

UNIVERSIDADE FEDERAL DO PARANÁ

BEATRIZ PIRES SANTANA

PROPRIEDADES INTERMODULARES: UMA DISCUSSÃO E EXPERIMENTAÇÃO EM MÚSICA E
LINGUAGEM

CURITIBA
2013

BEATRIZ PIRES SANTANA

PROPRIEDADES INTERMODULARES: UMA DISCUSSÃO E EXPERIMENTAÇÃO EM MÚSICA E
LINGUAGEM

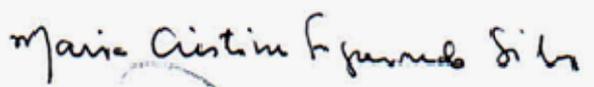
Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Linguística, no Curso de Pós-Graduação em Letras, Setor de Ciências Humanas, Letras e Artes da Universidade Federal do Paraná.

Orientadora: Profa. Dra. Maria Cristina Figueiredo Silva

CURITIBA
2013

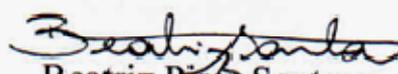


Ata sexcentésima décima sétima, referente à sessão pública de defesa de dissertação para a obtenção de título de mestre a que se submeteu a mestranda **BEATRIZ PIRES SANTANA**. No dia dez de outubro de dois mil e treze, às quatorze horas, na sala 1005B, 10.º andar, no Edifício Dom Pedro I, do Setor de Ciências Humanas da Universidade Federal do Paraná, foram instalados os trabalhos da Banca Examinadora, constituída pelos seguintes Professores Doutores: **MARIA CRISTINA FIGUEIREDO SILVA**, Presidente, **DANILO RAMOS** e **MAXIMILIANO GUIMARÃES**, designados pelo Colegiado do Curso de Pós-Graduação em Letras, para a sessão pública de defesa de dissertação intitulada: "PROPRIEDADES INTERMODULARES: UMA DISCUSSÃO E EXPERIMENTAÇÃO EM MÚSICA E LINGUAGEM", apresentada por **BEATRIZ PIRES SANTANA**. A sessão teve início com a apresentação oral da mestranda sobre o estudo desenvolvido. Logo após a senhora presidente dos trabalhos concedeu a palavra a cada um dos Examinadores para as suas arguições. Em seguida, a candidata apresentou sua defesa. Na sequência, a Professora **MARIA CRISTINA FIGUEIREDO SILVA** retomou a palavra para as considerações finais. Na continuação, a Banca Examinadora, reunida sigilosamente, decidiu pela aprovação da candidata. Em seguida, a senhora Presidente declarou **APROVADA** a candidata, que recebeu o título de **Mestre em Letras**, área de concentração **Estudos Linguísticos**, devendo encaminhar à Coordenação em até 60 dias a versão final da dissertação. Encerrada a sessão, lavrou-se a presente ata, que vai assinada pela Banca Examinadora e pela candidata. Feita em Curitiba, no dia dez de outubro de dois mil e treze. xxxxxxxxxxxxxxxxxxxxxx


Dr.ª Maria Cristina Figueiredo Silva


Dr. Danilo Ramos


Dr. Maximiliano Guimarães


Beatriz Pires Santana



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS HUMANAS
COORDENAÇÃO DO CURSO DE PÓS GRADUAÇÃO EM LETRAS

PARECER

Defesa de dissertação da mestranda BEATRIZ PIRES SANTANA para obtenção do título de **Mestre em Letras**.

Os abaixo assinados MARIA CRISTINA FIGUEIREDO SILVA, DANILO RAMOS e MAXIMILIANO GUIMARÃES arguiram, nesta data, a candidata, a qual apresentou a dissertação:

“PROPRIEDADES INTERMODULARES: UMA DISCUSSÃO E EXPERIMENTAÇÃO EM MÚSICA E LINGUAGEM”

Procedida a arguição segundo o protocolo que foi aprovado pelo Colegiado do Curso, a Banca é de parecer que a candidata está apta ao título de **Mestre em Letras**, tendo merecido os conceitos abaixo:

Banca	Assinatura	APROVADA Não APROVADA
MARIA CRISTINA FIGUEIREDO SILVA		aprovada
DANILO RAMOS		APROVADA
MAXIMILIANO GUIMARÃES		aprovada

Curitiba, 10 de outubro de 2013

Teresa Cristina Wachowicz
Coordenadora

AGRADECIMENTOS

À minha família, por seu apoio durante o desenvolvimento deste trabalho e por seu apoio sempre.

A todos os meus amigos, sem os quais o processo de desenvolvimento deste trabalho teria sido muito mais árduo.

À minha orientadora, Maria Cristina Figueiredo Silva, por ter aceitado me orientar neste trabalho desafiador e por tê-lo feito da melhor forma que eu poderia esperar.

