados e categorizados necessitam de um processamento de diferenças no tempo da ordem de 10ms. Por outro lado, o reconhecimento com precisão de uma altura definida, bem como de nuances de timbre, exigem uma amostragem temporal maior, de maneira que seja possível um melhor detalhamento do espectro. Parafraseando o princípio da incerteza, da física quântica, formulado por Heisenberg, Joos (1948, apud Zatorre et al., 2002) propôs o "princípio da incerteza acústica", que afirma que não é possível realizar-se uma medição de um evento auditivo que seja precisa tanto no domínio temporal como no espectral.

A partir de resultados de investigações empregando métodos de medição e visualização de processos cerebrais *in vivo*, (ressonância magnética, tomografia computadorizada, entre outros), Zatorre e colegas apresentam evidências de que o córtex auditivo esquerdo é capaz de lidar com intervalos menores de tempo, se comparado ao direito; este, por sua vez, apresenta uma resolução espectral melhor. Estes dados estão de acordo com a idéia já bastante difundida de que o hemisfério direito é o hemisfério musical e o esquerdo é o da linguagem, uma vez que o uso da linguagem pressupõe a detecção de pequenas diferenças temporais e a música (no sentido que estamos adotando) emprega primordialmente sons com altura determinada. Ao mesmo tempo, permitem especificar as regiões do cérebro onde se dão estes processos e quais os detalhes deste processamento, bem como sua relação com os sons da linguagem e da música.

A comparação entre as habilidades linguísticas e musicais observadas em pacientes com lesões cerebrais em diferentes regiões do cérebro leva a constatações relevantes sobre o papel da altura na percepção da fala e da música. Investigações cruzadas da atividade cerebral em casos de amusia[7] e afasia mostram relações interessantes com respeito à especialização de processos cerebrais. Pacientes amúsicos não-afásicos estão associados com lesões bilaterais

7 Incapacidade congênita ou adquirida de reconhecer melodias.

do cerébro, enquanto que afásicos não-amúsicos apresentam em geral lesões somente no HE.

Peretz e Hyde (2003) apresentam evidências de que existem redes neurais distintas especializadas no processamento da altura (notas, intervalos) e de sua organização (escalas, acordes), que podem ser inutilizadas seletivamente por lesões em diferentes pontos do córtex auditivo. Deficiências no funcionamento destas redes também podem ser determinadas geneticamente, o que resulta na chamada amusia congênita. Estima-se que cerca de 5% de uma dada população apresentam amusia congênita, e apresentam as mesmas deficiências em identificar ou compreender a organização das alturas na música que os indivíduos com amusia adquirida por alguma lesão.

Do mesmo modo, lesões em regiões do córtex auditivo esquerdo, especializadas no processamento temporal fino, resultam em deficiências na percepção e compreensão da linguagem que afetam sensivelmente a fala, e que caracterizam o quadro conhecido como afasia. Há relatos de pacientes com lesões cerebrais que perderam a capacidade de reconhecer os sons da fala, mas continuaram reconhecendo melodias conhecidas, o que vai ao encontro da idéia da separação entre os circuitos que lidam com o processamento espectral e temporal. De acordo com outras observações relatadas, deficiências no processamento da altura não afetam significativamente o desempenho da fala em sujeitos com amusia congênita. Segundo Peretz e Hyde (*op. cit.*), indivíduos portadores de amusia percebem bem as variações de altura na fala, mas as autoras não apresentam resultados de experimentos que corroborem esta afirmação. Por outro lado, Schön et alii (2004) apontam para o fato de que sujeitos com treinamento musical percebem melhor as nuances de sentido e intenção na entoação, o que nos leva a supor que algum componente do processamento de melodias atua na interpretação da entoação. Consoante a este resultado, constatou-se nesta mesma investigação que a atividade cerebral na percepção da fala tende a ser bilateral nos músicos e predominantemente no HE para os não-músicos.

Outros resultados relevantes quanto à percepção de altura na linguagem são os resultados de experiências comparando atividade cerebral de falantes de línguas tonais e não-tonais. De maneira geral pode-se dizer que falantes de línguas não tonais apresentam atividade cerebral maior no hemisfério direito ao ouvirem palavras de uma língua não tonal, enquanto que, nas mesmas condições, falantes de línguas não-tonais apresentam atividade cerebral maior no hemisfério esquerdo. A consequência desta constatação seria de que a percepção do contorno melódico fora do contexto linguístico é uma tarefa executada no hemisfério direito. A respeito do uso acentuado do HE podemos lançar uma hipótese: a identificação dos lexemas pressupõe algum processo fonológico que envolve o

reconhecimento de traços distintivos, os quais podem incluir categorias de intervalos.

Um desdobramento interessante deste assunto é a associação que parece existir entre a percepção dos tons lexicais em línguas tonais, e o intrigante fenômeno do ouvido absoluto (DEUTSCH, 2006), bem como as investigações mais recentes da mesma autora sobre a influência da linguagem do sujeito na sua percepção do paradoxo do trítono.

Um trabalho interessante que chega muito próximo destas características é a investigação da percepção dos tons lexicais em línguas tonais, levada a cabo na Unidade de Neurociência Cognitiva da Universidade de Montreal (KLEIN et al., 2001). A percepção dos tons lexicais do mandarim ativou áreas específicas do HE em falantes nativos desta língua, enquanto falantes nativos de inglês (uma língua não-tonal) apresentaram ativação no HD. Este e outros trabalhos na mesma área evidenciam a predominância do HD para percepção de timbre, harmonia e contorno melódico, enquanto o HE é ativado preferencialmente em tarefas de discriminação de intervalos e de ritmo. É praticamente consensual também que o HE é por excelência o hemisfério da linguagem.

Um recente estudo (FEDORENKO et alii, 2009) propõe um experimento em busca de evidências de compartilhamento de redes neurais em processos musicais e da linguagem. Seu ponto de partida é a hipótese proposta por Patel (2003), de que existe um processamento integrado de música e linguagem no nível sintático. O experimento consistia em medir estatisticamente a compreensão dos sujeitos em frases cantadas que apresentavam incongruências ou dificuldades lingüísticas e musicais no mesmo ponto, comparadas quando estes pontos que exigem mais do processamento on-line. As observações indicam uma diminuição da compreensão do enunciado quando as dificuldades linguísticas e as incongruências musicais coincidem, o que é uma evidência a favor da hipótese do compartilhamento do processamento on-line (ou pelo menos de uma parte dele) das alturas pela música e pela fala. Entretanto, os resultados não são conclusivos quanto à natureza e ao nível hierárquico dos sistemas compartilhados. O texto aponta como uma interpretação possível destes resultados a competição de dois sistemas independentes (cognição musical e lingüística) pelos recursos disponíveis de processamento on-line na organização sintática de frases e melodias. Há também relações ainda não esclarecidas com a memória de trabalho e seu papel na organização sintática dos enunciados (frases e melodias).

A partir dos textos analisados, podemos listar as seguintes constatações:

- tanto a escuta de música como a percepção da fala envolvem processamentos dinâmicos dos parâmetros sonoros, e alguns destes processamentos podem compartilhar a mesma rede neural;

- a percepção de tons e dos sons normalmente empregados na fala (especialmente as consoantes) são processadas independentemente, em locais diferentes do córtex auditivo e, de maneira geral, há uma predominância do hemisfério direito e esquerdo respectivamente;
- identificação de intervalos exige uma maior precisão na definição da frequência fundamental, e que está provavelmente conectada com o fenômeno da equivalência de oitava, que origina as classes de altura;
- o córtex auditivo apresenta uma especialização no processamento do som em cada hemisfério, sendo que o do HE possui uma maior resolução temporal e o do HD uma maior resolução espectral;
- os sons da fala, especialmente as consoantes, exigem uma maior resolução temporal para serem reconhecidos, e são, por este motivo, processados no HE;
- o processamento de timbre e altura se dá predominantemente no HD, porém há evidências de que o HE também é capaz de lidar com contornos melódicos, apesar de não ter a mesma precisão na discriminação dos intervalos;
- a identificação dos tons lexicais de línguas tonais envolve atividade no HE para os falantes nativos, porém no HD para os não falantes, o que pode indicar que a precisão intervalar não é significativa;
- a audição de música em músicos treinados apresenta atividade na área de Broca, reconhecidamente responsável por processamentos lingüísticos, o que parece indicar um compartilhamento do processamento dinâmico do material que chega até o cérebro (verificação de regras sintáticas);
- sujeitos com treinamento musical percebem melhor as nuances da entoação.

Considerando as evidências acima, podemos colocar uma questão relacionada com a investigação do fenômeno da entoação. Se o processo de discriminação do contorno melódico de uma frase entoacional é parte de um processo linguístico, então estaria localizado também no HE? (Pois no caso do reconhecimento dos tons lexicais é o que ocorre.) Neste caso o determinação da altura não é importante, mas a identificação de algum traço distintivo, que tanto pode ser de forma (Gestalt) ou de intervalo, ou, até mesmo, camadas de traços em diferentes parâmetro, formando uma textura polifônica. No processamento fonológico, cada segmento é posto em relação com diferentes eixos paradigmáticos de acordo com a presença ou ausência de um determinado traço distintivo. Existem, então, traços distintivos na

entoação? Ao identificarmos determinadas nuances de sentido, intenção e contexto na fala de alguém, estamos identificando categorias? Devemos lembrar também que o processamento da identificação de intervalos no músico treinado se dá predominantemente no HE. Por outro lado, a conclusão de que todo o processamento linguístico é concentrado no HE parece pouco elegante. Poderíamos neste caso sugerir que se façam investigações quanto a possíveis evidências de atividade especial no HD relacionada com algum processamento sintagmático na música e na linguagem, que de alguma maneira estaria relacionado com a identificação de uma forma no tempo (seja uma melodia ou o contorno de uma frase entoacional).

3.6. Investigação conjunta da fala e do canto

Vimos nas seções anteriores que tanto numa perspectiva diacrônica como sincrônica, fala e canto estão intimamente relacionados, sendo difícil em muitas situações distinguir com clareza quando se trata de um ou de outro. Há também casos intermediários em que falamos de uma "fala cantada" ou de um "canto falado". Considerando a proposta apresentada de uma categoria do trabalho humano que chamamos "práxis vocal", acreditamos que seja mais produtivo investigar a fala e o canto num mesmo quadro teórico e assim as inúmeras formas e funções da entoação estariam contempladas dentro de uma mesma abordagem.

Restaria determinar quais métodos seriam aplicáveis e se há alguma possibilidade de interação deles. Considerando a equivalência acústica de fala e canto (ambos são produzidos pela alternância de segmentos de vozeamento com consoantes, sonoras ou não, e são emitidos em trechos limitados por breath-groups), a investigação dos parâmetros acústicos por métodos estatísticos pode identificar semelhanças e diferenças significativas entre ambos. Estes métodos já são usados com sucesso tanto na área conhecida como etnomusicologia computacional, como na própria linguística, no campo dos estudos prosódicos. Pela aproximação e combinação de métodos e pelo cruzamento de dados seria possível esclarecer quais características acústicas distinguem o canto da fala.

Se a análise acústica por métodos estatísticos pode revelar detalhes sobre a forma dos enunciados, existe outra possível abordagem que pode buscar identificar as semelhanças e diferenças de ambos quanto ao uso. Uma investigação deste tipo poderia contribuir para que avançássemos no impasse sobre as funções da entoação, estabelecendo um diálogo com as noções de gênero e estilo e buscando encontrar relações entre o contexto em que o enunciado

(seja fala ou canto) é proferido e a maneira como este é interpretado.

Certamente a os resultados de ambas abordagens analisados de forma complementar levarão a uma compreensão mais rica e profunda do fenômeno das manifestações vocais. Acreditamos também que esta combinação coloca o problema da entoação numa perspectiva mais ampla e por isso pode oferecer saídas para os impasses e problemas epistemológicos apontados no primeiro capítulo.

4. Contribuições da etnomusicologia ao estudo da entoação

4.1. Etnomusicologia

Desde o estabelecimento da etnomusicologia como ciência autônoma, a partir dos estudos de Charles Seeger e Curt Sachs, entre outros, o problema de se analisar as manifestações musicais de culturas não-ocidentais evidencia a ineficácia das categorias que são relevantes para a musica ocidental quando aplicadas para descrever a música de outras culturas. O caso mais notável é o do emprego da notação ocidental para se registrar manifestações musicais de outras culturas (WILL, 1999), porém o problema se aprofunda ao considerarem-se os múltiplos e diversos usos do som nas diferentes culturas e, mais ainda, a impossibilidade de se formular uma definição universal de música.

Um investigador que mergulhar numa cultura estranha à sua deve aprender a reconhecer e identificar os diferentes gêneros musicais e seu correspondentes papeis no funcionamento da sociedade. As primeiras gerações de etnomusicólogos já dispunham de recursos para registrar os sons em gravações, mesmo que precários, mas as análises do material enfrentavam as barreiras culturais como a da notação, por exemplo.

