

“JÚLIO DE MESQUITA FILHO”

Faculdade de Ciências e Letras
Campus de Araraquara - SP

ANDRÉ LUIZ MACHADO

**DIFERENÇAS NA PERCEPÇÃO DE EMOÇÕES
EM ENUNCIADOS VERBAIS DECLAMADOS E
CANTADOS**



ARARAQUARA – S.P.
2017

ANDRÉ LUIZ MACHADO

DIFERENÇAS NA PERCEPÇÃO DE EMOÇÕES EM ENUNCIADOS VERBAIS DECLAMADOS E CANTADOS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – UNESP/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática

Orientador: Gladis Massini-Cagliari

Bolsa: Capes

ARARAQUARA – S.P.

2017

Machado, André Luiz
DIFERENÇAS NA PERCEPÇÃO DE EMOÇÕES EM ENUNCIADOS
VERBAIS DECLAMADOS E CANTADOS / André Luiz Machado –
2017
109 f.

Dissertação (Mestrado em Linguística e Língua
Portuguesa) – Universidade Estadual Paulista "Júlio
de Mesquita Filho", Faculdade de Ciências e Letras
(Campus Araraquara)
Orientador: Gladis Massini-Cagliari

1. Linguagem verbal. 2. Música. 3. Fala. 4. Emoção.
5. Percepção. I. Título.

Ficha catalográfica elaborada pelo sistema automatizado
com os dados fornecidos pelo(a) autor(a).

ANDRÉ LUIZ MACHADO

DIFERENÇAS NA PERCEPÇÃO DE EMOÇÕES EM ENUNCIADOS VERBAIS DECLAMADOS E CANTADOS

Dissertação de Mestrado, apresentada ao Programa de Pós em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – UNESP/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática

Orientadora: Gladis Massini-Cagliari

Bolsa: Capes

Data da defesa: 05/05/2017

MEMBROS COMPONENTES DA BANCA EXAMINADORA:

Presidente e Orientadora: Prof^a. Dr^a. Gladis Massini-Cagliari

Universidade Estadual Paulista “Júlio de Mesquita Filho” – Faculdade de Ciências e Letras, Câmpus de Araraquara – SP

Membro Titular: Prof. Dr. Waldemar Ferreira Netto

Universidade de São Paulo - Faculdade de Filosofia Letras e Ciências Humanas

Membro Titular: Prof^a Dr^a Anise de Abreu Gonçalves d’Orange Ferreira

Universidade Estadual Paulista “Júlio de Mesquita Filho” – Faculdade de Ciências e Letras, Câmpus de Araraquara – SP

Local: Universidade Estadual Paulista
Faculdade de Ciências e Letras
UNESP – Campus de Araraquara

[para'ɥilɨ̃ ɪ'kaɪ̃ɸi'fazẽ̃meʊz'dʒiasmaisfe'lizɪs]

AGRADECIMENTOS

Por mais que eu me esforce, certamente esta lista de agradecimentos estará incompleta: todas as pessoas que conheci ao longo desses 27 anos contribuíram, em algum grau, para eu me tornar quem eu sou hoje e, indiretamente, para este trabalho que ora finalizo. Certo de que lembrar todos que merecem e ser igualmente justo é uma batalha perdida, prossigo.

Agradeço aos meus pais, Adalberto e Marli, por sempre terem dado o melhor de si. Nunca foi fácil, *nós* não somos fáceis, mas nos esforçamos ao máximo e seguimos em frente. Amo vocês.

Minha mais sincera gratidão e admiração pela minha orientadora, Gladis, pela confiança em um projeto (e um orientando) tão excêntrico e por ter sido compreensiva e prestativa com meus eventuais deslizes. Sua competência e disposição para o trabalho são uma inspiração, e se um dia eu chegar a ser professor doutor, quero ser como você. Agradeço a Anise d'Orange Ferreira, que desde o começo foi incrivelmente generosa e prestativa, e a Cristina Martins Fargetti, cujo entusiasmo, carinho e leitura atenciosa me deixaram lisonjeado. Muito obrigado!

João Jorge e José Francisco pelos vocais, Fábio e Edward Estlin Cummings pelas palavras. A todos aqueles que responderam os questionários, boas vibrações em agradecimento. Obrigadíssimo, Rômulo Bartolozzi, cuja ajuda inestimável permitiu que este trabalho fosse concluído, e Gabi Orselli Fernandes, que me ensinou a usar o Excel e para quem eu ainda devo uma cerveja. Meus colegas de mestrado de 2015: Adriel e seu *deboísmo* inabalável, e em especial, Carlos, pelas longas conversas durante as madrugadas, pelas discussões sobre a academia, ciência, astrologia, o universo e tudo mais – obrigado!

A Emi-chan, Katia, Michele e Thaynary, pessoas queridas que estiveram comigo ao longo do último ano e tornaram um dos períodos mais difíceis da minha vida mais leve, meu amor e gratidão. Obrigado Renan e Maiko, que tornam Itápolis uma cidade melhor quando estão por perto.

Às duas pessoas a quem esse trabalho é dedicado, que mudaram a minha vida:

Karin, que vai ser sempre o meu bebê. Que me surpreende a cada dia, que me derrete todo sem o menor esforço, que desperta em mim o desejo de proteger de toda a maldade e de mostrar todos os encantos do mundo. Te amo.

Willian. Nada do que eu pudesse escrever aqui faria jus ao que eu sinto. Obrigado por fazer com que eu me reencontrasse, por estar ao meu lado nos meus piores momentos e por me devolver o paraíso que eu achei que tinha perdido para sempre. Obrigado por estar na minha vida. Te amo.

And in the end the love you take is equal to the love you make.¹

¹ Lennon e McCartney, 1969

So I say thank you for the music
The songs I'm singing
Thanks for all the joy they're bringing
Who can live without it?
I ask in all honesty – what would life be?
Without a song or a dance what are we?
So I say thank you for the music
For giving it to me

(ABBA)

RESUMO

A linguagem verbal é um fenômeno exclusivo da espécie humana, encontrado em todas as culturas do mundo. A música também está presente em todas as culturas conhecidas, e compartilha com a língua oral não apenas sua origem evolutiva, mas também sua materialidade elementar: o som. Este trabalho se propõe a examinar possíveis concordâncias, sobreposições e discrepâncias entre música e língua na expressão de emoções. Mais especificamente, buscou-se verificar, através de um teste de percepção, se um mesmo enunciado verbal é interpretado de forma distinta quando declamado e quando musicado. Para tanto, foram realizadas gravações de dois textos sendo declamados e cantados. Através de uma plataforma *online*, foi pedido aos informantes que ouvissem e avaliassem as gravações por meio de sete escalas bipolares de adjetivos (por exemplo, Alegre – Triste). Comparando-se os escores obtidos pelas diferentes gravações de um mesmo enunciado verbal, foi possível verificar se os informantes julgaram que o texto declamado expressava uma carga emocional igual ou diferente da expressada pelas versões em que o mesmo texto tinha sido musicado. Para garantir a validade da discussão e das conclusões, testes estatísticos foram utilizados para assegurar que os resultados encontrados não poderiam ser atribuídos a puro acaso. Como esperado, verificou-se que há uma diferença significativa na forma como um mesmo enunciado verbal é percebido quando cantado e quando declamado, no que se refere à percepção de emoções. As diferenças foram mais sensíveis ao julgar se o enunciador parecia estar alegre ou triste, excitado ou deprimido, e se parecia estar sério ou brincando. Essa diferença de percepção pode ser explicada pela diferença nos parâmetros acústicos e prosódicos entre as gravações, acarretadas principalmente pela transposição da fala para o canto.

Palavras – chave: Fala. Música. Emoção. Percepção. Linguagem verbal.

ABSTRACT

Verbal language is a phenomenon exclusive to humankind, found in all cultures around the world. Music is also present in all known cultures, and shares with oral language not only its evolutionary origin, but also its elementary materiality: sound. This work aims to examine possible concordances, overlaps and discrepancies between music and language in the expression of emotions. More specifically, we tried to verify, by means of a perception test, if the same verbal utterance is interpreted differently when recited and when turned into music. In order to do that, two texts were recorded being recited and sung. Through an online platform, informants were asked to listen and evaluate the recordings using seven bipolar adjective scales (eg Joyful - Sad). Comparing the scores obtained by the different recordings of the same verbal utterance, it was possible to verify if the informants judged that the recited text expressed an emotional tone equal to or different from that expressed by the versions in which the same text had been sung. To ensure the validity of the discussion and conclusions, statistical tests were used to ensure that the results found could not be attributed to pure chance. As expected, there was a significant difference in the way a verbal utterance is perceived when sung and when recited, regarding the perception of emotions. The differences were more distinguishable in judging whether the enunciator seemed to be joyful or sad, excited or depressed, and whether they seemed to be serious or playful. This difference in perception can be explained by the difference in the acoustic and prosodic parameters between the recordings, caused mainly by the transposition from speech to song.

Keywords: Speech. Music. Emotion. Reception. Verbal communication.

LISTA DE FIGURAS

Figura 1 Ilusão de óptica	39
Figura 2: Janela do programa Melodyne Editor.....	50
Figura 3: “Meu coração dança / nos intervalos da razão”	52
Figura 4: “E cai, pois não aprende / a coreografia”.....	53
Figura 5: “Meu coração dança... /Meu coração é dançarino”	53
Figura 6: “Desengonçado e divino / De vinho, fossa e de lua”	54
Figura 7: “Dançando em becos pelas ruas / Mas cai pois não aprende”	54
Figura 8: “E cai, pois não aprende / Não, ele nunca aprende”	55
Figura 9: “A coreografia...”	55
Figura 10: “Dançando à noite pelos bares / Dançando em todos os lugares”	56
Figura 11: “Meu coração transforma o tropeço / Num sobressalto /Num recomeço”	56
Figura 12: “Mesmo num samba / De despedida / Meu coração dança a vida”	57
Figura 13: “Vindo das cinzas”	58
Figura 14: “Um perfume me traz aqui”	59
Figura 15: “Corpos, ossos, beijo o escuro da dor”	59
Figura 16: “Põe-se o sol”	60
Figura 17: “Dorme o mar”	60
Figura 18: “A lua a ruir”	61
Figura 19: “Se desfaz”.....	61
Figura 20: “O amor se acende”	62
Figura 21: “Contra os olhos, rosto molhado”	62
Figura 22: “Me embalada no seu seio”	63
Figura 23: “Mergulhe os dedos do seu corpo em mim”	63
Figura 24: “Alta a maré em seus alcances sombrios”	64
Figura 25: “O aroma das noites vem para brilhar”.....	64
Figura 26: “Pétalas dançando no meu olhar”	65
Figura 27: “As mãos do oceano vêm me entregar”.....	65
Figura 28: “Seus braços me acariciar”	66
Figura 29: “Suave maestria apressa o meu pulso”	66
Figura 30: “Firme suavidade da carne”	67
Figura 31: “Sempre meu guardado mistério”	67
Figura 32: “Junto ao sopro das estrelas”	68
Figura 33: "Suas asas vão me levar"	68
Figura 34: Um dos questionários utilizados no experimento.....	73
Figura 35: Um das questões referentes a aspectos demográficos dos informantes	75
Figura 36: Aplicação da prova χ^2 no programa SOFA.....	91
Figura 37: Exemplo de gráfico gerado na aplicação do Coeficiente de Correlação por Postos de Spearman, a partir da amostra referente à gravação do texto <i>Arritmia</i> sendo declamado... 93	

LISTA DE TABELAS

Tabela 1: <i>Arritmia</i> : Média aritmética dos valores das respostas fornecidas pelos informantes	86
Tabela 2: <i>Arritmia</i> : Probabilidade de ocorrência dos dados obtidos sob H0, em cada questão, de acordo com a prova Kruskal-Wallis (gl=2), Mann-Whitney unilateral (MWu) e Mann-Whitney bilateral (MWb)	87
Tabela 3: <i>Cummings</i> : Probabilidade de ocorrência dos dados obtidos sob H0, em cada questão, de acordo com a prova Kruskal-Wallis (gl=2), Mann-Whitney unilateral (MWu) e Mann-Whitney bilateral (MWb).....	88
Tabela 4: Ocorrência de diferenças significativas nos dados, comparação entre <i>Arritmia</i> e <i>Cummings</i>	89
Tabela 5: Correlações e fatores demográficos.....	94

SUMÁRIO

AGRADECIMENTOS	5
EPÍGRAFE	7
RESUMO	8
ABSTRACT	9
LISTA DE FIGURAS	10
LISTA DE TABELAS.....	11
SUMÁRIO.....	12
INTRODUÇÃO	13
1. DEFINIÇÕES E SOBREPOSIÇÕES	15
1.1. Música, Língua e Linguagem	15
1.2. Por que unir música e língua?	18
1.3. Sobre a relação entre a linguagem musical e a linguagem verbal em composições musicais e seus sentidos.....	25
1.4. Considerações finais.....	28
2. FUNDAMENTAÇÃO TEÓRICA.....	29
2.1 Linguística	29
2.1.1 Prosódia.....	29
2.2 Música.....	32
2.2.1 Melodia e ritmo.....	35
2.2.2 Intervalos.....	37
2.2.3 Harmonia	38
2.2.4 Escalas, modos e tonalidades	38
2.3 Emoções	41
2.4 Considerações finais.....	43
3. METODOLOGIA	44
3.1 Preparação e gravações	44
3.1.2 Partitura das composições e espectrogramas das gravações	50
3.2 Elaboração e aplicação do experimento	69
3.3 Metodologia de análise de dados.....	77
3.4. Considerações finais.....	84
4. RESULTADOS E ANÁLISES	85
5. DISCUSSÃO	96
CONCLUSÕES	99
REFERÊNCIAS BIBLIOGRÁFICAS.....	101

INTRODUÇÃO

Esta dissertação de Mestrado trata da relação entre música e linguagem verbal. Trata, mais especificamente, da percepção da carga emocional de um texto quando enunciado por meio da linguagem puramente verbal e quando apresentado como um dos elementos de uma composição musical.

Com esse objetivo, foi realizado um teste de percepção. Através de uma plataforma online, voluntários avaliaram gravações de um mesmo enunciado verbal em duas situações diferentes: sendo declamado e sendo cantado, na forma de uma composição musical com intervalos de terça maior ou menor. Uma vez que, do ponto de vista lexical e sintático, os enunciados eram idênticos, eventuais discrepâncias nas avaliações de um texto declamado e musicado poderiam provavelmente ser atribuídas à própria musicalização do texto.

Nenhuma pesquisa pode ser concebida e executada sem uma perspectiva teórica prévia. Uma vez que este trabalho se ocupa de linguagens em suas manifestações sonoras, faz-se necessário recorrer ao arcabouço teórico da Fonética e da Fonologia.

Ao longo do período de desenvolvimento desta pesquisa, diversas obras serviram de aporte para a elaboração do experimento e para a avaliação dos resultados obtidos. Basicamente, parte-se das ideias de Chomsky (1971, 1994) que, grosso modo, encara a linguagem como uma competência enraizada em nossa biologia e que seria compartilhada por todos os seres humanos. A ideia da linguagem como um fenômeno de origem biológica e não apenas cultural abre caminho para elucubrações acerca da origem e da evolução da língua, sendo que Falk (2009) e Levman (1992) podem ser citados como expoentes dessa vertente. Por fim, tem-se que diversos trabalhos têm aproximado linguagem verbal e música, seja como *corpus* de análise, como fazem Tatit (1997) e Dietrich (2008), seja utilizando elementos da música ou peças musicais como fatores em teorizações sobre a linguagem e a cognição, como Allen (1967) e Jackendoff e Lerdahl (1980) e Lerdahl e Jackendoff (1983). Mais pertinentes à investigação do presente trabalho, no entanto, são trabalhos como os de Cook, Fujisawa e Takami (2004) e Curtis e Bharucha (2010), que apontam similaridades na forma como linguagem verbal e música utilizam elementos acústicos expressivamente, e os de Coutinho e Dibben (2013) e Steinbeins e Koelsch (2008), apontando para o uso de recursos cognitivos em comum no processamento da fala e da música.

Não obstante o supracitado experimento, que constitui a espinha dorsal da pesquisa de mestrado, este trabalho tem também o objetivo de aduzir evidências da prolificidade das

aproximações entre linguística e música, além de levantar hipóteses às quais, ao menos por enquanto, não propomos soluções conclusivas, não permitidas pelo escopo deste trabalho. Não obstante, acreditamos que essas hipóteses podem levar a desdobramentos teóricos, metodológicos e pragmáticos bastante interessantes no futuro.

A primeira seção desta dissertação apresenta definições daquilo a que, ao longo do texto, nos referiremos como “música”, “língua” e “linguagem”. Os limites entre alguns desses conceitos podem ser tênues e suas relações complexas, mas argumentaremos no sentido de que, assim como a língua, a música também é uma linguagem. A segunda seção apresenta a fundamentação teórica para este trabalho, explicitando, de forma sucinta, os conceitos-chave que sustentam a metodologia do experimento, a análise e a discussão dos resultados obtidos: prosódia - frequentemente considerada a “música” da língua -, ritmo, melodia e princípios básicos de harmonia, e o próprio conceito de emoções. A seção de metodologia descreve o processo de preparação e execução do experimento, a composição e a gravação das canções, a elaboração e aplicação do questionário *online* que configurou o teste de percepção, bem como a análise estatística dos dados – mais especificamente, a escolha de provas estatísticas para verificação da pertinência dos resultados. Tais dados serão apresentados na seção seguinte, na qual também veremos quais conclusões é possível tirar dos resultados obtidos. Basicamente, conclui-se que a transposição do texto da fala para o canto de fato acarreta diferenças na carga emocional percebida pelos ouvintes. A seção de discussão tem o intuito de esboçar explicações para as conclusões tiradas, e nela aponta-se uma semelhança no funcionamento das “gramáticas” da língua e da música tonal, mais especificamente no que se refere à marcação de tonicidade na fala e a atribuição da tonalidade a uma composição na música. Finalmente, a seção de conclusões fecha esta dissertação com um resumo dos resultados alcançados, suas implicações e a apresentação de encaminhamentos para possíveis trabalhos futuros.

1. DEFINIÇÕES E SOBREPOSIÇÕES

1.1. MÚSICA, LÍNGUA E LINGUAGEM

Da mesma forma que não compartilhamos apenas informações lexicais e gramaticais, mas também pensamentos, compartilhamos também a expressividade que as pessoas depositam na musicalidade que nos rodeia. (BENFATTI, 2010, p. 136)

O *Dicionário Aurélio de Língua Portuguesa* (FERREIRA, 1995) define a palavra “língua”, entre outras acepções, como “sistema de comunicação comum a uma comunidade linguística”. Tal definição remete, indiretamente, à palavra “linguagem”, definida no mesmo dicionário como “expressão do pensamento pela palavra, pela escrita ou por meio de sinais”. Tais definições mostram que o termo “linguagem” está associado, primeiramente, à língua. Mas poderíamos tomar essa definição como definitiva? Seriam as linguagens expressões *apenas* dos nossos pensamentos?

Ao produzir um enunciado verbal, os falantes revelam mais do que as palavras dizem à primeira vista. Massini-Cagliari e Cagliari (2001) mostram que mesmo um enunciado simples como “Ela foi ao cinema ontem” pode ser interpretado de várias maneiras. Dependendo da posição do acento frasal, podemos julgar que a informação mais importante se refere ao sujeito (“ELA foi ao cinema ontem”), à ação (“Ela FOI ao cinema ontem”), ao destino (“Ela foi ao CINEMA ontem”) ou mesmo ao tempo (“Ela foi ao cinema ONTEM”). A entoação, o timbre de voz e a velocidade de fala podem nos revelar o estado emocional do falante em relação à informação que ele expressa. Essas variáveis estão imbricadas ao conteúdo semântico e sintático das palavras que, ao mesmo tempo em que influencia, é influenciado por esses conteúdos. A língua, enquanto sistema de comunicação, expressa mais do que apenas pensamentos. Para além do significado e do significante, há muitos outros elementos imbuídos no ato de falar. A língua, então, não expressa apenas nossos pensamentos.

Outras manifestações humanas, como a dança e a pintura, evidentemente não são línguas. Ainda assim, são frequentemente chamadas de linguagens. Ora, linguagem implica expressar *algo*. Que algo essas outras manifestações expressam? O que faz delas *linguagens*? Tomemos a música como exemplo.

Tradicionalmente, concebeu-se o signo com a relação entre o plano do significante, que poderíamos tomar como a materialidade sonora ou gráfica de uma palavra, e um significado, o conceito mental para o qual aquela materialidade aponta - as noções semióticas de plano da expressão e plano de conteúdo são derivadas dessa concepção de signo. Independentemente da vagueza semântica de um enunciado musical, mesmo um leigo provavelmente concordaria que a música – seja ela instrumental ou cantada - é *expressiva*. Ou seja, os sons de uma composição musical evocam sentimentos, ideias, memórias. Portanto, pode-se dizer com razoável confiança que a música possui um significante, uma vez que tem materialidade (no caso, sonora), e também pode evocar significados. Disso, levantam-se questões complexas. Se um signo só pode existir como simbiose entre significante e significado, a capacidade da música de evocar, ainda que de forma vaga, conceitos e sentimentos, constitui um fenômeno intrigante, que pode e deve ser explorado pela Linguística, uma vez que toca reflexões profundas acerca da natureza da comunicação humana.

Uma conciliação de conceitos pode ser alcançada a partir das palavras do cientista cognitivo Philip Lieberman:

Linguists have been somewhat anthropocentric in defining language to be necessarily human language. We will instead propose an operational definition of language. *A language is a communications system that is capable of transmitting new information.* In other words, we're operationally defining a language as a communication system that places no inherent restriction on the nature or quality of the information transmitted. It is obvious that this definition does not require that all languages have all of the properties of human language.² (LIEBERMAN, 1975, p. 6)

Para o que a língua inglesa usa o termo “*language*”, o português tem dois termos correntes: “língua” e “linguagem”. Tendo em vista a discussão acima, adotemos aqui a definição de Lieberman como a definição de “linguagem”, incluindo aí todas as linguagens não-verbais, e utilizemos “língua” para se referir a uma linguagem específica, a linguagem verbal humana.

² Ao longo deste trabalho, apresentaremos a tradução dos trechos citados em língua estrangeira como notas de rodapé. Todas as traduções são de nossa autoria e são aqui apresentadas puramente no intuito de facilitar a leitura da presente dissertação. Assim, não devem ser tomadas como correções ou substituições a possíveis traduções já existentes. “Linguistas têm sido um tanto antropocêntricos ao definir linguagem como necessariamente a linguagem humana. Proporemos, ao invés disso, uma definição operacional de linguagem. Uma linguagem é um sistema de comunicação capaz de transmitir informações novas. Em outras palavras, estamos operacionalmente definindo linguagem como um sistema de comunicação que não impõe restrição inerente à natureza ou à qualidade da informação transmitida. É óbvio que essa definição não exige que todas as linguagens tenham todas as propriedades da linguagem humana.”

De acordo com o exposto até o momento, parece ponto pacífico dizer que, apesar de suas diferenças com a língua (linguagem verbal), a música também é um tipo de linguagem, pois é expressiva, como já vimos acima (ver FERREIRA, 1993, p.4, que chega à mesma conclusão). No entanto, postular o status de linguagem da música não soluciona outro problema: a própria definição do que é – e, por exclusão, o que não é – música.

A complexidade dessa questão é sucintamente explorada por Bryan G. Levman em seu texto *The Genesis of Music and Language*:

Much of the question of whether music or speech came first revolves around the definition of the word "music." (...) The issue of what constitutes an adequate definition of music is highly contentious, and many would dispute any proposal as too limiting. Jean-Jacques Nattiez, for example, claims that "what is musical in reality, is every phenomenon which a cultural group agrees to regard as such" (1971:97), and certainly the history of music in the twentieth century, where random noise, environmental cacophony, and the sounds of nature have entered the musical repertoire, bears this position out. Ethologists would argue that the "songs" of the humpback whales are just as much music as the "songs" of humans and some ethnomusicologists might agree. (...) As early as 1941 George Herzog proposed "patterning of sound" for an (unsatisfactory, in his view) definition of music, and pointed out that "there seems to be no criterion for any theoretical separation of the vocal expression of animals from human music" (1941:4).³ (LEVMAN, 1992, p. 149-150)

De fato, parece quase impossível criar uma definição que abarque elementos tão díspares quanto a cantiga *Se Essa Rua Fosse Minha* e o melodrama *Pierrot Lunaire* de Arnold Schönberg. Mencione-se, então, uma peça como *4'33"*, de John Cage⁴, e uma definição parecerá uma utopia sem possibilidade de concretização. Uma discussão bastante didática acerca da natureza da música e de sua possível definição se encontra no livro do compositor e educador musical canadense Murray Schafer, *O Ouvindo Pensante*. Depois de relatar uma série de experiências em sala de aula em que questiona e incita os alunos acerca de uma possível definição do que é e do que não é música, Schafer chega, junto com seus alunos, a uma

³ Grande parte da questão de saber se música ou fala veio em primeiro lugar gira em torno da definição da palavra "música". (...) O problema do que constitui uma definição adequada de música é altamente contenciosa, e muitos contestariam qualquer proposta como demasiado limitativa. Jean-Jacques Nattiez, por exemplo, afirma que "o que é musical na realidade, é todo fenômeno que um grupo cultural aceita considerar como tal" (1971: 97) e, certamente, a história da música no século XX, onde o ruído aleatório, a cacofonia ambiental e os sons da natureza entraram no repertório musical, sustenta essa posição. Os etólogos argumentariam que as "canções" das baleias jubarte são tão música quanto as "canções" dos humanos e alguns etnomusicologistas podem concordar. (...) Já em 1941, George Herzog propôs a "combinação do som" como uma definição (insatisfatória, na sua opinião) da música, e assinalou que "não parece haver critério para qualquer separação teórica da expressão vocal dos animais música humana" (1941: 4). (LEVMAN, 1992, páginas 149-150)

⁴ Também conhecida como "Quatro minutos e meio de silêncio", *4'33"* é uma peça de 1952 que pode ser executada com qualquer instrumento ou combinação de instrumentos, mas a partitura instrui o(s) executante(s) a não tocar nenhum instrumento durante os três movimentos da peça, de modo que os ouvintes ouçam apenas os sons ambientes enquanto a peça é "executada".

solução aparentemente razoável: “MÚSICA É UMA ORGANIZAÇÃO DE SONS (RITMO, MELODIA ETC.) COM A INTENÇÃO DE SER OUVIDA.”⁵ (SCHAFER, 1991, p. 35)

Tal definição é ao mesmo tempo suficientemente abrangente para abarcar as composições mencionadas anteriormente, mas restritiva para diferenciar a música produzida por seres humanos das vocalizações de animais e de sons incidentais aleatórios. É esta definição de música, portanto, que será utilizada neste trabalho. Obviamente, considerando o contexto de produção desta dissertação, as discussões aqui levantadas se referem, mais especificamente (porém não exclusivamente), à música ocidental tonal, por ser este o tipo de música à qual temos, de modo geral, mais exposição. Isso significa que as ideias aqui expostas só podem ser expandidas e aplicadas a outras tradições musicais (como a indígena, a japonesa e a indiana) com bastante cautela, uma vez que tradições distintas costumam divergir, por exemplo, na forma como dividem o espectro sonoro em escalas e na forma como trabalham a polifonia. Não obstante essas ressalvas, é necessário lembrar que todas as tradições musicais consistem, em última instância, a uma forma de organizar os sons ao longo do tempo e que, dado as características físicas do elemento sonoro, aspectos como o ritmo e intensidade se aplicam a todas as tradições, e há evidências suficientes para concluir que certos aspectos musicais são – pelo menos até certo grau – universais, como será explicitado mais adiante.

1.2. POR QUE UNIR MÚSICA E LÍNGUA?

O canto é um modo de expressão humana de origem já longínqua no tempo. Quando o homem começou a cantar? Não se sabe ao certo. A resposta é mais ou menos igual àquela que serve para responder a outras questões parecidas: Quando o homem começou a falar? Quando o homem começou a escrever? E a tocar instrumentos? E por aí poderíamos desfiar um rol de perguntas sobre as habilidades humanas.

Ao cumprir a tarefa de definir o objeto desta tese, as perguntas, algumas especulativas, não cessam de aparecer e adiam a definição. E o canto com texto, desde quando existiria? E se o homem cantou antes de falar? Ou se a fala nas suas variações de frequência fundamental foi repetindo melodias e germinou o canto, ou se o canto já existia antes das palavras e abrigou-as servindo de substância para que elas pudessem se materializar? (MEDEIROS, 2002, p. 20)

É difícil dizer com precisão o que faz com que língua e música frequentemente apareçam relacionadas. Como já discutido na seção 2.1, a música, bem como as diversas

⁵ Apesar de visualmente destoante, esse trecho, no livro de Schafer referenciado, encontra-se grafado em letras maiúsculas. Mantivemos, assim, a forma original.

línguas do mundo, são linguagens. Porém, também as artes visuais (incluídas aí escultura, pintura, desenho, etc.), os gestos, códigos de programação de software, entre muitos outros, podem ser citados como exemplos de linguagem. Em que pesem as obras de artistas como Wassily Kandinsky⁶, música e língua parecem ter uma relação muito mais próxima do que, por exemplo, música e linguagem de programação ou língua e dança. Podem servir como evidências da afirmação anterior:

- Afirmações como “Não gosto de música em inglês porque não entendo o que a música diz” e “Se eu não entendo a letra, a música não faz sentido”. Partindo do senso comum, não é raro encontrar quem tome por “sentido da letra da música” pelo mais amplo “sentido da música”;
- A própria nomenclatura musical: composições são divididas em frases, por exemplo;
- A análise da prosódia das línguas é apoiada em termos que são originários da música (ver Massini-Cagliari e Cagliari, 2001, e Massini-Cagliari, 1999);
- O fato de que, em algumas línguas, como a Venda, há sobreposição semântica entre as palavras e os conceitos de “música” e “poesia” (BLACKING, 1982, p. 18), e a “poesia” deriva da linguagem verbal;
- A poesia, por sua vez, frequentemente faz uso de métrica, sendo que a regularidade métrica também é um aspecto quase onipresente na música, pelo menos em se tratando da música popular;
- Talvez a maior parte da música à qual estamos, em geral, expostos, seja música com letra, o que reforça a relação entre as duas linguagens;
- As letras de música frequentemente utilizam de recursos associados à linguagem poética, como a rima, o que faz muitas pessoas considerarem que letras de música são poesia (o fato de em 2016 a Academia Sueca ter concedido o prêmio Nobel de Literatura ao cantor e compositor norte-americano Bob Dylan reforça essa visão).

No caso dos estudos linguísticos, as aproximações entre as linguagens verbal e musical têm se mostrado, ao longo dos anos, uma prolífica fonte de reflexões e descobertas. Como exemplos, podemos citar estudos sobre o uso da música como ferramenta didática no ensino de línguas, como em ROCHA (2012) e SANTOS e PAULUK (2008). Também

⁶ Em russo, Василий Кандинский, artista plástico nascido na segunda metade do século XIX, tinha na música fonte de inspiração e referência para a pintura. Suas obras abstratas frequentemente faziam referência à música, seja em seus títulos, seja na intenção. Em suas telas, Kandinsky ocasionalmente tentava transfigurar a linguagem musical em cores e formas.

merecem menção os trabalhos de TATIT (1997) e Tartit e Lopes (2008) que, com sua Semiótica da Canção, adaptam o *modus operandi* da semiótica para a aplicação na análise de composições musicais.

No entanto, talvez em nenhuma outra área da linguística a aproximação entre essas duas linguagens tenha sido tão prolífica quanto na fonologia. Partituras musicais nas quais estejam registradas as letras das canções são um excelente material para o estudo de fases passadas do sistema fonológico de uma língua. Neste sentido, merecem destaque os trabalhos de linguística histórica de Massini-Cagliari e Cagliari (1998) e Massini-Cagliari (2008 a, b, 2009 a, b, 2010 a, b, 2011) e COSTA (2010), nos quais as proeminências musicais são usadas como base para o estudo do ritmo do galego-português. A própria análise fonológica lança mão de conceitos relacionados à música. Por exemplo, na análise do ritmo das línguas, parte essencial do trabalho é justamente definir o tipo de pé rítmico que a língua estudada apresenta:

Ao lado da permanência de diversos rótulos (por exemplo, “iambo”, “troqueu” etc.), que, embora com o mesmo nome, redefinidos, designam unidades muito diferentes dos “metros” gregos e latinos, deve ser destacada a recorrência da concepção de PÉ como *estruturador principal do ritmo*. Assim era concebido o PÉ desde a antiguidade clássica (uma vez que a prosa também era considerada “metrificada”) assim era concebido dentro das teorias fonéticas apresentadas e assim o é até hoje, dentro dos diversos movimentos da fonologia. (MASSINI-CAGLIARI, 1999, p. 131)

À primeira vista, o termo “pé” pode parecer muito distante da música. Afinal, podemos nos lembrar do ritmo, da melodia, da tonalidade de uma composição, mas não falamos em pés. Apesar de não ser tão evidente, a relação – uma vez explicitada – torna-se clara:

Quanto à origem do rótulo PÉ, Allen (1973, p. 122) afirma, citando um dos trabalhos sobre métrica grega de Dale, que este termo parece ter sido tomado dos movimentos do pé humano, isto é, do movimento progressivo e alternante de levantamento e abaixamento. Por sua vez, Ravizza, além de ligar a origem do rótulo PÉ ao movimento dos pés humanos, estabelece também uma conexão com a música, relacionando, já desde as origens, as noções de “PÉ métrico” e “compasso musical”: “O verso latino [...] não abrange uma série regular de sílabas acentuadas, mas encerra uma série de compassos, que em linguagem métrica se chamam *pés*, talvez assim chamados porque antigamente na dança marcava-se com os pés a divisão dos compassos” (Ravizza, 1940, p. 415, rodapé). (MASSINI-CAGLIARI, 1999, p. 114)

A aproximação entre música e linguística também pode servir como um meio para alcançar um melhor entendimento da cognição humana. Jackendoff e Lerdahl (1983)

desenvolveram uma teoria gerativa da música tonal, baseados na gramática gerativa de Noam Chomsky. Curiosamente, o paralelo traçado pelos dois autores entre música e linguística parece ser mais metodológico do que explicativo: ao utilizar o *modus operandi* da fonologia métrica para construir uma “gramática” da música tonal, Lerdahl e Jackendoff advogam não necessariamente a favor de uma aproximação entre música e língua, mas de uma teoria cognitiva mais ampla que englobaria música, língua e outros processos relacionados à percepção temporal:

So we return to the question [...]: why should there be such parallelism? Given that both theories are attempts to account for human cognitive abilities, the existence of parallelism between them implies a claim that these areas are a respect in which human musical and linguistic capacities overlap. [...] However, if this claim is true, it would be surprising if music and language were the ONLY human abilities so structured. Rather, we should be led to look for something closely analogous to time-span structure in many human abilities under the rubric of "temporal patterning", from event perception to motor control to the planning of extended strategies of behavior. In particular, we should expect the notion of HEAD / ELABORATION to figure prominently in psychological theories of temporal organization.⁷ (JACKENDOFF, LERDAHL, 1980, p. 41)

Considerando o trecho de Jackendoff e Lerdahl (1980) citado, o paralelismo entre música e língua parece ocorrer não devido à natureza específica dessas duas linguagens, mas pela estrutura cognitiva do ser humano. Ainda que a explicação dos dois autores para tal paralelismo não deva ser descartada, há evidências suficientes para concluir que o paralelismo entre língua e música não se deve apenas à forma como o cérebro processa padrões temporais. De fato, sabe-se que música e língua estão ligadas também do ponto de vista *evolutivo*. A natureza exata dessa ligação, no entanto, ainda é alvo de debates:

Authors have adopted three primary positions: (1) that language and music developed along separate paths and are in effect two completely different faculties, (2) that music developed out of language, or at least was chronologically later than language, and (3) that language developed out of

⁷ “Então voltamos à questão [...]: por que há tal paralelismo? Uma vez que ambas as teorias são tentativas de conceber as habilidades cognitivas humanas, a existência de paralelismo entre elas implica uma alegação de que essas áreas são um aspecto em que as capacidades musicais e linguísticas humanas se sobrepõem. [...] No entanto, se essa afirmação é verdadeira, seria surpreendente se música e língua fossem as ÚNICAS habilidades humanas assim estruturadas. Ao contrário, devemos ser levados a procurar algo análogo à estrutura de tempo-alcance em muitas capacidades humanas sob a rubrica de "padrões temporais", desde a percepção de eventos ao controle motor para o planejamento de estratégias estendidas de comportamento. Em particular, devemos esperar que a noção de CABEÇA / ELABORAÇÃO figure em destaque em teorias psicológicas de organização temporal.”

music, or both developed from a common "proto-faculty."⁸ (LEVMAN, 1992, p. 147)

Discutir a origem da língua e da música e, por consequência, qual entre as duas linguagens é anterior, foge do escopo do trabalho aqui proposto, ainda que vários estudiosos se debrucem sobre esse assunto. Um bom exemplo é o húngaro Iván Fónagy, mencionado por Levman, em "The Genesis of Music and Language":

Fonagy's convincing argument that human proto-language was iconic, that is, that the prosodic and physical components were directly expressive of inner mental states, accounts for the popularity of the view that music primarily expresses emotion, while speech expresses rational thought. In Fonagy's view, humankind's first language - the common source out of which music and speech were later to develop - was clearly "musical" and the musical or prosodic elements carried the meaning directly. What we understand as speech was a later development of double articulation which evolved to express more complicated concepts which pitch variation alone was incapable of expressing. Double articulation, or duality of patterning, refers to the sonic division of a language into individual sound particles-phonemes-which have no intrinsic meaning in themselves but can be combined into new units of meaning known as morphemes. The development of double articulation marks the exact point where language ceases to be isomorphic and becomes symbolic. It is also of course the point where music and speech begin to separate and go their different ways. Yet even after their separation, music and speech continue to be intimately connected through their common frequency component.⁹ (LEVMAN, 1992, p. 153)

Por hora, é suficiente reconhecer que essas duas formas de expressão humana estão intimamente relacionadas biologicamente. A gênese da língua e da música é uma área prolífica para especulações e teorizações, e dada a estimativas qualitativas e quantitativas acerca da natureza dessas duas linguagens.