Ao professor Maximiliano Guimarães, por ter me desafiado diversas vezes durante o desenvolvimento deste trabalho e sem o qual este trabalho teria um formato completamente diferente.

Ao linguista Andrew Nevins, por ter se disposto a conversar sobre meu projeto pessoalmente, por ter sido muito solícito e por ter generosamente me dado excelentes ideias para a execução dos experimentos, mesmo sabendo que meu trabalho desafia um trabalho feito por ele, em parceria com Ansgar Endress.

Ao professor da UFPB José Ferrari Neto, por ter gentilmente me auxiliado na análise estatística dos dados dos experimentos.

Aos linguistas Ruan Mariano e Izabel Seara, pela disponibilização de *scripts* criados no programa Praat, em cujos comandos me baseei para criar os *scripts* dos experimentos aqui relatados.

Ao LEA (Laboratório de Estatística Aplicada da UFPR) e, mais particularmente, à professora Fernanda Bühner Pizzato e à aluna Andressa Avendaño Forbellone, pela realização da análise estatística dos experimentos.

Aos membros da banca examinadora, Maximiliano Guimarães, Danilo Ramos e Caetano Galindo (suplente), pela disponibilidade de participar.

A todos os 60 estudantes da UFPR que aceitaram participar voluntariamente dos experimentos.

RESUMO

O número de estudos sobre as relações entre música e linguagem vem crescendo muito nas últimas décadas. Estes estudos estão se tornando cada vez mais difundidos no ramo da psicologia cognitiva, da neurociência, da musicologia e mesmo da linguística. Um estudo realizado pelos linguistas Andrew Nevins e Ansgar Endress enquadra-se nessa área de estudo, sob a perspectiva da linguística formal e experimental. O presente trabalho é uma resposta ao trabalho supracitado, numa tentativa de enquadrar sua discussão teórica no âmbito da interdisciplinaridade entre música e linguagem.

O estudo de Nevins e Endress apresenta os resultados de dois experimentos. O primeiro deles é um experimento que envolve linguagem e que busca verificar as generalizações relacionadas a um fenômeno cujos dados aparentam ser desordenados. O segundo deles é um experimento que envolve música e que busca verificar se as generalizações encontradas para explicar o fenômeno linguístico em questão estão também disponíveis para o processamento musical dos seres humanos ou se são generalizações específicas à linguagem. Segundo os autores, os resultados dos experimentos supracitados apontam para a hipótese de que música e linguagem diferenciam-se no aspecto investigado.

O presente trabalho discute os experimentos dos autores e traz novos resultados advindos da recondução dos mesmos experimentos, com modificações motivadas por uma discussão teórica musical prévia que fundamenta o delineamento do experimento que envolve música. Além disso, propõe-se uma discussão teórica sobre a relação entre música e linguagem prévia que fundamenta a comparação entre nossas capacidades linguística e musical e, conseqüentemente, a comparação entre os dois experimentos.

Palavras-chave: Música. Linguagem. Linguística formal. Gerativismo. Cognição.

ABSTRACT

The number of studies on the relations between music and language has been a growing field of research over the past decades. These studies are becoming increasingly widespread in the branch of cognitive psychology, neuroscience, musicology, and even linguistics. A study carried out by Andrew Nevins and Ansgar Endress falls into this area of research, from the standpoint of formal and experimental linguistics. The present study is a response to the above mentioned study, aiming to frame its theoretical discussion into the scope of the interdisciplinary approach of music and language.

The study by Nevins and Endress presents the results of two experiments. The first one is a linguistic experiment which aims to verify the generalizations related to a phenomenon whose data appear to be chaotic. The second one is a musical experiment which aims to verify whether or not the generalizations found to account for the concerned linguistic phenomenon are too available to musical processing, or whether or not the generalizations are specific to language. According to the authors, the experiments' results point out to the hypothesis that music and language differ in the aspect under investigation.

The present work discusses the authors' experiments, and brings new results derived from a new conduction of these experiments, with changes motivated by a musical theoretical discussion that grounds the design of the musical experiment. Moreover, we propose a theoretical discussion on the relations between music and language that grounds the comparison between the human capacity for music and language and, consequently, the comparison between the two experiments.

Key-words: Music. Language. Formal linguistics. Generative theory. Cognition.