Segundo Blacking (apud TZANETAKIS et al, 2007) objetivo principal da etnomusicologia é explicar a música e as práticas musicais das diferentes culturas com relação às suas estruturas sociais em termos dos fatores musicais que envolvem a execução e a apreciação. Neste sentido a etnomusicologia se distancia de uma antropologia da música, já que seu discurso deve se pautar primordialmente pelo conhecimento musical.

Percebe-se, então, que a exigência fundamental da disciplina (evitar concepções eurocêntricas que podem distorcer a compreensão da cultura do outro) se choca com a exigência de se produzir um discurso musicalmente qualificado. Este tem sido o desafio principal da área, e assemelha ao que acontece no estudo da entoação: como investigar o comportamento da melodia sem conhecer as suas categorias a priori?

Além dessa aproximação num impasse epistemológico, a etnomusicologia também se vê às voltas com o problema de se distinguir o que é fala e o que é canto. É comum encontrar exemplos de fala que poderíamos, do ponto de vista da nossa cultura, considerar como canto, e vice-versa. E a análise puramente acústica não permite a distinção, porque esta está ligada essencialmente ao seu uso no seio da sociedade.

Nas últimas décadas, o emprego de métodos computacionais tem permitido a análise

de grandes massas de dados por meio de métodos estatísticos e algoritmos de busca de padrões. Estes processos permitem a descoberta de formas de organização do som a partir do próprio som, e têm sido decisivos para a solução de problemas como a identificação de gêneros e estilos musicais em culturas não-ocidentais.

Acreditamos que estes mesmos processos podem ser empregados na investigação da entoação, tanto dentro de uma língua ou dialeto, como na comparação entre línguas. Neste capítulos expomos os princípios e práticas destas metodologias que caracterizam o que chamamos atualmente de etnomusicologia computacional (TZANETAKIS et al, 2007), avaliando a possibilidade de aplicação na elucidação dos problemas levantados nos capítulos anteriores.

4.2. Fronteiras entre fala e canto

List (1963) observa que a fala e o canto compartilham três características fundamentais que as distinguem de todas as outras manifestações sonoras humanas: 1) são produzidas pela voz; 2) são linguisticamente significantes; e 3) são melódicas (LIST, 1963, p.1). Nenhuma outra forma de comunicação pelo som desenvolvida pelo homem apresenta essas três características simultaneamente.

A partir desta constatação, é possível considerar o conjunto de todas as manifestações da fala ao canto como uma ampla categoria da atividade humana, o que concorda com as ideias apresentadas anteriormente a respeito da práxis vocal e também com a noção de um espectro contínuo da fala ao canto proposto por Brown (cf. cap. III). Assim, o problema que se apresenta é o da classificação destas manifestações, uma topologia do campo da práxis vocal.

List argumenta que o emprego das duas primeiras características (produção vocal e significação) envolvem problemas muito complexos, sem tecer maiores explicações e, por isso, uma classificação que levasse em consideração as característica melódicas permitiria a distinção e comparação das mais diversas manifestações vocais. Cremos que o motivo principal que levou o autor a optar pela melodia como base da sua classificação é que esta pode ser reduzida a duas grandezas simples, frequência fundamental e tempo, enquanto que as outras duas seriam muito mais difíceis (se não impossíveis) de serem parametrizadas.

Quando se refere à fala (*speech*), List tem em mente a fala espontânea (ele usa o termo *casual*), que distingue do que denomina "*heightened speech*" que poderíamos traduzir por fala

elevada, intensificada ou incrementada[8] e que ele considera como formas intermediárias entre a fala e o canto (*op. cit.*, p. 3). Esta fala alterada abrange casos como discursos, sermões, narrativas e representações dramáticas, entre outros. Canto[9], por outro lado, é definido como uma forma que exibe alturas relativamente estáveis e uma estrutura escalar com, no mínimo, sete tons (escala heptatônica[10]) e com uma mínima influência da entoação da fala casual.

Antes de propor o seu sistema de classificação, List lembra que qualque proposta de classificação é influenciada pela cultura de quem propõe. Algumas culturas distinguem inúmeras formas de fala alterada enquanto outras consideram todas uma categoria só; outras nem têm nome para o que chamamos de música. As classificações e respectivas nomenclaturas variam de acordo com a função social de cada forma (ou tipo, de fala alterada). List cita o exemplo da fala do leiloeiro, que apesar de não ser considerada canto, tem algumas características melódicas de um canto monotônico como a salmodia, por exemplo[11].

A proposta de List se baseia na observação de que a alteração da fala que caracteriza o canto e outras manifestações intermediárias acontece de duas maneiras opostas. Uma é a negação ou a planificação (nivelamento) da curva de F0, em direção a uma fala monotônica; a a outra é o exagero do movimento melódico da entoação, como o que aparece, por exemplo, na fala teatral. A partir destes dois tipos de alteração da fala, List constroi um gráfico no qual pode situar as diferentes formas de entoação conhecidas. Vejamos a explicação do autor:

> "The classification system which we shall now develop is based to a great extent upon these two divergent modifications of speech intonation. The chart or graph (…) is analogous to a hemispheric map of the world. At the north pole is placed casual speech. Song, as previously defined, is found at the south pole. The forms found north of the equator are those which seem to have more characteristics of speech than song. The forms south of the equator exhibit to a greater extent the traits associated with song. Movement to the east represents the diminution or negation of the influence of speech intonation. Movement to the west represents either the expansion of intonational contours or of scalar structures. Modification is continuous along the diagonals and each area marked extends its influence with decreasing magnetism in the three possible directions. Thus, the use of lines of latitude and longitude permits the placement of forms at mid-points both horizontally and vertically" (LIST, 1963, p. 7).

Abaixo reproduzimos o gráfico idealizado por List.

---

8 Nenhuma destas traduções corresponde à ideia de uma alteração do padrão usual da fala, pelo que optaremos pela expressão "fala alterada".

9 O termo empregado por List é *song*, que se traduz mais adequadamente por "canção"; no entanto, este termo em português se refere a um gênero específico de música vocal, e por isso optamos por usar a expressão mais abrangente "canto".

10 Não encontramos um motivo plausível para o autor excluir as formas cantadas que empregam escalas de cinco e seis tons, bastante comuns na América, África e Extremo Oriente; de qualquer forma essa exclusão não compromete a lógica da classificação proposta.

11 Aqui falta um termo adequado para traduzirmos *chant*, que se refere a certos usos de uma fala recitada em estilo monotônico em situações ritualísticas, especialmente no serviço religioso.

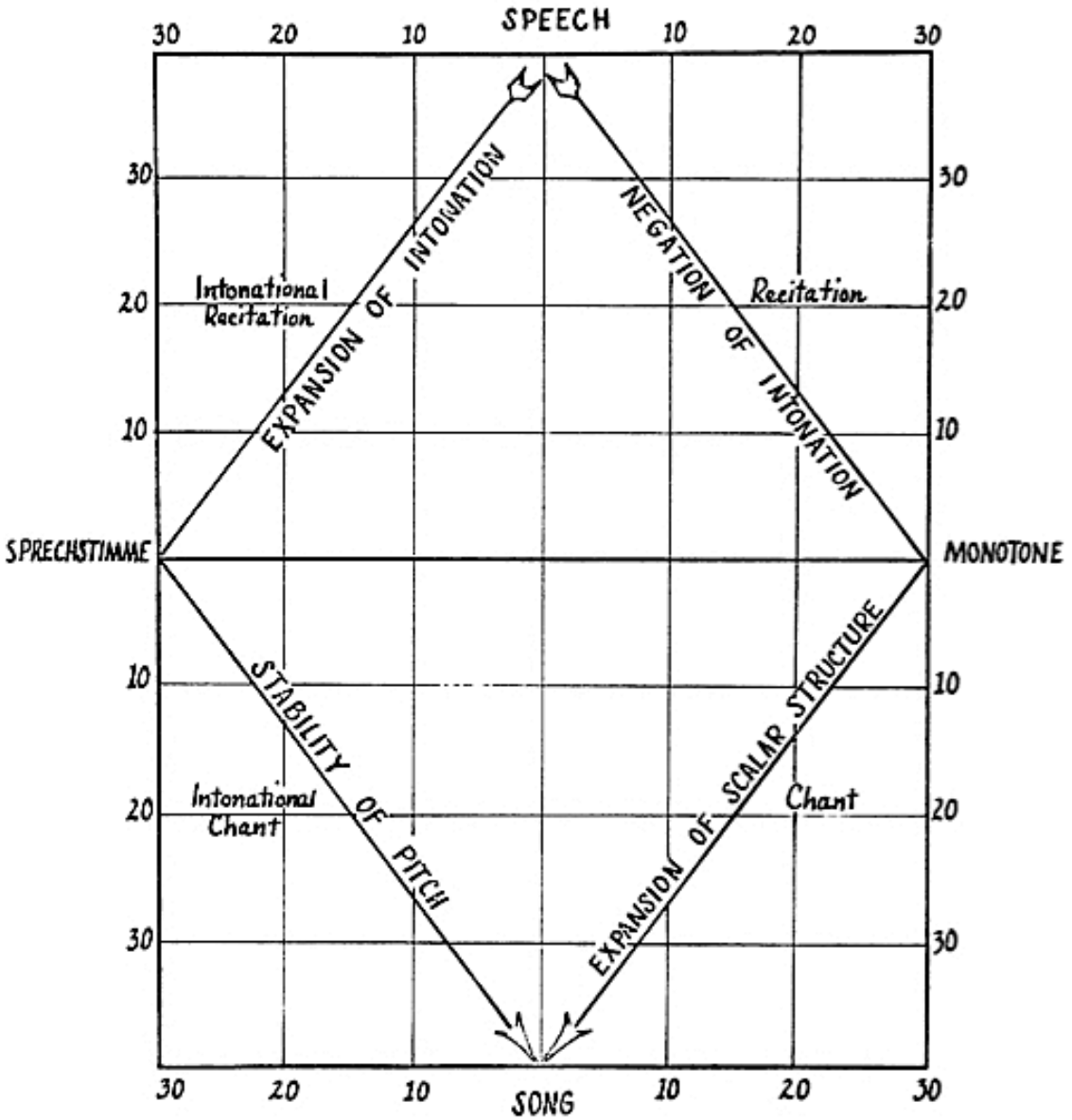


Figura 10 – Gráfico usado para a classificação das formas intermediárias entre fala e canto (LIST, 1963, p. 9)

A partir deste gráfico, o autor se propõe situar as manifestações vocais segundo suas características melódicas, localizando em diferentes regiões do plano manifestações tão diferentes como o recitativo, as narrativas, os cantos litúrgicos e o *sprechgesang*[12] (*op. cit.*, p. 9 e ss.).

12 Forma de declamação melódica característica do melodrama alemão no final do século XIX, cuja notação foi sistematizada por Arnold Schoenberg em sua célebre obra Pierrot Lunaire; literalmente, "canto falado"

Podemos enumerar uma série de problemas na proposta de classificação feita por List. A primeira é a restrição que faz a algumas formas vocais, especialmente as vocalizações sem sequências lexicais, o que exclui formas tradicionais de expressão vocal, como chamados e aboios, por exemplo. Mais crítico, porém, é uma certa falta de rigor matemático na atribuição dos eixos e dos parâmetros (expansão e negação da entoação são características baseadas em impressões, por demais subjetivas e dificilmente quantificáveis). O diagrama, em si, na forma de um mapa-múndi, dificulta a percepção de eventuais independências entre os parâmetros. Finalmente, o modelo não inclui as durações das sílabas como parâmetro ou critério de classificação, quando sabemos que é uma característica importante quando distinguimos a fala do canto. Existe, porém, uma premissa implícita que, se não afeta a abordagem etnomusicológica, certamente entra em conflito com o estado atual dos estudos prosódicos: List trata a fala espontânea (casual speech) como um modo neutro da fala, uma espécie de "grau zero da fala", que sabemos, não pode ser sustentado. Voltaremos a este ponto na discussão sobre fala espontânea, no capítulo V.

Não obstante todas essas deficiências, o modelo de List merece atenção pela iniciativa de propor uma topologia do conjunto das manifestações vocais baseada, em princípio, em propriedades acústicas pelo menos em tese quantificáveis. Um modelo deste tipo auxilia na comparação das diferentes formas (e, acrescentaríamos, gêneros e estilos) observadas dentro de uma mesma cultura, e também entre diferentes culturas. O autor encerra o artigo afirmando que é possível melhorar a proposta adotando uma definição mais rigorosa dos parâmetros, com dados mais precisos e incluindo novos parâmetros. Segundo sua perspectiva, com uma representação mais precisa, cartesiana, com tantos eixos quanto sejam necessários, constituindo um espaço multi-dimensional, seria possível situar com mais exatidão as diferentes formas das manifestações vocais. Deste processo resultariam agrupamentos destas formas, e estes serviriam de base para uma eventual tipologia (LIST, 1963,p. 14). Num certo sentido, o desenvolvimento posterior da etnomusicologia seguiu este caminho, gradualmente implementando formas mais sofisticadas de representação deste espaço, como veremos a seguir.