⁸ Autores adotaram três posições principais: (1) que língua e música se desenvolveram ao longo de caminhos separados e são, em efeito, duas faculdades completamente diferentes, (2) que a música se desenvolveu a partir da língua, ou pelo menos foi cronologicamente posterior à língua, e (3) que a língua se desenvolveu a partir da música, ou ambas se desenvolveram a partir de uma "proto-faculdade" comum.

⁹ O argumento convincente de Fonagy de que a proto-linguagem humana era icônica, isto é, que os componentes prosódicos e físicos eram diretamente expressivos de estados mentais internos, explica a popularidade da visão de que a música expressa principalmente emoção, enquanto a fala expressa o pensamento racional. Na visão de Fonagy, a primeira linguagem da humanidade - a fonte comum a partir da qual música e fala se desenvolveriam mais tarde - era claramente "musical", e os elementos musicais ou prosódicos expressavam o significado diretamente. O que nós entendemos como fala foi um desenvolvimento posterior de dupla articulação que evoluiu para expressar conceitos mais complicados que a variação de altura melódica sozinha era incapaz de expressar. Dupla articulação, ou dualidade de padrões, refere-se à divisão sônica de uma linguagem em partículas sonoras individuais - fonemas - que não têm significado intrínseco em si mesmas, mas podem ser combinadas em novas unidades de significado conhecidas como morfemas. O desenvolvimento da articulação dupla marca o ponto exato em que a língua deixa de ser isomórfica e se torna simbólica. Claro, é também o ponto em que música e fala começam a se separar e a seguir caminhos distintos. Ainda assim, mesmo depois de sua separação, música e fala continuam a ser intimamente conectadas através de seu componente de frequência comum.

The issue of pitch is fundamental to the question of the relatedness of music and speech. If my hypothesis is correct, then pitch at one time must have played as important a role in humankind's proto-language as it does today in music. **Clearly there are other differences between music and speech of a rhythmic, dynamic, or articulatory nature, but most would view these as differences of degree rather than kind**, and they will not be dealt with here in a systematic fashion.¹⁰ (LEVMAN, 1992, p. 152, grifos nossos)

Esta peculiar relação não se limita às suas origens. Ao contrário, o processamento e a produção de estímulos sonoros tanto musicais quanto de linguagem verbal mobilizam recursos comuns. Para além do truísmo de que o canto e a fala são ambos produzidos pelo sistema fonatório e apreendidos pelos ouvidos, há similaridades nos recursos cognitivos/cerebrais mobilizados na interpretação desses estímulos sonoros. Sascha Frühholz, Wiebke Trost e Didier Grandjean (2014) propõem um modelo de explicação da função do sistema límbico no processamento de emoções na voz e na música, inclusive a instrumental – um detalhe importante, que garante que esta relação se dá em um nível mais profundo, e não através de uma mediação do processamento dos elementos linguísticos presentes na música cantada.

Specifically, recent studies have shown that auditory stimuli such as vocal expressions [...] and musically expressed emotions [...] can elicit activity in the amygdala, indicating that both vocal emotions and musical emotions share similar ways of expression, and involve similar brain mechanisms for the decoding of and adaptive response to emotional cues.¹¹ (FRÜHHOLZ; TROST; GRANDJEAN, 2014, p. 2)

É possível, ainda que não haja evidências suficientes para tirar tal conclusão, que esses mecanismos cerebrais em comum estejam na raiz de outro fenômeno, este já apoiado por evidências mais abundantes: alterações no estímulo sonoro provocam o mesmo tipo de efeito de percepção tanto na música quanto na fala:

Using a three-dimensional model of emotions to study three particular acoustic cues, Ilie and Thompson (2006) found a variety of effects and interactions: intensity was found to influence both prosody and music in the same way (greater intensity was associated with higher ratings of valence,

¹⁰ O tópico da altura melódica é fundamental para a questão do parentesco da música e da fala. Se minha hipótese está correta, então a altura melódica, a uma certa altura, deve ter exercido um papel importante na proto-linguagem da humanidade tal qual exerce hoje na música. Claramente, há outras diferenças entre música e fala de natureza rítmica, dinâmica ou articulatória, mas a maioria veria essas diferenças de grau em vez de tipo, e elas não serão abordadas aqui de maneira sistemática.

¹¹ Especificamente, estudos recentes mostraram que estímulos auditivos como expressões vocais (...) e expressões expressadas musicalmente (...) podem evocar atividade na amígdala, indicando que tanto emoções vocais quanto emoções musicais compartilham caminhos de expressão similares, e envolvem mecanismos cerebrais similares para a decodificação de e resposta adaptativa a sinais emocionais.

tension and energy), rate had varying effects on both domains (faster music and speech was associated with higher energy, but whereas fast speech was judged as less pleasant than slow speech, fast music was judged as more tense than slow music), and pitch height influenced the two domains in opposite directions (high pitched speech and low pitched music were both attributed higher ratings of valence). Evidence that prosody and music share processing resources at an intermediate level in the auditory pathway comes from a study of perceptual discrimination in music and speech by two amusic subjects, which showed that performance was similar across domains although the participants had different perceptual deficits (Patel et al., 1998).¹² (COUTINHO, DIBBEN, 2013, p. 659)

Eduardo Coutinho e Nicola Dibben (2013), por exemplo, criaram um modelo computacional capaz de prever as emoções que estímulos acústicos evocariam nos informantes através de um grupo relativamente limitado de características psicoacústicas (entre elas, intensidade, andamento/velocidade de fala, e melodia/contorno prosódico). Evidentemente, a correspondência entre os dois domínios (musical e linguístico) não é exata. Como os próprios Coutinho e Dibben (2013, p. 680) existem “pistas” acústicas limitadas ao domínio musical, outras restritas ao domínio da fala, e ainda um terceiro grupo de pistas que seria compartilhado por esses dois domínios.

This distinction between shared and domain-specific attributes of music and speech is highlighted by Ilie and Thompson (2011, pp. 260_261) in their overview of emotional communication in the auditory domain. They speculated that cues which are shared across auditory domains are likely to resist enculturation in comparison to domain-specific cues, and that enculturation of these domain-specific cues may lead to fractionation of emotional communication systems, thus accounting for cross-cultural differences in emotional coding. For example, they argue that the association of higher pitch with motherese during child development may lead to high pitch in speech being associated with greater pleasantness than high pitch in music. Thus, while we argue that our results provide evidence of the importance of acoustic cues to emotional responses to music we recognise that cues are also subject to historical, social and technological contingencies¹³ (COUTINHO; DIBBEN, 2013, p. 680)

¹² Usando um modelo tridimensional de emoções para estudar três pistas acústicas particulares, Ilie e Thompson (2006) encontraram uma variedade de efeitos e interações: descobriu-se que a insensibilidade influenciava tanto prosódia quanto a música da mesma maneira (maior intensidade estava associada com avaliações maiores de valência, tensão e energia), velocidade de emissão teve efeitos variáveis em ambos os domínios (música mais rápida e fala mais rápida foram ambas associadas com maior energia, mas enquanto fala rápida foi julgada como menos agradável do que fala lenta, música rápida foi julgada como mais tensa do que música lenta) e a altura melódica influenciou os dois domínios em direções opostas (fala mais aguda e música mais grave receberam ambas avaliações mais altas de valência). Evidência de que prosódia e música compartilham recursos de processamento em um nível intermediário no caminho auditivo vem de um estudo de discriminação perceptiva na música e na fala por dois sujeitos amusicos, que mostrou que a performance era similar entre os domínios embora os participantes tivessem déficits perceptivos diferentes.

¹³ Esta distinção entre os atributos compartilhados e específicos do domínio da música e da fala é destacada por Ilie e Thompson (2011, p. 260 e 261) em sua visão geral da comunicação emocional no domínio auditivo. Eles especularam que pistas que são compartilhadas entre os domínios auditivos são susceptíveis de resistir a enculturação em comparação com pistas específicas do domínio, e que a enculturação dessas pistas específicas

Note-se que, ainda que a correspondência entre os dois domínios não seja uma relação de “um para um”, a correlação encontrada é notável e aponta para uma relação muito mais íntima do que se poderia supor à primeira vista.

Por último, mas não menos importante, lembramos que a grande maioria das línguas naturais, essencialmente verbais em sua origem, utiliza-se da mesma materialidade da música – o som – na produção de sentidos. Esse fato, por si só, já poderia justificar a aproximação entre as linguagens.

1.3. SOBRE A RELAÇÃO ENTRE A LINGUAGEM MUSICAL E A LINGUAGEM VERBAL EM COMPOSIÇÕES MUSICAIS E SEUS SENTIDOS

Qualquer perspectiva teórica à qual se adira no intuito de abordar os significados de um enunciado implicará limitações, descontinuidades e pontos discutíveis. No que se refere à música (e, mais especificamente, às letras de música), um trabalho de análise estaria incompleto caso não houvesse, pelo menos, uma breve problematização da relação entre música e texto. Tal questão é abordada por Blacking (1982), como se vê no fragmento a seguir:

Combining words and music is not just a matter of uniting text and melody in an agreeable harmony. It is a basic problem of human expression which has far-reaching implications: can the music and the speech modes be combined with equal attention to both without subordinating one to the other? Will words detract attention from music? Or vice versa? and if so, why? Will the physical strength and cultural priority of the brother dominate, or the moral force and natural superiority of the sister? Will either words or music dominate under all circumstances, or only under special conditions? ¹⁴
(BLACKING, 1982, p.18)

do domínio pode levar ao fracionamento dos sistemas de comunicação emocional, explicando as diferenças transculturais na condificação de emoções. Por exemplo, eles argumentam que a associação de altura melódica mais alta com o manhês durante o desenvolvimento da criança pode levar a alta altura na fala a ser associada com maior prazer do que o tom alto na música. Assim, ainda que argumentemos que nossos resultados fornecem evidência da importância das pistas acústicas para as respostas emocionais à música, reconhecemos que as pistas também estão sujeitas a contingências históricas, sociais e tecnológicas.

¹⁴ “Combinar palavras e música não é apenas uma questão de unir texto e melodia em uma harmonia agradável. É um problema básico de expressão humana que tem implicações de longo alcance: podem música e fala serem combinadas com igual atenção a ambas sem subordinar uma à outra? Irão as palavras desviar a atenção da música? Ou vice-versa? E se sim, por quê? Irá a força física e a prioridade cultural do irmão dominar, ou a força moral e superioridade natural da irmã a fazê-lo? Irão as palavras ou a música dominar sob todas as circunstâncias, ou apenas sob condições especiais?”

Em um nível mais superficial da relação entre palavra e música, pode-se mencionar que o contorno prosódico do texto tende – ou mesmo deve – coincidir com o contorno melódico. Um bom exemplo dessa concordância seriam os recitativos, que “trazem marcas entoacionais da fala: uma pergunta termina numa quarta ascendente e uma declaração termina num intervalo estreito, descendente” (MEDEIROS, 2002, p. 27, nota de rodapé). É possível que essa concordância se deva pelo fato de que música e fala são processadas, até certo ponto, pelos mesmos recursos cognitivos. Essa é a conclusão à qual chegam Coutinho e Dibben (2013), após testar um modelo computacional capaz de discriminar emoções expressas em sinais de áudio a partir de parâmetros acústicos:

Our study provides supporting evidence for the idea that emotional content of music and speech is decoded, at least partially, by a shared processor that responds to psychoacoustic features regardless of the type of sound source (Ilie & Thompson, 2006, 2011; Juslin & Laukka, 2003), and specifies more precisely the particular acoustic features involved.¹⁵(COUTINHO; DIBBEN 2013, p. 678-679)

Mas há níveis mais profundos de análise. Evidentemente, tratar da organização formal de um texto, a formação discursiva à qual ele pertence e os recursos estilísticos dos quais ele lança mão pode levar a resultados bastante distintos de uma investigação voltada para o que seus leitores efetivamente apreendem desse texto. Teorias como a Semiótica da Canção de Luiz Tatit (1997) não atentam para - tampouco problematizam - o fato de que, ao ouvir uma composição, um ouvinte pode não apreender a letra da música em sua totalidade, não apreender a estrutura musical em sua totalidade, ou não se aperceber da relação entre melodia, harmonia, ritmo e letra. E, caso apreendam, voltamos ao questionamento de Blacking: teriam ambas as linguagens, musical e verbal, o mesmo peso, ou uma delas prevaleceria? Uma das formas de lidar com essa questão poderia ser analisar a atividade cerebral e verificar em quais regiões do cérebro enunciados verbais e musicais são processados.

On balance, the evidence of the lateralization of brain function suggests that speech and music may be produced with the help of two different, though relatable, systems, and that they cannot be united on equal terms. They can be combined in song, but complementarity is not the same as unity (e.g. you could not unite the attention required to drive a car in heavy traffic and to carry on a complicated discussion without almost certain disaster, but you

¹⁵ Nosso estudo fornece evidências de apoio para a idéia de que o conteúdo emocional da música e da fala é decodificado, pelo menos em parte, por um processador compartilhado que responde a características psicoacústicas, independentemente do tipo de fonte sonora (Ilie & Thompson, 2006, 2011; Juslin & Laukka, 2003), e especifica mais precisamente as características acústicas particulares envolvidas.

could combine them with unequal attention to each task.).¹⁶ (BLACKING, 1982, p. 20)

Por mais sofisticados que os recursos cognitivos dos seres humanos sejam, tais recursos são limitados. Há limites para a nossa capacidade de atenção, processamento e memorização – disso resulta, por exemplo, a proibição de falar ao celular enquanto dirigimos. Reconhecendo esse fato, Blacking (1982, p. 21) finaliza:

Music and speech (considered as ideal types) may be the products of two different cognitive systems or modes of thought, and it may be impossible to merge one with the other without subordinating one to the other. Perhaps it is impossible to give full and equal attention to the content of music and speech at the same time.¹⁷

O texto de Blacking data de 1982. Nas décadas que se seguiram, muitos outros estudos sobre a atividade cerebral quando indivíduos são expostos a estímulos verbais e/ou musicais foram feitos, facilitados inclusive pelo avanço tecnológico. Outros trabalhos citados nesta dissertação podem, num primeiro momento, parecer contradizer as reflexões de Blacking (1982). No entanto, chamamos a atenção para o fato de que a maioria dos trabalhos aqui mencionados trata da percepção das *emoções* expressas e evocadas pela música. Tome-se o trabalho de Frühholz, Trost e Grandjean (2014, p. 13), por exemplo:

Taken together, our proposed network summarizes and points out that musical emotions and vocal expressions have common processing pathways for auditory processing, but that they are partly distinct in terms of emotional processing in the TLS given the differential nature of their communicational value.¹⁸

Uma análise musical exaustiva não se limita ao aspecto emocional de uma composição. Tome-se uma música cantada: pode-se questionar se é possível apreender os aspectos estruturais e técnicos em sua plenitude ao mesmo tempo em que se presta atenção na narrativa sendo entoada, ou se é possível atentar para todos os elementos estilísticos e poéticos da letra enquanto se presta atenção na progressão harmônica e nos *arpeggios* que acompanham a linha melódica, por exemplo. Isso para não entrar na questão do que significa,

¹⁶ “De forma balanceada, a evidência de lateralização de função cerebral sugere que fala e música podem ser produzidas com a ajuda de dois sistemas diferentes, ainda que relacionáveis, e que eles não podem ser unidos em termos iguais. Eles podem ser combinados em canção, mas complementaridade não é o mesmo que unidade (por exemplo, não se pode unir a atenção exigida para dirigir um carro em trânsito intenso e levar uma discussão complicada sem um desastre quase certo, mas seria possível combina-las com atenção desigual a cada tarefa.)”

¹⁷ “Música e fala (consideradas como tipos ideais) podem ser o produto de dois sistemas cognitivos ou modos de pensamento diferentes, e pode ser impossível fundir uma à outra sem subordinar uma à outra. Talvez seja impossível dar atenção plena e igual ao conteúdo de música e de fala ao mesmo tempo.”

¹⁸ Em conjunto, a rede proposta resume e aponta que emoções musicais e expressões vocais têm caminhos de processamento comuns pra o processamento auditivo, mas que elas são parcialmente distintas em termos de processamento no sistema límbico temporal, dada a natureza diferencial de seu valor comunicacional.

exatamente, perceber: é perceber conscientemente cada aspecto, é reagir racional e emocionalmente a uma obra, é saber descrever metalinguisticamente o que se ouviu? Existem camadas de leitura, diferentes maneiras de se abordar uma peça, e a cada audição é possível dirigir a atenção a um aspecto específico. Esses fatores fazem com que o fenômeno da audição musical seja muito mais complexo do que normalmente nos damos conta, e tentar abarcar todas essas variáveis em um único trabalho investigativo não é viável.

Não temos a pretensão de, neste trabalho, oferecer uma resposta ou uma solução mágica para lidar com o problema, mas julgamos necessário chamar a atenção para esse aparente impasse que, aliás, foi decisivo para a inclusão da questão referente ao envolvimento musical dos informantes nos questionários do experimento, uma vez que se supôs que menor envolvimento musical poderia implicar uma tendência a “negligenciar” aspectos puramente musicais das gravações em favor do conteúdo verbal, possivelmente mais concreto para a maioria dos participantes. Para mais detalhes acerca de fatores populacionais dos participantes e as questões que buscaram delimitar esses fatores, remetemos à seção número 3, de Metodologia.

Note-se que o questionamento acima, apesar de pertinente nos casos em que se busca depreender o sentido das composições, não configura necessariamente um problema para todos os estudos linguísticos que lançam mão de *corpus* musical. Estudos acerca de fenômenos fonológicos, por exemplo, podem prescindir de reflexão aprofundada sobre o assunto.

1.4. CONSIDERAÇÕES FINAIS

Nessa primeira seção, ainda que não tenhamos a pretensão de esgotar o debate, procuramos uma definição funcional de língua, linguagem e música. Abordamos as interseções, sobreposições e relações entre música e língua, defendendo que as aproximações entre essas duas linguagens, além de naturais, podem ser bastante prolíficas. Por fim, problematizamos brevemente essas aproximações do ponto de vista científico, não para desmerecê-las, mas no intuito de atentar para uma questão pertinente.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 LINGUÍSTICA

Uma vez que este trabalho trata com a língua em sua materialidade sonora, as interpretações e análises de resultados são apoiadas na fonética e na fonologia: Cagliari (2002), Lieberman e Blumstein (1988), Massini-Cagliari (1992) e Massini-Cagliari e Cagliari (2001). Chomsky (1971, 1994) e Lieberman (1975), por sua vez, são úteis para pensar as implicações dos resultados deste trabalho dentro de um contexto linguístico e cognitivo mais amplo.

2.1.1 PROSÓDIA

Historicamente, o termo *prosódia* era usado pelos gregos para se referir aos traços da fala que não podiam ser indicados na ortografia (COUPER-KUHLEN, 1986, p. 1). Massini-Cagliari (2015, p. 20) considera que a fala tem melodia (entoação, tons) e pulsação (acento, ritmo e duração) e toma esses fatores como prosódicos. Em uma definição sucinta, pode-se dizer que a prosódia inclui todos os aspectos acústicos da fala não limitados ao segmento e que possuem funções linguísticas (CÓRDULA, 2013, p. 26). Entre esses aspectos, pode-se citar a altura melódica (variando entre sons graves e agudos), a intensidade (sons mais fortes ou mais suaves) e o timbre (a qualidade do som). A manipulação destes aspectos tem uma função comunicativa crucial: através dela, podemos diferenciar uma pergunta de uma afirmação, definir o tópico frasal e expressar atitudes do falante, por exemplo, seu estado emocional (para mais detalhes, ver Massini-Cagliari e Cagliari, 2001).

Considerando que os sons da fala podem ser analisados fonética ou fonologicamente (e que os dois tipos de análise são complementares), é possível também pensar em definições fonéticas e fonológicas para a prosódia. De um ponto de vista mais físico e fonético, uma boa definição seria que “Speech prosody is the pattern of acoustic changes within spoken utterances that communicate meaning independently of verbal comprehension”¹⁹ (COUTINHO, DIBBEN, 2013, P. 658). Uma explicação mais fonológica, ou seja, levando em consideração a função desses fatores no sistema linguístico, pode ser encontrada em Gebara (1976, p. 8):

¹⁹ Prosódia é o padrão de mudanças acústicas em enunciados falados que comunicam significado independentemente da compreensão verbal.

No nível dos elementos vocais [...], dois componentes, pelo menos, devem ser notados: um componente verbal ou segmental, responsável pela estrutura da sentença tradicionalmente considerada: fonologia, gramática e léxico; e um componente não verbal, constituído por subcomponentes prosódicos (ou supra-segmentais) e paralingüísticos, que são aspectos da cadeia sonora que têm relação variável com os elementos constitutivos do componente verbal do discurso, no sentido de que a cada sequência segmental não corresponde necessariamente uma determinada sequência supra-segmental.

Frequentemente se diz que a prosódia é a “música” da língua. Assumindo essa analogia, poderíamos dizer que humanos são seres bastante musicais, já que a prosódia é o primeiro traço da linguagem verbal que os bebês adquirem – de fato, o choro de bebês recém-nascidos reflete os padrões entoacionais da língua de seus pais (MAMPE et al., 2009). Não apenas a prosódia é adquirida precocemente, mas também parece estar ligada à musicalidade, aspectos prosódicos de certas línguas parecem estar evidentes em músicas de países em que tais línguas são faladas - ou vice versa, uma vez que as correlações encontradas não explicitam uma direção causal (ver PATEL, IVERSEN e ROSENBERG, 2006). A manipulação da altura melódica e intensidade sonora durante a fala é mantida pelos seres humanos ao longo da vida; ainda assim, parece haver uma diferença na importância atribuída a essa “música” em diferentes faixas etárias.

A compreensão, por parte de crianças, das emoções na fala foi explorada em três experimentos por Morton e Trehub (2001). No primeiro experimento, adultos e crianças entre quatro e dez anos julgaram a felicidade ou a tristeza expressas por um falante através do conteúdo semântico das palavras e da paralinguagem afetiva. Diante de conflitos (por exemplo, quando uma situação feliz era descrita usando paralinguagem triste), as crianças privilegiaram o conteúdo semântico, ao contrário de adultos, que se basearam na paralinguagem. Foram observadas mudanças graduais, com as crianças de quatro anos focando quase exclusivamente em conteúdo semântico e adultos focando quase exclusivamente na paralinguagem. Também foi notado que as crianças rotularam a paralinguagem afetiva com exatidão quando expostas a um enunciado em língua estrangeira, com o conteúdo semântico obscurecido. Os autores julgam que os dados são consistentes com a compreensão limitada das crianças acerca das funções comunicativas da paralinguagem afetiva (MORTON, TREHUB, 2001).

Um dos aspectos da prosódia é a entoação, que se refere à altura melódica da fala (CÓRDULA, 2013, p. 26), ou seja, se refere à variação da frequência fundamental (F_0) da voz durante a produção da fala. Devido a essa variação, é possível fazer uma analogia e dizer que

os padrões entoacionais são como linhas melódicas de uma língua. Junto à intensidade, velocidade de fala e outros parâmetros, a entoação complementa ou modifica o sentido das palavras enunciadas.

While expressions of anger, fear, and happiness, for example, are accompanied by increased vocal intensity including an enhanced intensity variation, expressions of sadness are usually accompanied by low intensity. Similarly, angry, happy, and fearful voices show higher levels of pitch as mainly determined by the F0 as well as pitch variability compared to sad voices.²⁰ (FRÜHHOLZ, TROST, GRANDJEAN, 2014, p. 10)

Obviamente, línguas diferentes fazem usos distintos dos parâmetros prosódicos. Em línguas tonais, por exemplo, variações de altura melódica (padrões entoacionais) agem sobre um item lexical, fazendo com que uma mesma sequência de fonemas tenha significado distintos; por outro lado, em línguas entoacionais, a variação da altura melódica abre sobre a frase, mantendo o significado lexical e morfológico, mas alterando a função sintática e semântica/pragmática (CÓRDULA, 2013, p. 26). É interessante notar, no entanto, que apesar das diferenças entre os sistemas fonológicos entre as línguas, há evidências de que certos aspectos sejam universais, especialmente no que se refere à expressão de emoções:

Similarly to facial expressions (Ekman, 1992), certain aspects of emotional expression in speech prosody appear to be universal (e.g., Scherer, Banse, & Wallbott, 2001) and convincing evidence points to similar prosodic codes used across cultures to convey similar emotions (e.g., Thompson & Balkwill, 2006). This is apparent, for instance, in our capacity to decode emotional meaning even in unfamiliar languages.²¹ (COUTINHO, DIBBEN, 2013, p. 659)

Outro aspecto prosódico é o ritmo, com o qual podemos traçar um paralelo imediato com a música. Ritmo se refere à organização dos sons no tempo, e aos padrões de proeminências que emergem dessa organização. Como menciona Cagliari (1981, p. 123), “o ritmo se manifesta através do movimento de um fenômeno que se desdobra no tempo, pondo em relevo repetidamente algum aspecto desse mesmo fenômeno”. Ainda que a fala não possua um padrão rítmico tão marcado quanto, por exemplo, o de uma valsa ou de um poema

²⁰ Enquanto expressões de raiva, medo e felicidade, por exemplo, são acompanhadas por intensidade vocal aumentada, incluindo uma variação de intensidade aumentada, expressões de tristeza são geralmente acompanhadas por baixa intensidade. Da mesma forma, as vozes irritadas, felizes e temerosas mostram níveis mais altos de altura melódica, como determinado principalmente pelo F0 e também pela variabilidade de altura melódica em comparação com vozes tristes.

²¹ Similar a expressões faciais (Ekman, 1992) certos aspectos da expressão emocional na prosódia parecem ser universais (por exemplo, Scherer, Banse e Wallbott, 2001) e evidências convincentes apontam para códigos prosódicos similares usados entre culturas para expressar emoções similares (por exemplo, Thompson & Balkwill, 2006). Isso é aparente, por exemplo, na nossa capacidade de decodificar significados emocionais mesmo em línguas desconhecidas.

metrificado, isso não significa que a fala não tenha um ritmo, pois “se a fala não tivesse ritmo, seria impossível para qualquer ser humano falar, pois ritmo é uma maneira que a linguagem tem para organizar no tempo o que deve ser dito (em termos segmentais)” (MASSINI-CAGLIARI, 1992, p. 11). O ritmo linguístico não é exato e regular e nem poderia ser marcado com um metrônomo, pois envolve questões como a expectativa dos falantes (e dos ouvintes)²²:

Neste sentido, o significado da palavra RITMO, em Lingüística, não se resume apenas a padrões muito rígidos de repetição de acentos ou durações, como nos tratados de metrificação, mas abrange a maneira como as manifestações linguísticas dos seres humanos são organizadas no tempo ao serem pronunciadas. Por outro lado, não se resume apenas a padrões quaisquer de repetição, mas abrange a noção de expectativa de uma eventual repetição de algum parâmetro no tempo. (MASSINI-CAGLIARI, 1992, p. 11)

O conceito de pé rítmico é a base para o estudo do ritmo linguístico, mas está intimamente ligado com a ideia de compasso musical:

Na ciência linguística, a noção de PÉ foi desde cedo incorporada aos estudos de fonética. Correspondia, em muito, ao que, em música, é chamado de *compasso: divisão da música em pequenas partes de duração igual ou variável* (Lacerda, 1976, p. 15). Desta forma, da hierarquia entre os tempos dos compassos, em música (o primeiro tempo de cada compasso é sempre considerado “forte” e os demais “fracos – ou, em alguns casos, de força relativa ou “mezzo-forte”), resulta, por analogia, a definição fonética de PÉ: *unidade de duração compreendida entre duas tônicas (incluindo a primeira e excluindo a segunda)* – Massini Cagliari (1992, p. 68 nota 1). (MASSINI-CAGLIARI, 1999, p. 118)

Nota-se, através desta breve discussão acerca dos elementos prosódicos da língua falada, o quanto a fonologia deve à música na elaboração de seus conceitos e nomenclatura.

2.2 MÚSICA

O presente trabalho tem um caráter alinhado, teoricamente, ao pensamento gerativista. O gerativismo se baseia na ideia de que todos os seres humanos são dotados de um mecanismo cognitivo predisposto geneticamente para a aquisição, processamento e produção da linguagem. Esse dispositivo seria comum a todos os seres humanos, do que se depreende

²² Para mais detalhes e para uma discussão aprofundada sobre a maneira como, mais do que o tempo, a intensidade sonora é um parâmetro acústico organizador do ritmo na língua, remetemos ao livro *Acento e Ritmo*, de Massini-Cagliari (1992).

que as habilidades linguísticas também apresentariam características comuns a todas as populações do mundo. Uma das tarefas da Linguística, assim, seria justamente encontrar esses elementos universais nas línguas do mundo, e elucidar o funcionamento comum desse mecanismo cognitivo humano. Se a música também é universal e utiliza elementos também utilizados pela linguagem, e levando em consideração que música e língua têm a mesma origem evolutiva, parece razoável supor que, de alguma forma, a disposição inata dos seres humanos para a linguagem seja compartilhada, em algum grau, com a disposição para desenvolver habilidades musicais. Essa suspeita é reforçada pelos estudos acerca da origem da linguagem verbal e da música, que mostram diversas aproximações possíveis entre essas duas manifestações humanas.

Although frequently regarded as subjective and culturally grounded, there is convincing evidence that music can express emotions that are recognised universally (e.g., Balkwill, Thompson, & Matsunaga, 2004; Fritz et al., 2009), a phenomenon that is associated with acoustic profiles that transcend cultural boundaries.²³ (COUTINHO; DIBBEN, 2013, p. 659)

Um dos pilares do gerativismo é a existência de universais linguísticos, isto é, características comuns a todas as línguas humanas, devido à natureza inata da faculdade linguística, que por sua vez seria geneticamente determinada. Ainda que a perspectiva gerativa não seja aceita por todos e que lucubrações teóricas acerca dos mecanismos que governam a língua estejam longe de acabar, a noção de que os recursos cognitivos e biológicos humanos exercem influenciam (e, até certo ponto, moldam) a língua parece ser um ponto pacífico, mesmo para críticos do gerativismo. Tomemos, como exemplo, o linguista Marcos Bagno, que em sua *Gramática Pedagógica do Português Brasileiro*, explicita a concepção de língua presente na obra. Após uma longa discussão na qual, entre outros tópicos, associa a Gramática Tradicional, o Estruturalismo de Saussure e o Gerativismo de Chomsky a uma visão platônica da língua, Bagno não deixa de conformar-se com *certa* universalidade cognitiva:

Os chamados universais linguísticos não são uma herança geneticamente transmitida em nosso DNA. Eles resultam da interação entre as capacidades cognitivas dos seres humanos (que são as mesmas para toda a espécie), as restrições psicofísicas dessas capacidades (o sistema articulatório e respiratório, os limites do processamento mental e da memória etc.) e o mundo físico, com o qual os humanos têm de interagir para dar sentido à

²³ Embora freqüentemente considerada subjetiva e fundamentada na cultura, há evidências convincentes de que a música pode expressar emoções que são reconhecidas universalmente (por exemplo, Balkwill, Thompson & Matsunaga, 2004, Fritz et al., 2009), um fenômeno que está associado com perfis acústicos que transcendem fronteiras culturais.

existência e manter a coesão social. Uma vez que todos os seres humanos têm de enfrentar os mesmos desafios de conhecimento e experiência com os recursos biológicos de que dispõem, as línguas têm de apresentar um grau de similitude. [...] A possibilidade de traduzir de uma língua para outra, por mais diferentes que seja, possibilidade documentada desde os tempos mais remotos da civilização, é a prova empírica cabal da universalidade sociocognitiva dos mecanismos da linguagem humana. (BAGNO, 2012, p. 78-79)

Quaisquer que sejam as causas dos universais linguísticos, e sem intenção de apelar à polarização antiquada e reducionista de *nature versus nurture*²⁴, é inegável que a linguagem verbal, enquanto manifestação humana, não é uma convenção social arbitrária: ao contrário, por exemplo, de regras de etiqueta (“é feio colocar os cotovelos na mesa”), as línguas do mundo não são puramente construções sociais.

A música, por outro lado, parece ser uma linguagem muito mais arbitrária, pelo menos à primeira vista. Parece difícil conceber um motivo razoável para que acordes maiores sejam relacionados a sentimentos alegres e acordes menores à tristeza, além do uso consagrado pela tradição. Pessoas que passaram a vida imersas em uma tradição musical que utiliza acordes menores para simbolizar a tristeza aprenderiam, com o tempo, a associar acorde e sentimento. Seguindo essa lógica, alguém que crescesse em uma tradição musical radicalmente diferente não faria a mesma associação e teria uma percepção diferente dos acordes.

No entanto, há evidências de que esse não seja necessariamente o caso: um estudo (SMITH, L. D, WILLIAMS, R. N, 1999) em especial sugere que a percepção de significados no discurso musical seja um fenômeno universal, e não baseado na cultura. Em um experimento, meninos brancos da África do Sul fizeram desenhos como resposta a quatro intervalos musicais. Em um segundo experimento, seis intervalos musicais serviram como estímulo para que informantes de ambos os sexos oriundos de populações brancas e negras, rurais e urbanas, fizessem seus desenhos. Observou-se que o conteúdo dos desenhos baseados nos intervalos musicais era similar entre as diferentes culturas: consonâncias foram percebidas como geralmente positivas, dissonâncias como geralmente negativas. Foi percebida também uma dimensão de atividade, com intervalos de sétimas, quintas e oitavas interpretados como mais ativos e terças e quartas (perfeitas ou aumentadas) como mais passivas.

Although skill in listening can be acquired, as can skill in playing, being responsive to musical discourse is not a matter of mastery of notation, but of receptiveness to meaning and message. This study has shown, at least to a

²⁴ Em tradução livre, "natureza versus criação". Esta dicotomia se refere à diferença de importância que se pode dar às qualidades inatas de uma pessoa (em especial aquelas determinadas por fatores biológicos) em relação às experiências pessoais (sua história, o meio em que vive e a forma como foi criada), na tentativa de explicar o comportamento humano e questões sociais.

limited degree, that even young and unsophisticated listeners associate even something as basic as two notes with certain general emotions. Although the suggestion that the purpose of music is to express emotion might be considered at best romantic, and rather superficial, we should remember, as argued in the introduction to this study, that emotion and cognition are interwoven, rather than separate, and in some way inimical, activities. What we are suggesting is that even small musical fragments have their own emotional and cognitive colors, so participating in musical activity, whether as a composer, performer or listener, is to engage in a powerful and basic means of communication.²⁵ (SMITH, WILLIAMS, 1999, p. 408)

Do ponto de vista musical, este trabalho se baseia na ideia de intervalos e tonalidade. As explicações a seguir são adaptadas de Schreuder (2006), Curtis e Bharucha (2010) e Bennet (1986). Uma vez que não se trata de um trabalho sobre música em si, não discutiremos os conceitos a seguir em grande profundidade. A proposta não é mobilizar todo o arcabouço teórico e discutir os conceitos de música, mas elucidar o básico desses conceitos para auxiliar a compreensão do trabalho por parte de leitores que não possuem formação musical.

2.2.1 MELODIA E RITMO

Para a maioria das pessoas, a **melodia** é o componente mais importante numa peça musical. Todo o mundo sabe, naturalmente, o que é melodia, palavra muito comum, cujo significado, no entanto, é difícil ser precisado com exatidão. Um dicionário musical sugere a seguinte definição: “sequência de notas, de diferentes sons, organizadas numa dada forma de modo a fazer sentido musical para quem escuta”. Contudo, o modo de reagir a uma melodia é questão muito pessoal. Aquilo que faz “sentido musical” para um pode ser inaceitável para outro, e o que se mostra interessante e até belo para uma pessoa pode deixar uma outra inteiramente diferente. (BENNET, 1986, p. 11)

Quando se trata de obras de arte, hierarquizações são sempre problemáticas, e a ideia de que a melodia é o componente mais importante de uma peça musical, como exposto por Bennet no fragmento acima, é bastante discutível. O que se pode dizer é que uma melodia é definida como uma sequência de sons de altura definida, dispostos ao longo do tempo de acordo com dadas convenções e limitações culturais (RINGER, 2017). A esses sons se dá o

²⁵ Embora a habilidade de ouvir possa ser adquirida, assim como a habilidade de tocar, ser responsivo ao discurso musical não é uma questão de domínio da notação, mas de receptividade ao significado e à mensagem. Este estudo mostrou, pelo menos em um grau limitado, que mesmo ouvintes jovens e não sofisticados associam até mesmo algo tão básico quanto duas notas a certas emoções gerais. Embora a sugestão de que o propósito da música seja expressar emoção possa ser considerada, na melhor das hipóteses, romântica e superficial, devemos lembrar, como argumentado na introdução deste estudo, que a emoção e cognição estão atividades entrelaçadas, em vez de separadas, e em certos aspectos contrárias. O que estamos sugerindo é que mesmo pequenos fragmentos musicais têm suas próprias cores emocionais e cognitivas, de modo que participar da atividade musical, seja como compositor, intérprete ou ouvinte, é envolver-se em um poderoso e básico meio de comunicação.

nome de notas, ou seja, uma nota musical nada mais é do que um som com uma frequência de onda específica. As frequências de ondas determinam o quão grave ou agudo um som é. Logo, é possível dizer que uma melodia é a sequência estabelecida de sons sucessivos, com duração e frequência específicas.