SUMÁRIO

1. QUESTÕES PRELIMINARES SOBRE MÚSICA E LINGUAGEM.....	9
1.1 INTRODUÇÃO.....	9
1.2 A TEORIA GERATIVA... ..	13
1.2.1 ... da linguagem	13
1.2.2 ... da musica tonal	17
1.3 O FORMALISMO DA TEORIA... ..	19
1.3.1 ... da linguagem	20
1.3.2 ... da música tonal.....	25
1.4 MÚSICA, LINGUAGEM E SUA INTERDISCIPLINARIDADE.....	28
1.4.1 Aquisição.....	30
1.4.2 Patologias relacionadas à música e à linguagem.....	33
1.4.3 Ritmo.....	35
1.5 CONSIDERAÇÕES FINAIS.....	37
2. DESCRIÇÃO E PROBLEMATIZAÇÃO DE NEVINS E ENDRESS	
(2007).....	39
2.1 INTRODUÇÃO.....	39
2.2 DESCRIÇÃO DOS EXPERIMENTOS DE NEVINS E ENDRESS	
(2007).....	39
2.2.1 Línguas secretas.....	39
2.2.2 Experimento linguístico.....	44
2.2.3 Experimento musical.....	48
2.3 DISCUSSÃO.....	51
2.3.1 Uma discussão musical sobre os estímulos musicais.....	52
2.3.2 Uma discussão de interface sobre os estímulos musicais.....	55
2.3.3 Katz e Pesetsky (2011)	57
2.3.4 Por que um alinhamento <i>bottom-up</i> ?	67
2.3.4.1 A atribuição de estrutura métrica.....	68
2.3.5 Os dois vieses de Nevins e Endress (2007).....	70
2.3.5.1 A falsa inversão.....	71
2.3.5.2 A não-ambiguidade dos dados.....	73

2.3.5.3 E se a falsa inversão não for falsa?	76
3. EXPERIMENTOS	79
3.1 INTRODUÇÃO.....	79
3.2 DIRETRIZES DOS EXPERIMENTOS.....	80
3.3 METODOLOGIA DOS EXPERIMENTOS MUSICAIS.....	82
3.3.1 Experimento 1.....	84
3.3.1.1 Estímulos.....	84
3.3.1.2 Participantes.....	88
3.3.1.3 Equipamento.....	89
3.3.1.4 Procedimento.....	89
3.3.1.5 Resultado.....	93
3.3.2 Experimento 2.....	94
3.3.2.1 Estímulos.....	94
3.3.2.2 Participantes.....	95
3.3.2.3 Equipamento.....	95
3.3.2.4 Procedimento.....	95
3.3.2.5 Resultado.....	96
3.3.3 Experimento 3.....	96
3.3.3.1 Estímulos.....	96
3.3.3.2 Participantes.....	98
3.3.3.3 Equipamento.....	98
3.3.3.4 Procedimento.....	98
3.3.3.5 Resultado.....	99
3.3.4 Experimento 4.....	99
3.3.4.1 Estímulos.....	99
3.3.4.2 Participantes.....	100
3.3.4.3 Equipamento.....	101
3.3.4.4 Procedimento.....	101
3.3.4.5 Resultado.....	104
3.4 METODOLOGIA DO EXPERIMENTO LINGUÍSTICO.....	105
3.4.1. Experimento 5.....	105
3.4.1.1 Estímulos.....	105

3.4.1.2 Participantes.....	107
3.4.1.3 Equipamento.....	107
3.4.1.4 Procedimento.....	108
3.4.1.5 Resultado.....	112
4. DISCUSSÃO.....	113
4.1 INTRODUÇÃO.....	113
4.2 ANÁLISE QUALITATIVA DOS RESULTADOS.....	114
4.3 DISCUSSÃO.....	116
4.3.1 Sobre música.....	116
4.3.2 Sobre linguagem.....	121
4.3.3 Sobre música e linguagem.....	121
4.4 CONSIDERAÇÕES FINAIS.....	124
REFERÊNCIAS.....	126
APÊNDICE A – TABELAS DE DADOS.....	131
APÊNDICE B – FICHA INFORMATIVA DOS PARTICIPANTES.....	141
ANEXO	142

1. QUESTÕES PRELIMINARES SOBRE MÚSICA E LINGUAGEM

1.1 Introdução

A música e a linguagem são duas faculdades humanas cuja comparação sempre esteve presente no percurso histórico da reflexão e do estudo sobre as diversas facetas que as envolvem.¹ A tendência a essa aproximação entre música e linguagem se torna compreensível após a observação de que música e linguagem apresentam algumas propriedades em comum.

Primeiramente, a música e a linguagem parecem ser faculdades *exclusivamente* humanas. Há quem possa refutar tal afirmação após uma ligeira reflexão sobre o canto dos pássaros e sobre o sistema comunicativo das abelhas, entre outros. Entretanto, há uma extensa literatura que discute as propriedades fundamentais que parecem fazer da música e da linguagem produtos exclusivamente humanos, sendo os sistemas comunicativos de outras espécies sistemas qualitativamente distintos, por não compartilharem das mesmas propriedades fundamentais (FITCH; HAUSER, 2004; TOMASELLO, 2003; PINKER, 1987; PEPPERBERG, 2002; CROSS, 2001; HAUSER; MCDERMOTT, 2003; PATEL, 2008, cap. 7). Por exemplo, a função de cada um desses sistemas é distinta entre as espécies: o canto dos pássaros é um comportamento reprodutivo determinado biologicamente (CATCHPOLE; SLATER, 1995), e não um ato estético volitivo, como é a música humana; e o sistema comunicativo das abelhas tem única e exclusivamente a função de comunicar às outras abelhas a direção e a distância do alimento encontrado (von FRISCH, 1953), não tendo, portanto, toda a complexidade e a versatilidade de uso da linguagem humana. Essa e outras propriedades nos fazem crer que a música e a linguagem sejam atividades