4.3. Cantometria

Lomax (1962) propõe um sistema de classificação das manifestações musicais com o sugestivo nome de cantometria (*cantometrics*). Este modelo foi desenvolvido a partir de uma incumbência de analisar e classificar um acervo de aproximadamente 400 conjuntos de

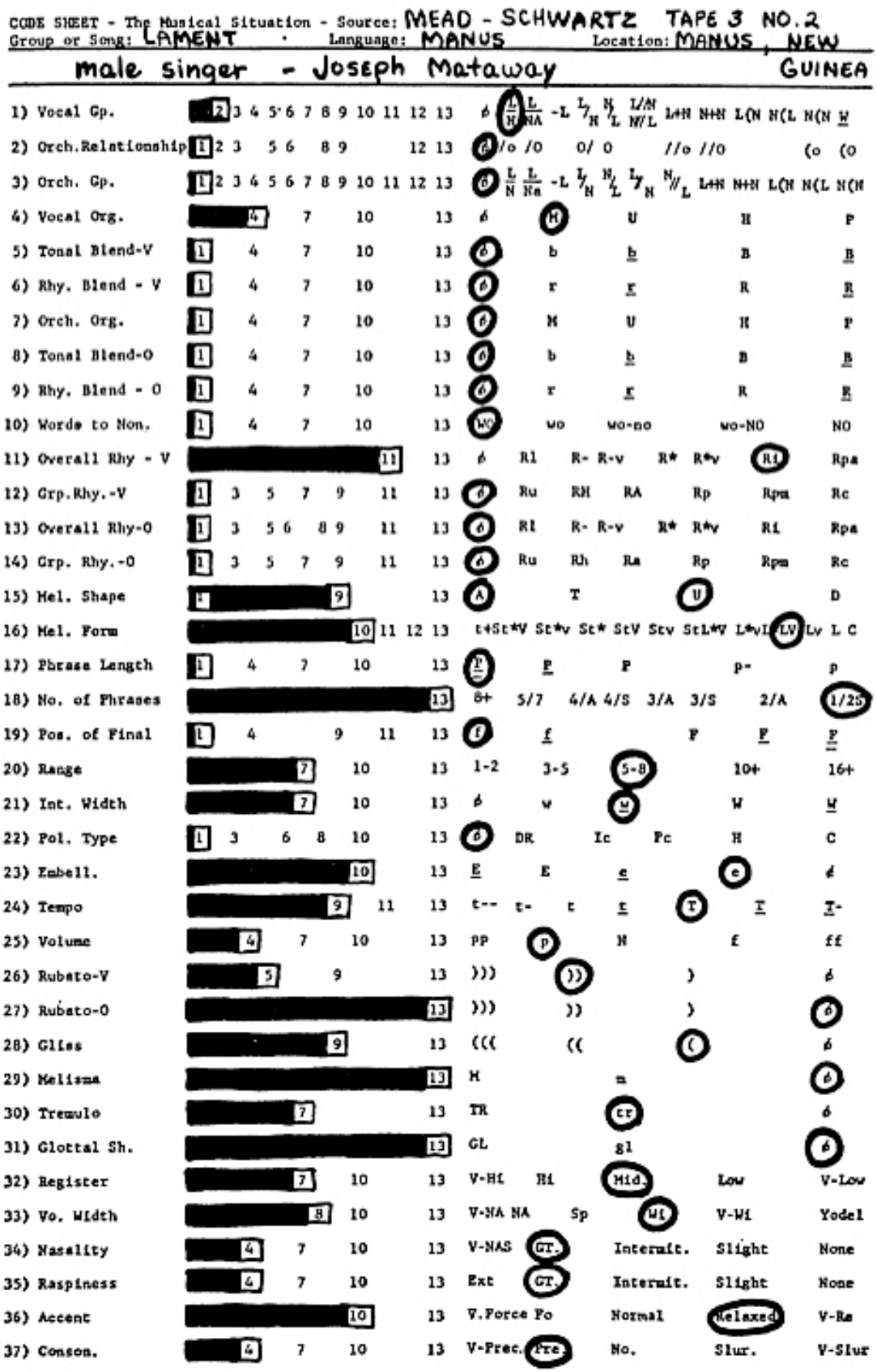
gravações em disco e em fita magnética de exemplos musicais advindos de cerca de 250 culturas diferentes, pertencente à Biblioteca do Congresso Nacional dos Estados Unidos. A ideia era desenvolver um tipo de análise que permitisse comparações objetivas entre as diferentes culturas, tendo como hipótese de trabalho existência de correlações entre as formas e práticas musicais e as estruturas sociais do grupo que as produz. Lomax também pretendia se valer da vantagem de poder analisar o som diretamente das gravações, sem a intermediação da notação musical convencional (ocidental), que, além de não dar conta das sutilezas da interpretação particular de cada grupo, introduz implicitamente categorias não necessariamente significativas para aquela cultura.

O método consiste em registrar sistematicamente características gerais de estilo da prática musical analisada em uma tabela padronizada, segundo critérios definidos divididos pelos elementos comuns da música. Vejamos como Lomax explica o método.

> "Cantometrics is a system for rating a song performance in a series of qualitative judgments (…). [It] takes into account the phenomena described by European music notation – melody, rhythm, harmony, interval size, etc. – but it looks beyond these European basics at many other factors present in and (as far as we could tell by intensive listening) generic to the song style of other areas. These factors include (...) the type and the degree of melodic, rhythmic and vocal embellishment in a sung performance; and the qualities of the singing voicenormally effected by the chosen singers in a culture. (…) These judgments are recorded in a series of 37 rating scales on a standard data sheet. Each one of these parameters or lines contains from three to thirteen points, each point being the locus of a proximate judgment in relation to the other points in the same line. The number of levels was limited to 37 by the size of the coding sheet, and the number of points was limited by the thirteen punches available in a column on an IMB card. No more points were included on any line than we felt could be handled by an attentive listener. These 37 lines, with 219 points, are set forth in a symbolic map on the right side of the coding sheet. The symbols, which are abbreviations for the distinctions made in each line, greatly facilitate learning and using the system. The listener records his judgments on the symbolic map and then transfers them to a number map on the left, which also serves as an IBM data sheet. Here the numbers are arranged and spaced so that they match the corresponding lines of symbols. (…) The results of this notation may be compared and then averaged with material from the same culture, until, within a short working period, a master profile in numerical or linear form is ready for cross-cultural comparison" (LOMAX, 1962, p. 427)

Segundo Lomax, o método permite a qualquer observador treinado fazer registros objetivos, similares e comparáveis sobre qualquer música que ele ouvir, seja ao vivo ou em gravação em uma ficha padronizada (fig. 11). Cada linha corresponde a uma das características da manifestação musical analisada, que compreendem desde aspectos sociais e de organização da prática musical, até traços especificamente musicais, como andamento, ritmo, forma melódica, ornamentação, presença de melismas, entre outros.

Figure 1: Sample Cantometric Coding Sheet



CODE SHEET - The Musical Situation - Source: MEAD - SCHWARTZ TAPE 3 NO. 2
Group or Song: LAMENT · Language: MANUS Location: MANUS, NEW GUINEA
male singer - Joseph Mataway

1) Vocal Gp.	1	2	3	4	5	6	7	8	9	10	11	12	13	ø	L/N	L/NA	-L	L/N	N/L	L/N N/L	L+N	N+N	L(N	N(L	N(N	W
2) Orch.Relationship	1	2	3		5	6		8	9			12	13	ø	/o	/O		O/	O		//o	//O			(o	(O
3) Orch. Gp.	1	2	3	4	5	6	7	8	9	10	11	12	13	ø	L/N	L/Na	-L	L/N	N/L	L/N	N/L	L+N	N+N	L(N	N(L	N(N
4) Vocal Org.				4			7			10			13	ø		H			U				H			P
5) Tonal Blend-V	1			4			7			10			13	ø		b			b				B			B
6) Rhy. Blend - V	1			4			7			10			13	ø		r			r				R			R
7) Orch. Org.	1			4			7			10			13	ø		H			U				H			P
8) Tonal Blend-O	1			4			7			10			13	ø		b			b				B			B
9) Rhy. Blend - O	1			4			7			10			13	ø		r			r				R			R
10) Words to Non.	1			4			7			10			13	WO		wo			wo-no				wo-NO			NO
11) Overall Rhy - V											11		13	ø		Rl		R-	R-v		R*	R*v		Rl		Rpa
12) Grp.Rhy.-V	1		3		5		7		9		11		13	ø		Ru			RH		RA		Rp		Rpm	Rc
13) Overall Rhy-O	1		3		5	6		8	9		11		13	ø		Rl		R-	R-v		R*	R*v		Rl		Rpa
14) Grp. Rhy.-O	1		3		5		7		9		11		13	ø		Ru			Rh		Ra		Rp		Rpm	Rc
15) Mel. Shape	1								9				13	A				T				U				D
16) Mel. Form										10	11	12	13	t+St*V		St*v		St*	StV		Stv	StL*V		L*vL	LV	Lv L C
17) Phrase Length	1			4			7			10			13	P		P			P				p-			p
18) No. of Phrases													13	8+		5/7		4/A	4/S		3/A	3/S		2/A		1/2S
19) Pos. of Final	1			4					9		11		13	f		f						F		F		F
20) Range							7			10			13	1-2		3-5			5-8				10+			16+
21) Int. Width							7			10			13	ø		w			W				W			W
22) Pol. Type	1		3			6		8		10			13	ø		DR			Ic		Pc		H			C
23) Embell.										10			13	E		E			e				e			é
24) Tempo									9		11		13	t--		t-		t	t			T		T		T-
25) Volume				4			7			10			13	PP		p			N				f			ff
26) Rubato-V				5					9				13	)))				))				)				ó
27) Rubato-O													13	)))				))				)				ó
28) Glies									9				13	(((				((				(				ó
29) Melisma													13	M					m							ó
30) Tremulo							7						13	TR					tr							ó
31) Glottal Sh.													13	GL					gl							ó
32) Register							7			10			13	V-Hi		Hi			Mid.				Low			V-Low
33) Vo. Width								8		10			13	V-NA		NA		Sp		Wi			V-Wi			Yodel
34) Nasality				4			7			10			13	V-NAS			GT.		Intermit.				Slight			None
35) Raspiness				4			7			10			13	Ext			GT.		Intermit.				Slight			None
36) Accent										10			13	V.Force		Fo			Normal				Relaxed			V-Re
37) Conson.				4			7			10			13	V-Prec.			Pre.		No.				Slur.			V-Slur

Figura 11 – Exemplo de ficha para registro de informações do sistema cantometrics (LOMAX, 1962)

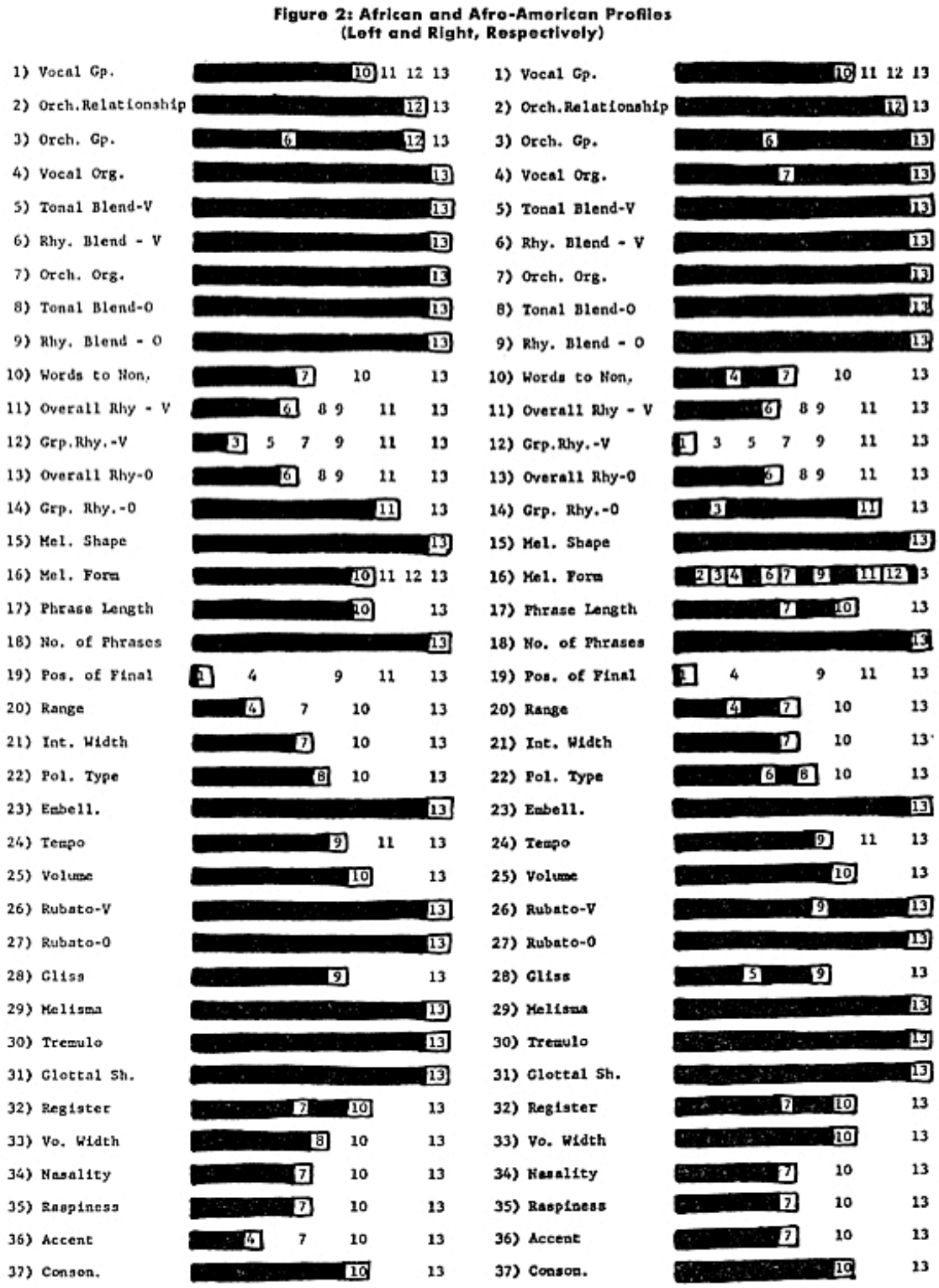


Figura 12 – Comparação entre dois perfis produzidos pelo sistema cantometrics
(LOMAX, 1962, p. 432)

Os parâmetros que podem ser colocados em correspondência com algum número, como, por exemplo, andamento, intensidade, registro e âmbito da melodia, são registrados em uma escala (rating) do maior para o menor, ou vice-versa, conforme a conveniência. Já as

características que não são quantificáveis, como forma melódica, ornamentação e melismas, são marcadas na escala conforme a maior ou menor incidência na amostra analisada. O resultado formatado dos registros na ficha torna-se um perfil (profile), que pode ser empregado em análises comparativas de músicas de diferentes culturas (fig. 12).