Uma explicação talvez mais acessível, considerando o senso comum, seria que a melodia é aquilo que, em geral, as pessoas cantam. Quando se pede a alguém que cante “Atirei o pau no gato”, é quase certo que a pessoa entoará (ou tentará entoar) a melodia dessa canção. Se alguém tenta evocar a Quinta Sinfonia de Beethoven em Dó menor com a voz, solfejando os primeiros cinco compassos (os incisivos mais icônicos da peça), através de um “Tam-tam-tam-tam” ou “Pam-pam-pam-pam”, estará entoando a melodia da sinfonia. Quando cantarolamos, ainda que simples vocalizações (como “Lá lá lá lá lás”), estamos entoando melodias.

Mesmo que seja discutível dizer que a melodia sozinha é a parte mais importante da música, é possível dizer que a melodia e ritmo são os dois principais parâmetros da estrutura musical.

Changes in instrumentation, orchestration or dynamics (i.e. changes to the secondary parameters) are understood as different arrangements of the same musical work, whereas alterations in pitch or rhythm may result in a new, different work. If pitch is concerned with the disposition of the frequencies of musical notes, then rhythm is concerned with the description and understanding of their duration and durational patternings. These durations may be more or less regular, may or may not give rise to a sense of beat or tempo, and may be more or less continuous, but as all music involves duration(s), all music necessarily has some manner of rhythm.²⁶ (LONDON, 2017)

A melodia se diferencia da harmonia por se referir a sons sucessivos, mas não sobrepostos, e do ritmo, que se refere aos “diferentes modos pelos quais um compositor agrupa os sons musicais, principalmente do ponto de vista da duração dos sons e de sua acentuação” (BENNET, 1986, p. 12). Frequentemente, o ritmo é uma unidade de padrões que se repete de forma cíclica ao longo da composição. Até certo ponto, a separação entre ritmo e melodia é uma abstração conceitual: é impossível entoar uma melodia sem ritmo. No entanto, é possível haver uma música com ritmo, mas sem melodia – basta pensar na música produzida

²⁶ Mudanças na instrumentação, orquestração ou dinâmica (isto é, mudanças nos parâmetros secundários) são entendidas como diferentes arranjos da mesma obra musical, enquanto alterações na afinação ou ritmo podem resultar em uma obra nova, diferente. Se a afinação tem a ver com a disposição de frequências das notas musicais, então o ritmo está relacionado com a descrição e a compreensão de suas durações e padrões de duração. Essas durações podem ser mais ou menos regulares, podem ou não dar origem a um senso de batida ou andamento, e podem ser mais ou menos contínuas, mas, uma vez que toda música envolve duração (durações), toda música necessariamente tem alguma forma de ritmo.

pelos tambores taiko japoneses cujo som, apesar de não terem altura melódica definida e, portanto, não constituir melodia, certamente é música.

3.2.2 INTERVALOS

Um intervalo nada mais é do que a distância entre duas notas específicas. Cada nota tem uma frequência própria. Por exemplo, tome-se a nota Lá (ou A, dependendo da notação musical utilizada) indicando um som de frequência 440 Hz²⁷ e suas divisões e multiplicações por dois (sons de 880 Hz e 220 Hz, apenas uma oitava acima e uma oitava abaixo, respectivamente, do Lá 440 Hz). Dentro de uma oitava, duas notas Lá (digamos, Lá vibrando a 440 Hz e Lá a 880 Hz) estão a doze semitons de distância uma da outra - isso pode ser facilmente observado em um piano, no qual as duas notas Lá estarão a cinco teclas pretas e sete teclas brancas de distância. A razão da frequência entre dois semitons é igual a $\sqrt[12]{2}$, aproximadamente 1,0595. Tomando duas notas quaisquer, dependendo da distância entre elas, teremos um intervalo diferente. Um intervalo de quatro semitons (por exemplo, entre Dó natural e Mi natural) é chamado de terça maior, e um intervalo de três semitons (por exemplo, entre Dó natural e Mi bemol) é chamado de terça menor. Intervalos de terça menor, em música, estão frequentemente associados a tristeza, enquanto intervalos de terça maior são associados com alegria²⁸. Essa diferença nos intervalos também é determinante no que se refere à tonalidade de uma composição.

²⁷ A nomenclatura da afinação é, em última análise, arbitrária, o que não significa que não existam padrões e tradições. O padrão Cammerton, por exemplo, atualmente implica a nota Lá a 400 Hz, ainda que este valor tenha oscilado ao longo dos séculos (HAYNES, 2017). Nos dias atuais, é mais comum que orquestras sigam o temperamento Lá = 442 Hz, ou eventualmente 443 Hz, mas o valor 440 Hz ainda costuma ser o utilizado em exposições didáticas, como a desta seção.

²⁸ Evidentemente, como em qualquer forma de arte, esta não é uma regra inviolável. Compositores podem brincar com essas associações, e é possível encontrar músicas tristes escritas em tom maior ou alegres escritas em tom menor, por exemplo.

2.2.3 HARMONIA

A harmonia ocorre quando duas ou mais notas de diferentes sons são ouvidas ao mesmo tempo, produzindo um acorde. Os acordes são de dois tipos: consoantes, nos quais as notas concordam umas com as outras, e dissonantes, nos quais as notas dissoam em maior ou menor grau, trazendo o elemento de tensão à frase musical. Usamos a palavra “harmonia” de duas maneiras: para nos referirmos à seleção de notas que constituem determinado acorde e, em sentido lato, para descrevermos o desenrolar ou a progressão dos acordes durante toda uma composição. (BENNET, 1986, p. 11)

Ao contrário da melodia, que se refere a notas musicais sucessivas, a harmonia se refere às notas quando tocadas simultaneamente. A escolha dos acordes em uma determinada composição é, em geral, influenciada pelo modo ou pela tonalidade dessa composição.

2.2.4 ESCALAS, MODOS E TONALIDADES

Para se compreender o que são modos e tonalidades, primeiramente é necessário que se explique o que são escalas. Cada tradição musical tem um conjunto de escalas que podem ser utilizadas para compor uma melodia. Uma escala nada mais é do que uma forma de dividir o espectro sonoro.

A escala diatônica, que é a principal escala de música tonal, é uma escala formada por sete tons cujos intervalos podem ser de um tom inteiro ou de um semitom, ao contrário da escala cromática que é composta por doze notas cujos intervalos são todos de semitons. As escalas diatônicas são, portanto, doze conjuntos de sete notas dentro do círculo cromático com um padrão fixo de intervalos [...]. A oitava nota da escala diatônica é igual à primeira, num registro mais agudo, i. e., o intervalo existente entre elas é o de oitava. Se a escala for do modo maior, o intervalo semitonal ocorre entre o terceiro e o quarto grau da escala (III-IV), e entre o sétimo e oitavo grau da escala (VII-VIII). A escala sobre o tom dó (C) é o exemplo de escala natural (i.e. sem sustenidos ou bemóis) em modo maior: dó, ré, mi, fá, sol, lá, si e dó (8ª). Os intervalos semitoniais ocorrem entre mi e fá e entre si e dó. Se a escala for do modo menor, os intervalos semitoniais ocorrem entre o segundo e o terceiro grau da escala (II-III) e entre o quinto e o sexto (V-VI) graus da escala. A escala de lá (A) é o exemplo de escala natural em modo menor: lá, si, dó, ré, mi, fá, sol e lá (8ª). (FERREIRA, 1993, p. 19-20)

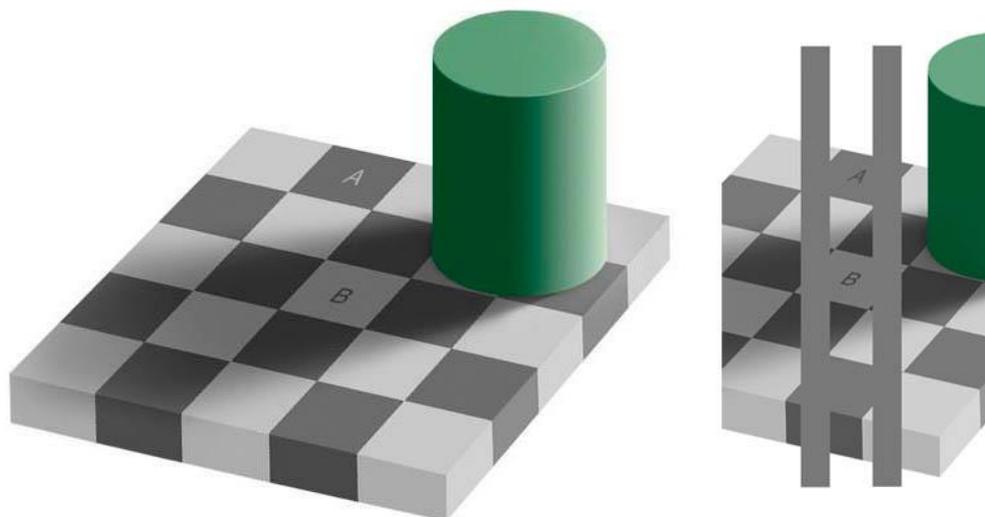
Para os leigos, uma boa analogia seria o de uma paleta de cores: ao pintar um quadro, um artista tem à sua disposição um conjunto de tintas de diversas cores que vão ser utilizadas na composição da imagem; diferentes combinações de cores produzem diferentes efeitos de

sentido e sensações no observador. Da mesma forma, cada escala pode ser tomada como uma paleta de frequências/notas, análogas às cores da tinta. A escala de Ré maior, por exemplo, é composta das “cores” Ré natural, Mi natural, Fá sustenido, Sol natural, Lá natural, Si natural e Dó sustenido. Se uma música é composta utilizando a escala de Ré maior, dizemos que o tom da música é Ré. Se uma música é composta utilizando a escala de Dó sustenido maior, dizemos que o tom da música é Dó sustenido maior, e assim por diante.

É importante notar que “cores iguais” não implicam necessariamente “expressões idênticas”. Podemos ter dois quadros, ambos pintados com tons de azul e tons de vermelho. Mas se um dos quadros usa predominantemente tons de azul e uns poucos de vermelho, certamente causará um efeito distinto de outro quadro que use muitos tons de vermelho e pouco azul. O mesmo ocorre com as escalas: assim como as cores de uma pintura, seu efeito só é percebido quando analisamos o contexto em que aparecem.

Voltando à analogia da paleta de cores: uma mesma cor pode ser percebida como mais clara ou mais escura dependendo do contexto em que se encontra, ou seja, das outras cores que a cercam. Tomemos, como exemplo, a imagem abaixo:

Figura 1 Ilusão de óptica



Fonte: <http://hypescience.com/incriveis-ilusoes-de-optica-que-irao-confundi-lo/>

Os quadrados marcados com a letra A e com a letra B têm exatamente a mesma cor, o mesmo tom de cinza. Ainda assim, enxergamos como dois tons de cinza diferentes. Isso acontece porque nosso cérebro não é apenas um decodificador do estímulo visual: interpretamos o estímulo de acordo com o contexto em que ele ocorre, e o contexto dos quadradinhos na imagem faz com que nosso cérebro se esforce para perceber as cores dos quadrados A e B como diferentes.

O mesmo ocorre com as notas musicais: uma onda de frequência 440 Hz é, a princípio, interpretada como um Lá natural, mas pode ser interpretada como Si bemol, dependendo do contexto em que ocorrer. Em outras palavras, uma mesma frequência exerce “funções” diferentes e, assim, recebe nomes diferentes para diferenciar essas funções, dependendo do contexto em que ocorre. Em música, esse contexto é definido por todas as outras notas utilizadas na composição e as relações que se estabelecem entre elas. Uma vez que, como mencionamos anteriormente, as notas utilizadas em uma composição são, de forma geral, definidas pelo tom de uma composição, pode-se dizer que, dependendo da escala, cada nota terá uma função diferente²⁹. Essa questão também é abordada por Ferreira (1993, p. 21):

Observa-se, assim, que a música tonal ocidental se baseia, do ponto de vista teórico-prático, na hierarquia de relações intervalares da escala diatônica e na supremacia da tônica sobre as demais notas da escala, para se determinar a tonalidade de uma música. Nesta hierarquia figuram, também, os elementos da tríade: o primeiro, o terceiro e o quinto graus.

Para os propósitos deste trabalho, basta esclarecer duas funções: a Tônica (1º grau) e a Mediante (3º grau). Tomemos a sequência de sete notas da escala de Dó: Dó, Ré, Mi, Fá, Sol, Lá, Si. A primeira nota dessa sequência corresponde ao primeiro grau (tônica), a segunda nota ao segundo grau, a terceira nota ao terceiro grau (mediante) e assim por diante. Quando, em uma escala, o intervalo entre a tônica e a mediante é de quatro semitons, um intervalo de terça maior, dizemos que a escala é maior, e uma música composta nessa escala está em modo maior. Quando o intervalo entre a tônica e a mediante é de três semitons, um intervalo de terça menor, dizemos que a escala é menor, e uma música nessa escala está composta em modo menor. Seguindo o exemplo, a escala de Dó maior tem Dó e Mi naturais (uma distância de quatro semitons), enquanto a escala de Dó menor tem Mi bemol, reduzindo a distância em

²⁹ Esta é uma simplificação para fins didáticos. Seria mais preciso, do ponto de vista da teoria musical, dizer que essa relação entre notas e tonalidade é circular: ao mesmo tempo em que as notas utilizadas sugerem um tom, são sugeridas por ele. Um compositor pode escolher o tom em que deseja compor, e essa escolha por si só sugerirá as notas a serem utilizadas, ou pode compor sem se preocupar com o tom, e a composição pronta, quando ouvida, será interpretada pelos ouvintes como dentro de uma tonalidade específica.

um semitom (totalizando, assim, três semitons). Lembrando que intervalos de terça menor e músicas em tom menor são associadas à tristeza (ver, por exemplo, Curtis e Bharucha, 2010).

É importante reforçar que, para que seja atribuída uma tonalidade e um modo a uma composição, não basta que esta composição apresente um determinado conjunto de notas. As escalas de Dó maior natural e Lá menor natural, por exemplo, apresentam as mesmas notas (ou as mesmas cores, se mais uma vez pensarmos na metáfora da paleta de cores). A diferença entre as duas escalas mencionadas está na relação estabelecida por estas notas. O estabelecimento das relações entre as notas é o cerne do trabalho de Ferreira (1993), que observou a maneira como ouvintes com maior ou menor experiência musical organizam (hierarquizam) as notas em composições musicais tonais e atonais. Ao revisar a bibliografia sobre o assunto disponível, Ferreira (1993, p. 27-28) conclui que:

Todos esses trabalhos parecem indicar que indivíduos de diferentes idades e níveis de experiência musical percebem o padrão tonal, todavia, esta percepção pode ser facilitada pelo tipo de contorno e direção melódica sem muitas mudanças, pela estrutura rítmica em fase com a melodia, por uma segmentação temporal estável e pela proximidade intervalar.

Conclui-se, portanto, que:

[...] a condição tonal possivelmente deve ser dada por uma ordem temporal das notas. Não basta, assim, o fato das notas pertencerem a uma escala, ou seja, possuírem uma determinada frequência acústica que as aproxima através de relações artificialmente estabelecidas, para que se perceba a tonalidade. Esta percepção, bem como a hierarquia das relações, passa a depender de uma ordem temporal das notas na música, ao que se pode acrescentar, uma ordem encontrada nas composições musicais e que, do ponto de vista gestáltico 'deve' possuir uma "boa forma". Além disso, não só determinadas sequências intervalares, mas também fatores extratonais devem contribuir para fornecer a "boa forma" e ressaltar o centro tonal. (FERREIRA, 1993, p. 12-13)

2.3 EMOÇÕES

Uma vez que este trabalho trata de expressões musicais e verbais de emoções, faz-se necessário, também, definir o que se entende por emoções. Em psicologia, o conceito de “emoção” tem se mostrado repetidamente vago e inadequado e, ainda assim, permanece (ENGELMANN, 1978, p.15). Ainda que qualquer nomenclatura adotada seja evidentemente limitada e discutível, problematizar o que se entende por “emoção” e/ou qual o melhor

vocábulo para se referir a esse conceito está fora do escopo desse trabalho. Reconhecemos que vários sistemas de classificação de emoções foram criados ao longo do tempo, variando tanto na concepção de emoções quanto em quais emoções seriam incluídas nesses sistemas³⁰. Um exemplo digno de nota, mencionado por Frühholz, Trost e Grandjean (2013, p. 03) são as Escalas Musicais Emocionais de Geneva (Geneva Emotional Music Scales, ou GEMS), um sistema elaborado para mensurar emoções especificamente evocadas pela música. No entanto, considerando a bibliografia disponível ao pesquisador na época da elaboração do teste de percepção, bem como as bases sobre as quais o teste de percepção foi criado (ver seção número 3, de Metodologia), optamos por nos ater às definições de Engelmann. Assim, o conceito de “emoção” ao qual nos referimos ao longo deste trabalho corresponde ao que Arno Engelmann, na tentativa de classificar e traçar correspondências entre relatos verbais e estados subjetivos, chama coletivamente de “emoções e/ou...” (ENGELMANN, 1978, p. 38).³¹

É sabido que as emoções experimentadas pelos indivíduos se manifestam através de vários movimentos e modificações no organismo, observáveis direta ou indiretamente (ENGELMANN, 1978, p. 59). No que se refere à linguagem, um falante pode utilizar diferentes padrões entoacionais para fazer acréscimos ou pequenas modificações de significado nos enunciados, expressando, por exemplo, raiva, dúvida ou tristeza (MASSINI-CAGLIARI, CAGLIARI, 2001, p. 118-119, HALLIDAY, GREAVES, 2008, p.50). Entretanto, ainda não foram encontradas variações acústicas que sejam indicadores *exclusivos* de emoções (ENGELMANN, 1978, p. 73).

É importante ressaltar que, assim como o trabalho de Coutinho e Dibben (2013), a presente pesquisa se volta para as emoções expressas nos estímulos sonoros avaliados, e não para as emoções que os informantes sentiram ao serem expostos a tais estímulos.

This study focused on emotion portrayed by music and speech stimuli, as distinct from the emotion felt by the listeners. This is an important distinction since emotion recognised and emotion felt are not necessarily identical (Gabrielsson, 2002). Indeed, there is preliminary evidence that perception of emotion in music and speech differs from the emotion induced, although with no clear pattern of responses (Ilie & Thompson, 2011). (COUTINHO, DIBBEN, 2013, p. 661)

³⁰ Para mais referências acerca de emoções expressas na fala e na música e sistemas de classificação para essas emoções, remetemos a Frühholz et al. (2013, p. 3-4).

³¹ Para uma discussão detalhada acerca da natureza, dos conceitos, das caracterizações teóricas e de fenômenos relativos a emoções, inclusive um histórico das diversas nomenclaturas empregadas pelos autores ao longo dos séculos ao estudar e pensar sobre as emoções, remetemos a ENGELMANN (1978), capítulos 1 e 2.

2.4 CONSIDERAÇÕES FINAIS

Ao longo desta seção, apresentamos os conceitos linguísticos e musicais que nos serviram de base para a elaboração e a execução do experimento e posterior análise dos dados nele obtidos. Tais conceitos serão retomados ao longo das seções de metodologia, resultados e análises e discussão.

3. METODOLOGIA

O trabalho de pesquisa aqui apresentado tem como elemento central um teste de percepção, no qual foi pedido aos informantes que avaliassem o conteúdo emocional expresso em gravações de enunciados sendo declamados e/ou cantados. A presente seção detalha a metodologia de elaboração e aplicação do experimento e de análise de dados.

3.1 PREPARAÇÃO E GRAVAÇÕES

Como base para as etapas posteriores ao experimento, dois textos foram escolhidos. Pensou-se que era importante que os textos fossem inéditos, a fim de controlar algumas variáveis: era importante evitar que os informantes tivessem memórias afetivas relacionadas aos textos, o que poderia comprometer os julgamentos e a confiabilidade dos resultados obtidos nessa pesquisa. Outro aspecto importante a ser pensado era a própria estrutura interna dos textos. Considerando que os textos deveriam ser não apenas declamados, mas também musicados e, posteriormente, cantados, optou-se pela utilização de dois poemas, uma vez que a estrutura típica desse tipo de texto se presta facilmente à musicalização.

Textos de vários alunos de graduação e pós-graduação da Faculdade de Ciências e Letras de Araraquara foram lidos e cogitados, e um deles foi escolhido para ser usado neste trabalho. O poema intitulado *Arritmia*, de autoria de Fábio Gerônimo Mota Diniz, foi escolhido como o primeiro texto. A partir desse texto, foi composta uma melodia em mi menor. O caso do segundo texto difere do primeiro no sentido em que a melodia precede as palavras. O segundo texto foi escrito como a letra para uma melodia original em lá menor, já existente. Essa letra é livremente baseada em poemas do norte-americano E. E. Cummings, que não intitulava seus poemas. Nas suas obras publicadas, os poemas de Cummings costumam ser identificados por números e pelo primeiro verso. Reproduzimos, a seguir, os poemas que inspiraram a letra da segunda composição.

III

as is the sea marvelous
 from god's
 hands which sent her forth
 to sleep upon the world

and the earth withers
 the moon crumbles
 one by one
 stars flutter into dust

but the sea
 does not change
 and she goes forth out of hands and
 she returns into hands

and is with sleep

love,
 the breaking
 of your
 soul
 upon
 my lips
 (CUMMINGS, 1954, p. 33)

IV

if i believe
 in death be sure
 of this
 it is

because you have loved me,
 moon and sunset
 stars and flowers
 gold crescendo and silver muting

of searides
 i trusted not,
 one night
 when in my fingers
 dropped your shining body
 when my heart
 sang between your perfect
 breasts

darkness and beauty of stars
 was on my mouth petals danced
 against my eyes
 and down

the singing reaches of
 my soul

As gravações referentes a esse segundo texto, portanto, serão doravante referidas simplesmente como *Cummings*, de forma a identificar o autor do texto original que inspirou a composição.

Os textos que foram declamados e musicados para utilização nesse estudo são apresentados a seguir.

(3.1)

Texto ° 01 – Arritmia

Meu coração dança
Nos intervalos da razão
E cai, pois não aprende
A coreografia

Meu coração dança...

Meu coração é dançarino
Desengonçado e divino
De vinho, fossa e de lua
Dançando em becos pelas ruas
Mas cai, pois não aprende
E cai, pois não aprende
Não, ele nunca aprende
A coreografia

Dançando à noite pelos bares
Dançando em todos os lugares
Meu coração transforma o tropeço
Num sobressalto
Num recomeço

Mesmo num samba
De despedida
Meu coração dança a vida

(3.2)

Texto n° 02 – Cummings

Vindo das cinzas
Um perfume me traz aqui
Corpos, ossos beijo o escuro da dor
Põe-se o sol
Dorme o mar
A lua a ruir

Se desfaz
 O amor se acende
 Contra os olhos, rosto molhado
 Me embala no seu seio
 Mergulhe os dedos do seu corpo em mim
 Alta a maré em seus alcances sombrios
 O aroma das noites vem para brilhar
 Pétalas dançando no meu olhar
 As mãos do oceano vêm me entregar
 Seus braços me acariciar
 Suave maestria apressa o meu pulso
 Firme suavidade da carne
 Sempre meu guardado mistério
 Junto ao sopro das estrelas
 Suas asas vão me levar

É importante notar que, em ambos os casos, buscamos certa “inespecificidade” emocional: nossa intenção era de que os textos pudessem ser interpretados de várias maneiras diferentes. Evitamos, em ambos, palavras como “tristeza” e “alegria”, que poderiam guiar a interpretação dos informantes. Em *Arritmia*, o texto em si não permite ter certeza, por exemplo, se o enunciador está lamentando o “comportamento” de seu coração ou apenas comentando esse comportamento (como quem comenta, achando graça, as travessuras de uma criança). Em *Cummings*, palavras que podem evocar sentidos macabros, como “ossos”, “dor” e “escuro”, são justapostas a “acende”, “perfume”, “suave” e “estrelas”, de carga semântica mais positiva. Tal escolha estilística teve o intuito de impedir que o texto em si fosse definitivo para a interpretação dos informantes, aumentando a importância da melodia (e da entoação na gravação declamada).

As melodias foram gravadas em estúdio, a fim de garantir a qualidade técnica da gravação. Cada composição foi cantada por um vocalista diferente, *Arritmia* por um profissional, *Cummings* por um amador³². Uma terceira pessoa declamou as letras das composições em duas outras gravações, também realizadas em estúdio. Fez-se necessário, então, decidir se as melodias cantadas teriam acompanhamento instrumental, já que a tonalidade, na música ocidental, está relacionada à harmonia, e a harmonia, por sua vez, se refere à relação entre sons executados simultaneamente. Apesar de sua natureza

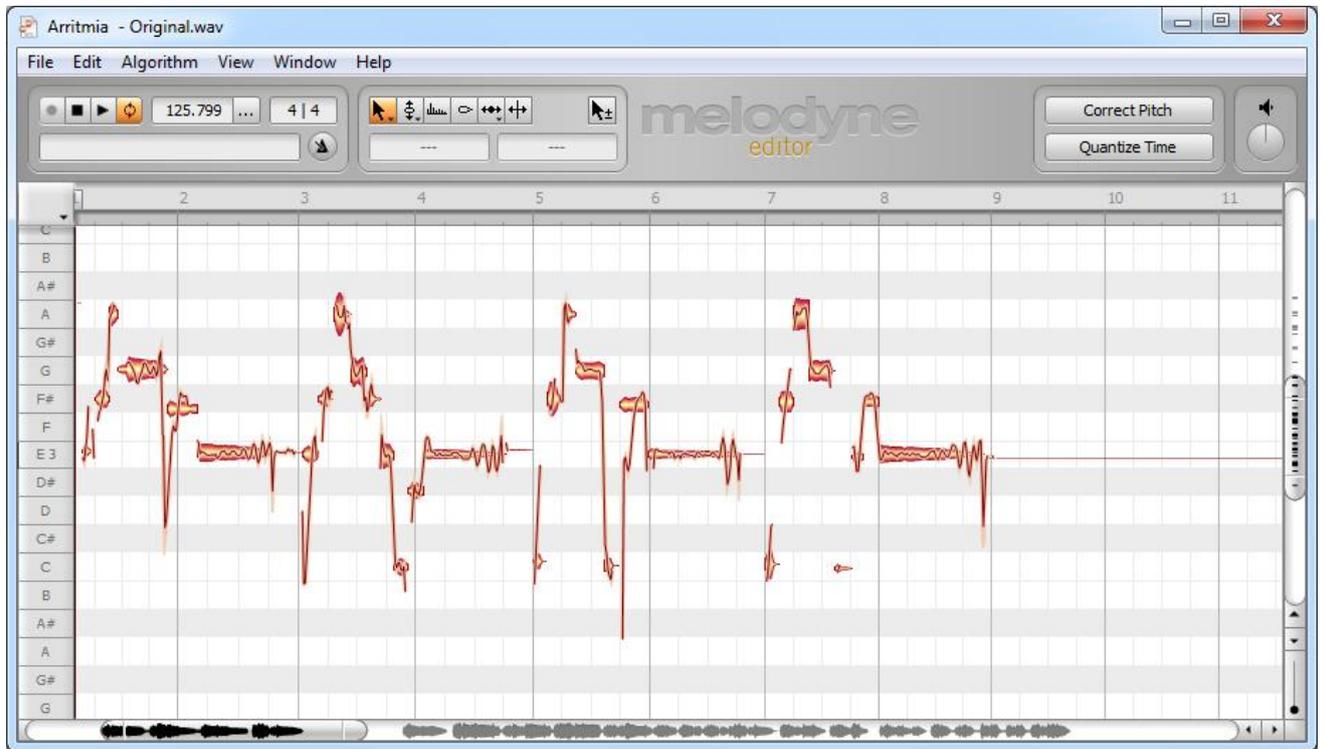
³² Originalmente, foi feita uma gravação de *Cummings* sendo cantada pelo vocalista de *Arritmia*. No entanto, o vocalista teve dificuldades com algumas passagens e o resultado da gravação não foi tão bom quanto o esperado. Posteriormente, foi realizada uma nova gravação da composição, com o vocalista amador. Infelizmente, este também teve dificuldades, e subtonou várias vezes ao longo da música. Para não estourar os prazos pré-estabelecidos, foi necessário escolher uma das gravações obtidas. Após consultar a orientadora, o autor deste trabalho recorreu à professora doutora Cristina Martings Fargetti, que aconselhou o uso da gravação do segundo vocalista no experimento.

multidisciplinar, o presente trabalho é, em seu âmago, linguístico. Considerando, também, as experiências nos quais as suspeitas que levaram a esse trabalho se basearam (CURTIS; BHARUCHA, 2010, bem como SCHREUDER, 2006), nas quais foi verificada uma correlação entre o uso de intervalos melódicos na música e na fala, optou-se por não utilizar qualquer acompanhamento instrumental. Não obstante, quis-se evitar a chance de que os resultados aqui obtidos pudessem ser atribuídos exclusivamente a aspectos puramente musicais, como a linha de um instrumento no acompanhamento. Tomou-se, assim, a decisão de que as composições seriam gravadas *a capella*, ou seja, somente voz, sem acompanhamento instrumental. É certo que a maioria esmagadora das pessoas não costuma ouvir música sem acompanhamento harmônico. A possível estranheza advinda da exposição a essas composições fora do comum foi fonte genuína de preocupação durante esta pesquisa. De fato, acompanhamentos percussivos foram criados através de sintetizadores digitais e mixados com as melodias vocais, no intuito de reduzir esse estranhamento. Os instrumentos percussivos foram uma alternativa a instrumentos como piano e violão, por não apresentarem timbre definido, de modo que qualquer variação observada nos resultados deveria, provavelmente, ser atribuída à melodia vocal, e não ao acompanhamento. O resultado dessa mixagem, no entanto, não foi satisfatório: observou-se que, enquanto as gravações a capella não geravam tanto estranhamento quanto se poderia pensar à primeira vista, os acompanhamentos percussivos acabavam sendo uma distração da melodia e da letra da música. A decisão mais sensata, então, pareceu ser a de descartar os acompanhamentos percussivos e utilizar as gravações originais, apenas com a voz.

As gravações das duas melodias foram, então, editadas utilizando o software Melodyne editor, versão 2.1.2.. A interface desse programa “enquadra” as ondas de um sinal sonoro em notas musicais, de acordo com as frequências fundamentais desse sinal, de modo que é possível, assim, visualizar claramente quais trechos da gravação correspondem a quais notas musicais. O programa também permite alterar a altura de trechos específicos do som e, conseqüentemente, a nota musical. É possível, por exemplo, selecionar trechos cuja frequência fundamental corresponda à da nota Dó e aumentar a altura desses trechos, transformando a nota Dó em Ré. No caso da primeira composição, *Arritmia*, as notas Sol foram aumentadas em um semitom; no caso da segunda composição, *Cummings*, as notas Dó, Fá e Sol foram aumentadas em um semitom. Dessa forma, nos dois casos, buscou-se emular uma melodia composta em modo maior. Em outras palavras, *Arritmia*, originalmente escrita

em Mi menor, foi “transposta” para Ré maior, e *Cummings*, originalmente em Lá menor, foi “transposta” para Lá maior³³.

Figura 2: Janela do programa Melodyne Editor



Ao final do processo de edição das gravações no software Melodyne, dispúnhamos de seis gravações, sendo três versões diferentes de cada texto: Arritmia na versão declamada, Arritmia musicada em modo menor e Arritmia musicada em modo maior; Cummings declamada, Cummings musicada em modo maior e Cummings musicada em modo menor.

3.1.2 PARTITURA DAS COMPOSIÇÕES E ESPECTROGRAMAS DAS GRAVAÇÕES

Para que se tenha mais clareza acerca dos estímulos aos quais os participantes do experimento foram expostos, seguem-se as partituras e os espectrogramas das gravações, divididos por frases. Em cada composição, as figuras mostrarão, respectivamente:

³³ O uso de aspas aqui se deve ao fato de que não foi feita uma transposição propriamente dita, apenas alteraram-se digitalmente os intervalos na expectativa de que isso acarretasse uma mudança na percepção da tonalidade das composições por parte dos informantes. Para detalhes sobre a efetividade desse método, ver a seção de Discussão dos resultados.

- 1-) A partitura original da canção, utilizada pelos vocalistas na hora da gravação em estúdio,
- 2-) O espectrograma da gravação em áudio alterada, com intervalo de terça maior,
- 3-) O espectrograma da gravação de áudio original, em modo menor,
- 4-) O espectrograma da gravação de áudio com o texto sendo declamado.

Composição nº 1 – Arritmia

Arritmia foi originalmente escrita e gravada em Mi (Em) menor. Posteriormente, a gravação foi alterada digitalmente para que as notas Sol (G) se transformassem em Sol sustenido (G#), no intuito de emular uma tonalidade de Ré maior. Nas figuras abaixo, A-M indica um espectrograma da gravação alterada, A-m um espectrograma da gravação original, e A-d um espectrograma da gravação em que o texto foi declamado.

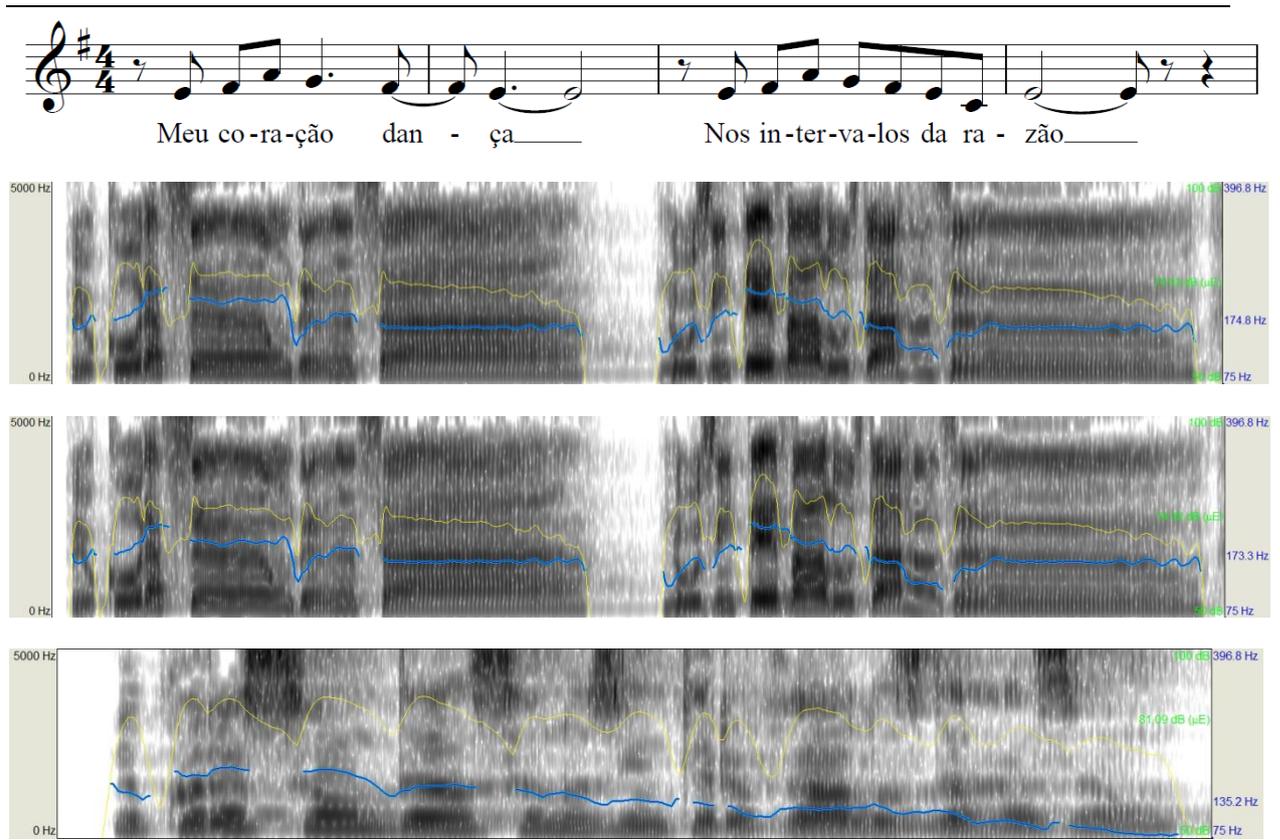
Com relação à versão declamada, é possível perceber que os versos foram tomados como “enunciados” prosódicos (no sentido de Nespor e Vogel, 1986). Assim, foram segmentados como se todos os versos fossem assertivos. Isso explica o fato de que todos os trechos segmentados, relativos às figuras 3-12, apresentarem padrão descendente, uma vez que este é o padrão esperado para enunciados afirmativos no português brasileiro (Massini-Cagliari e Cagliari, 2001).

No que se refere às gravações do texto sendo cantado, a comparação entre os espectrogramas relativos à gravação da melodia original e da alterada mostram que a diferença melódica se concentra justamente na nota alterada, uma vez que o restante da melodia se mantém, em termos de altura e de desenho melódico. A diferença entre os dois contornos, assim, é mínima.