¹ Como exemplo de literatura da aproximação comparativa entre música e linguagem, temos Quintiliano que, no século II, em suas *Instituições Oratórias* (obra teórica e pedagógica sobre a retórica), dedica um capítulo à comparação entre a arte do orador e a expressividade da música. Temos o teórico musical Joachim Burmeister que, no século XVII, escreve sobre o sistema retórico-musical, inteiramente baseado no sistema retórico-linguístico. Temos também Bernstein (1976), que busca, na música, um correlato dos componentes fonológico, sintático e semântico da linguagem. Temos, ainda, Medeiros (2006), que aponta para a existência de 32 termos linguísticos que são emprestados dos estudos musicais. Ainda, o personagem Adrian Leverkühn do romance *Dr. Fausto* de Thomas Mann iguala a música à fala (MANN, 2000, p. 229). Para além de referências presentes na literatura, é possível perceber como a aproximação entre música e linguagem está presente no imaginário popular através de lugares-comuns como "a música é a linguagem universal" ou "a música é a linguagem da alma".

manifestas apenas pela espécie humana, incomparáveis aos sistemas comunicativos que algumas espécies de animais apresentam.

Em segundo lugar, música e linguagem são não só atividades *exclusivamente* humanas, como são também *universais* dentro da espécie, dado que todas as culturas existentes até então conhecidas manifestam tais habilidades, independentemente de quais outros aspectos da cultura estão ausentes (NETTL, 2000).

Outra propriedade em comum entre música e linguagem é o fato de se caracterizarem pela organização de unidades sonoras no tempo. Essa característica dá origem a uma gama de outras propriedades que música e linguagem necessariamente compartilham, algumas delas derivadas de aspectos externos, como a necessidade de lidar com as propriedades acústicas do som (altura, intensidade, duração e timbre) e outras derivadas da maneira como o cérebro organiza *inputs* sonoros, como a organização rítmica das unidades sonoras que formam complexos musicais e linguísticos, ou seja, a existência de padrões de alternância entre unidades mais proeminentes e unidades menos proeminentes, em que se baseiam os estudos da fonologia métrica na linguística (ver LIBERMAN; PRINCE, 1977) e da estrutura métrica na música (ver COOPER; MEYER, 1960).

Na linguagem, a organização rítmica se manifesta claramente dentro das palavras. Quando são formadas por duas ou mais sílabas, uma delas deve ser tônica (mais proeminente) e todas as outras devem ser átonas (menos proeminentes): "médico", por exemplo, é uma palavra cuja primeira sílaba ("mé") é mais proeminente do que as outras ("di" e "co"), o que é, neste caso, representado na escrita através do sinal de acentuação. Entretanto, o ritmo na linguagem não se limita a apenas esse aspecto. Dentro das palavras, quando há diversas sílabas átonas adjacentes, gera-se um padrão de alternância entre sílabas átonas menos proeminentes e sílabas átonas mais proeminentes que, embora mais proeminentes do que as outras sílabas átonas, ainda são menos proeminentes do que as sílabas tônicas: "candelabro", por exemplo, tem a sílaba "la" como sílaba tônica, sendo todas as outras átonas; dessas átonas, a mais proeminente é a sílaba "can" (que possui, assim, o chamado "acento secundário"), o que gera um padrão rítmico binário na palavra. Adicionalmente, quando diversas palavras são concatenadas em uma frase, as sílabas tônicas dessas palavras também formam um padrão de alternância entre sílabas tônicas mais e menos proeminentes: em "o médico comprou um candelabro", por exemplo, a sílaba tônica de "candelabro" é mais

proeminente do que a sílaba tônica de "médico". Tais alternâncias são fruto não só de princípios cognitivos gerais sobre padronização rítmica, como o Princípio de Alternância Rítmica – que exige que entre dois pulsos fortes sucessivos intervenha pelo menos um e no máximo dois pulsos fracos (SELKIRK, 1984, p. 12; LIBERMAN, 1975, p. 313) – como também de regras que interagem com os diversos componentes da linguagem, como a sintaxe (ver SELKIRK, 1984, cap. 4).