Por meio destas análises comparativas é possível, segundo o autor, identificar traços comuns entre diferentes culturas musicais e se obter uma estimativa de maior ou menor proximidade entre elas. Embora não explicite no texto, podemos concluir que o mesmo método também pode ser usado para comparar diferentes manifestações dentro de uma mesma cultura, e pode ser útil para descrever gêneros e estilos. Grauer (1965), que foi assistente de Lomax, publicou estudos ulteriores a partir da metodologia cantometrics e reuniu evidências do agrupamentos (clusters) de estilos musicais de culturas correlatas nas tabelas resultantes das análises.

Na verdade trata-se de uma descrição paramétrica de estilo, um tipo de análise que foi proposto por diferentes musicólogos, não só no âmbito da pesquisa etnomusicológica, mas também para o estudo do repertório ocidental (BENT, 1986). Usamos a proposta de Lomax para ilustrar e expor os princípios destes tipos de método por seu pioneirismo e também perla proximidade dos problemas da etnomusicologia com a questão dos gêneros e estilos. A preocupação em se elaborar uma tipologia de aspectos melódicos motivou trabalhos como os de Adams (1976), cujo desenvolvimento posterior teve grande avanço com a possibilidade de análise e transcrição de melodias por computador, mas cujos modelos teóricos sofreram grande impacto pela imensa capacidade de cálculo, que a um só tempo possibilitou e exigiu modelos matemáticos cada vez mais complexos.

Nos anos 70 Lomax também propôs a aplicação de método semelhante para descrever a fala de culturas estrangeiras registradas em gravações, que ele denominou *parlametrics* (LOMAX et al., 1977). Também foi elaborada uma ficha com parâmetros a serem marcados em uma escala pelo observador/ouvinte, desta vez com 50 linhas. A maioria dos parâmetros ou propriedades da fala que constam deste modelo são informações sobre o comportamento geral da fala na conversa, incluindo afetos e atitudes, mas há também informações sobre a prosódia, como padrões de acento, regularidade das sílabas, âmbito melódico, entre outros. A motivação fundamental deste novo método era fundamentalmente dar suporte a investigações sobre as relações sociais nos grupos estudados, não havendo comentários relevantes a respeito dos aspectos linguísticos.

Feld e Fox (1994, p. 35 e ss.) consideram que a importância das interações entre música e linguagem levaram ao desenvolvimento do que eles chamam etnografia músico-

linguística, citando entre outros, os trabalhos de List e Lomax.. Nos termos que estamos empregando neste trabalho, esta etnografia músico-linguística seria a investigação e comparação entre as culturas da organização da sua práxis vocal e a correspondente relação com as estruturas sociais, e a premissa básica desta abordagem é a hipótese de um *continuum* da fala ao canto, como a que estamos considerando. Para os autores, esta tendência contribui para a ênfase crescente das dimensões poéticas e pragmáticas da performance que se observa nos estudos da antropologia sociocultural, bem como para a crítica a uma concepção discursiva da construção social do significado (FELD; FOX, 1994, p. 43).

Aproximando-se o século XXI, a investigação sistemática de estilos e gêneros ganhou novo impulso com as possibilidades trazidas pelas análises automáticas feitas por computador, criando, por assim dizer, um novo campo, o da etnomusicologia computacional.

4.4. Etnomusicologia computacional

Embora não haja ainda um consenso estabelecido a respeito de uma disciplina independente que se possa chamar de etnomusicologia computacional, o fato é que desde o final do século XX o desenvolvimento de novas tecnologias para exploração da música registrada em arquivos de áudio digital tem apresentado um potencial para ampliar em muito a capacidade de analisar, classificar e comparar a música das mais diferentes culturas (TZANETAKIS et al., 2007). O interesse comercial que a música desperta, e a necessidade de lidar com uma quantidade imensa de gravações disponíveis circulando pela internet, faz com que haja um grande número de pesquisas visando ao desenvolvimento tecnológico para resolver este problema de ordem prática.

Já existe uma linha de investigação que emprega equipes interdisciplinares para desenvolver métodos para recuperar informações sobre a música registrada em arquivos de áudio digital, conhecida como Music Information Retrieval (recuperação de informações musicais, doravante MIR). Estes trabalhos são bastante recentes – a grande maioria das técnicas e métodos surgiu já no século XXI – mas seu desenvolvimento é exponencial. Os pesquisadores desta área formaram uma sociedade que mantêm encontros anuais há quinze anos, a International Society for Music Information Retrieval (ISMIR, 2014).

As técnicas de MIR empregam os métodos mais avançados de processamento de sinal, sistemas especialistas (machine-learning) e algoritmos que emulam a percepção humana, de maneira a detectar e analisar informações em um nível de detalhe além da percepção humana

e, ao mesmo tempo, realizar estas análises em uma massa de dados imensa (TZANETAKIS et al., 2007, p.6).

Nos artigos encontrados nos anais das conferências da ISMIR as investigações etnomusicológicas são cada vez mais frequentes, e há um constante intercâmbio entre as tecnologias e as abordagens etnomusicológicas. Mesmo assim, a maioria dos trabalhos ainda tem os modelos e padrões da música ocidental como objeto de investigação. Estes modelos estão, via de regra, baseados na notação musical convencional do ocidente, e se dedicam a identificação de padrões melódicos e reconhecimento de estilo (ANAN et al., 2011, HILLEAERE et al., 2012; MÜLLENSIEFEN; FRIELER, 2004; JUHÁSZ, 2009). Os dados neste caso já são as alturas discretas do sistema musical e suas respectivas durações, derivadas de uma partitura ou extraídas automaticamente do arquivo de áudio.

Existem, entretanto, outras técnicas que buscam obter informações diretamente do arquivo de áudio, sem empregar a mediação de um sistema de alturas conhecido, como as escalas da música ocidental. Estas técnicas são úteis para se investigar a música de culturas que não apresentam escrita musical e, especialmente, não possuem a noção de uma composição acabada, uma obra, sendo a improvisação e a variação na performance fatores fundamentais da prática musical que devem ser considerados. Com métodos estatísticos e algoritmos de busca, estas técnicas permitem identificar e classificar unidades musicais diferentes das ocidentais, como as fórmulas melódicas da música do Oriente (TZANETAKIS et al., 2007, p. 11; cf.; KODURI et al., 2012).

Uma das técnicas frequentemente usadas é o cálculo da estimativa da função densidade de probabilidade (probability density function, doravante PDF) para identificar categorias discretas em um determinado parâmetro, como a F0, por exemplo. O gráfico abaixo (fig. 13) representa uma estimativa média da probabilidade (eixo vertical) de se encontrar uma determinada frequência numa amostra gravada em arquivo de áudio digital estilo da música carnática, um estilo característico do sul da Índia (KODURI et al., 2012); os picos da curva permitem determinar a afinação dos graus da escala numa escala de centésimos de semitom (cents, no eixo horizontal).

Na música indiana, e em particular na chamada música carnática, não há uma padronização absoluta da afinação dos graus da escala, como acontece na música ocidental. Este método se baseia na estratégia de determinar estatisticamente quais seriam os alvos potenciais de um determinado parâmetro acústico (no caso, aqui, a F0), e pode ser empregado tanto na análise de música como da fala (cf. também TZANETAKIS et al., 2002).

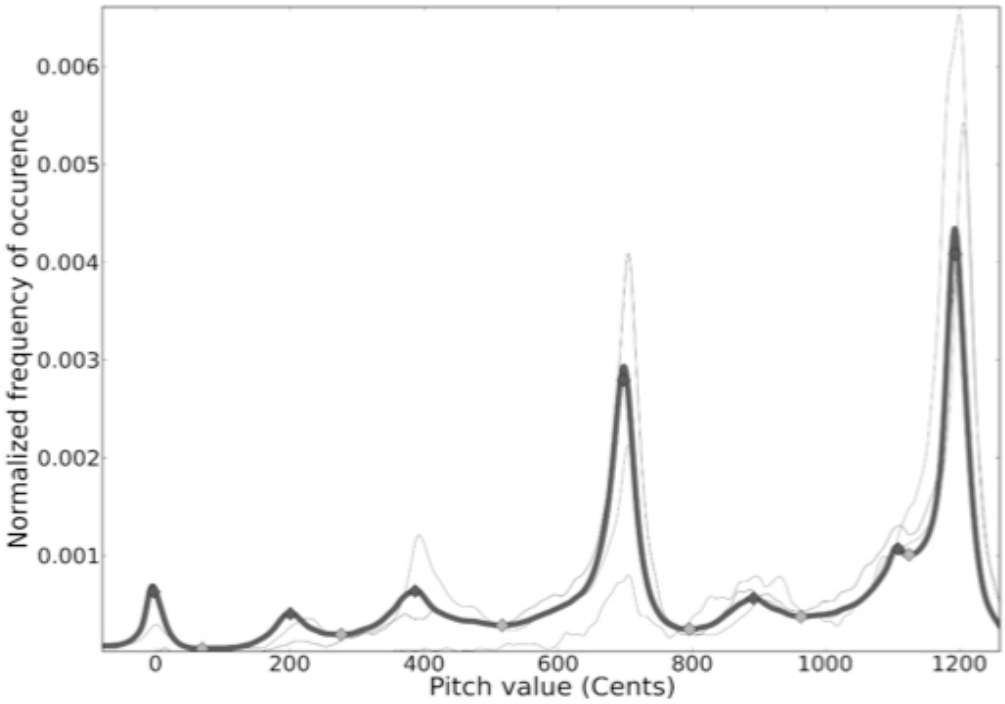


Figura 13 – Histograma de alturas do raaga Kaliyani; as abscissas dos picos correspondem à afinação dos graus da escala empregada, em centésimos de semitom a partir do primeiro grau (KODURI et al., 2012).

Uma outra técnica mais sofisticada é o uso de interfaces de inteligência artificial, como mapas auto-organizativos e sistemas especialistas, que "aprendem a reconhecer padrões". Ness e colegas (2002) apresentam um método em que um sistema especialista é alimentado com informações e anotações de especialistas na área de interesse que acessam a interface pela internet. Segundo os autores, em muitos domínios, os próprios pesquisadores ainda não formularam as questões mas estão diante de coleções de dados imensas. A técnica conhecida como *bootstrapping* consiste em segmentar e anotar manualmente trechos de gravações que podem corresponder a unidades significativas para que depois o sistema aprenda a reconhecer padrões semelhantes. Ness e colegas (*op. cit.*) mostram a possibilidade de uso em dois domínios completamente diferentes, a saber, o estudo de lamentos e salmodias da tradição judaica e o canto das orcas. Destacamos aqui o uso da estimativa da PDF para encontrar as alturas mais frequentes de cada recitante (sua "escala" particular) e a posterior discretização dos gestos vocais ou fórmulas melódicas em termos destes tons, como pode ser observado na figura 14 (NESS et alii, 2008).