Comparando as gravações cantadas com a declamada, é importante mencionar que o texto cantado apresenta muito mais variações de altura melódica e de intensidade do que a versão declamada. Deste modo, a menor variação de altura melódica pode ter levado os informates a interpretarem a versão declamada como mais triste ou mais calma (levanto em consideração Frühholz, Trost e Grandjean, 2014, p. 10)

Em cada espectrograma, a linha amarela mostra a amplitude do sinal sonoro, e a linha azul representa a altura melódica.

Figura 3: “Meu coração dança / nos intervalos da razão”



Na figura 3, a diferença mais importante é que, enquanto as gravações cantadas apresentam uma pausa entre o texto dos dois primeiros compassos, a versão declamada não. O contorno de entonação descendente, típico para uma frase afirmativa no Português Brasil, também é mais evidente na versão declamada.

Na figura 4, novamente, há diferença nas pausas (depois de “aprende” nas versões musicadas, depois de “cai”, na versão declamada) e no contorno melódico – as versões musicadas se afastam do padrão descendente típico.

Já na figura 5, também vemos claramente que o padrão entoacional da versão declamada não é o mesmo das versões cantadas, ainda que, aqui, o local das pausas coincida.

Figura 4: “E cai, pois não aprende / a coreografia”

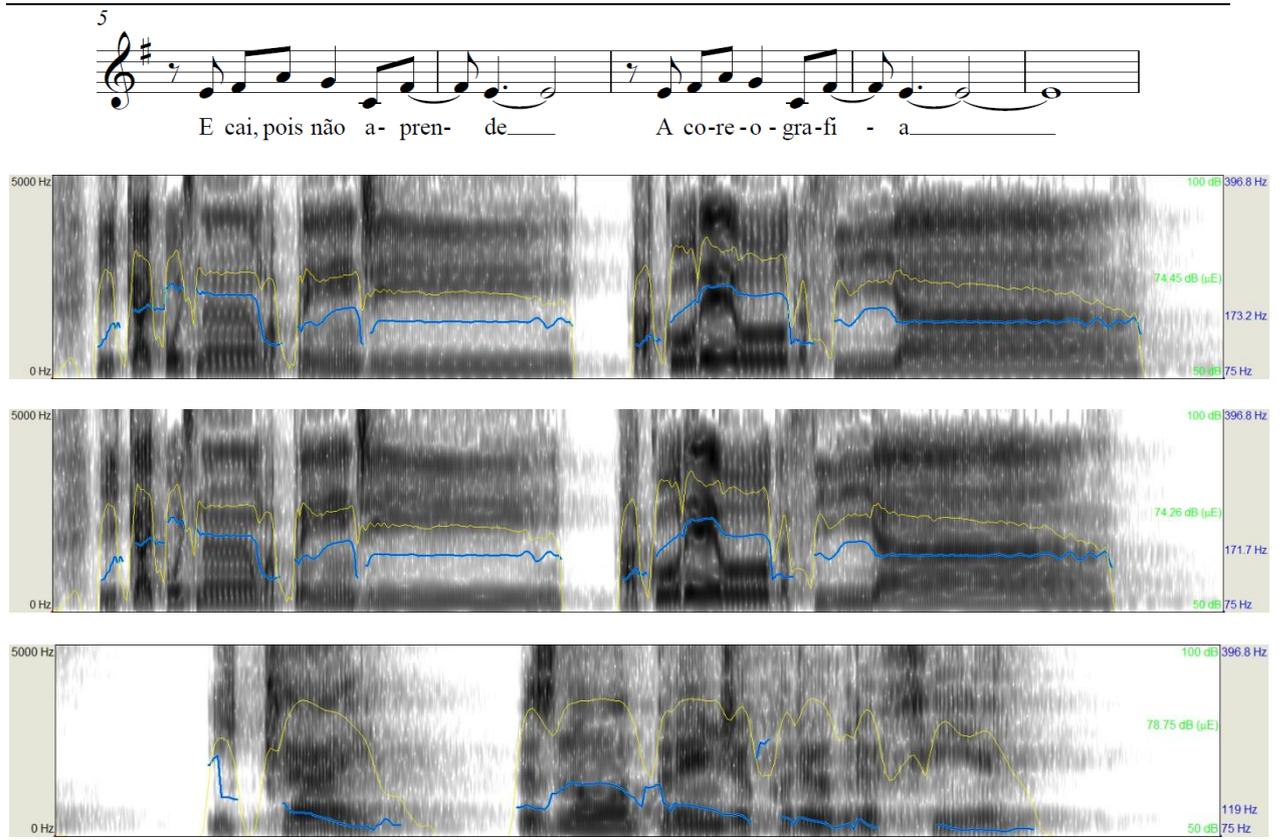


Figura 5: “Meu coração dança... / Meu coração é dançarino”

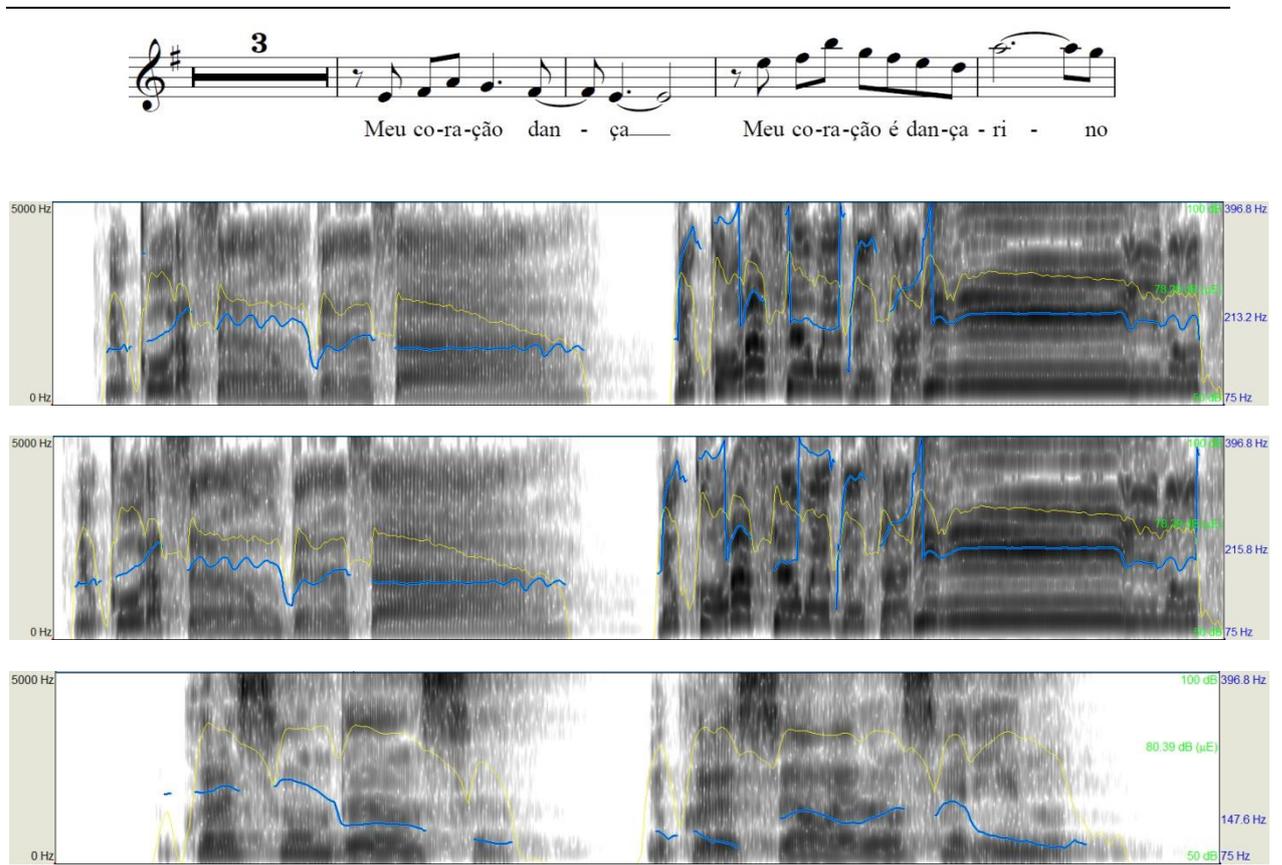


Figura 6: “Desengonçado e divino / De vinho, fossa e de lua”

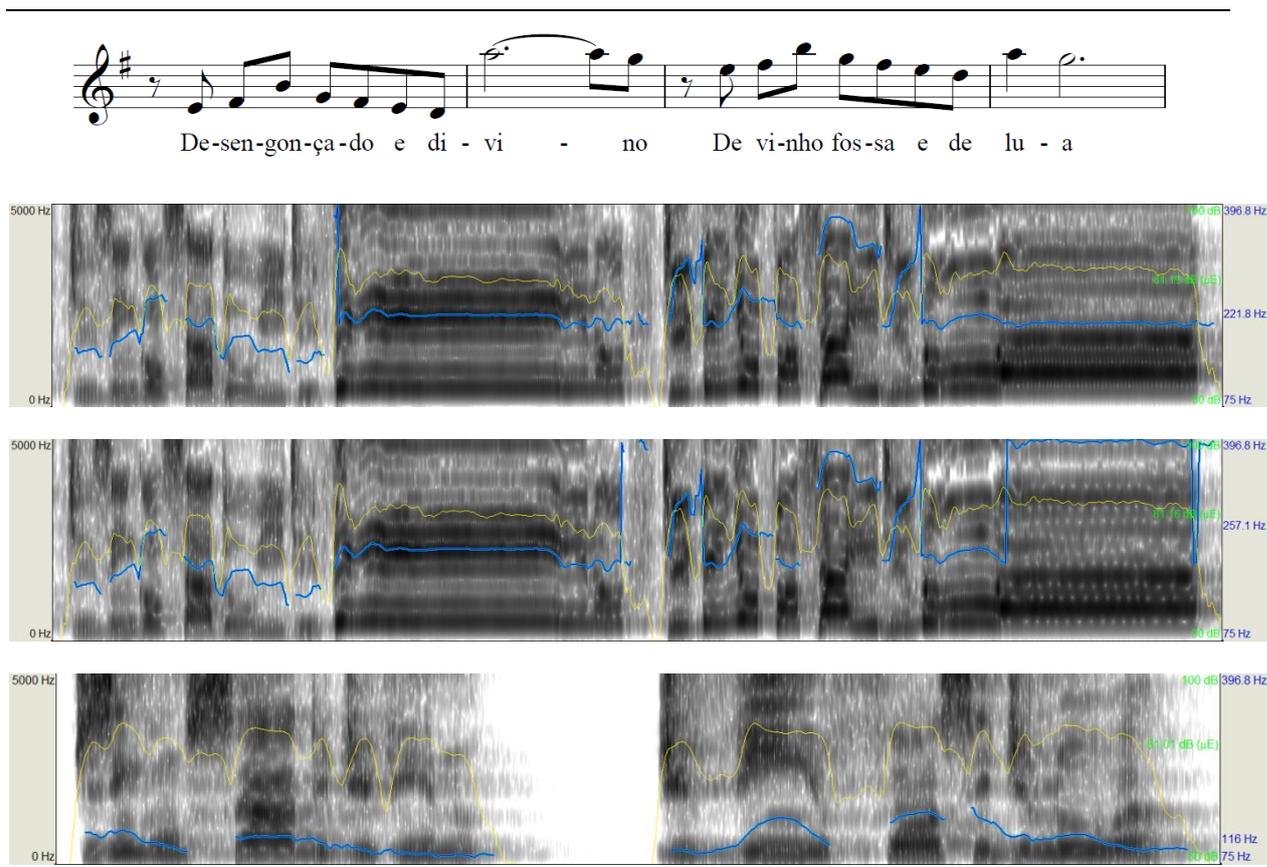


Figura 7: “Dançando em becos pelas ruas / Mas cai pois não aprende”

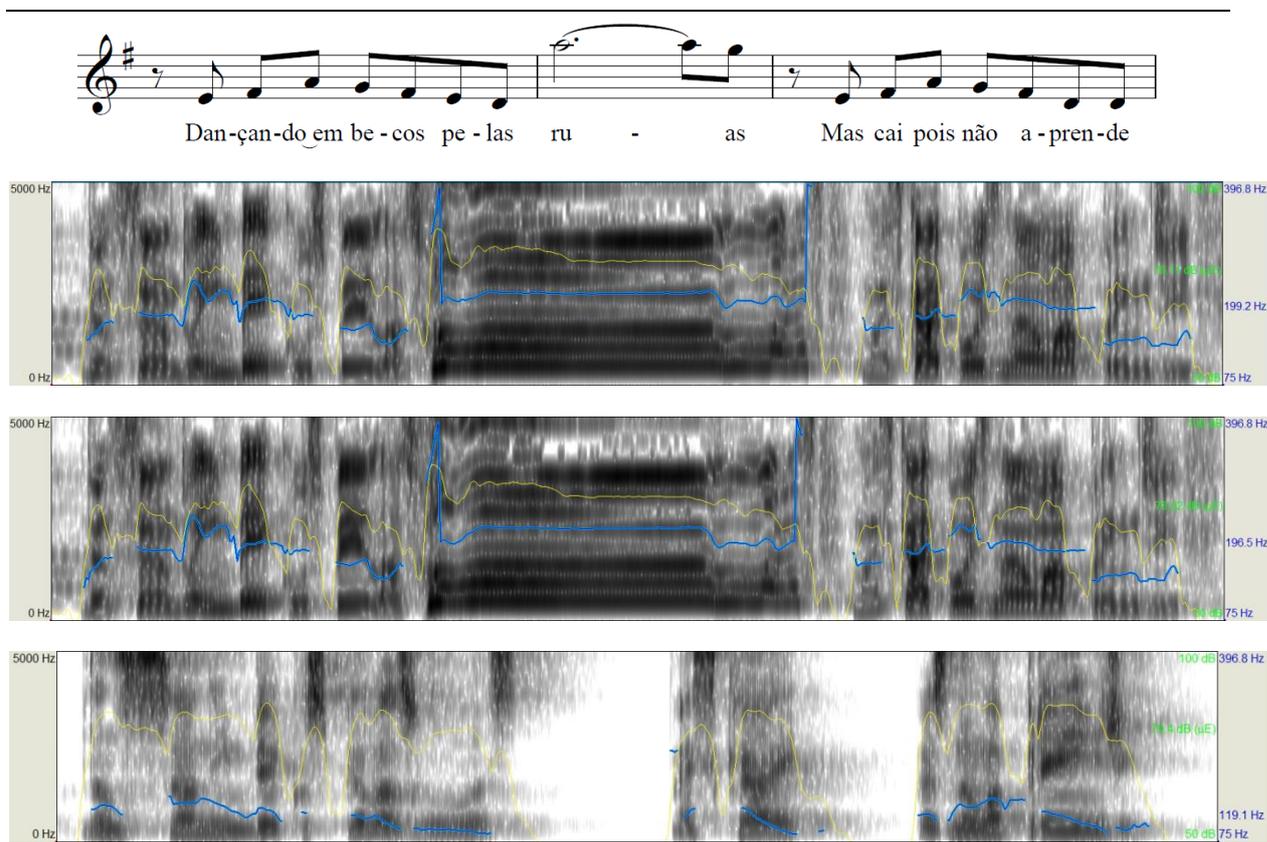


Figura 8: “E cai, pois não aprende / Não, ele nunca aprende”

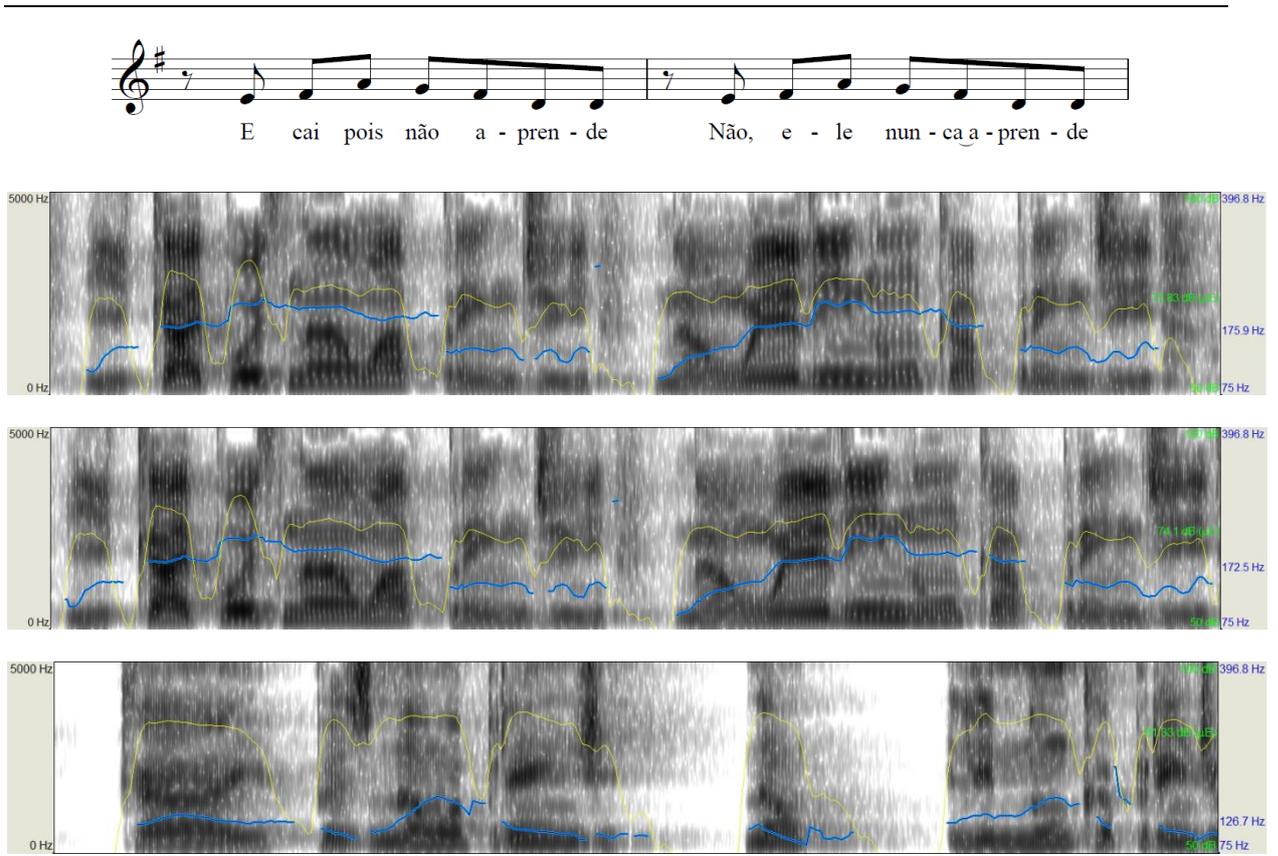


Figura 9: “A coreografia...”

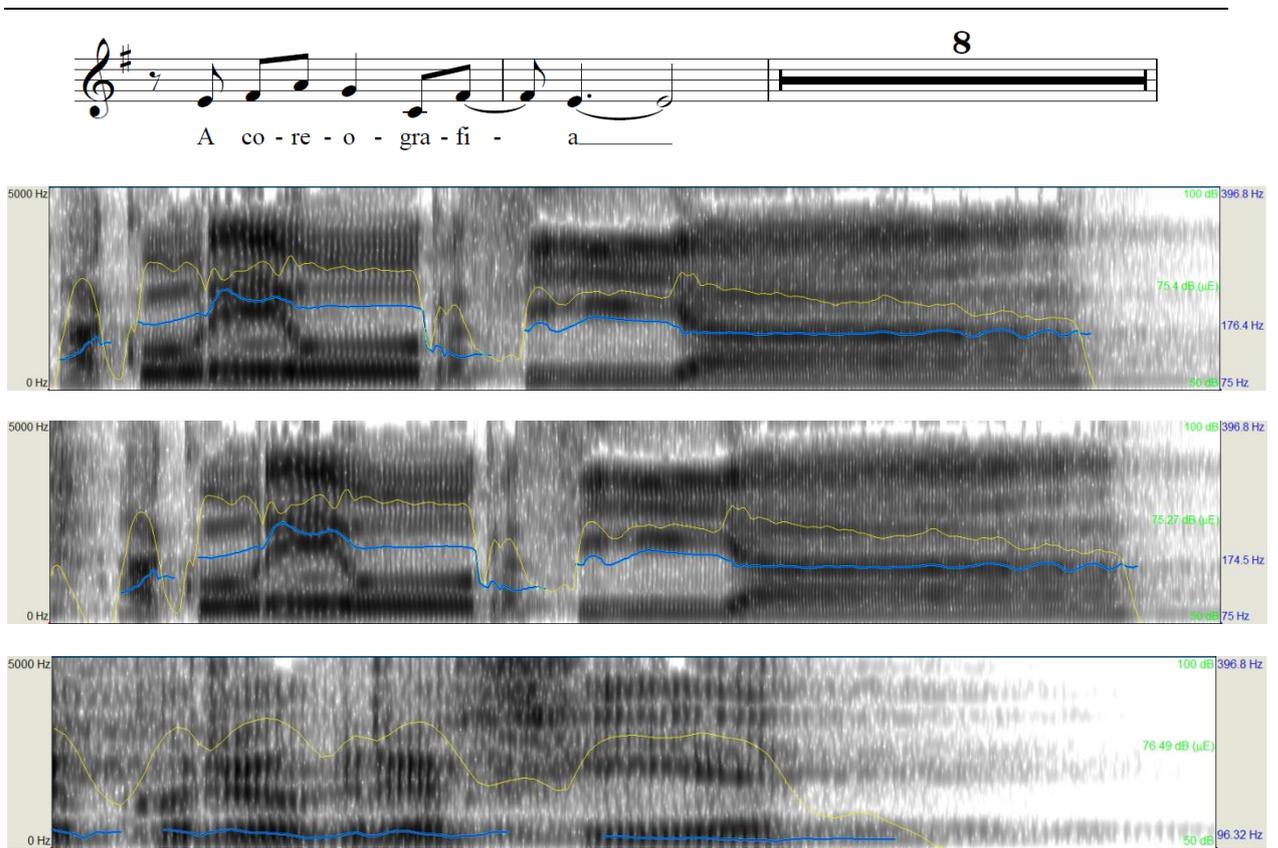


Figura 10: “Dançando à noite pelos bares / Dançando em todos os lugares”

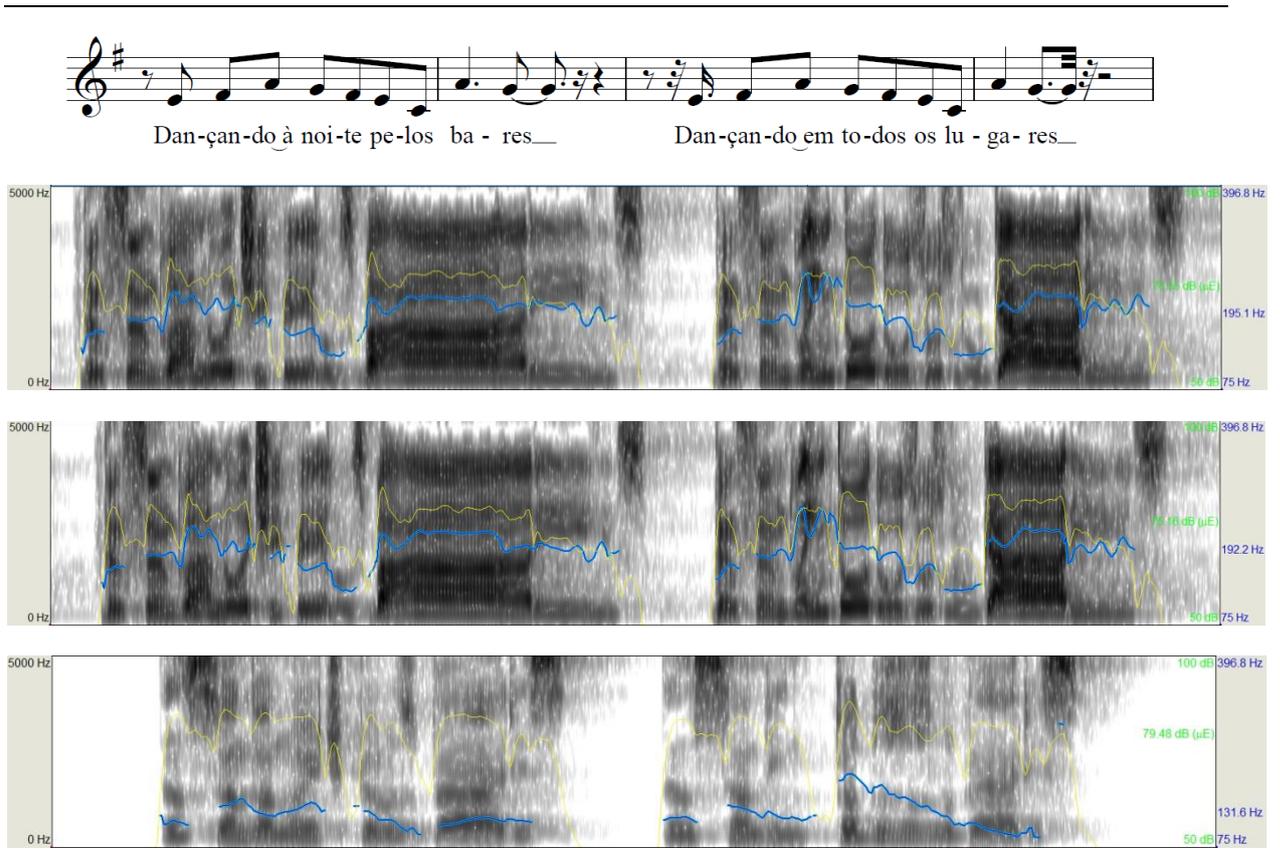


Figura 11: “Meu coração transforma o tropeço / Num sobressalto / Num recomeço”

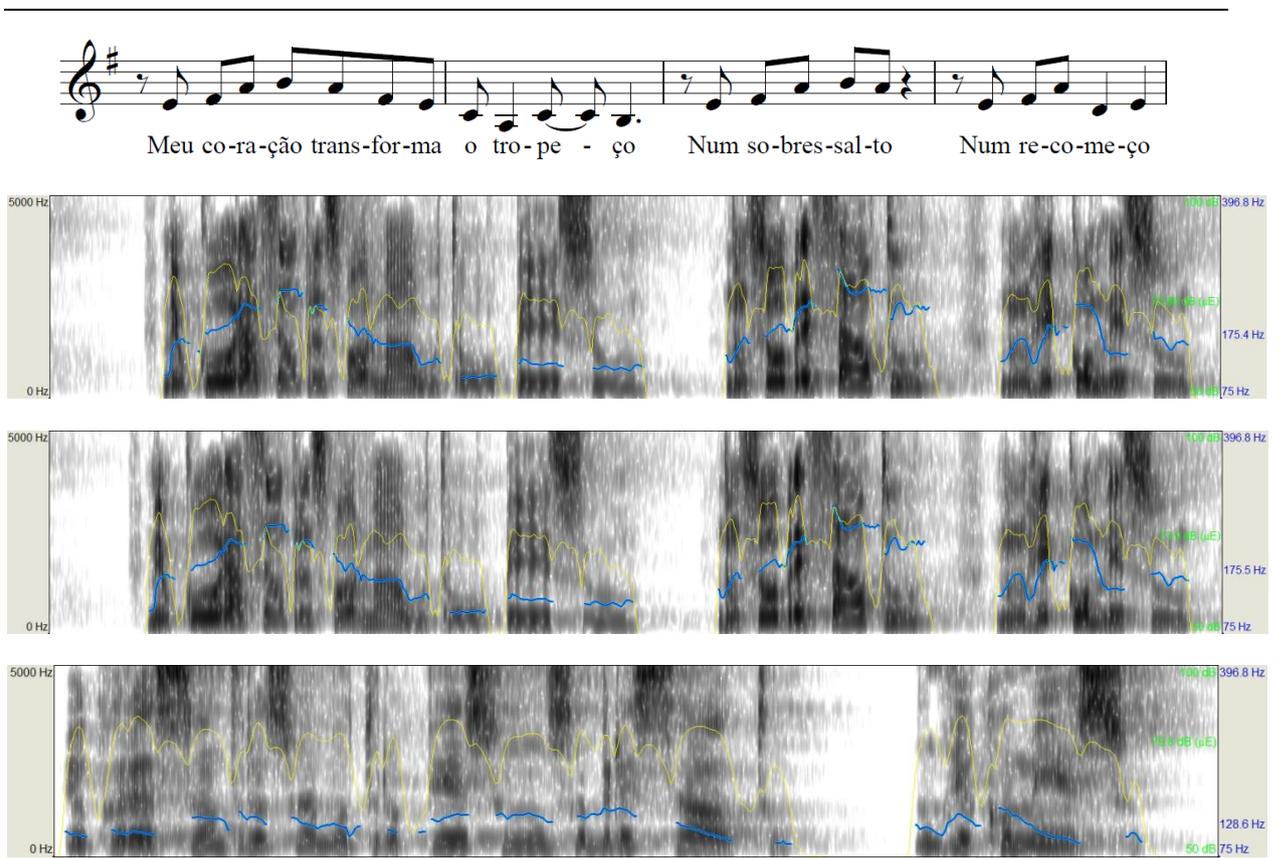
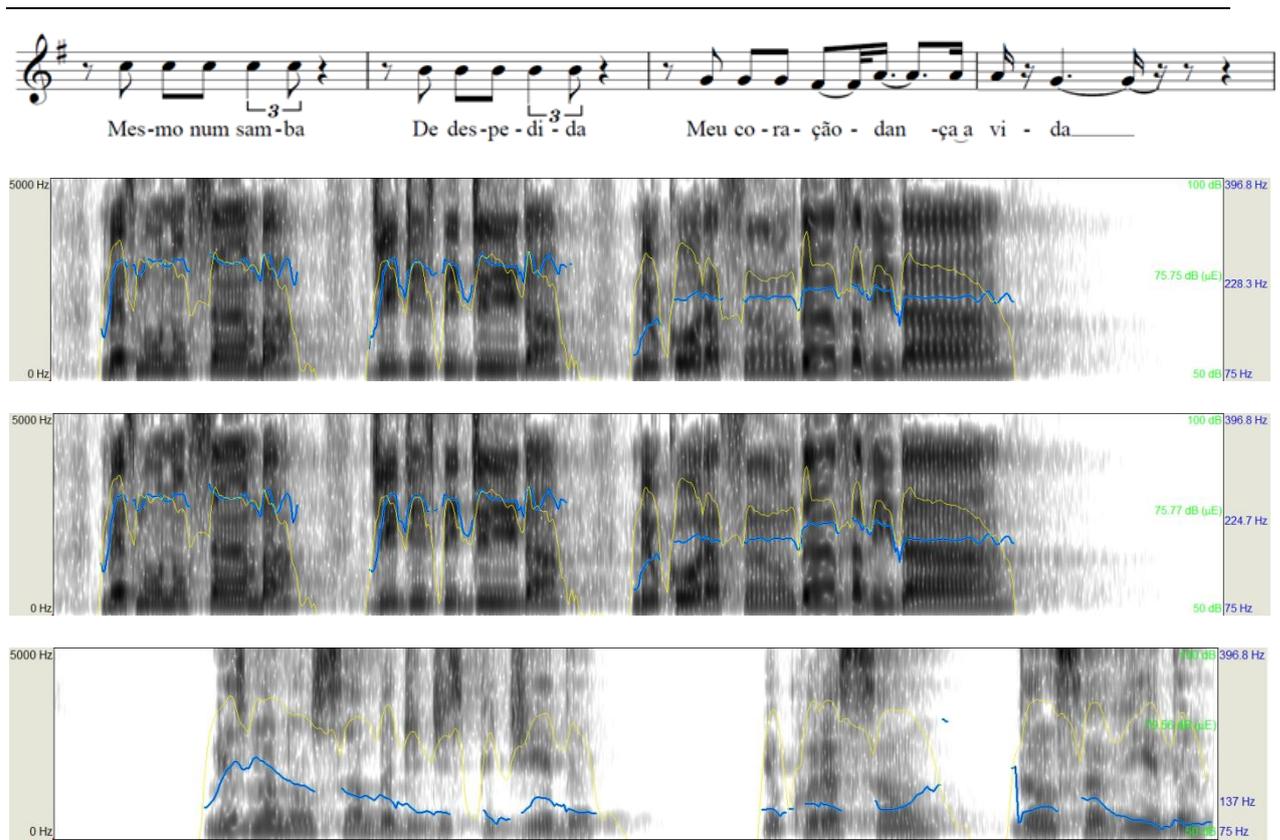


Figura 12: “Mesmo num samba / De despedida / Meu coração dança a vida”



Enquanto nas versões musicadas o texto dos três compassos é cantado em três “blocos” distintos, com uma pausa bem marcada entre cada um deles, na versão declamada são enunciados como dois: “Mesmo num samba de despedida [pausa] meu coração dança a vida”. Também há uma diferença considerável na altura melódica.

Composição nº 2: Cummings

Cummings foi originalmente escrita e gravada em Lá menor (Am). Posteriormente, a gravação foi alterada digitalmente para que as notas Fá (F), Sol (G) e Dó (C) se transformassem em Fá sustenido (F#), Sol sustenido (G#) e Dó sustenido (C#), emulando uma tonalidade de Lá maior. Nas figuras abaixo, C-M indica um espectrograma da gravação alterada, C-m um espectrograma da gravação original, e C-d um espectrograma da gravação em que o texto foi declamado. Em cada espectrograma, a linha amarela mostra a amplitude do sinal sonoro, e a linha azul representa a altura melódica.

No caso de *Cummings*, não são percebidas diferenças nos locais de pausa como em *Arritmia*, mas as diferenças no contorno melódico e prosódico também são significativas.

Mais especificamente, nas figuras 14, 17, 18, 19, 25, 27, o padrão descendente da declamação contraria o movimento da linha melódica. Nas demais figuras, os padrões do canto e da fala combinam – ascendente nas figuras 16 e 23, descendente nas demais. 6 das 21 figuras (quase um terço do total) apresenta divergência no padrão entoacional da música com a fala.