Na música, também temos padrões de alternância entre unidades mais e menos proeminentes, o que, inclusive, justifica a representação gráfica padrão da música tonal ocidental através da divisão de grupos de sons em *compassos*, que são unidades estabelecidas pela maneira como a superfície musical alterna suas proeminências: a unidade inicial de um compasso é, por *default*, a unidade mais proeminente do compasso todo. Assim como na linguagem, a organização rítmica da música não se limita a tal aspecto único e apresenta diversas outras camadas: quando um compasso apresenta várias unidades menos proeminentes, elas geram entre si uma alternância de proeminências – por exemplo, em compassos quaternários, o terceiro tempo é, também por *default*, mais proeminente do que o segundo e o quarto, embora menos proeminente do que o primeiro.² Quando concatenados diversos compassos, a primeira unidade de cada um deles também gera um padrão de alternância de proeminências com a primeira unidade dos outros compassos, como em uma frase musical simples de quatro compassos, em que o primeiro tempo do primeiro compasso é, no geral, mais proeminente do que o primeiro tempo dos outros compassos. Também como a linguagem, tais alternâncias obedecem a regras que interagem com as diversas propriedades presentes no estímulo musical, como a estabilidade harmônica³ (ver LERDAHL; JACKENDOFF, 1983, p. 90).

² O padrão do compasso quaternário é uma estrutura com alternância binária em dois níveis distintos: no primeiro nível, entre os tempos um e dois e entre os tempos três e quatro, temos a alternância forte-fraco. No segundo nível, tomando-se os dois tempos fortes do primeiro nível, que são os tempos um e três, também temos a alternância forte-fraco. Na música tonal ocidental, esta costuma ser a regra: a periodicidade dos pulsos é reforçada nível a nível (LERDAHL; JACKENDOFF, 1983, p. 20), o que é representado através de compassos no sistema de escrita musical.

³ A estabilidade harmônica é uma propriedade fundamental ao processamento musical e é a base para a descrição da estrutura abstrata que subjaz à superfície da música. Nessa estrutura abstrata, os itens hierarquicamente mais importantes são aqueles que são harmonicamente mais estáveis. Na percepção musical, a estabilidade está relacionada à sensação de completude e à percepção de pontos de referência. Além disso, os pontos de maior estabilidade são sentidos como a base da música, o que permite que, mantendo-se a base, variações, ornamentos e instrumentações modifiquem a superfície musical, sem que se perca a identidade da música.

Por fim, outra característica que une música e linguagem é o fato de ambas caminharem muitas vezes juntas, dado que, embora nem toda cultura produza música com instrumentos musicais, toda cultura pratica o canto (LEVITIN, 2008). Assim, letra e música, canto e poesia são sobrepostos em uma única camada estética, aumentando, assim, a proximidade superficial que a muitos instiga entre essas duas atividades humanas. Embora esta propriedade muito provavelmente não se refira a uma proximidade natural ou cognitiva da música e da linguagem, ela é um elemento a mais para motivar a tendência à especulação intelectual, já muito comum, entre música e linguagem.

Não distante de tantos outros que se viram instigados pela relação entre a música e a linguagem, o presente trabalho investiga, por meio de um fenômeno particular, possíveis aspectos em comum entre a música e a linguagem. O fenômeno em questão é a manipulação do ordenamento de unidades linguísticas, o que é manifesto na linguagem natural por meio das chamadas *línguas secretas* (ver 2.2.1). Encontrando-se os princípios que subjazem a esse fenômeno linguístico, será investigada a possibilidade de sua manifestação em estímulos musicais. Tal investigação tem um caráter cognitivo, ou seja, baseia-se em teorias sobre as capacidades musical e linguística do homem, cujos principais tópicos de pesquisa são as maneiras como o homem adquire essas capacidades, o surgimento dessas capacidades no percurso evolutivo da espécie humana e as formas como ouvimos, processamos, entendemos e produzimos estímulos musicais e linguísticos. Antes de imergirmos na pesquisa aqui relatada, cabe a esta introdução elucidar os pressupostos teóricos que alicerçaram a condução do trabalho. A descrição das teorias será simplificada e fragmentada, dado que a base para o presente trabalho consiste em teorias de duas áreas distintas, cujo conhecimento prévio não pode ser pressuposto do leitor que, em sua grande maioria, é familiarizado apenas com uma delas. A teoria linguística descrita não se aterá a nenhum modelo teórico específico; a descrição consta de elementos de diferentes modelos do programa gerativo-transformacional (CHOMSKY, 1957, 1981, 1995b; CHOSMY; LASNIK, 1993) e da fonologia suprasegmental gerativa (SELKIRK, 1984), devido ao caráter introdutório da descrição. A teoria musical descrita, por outro lado, segue unicamente o modelo de Lerdahl e Jackendoff (1983).

1.2 A teoria gerativa...

1.2.1 ...da linguagem

O estudo da linguagem pode ser realizado sob diversas perspectivas investigatórias: teológica (exemplificada pelo mito da Torre de Babel), filosófica (filosofia da linguagem), histórica (filologia), social (sociolinguística) ou matemática (lógica). No caso da teoria gerativa da linguagem – à qual também nos referiremos como "gerativismo" ou "gramática gerativa" – a perspectiva é cognitiva, ou seja, estuda a linguagem como um produto da mente humana. Essa teoria foi fundada por Noam Chomsky nos anos 50, o que revolucionou o estudo da linguagem que, até então, olhava para seu objeto sob uma perspectiva sobretudo social.