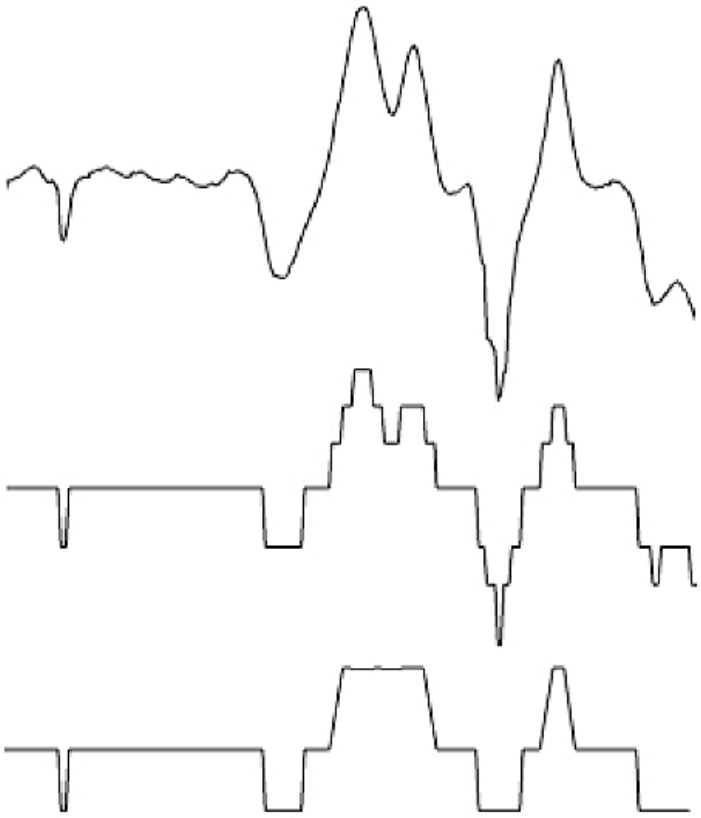


Figure 4: Melodic contours at different levels of abstraction (top: original, middle: quantized, bottom: simplified using 3 most prominent scale degrees

Figura 14 – Contornos melódicos obtidos pelo sistema Cantillion em diferenes níveis de abstração: em cima, a curva de F0 original; no meio a quantização ou discretização do contorno em termos dos graus da escala obtidos pelo histograma de alturas; e embaixo um esquema simplificado empregando os três graus mais significativos da escala (NESS et alii, 2008)

Note-se que neste processo os padrões são obtidos independentemente de uma notação prévia, o que permite a descoberta de detalhes ou padrões ainda não descritos pela notação tradicional. Se uma imagem vale por mil palavras, não podemos deixar de antever no diagrama da figura 14 as possibilidades de uso de um tal método na investigação de tons (lexicais ou não) na fala.

A descrição de estilos e gêneros por parâmetros acústicos, conceitualmente idêntica às propostas de List e Lomax (cf. supra) hoje pode ser aplicada a quantidades enormes de dados, usando técnicas como as redes neurais e os mapas auto-organizados, que encontram relações de similaridade e permitem construir uma topologia das manifestações musicais modeladas em um espaço multi-dimensional, que pode ser representada em um espaço bi-dimensional

(*idem*, *ibidem*, p. 12). Métodos sofisticados de cálculo numérico também podem executar tarefas tais como identificar um estilo de canto (KAKO et al., 2009).

As possibilidades de aplicação das técnicas de MIR à investigação etnomusicológica são inúmeras, mas ainda estão no estágio de "soluções à procura de um problema" pois na maioria das vezes as técnicas são testadas em *corpora* sem um objetivo musicalmente definido (TZANETAKIS et al., 2007, p. 12). No entanto, as perspectivas são promissoras e, considerando o fato exposto no início deste capítulo, de que a investigação da fala e da música vocal partilham problemas técnicos e conceituais, num futuro bem próximo estas técnicas ajudarão a revelar detalhes e reconhecer padrões da entoação de grandes *corpora* de enunciados orais. De fato, a técnica de histograma de alturas por convolução, característica da MIR, já é empregada na identificação de estilos de fala, e foi incorporada no algoritmo ExProsodia (FERREIRA NETTO et al., 2013).

4.5. Um problema persiste

Refletindo sobre o impacto das novas tecnologias na investigação etnomusicológica, Will (1999) observa que mesmo louvando a possibilidade de se libertar da notação convencional, permitida pela análise automática dos arquivos de áudio, a abordagem dos pesquisadores ainda é em grande medida influenciada ou balizada pelas categorias implícitas que derivam da notação. Para o autor, esta contradição reflete a ausência de uma ferramenta analítica alternativa, que acabou criando assim um "ponto cego" metodológico (op. cit., p. 4).

Will argumenta que esta situação se deve à onipresença da escrita no mundo que, ao longo dos séculos passou de simples recurso mnemônico a uma força poderosa que molda nossos pensamentos e influencia a fala e o uso da linguagem como um todo. Para ele, houve mudanças profundas nas habilidades cognitivas e nos processos de pensamento na passagem de culturas orais para culturas escritas[13].

Há evidências de que poetas de tradições orais tiveram suas habilidades prejudicadas ao aprenderem a ler e escrever, pelo fato de que a escrita introduz em sua mente a ideia de um texto controlando a narrativa e assim interfere no processo de composição oral (idem, ibidem, p. 5). Também a capacidade de segmentar o fluxo contínuo da fala é afetada pela aquisição da

13 Apesar de o autor não citar, mencionamos aqui o extenso estudo de Marshall McLuhan (19..), que parte da concepção dos meios de comunicação como extensões do homem.

escrita. Entre outros exemplos, Will cita o fato já demonstrado de que chineses que só leem a escrita em ideogramas, mas não leem o *pinyin*, uma adaptação do alfabeto romano usada para transcrição fonética do mandarim padrão, não conseguem detectar e isolar os fonemas que compõem uma palavra.

Estas constatações apontam para um problema epistemológico quando se trata de estudar a língua em uso na forma falada. A maneira como concebemos a linguagem é profundamente influenciada pela escrita, e mesmo que estejamos conscientes deste fato, não conseguimos nos livrar deste viés metodológico, assim como os etnomusicólogos não conseguem abrir mão da representação visual baseada na notação musical cocnvencional. Conforme aponta Will, "a forma escrita de uma língua provê um modelo, ou um conceito, no qual as categorias dos constituintes da forma falada, em combinação com a familiaridade do sistema de escrita, determina a consciência que é possível de se ter sobre a estruturação sintagmática de uma língua" (*loc. cit.*).

Will relembra que qualquer enunciado oral tem dois componentes: o que está sendo dito e uma indicação de como este deve ser entendido. Os sistemas de escrita tendem a lidar somente com primeiro componente. Assim, a escrita tende a diminuir nossa consciência a respeito das particularidades da língua falada, que não estão representadas graficamente.

É curioso encontrar uma explanação tão lúcida sobre um problema epistemológico premente nos estudos da linguagem no discurso de um etnomusicólogo. O objetivo do autor ali é usar este argumento para realizar uma crítica aos métodos que empregam recursos tecnológicos inovadores mas carregam implicitamente em suas premissas a organização do discurso musical baseada na notação ocidental. Mas seus argumentos nos levam a refletir sobre o problemas semelhante que encontramos no atual panorama dos estudos prosódicos.

Will considera ser essencial o uso das novas tecnologias, mas que a investigação deve partir tanto quanto possível da análise dos dados brutos, mediadas pelo conhecimento de que já se dispõe a respeito da percepção e cognição humanas, sem deixar de considerar as contingências culturais da percepção auditiva. Só assim será possível identificar e analisar os traços cognitivos significativos de uma certa cultura musical (ou linguagem, acrescentamos) e, por meio destes, descrever satisfatoriamente um comportamento musical (linguístico, cultural).

Por fim, o autor ainda faz a importante ressalva de que os dados brutos, as gravações, mesmo sendo adequadamente tratados ainda são incompletos no que diz respeito à compreensão do signo musical (ou linguístico, acrescentamos) como fato social total, pois isolam o som do contexto em que foi produzido. Ficamos sujeitos à atitude pós-moderna,

sintetizada na célebre definição de música enunciada pelo compositor italiano Luciano Berio: música é tudo aquilo que se ouve querendo que seja música (BERIO; DALMONTE, 1981). Esta fórmula resume a alienação e a fragmentação do conhecimento no mundo contemporâneo, e revela que, ao fim e ao cabo, as categorias não podem ser encontradas somente a partir do sinal acústico. O contexto em que um enunciado é produzido, seja linguístico, seja musical, determina as categorias, orienta a percepção e direciona a sua interpretação. Os novos métodos abrem vias fantásticas para a investigação do som em suas minúcias, mas continuarão a ser "soluções à procura de um problema" enquanto a dinâmica do contexto em que a linguagem – e em especial a fala – é produzida estiver fora do quadro teórico do programa de pesquisa, a não ser que nos contentemos com uma simples descrição do que acontece.

Considerando as evidências que reunimos nesses quatro primeiros capítulos, acreditamos que podemos ir além de uma simples descrição, e devemos buscar, nas propriedades notáveis da fala, um caminho para o aprofundamento do conhecimento da natureza da linguagem, e, naturalmente, do homem.

5. Reflexões sobre a natureza da entoação

5.1. Ritual e protocolo

A origem da palavra protocolo está ligada ao selo ou lacre que fechava os rolos de pergaminho na correspondência oficial da administração romana, indicando a origem e autenticidade do documento, e consequentemente a hierarquia envolvida e as providências que deveriam ser tomadas quanto ao seu conteúdo. *Protocollum* é uma palavra derivada do grego tardio *protokollon*, que quer dizer simplesmente "primeira cola" (HOUAISS, 2001).

Ao longo dos séculos a expressão assumiu múltiplos significados, que podem ser agrupados em duas acepções básicas: uma ligada ao correto encaminhamento de uma informação dentro de uma instituição (registro de protocolo, seção de protocolo), e outra ligada à ideia de cerimonial, de normas de procedimento (protocolo oficial, protocolo diplomático, "quebrar" o protocolo, etc.). Considerando as duas acepções a uma certa distância, é possível perceber que ambas estão unidas pela ideia de *regras de comportamento necessárias ao mútuo entendimento (ou concordância) no confronto entre duas entidades (pessoas ou instituições)*. Há portanto uma similaridade os dosi sentidos da palavra no que diz respeito ao fluxo de informação num contexto social, ou seja, não importa somente o conteúdo do que se transmite, mas também o modo como deve ser recebido e a sua influência sobre o destinatário daquela mensagem. Estamos no universo da pragmática.

Aproveitando nossa incursão aos dicionários, acabamos por constatar que a palavra "pragmática" também se refere ao conjunto de regras que regulam cerimoniais da corte e da igreja. No Direito tem também o sentido de um "conjunto de regras relacionadas com a prática social e jurídica, em oposição a palavras e fórmulas" (*idem*, *ibidem*).

Chama a atenção esta aproximação entre pragmática e protocolo, ainda mais se considerarmos a definição semiótica da primeira: "estudo das relações causais entre as palavras, expressões, símbolos e seus usuários" (HOUAISS, 2001). Morris (apud CARLSON, 2010, p. 75) definiu a pragmática como "a ciência da relação entre signos e intérpretes". Trazendo de volta a concepção de competência linguística de Bourdieu, estamos falando dos aspectos fundamentais da comunicação linguística relacionados à conquista do direito de dizer, de ser ouvido, de "impor recepção".

No mundo contemporâneo, da internet, o termo "protocolo" passou a ter uma terceira acepção, não muito distante das outras duas, mas com uma natureza técnica, precisa, fria

como as linguagens de computador. Douglas Comer define protocolo como sendo "uma descrição formal do formato das mensagens e das regras que duas ou mais máquinas [computadores] devem seguir para trocar essas mensagens" (BACCALA, 1997). Os protocolos podem ser de dois tipos: ou eles são textos para serem compreendidos por seres humanos, ou eles são parte de um código de programação, para serem entendidos por computadores. Nos dois casos é preciso que eles especifiquem a interpretação precisa de cada bit de cada mensagem trocada através de uma rede (*idem*, *ibidem*). Os protocolos são necessários toda vez em que dois computadores ou programas diferentes precisam concordar em como eles comunicar a informação entre eles (*idem*, *ibidem*).

Nesse contexto, protocolos são uma parte da informação que a organiza e indica a sua interpretação. São elementos que fazem parte de um arquivo ou mensagem que identificam a natureza (e às vezes a origem) destes e também como devem ser decodificados. O protocolo, na área da tecnologia da informação tem assim uma dupla função: organizar a informação e indicar o que deve ser feito com ela.

Voltando às noções de protocolo como cerimonial e organização da informação, podemos propor que, assim como o protocolo tem função cerimonial, este determina também a forma como o conteúdo linguístico é apreendido (impõe recepção) e quais serão as consequências para os interlocutores. E não encontramos impedimento para afirmar que a entoação (e a prosódia, provavelmente) desempenha o papel de um protocolo, um elemento da enunciação que determina sua interpretação e direciona a reação do interlocutor.

Essa nova perspectiva abre caminho para uma distinção objetiva entre manifestações orais, combinada com a correspondência de seus estilos com gêneros estabelecidos do discurso oral. E, traçando um paralelo com o estudo de Greimas sobre a gestualidade humana, podemos assumir uma divisão fundamental neste campo, a saber, a divisão entre manifestações orais práticas e míticas (algo semelhante à divisão proposta por Bakhtin entre gêneros primários e secundários).