Figura 13: “Vindo das cinzas”

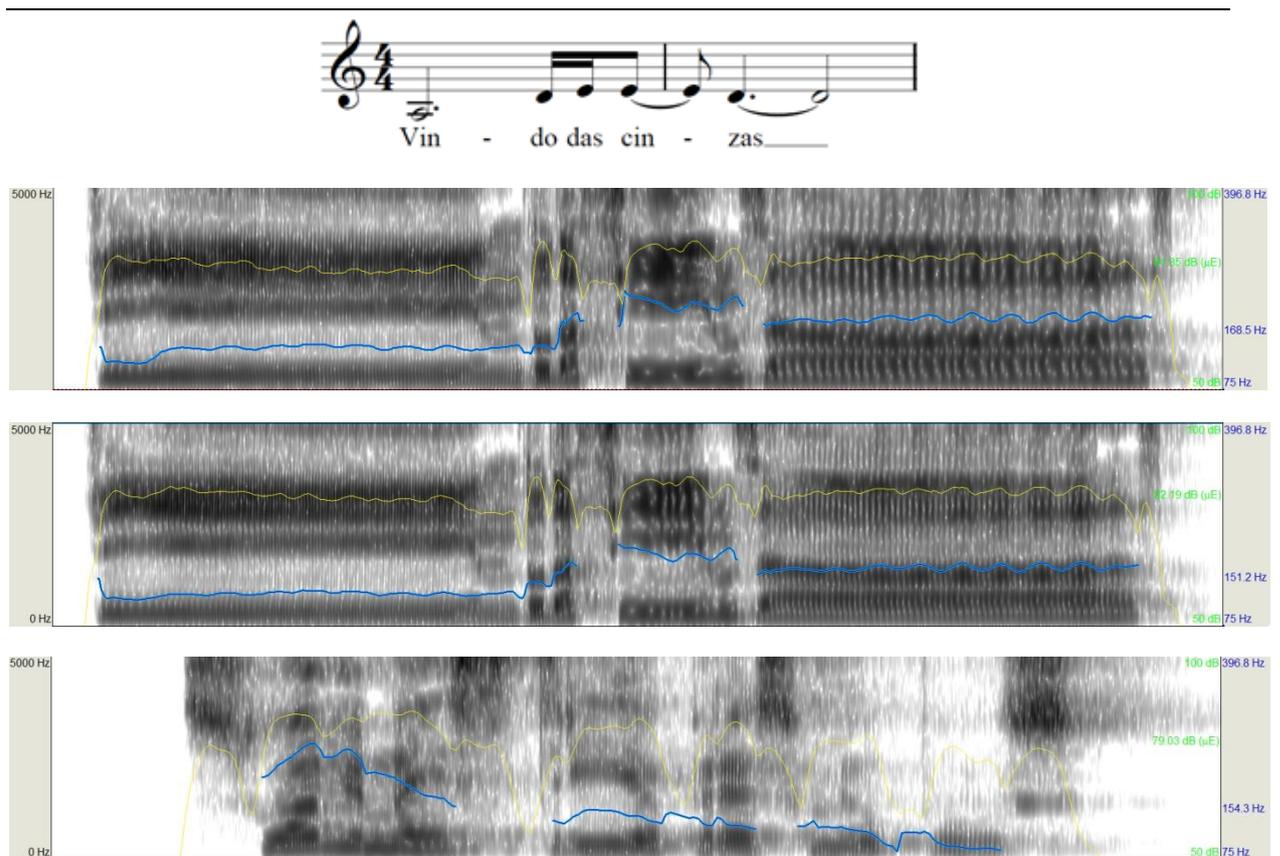


Figura 14: “Um perfume me traz aqui”

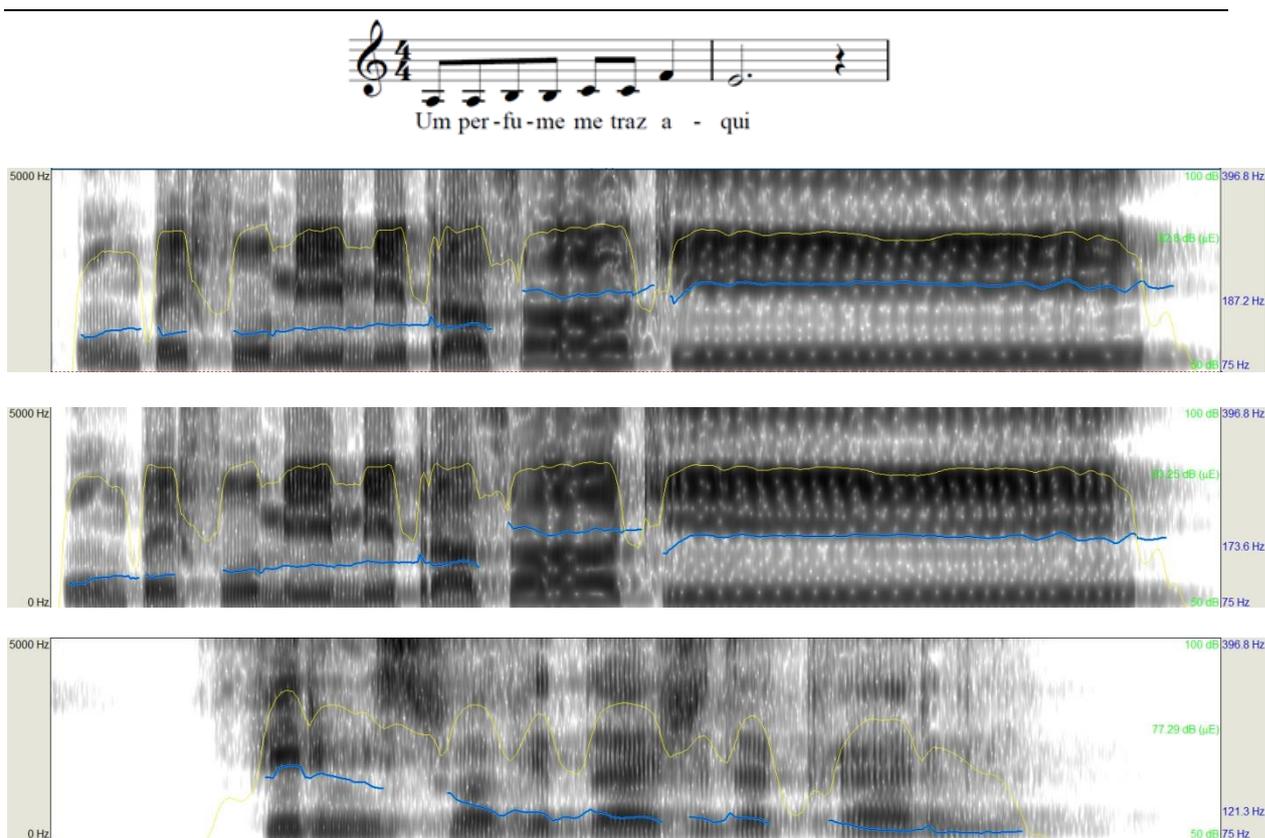


Figura 15: “Corpos, ossos, beijo o escuro da dor”

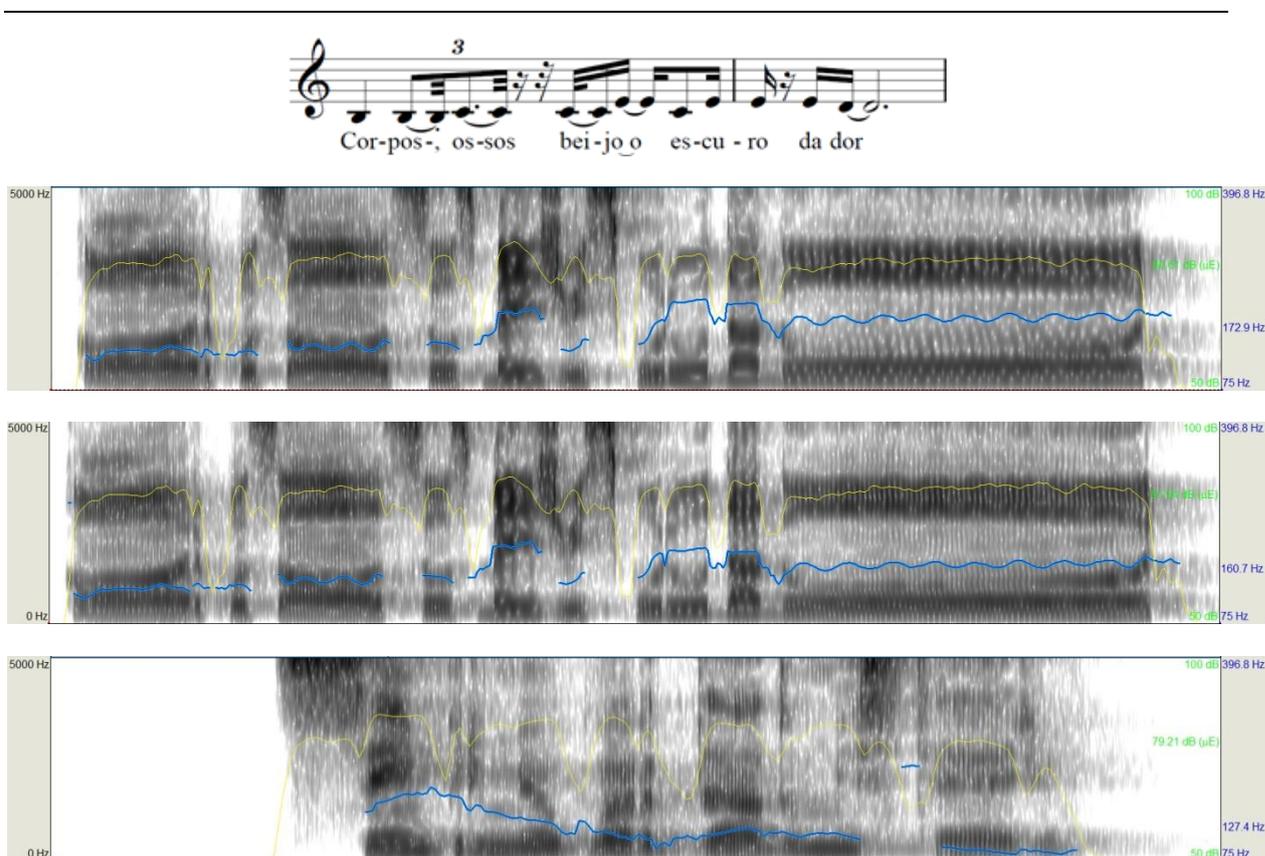


Figura 16: “Põe-se o sol”

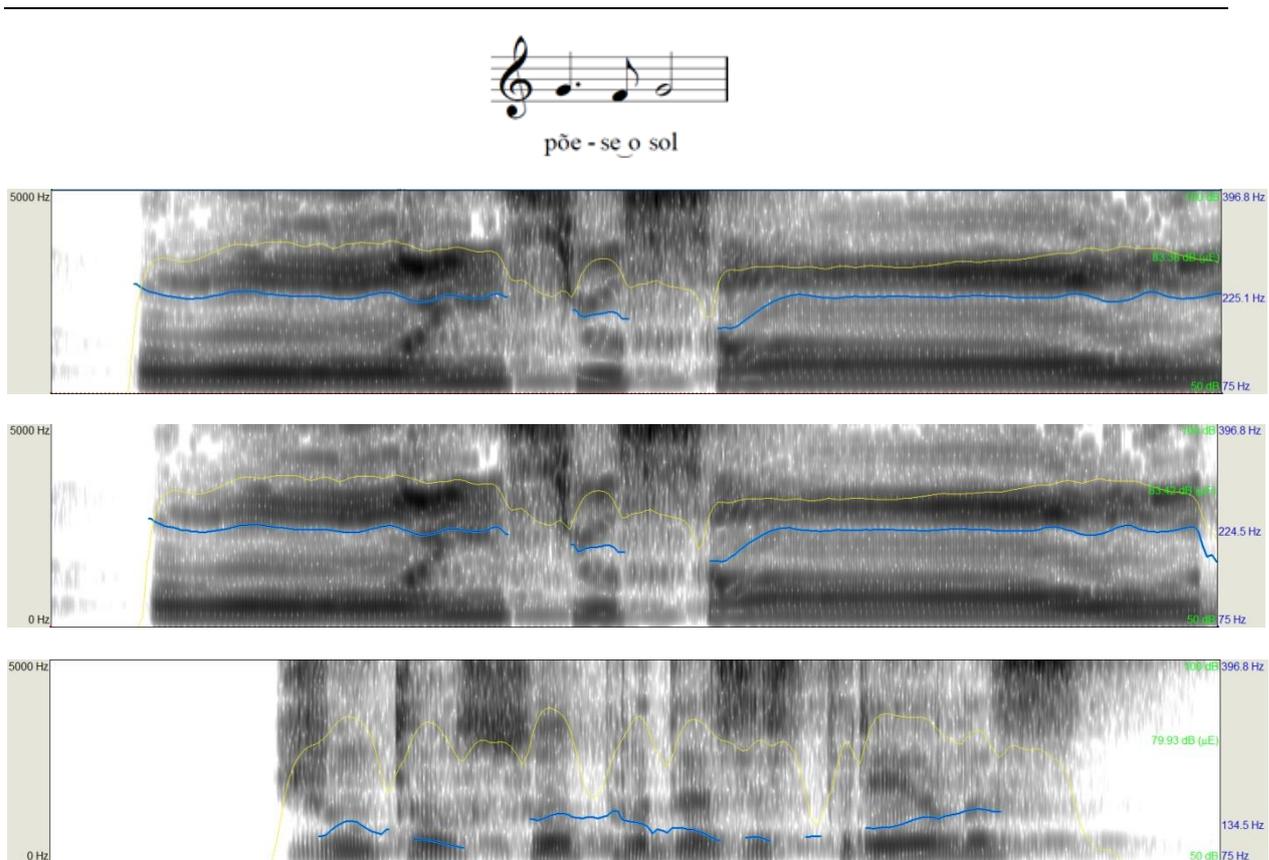


Figura 17: “Dorme o mar”

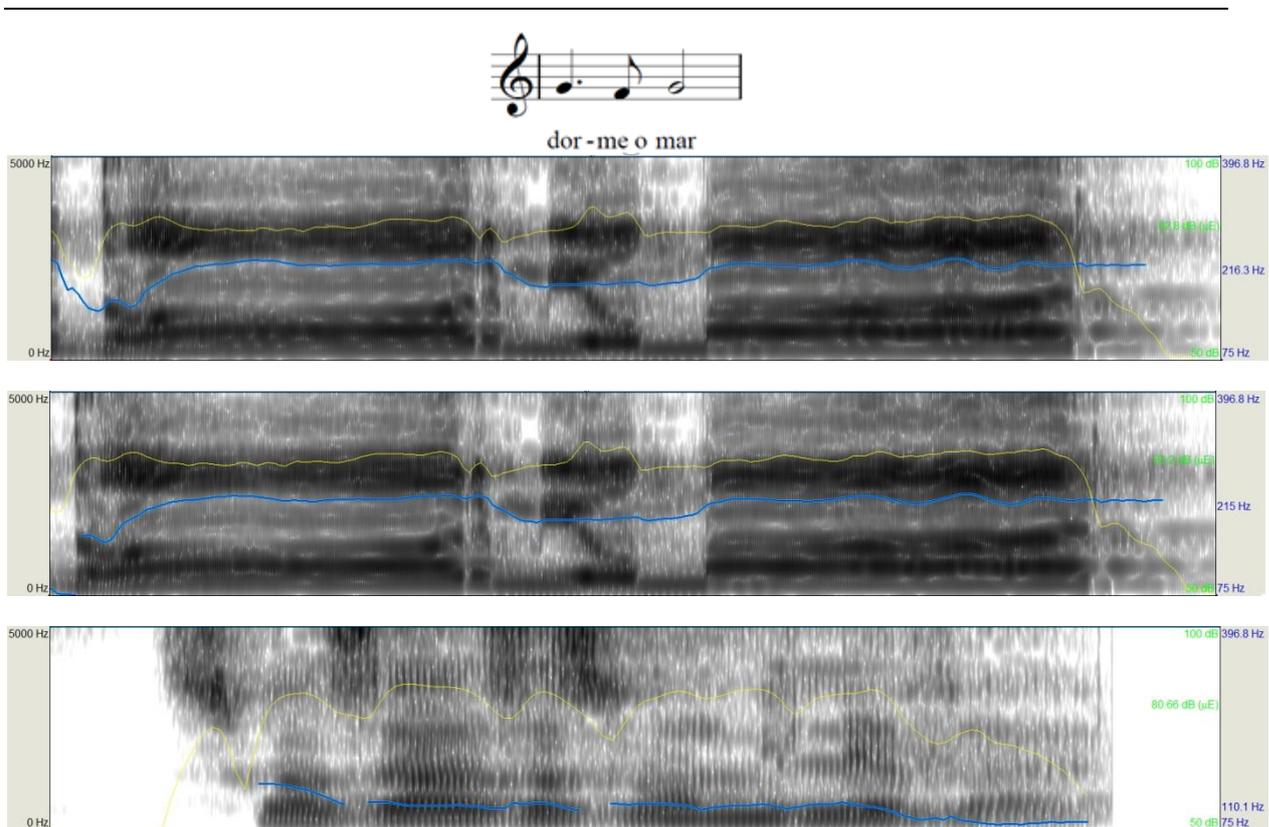


Figura 18: “A lua a ruir”

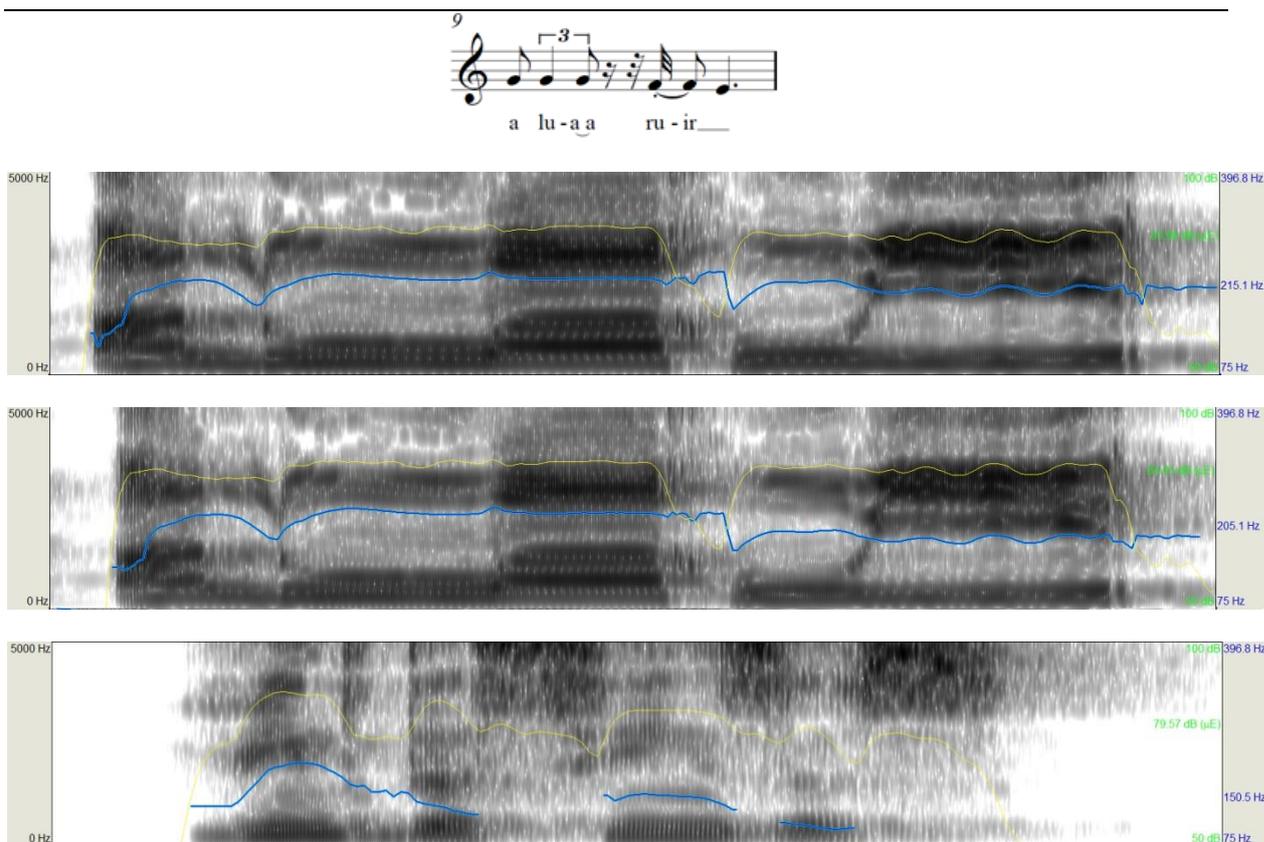


Figura 19: “Se desfaz”

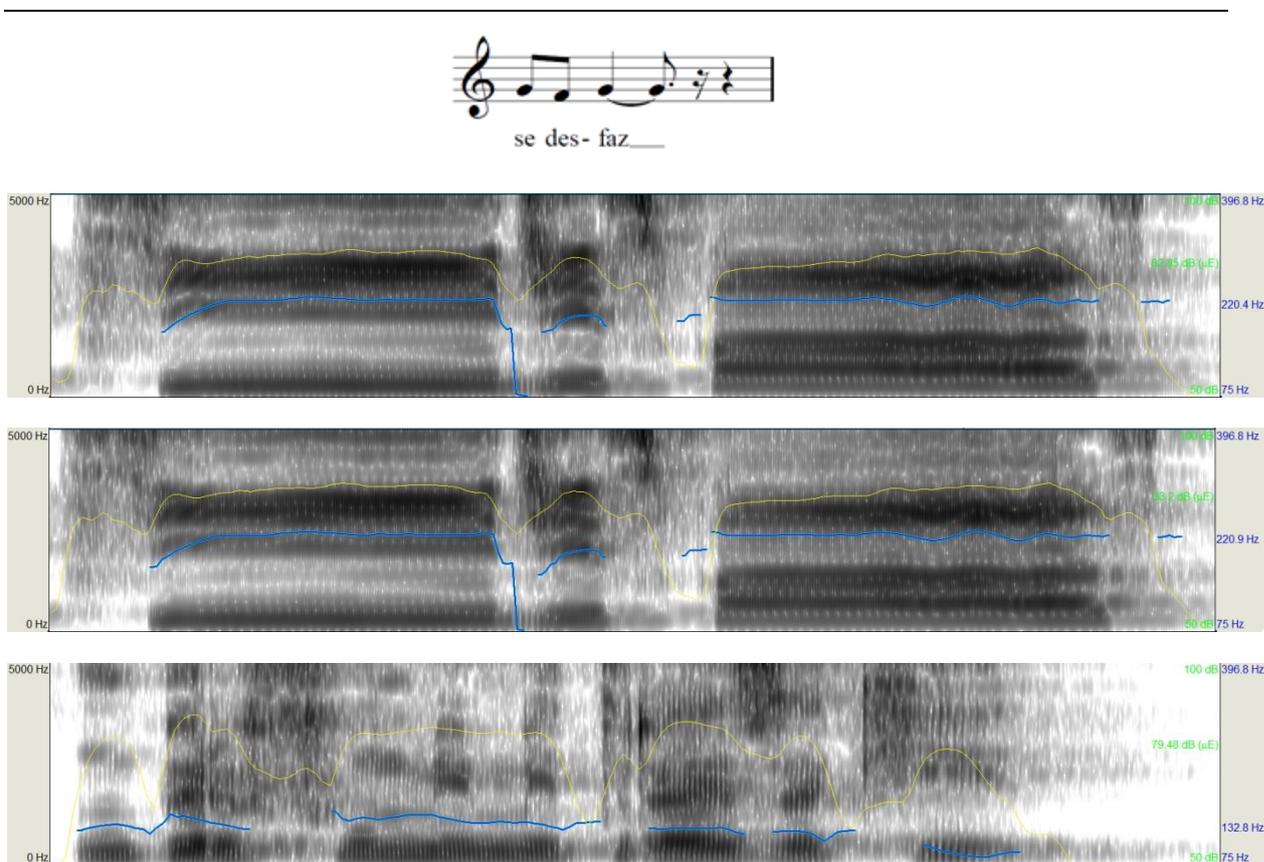


Figura 20: “O amor se acende”

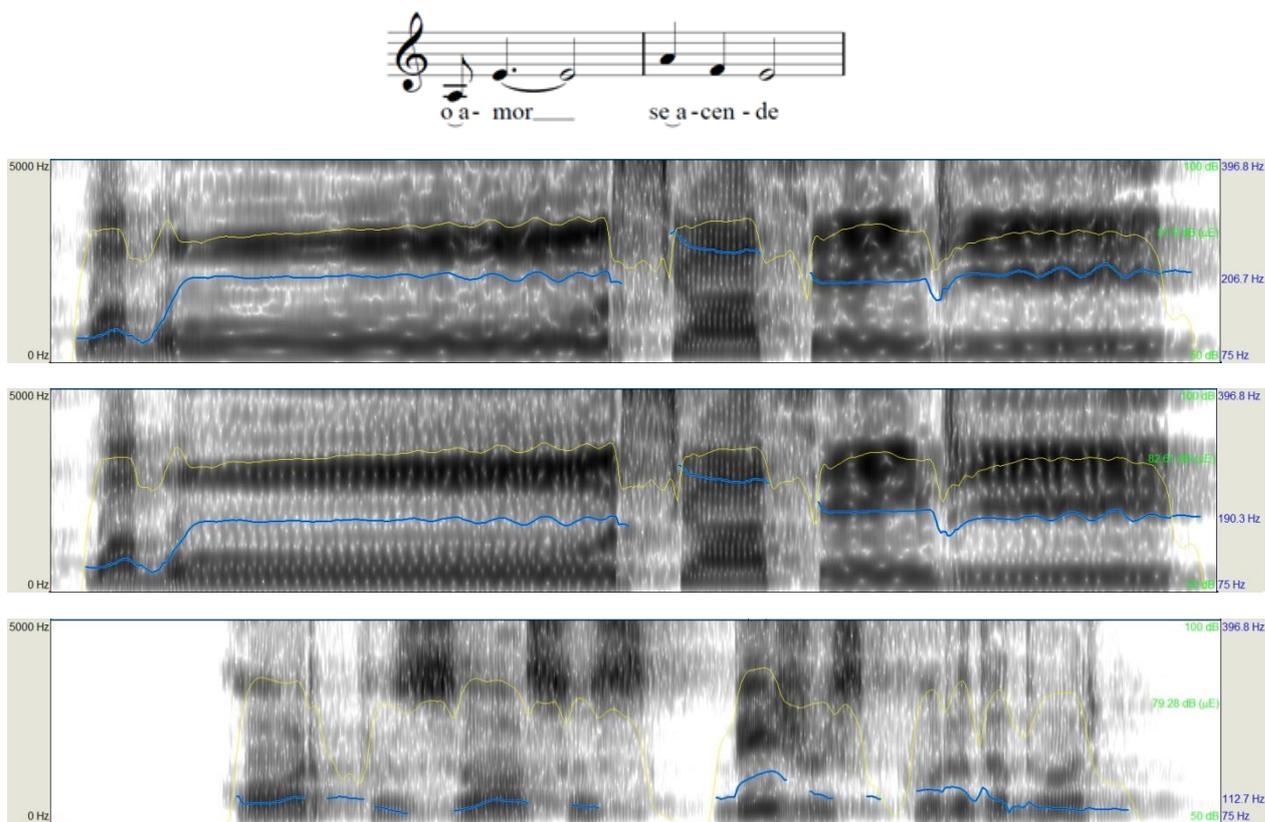


Figura 21: “Contra os olhos, rosto molhado”

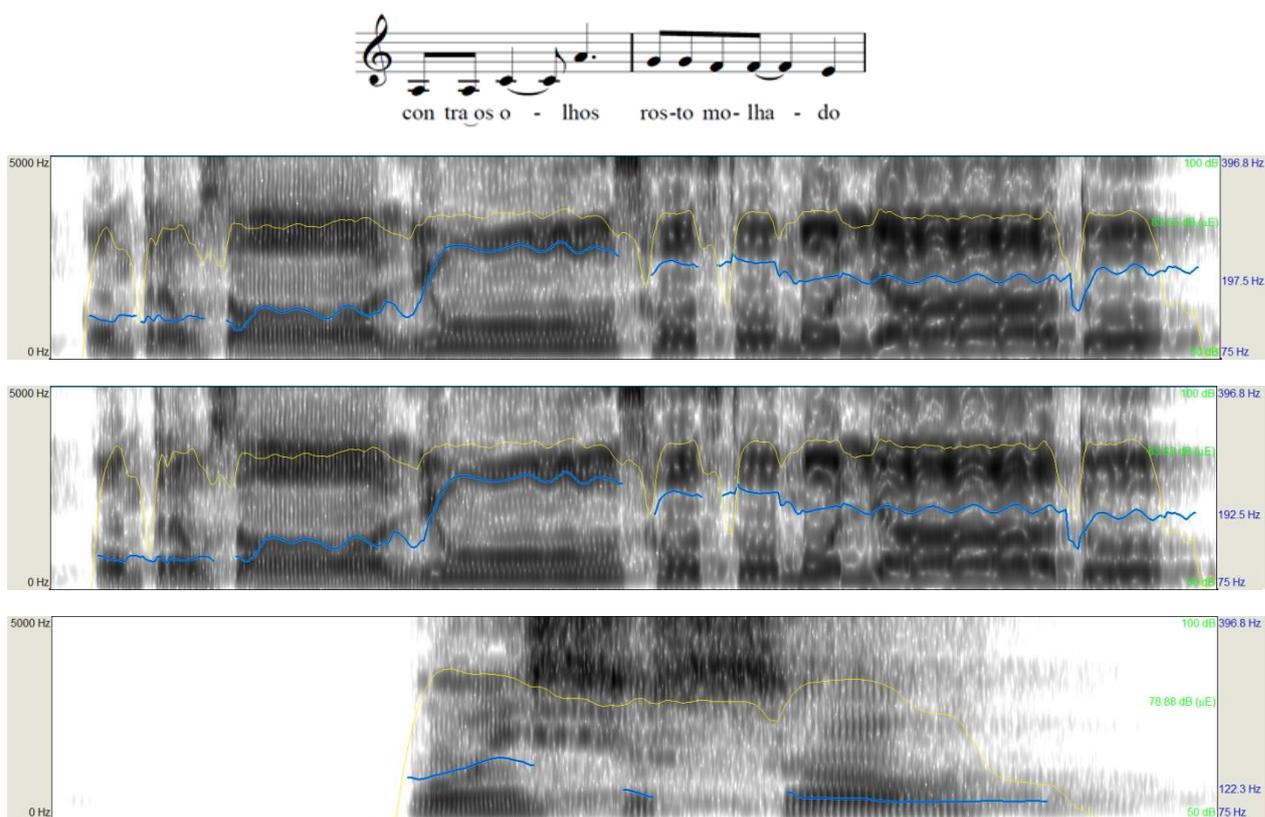


Figura 22: “Me embalada no seu seio”

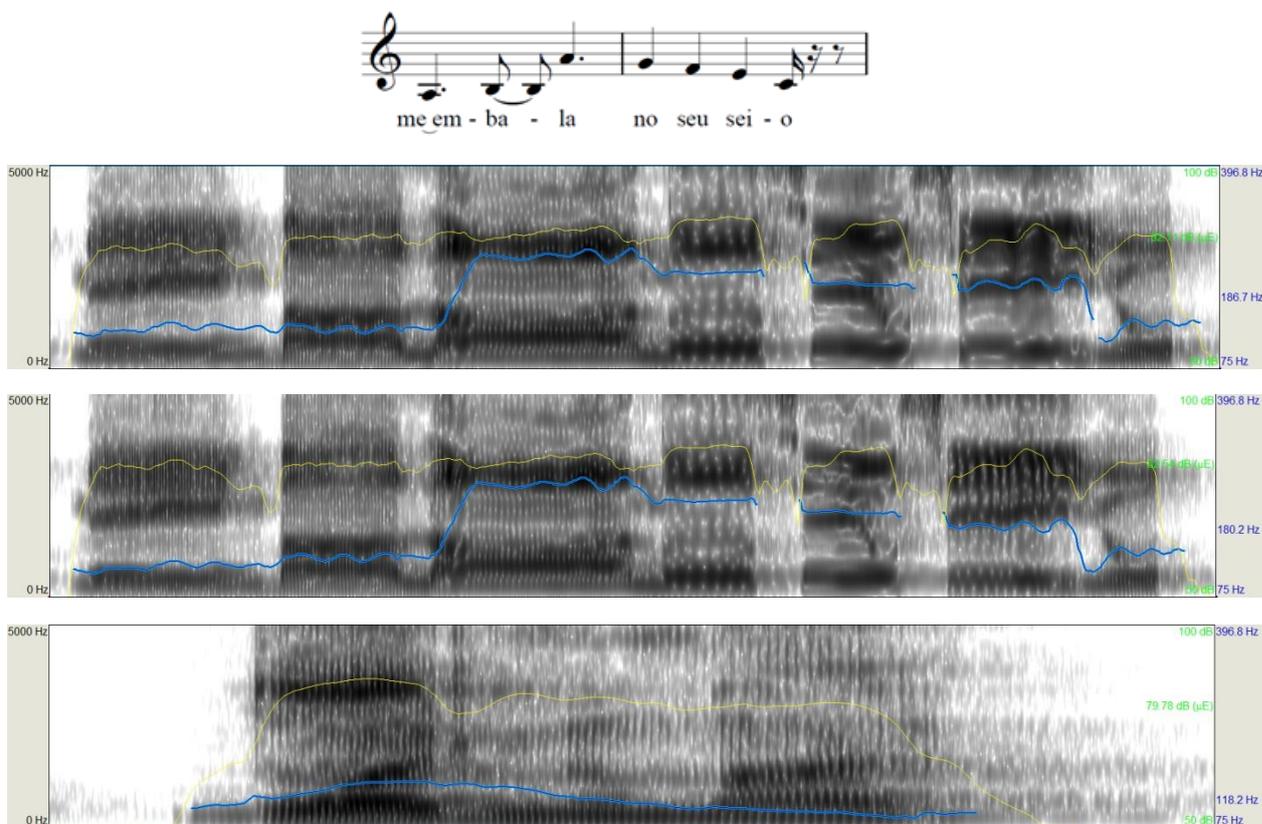


Figura 23: “Mergulhe os dedos do seu corpo em mim”

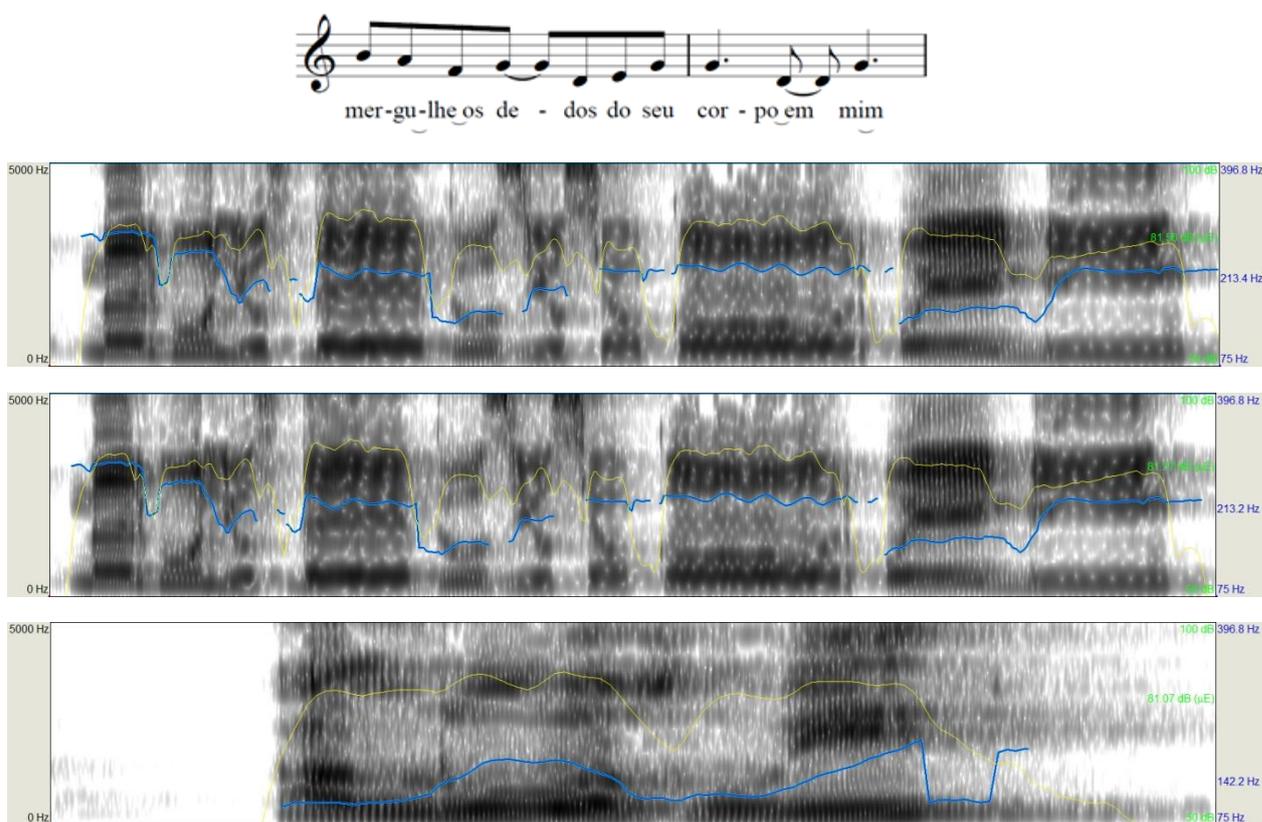


Figura 24: “Alta a maré em seus alcances sombrios”

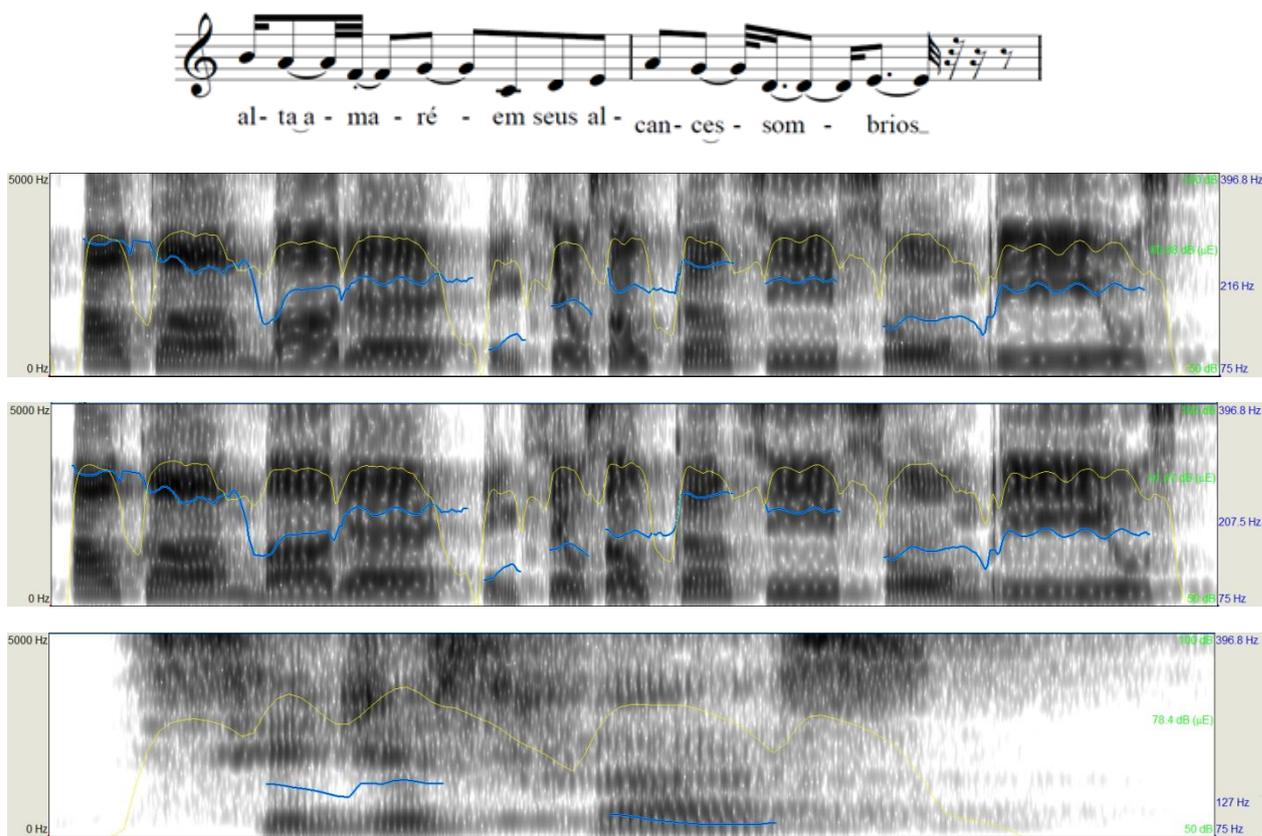


Figura 25: “O aroma das noites vem para brilhar”