Para uma teoria que estuda a linguagem sob uma perspectiva mental, a principal questão que se coloca sobre seu objeto de estudo é: o que um falante *sabe* quando sabe falar uma língua? Ou seja, de que mecanismos cognitivos um falante lança mão quando faz uso da linguagem? É bastante incontroverso o fato de que a capacidade linguística do ser humano não se resume a um inventário mental que contenha todas as sentenças possíveis da língua nativa do falante. Dito de outro modo, saber falar uma língua não é conhecer todas as sentenças dessa língua. Primeiramente porque, a todo momento, os falantes de uma língua produzem e entendem sentenças que jamais tinham ouvido antes, o que mostra que deve haver algum componente *gerativo* permitindo a geração dessas novas sentenças. E em segundo lugar porque, devido a uma propriedade da língua denominada *recursividade*, o número de sentenças possíveis de uma língua é infinito, conforme mostra o exemplo (1) abaixo, o que torna impossível para um falante ter ouvido todas elas e ter todas elas armazenadas em sua memória.

(1)

- a. O João comeu um bolo.
- b. A Maria disse que [o João comeu um bolo].
- c. A Joana acha que [a Maria disse que [o João comeu um bolo]].
- d. O Pedro perguntou se [a Joana acha que [a Maria disse que [o João comeu um bolo]]].
- e. O Luiz afirmou que [o Pedro perguntou se [a Joana acha que [a Maria disse que [o João comeu um bolo]]]].

...

A progressão mostrada acima é, em princípio, infinita. Sempre é possível tomar a sentença anterior como *input* para o acréscimo de uma estrutura que contenha um sujeito (como "O Pedro"), um verbo que permite como complemento uma sentença (como "dizer") e um conectivo (como "que") que conecte a sentença já existente com esses elementos novos; e a sentença formada pode sempre ser um novo *input* para a formação de uma nova sentença com os elementos supracitados. O que nos impediria de construir uma sentença infinita desse tipo são fatores externos à linguagem, como tempo e memória. Entretanto, nossa capacidade linguística permite que uma sentença infinita seja formada, utilizando-se recursivamente o mesmo mecanismo descrito acima. Se o conhecimento linguístico humano não pode ser a simples memorização de todas as sentenças de sua língua, seria plausível hipotetizar, portanto, que deve haver em nossa memória um inventário com um número *finito* de unidades linguísticas – de palavras, por exemplo – que o falante seleciona para formar as infinitas sentenças de sua língua.

Entretanto, apenas um conjunto finito de unidades a serem concatenadas ainda não é suficiente para caracterizar a capacidade linguística, pois dado um subconjunto qualquer de palavras, nem todas as combinações entre elas são possíveis; e os falantes são capazes de reconhecer sequências ruins de palavras e sequências boas de palavras. Por exemplo, a sequência de palavras "O Pedro viu um balão no céu" é uma sentença bem-formada da nossa língua, mas a sequência "Balão no viu o um Pedro céu" não é uma sentença boa da língua e todo falante é capaz de reconhecer isso. Da mesma forma, embora "João ama Maria" e "Maria ama João" sejam duas sentenças boas da língua e sejam construídas a partir das mesmas palavras, o significado das duas sentenças é diferente. Existem regras e restrições regendo a geração de sentenças da língua, que é a *gramática* da língua. Assim, o principal objetivo da teoria gerativa é, primeiramente, descrever a gramática de uma língua, ou seja, descrever o conjunto de regras que seja suficientemente poderoso para gerar *todas* as sentenças dessa língua e ao mesmo tempo suficientemente restritivo para bloquear a geração de sequências que não formam sentenças da língua.

Outra questão que uma teoria da linguagem de caráter cognitivo deve colocar sobre seu objeto é: descobrindo-se o que um falante *sabe* quando sabe falar uma língua (ou seja, tendo em mãos a descrição da gramática dessa língua), de onde vem esse conhecimento? Existem dois extremos que podem ser hipotetizados ao nos colocarmos

essa questão. O primeiro deles é que esse conhecimento é inteiramente advindo de estímulos externos e o segundo deles é que o conhecimento linguístico é inteiramente inato. A seguir discutiremos essas duas hipóteses.