Assim como os movimentos e atitudes dos participantes de um ritual (religioso ou não) ou uma cerimônia são determinados por regras, da mesma forma os movimentos melódicos de uma enunciação mítica são também convencionados e obedecem à prescrições de gênero e estilo. Lembre-se aqui da concepção de Searle para a linguagem: um comportamento regido por regras. Ora, um conjunto de regras para o comportamento em cerimônias é denominado protocolo. Podemos também estender a analogia e pensar na entoação da fala cotidiana como uma espécie de protocolo melódico para os enunciados proferidos em situações sociais específicas. Vale lembrar que aqui estamos pensando ritual na

acepção mais ampla possível, que pode ser um ritual religioso, mas também estético, lúdico, etc.

Aqui, por meio desta aproximação assaz inusitada, pensamos ter encontrado o elo perdido entre a enunciação prática e a enunciação mítica. As enunciações míticas, embora façam parte da práxis oral-verbal, não têm o objetivo imediato de interferir nas relações pessoais; expressam, antes, um desejo de transformação da realidade por meio da representação e da mimese. As enunciações práticas pertencem à esfera do cotidiano; são pedidos, ordens, sugestões, perguntas, e mesmo os comentários e declarações aparentemente inocentes têm como objetivo influir no comportamento do interlocutor – são, acima de tudo, ações, atos de fala.

As enunciações míticas têm uma natureza ligada ao ritual e por este motivo têm um alto grau de convenção. Aqui os estilos são bem definidos e os gêneros relativamente bem delimitados, não só quanto às possibilidades lexicais e melódicas, mas também quanto às situações sociais em que podem acontecer. As relações entre texto e melodia são definidas a priori, mesmo que permitam alguma flexibilidade na realização, como é o caso do salmodiar, por exemplo. De qualquer modo, existem regras que prescrevem como o texto deve receber a melodia (ou a melodia receber o texto?).

Na enunciação prática, o protocolo tem a natureza de um protocolo de comunicação; na enunciação mítica, ele é um protocolo de cerimonia. Em ambos os casos podemos dizer que são elementos de uma pragmática, ora convencional, cerimonial, ora cotidiana, livre. Esses protocolos de entoação têm, como já sabemos, uma dupla função, envolvendo produção e recepção do enunciado. No lado da produção, as regras determinam a formação da melodia e a colocação do texto; no lado da recepção, o conhecimento intuitivo das regras pelo ouvinte orientam a decodificação e a interpretação do que o emissor pretende com aquele enunciado.

Percebe-se que a aproximação semântica entre protocolo e pragmática coloca novas perspectivas para a interpretação da relação entre entoação e gênero do discurso oral, sendo possível estabelecer novas conexões com o conceito de práxis vocal. Sendo produzidas acusticamente da mesma maneira, tanto a fala cotidiana como outras situações de fala mais formais são semelhantes ao canto – cada estilo é adequado a uma certa situação: algumas mais informais, improvisadas, outras formais, quase rituais, seguindo um rigoroso protocolo.

5.2. Intenção e gênero de discurso

A possibilidade de se descrever o estilo melódico de uma manifestação vocal qualquer por meio de parâmetros acústicos corrobora a proposta de uma super-categoria do trabalho humano que seria a práxis vocal (ou oral), um subconjunto da práxis sonora. Essa práxis oral compreende desde o canto em todas as suas formas até a fala cotidiana, passando por situações intermediárias. A divisão entre gêneros primários e secundários, proposta por Bakhtin, dá lugar a um continuum em que é difícil estabelecer fronteiras.

E mais: a parametrização do estilo permite a idealização de um espaço multi-dimensional no qual podemos localizar essas diferentes manifestações segundo relações de similaridade, à maneira da proposta esboçada por List.

Nessa perspectiva, podemos imaginar estratégias de investigação da entoação empregando técnicas de identificação de estilo a partir de arquivos de áudio digital, como acontece nos trabalhos de MIR. Como essas técnicas são automatizadas, é possível analisar uma grande quantidade de dados, previamente anotados quanto à situação de fala a que correspondem. Aos poucos, sistemas especialistas aprenderão a reconhecer estilos já analisados e até identificar estilos por similaridade melódica entre as gravações.

Estas técnicas provavelmente já estão em testes, tamanha é a velocidade com que se desenvolvem as tecnologias voltadas para a comunicação, especialmente a comunicação homem-máquina. Mas a percepção de intenção ainda continua um quebra-cabeças que exige um imensa capacidade de processamento e memória que, por enquanto, só o ser humano possui.

Mas esse mesmo ser humano na busca incansável pelo conhecimento poderá decifrar como ele próprio reconhece as sutilezas da comunicação e transferir este conhecimento para uma máquina. Acreditamos que a parametrização dos estilos (seja de fala ou canto) e a classificação dos respectivos gêneros é um caminho que aponta para esta realidade.

E segundo as evidências reunidas neste trabalho, somos levados a concordar com Bakhtin e Bourdieu no sentido em que o fato primordial na interação das pessoas pela linguagem é a vontade de comunicar, o desejo de se fazer entender, e é desse impulso que nasce a entoação, procurando então o gênero e o estilo que satisfaçam sua vontade comunicativa para nela inserir o que costumamos chamar de conteúdo linguístico. Chegamos, pois a uma reflexão que se situa no campo da filosofia da linguagem, corroborando o que sustentamos ao longo desse percurso, que para decifrar a entoação é necessário um mergulho profundo em busca da natureza da linguagem.

5.3. O paradoxo da fala espontânea

> "Mas como? dirá alguém, estes acentos tão plangentes, tão dolorosos, que esta mãe arranca do fundo de suas entranhas, e com os quais as minhas são tão violentamente sacudidas, não é o sentimento atual que os produz, não é o desespero que os inspira? De modo algum; e a prova é que são medidos, que fazem parte de um sistema de declamação; que mais baixos ou mais agudos do que a vigésima parte de um quarto de tom, são falsos; que estão sujeitos a uma lei de unidade; que são, como na harmonia, preparados e preservados: que satisfazem todas as condições requeridas apenas através de um longo estudo; que concorrem para a solução de um problema proposto; que, para ser levados ao ponto justo, foram ensaiados cem vezes e que, apesar desses freqüentes ensaios, ainda lhes falta algo (…)" (DIDEROT, 1987, p. 361)

O trecho acima transcrito pertence ao clássico da literatura sobre o fazer teatral "Paradoxo do Comediante", escrito por Denis Diderot em 1769 e considerado um marco da reflexão sobre o trabalho do ator. O argumento central do texto é o de que o ator não deve se envolver com os sentimentos suscitados pelo texto, ou, em outras palavras, não deve sentir o que a personagem sente; ao contrário, deve se esmerar em criar no palco o artifício exato que faça o espectador sentir a referida emoção. Estes artifícios devem ser praticados exaustivamente para que sejam executados à perfeição; portanto, seguem um certo conjunto de convenções. Mas de onde resultam estas convenções? Da imitação das reações somáticas (incluindo alterações na fala) correspondentes a essas emoções, e posterior sistematização e exagero desses traços característicos.

Por "emoções" aqui entendemos que o autor emprega uma noção bastante abrangente que pode incluir desde estados psicológicos até convenções de polidez, atitudes proposicionais e modalidades. E a sua sistematização resulta da observação do efeito de cada gesto, cada entoação, cada alteração de qualidade vocal em si e nos outros, seguida de diligente repetição da sua imitação. Assim, um sistema de declamação é criado, constituído de convenções, e, portanto, regras; porém seus enunciados estão sujeitos a critérios de verdade que são ditados pela natureza, ou, mais especificamente, pelo comportamento da fala em situações não-encenadas. Estamos, portanto, diante de mais uma manifestação do grande impasse que resulta quando refletimos acerca da relação (e consequente distinção) entre natureza e cultura, especialmente no domínio da comunicação humana e, mais especificamente, da linguagem. Chegando ao final deste percurso que realizamos ao longo da tese, contemplamos este *magnum mysterium* do ponto de vista dos estudos prosódicos, com o objetivo de refletir sobre a adequação do conceito de fala espontânea e a distinção tão evidente, acusticamente (mas somente acústica?), entre a leitura em voz alta, a fala decorada e

a fala não-decorada. Estas diferenças podem ser descritas por meio dos parâmetros de altura, duração e frequência; nossa abordagem, entretanto, privilegia a altura dentre os demais, por razões que ficaram claras mais adiante.

Trazendo a questão apresentada para o domínio do uso da fala na comunicação cotidiana, podemos iniciar o reflexão com uma pergunta: como percebemos, com relativa facilidade, na curva melódica que alguém está lendo ou falando um texto decorado, e não falando de improviso? Por que a melodia resultante nestas situações nos parece, com frequência, "falsa"? Quando assistimos a um espetáculo de teatro com atores inexperientes, é bastante comum haver momentos em que perdemos o sentido de uma frase, não por deficiência na dicção, ou por uma voz débil, mas por causa de incoerências na entoação, que acontecem porque o ator impõe uma melodia que não corresponde à situação em que a personagem está falando. Ora, se acontecem incongruências é porque existe uma (ou talvez mais de uma, mas não qualquer uma) entoação apropriada para aquela sentença naquela situação. Porém, como podemos saber que aquela curva melódica que ouvimos é incoerente com o que está sendo dito, se estamos ouvindo (ou pelo menos, mesmo que estejamos ouvindo) pela primeira vez e não há um termo de comparação?

Há duas respostas possíveis, que provavelmente estão conectadas de alguma forma. A primeira é a hipótese da existência de regras de boa formação (*well-formedness rules*) para a constituição da melodia da fala, regras estas que aparentemente são violadas com frequência na leitura em voz alta e na fala decorada, mas que podemos perceber, mesmo que intuitivamente. Aliás, é justamente aí que reside o paradoxo: para reconhecermos as violações das regras de boa-formação, é necessário conhecê-las, mesmo que intuitivamente. Mas, aparentemente não conseguimos, a não ser com muito treino e observação (da fala espontânea...), reproduzir o funcionamento natural da língua numa situação de fala forçada ou preparada (leitura em voz alta ou fala decorada)

No contexto dos estudos prosódicos, a fala espontânea é considerada um dos modos possíveis em que a língua pode se manifestar. Entretanto, é raro encontrar uma reflexão mais aprofundada sobre uma definição mais estreita. Percebe-se uma certa hesitação dos estudiosos da prosódia em se analisar a fala espontânea, em parte porque seu comportamento é muito difícil de controlar experimentalmente, e em parte por que o sua aparência é quase a de um caos total.

A área do conhecimento que mais avançou neste sentido é aparentemente o conjunto de disciplinas que servem ao grande projeto que podemos chamar de Reconhecimento Automático da Fala (em inglês *Automatic Speech Recognizing*, ASR)

> "Do ponto de vista da enunciação, a fala espontânea pode ser definida como um 'enunciado concebido e percebido ao longo de (durante a) sua enunciação' (Luzzati, 2004), ou seja, um enunciado produzido para um interlocutor real por um enunciador que improvisa; isto implica que as correções se traduzem necessariamente por prolongamentos da mensagem, A fala preparada (aquela empregada pelos jornalistas que apresentam as informações radiofônicas ou televisivas) é uma fala produzida para um interlocutor mais ou menos fictício, por um enunciador que tem a habilidade (...) de produzir enunciados que não necessitam ser retomados ou corrigidos, ou que é capaz de mascarar os erros. Deste ponto de vista, compreende-se que podemos falar igualmente de uma fala conversacional, não-premeditada ou co-construida."[14] (Bazillon et alii, 2008)

Vê-se porém que a acepção de fala espontânea é a de uma fala com erros, disfluente, que foge à norma linguística. Mas é justamente o contrário que defendemos aqui. Apesar de sua aparente precariedade, na fala espontânea é que podemos observar a entoação em sua plenitude, e se não conseguimos ainda decifrar a sua lógica, o problema não está no objeto , e sim no método de investigação e nas premissas que o orientam.

Em seu célebre artigo, "In defense of lab speech", Yi Xu emprega a definição de Beckman.

> "Spontaneous speech, according to Beckman (1997:7), is "speech that is not read to script". She further distinguishes between ten different types of spontaneous speech recordings, ranging from *unstructured narrative* to *instruction monologues*. The dividing line between lab speech and spontaneous speech can sometimes be blurred. For example, even when recording unscripted speech, certain levels of control can be implemented. In what is referred to as *instruction monologues*, the speaker is asked to instruct a real or imaginary silent listener to perform a task. With this technique, some control over both content words and syntactic structure can be achieved (Beckman, 1997). To the extent the level of control is achieved, this type of speech could be labeled as lab speech as well" (XU, 2010).

Ou seja, fala espontânea é um fenômeno de múltiplas faces, dependendo muito das condições em que ocorre, porém não totalmente aleatório, visto que pode ser classificado e até ordenado em termos de um "grau de espontaneidade". Na discussão sobre a fala de laboratório, Xu defende que, apesar de o controle imposto pelas condições do experimento reduzir as variações que observamos na fala dita natural, ele é necessário para que se possam testar hipóteses que levem à descrição e explicação do fenômeno.