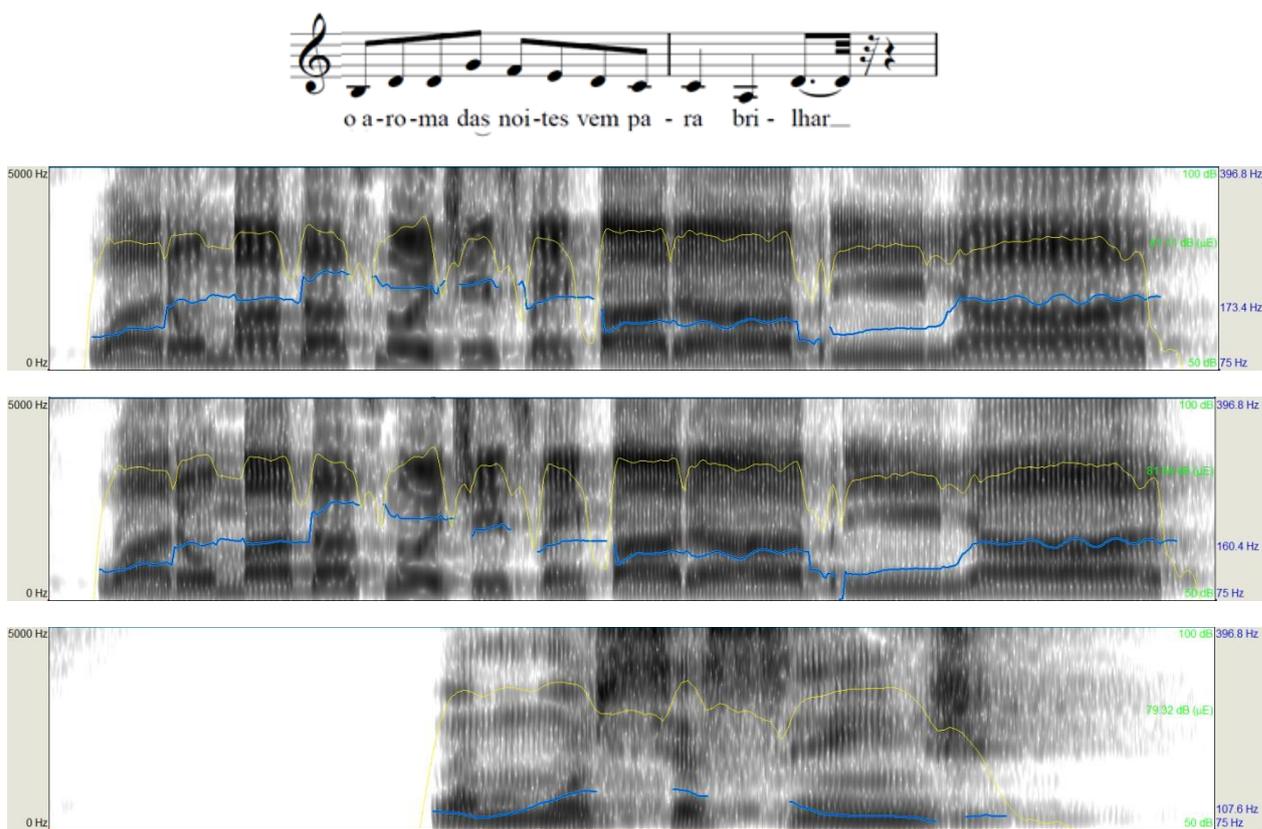


Figura 26: “Pétalas dançando no meu olhar”

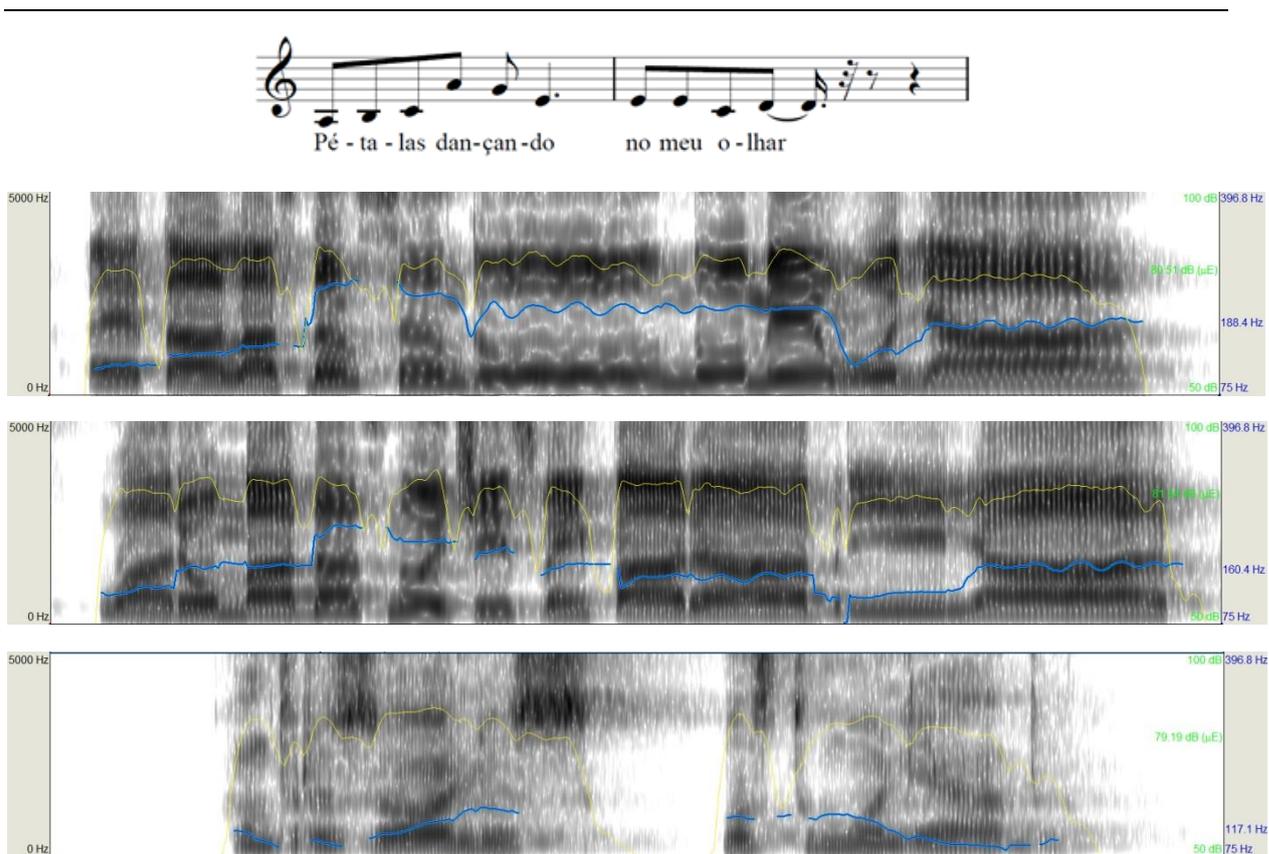


Figura 27: “As mãos do oceano vêm me entregar”

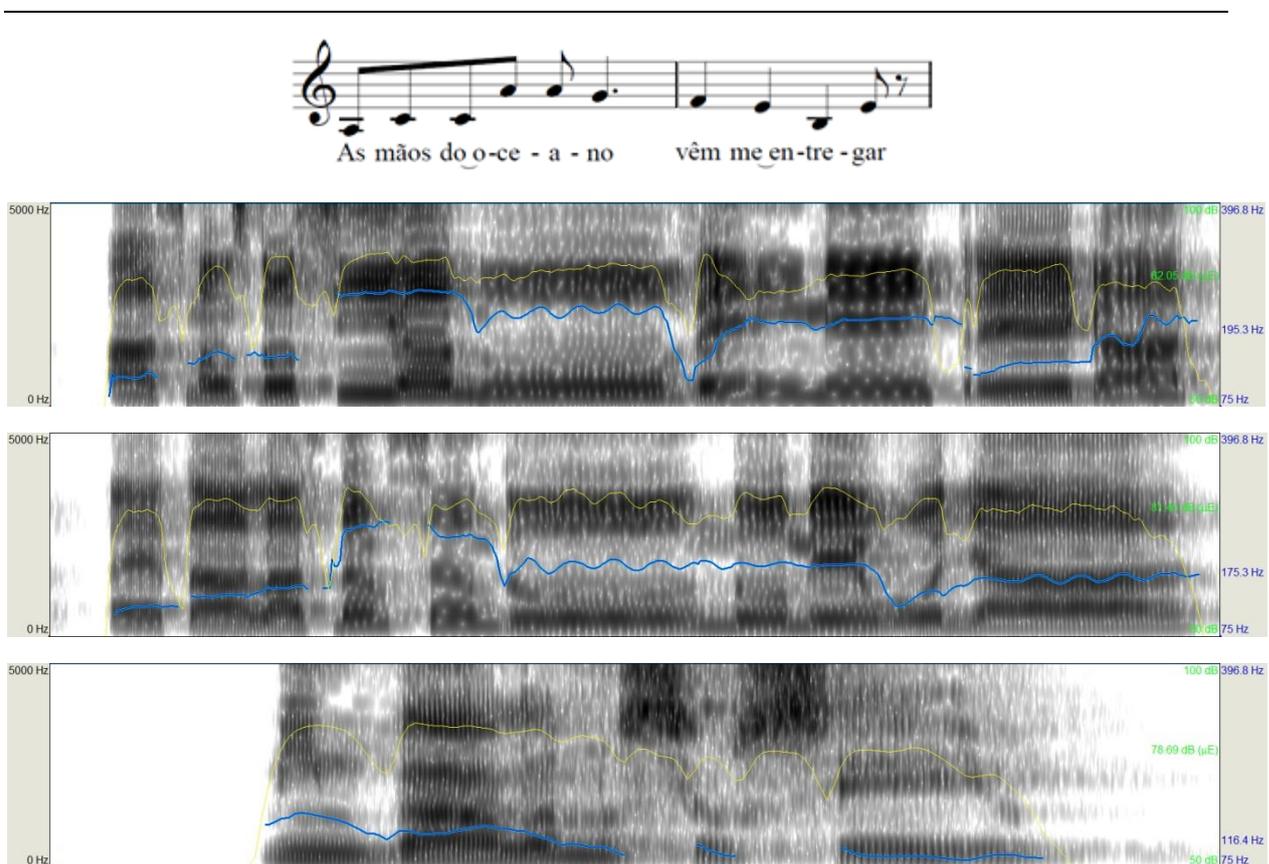


Figura 28: “Seus braços me acariciar”

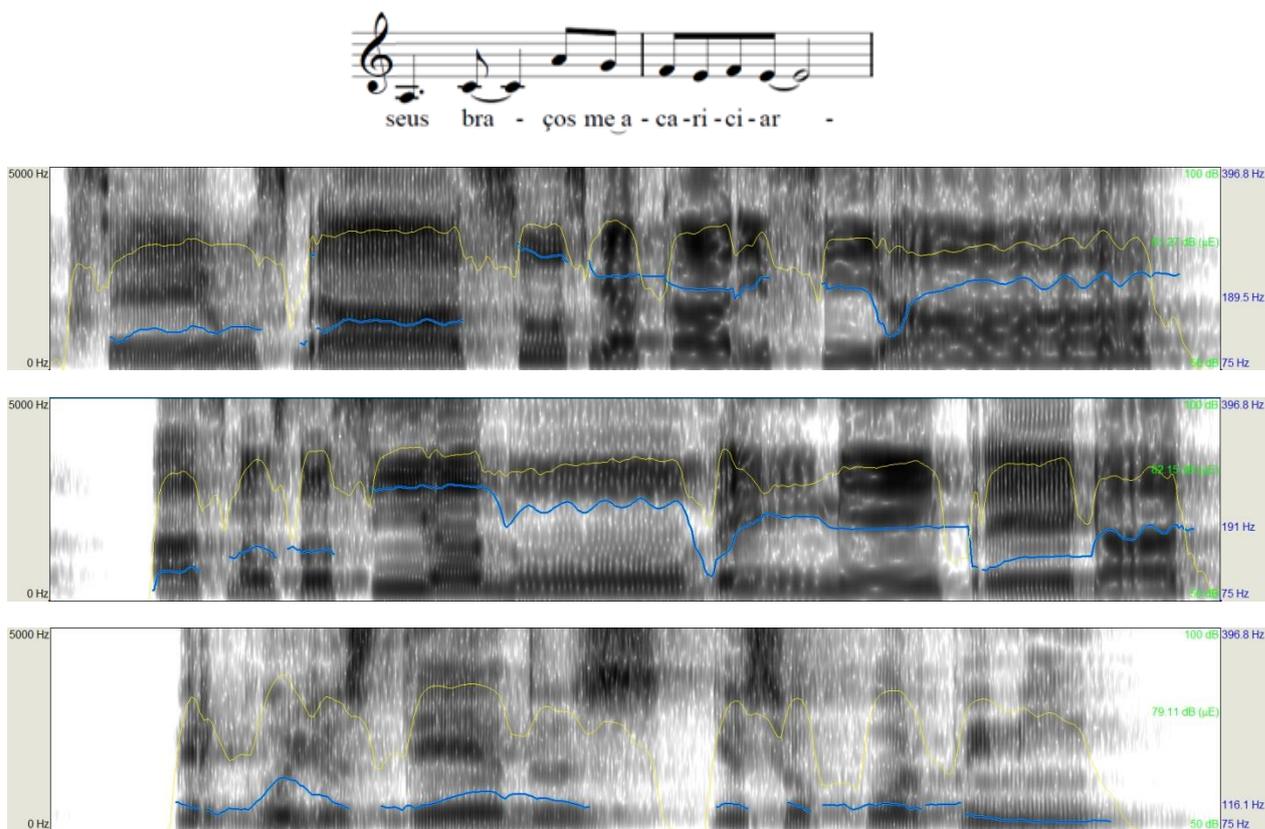


Figura 29: “Suave maestria apressa o meu pulso”

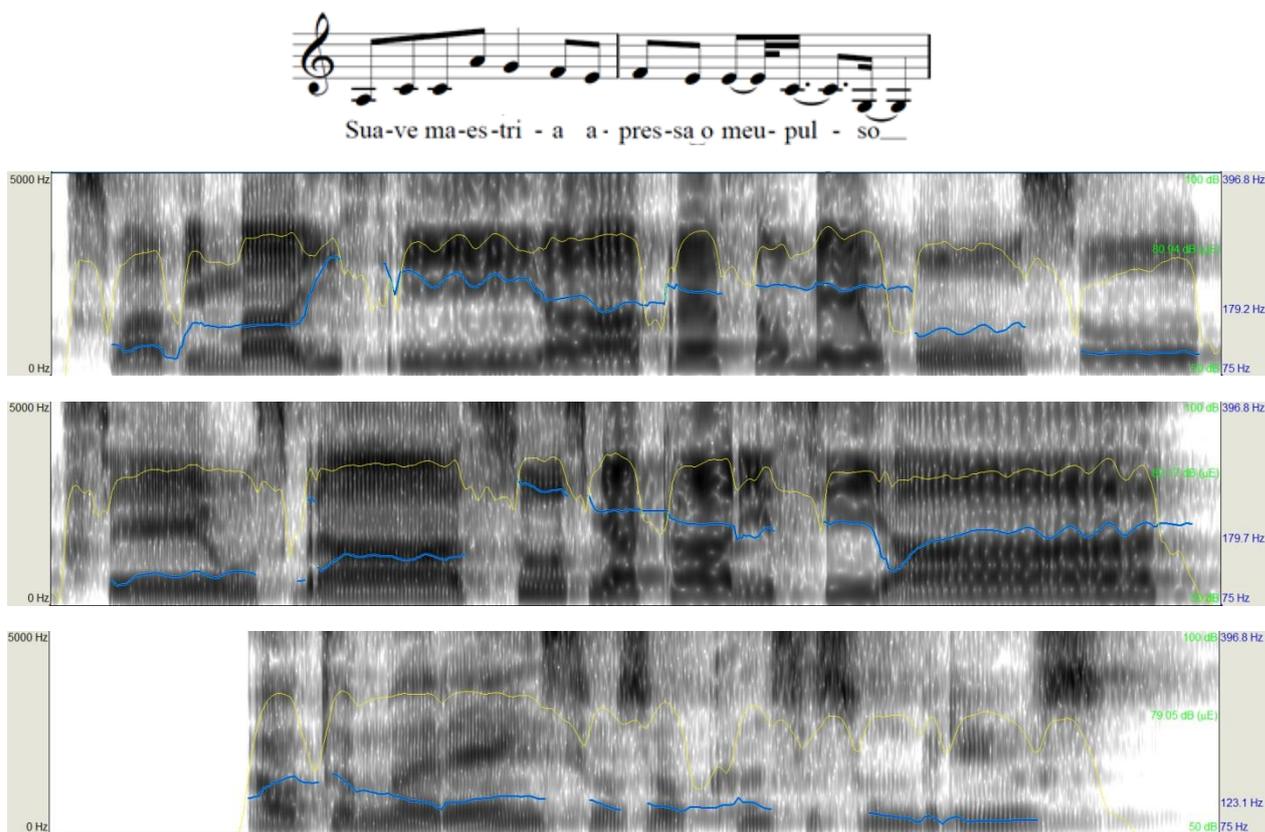


Figura 30: “Firme suavidade da carne”

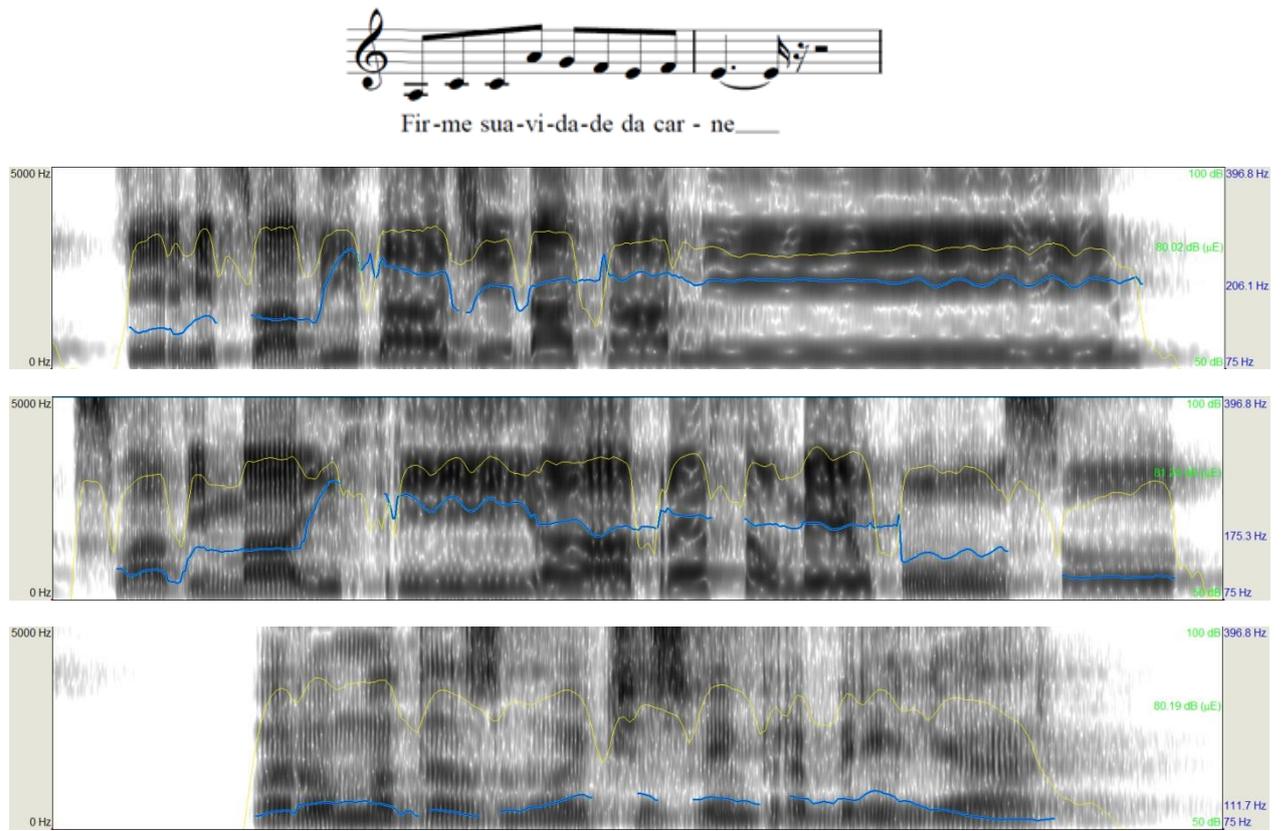


Figura 31: “Sempre meu guardado mistério”

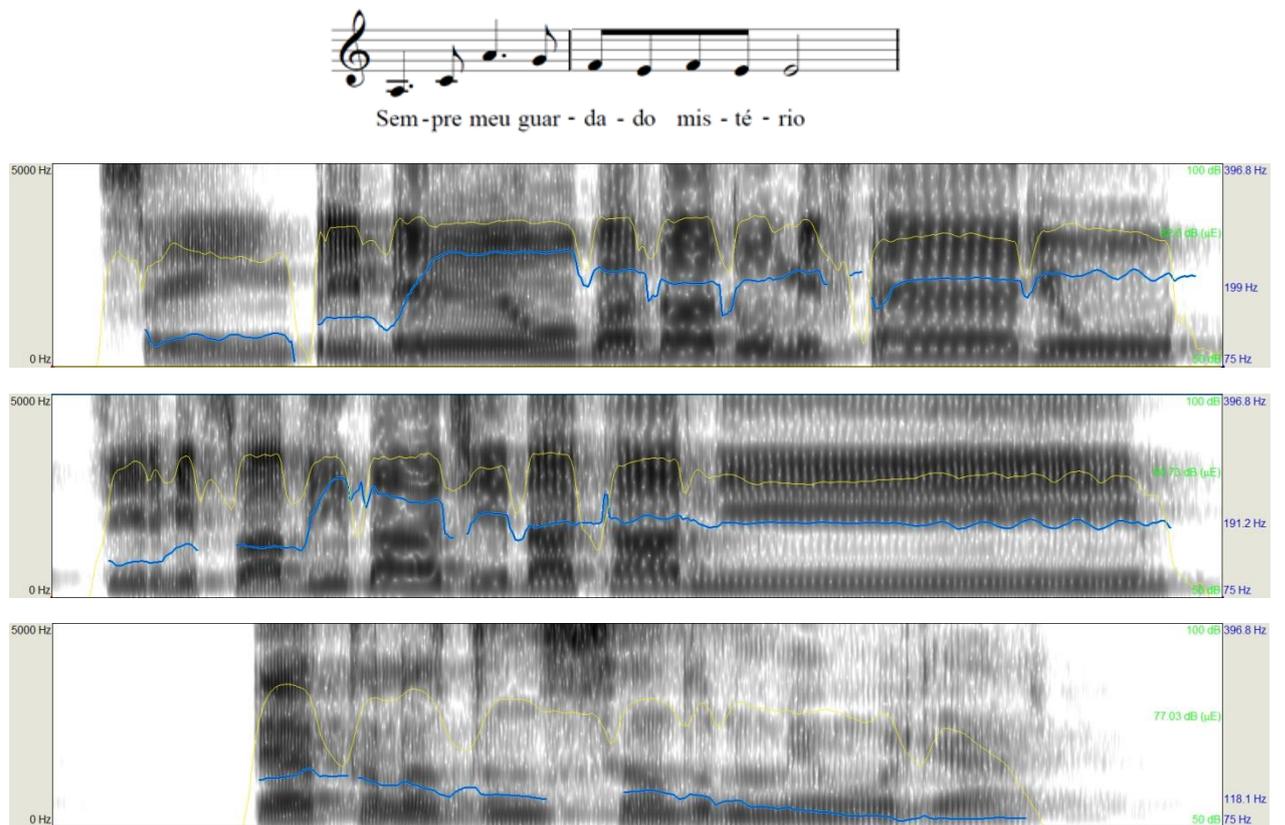


Figura 32: “Junto ao sopro das estrelas”

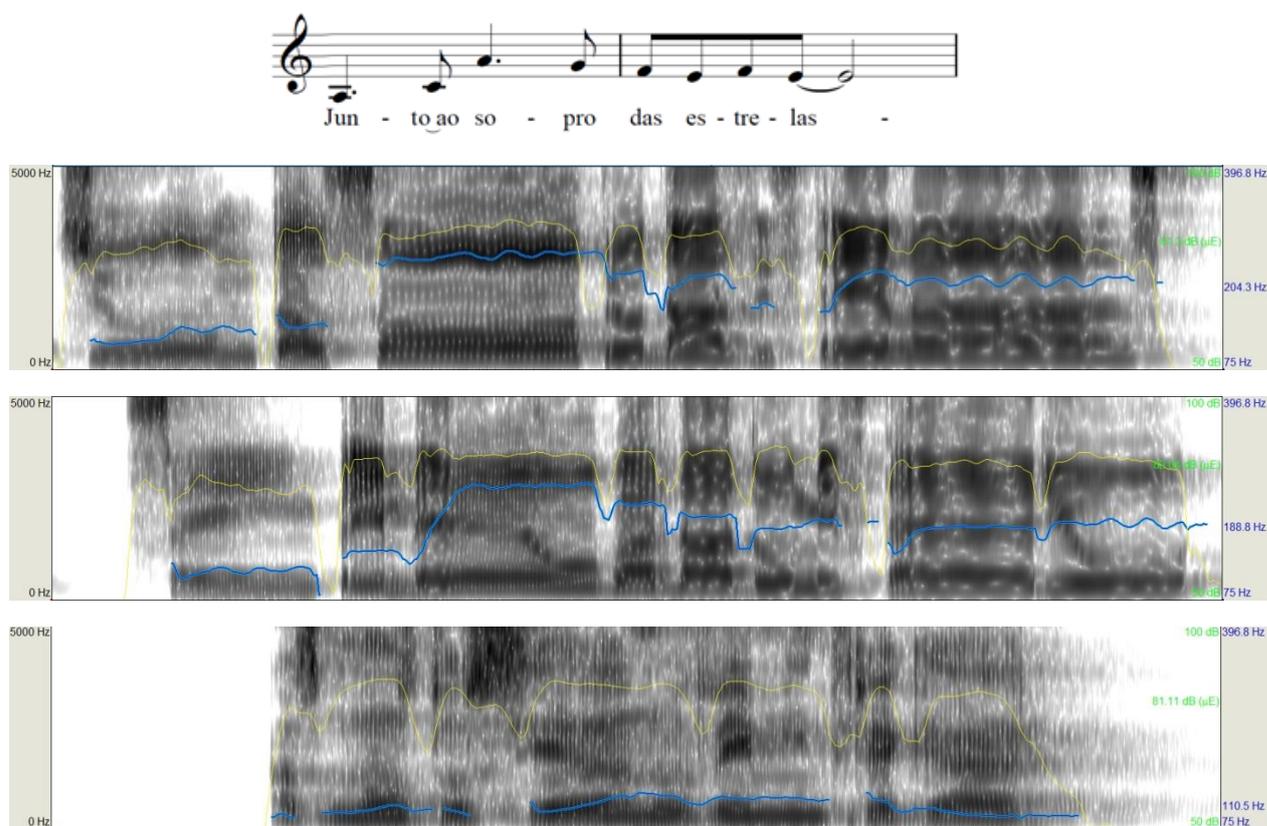
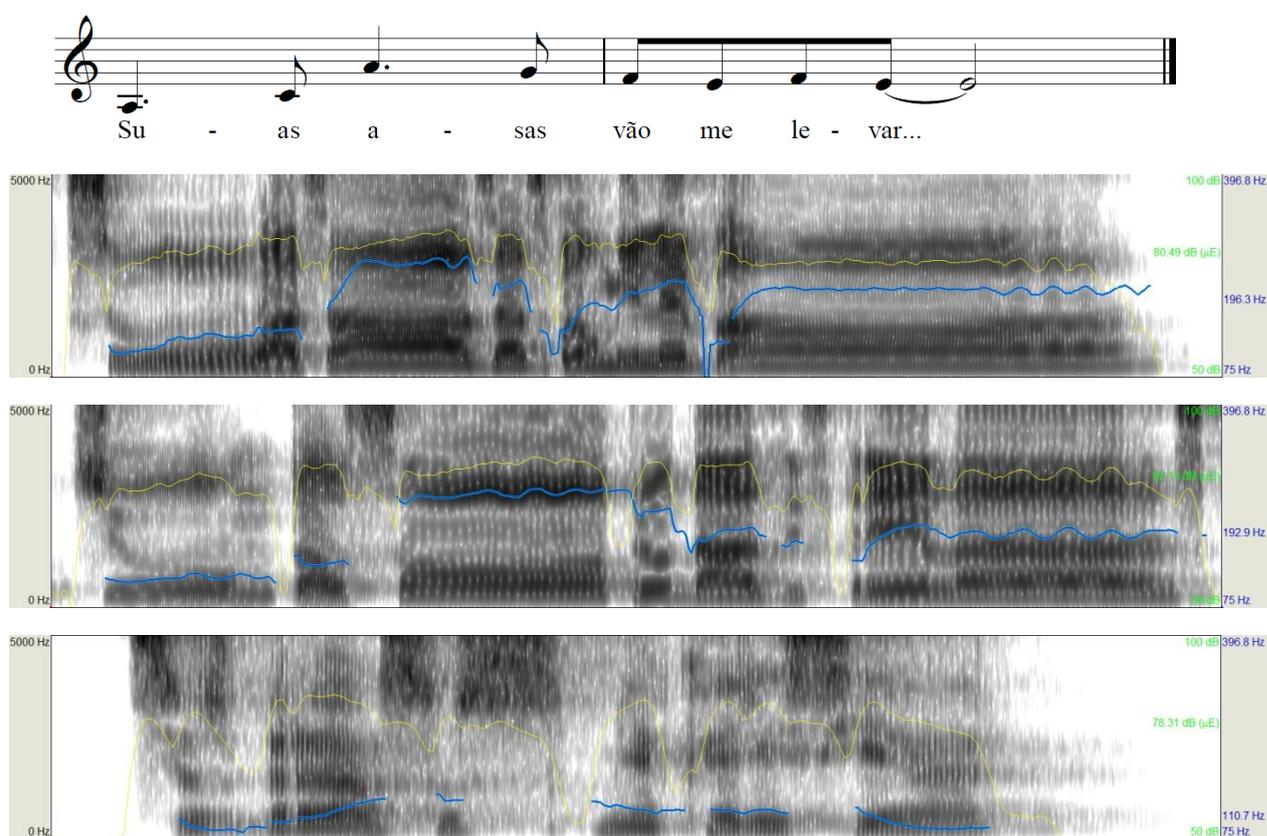


Figura 33: "Suas asas vão me levar"



3.2 ELABORAÇÃO E APLICAÇÃO DO EXPERIMENTO

O projeto que deu origem a esse trabalho foi aprovado pelo Comitê de Ética em Pesquisa CEP/FCLAr, registrado sob número CAAE 47797015.1.0000.5400, número do parecer 1.179.332, data da relatoria 01/09/2015.

Uma vez que seria muito difícil reunir um número considerável de informantes em uma data e local específicos para a realização do experimento, decidiu-se que o teste de percepção seria aplicado individualmente, *online*, por questão de comodidade. Isso foi possível através da plataforma Typeform, escolhida depois de extensa pesquisa e experimentação dos vários recursos disponíveis na internet, como Google Forms, Survey Monkey e Qualtrics.

Devido a limitações da plataforma, não era possível simplesmente anexar um arquivo de som ao formulário. O Typeform, no entanto, permitia que vídeos de sites de streaming, como YouTube e Vimeo, fossem anexados aos formulários. Sendo assim, foram criados seis vídeos, com o áudio de cada gravação. A única imagem exibida durante o vídeo era a mensagem “Por favor, ouça essa gravação até o final e faça sua avaliação.”, escrita em letras brancas sobre um fundo completamente preto. Os vídeos foram carregados na plataforma de streaming YouTube.

O questionário em si foi desenvolvido com base no extenso experimento realizado por Arno Engelmann e relatado em seu livro *Os estados subjetivos: uma tentativa de classificação de seus relatos verbais* (1978). No livro, Engelmann (1978) conduz uma pormenorizada discussão acerca dos diversos modelos de questionários existentes e conclui que, para os fins de sua investigação, o questionário baseado em escalas bipolares é o mais adequado. Em uma reflexão que perpassa vários capítulos, Engelmann (1978) também movimentou etimologia, lexicografia e experimentos de psicologia e biologia para chegar a dezesseis escalas bipolares e duas não bipolares (ativação e duração) que, segundo seu julgamento, são adequadas para definir os relatos verbais referentes a emoções. As escalas bipolares empregadas por Engelmann são:

FORTE - FRACO
 AGRADÁVEL - DESAGRADÁVEL
 ALEGRE - TRISTE
 CONTROLE COMPLETO SOBRE SI MESMO - FALTA DE
 CONTROLE SOBRE SI MESMO
 TENSO - RELAXADO
 ACEITANDO - REJEITANDO

CONTROLE COMPLETO SOBRE A SITUAÇÃO - FALTA DE
 CONTROLE SOBRE A SITUAÇÃO
 VONTADE DE ESTAR COM OUTRAS PESSOAS - VONTADE
 DE FICAR SOZINHO
 BOM - MAU
 ATIVO - PASSIVO
 DOMINANTE - SUBMISSO
 QUERENDO ALGO - NÃO QUERENDO NADA
 DESEJANDO CONSERVAR - DESEJANDO DESTRUIR
 LEVANDO A SÉRIO - LEVANDO NA BRINCADFEIRA
 PREOCUPADO COM OS OUTROS - PREOCUPADO CONSIGO
 MESMO

(ENGELMANN, 1978, p. 170)

Dessas dezesseis escalas, cinco foram selecionadas, uma adaptada e duas combinadas para serem usadas neste trabalho. A escolha dessas escalas foi baseada nos experimentos de Curtis e Bharucha (2010), que usavam “Angry”, “Happy”, “Pleasant” e “Sad” (em português, “Irritado”, “Feliz”, “Agradável” e “Triste”, respectivamente) como definições de emoções a serem expressas pelas atrizes cuja fala foi analisada, pelas associações frequentemente levantadas em teoria musical no que se refere a intervalos e harmonia (comumente evocando tensão, relaxamento, tristeza e alegria). As escalas selecionadas e adaptadas para esse trabalho foram:

(3.3)

- Agradável – Desagradável;
- Excitado – Deprimido, baseada na escala de ativação não-bipolar elaborada por Engelmann (1978, p. 168);
- Alegre – Triste;
- Tenso – Relaxado;
- Interessado – Desinteressado, adaptação da escala ACEITANDO – REJEITANDO, que em Engelmann se refere a estar ou não aberto e interessado a novos estímulos na situação (ENGELMANN, 1978, p . 155, 157);
- Controle completo sobre si mesmo e da situação - Falta de controle sobre si mesmo e sobre a situação, combinação das escalas CONTROLE COMPLETO SOBRE A SITUAÇÃO - FALTA DE CONTROLE e CONTROLE COMPLETO SOBRE SI MESMO - FALTA DE CONTROLE SOBRE SI MESMO originais de Engelmann;
- Levando a sério - Levando na brincadeira.

Ainda que as escalas selecionadas fossem bipolares, durante a elaboração do questionário, optamos por rotular o ponto do meio da escala, correspondendo à casa de número 3, como neutro (*Nem muito agradável, nem muito desagradável, Nem muito excitado, nem muito deprimido*, etc). Essa decisão foi tomada no intuito de estabelecer um parâmetro para os informantes: abaixo de 3, na direção de um dos polos, acima de 3, na direção de outro, 3 no meio do caminho. Dessa forma, os informantes teriam a ideia de simetria entre os graus da escala, evitando possíveis questionamentos como “A partir de qual casa deixa de ser agradável e passa a ser desagradável?”.

Foram criados três formulários diferentes, cada um dividido em duas partes. Na primeira parte, os informantes eram expostos a uma das gravações de *Arritmia*, na segunda, a uma das versões de *Cummings*. Mais especificamente:

(3.4)

- Formulário A – Arritmia declamada e Cummings menor
- Formulário B – Arritmia menor e Cummings declamada
- Formulário C – Arritmia maior e Cummings maior

Os três formulários tinham uma tela de boas-vindas comum, com o título “Música, emoção e linguagem verbal”, e as seguintes instruções:

Olá, seja bem-vindo!

Por favor, leia atentamente as instruções a seguir antes de continuar.

Esta é uma pesquisa sobre a maneira como as pessoas percebem diferentes emoções quando elas são expressas em um enunciado verbal e/ou em uma melodia cantada. Sua participação levará por volta de cinco minutos.

Você vai ouvir duas gravações de aproximadamente um minuto cada. Depois de ouvir, você vai encontrar uma lista de termos relacionados ao estado emocional expresso na gravação, em escalas bipolares. Você só precisa classificar a gravação de acordo com o sentimento que você acha que ela expressa.

Quando acabar, responda algumas perguntas para que possamos traçar o perfil demográfico das pessoas que participaram dessa pesquisa.

Sua participação nesta pesquisa é voluntária e sua identidade não será divulgada.