Com o desenvolvimento das ciências cognitivas, hoje é pouco discutível que a primeira hipótese deva ser descartada. Muitas evidências apontam para o fato de que os seres humanos não nascem uma *tábula rasa*, sem um "conteúdo mental embutido", isto é, não é mais possível defender que todo o conhecimento surja da experiência e da percepção (PINKER, 2002; para visão contrária, ver HOWE; DAVIDSON; SLOBODA, 1998). No caso da linguagem, particularmente, uma forte evidência de que não é possível que as crianças aprendam a falar sua língua materna apenas através da experiência, seja por imitação e reforço ou por estímulo e resposta, é o fato de que as crianças são capazes de formular regras a respeito da língua que estão adquirindo, o que é visível na produção linguística da criança. Por exemplo, qualquer um que já teve contato com uma criança em fase de aquisição já a viu produzindo estruturas como "se eu *sesse* você" ou "eu *fazi* um desenho". As formas em itálico não são formas presentes no português adulto e, portanto, não é por imitação que as crianças as produzem, já que não ouvem ninguém as produzindo. Quando um adulto ouve tais produções de uma criança, sua reação em geral é corrigi-la repetindo a ela a mesma frase, mas com a forma da gramática adulta ("se eu *fosse* você" e "eu *fiz* um desenho"); portanto, não é por estímulo e resposta que as crianças produzem essas formas. E, não obstante, as crianças permanecem produzindo essas formas por algum tempo. Ao analisar a estrutura morfológica dessas formas verbais infantis, vemos que o que está acontecendo é que as crianças estão formulando regras a respeito da estrutura morfológica dos verbos de sua língua e aplicando-as não apenas nos contextos adequados, mas também em verbos que não obedecem a essas regras (verbos irregulares). "Ser" e "fazer" são verbos da classe da segunda conjugação (ou seja, que acabam em *-er*, sendo *-e-* a vogal temática e *-r* o sufixo de infinitivo), assim como os verbos "comer", "beber", "ler", "receber", etc. Para conjugar um verbo da segunda conjugação na primeira pessoa do singular do pretérito imperfeito do subjuntivo, a regra é retirar o sufixo de infinitivo e inserir o sufixo "-sse": "se eu *come-sse*", "se eu *bebe-sse*", "se eu *le-sse*", "se eu *recebe-sse*" e assim por diante. Por esse motivo, a criança produz a forma *se-sse* nesse contexto; pois ela está aplicando a regra geral a um contexto que, em realidade, é irregular. O mesmo ela faz com o verbo "fazer". Para conjugar um verbo da segunda conjugação na primeira pessoa do singular

do pretérito perfeito do indicativo, a regra é retirar a vogal temática e o morfema de infinitivo e inserir o sufixo *-i*: "eu *com-i*", "eu *beb-i*", "eu *l-i*", "eu *receb-i*" e assim por diante. Por esse motivo, a criança produz a forma *faz-i*. O que a criança faz é deduzir uma regra geral a partir dos dados que ouve e aplicar essa regra em todos os contextos que satisfazem a descrição estrutural pertinente (respeitando informação de tempo, modo, aspecto, pessoa, número, etc.) – e isso ela faz sem ter sido ensinada a fazê-lo. Se a criança não foi explicitamente instruída a fazer essa análise morfológica e se todas as crianças apresentam o mesmo comportamento perante a língua, esses dados demonstram que essa capacidade é de alguma forma dada biologicamente (seja essa uma capacidade geral para análise de quaisquer estruturas, seja essa uma capacidade específica à decodificação de estruturas linguísticas).

Quanto à segunda hipótese, a respeito de a linguagem ser inteiramente inata, essa é mais facilmente descartada, à simples menção do fato de que existem diversas línguas diferentes no mundo. As diferenças entre as línguas naturais não incidem apenas na forma diferente das palavras, ao contrário do que a tradução da sentença "o menino ama a menina" para o inglês, por exemplo, nos faria pensar. Essa sentença dita em inglês é "*the boy loves the girl*", e cada palavra da sentença em português corresponde à palavra que ocupa a mesma posição da sentença em inglês: *the* = o, *boy* = menino, *loves* = ama, *the* = a, *girl* = menina. Entretanto, se for feita uma tradução palavra a palavra dessa mesma sentença para uma língua como o pemon, que é uma língua caribe, seu significado será de que é a menina quem ama o menino, dado que, nessa língua, a ordem objeto-verbo-sujeito é mais comum do que a ordem sujeito-verbo-objeto⁴ (cf. de ARMELLADA, 1999). Assim, entre línguas diferentes, existem diferenças não só na forma das palavras (isto é, de natureza lexical), como também na ordem das palavras (de natureza sintática) e ainda nos sons utilizados (fonológica) e na maneira como se formam as palavras (morfológica). Dado que cada criança adquire a língua que é falada a seu redor, torna-se óbvio o fato de que o que desencadeia a aprendizagem da

⁴ Em realidade, a língua pemon apresenta marcação de caso. Em sentenças transitivas (que apresentam sujeito e objeto), o sujeito recebe o sufixo *-da* e ao objeto não se acresce nenhum material fonológico (o que algumas teorias linguísticas traduziriam como presença de um morfema zero). Portanto, obedecendo-se a morfologia dessa língua, os elementos sujeito, verbo e objeto poderão ser identificados quanto à sua função, independentemente da ordem em que sejam apresentados, já que morfológicamente todos eles revelam sua função na sentença (embora, é claro, algumas ordens soem artificiais para os falantes da língua, dado que existem as ordens canônicas, e assim outras ordens ou são impossíveis ou servem para enfatizar algum elemento da sentença). Não obstante, o problema da tradução palavra a palavra se mantém, já que uma tradução desse tipo ignora a noção de sujeito e, portanto, falharia em adicionar à palavra "menino" o sufixo *-da*.

linguagem em uma criança são os estímulos a que elas são expostas na infância (que são chamados de *dados linguísticos primários*). Logo, a linguagem também não pode ser inteiramente inata.