Vejamos o que o autor diz a respeito da fala soar não-natural (ou antinatural; repare-se

---

14 D'un point de vue énonciatif, la parole spontanée peut se définir comme un « énoncé conçu et perçu dans le fil de son énonciation » (Luzzati, 2004), c'est-à- dire un énoncé produit pour un interlocuteur réel par un énonciateur qui improvise ; cela implique que les corrections ne peuvent se traduire que par un prolongement du message. La parole préparée (celle qu'emploient les journalistes présentant les informations radiophoniques ou télévisées) est une parole produite pour un interlocuteur plus ou moins fictif, par un énonciateur qui en possède la maîtrise, qui est capable de produire des énoncés qui n'ont plus à être repris ou corrigés, ou qui est capable de le masquer. De ce point de vue, on comprend qu'on puisse parler également de parole conversationnelle, non préméditée ou co- construite.

que ele emprega também o termo estilo, para se referir às mudanças na melodia, relacionando-o com a situação de fala):

> "So, if a person's speech recorded in the lab does sound unnatural, it is neither because it is not conversational, nor because it is non-human. What, then, may have made the person's speech sound unnatural? Is it because s/he is doing something that has never been done in his/her life? Probably not. It is more likely that the person has spontaneously assumed a speaking style that is appropriate for the occasion, i.e., reading aloud text in a formal setting. Such a style shift is not artificial, but a rather *natural* adjustment to the situation" (XU, 2010).

Aqui temos um outro aspecto do paradoxo da fala espontânea: como julgar se uma fala é natural ou não, se a fala é sempre dinâmica e se adapta muito rapidamente às mudanças de contexto.

O objetivo aqui era expor o paradoxo, e não propor alguma solução para elucidá-lo. Provavelmente é mais um nó górdio da ciência, como o teorema de Gödel e o princípio da incerteza de Heisenberg. O fato é que a constatação do paradoxo da fala espontânea coloca uma realidade, segundo nossa opinião, intransponível, que pode ser resumida da seguinte maneira: não é possível imaginar uma entoação neutra, que sirva de baliza para avaliar as outras melodias observadas. Como afirmou Bourdieu, sempre que há um enunciado há uma pessoa se colocando em relação a outra.

Acreditamos que toda investigação da entoação precisa levar este dado em conta ao planejar seus experimentos e propor seus modelos teóricos. Ficamos praticamente divididos entre duas afirmações diametralmente opostas: ou não há fala espontânea, ou toda fala tem algum grau de espontaneidade. Um paradoxo, enfim

5.4. A precedência da entoação

Chegando ao final deste percurso, embora tenhamos dito não pretender elucidar o paradoxo, apresentamos uma possível hipótese que, no entanto, até onde podemos perceber, não pode ser falseada, e consequentemente não seria possível de submeter (ainda) a um teste experimental.

Se pensarmos na origem da poesia, em que um esquema rítmico-melódico era preenchido pela cadeia segmental, ou então no canto dos repentistas do Nordeste brasileiro, que inventam na hora os versos para uma melodia pré-estabelecida, perceberemos que a situação original da poesia e do canto é a de ter-se determinada a priori a curva melódica, para nela se encaixar em seguida a cadeia segmental. No caso do repente, por exemplo, há uma

diversidade de gêneros com diferentes padrões silábicos e melódicos, cada um recomendado para diferentes temas (martelo agalopado, galope à beira mar, dez pés a quadrão, mourão, gemedeira, etc.). Exceto no caso da composição da tradição ocidental erudita, em que a construção da melodia é feita a partir do texto, na maioria das outras culturas a melodia e o padrão rítmico precedem a composição do texto.

Concordando com Bakhtin, quando afirma que "quando escolhemos as palavras, partimos do conjunto projetado no enunciado, e esse conjunto que projetamos e criamos é sempre expressivo, e é ele que irradia a sua expressão" (cf. seção 4.2) vemos como mais provável, para a situação de fala espontânea, o uso de certas fórmulas melódicas (que o ouvido atento percebe constantemente no cotidiano) adequadas às diferentes situações de fala, relações de hierarquia, vontade comunicativa, enfim, um gênero corresponde a uma intenção, que é o conjunto de onde parte o processo de enunciação oral espontânea (esse entendido com uma resposta real do falante a um estímulo real, numa situação real).

Quando não há o estímulo real, simplesmente conectamos as palavras sem ter um fio que as suporte, e o resultado acaba sendo opaco, sem vida, como a fala do ator inábil. O ator habilidoso aprende truques para fornecer a melodia que sustentará as palavras: ou aprende melodias que funcionam, ou emprega algum subterfúgio cognitivo que cria a ilusão de que ele está, de fato, concebendo aquele enunciado ali, naquele momento. São os truques do ofício.

Apesar de não fornecer nenhum resultado concreto neste trabalho, a exposição deste paradoxo traz uma contribuição importante para a compreensão do funcionamento da fala, que ainda pode se valer do desenvolvimento tecnológico para, num futuro não muito distante, testar as hipóteses que dele decorrem.

Conclusão

A entoação é, muito provavelmente, o elemento da linguagem mais difícil de se delimitar enquanto objeto de investigação científica. As múltiplas informações que estão codificadas por meio da variação da altura e o caráter gradiente de seus traços característicos levantam controvérsias quanto à própria natureza do fenômeno e desafiam os pesquisadores com impasses epistemológicos complexos que muitas vezes passam despercebidos nas discussões metodológicas.

Não obstante essas dificuldades, o conhecimento a respeito da entoação avançou muito nas últimas décadas, em parte pela inovação tecnológica que permite análises por computador de uma grande quantidade de amostras de fala disponíveis em gravações digitais. Por outro lado a teoria métrica-autossegmental teve sucesso na incorporação de fatos relacionados à variação de F0 a um modelo fonológico, o que abriu caminho para uma descrição fonológica da entoação.

Mesmo assim, ainda é difícil criar situações experimentais que permitam testar empiricamente hipóteses a respeito de aspectos isolados da entoação. Os diferentes papeis que a variação da F0 desempenha na comunicação oral-verbal estão de tal modo interligados que há sempre uma influência mútua entre eles e é praticamente impossível, pelo menos até o momento, estabelecer correspondências unívocas entre determinadas melodias e determinadas utilizações da língua. Quando levada para o laboratório, a fala perde a naturalidade que dá a riqueza da sua melodia. Este fato foi poucas vezes mencionado na literatura especializada a que tivemos acesso, e talvez não seja possível testá-lo empiricamente enquanto não houver um modelo teórico que explique o funcionamento da entoação na fala espontânea.

Porém, nossa experiência com o teatro e com o treinamento do ator nos mostra evidências de que existe uma lógica no comportamento aparentemente caótico da entoação na fala espontânea. O fato de reconhecermos que certa fala de um ator está “com a intenção errada” é, para nós, uma evidência de que existem regras de boa-formação da melodia da fala. Ao mesmo tempo, a percepção de uma “intenção errada” é indício de que aquele enunciado não foi concebido no momento da sua realização, mas foi preparado, antecipado, e por isso resulta “artificial”. Na fala espontânea observamos disfluências, palavras erradas e interrupções, mas muito raramente temos a sensação de que o falante “errou a intenção”, como acontece com os atores na fala decorada.

É isto o que denominamos o paradoxo da fala espontânea: só na fala espontânea

podemos testar a língua de fato em funcionamento, porém não é possível concluir empiricamente que um determinado enunciado é um exemplo de fala espontânea, e, mesmo que fosse possível, os elementos que constituem a melodia da fala estão de tal forma imbricados que é muito difícil isolar algum aspecto. Podemos, no máximo, produzir condições que normalmente resultem em fala espontânea (como as tarefas de nomear e descrever figuras, por exemplo) mas o quão preparado o falante está ao proferir um enunciado é uma variável que dificilmente pode ser controlada no contexto experimental.

Esse aspecto negativo de nossa conclusão é, acreditamos, temporário. O desenvolvimento de novas tecnologias para a análise automática de arquivos de áudio digital com métodos estatísticos já apresenta perspectivas de, em breve, vasculhar grandes massas de dados à procura de padrões rítmicos e melódicos. Porém, a reflexão sobre a natureza destes padrões e de sua influência mútua é fundamental para que esta exploração seja eficaz. Vimos que a língua desempenha seu papel nos atos de fala, na interação dos seres humanos, na luta pelo poder, e é aí que ela mostra sua face, ou melhor, suas faces. A linguagem tem um dinamismo inigualável por qualquer outro comportamento humano. Continuamente se modifica, moldando-se de acordo com o contexto, com o interlocutor, com as necessidades e intenções na comunicação. É mister fazer-se entender; "impor recepção", diria Boudieu.

Consideramos portanto que a fala deve ser estudada como um comportamento e, como tal, não pode ser dissociado das condições que promovem o seu aparecimento. E esta consideração tem um outro desdobramento que talvez seja a conclusão mais importante deste trabalho. Assim como o paradoxo da fala espontânea (e, de fato, está intimamente ligada a ele), essa conclusão não pode ser testada empiricamente, pelo menos por enquanto, mas as evidências apontam para a formulação dela como hipótese em busca de, no futuro, encontrar condições de falseabilidade.

Na verdade, Bakhtin já enunciou despretensiosamente essa hipótese, como vimos no segundo capítulo. Disse ele que "não enfileiramos as palavras, não vamos de uma palavra a outra, mas é como se completássemos com as devidas palavras a totalidade"(cf. cap. II). Nesta frase está implícita a ideia de que o perfil melódico, a *gestalt* prosódica, é que ordena a conexão dos constituintes. Aparentemente, tudo se passa como se a sequência fosse: intenção comunicativa (reação a um contexto específico) – escolha do perfil melódico (mediado pelo conhecimento dos gêneros) – determinação dos pontos de contato com a cadeia segmental (pelas regras do acento) – execução do programa produzindo a fala em tempo real.

Como já foi dito, esta hipótese ainda não pode ser testada, mas acreditamos que não estamos muito longe disso. Já existe um estudo sistemático do funcionamento do cérebro no

domínio da linguagem, e muito já se conhece sobre a condução de tarefas simultâneas e coordenadas em diferentes áreas do cérebro.

Ferreira Netto (2006) argumentando a partir dos resultados conhecidos sobre a lateralização das tarefas relacionadas com a percepção auditiva no cérebro, coloca a hipótese de um duplo foco resolutivo, que corresponderia, em linhas gerais, segundo nossa interpretação, a dois mecanismos de controle da F0, um ligado ao acento de palavra e outro ligado à forma geral do enunciado. Se, num futuro próximo, for possível localizar no cérebro quais áreas são responsáveis por cada um desses mecanismos, poderão ser idealizados experimentos para descobrir se um é ativado antes do outro. Caso seja possível, poderemos comprovar se há de fato a precedência da entoação em relação à ordenação dos segmentos e sintagmas num enunciado, e assim investigarmos mais a fundo os meandros do paradoxo da fala espontânea.

Porém, um experimento dessa natureza ainda não revelaria a lógica da entoação, apenas seria uma proposta de explicação do funcionamento da fala. Pelos conhecimentos que pude coligir ao longo de mais de quatro anos de pesquisa (e pelo menos dez anos de observação da fala), dos quais apenas uma parte está nesta tese, ainda há uma longa estrada a percorrer até começarmos a divisar no horizonte a última fronteira da ciência: o comportamento humano.

"*when skies are hanged and oceans drowned, the single secret will still be man*".

e. e. cummings

Referências

ARAÚJO, Samuel; PAZ, Gaspar. Música, linguagem e política; repensando o papel de uma práxis sonora. Terceira Margem. Número 25, pp. 211-231. Rio de Janeiro, 2011

ADAMS, Charles R. Melodic contour typology. Ethnomusicology, v. 20, nº 2, pp. 179-215. University of Illinois Press, 1976. Disponível em <http://www.jstor.org/stable/851015>. Acessado em 11/03/2013

ANAN, Yoko; HATANO, Kohei; BANNAI, Hideo; TAKEDA, Masayuki. Music genre classification using similarity functions. 12th International Society for Music Information Retrieval Conference (ISMIR), 2011.

AUSTIN, John. L. How to do things with words. Oxford: Oxford University Press, 1962.

BACCALA, Brent. Connected: an internet encyclopedia. Site. 1997. Disponível em: <http://www.freesoft.org> Acesso: Maio/2014.

BAKHTIN, Mikhail. Estética da criação verbal. São Paulo: Martins Fontes, 2003.

BARBOSA, Plínio A. From syntax to acoustic duration: a dynamical model of speech rhythm production. Speech Communication, 2007

BAZILLON, Thierry et al. La parole spontanée: transcription et traitement. Revue TAL. V. 49, nº3, pp. 1-26, 2008

BENT, I.D. Analysis. In: The New Grove Dictionary of Music and Musicians (ed. Stanley Sadie), Oxford: Oxford University Press, 1986

BERIO, L.; DALMONTE, R. Entrevista sobre a música contemporânea. São Paulo: Civilização Brasileira, 1981

BOLINGER Dwight. Intonation and Gesture. American Speech, Vol. 58, No. 2, 1983, pp. 156-174. Duke University Press

BOURDIEU, Pierre. The economics of linguistic exchange. Social Sciences Information. V. 6, nº 6, pp. 645-668

BROWN, Steven. The ‘musilanguage’ model of music evolution. In: WALLIN, N. L.; MERKER, B.; BROWN, S. (org.). The origins of music (pp. 271-300). Cambridge, MA: MIT Press, 2000

CALENGANO, Mauro. "Imitar col canto chi parla": Monteverdi and the Creation of a Language for Musical Theater. University of California/ American Musicological Society, 2002. Disponível em: <http://www.jstor.org/stable/831883>. Acesso em 21/09/2010

CARLSON, Marvin. Performance: uma edição crítica. Belo Horizonte: Editora da UFMG, 2010.