Muito obrigado pela sua participação!

Ao clicar no botão “Começar” ou simplesmente teclar “Enter”, o informante era levado à página de questões. Sob o título de “Gravação nº 1”, as instruções pediam aos informantes que ouvissem o áudio incorporado na página e respondessem a questões, lembrando que não havia resposta certa ou errada. Cada escala era, então, apresentada na forma de questão, e as respostas deveriam ser fornecidas através de uma escala bipolar de cinco casas numeradas. As casas de número 1, 3 e 5 foram rotuladas para orientar os informantes (por exemplo, com as expressões. “Alegre”, “Nem muito alegre, nem muito triste”, e “Triste”, respectivamente). Por questões de espaço e para evitar desconfigurar o formulário, apenas as casas 1 e 5 relativas às questões F e G foram rotuladas. As questões apresentadas aos informantes, bem como os rótulos das casas de número 1, 3 e 5 das respostas, foram as seguintes:

(3.5)

- a) Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?

Agradável, Nem muito agradável, nem muito desagradável, Desagradável

- b) O enunciador da gravação parece estar mais excitado, cheio de energia, ou deprimido, sem ânimo?

Excitado, Nem muito excitado, nem muito deprimido, Deprimido

- c) O enunciador da gravação parece estar alegre ou triste?

Alegre, Nem muito alegre, nem muito triste, Triste

- d) O enunciador parece estar tenso ou relaxado?

Tenso, Nem muito tenso, nem muito relaxado, Relaxado

- e) O enunciador parece estar interessado ou desinteressado em relação à situação em que ele se encontra?

Interessado, Nem muito interessado, nem muito desinteressado, Desinteressado

- f) O enunciador parece estar no controle de si mesmo e da situação ou parece não ter controle algum?

Controle completo sobre si mesmo e/ou a situação, Falta de controle sobre si mesmo e/ou a situação

- g) O enunciador parece estar levando a situação a sério ou levando a situação na brincadeira?

Levando a sério, Levando na brincadeira

Figura 34: Um dos questionários utilizados no experimento

1 → **Gravação nº 1**

Ouçã o áudio abaixo e responda as questões que se seguem. Lembre-se que não existe resposta certa ou errada. O que nos importa é a **sua opinião**.

Por favor, ouça esta gravação até o final e faça sua avaliação.



a. Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?^{*}

1	2	3	4	5	
Agradável		Nem muito agradável, nem muito desagradável			Desagradável

0 de 19 respondidas

Crie o seu próprio typeform...

↑

↓

Após responder a uma questão, o formulário automaticamente mudava para a próxima questão. Vale ressaltar, no entanto, que era possível subir a página e ouvir a gravação novamente, assim como alterar as respostas dadas às questões anteriores. Depois da última questão, mas ainda na mesma página, a segunda gravação era apresentada, com as mesmas instruções, seguida do novo conjunto de questões, igualmente idênticas ao primeiro conjunto em enunciado e apresentação.

No intuito de verificar possíveis correlações entre as variações nas avaliações e fatores demográficos, após avaliar as duas gravações, era pedido ao informante que respondesse algumas perguntas pessoais. As perguntas, bem como suas referentes opções de respostas, eram:

(3.6)

a) Com qual gênero você se identifica?

Feminino

Masculino

Não-binário

Outro / Prefiro não responder

b) Qual é a sua idade?

Menos de 18 anos

18 a 24 anos

25 a 34 anos

35 a 44 anos

45 a 54 anos

55 anos ou mais

c) Qual é o seu nível de escolaridade?

Não completei nenhum nível de educação formal

Ensino fundamental

Ensino médio

Ensino superior / técnico

Pós-graduação

d) Você toca algum instrumento musical ou canta (em coral, banda, etc), ainda que não profissionalmente? Selecione mais de uma opção, se necessário.

Sim, canto

Sim, toco um instrumento

Já cantei e/ou toquei um instrumento musical, mas não pratico mais

Não, nunca cantei nem toquei instrumentos musicais

Nunca cantei nem toquei instrumentos musicais, mas gostaria de aprender

Figura 35: Um das questões referentes a aspectos demográficos dos informantes

3 → Conte-nos um pouco sobre você

d. Você toca algum instrumento musical ou canta (em coral, banda, etc), ainda que não profissionalmente? Selecione mais de uma opção, se necessário.*

Podem ser selecionadas várias opções

A Sim, canto

B Sim, toco um instrumento

C Já cantei e/ou toquei um instrumento musical, mas não pratico mais

D Não, nunca cantei nem toquei instrumentos musicais

E Nunca cantei nem toquei instrumentos musicais, mas gostaria de aprender

1 de 19 respondidas

Crie o seu próprio typeform...

^

v

Por fim, havia uma última questão, na qual os informantes eram convidados a deixar suas impressões acerca do experimento, caso desejassem. As instruções foram:

Você terminou sua participação nesta pesquisa. Se você tiver algum comentário, crítica, sugestão ou elogio, sinta-se à vontade para escrever abaixo. Não se esqueça de clicar no botão "Enviar", mesmo que você não queira escrever mais nada. Muito obrigado por participar!

Como um adendo à elaboração dos formulários, admite-se que o ideal seria incluir no experimento um teste auditivo, para se certificar de que todos os respondentes tivessem audição perfeita ou, pelo menos, boa o suficiente para permitir a participação tranquila no teste. No entanto, devido às circunstâncias em que a coleta de dados foi realizada, isso não foi possível, restando ao pesquisador confiar no bom-senso dos respondentes, na fé de que nenhum dos informantes era surdo e/ou deu respostas aleatórias às perguntas no formulário.

A pesquisa foi divulgada através da rede social Facebook, tanto no perfil pessoal dos envolvidos, quanto por meio de mensagens nos grupos da Faculdade de Ciências de Letras de Araraquara e do Programa de Pós-Graduação em Linguística e Língua Portuguesa. Ao todo, 102 pessoas participaram da pesquisa: 45 informantes responderam o formulário A, 27 informantes responderam o formulário B e 30 informantes responderam o formulário C. Para que o número de respondentes a cada formulário não fosse excessivamente discrepante, minimizando também a chance de que um ou outro formulário não fosse respondido, todas as postagens de divulgação no Facebook apresentavam *links* para os três formulários. A ordem desses *links*, no entanto, foi deliberadamente alterada em cada postagem: enquanto uma postagem apresentava a ordem de formulários B-A-C, outra apresentava a ordem A-C-B, e assim por diante. Outra estratégia adotada pelo pesquisador, antes mesmo da divulgação dos formulários pela rede social, foi pedir a amigos e familiares, um a um, que respondessem a um questionário específico. Deste modo, mantendo as anotações de qual formulário tinha sido enviado a qual pessoa, foi possível se certificar que, à época da divulgação mais ampla da pesquisa, cada um dos formulários já contasse com aproximadamente 10 respondentes. Sessenta dias após o lançamento dos formulários e o início da divulgação do experimento, os resultados foram coletados. A etapa seguinte do experimento foi analisar os dados.

3.3 METODOLOGIA DE ANÁLISE DE DADOS

A plataforma Typeform oferece a possibilidade de exportar os resultados de cada formulário no formato de planilhas com extensão .xlsx ou .csv, as quais podem ser abertas através do programa Excel, parte do pacote Office, da Microsoft.

A princípio, foi calculada a média aritmética das respostas. Esse método, no entanto, não se mostrou efetivo para que qualquer conclusão pudesse ser tirada (para mais detalhes, ver a seção de resultados). Fez-se necessário, assim, o emprego de provas estatísticas.

A análise estatística dos dados do experimento realizado nesse trabalho é baseada em Pereira (1999) e, principalmente, em Siegel (1975), que fornece um roteiro para análises estatísticas:

O primeiro passo, ou estágio, no processo de tomada de decisão, é definir a hipótese de nulidade (H_0). A *hipótese de nulidade* é uma hipótese de que não haja diferenças. ⁽¹⁾ Formula-se usualmente com o exposto propósito de ser rejeitada. Se é rejeitada, pode-se aceitar a hipótese alternativa (H_1). A *hipótese alternativa* é a definição operacional da hipótese de pesquisa do pesquisador. A *hipótese de pesquisa* é a predição deduzida da teoria que está sendo comprovada. (SIEGEL, 1975, p. 7)

Delineiam-se, assim, as seguintes hipóteses para esse trabalho:

- H_0 : Aspectos musicais, como melodia, ritmo e tonalidade, não exercem qualquer influência sobre a percepção e interpretação de enunciados verbais.
- H_1 : Aspectos musicais, como melodia, ritmo e tonalidade, exercem influência sobre a percepção de enunciados verbais.

Considerando que a decisão pela rejeição de uma ou outra hipótese se baseia nos dados obtidos no experimento, e tais dados são constituídos de valores numéricos entre 1 a 5, pode-se também expressar as hipóteses da seguinte forma.

(3.7)

$$H_0 = \mu_1 \neq \mu_2$$

em que:

μ_1 = a média da população 1, para determinada resposta, e

μ_2 = a média da população 2, para determinada resposta.

Como exemplo e para fins de elucidação, tomemos o texto número 1, *Arritmia*, em suas três versões. A partir dessas três versões, temos três amostras populacionais: a população

exposta ao texto sendo declamado, a população exposta ao texto musicado em modo maior, e a população exposta ao texto musicado em modo menor. Chamemos essas três populações de 1, 2 e 3, respectivamente. As três populações responderam a um questionário com questões idênticas, cujas respostas consistiam de valores numéricos entre 1 e 5. Tomando a primeira pergunta do questionário, “Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?”, teremos uma média numérica das respostas da população 1 (μ_1), uma média numérica das respostas da população 2 (μ_2) e uma média numérica das respostas da população 3 (μ_3). H_0 será considerada falsa se, ao longo da análise dos resultados obtidos, verificar-se que $\mu_1 < \mu_2$, ou que $\mu_1 > \mu_2$, ou que $\mu_1 < \mu_3$, ou ainda que $\mu_1 > \mu_3$.

O próximo passo é a escolha de uma prova estatística que nos ajude a acatar ou rejeitar H_0 . As provas estatísticas dividem-se em dois modelos estatísticos: paramétrico e não-paramétrico. Os modelos estatísticos paramétricos requerem que uma série de condições sejam satisfeitas para que os dados sejam adequadamente analisados por uma prova paramétrica. São, no mínimo, as seguintes:

1. As observações devem ser independentes. Isto é, a escolha de determinado elemento para inclusão na amostra não deve influir na escolha de outros elementos. E o valor numérico atribuído a cada caso não deve influir nos valores atribuídos aos outros casos.
2. As observações devem ser extraídas de populações com distribuição normal.
3. As populações devem ter a mesma variância (ou, em casos especiais, deve existir uma relação conhecida entre as variâncias).
4. As variáveis em jogo devem ser medidas pelo menos em escala intervalar, de modo que seja possível utilizar as operações aritméticas (adição, multiplicação, obtenção de médias, etc.) sobre os valores. (SIEGEL, 1975, p. 20)

Foquemo-nos, aqui, na exigência de número 4. Essa exigência se refere ao nível de mensuração utilizado em determinada pesquisa. Siegel (1975, p. 23-33) discute quatro níveis de mensuração: nominal, ordinal, intervalar e de razão. As descrições e os exemplos de cada escala, apresentadas a seguir, são adaptados desse autor.

Escala nominal, também chamadas classificadoras, constituem o nível mais baixo de mensuração, na qual números ou outros símbolos são usados simplesmente para classificar um objeto, pessoa ou característica. Os números das camisas de futebol dos jogadores, por exemplo, encontram-se em uma escala nominal: esses números identificam os jogadores, mas não nos dizem nada sobre eles.

Escala ordinal, ou escalas por postos, por outro lado, ocorrem quando os elementos em uma categoria não são apenas diferentes dos elementos em outras categorias da mesma

escala, mas guardam certo tipo de *relação* entre si. Um bom exemplo de escala ordinal é o sistema de graduação do serviço militar: Sargento > Cabo > Soldado. Note-se que os elementos dessa escala se definem entre si, e não a um referencial externo. Sabe-se que um soldado está abaixo de um cabo, mas não é possível medir, de forma objetiva, exatamente a distância entre eles.

Quando a escala tem todas as características de uma escala ordinal e, além disso, se conhecem as distâncias entre dois números quaisquer da escala, obtém-se uma mensuração no sentido de uma **escala intervalar**, na qual a fixação das classes de objetos é precisa a ponto de sabermos exatamente quão grandes são os intervalos (distâncias) entre todos os elementos da escala. A temperatura, por exemplo, é medida em uma escala intervalar. A unidade de medida e o ponto zero, em uma escala intervalar, são arbitrários. No caso da temperatura, utilizando a escala centígrada, estabeleceu-se arbitrariamente que o congelamento da água ocorre a zero graus centígrados e a fervura a 100 graus centígrados. Ainda assim, é possível dizer que a diferença entre as temperaturas de 30°C e 10°C é a mesma diferença entre as temperaturas de 70°C e 50°C. Em outras palavras, é possível medir e afirmar, de forma objetiva e precisa, que a distância/diferença entre 10°C e 30°C é a mesma distância entre 50°C e 70°C, desde que essa distância seja medida em °C.

Finalmente, quando uma escala tem todas as características de uma escala de intervalos e, além disso, tem um ponto zero não-arbitrário na origem, temos uma **escala de razões**. Medimos massa ou peso em uma escala de razões. A escala métrica, em gramas e quilogramas, tem um ponto zero verdadeiro, a ausência total de massa. Assim, pode-se dizer que, em uma escala de razões, a unidade de medida é arbitrária, mas o zero verdadeiro não. Tal fato permite que qualquer prova estatística seja aplicável aos dados.

Voltemos, agora, para os dados obtidos através do experimento. A escolha pela prova estatística adequada implica, primeiramente, a verificação da escala de mensuração na qual os dados obtidos se encontram. Apesar de os dados encontrarem-se na forma de números, julgamos inadequado interpretá-los dentro de uma escala intervalar ou de razões, pelos seguintes motivos:

- a- As alternativas de respostas são referências instintivas, porém vagas do ponto de vista da mensuração. Enquanto seres humanos, temos uma intuição bastante forte acerca da diferença entre, por exemplo, “triste” e “alegre”, e “interessado” ou “desinteressado”. No entanto, a distância entre esses dois polos (e dos pontos intermediários entre eles) não pode ser medida objetivamente: não há unidade de medida adequada para mensurar a distância entre “levando a sério” e “levando na

brincadeira”. É plausível, ainda, que alguns informantes julguem as distâncias entre os polos de forma diferente: um informante pode enxergar mais nuances entre “tenso” e “relaxado” e achar que esses dois termos estão mais distantes entre si do que outro informante poderia considerar, bem como um informante pode julgar que a distância entre “levar a sério” e “levar na brincadeira” é maior ou menor do que a distância entre “agradável e desagradável”. Essa flexibilidade impede que tomemos os dados como em uma escala de razões ou intervalar.

- b- A natureza do fenômeno que se quer investigar implica que as respostas obtidas através do questionário não são informações coletadas de forma totalmente objetiva, mas sim resultado de interpretações e mensurações dos informantes. Quando perguntado se “O enunciador da gravação parece estar mais excitado, cheio de energia, ou deprimido, sem ânimo?”, antes de responder, o informante precisa parar um instante e conceituar, internamente, o que significa “excitado, cheio de energia” e o que significa estar “deprimido, sem ânimo”. Existem, é certo, fenômenos que podem ser observados e medidos de forma bastante objetiva, em se tratando de estados subjetivos e emoções: os batimentos cardíacos, a dilatação da pupila, o suor e a tensão muscular, por exemplo, são fenômenos que podem ser associados a certos estados emocionais, não são, porém, o *estado emocional em si*, necessariamente. Ao pedir que um falante analise a gravação, o falante percebe fenômenos objetivos (como a entonação, o timbre, as alturas, os acentos e os fones) e, a partir deles, constrói uma representação mental do estado emocional no qual o enunciador da gravação supostamente se encontra. Assim, julgar que os dados obtidos possam ser tomados como em uma escala intervalar ou de razões é, no mínimo, temerário, pois implicaria fazer pressuposições sobre a objetividade e a regularidade do julgamento dos informantes.

Ao mesmo tempo, os dados obtidos não podem ser considerados como parte de uma escala meramente nominal devido ao fato de que eles não representam apenas “identificações”. Certamente, há um contínuo entre “Agradável” e “Desagradável” e, ainda que não se possa medir exatamente o quanto, não se pode discutir que os pontos estabelecidos dentro desse contínuo, por mais arbitrários que sejam, guardam uma relação entre si. Assim, o nível de mensuração dos dados obtidos no experimento desse trabalho é o nível de uma escala ordinal. A partir desse dado, torna-se possível escolher uma prova estatística adequada.

As provas estatísticas foram realizadas utilizando o software SOFA, versão 1.4.6. Para execução das provas, os dados foram importados diretamente das planilhas .csv fornecidas

pela plataforma Typeform. Duas provas oferecidas pelo programa foram escolhidas: para comparar as três versões de cada texto ao mesmo tempo, Kruskal-Wallis, e U de Mann-Whitney para comparar as versões de cada texto em duplas. Note-se que as comparações foram realizadas entre versões do mesmo texto, ou seja, as diversas gravações de cada texto foram comparadas entre si, mas nenhuma das versões de Arritmia foi comparada com nenhuma versão de Cummings. Temos, portanto, amostras independentes, pois os dados referentes às três versões do primeiro texto não nos dizem nada sobre os dados referentes às três versões do segundo texto.

As provas utilizadas na execução deste trabalho fornecem as probabilidades de ocorrência dos resultados observados sob H_0 , ou seja, se H_0 for verdadeira. Para decidir se esta probabilidade é significativa ou não a ponto de nos permitir tirar conclusões, lembramos um conceito abordado por SIEGEL (1975, p. 8-12), que menciona a importância de determinar um nível de significância (α), que indicaria a probabilidade de rejeitar H_0 quando H_0 fosse verdadeira. Este valor deve ser estabelecido pelo pesquisador de acordo com suas próprias estimativas, antes da pesquisa. Ainda que não faça sugestões de valores α a serem utilizados em provas estatísticas, o valor utilizado nos vários exemplos de Siegel tem $\alpha = 0,05$, o que nos leva a crer que este é um valor razoável para diversas situações³⁴. Sendo assim, optamos por julgar um valor de probabilidade como relevante se este valor for, também, de 0,05 ou menos.

A prova Kruskal-Wallis fornece a probabilidade da ocorrência dos dados observados sob H_0 . Em outras palavras, e mais especificamente no caso aqui relatado, a prova Kruskal-Wallis fornece a probabilidade de ocorrência dos dados observados caso a hipótese 0 (de que a versão da gravação à qual o informante foi exposto não influencia as respostas) seja verdadeira. Um valor de probabilidade $P=0,70$ indica que os valores obtidos têm 70% de chance de ocorrer se H_0 é verdadeira, o que nos permite acatar H_0 e concluir que a variação nos dados provavelmente se deve a puro acaso. Uma probabilidade $P=0,01$, por outro lado, indica uma probabilidade de 1% de chance de que tais dados sejam obtidos no caso de H_0 ser verdadeira, e um valor P tão baixo nos permite concluir, com razoável segurança, que a variação nos dados não se deve a puro acaso e que H_0 deve ser, portanto, rejeitada. Consideramos que valores de P inferiores a 0,05 (5% de chance de ocorrência sob H_0) indicam variações significativas.

³⁴ Este mesmo valor de α foi escolhido, por exemplo, no trabalho de Medeiros (2002, p. 28, nota de rodapé), parte da bibliografia da presente dissertação de mestrado, e em Massini-Cagliari (1992, p. 51-52).

Paralelamente a Kruskal-Wallis, outra prova estatística foi utilizada no intuito de verificar os resultados obtidos: a prova U de Mann-Whitney. Essa prova, que permite comparar duas amostras por vez, permite verificar tanto a probabilidade de distribuições unilaterais (ou seja, em um sentido específico, por exemplo, A é significativamente maior do que B) quanto de distribuições bilaterais A é significativamente maior *ou* menor do que B). Mais uma vez, consideramos que valores de P inferiores a 0,05 (5% de chance de ocorrência sob H_0) indicam variações significativas.

Para verificar possíveis correlações entre desvios nas avaliações e aspectos demográficos, duas provas foram utilizadas. A primeira, χ^2 (Chi-quadrado), é utilizada para verificar se os dois resultados são independentes ou relacionados e poderia responder, por exemplo, se há uma relação entre o nível de escolaridade e o quão alegre ou triste uma das gravações foi considerada. Infelizmente, os resultados obtidos por esta prova não puderam ser considerados válidos (ver seção de resultados para mais detalhes). A alternativa encontrada à prova χ^2 foi o Coeficiente de Correlação por Postos de Spearman (R), que verifica se há uma relação (a grosso modo) linear entre duas variáveis. Em outras palavras, a prova R ajuda a verificar se suposições como “Expostos à versão do texto musicada em modo maior, quanto mais alto o nível de envolvimento musical dos informantes, maior o escore obtido na questão referente ao polo Agradável x Desagradável”. O programa SOFA também gera, para cada aplicação do Coeficiente de Correlação por Postos de Spearman, um gráfico correspondente.

Para tanto, foi necessário alterar a forma de exibição de alguns dados demográficos. Quando se fala em faixas etárias (18 a 24 anos, 25 a 34 anos, etc) e escolaridade (Nenhum nível completo, Nível fundamental completo, Ensino médio completo, etc), estamos trabalhando no nível de mensuração ordinal (uma vez que os intervalos não são regulares, mas há uma evidente diferença de “hierarquia” entre “Ensino médio” e “Ensino superior”). Portanto, foi possível “converter” essas informações, originalmente coletadas nos questionários na forma de texto, em números. Ao menor nível de escolaridade foi atribuído o valor 1, e ao maior, 5. Assim, temos os valores:

(3.8)

Escolaridade:

- 1 – Não completei nenhum nível de educação formal
- 2 – Ensino fundamental
- 3 – Ensino médio
- 4 – Ensino superior / técnico

5 – Pós-graduação

Nas questões relativas à educação musical, os dados foram interpretados de forma a tentar quantificar o nível de envolvimento musical ativo dos informantes. O valor 1, o mais baixo na mensuração, foi atribuído às respostas “Não, nunca cantei nem toquei instrumentos musicais”, já que esses informantes podem ouvir e apreciar música no seu dia-a-dia, mas não se envolvem ativamente com ela, estudando, tocando ou cantando, por exemplo. Às respostas “Nunca cantei nem toquei instrumentos musicais, mas gostaria de aprender”, foi atribuído valor 2, por considerar que é possível que esses informantes, apesar de não se envolverem sistematicamente com música, podem ter um envolvimento ativo acima da média, por exemplo, tendo lido a respeito de teoria musical, uma vez que já demonstram interesse em tocar ou cantar. Às respostas “Já cantei e/ou toquei um instrumento musical, mas não pratico mais” foi atribuído valor 3: quem já cantou ou tocou um instrumento conta com uma bagagem de envolvimento musical maior do que quem nunca tocou ou cantou, mas menor do que quem pratica até hoje. O valor 4 foi atribuído às respostas “Sim, canto” e “Sim, toco um instrumento”, desde que não ocorressem simultaneamente, e o valor 5 foi atribuído quando os informantes respondiam “Sim, canto” e também “Sim, toco um instrumento”. Em caso de informantes que assinalavam respostas de níveis diferentes, por exemplo, “Sim, canto” e também “Já cantei e/ou toquei um instrumento musical, mas não pratico mais”, tomou-se a decisão de considerar o menor valor assinalado. Assim, temos:

(3.9)

Nível de envolvimento musical:

- 1 – Não, nunca cantei nem toquei instrumentos musicais
- 2 – Nunca cantei nem toquei instrumentos musicais, mas gostaria de aprender
- 3 – Já cantei e/ou toquei um instrumento musical, mas não pratico mais
- 4 – Sim, canto *ou* Sim, toco um instrumento
- 5 - Sim, canto *e* Sim, toco um instrumento

Por fim, no caso da idade, para manter os valores entre 1 e 5, unimos duas categorias de resposta (“Menos de 18 anos” e “18 a 24 anos”) e obtivemos os seguintes valores:

(3.10)

Idade

- 1 – Menos de 24 anos
- 2 – 25 a 34 anos
- 3 – 35 a 44 anos
- 4 – 45 a 54 anos
- 5 – Mais de 55 anos

Essas pequenas adaptações na exibição dos dados possibilitaram a busca por possíveis correlações entre as avaliações do experimento e os dados demográficos dos informantes.

3.4. CONSIDERAÇÕES FINAIS

Nesta seção, explicitamos a metodologia de preparação, aplicação e análise dos resultados do teste de percepção, ponto principal deste trabalho. Mostramos os critérios que guiaram a escolha e escrita dos textos utilizados para as gravações. Explicamos como as gravações foram realizadas e posteriormente editadas no *software* Melodyne. Foram apresentadas as partituras que nortearam as gravações, bem como os espectrogramas oriundos das três gravações (declamada, com intervalo de terça maior e com intervalo de terça menor) de cada texto. Discorremos sobre os motivos que nos levaram a formatar o questionário da maneira como o fizemos, bem como os princípios estatísticos que nortearam a análise dos dados. Mencionamos, também, quais provas estatísticas foram utilizadas e como tentamos lidar com as dificuldades encontradas.

4. RESULTADOS E ANÁLISES

As perguntas nos formulários respondidos pelos informantes eram as mesmas, independente de a qual gravação o informante havia sido exposto. Assim, para cada pergunta, há três amostras de respostas, relativas a cada gravação do texto (declamado, cantado com intervalo de terça menor e com intervalo de terça maior). Cada amostra consiste em um conjunto de N respostas dadas a cada questão. Naturalmente, há uma variação nos valores das respostas *dentro* de cada amostra e, também, *entre* as amostras. Uma vez que nosso interesse é verificar se as diferentes gravações do mesmo texto acarretaram diferentes reações nos informantes, interessam-nos as possíveis diferenças *entre* as amostras. A Tabela 01 mostra as médias aritméticas das respostas para cada questão.

Pode-se perceber, ao olhar os dados expostos na tabela, que houve *alguma* variação interamostral. No entanto, os valores são pequenos, evidenciando que é temerário (na melhor das hipóteses) tirar qualquer conclusão a partir desses valores. Não se pode negar que há variações, mas seriam essas variações *significativas*? Em outras palavras, pode-se concluir que as diferentes gravações acarretaram respostas diferentes, ou seriam essas diferenças frutos de mero acaso? Considerando, ainda, que N não é igual para as três amostras, fica claro que uma simples média aritmética não é o suficiente para tirar conclusões. Foram utilizadas, portanto, as provas de Kruskal-Wallis e a prova U de Mann-Whitney (para uma breve explicação sobre cada uma das provas, remetemos à seção de procedimentos metodológicos).

A Tabela 02 mostra os valores de P obtidos nas diversas comparações entre as amostras referentes ao primeiro texto, “Arritmia”. A primeira coluna mostra o valor de P quando as amostras obtidas a partir das três gravações são comparadas, a segunda coluna mostra os valores quando são comparadas as amostras obtidas a partir das duas gravações musicadas, a coluna seguinte mostra os valores quando são comparadas as amostras obtidas a partir das versões com o texto sendo declamado e sendo cantado em modo maior; finalmente, a última coluna mostra os resultados quando são comparadas as versões com o texto sendo declamado e cantado em modo menor. Os valores de P que indicam variações significativas estão grifados.

Tabela 1: Arritmia: Média aritmética dos valores das respostas fornecidas pelos informantes

Arritmia			
	Declamada (N=45)	Menor (N=27)	Maior (N=30)
Questão 1: Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?	2,577778	2,592592593	2,5
Questão 2: O enunciador da gravação parece estar mais excitado, cheio de energia, ou deprimido, sem ânimo?	3,111111	2,222222222	1,9
Questão 3: O enunciador da gravação parece estar alegre ou triste?	3,155556	2,296296296	2,1
Questão 4: O enunciador parece estar tenso ou relaxado?	3,355556	3,592592593	3,6
Questão 5: O enunciador parece estar interessado ou desinteressado em relação à situação em que ele se encontra?	2,222222	2,185185185	2,066667
Questão 6: O enunciador parece estar no controle de si mesmo e da situação ou parece não ter controle algum?	2,6	2,666666667	2,733333
Questão 7: O enunciador parece estar levando a situação a sério ou levando a situação na brincadeira?	1,755556	2,740740741	2,766667

Como se vê, os resultados da aplicação da prova *U*, em sua grande maioria, reforçam os resultados da prova Kruskal-Wallis. Os dados sugerem que diferenças significativas na percepção do enunciado ocorrem mais quando se comparam as versões “Declamada” e “Musicada com intervalo de terça maior”, em vez de “Declamada” e “Musicada com intervalo de terça menor” ou mesmo quando se comparam as duas versões musicadas. Dessa forma, pode-se pensar que o fator decisivo para a alteração da percepção é o texto ser ou não musicado, e não o modo no qual o texto foi musicado.

A Tabela 03 segue o modelo da tabela anterior, mas mostra os valores de P obtidos nas diversas comparações entre as amostras referentes ao segundo texto, “Cummings”.

Tabela 2: Arritmia: Probabilidade de ocorrência dos dados obtidos sob H_0 , em cada questão, de acordo com a prova Kruskal-Wallis ($gl=2$), Mann-Whitney unilateral (MWu) e Mann-Whitney bilateral (MWb)

Probabilidade de ocorrência sob H_0, na comparação ($\alpha = 0,05$)				
	Entre as 3 amostras	Entre maior e menor	Entre declamada e maior	Entre declamada e menor
1 - Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?	P=0,9358 (KW)	P=0,7986 (KW) P=0,4022 (MWu) P=0,8044 (MWb)	P=0,7248 (KW) P=0,3667 (MWu) P=0,7334 (MWb)	P=0,9330 (KW) P=0,4676 (MWu) P=0,9351 (MWb)
2 - O enunciador da gravação parece estar mais excitado, cheio de energia, ou deprimido, sem ânimo?	P=0,0003 (KW)	P=0,1937 (KW) P=0,1077 (MWu) P=0,2155 (MWb)	P=0,0002 (KW) P=0,0001 (MWu) P=0,0003 (MWb)	P=0,0084 (KW) P=0,0059 (MWu) P=0,0118 (MWb)
3 - O enunciador da gravação parece estar alegre ou triste?	P=0,0185 (KW)	P=0,4943 (KW) P=0,2561 (MWu) P=0,5123 (MWb)	P=0,0077 (KW) P=0,0054 (MWu) P=0,0109 (MWb)	P=0,0590 (KW) P=0,0352 (MWu) P=0,0705 (MWb)
4 - O enunciador parece estar tenso ou relaxado?	P=0,0480 (KW)	P=0,9474 (KW) P=0,4745 (MWu) P=0,9490 (MWb)	P=0,0244 (KW) P=0,0149 (MWu) P=0,0297 (MWb)	P=0,0652 (KW) P=0,0366 (MWu) P=0,0732 (MWb)
5 - O enunciador parece estar interessado ou desinteressado em relação à situação em que ele se encontra?	P=0,4927 (KW)	P=0,7444 (KW) P=0,3777 (MWu) P=0,7553 (MWb)	P=0,2394 (KW) P=0,1292 (MWu) P=0,2584 (MWb)	P=0,4912 (KW) P=0,2537 (MWu) P= 0,5073 (MWb)
6 - O enunciador parece estar no controle de si mesmo e da situação ou parece não ter controle algum?	P=0,7011 (KW)	P=0,7744 (KW) P=0,3899 (MWu) P=0,7797 (MWb)	P=0,3715 (KW) P=0,1935 (MWu) P=0,3869 (MWb)	P=0,7242 (KW) P=0,3657 (MWu) P=0,7315 (MWb)
7 - O enunciador parece estar levando a situação a sério ou levando a situação na brincadeira?	P=0,813 (KW)	P=0,9478 (KW) P=0,4745 (MWu) P=0,9490 (MWb)	P=0,0564 (KW) P=0,0326 (MWu) P=0,0652 (MWb)	P=0,0673 (KW) P=0,0385 (MWu) P=0,0771 (MWb)

Tabela 3: *Cummings*: Probabilidade de ocorrência dos dados obtidos sob H_0 , em cada questão, de acordo com a prova Kruskal-Wallis ($g_l=2$), Mann-Whitney unilateral (MWu) e Mann-Whitney bilateral (MWb)

Probabilidade de ocorrência sob H_0, na comparação ($\alpha = 0,05$)				
	Entre as 3 colunas	Entre maior e menor	Entre declamada e maior	Entre declamada e menor
1 - Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?	P=0,0115 (KW)	P=0,1219 (KW) P=0,0657 (MWu) P=0,1314 (MWb)	P=0,0016 (KW) P=0,0011 (MWu) P=0,0023 (MWb)	P=0,1250 (KW) P=0,0698(MWu) P=0,1396 (MWb)
2 - O enunciador da gravação parece estar mais excitado, cheio de energia, ou deprimido, sem ânimo?	P < 0,0001 (KW)	P=0,0001 (KW) P=<0,0001 (MWu) P=<0,0001 (MWb)	P=0,2023 (KW) P=0,1201 (MWu) P=0,2401 (MWb)	P=0,0003 (KW) P=0,0002 (MWu) P=0,0005 (MWb)
3 - O enunciador da gravação parece estar alegre ou triste?	P=0,0023 (KW)	P=0,0009 (KW) P=0,0007 (MWu) P=0,0014 (MWb)	P=0,0259 (KW) P=0,0189 (MWu) P=0,0377 (MWb)	P=0,1880 (KW) P=0,1066 (MWu) P=0,2133 (MWb)
4 - O enunciador parece estar tenso ou relaxado?	P=0,2295 (KW)	P=0,2169 (KW) P=0,1150 (MWu) P=0,2300 (MWb)	P=0,0909 (KW) P=0,0507 (MWu) P=0,1014 (MWb)	P=0,5810 (KW) P=0,2963 (MWu) P=0,5926 (MWb)
5 - O enunciador parece estar interessado ou desinteressado em relação à situação em que ele se encontra?	P=0,4570 (KW)	P=0,1699 (KW) P=0,0981(MWu) P=0,1962 (MWb)	P=0,6179 (KW) P=0,31870 (MWu) P=0,6373 (MWb)	P=0,7362 (KW) P=0,3767 (MWu) P=0,7535 (MWb)
6 - O enunciador parece estar no controle de si mesmo e da situação ou parece não ter controle algum?	P=0,0128 (KW)	P=0,1669 (KW) P=0,0891 (MWu) P=0,1782 (MWb)	P=0,0031 (KW) P=0,0022 (MWu) P=0,0043 (MWb)	P=0,0621 (KW) P=0,0371 (MWu) P=0,0742 (MWb)
7 - O enunciador parece estar levando a situação a sério ou levando a situação na brincadeira?	P=0,0132 (KW)	P=0,4327 (KW) P=0,2245 (MWu) P=0,4490 (MWb)	P=0,0195 (KW) P=0,0143 (MWu) P=0,0286 (MWb)	P=0,0062 (KW) P=0,0045 (MWu) P=0,0090 (MWb)

Também nesse caso os resultados da aplicação da prova U , em geral, reforçam os resultados da prova Kruskal-Wallis. No caso do segundo texto, “Cummings”, os dados mais

uma vez sugerem que as maiores diferenças na percepção são observadas quando se compara a versão declamada com uma das versões musicadas, e não quando se comparam as versões musicadas entre si. Isso fortalece a ideia de que outros elementos musicais que não o modo sejam mais influentes na percepção do enunciado.

A Tabela 04 justapõe os resultados da prova de Kruskal-Wallis obtidos nas comparações de amostras relativas aos dois textos.

Tabela 4: Ocorrência de diferenças significativas nos dados, comparação entre *Arritmia* e *Cummings*

Probabilidade de ocorrência sob H_0, na comparação ($\alpha = 0,05$)				
	Entre as 3 colunas	Entre maior e menor	Entre declamada e maior	Entre declamada e menor
1 - Na sua opinião, essa gravação expressa um sentimento agradável ou desagradável?	P=0,9358 <u>P=0,0115</u>	P=0,7986 P=0,1219	P=0,7248 <u>P=0,0016</u>	P=0,9330 P=0,1250
2 - O enunciador da gravação parece estar mais excitado, cheio de energia, ou deprimido, sem ânimo?	<u>P=0,0003</u> <u>P < 0,0001</u>	P=0,1937 <u>P=0,0001</u>	<u>P=0,0002</u> P=0,2023	<u>P=0,0084</u> <u>P=0,0003</u>
3 - O enunciador da gravação parece estar alegre ou triste?	<u>P=0,0185</u> <u>P=0,0023</u>	P=0,4943 <u>P=0,0009</u>	<u>P=0,0077</u> <u>P=0,0259</u>	P=0,0590 P=0,1880
4 - O enunciador parece estar tenso ou relaxado?	<u>P=0,0480</u> P=0,2295	P=0,9474 P=0,2169	<u>P=0,0244</u> P=0,0909	P=0,0652 P=0,5810
5 - O enunciador parece estar interessado ou desinteressado em relação à situação em que ele se encontra?	P=0,4927 P=0,4570	P=0,7444 P=0,1699	P=0,2394 P=0,6179	P=0,4912 P=0,7362
6 - O enunciador parece estar no controle de si mesmo e da situação ou parece não ter controle algum?	P=0,7011 <u>P=0,0128</u>	P=0,7744 P=0,1669	P=0,3715 <u>P=0,0031</u>	P=0,7242 P=0,0621
7 - O enunciador parece estar levando a situação a sério ou levando a situação na brincadeira?	P=0,813 <u>P=0,0132</u>	P=0,9478 P=0,4327	P=0,0564 <u>P=0,0195</u>	P=0,0673 <u>P=0,0062</u>
Total	3 5	0 2	3 4	1 2

Como pode ser observado na Tabela 04, verificaram-se 20 ocorrências de diferenças estatisticamente significativas (7 em *Arritmia*, 13 em *Cummings*). Quando se compara a distribuição dessas ocorrências, fica clara a tendência para que essas diferenças se concentrem

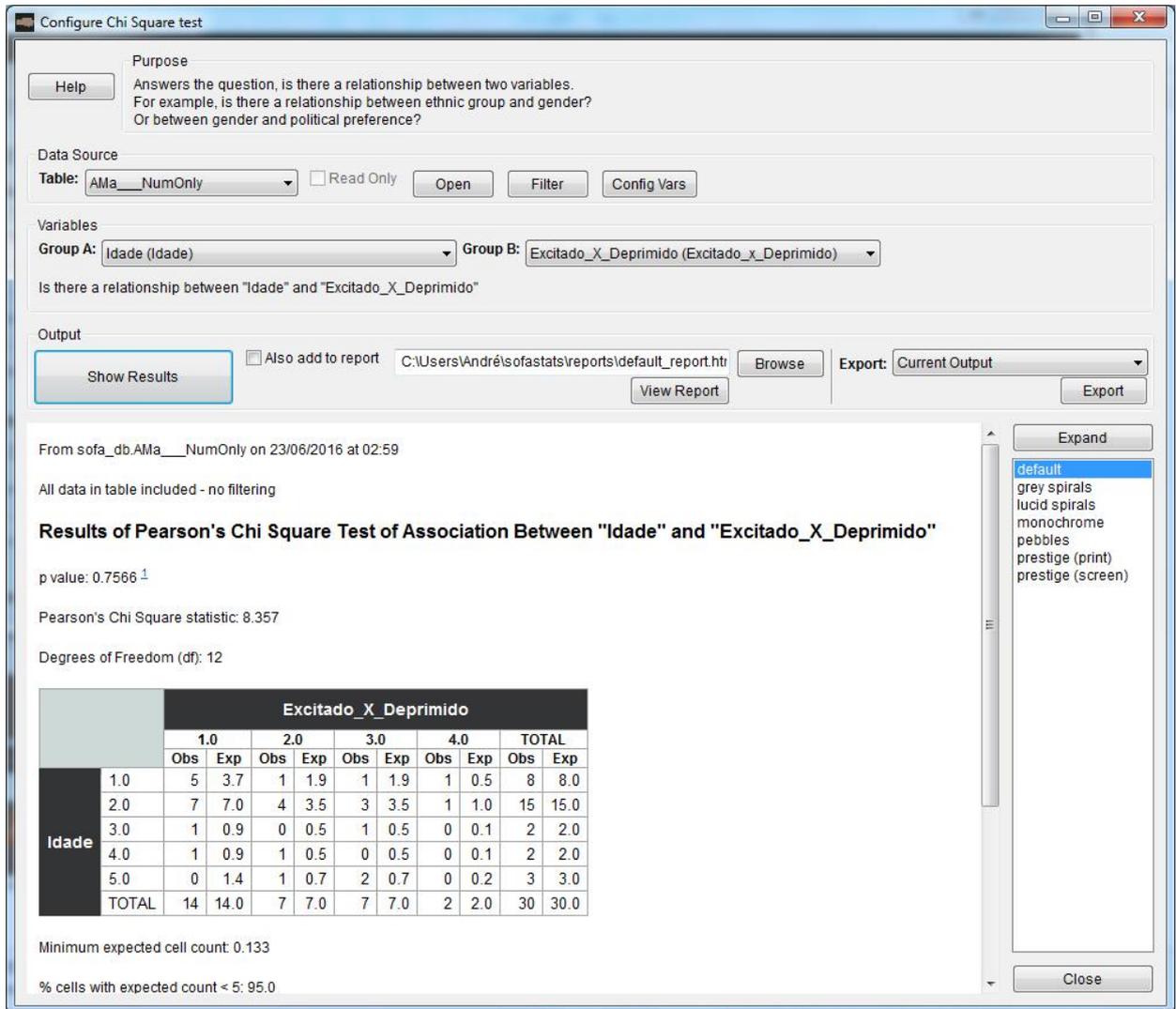
na comparação entre as amostras “Declamada” e “Maior” dos respectivos textos, e também na comparação entre as três versões desses textos. Considerando que cada uma das questões pode ser simplificada em uma escala bipolar de termos opostos, é possível mapear em quais CASOS foram percebidas diferenças de percepção e em qual situação.

Na comparação entre a versão declamada e uma das versões musicadas (com intervalo de terça maior ou menor), observam-se desvios consideráveis nas escalas Agradável/Desagradável (Questão 1), Cheio de Energia/Deprimido (Questão 2), Alegre/Triste (Questão 3), Tensão/Relaxamento (Questão 4), Controle/Falta de Controle (Questão 6) e A Sério/Na Brincadeira (Questão 7). Na comparação entre as duas versões musicadas, as quais se distinguem entre si pelo modo, houve desvios consideráveis nas escalas Cheio de Energia/Deprimido (2) e Alegre/Triste (3). Olhando para os resultados como um todo, nota-se que, em especial nas escalas das questões 2, 3 e 7, houve variações significativas tanto nas comparações referentes ao primeiro texto quanto nas referentes ao segundo.

No intuito de tirar o máximo proveito possível dos dados obtidos e evitar que possíveis correlações passassem despercebidas, buscou-se verificar se fatores como idade, escolaridade e gênero pareciam se relacionar com os valores obtidos nas respostas para as perguntas, uma série de aplicações da prova Chi-Quadrado (ou Chi Square, χ^2) foi realizada com o auxílio do software SOFA Statistics. A prova χ^2 mostra se os resultados de duas variáveis são independentes ou relacionadas. A ideia era verificar se havia relação entre fatores demográficos e os julgamentos em cada uma das questões. Um exemplo de aplicação da prova no programa pode ser visto na imagem, que pode ser considerada uma resposta à pergunta “Há alguma relação entre a idade do informante e sua avaliação sobre se o enunciador parece estar mais excitado e cheio de energia ou deprimido e sem ânimo?”.

Infelizmente, durante as aplicações, verificou-se que os resultados obtidos por χ^2 não podem ser considerados válidos: a prova χ^2 calcula e usa o número de ocorrências esperadas e o número de ocorrências observadas para verificar a probabilidade de ocorrência dos resultados obtidos no caso de H_0 ser verdadeira e, para tanto, nenhum dos valores de frequência esperada deve ficar abaixo de 1 (SIEGEL, p. 202). Um exemplo de tabela mostrando a relação entre valores esperados e observados para uma questão específica, de acordo com o gênero dos informantes, pode ser visto abaixo:

Figura 36: Aplicação da prova χ^2 no programa SOFA



Como se vê, há pelo menos onze células nas quais os valores esperados se encontram abaixo de 1 e, portanto, mitigam a validade dos resultados obtidos com χ^2 . Mesmo agrupando as variáveis (por exemplo, juntando as categorias “Não-binário” e “Outro/Prefiro não responder” ao verificar a influência do gênero, ou agrupando “Sim, canto” e “Sim, toco um instrumento” para analisar a influência da educação musical nas respostas), não foi possível evitar que pelo menos uma célula tivesse o valor esperado abaixo de 1 e, portanto, invalidasse os resultados de χ^2 .

Aplicou-se, portanto, uma prova alternativa a χ^2 , o Coeficiente de Correlação por Postos de Spearman (R), para verificar se há uma relação linear entre os escores obtidos em cada questão e os fatores demográficos.

Os resultados, para cada amostra, foram (lembrando que $\alpha = 0,05$):

4.1 Arritmia – Declamada

Idade: Não foram encontradas correlações

Envolvimento musical: Foram encontradas correlações nos polos Interessado x Desinteressado ($p=0,01818$) e A Sério x Na Brincadeira ($p < 0,001$).

Nível de Escolaridade: Não foram encontradas correlações

4.2 Arritmia – Maior

Idade: Não foram encontradas correlações.

Envolvimento musical: Foi encontrada correlação no polo A Sério x Na Brincadeira ($p=0,04092$)

Nível de Escolaridade: Não foram encontradas correlações.

4.3 Arritmia – Menor

Idade: Não foram encontradas correlações.

Envolvimento musical: Não foram encontradas correlações.

Nível de Escolaridade: Não foram encontradas correlações significativas.

4.4 Cummings – Declamada

Idade: Não foram encontradas correlações.

Envolvimento musical: Foram encontradas correlações nos polos Alegre x Triste ($p=0,02734$) e A Sério x Na Brincadeira ($p=0,04408$)

Nível de Escolaridade: Não foram encontradas correlações significativas.

4.5 Cummings – Maior

Idade: Não foram encontradas correlações significativas.

Envolvimento musical: Não foram encontradas correlações.

Nível de Escolaridade: Não foram encontradas correlações.

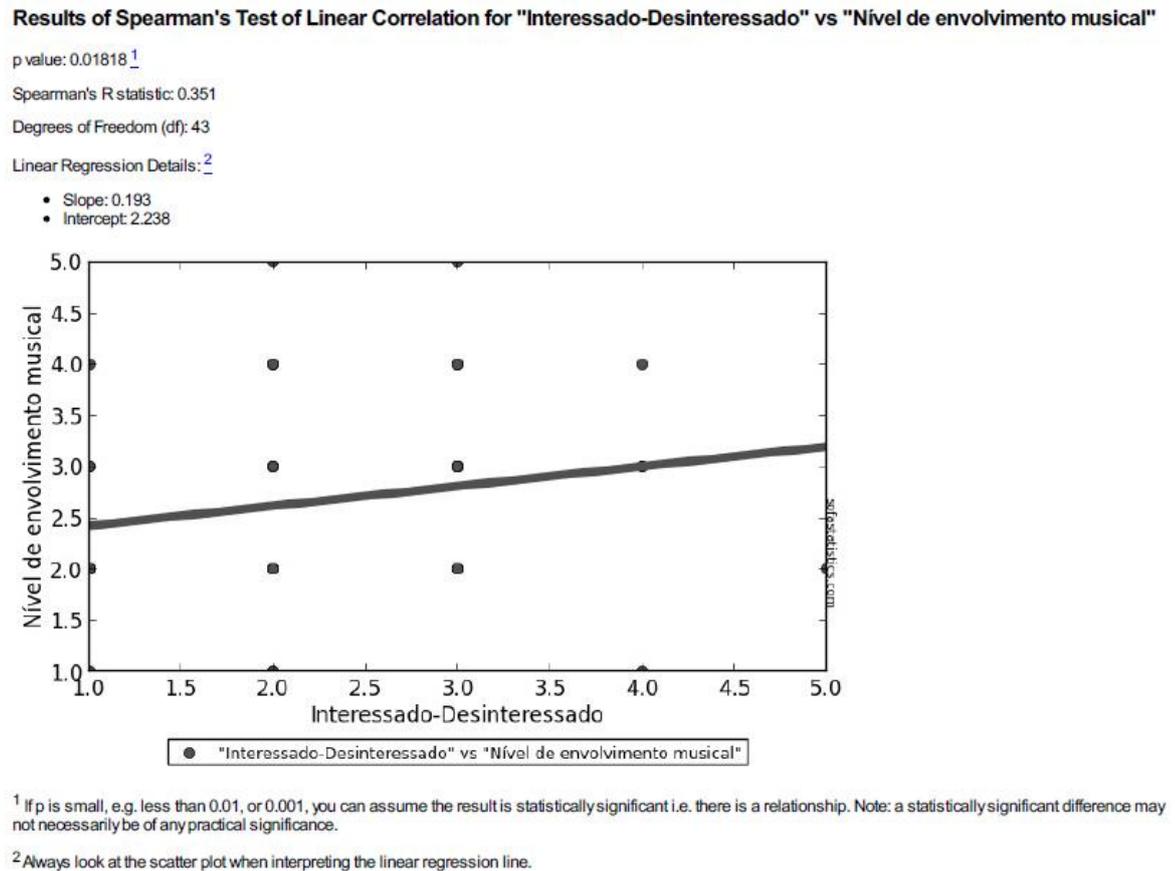
4.6 Cummings – Menor

Idade: Não foram encontradas correlações.

Envolvimento musical: Não foram encontradas correlações significativas.

Nível de Escolaridade: Foram encontradas correlações nos polos Alegre x Triste ($p=0,02494$) e Excitado x Deprimido ($p=0,01431$).

Figura 37: Exemplo de gráfico gerado na aplicação do Coeficiente de Correlação por Postos de Spearman, a partir da amostra referente à gravação do texto *Arritmia* sendo declamado



A tabela a seguir mostra em quais amostras essas correlações ocorreram, de acordo com os fatores demográficos (eixo Y, vertical) e as escalas nas quais ocorreram correlações (eixo X, horizontal).

Tabela 5: Ocorrências de correlações entre fatores demográficos (idade, envolvimento musical e escolaridade) e escalas bipolares (1 – 7)

Fatores demográficos x Escalas na qual há ocorrência de correlações								
	1	2	3	4	5	6	7	Total: Escala na qual há correlações (nº de ocorrências)
Idade	-	-	-	-	-	-	-	-
Envolvimento musical	-	-	Cd	-	Ad	-	Ad AM Cd	3 (1) 5 (1) 7 (3)
Escolaridade	-	Cm	Cm	-	-	-	-	2 (1) 3 (2)

Abreviaturas:

Ad – Arritmia, versão declamada

AM – Arritmia, versão Maior

Am – Arritmia, versão menor

Cd – Cummings, versão declamada

CM – Cummings, versão maior

Cm – Cummings, versão menor

Em suma, os resultados obtidos apontam para as seguintes conclusões:

- Parece haver, sim, influência da música na interpretação do texto. Essa influência parece estar mais no fato de o texto ser musicado do que no tom da composição, como se nota pela ocorrência das versões Ad, AM, Cd e Cm na tabela, mas não das versões Am e CM.
- A maior variação foi observada nos polos A sério x Na brincadeira (escala 7, com 3 ocorrências) e Alegre x Triste (escala 3, com 3 ocorrências).

- Dentre os fatores demográficos, o nível de envolvimento musical parece ser o que exerce maior influência sobre a percepção dos informantes (5 ocorrências), seguido da escolaridade (3 ocorrências).

5. DISCUSSÃO

Os resultados discutidos no final da seção anterior apontam para um fato curioso: mesmo tendo simulado uma composição em modo maior a partir da composição em modo menor, a partir da alteração do intervalo de terça de menor para maior, a diferença na avaliação entre as composições não foi tão significativa quanto a diferença entre qualquer uma das versões musicadas e a versão declamada. Isso pode parecer estranho, uma vez que composições em modo menor, assim como o próprio intervalo de terça menor isolado, são características de composições em geral interpretadas como mais tristes e melancólicas. Seria de se esperar, a princípio, que a houvesse uma grande diferença na percepção entre as duas gravações cantadas, especialmente no polo Alegre x Triste. Por que isso não aconteceu? Basicamente, porque a presença de intervalos de terça maior ou menor não é o bastante para atribuir uma tonalidade e um modo a uma composição. A mesma conclusão é apontada por Ferreira (1993, p. 28): “[...] o conjunto dos dados mostrou que a identificação de uma hierarquia tonal pode ser maior ou menor em função de outros fatores pertinentes ao contexto musical real [...]”.

De imediato, é importante mencionar que a ausência de um acompanhamento é um dos fatores que, provavelmente, atrapalhou a hierarquização tonal por parte dos ouvintes. Como não há outros instrumentos ou outras vozes acompanhando a música, não há formação de acordes ou progressão harmônica evidente. Isso já torna mais provável que a hierarquia dos intervalos, em si, já tenha sido relativizada na percepção dos informantes.

Outro fator pertinente é o contexto no qual as notas de grau III aparecem. A composição *Arritmia* foi escrita utilizando a escala de Mi menor, ou seja, o intervalo entre o primeiro grau (Mi) e o terceiro grau (Sol) é de uma terça menor. No entanto, não há, ao longo da canção, nenhum momento em que uma nota Mi seja seguida de uma nota Sol, há sempre pelo menos uma nota separando os (supostos) graus I e III da composição, prejudicando a hierarquização. Após a edição digital da composição para simular um tom de Ré maior, o problema se tornou ainda mais evidente. As poucas notas Rés ao longo da composição encontram-se, em sua maioria, nos tempos finais dos compassos, uma posição desfavorável,

já que em um compasso 4/4 o segundo e o quarto tempos são considerados fracos e, portanto, menos importante. Isso provavelmente impediu que a nota Ré fosse hierarquizada como tônica (grau I), como originalmente esperado.

Cummings também foi um caso problemático. Escrita em Lá menor natural, ao longo da composição há vários momentos em que a nota Lá (grau I) é seguida ou precedida da nota Dó (grau III), especialmente na segunda metade da composição. No entanto, muitas dessas ocorrências ocorreram em tempos fracos do compasso e em contextos menos pronunciados rítmica e melodicamente. Mesmo tendo sido posteriormente alterada digitalmente para simular um tom de Lá maior, com Fá, Sol e Dó sustentados, o mais provável é que a tonalidade não tenha sido interpretada como Lá maior, pelas razões expostas.

Os problemas levantados até agora reforçam o que foi discutido por Ferreira (1993): não basta apresentar intervalos maiores ou menores de uma determinada escala, é preciso que o contexto que engloba as notas contribua para a correta hierarquização dos graus.

Curiosamente, aqui se pode fazer uma comparação com o que acontece com a marcação de tonicidade no português brasileiro. Em *Acento e Ritmo* (1992), Massini-Cagliari discute os correlatos físicos (fonéticos) do acento (marcado fonologicamente) no português brasileiro, e conclui que não há um único correlato responsável pela marcação do acento, mas uma combinação de três aspectos acústicos, cada um com seu grau de importância:

Até aqui foram estudadas as principais marcas acústicas (correlatos físicos) do acento no português – variedade falada em Campinas, Brasil. Concluiu-se que as principais marcas do acento são, no nível lexical, uma maior duração na SÍLABA tônica, acompanhada (na maioria das vezes) de uma queda na intensidade na(s) SÍLABA(S) pós-tônica(s). À qualidade vocálica também foi atribuída grande responsabilidade na caracterização do acento. No nível frasal, concluiu-se que o acento é um resultado de padrões entoacionais específicos incidindo sobre SÍLABAS acentuadas no nível lexical – isto é, com as marcas acústicas de acento no nível lexical (duração, intensidade e qualidade vocálica) – constituindo também a “tônica” (no sentido de Halliday 1963, 1970) do padrão entoacional. (MASSINI-CAGLIARI, 1992, p. 31)

Mais adiante, Massini-Cagliari (1992, p. 38) reforça a presença dos três correlatos (duração, intensidade e qualidade vocálica) que contribuem com peso diferente na marcação do acento:

Diante dos fatos observados, pode-se concluir que o acento em português é uma proeminência atualizada foneticamente pela co-ocorrência de diversos fatores prosódicos.

No nível lexical, os principais correlatos do acento são (em ordem decrescente de importância): duração, intensidade e qualidade vocálica.

No nível frasal, o acento do português é caracterizado por uma variação do padrão entoacional que se sobrepõe a uma sílaba tônica em nível lexical. Em outras palavras, para que uma sílaba tônica em nível lexical o seja em nível frasal, é preciso que ela ocorra num determinado momento do padrão entoacional (relativo a todo o enunciado) em que esteja havendo variações relativamente grandes de F_0 em relação a todo o resto do enunciado.

Nota-se, assim, mais uma semelhança na organização das “gramáticas” da língua e da música tonal: assim como na língua, para a marcação da tonicidade, três fatores interagem no estímulo sonoro com peso maior ou menor, também na música, para a hierarquização dos graus e atribuição da tonalidade a uma composição, há também uma interação de fatores³⁵, entre eles, a altura melódica, a posição das notas dentro do compasso e a duração das notas.

Uma vez que a questão das diferenças nos intervalos de terça (maior ou menor) das composições não acarretou as diferenças de percepção esperadas a princípio, atribui-se a diferença nas avaliações a outros fatores:

- as linhas melódicas das composições não seguem os padrões entoacionais do português brasileiro: em *Arritmia*, há muitas variações bruscas na altura melódica e, em ambas as composições, é comum haver uma subida na altura melódica no final de frases afirmativas, ao contrário do que seria esperado para o português brasileiro;
- as pausas nas composições não coincidem totalmente com as pausas observadas nas gravações do texto sendo declamado;
- há muito mais variação de intensidade nas gravações cantadas do que na gravação declamada;
- a velocidade de emissão é bastante diferente quando o texto é declamado e cantado, sendo que as versões declamadas apresentam mais sílabas por minuto do que as versões cantadas.

Reconhece-se, assim, que o experimento deste trabalho não encontrou uma correlação direta entre intervalos de terça menor e tristeza, ao contrário dos trabalhos de Curtis e Bharucha (2010) e Schreuder (2006), que serviram de inspiração para esta pesquisa. Note-se, no entanto, que o resultado final é positivo. Concluimos que, de fato, as inevitáveis variações

³⁵ Para uma discussão mais aprofundada sobre os fatores que condicionam a hierarquização dos graus em música, ver Ferreira (1993).

de parâmetros acústicos interferem na percepção (e, por extensão, na interpretação) de um enunciado.

Também se espera que os dados e resultados apresentados levantem questionamentos acerca do senso comum e de algumas práticas cotidianas, inclusive dentro das salas de aula.

CONCLUSÕES

As relações entre música e língua, suas semelhanças, sobreposições e influências mútuas, foram o tema deste trabalho. Ao longo do texto, buscamos aduzir evidências de que música e língua, talvez por terem a mesma origem evolutiva, talvez por ambas lançarem mão do som para a expressão humana (mais provavelmente por ambos os motivos), estabelecem entre si uma relação bastante próxima. As duas formas de expressão humana mobilizam muitos recursos físicos e cognitivos em comum, e no que pese as diferenças entre as duas linguagens, parte considerável de suas “gramáticas” parece ser comum a ambas.

O trabalho de mestrado que ora se conclui apresentou vários desafios. À época do levantamento bibliográfico para a elaboração do projeto, no que pese a vasta produção relacionando língua, música e atitudes do falante, não conseguimos encontrar, no Brasil, trabalhos semelhantes a este, o que nos leva a crer que o trabalho aqui apresentado é inédito. Mesmo se considerando a produção acadêmica estrangeira, não encontramos experimentos semelhantes nem um modelo pronto a ser seguido. A consequência mais importante desta dificuldade é que o próprio levantamento bibliográfico se provou mais árduo do que se esperava: pesquisando por temas e palavras-chave, foi necessário construir aos poucos uma base sobre a qual trabalhar, sendo que, ao longo do processo, infelizmente foram várias as leituras que, apesar de tangenciarem o tema, não se mostraram úteis para o desenvolvimento da pesquisa. Também foi necessário criar uma metodologia de levantamento e análise de dados e testar diversas plataformas online até encontrar uma que satisfizesse nossas necessidades. As dificuldades encontradas, junto com o período relativamente curto do Mestrado (2 anos), impediram que atingíssemos a profundidade desejada. Não obstante, acreditamos que a realização deste trabalho abre precedentes para que outros pesquisadores desenvolvam pesquisas relacionadas, e pensamos que, no desenvolvimento de novas pesquisas sobre este tema, seus realizadores terão mais facilidade em evitar as armadilhas nas quais eventualmente incorremos e poderão aperfeiçoar a metodologia aqui testada. No mínimo, esperamos que este trabalho possa servir como uma base para aqueles que desejem

explorar as relações entre fala e música através de uma abordagem experimental, e que estes já comecem suas pesquisas mais bem munidos de referências do que nós.

Os resultados do teste de percepção integrante deste trabalho, apesar das ressalvas feitas na seção de Discussão, vão na mesma direção das conclusões de trabalhos mencionados ao longo desta dissertação. Ainda que uma relação direta entre os intervalos maiores e menores não possa ter sido estabelecida, devido a razões já discutidas anteriormente, é inegável que a transposição de um suporte (fala) para outro (música) acarretou diferenças na forma como o conteúdo semântico do texto foi percebido, no que se refere à sua carga emocional. Estes dados são mais uma prova da importância que elementos prosódicos - como a velocidade de fala, padrões rítmicos e entoacionais - exercem na comunicação.

Longe de colocar um ponto final na discussão, o trabalho aqui apresentado abre caminho para várias pesquisas e possíveis experimentos futuros. Baseando-se na metodologia desenvolvida para a pesquisa que ora se encerra, uma das possibilidades seria tentar quantificar a influência de parâmetros prosódicos. Por exemplo, poder-se-ia realizar um novo teste de percepção no qual um mesmo texto fosse gravado (declamado ou cantado) com o mesmo padrão entoacional/linha melódica, mas velocidade de fala/andamento distinto e, através de escalas bipolares, mensurar como os informantes percebem conteúdos emocionais expressos no texto, e o qual o valor da diferença de percepção que um aumento ou diminuição X na velocidade de fala/andamento acarreta. É oportuno frisar, também, a importância de que experimentos elencados nesta dissertação e realizados no exterior com falantes de outras línguas sejam replicados no Brasil com falantes de português brasileiro, para confirmar, refutar ou complementar as evidências já levantadas. Da mesma forma, seria interessante que o mesmo teste de percepção aqui aplicado fosse reproduzido em outros contextos linguísticos e culturais.

Outrossim, se podemos ter certeza de algo, é que a exploração das relações, dos limites e das ambiguidades entre música e língua ainda tem muito a oferecer.

REFERÊNCIAS BIBLIOGRÁFICAS

ALLEN, G. D. The place of rhythm in a theory of language. **Working Papers in Phonetics 10**. Phonetics Lab./Dept. of Linguistics/UCLA, 1967, p. 60-84.

ANDERSSON, B., ULVAEUS, B. Thank You for the Music. In: ABBA. **ABBA – The Album**. Polar Music International, 1977. 1 CD (40 minutos aprox..). Faixa 7.

BAGNO, M. **Gramática Pedagógica do Português Brasileiro**. São Paulo: Parábola Editorial, 2012.

BENFATTI, M. F. N. **Falando em música... Um ensaio sobre o papel dos fenômenos linguísticos em uma epidemiologia de representações musicais**. Dissertação (Mestrado em Linguística) - Universidade Federal do Paraná, Curitiba, 2010.

BENNET, R. **Uma Breve História da Música**. Rio de Janeiro: Jorge Zahar Ed., 1986.

BLACKING, J. The Structure of Musical Discourse: The Problem of the Song Text. **Yearbook of Traditional Music**. Vol. 14, 1982, p. 15-23.

CAGLIARI, L. C. **Elementos de Fonética do Português Brasileiro**. Teste (Livre Docente). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1981.

CAGLIARI, L. C. **Análise fonológica :introdução à teoria e à prática com especial destaque para o modelo fonêmico**. Campinas: Mercado de Letras, 2002

CHOMSKY, N. **Linguagem e Pensamento. Petrópolis:** Vozes, 1971.

CHOMSKY, N. **O conhecimento da língua: sua natureza, origem e uso.** Lisboa: Caminho, 1994.

CÓRDULA, M. S. M. **Entoação e sentidos: análise fonético-fonológica dos padrões entoacionais do português brasileiro e do inglês norte-americano no filme Shrek (2001).** São Paulo: Cultura Acadêmica, 2013.

COSTA, D. S. **A interface música e lingüística como instrumental metodológico para o estudo da prosódia do português arcaico.** Tese (Doutorado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, Universidade Estadual Paulista, 2010.

COOK, N.D.; FUJISAWA, T.; TAKAMI, K. Application of a Psycho-acoustical Model of Harmony to Speech Prosody. **Proceedings of Speech Prosody**, Nara, 2004, p. 147-150.

COUPER-KUHLEN, E. **An Introduction to English Prosody.** Londres: Edward Arnold, 1986.

COUTINHO, E., DIBBEN, N. Psychoacoustic cues to emotion in speech prosody and music. **Cognition and Emotion**, 27:4, 2013, p. 658-684.

CUMMINGS, E. E. **Poems :1923-1954.** Nova York: Harcourt, Brace & World, 1954.

CURTIS, M. E.; BHARUCHA, J. J. The Minor Third Communicates Sadness in Speech, Mirroring Its Use in Music. **Emotion**, v. 10, n. 3, 2010, p. 335-348. Disponível em <http://ase.tufts.edu/psychology/music-cognition/pdfs/Curtis&Bharucha2010Emotion.pdf>. Acessado em 07 Ago 2014.

DIETRICH, P. **Semiótica do Discurso Musical**. Uma discussão a partir das canções de Chico Buarque. Tese (Doutorado em Linguística). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2008.

ENGELMANN, A. **Os estados subjetivos: uma tentativa de classificação de seus relatos verbais**. São Paulo: Ática, 1978.

FALK, D. **Finding Our Tongues: Mothers, Infants & the Origins of Language**. Nova York: Basic Books, 2009.

FERREIRA, A. A. G. D. **A Hierarquia Perceptiva de Tons em Música Tonal e Atonal para Ouvintes Musicalmente Experientes e Inexperientes**. Tese (Doutorado em Psicologia). Instituto de Psicologia, Universidade de São Paulo, São Paulo, 1993.

FERREIRA, A. B. H. **Dicionário Aurélio básico da língua portuguesa**. São Paulo: Folha de São Paulo, 1995.

FRÜHHOLZ, S.; TROST, W.; GRANDJEAN, D. The role of the medial temporal limbic system in processing emotions in voice and music. In: **Progress in Neurobiology**, 123, 2014, p. 1- 17.

GEBARA, E. M. S. **Alguns aspectos da intonação no português**. Dissertação (Mestrado em Linguística). Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, 1976.

HALLIDAY, M. A. K., GREAVES, W. S. **Intonation in the Grammar of English**. Londres: Equinox Publishing Ltd., 2008.

HAYNES, B. Cammerton. In: **Grove Music Online**. Disponível em: <http://www.oxfordmusiconline.com/subscriber/article/grove/music/04666>. Acesso em 21 de março de 2017

JACKENDOFF, R.; LERDAHL, F.A **Deep Parallel Between Music and Language**. Bloomington: Indiana University Linguistics Club, 1980.

LENNON, J.; MCCARTNEY, P. The End. In: The Beatles. **Abbey Road**. Parlophone, 1969. 1 CD (47 minutos aprox.) Faixa 10.

LERDAHL, E.; JACKENDOFF, R. **A generative theory of tonal music**. Cambridge: The MIT Press, 1983.

LEVMAN, B. G. - The Genesis of Music and Language. **Ethnomusicology**, Vol. 36, No. 2. University of Illinois Press on behalf of Society for Ethnomusicology, 1992, p. 147-170.

LIEBERMAN, P., BLUMSTEIN, S. **Speech physiology, speech perception, and acoustic phonetics**. Cambridge: Cambridge University Press, 1988

LIEBERMAN, P. **On the Origins of Language: an introduction to the evolution of human speech**. Nova York: Macmillan Publishing Co., 1975.

LONDON, J. Rhythm. In: **Grove Music Online**. Oxford University Press. Disponível em: <http://www.oxfordmusiconline.com/subscriber/article/grove/music/45963>. Acesso em 21 de março de 2017.

MAMPE, B., FRIEDERICI, A. D., CHRISTOPHE, A., WERMKE, K. Newborn's Cry Melody Is Shaped by Their Native Language. In: **Current Biology** vol 19, n 23, 1994-1997, 2009, p. 1994-1997.

MASSINI-CAGLIARI, G. **Acento e ritmo**. São Paulo: Contexto, 1992.

_____ O Conceito de *Pé* como Unidade Rítmica: Trajetória. In: SCARPA, Ester M. (org) **Estudos de Prosódia**. São Paulo: Editora da Unicamp, 1999. (Coleção Pesquisas)

_____ **A música da fala dos trovadores**: estudos de prosódia do português arcaico, a partir das cantigas profanas e religiosas. Tese (Livre Docência em Fonologia) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2005.

_____ Das cadências musicais para o ritmo lingüístico: Uma análise do ritmo do Português Arcaico, a partir da notação musical das Cantigas de Santa Maria. **Revista da ABRALIN**, v. 7, n. 1, p. 9-26, jan./jun. 2008a.

_____ Interface Fonologia-Poesia-Música: Uma análise do ritmo lingüístico do Português Arcaico, a partir da notação musical das *Cantigas de Santa Maria*. **Estudos Lingüísticos**. São José do Rio Preto: GEL, 2008b. v. 37, n.1, p. 9-20.

_____ Contribuição para a análise do ritmo lingüístico das cantigas medievais profanas e religiosas a partir de uma interface Música-Lingüística. In: CONGRESSO DA ASSOCIAÇÃO INTERNACIONAL DE LUSITANISTAS, 9, 2009a, Madeira. **Lusofonia: Tempo de Reciprocidades**: atas. Porto: Edições Afrontamento, 2011. Vol. I, p. 41-53.

_____ Fonologia Histórica: estudando o ritmo lingüístico a partir de uma interface Lingüística-Música. In: CONGRESSO INTERNACIONAL DA ABRALIN, n. 6, João Pessoa. **VI Congresso Internacional da Abralín**: anais. João Pessoa: Idéia, 2009b, p. 1683- 1692.

_____ From Musical Cadences to Linguistic Prosody: How to Abstract Speech Rhythm of the Past. In: PARTRIDGE, J. (ed.) **Interfaces in language**. Newcastle upon Tyne: Cambridge Scholars, 2010a. p. 113-134.

_____ Sobre as relações entre proeminências musicais e poéticas na poesia trovadoresca profana e religiosa. In: PIRES, A. D.; FERNANDES, M. L. O. **Matéria de poesia. Crítica e criação**. Araraquara: FCL-UNESP Laboratório Editorial; São Paulo: Cultura Acadêmica, 2010b. p. 47-66.

_____ Análise do texto e da música da Cantiga de Santa Maria 70: Discutindo uma questão de edição. In: ENCONTRO INTERNACIONAL DE ESTUDOS MEDIEVAIS, 8, 2011, Cuiabá. **VIII Encontro Internacional de Estudos Medievais. As múltiplas expressões da Idade Média: Filosofia, Artes, Letras, História e Direito: Anais.** Cuiabá: EDUFMS, 2011

_____ **A música da fala dos trovadores:** Desvendando a prosódia medieval. São Paulo: Editora Unesp Digital, 2015.

MASSINI-CAGLIARI, G.; CAGLIARI, L. C. De sons de poetas OU Estudando fonologia através da poesia. **Revista da ANPOLL**, São Paulo, n. 5, p. 77-105, jul./dez., 1998.

_____—Fonética. IN MUSSALIM, F., BENTES, A. C. **Introdução à Linguística: domínios e fronteiras.** São Paulo: Cortez, 2001. vol. 1, p. 105-146.

MEDEIROS, B. R. **Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro.** Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2002.

MORTON, J. B., TREHUB, S. E. Children's Understanding of Emotion in Speech. **Child Development**, Vol. 72, Nº 3 (May - Jun. 2001). p. 834-843

NESPOR, M.; VOGEL, I. **Prosodic phonology.** Dordrecht: Foris, 1986.

PATEL, A.D., IVERSEN, J.R., ROSENBERG, J.C. Comparing the rhythm and melody of speech and music: The case of British English and French. **J. Acoust. Soc. Am**, nº 119, 3034–3047, 2006.

PEREIRA, J. C. R. **Análise de Dados Qualitativos: Estratégias Metodológicas para as Ciências da Saúde, Humanas e Sociais**. São Paulo: Editora da Universidade de São Paulo, 1999.

RINGER, A. L. Melody. **Grove Music Online**. Oxford University Press. Disponível em: <http://www.oxfordmusiconline.com/subscriber/article/grove/music/18357>. Acesso em 21 de março de 2017.

ROCHA, N. A. Articulações teórico-práticas em curso de formação de Língua Estrangeira: ensinando espanhol por meio de músicas. In: LEITE, Y. U. F. L. et al. **Articulações entre teoria e prática**. Araraquara: Junqueira e Marin Editores, 2012. Vol. II, p. 1-10

SANTOS, J. F.; PAULUK, I. **Proposições para o ensino de língua estrangeira por meio de músicas**. 2008. Disponível em: <http://www.diaadiaeducacao.pr.gov.br/portals/pde/arquivos/752-4.pdf>. Acesso em: 07 de fevereiro de 2017.

SCHAFFER, R. M. **O ouvido pensante**. São Paulo: Fundação Editora da Unesp, 1991.

SCHREUDER, M. **Prosodic Processes in Language and Music**. Tese (Doutorado em Letras). Rijksuniversiteit Groningen, Groningen, 2006.

SIEGEL, S. **Estatística não-paramétrica para as ciências do comportamento**. São Paulo: McGraw-Hill do Brasil, 1975

SMITH, L. D.; WILLIAMS, R. N. Children's Artistic Responses to Musical Intervals. **The American Journal of Psychology**, vol. 112, Nº 3 (Autumn, 1999). p. 383-410.

STEINBEINS, N.; KOELSCH, S. Shared Neural Resources between Music and Language Indicate Semantic Processing of Musical Tension-Resolution Patterns. In: **Cerebral Cortex**, vol. 18(5), p. 1169-1178, 2008. Disponível em: <http://cercor.oxfordjournals.org/content/18/5/1169.full> Acessado em 20 de Agosto de 2014.

TATIT, L. **Musicando a semiótica**: Ensaios. São Paulo: Annablume, 1997.

TATIT, L., LOPES, I. **Elos de melodia e letra**: análise semiótica de seis canções. Cotia: Ateliê Editorial, 2008.