CRESTI, Emanuela. The definition of focus in Language into Act Theory (LAcT). In: MELLO, Heliana; PANUNZIO, Alessandro; RASO, Tommaso (org.). Pragmatics and prosody: illocution, modality, attitudes, information patterning and speech anotation. Firenze: Firenze University Press, 2011

DIDEROT, Denis. O paradoxo do comediante. In: DIDEROT, Denis. Textos escolhidos (Coleção Os pensadores). São Paulo: Abril Cultural, 1987

DEUTSCH, Diana. The enigma of absolute pitch. Acoustics Today. V. 2, nº 4, p. 11-19. Acoustical Society of America, 2006

ECO, U. A Estrutura Ausente. São Paulo: Perspectiva, 1991

FEDORENKO, Evelina; PATEL, Aniruddh; CASASANTO, Daniel; WINAWER, Jonathan; GIBSON, Edward. Structural integration in language and music. Memory & Cognition. Nº 37, v.1. pp. 1-9. 2009

FELD, Steven; FOX, Aaron A. Music and language. Annual Review of Anthropology, v. 23, pp. 22-53. Annual Reviews, 1994. Disponível em <http://www.jstor.org/stable/2156005>. Acesso: 11/03/2013

FERREIRA NETTO, Waldemar. Variação de freqüência e constituição da prosódia da língua portuguesa. Tese de Livre-Docência/USP, 2006.

FERREIRA NETTO, Waldemar; CONSONI, Fernanda. Estratégias prosódicas da leitura em voz alta e da fala espontânea. Alfa Revista de Linguística. V. 52 (2), pp. 521-524. São Paulo: UNESP, 2008. Disponível em www.alfa.ibilce.unesp.br. Acesso em

FERREIRA NETTO, Waldemar; SOUZA, André R.; VIEIRA, Maressa de F.; PERES, Daniel O.; MARTINS, Marcus V. M. Desenvolvimento de algoritmo de análise automática da curva de frequência por meio de convoluções gaussianas do histograma de alturas. Anais do IV Congresso Brasileiro de Prosódia da Fala. Maceió: UFAL, 2013

FITZPATRICK , Jennifer. On intonational typology. In: SIEMUND, Peter (ed.). Methodological issues in language typology. Sprachtypologie und Universalienforschung. 2000

FÓNAGY, Ivan (2003). Des fonctions de l'intonation: essai de synthèse. Flambeau, 29:1-20 (http://ed268.univ-paris3.fr/lpp/pages/EQUIPE/vaissiere/fonagy/articles/).

FOX, Anthony (2000). Prosodic features and prosodic structure: the phonology of suprasegmentals. Nova Iorque: Oxford University Press.

GOLDSMITH, John A. The handbook of phological theory. Cambridge: Blackwell, 1995

GOODY, J.; WATT, I. As conseqüências do letramento. São Paulo: Paulistana, 2006.

GRAUER. Victor A. Some song-style clusters – a preliminary study. Ethnomusicology. v. 9, nº 3, pp. 265-271. University of Illinois Press, 1965. Disponível em

<http://www.jstor.org/stable/850237>. Acessado em 11/03/2013.

GREIMAS, A. J. Conditions d´une sémiotique du monde naturel. In: Langages, n˚10. Paris: Didier/Larousse, 1968, p. 4-35

GROVE ONLINE. Site. Oxford University. Disponível em <http://www.oxfordmusiconline.com>. Acesso: Maio/2011

HATFIELD, Hunter. Joshua Steele 1775: speech intonation and music tonality. Disponível em <www2.hawaii.edu/~hunterh/Docs/JoshuaSteel.pdf>. Acessado em agosto/2010

HENRIQUE, Luís. Acústica musical. Lisboa: Fundação Calouste Gulbenkian, 2002

HIRSCHBERG, Julia. The pragmatics of intonational meaning. Speech Prosody, 2002

HIRST, Daniel; DI CRISTO, Albert. Intonation systems: a survey of twenty languages. Cambridge: Cambridge University Press, 1998

HOUAISS, Antônio. Grande dicionário da língua portuguesa. Rio de Janeiro: Objetiva, 2001

ISMIR. International Society for Music Information Retrieval. Site da entidade. Disponível em: <http://www.ismir.net/>.

JAPIASSU, Hilton. Interdisciplinaridade e patologia do saber. Rio de Janeiro: Imago, 1976

JUHÁSZ, Zoltan. Motive identification in 22 folksong corpora using dynamic time warping and self organizing maps. Proceedings of the 10th International Society for Music Information retrieval Conference (ISMIR 2009). Disponível em <www.ismir.net>.

KAGER, René. The metrical theory of stress. In: GOLDSMITH, John A. The handbook of phological theory. Cambridge: Blackwell, 1995.

KAKO, Tatsuya; OHISHI, Yasunori; KAMEOKA, Hirokazu; KASHINO, Kunio; TAKEDA, Kazuya. Automatic identification for singing style based on sung melodic contour characterized in phase plane. ISMIR 2009

KLEIN, Denise; ZATORRE, Robert; MILNER, Brenda; ZHAO, Viviane. A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. NeuroImage. Nº 13, pp. 646-653. Academic Press, 2001.

KODURI, Gopala K.; SERRÀ, Joan; SERRA, Xavier. Characterization of intonation in Carnatic music by parametrizing pitch histograms. ISMIR 2012

LADD, Robert (1996). Intonational phonology. Cambridge: Cambridge University Press.

LIST, George. The Boundaries of Speech an Song. Ethnomusicology. v. 7, nº 1, pp. 1-16. University of Illinois Press, 1963. Disponível em: <http://www.jstor.org/stable/924141> Acesso: 05/08/2008.

LOMAX, Alan; BERKOWITZ, Norman; DENG, Dorothy; KULIG, Carol; MARKEL, Norman. A stylistic analysis of speaking. Language in society. v. 6, nº 1, pp. 15-47. Cambridge University Press, 1977. Disponível em <http://www.jstor.org/stable/4166897>. Acesso: 11/03/2013.

LOMAX, Alan. Song structure and social structure. Ethnology, Vol. 1, No. 4, pp. 425-451. University of Pittsburgh, 1962. Disponível em: <http://www.jstor.org/stable/3772850>. Acesso: 11/03/2013.

LUCENTE, Luciana. Aspectos dinâmicos da fala e da entoação no português brasileiro (Tese de Doutorado). Unicamp. Campinas, 2012.

MAINGUENEAU, Dominique. Pragmática para o discurso literário. São Paulo: Martins Fontes, 1996.

MARTINS, M. V. M.; NETTO, W. F. Prosódia e escalas de frequência: um estudo em torno da escala de semitons. ReVEL, v. 8, n. 15, 2010. Disponível em <www.revel.inf.br>.

MCLUHAN, Marshall. Os meios de comunicação como extensões do homem. Trad. Décio Pignatari. São Paulo: Cultrix, 1979

MENEZES, Flo. A acústica musical em palavras e sons. São Paulo: Ateliê Editorial/Fapesp, 2003.

MORAES, João Antônio. Intonation in Brazilian Portuguese. In: HIRST, Daniel; DI CRISTO, Albert. Intonation systems: a survey of twenty languages. Cambridge: Cambridge University Press, 1998

MÜLLENSIEFEN, Daniel; FRIELER, Klaus. Optimizing measures of melodic similarity for the exploration of a large folk song database. Proceedings of the 5th International Society for Music Information retrieval Conference (ISMIR 2004). Disponível em <www.ismir.net>. Acesso: Fevereiro/2012

NESS, Steven; WRIGHT, Mathew; MARTINS, L. Gustavo; TZANETAKIS, George. Chants and Orcas: Semi-automatic Tools for Audio Annotation and Analysis in Niche Domains. 1st International Workshop on Robust Multimedia Learning in Broad Domains, 2008

NIETZSCHE, Friedrich W. Para a Genealogia da Moral. Trad. Antônio Carlos Braga. São Paulo: Escala, s.d.

PASCOAL, Hermeto. Festa dos Deuses. CD. Polygram, 1992

PASCOAL, Hermeto. O som da aura. Disponível em <www.hermetopascoal.com.br>. Acesso em 24/04/2011

PEARL, Jonathan G. S. The Music of Language: The Notebooks of Leoš Janáček. Tese de Doutorado. Santa Barbara: University of California, 2005. Disponível em <http://www.musiclanguage.net>. Acessado em julho/2010

PERETZ, Isabelle; HYDE, Krista L. What is specific to music processing? Insights from congenital amusia. Trends in Cognitive Sciences. Nº8, v. 7, pp. 362-367. Elsevier, 2003

PIKE Kenneth L. On the Grammar of Intonation. Proceedings of the Fifth International Congress of Phonetic Sciences, Munster 1964.

RIZZO, Josefa Freixa Pascua. O papel da entoação do portugues brasileiro na descricao dos atos de fala,. Campinas 1981.

ROEDERER, Juan G. Introdução à física e à psicofísica da música. Trad. Alberto Luis da Cunha. São Paulo: Edusp, 2002. Edição original de 1975.

ROSS , Deborah; CHOI, Jonathan; PURVES, Dale. Musical intervals in speech. PNAS. v. 104, p. 23, 2007.

SCHAEFFER, P. Traité des objets musicaux. Paris: Seuil, 1966

SCHÖN, Danielle; MAGNE, Cyrille; BESSON, Mireille. The music of speech: Music training facilitates pitch processing in both music and language. Psychophysiology, Nº 41, pp. 341-349. Society for Psychophysiological Research. Blackwell Publishing, 2004

SEARLE, John R. Os actos de fala: um ensaio de filosofia da linguagem. Coimbra: Almedina, 1981

SOUZA, André R. Ação e Significação: em busca de uma definição de gesto musical (dissertação de mestrado). São Paulo: UNESP, 2004.

_____. A Percepção Musical da Fala e o Problema do Alinhamento da Curva de Frequência Fundamental. Anais do III Colóquio Brasileiro de Prosódia da Fala. Belo Horizonte: UFMG, 2011.

_____. Gesto musical: ação e significação. Anais do V SIMCAM. Goiânia:UFG, 2009.

_____. Musicologia e Linguística: história e perspectivas de uma cooperação. Anais do XX Congresso da ANPPOM. Florianópolis, 2010.

STANISLAVSKI, Konstantin. A construção da Personagem. Rio de Janeiro: Civilização Brasileira, 1996.

STEELE, Joshua. Prosodia Rationalis: or, an essay towards establishing the melody and measure of speech to be expressed and perpetuated by peculiar symbols. 2ª ed. Londres: J. Nichols, 1779. Disponível em <www.jstor.org>. Acesso: Fevereiro/2012

T'HART, J.; COLLIER, R.; COHEN, A. A perceptual study of intonation: an experimental-phonetic approach to speech melody. Cambridge: Cambridge University Press, 1990.

TZANETAKIS, George; KAPUR, Ajay; SCHLOSS, Andrew K.; WRIGTH, Matthew. Hesaplamalı Etnomüzikoloji, Computational Ethnomusicology journal of interdisciplinary music studies

TZANETAKIS, George; ERMOLINSKYI, Andrei; COOK, Perry. Pitch histograms in audio and symbolic Music Information Retrieval. IRCAM, 2002

TZANETAKIS, George; KAPUR, Ajay; SCHLOSS, W. Andrew; WRIGHT, Matthew. Computational Ethnomusicology. Journal of Interdisciplinary Music Studies. v. 1, nº 2, pp. 1-24. 2007

VAISSIÈRE, Jacqueline. Perception of intonation. In: PISONI, David B.; REMEZ, Robert E. The handbook of speech perception. Malden: Blackwell, 2006

VIOLA, Izabel Cristina. O gesto vocal: a arquitetura de um ato teatral (tese de doutorado).São Paulo: PUC, 2006.

WENNERSTROM, Ann. The music of everyday speech : prosody and discourse analysis. Nova Iorque: Oxford University Press, 2001.

WILL, Udo. The magic wand of ethnomusicology: re-thinking notation and its application in music analyses. Tradução inglesa pelo próprio autor de “La baguette magique d’ethnomusicologie. Re-penser la notation et lanalyse de la musique”. Cahiers de musiques traditionelles, Nr.12, pp.9-34, 1999.

XU, Yi. In defense of lab speech. Journal of Phonetics 38: 329-336, 2010

_____. The penta model of speech : Transmitting multiple communicative functions in parallel. Haskins Laboratories, New Haven, 2004

ZATORRE, Robert; BELIN, Pascal; PENHUNE, Virginia B. Structure and function of auditory cortex: music and speech. Trends in Cognitive Sciences. Nº1, v. 6, pp. 37-46. Elsevier, 2002