Não podendo a linguagem ser nem inteiramente inata e nem inteiramente adquirida através da experiência, deve haver alguma estrutura mental geneticamente determinada para aquisição da linguagem que seja comum a todos os seres humanos e que seja suficientemente abstrata para poder se manifestar de diversas formas distintas. O objetivo da teoria gerativa é, sobretudo, descrever o estado inicial da faculdade linguística do ser humano e a maneira como esse estado inicial vai se modificando através da exposição a estímulos linguísticos. O estado inicial da faculdade linguística é denominado *gramática universal* (GU) e restringe as possibilidades lógicas para estruturas linguísticas. Assim, supõe-se que todas as línguas humanas existentes e possíveis apresentem propriedades em comum que são determinadas pela gramática universal, que é um dispositivo inato. A essas propriedades em comum entre as línguas, dá-se o nome de *universais linguísticos*.⁵ A busca por esses universais, através da observação e comparação entre diversas gramáticas diferentes, é uma tarefa que tem o potencial de contribuir muito para o entendimento do estado inicial do processo de aquisição linguística do ser humano.

1.2.2 ...da música tonal

Assim como o estudo da linguagem, o estudo da música também pode ser realizado sob diversas perspectivas investigatórias. Na história dos estudos musicais, há uma grande e diversificada literatura de métodos analíticos de descrição de peças musicais (ver COOK, 1987, cap. 1) – a chamada *análise musical*, que consiste na observação de padrões de estruturas, harmonias e melodias e na "tradução" desses padrões em conceitos teóricos. Através desses sistemas analíticos, é possível extrair os princípios que subjazem e que regem um idioma musical particular. Encontrar os princípios por trás da construção de peças musicais de um idioma musical e,

⁵ A abordagem teórica designada Princípios e Parâmetros (CHOMSKY, 1981; CHOMSKY; LASNIK, 1993) busca dar conta das semelhanças e da diversidade entre as línguas do mundo através da postulação de um dispositivo linguístico que contém princípios, que são propriedades comuns a todas as línguas, e parâmetros, que são propriedades a que se pode atribuir valores diferentes, sendo que a atribuição de diferentes valores a diferentes parâmetros é o que explicaria a diversidade entre as línguas. Simplificadamente, por exemplo, a posição do núcleo dos sintagmas pode ser tanto inicial como final nas línguas. No português, a posição é inicial e, no japonês, a posição é final. Isso ocorre porque cada uma dessas línguas designa um valor diferente ao parâmetro relacionado à posição do núcleo no sintagma.

principalmente, verificar qual(is) desses princípios é(são) compartilhado(s) por outros idiomas musicais pode revelar muito sobre a maneira como nossa mente entende e processa música.

Com forte base no sistema analítico criado pelo teórico musical Heinrich Shenker no fim do século XIX e início do século XX, Lerdahl e Jackendoff (1983) fundam a chamada *teoria gerativa da música tonal*, cujo nome remete à teoria gerativa da linguagem, descrita na seção anterior (1.2.1). O motivo pelo qual Lerdahl e Jackendoff dão esse nome à teoria é que os autores adotam a natureza formal e o interesse psicológico da teoria linguística em sua teoria musical. Da mesma forma que a teoria descrita na seção anterior, a teoria musical aqui descrita se coloca a questão sobre o que um ouvinte experiente de um idioma musical *sabe* quando ouve, entende e aprecia uma peça musical de um idioma musical. A teoria gerativa da música tonal tem como objetivo a descrição da intuição musical de um ouvinte experiente através de um sistema de regras e princípios capazes de gerar todas as melodias do idioma musical tonal e, ainda, descartar sequências sonoras que não pertençam a tal idioma.⁶

Como dito no início deste capítulo, a música também parece ser uma faculdade única e universalmente humana. Quando nos deparamos com a expressão "ouvinte experiente", tendemos a atribuir essa expressão somente àqueles que foram musicalmente instruídos, que estudaram algum instrumento musical e que tocaram em concertos. Certamente, existe uma grande diferença entre aqueles que estudaram e aqueles que não estudaram música, como, por exemplo, a capacidade de ler partituras, a capacidade de nomear elementos musicais e de conversar sobre música fazendo uso de uma terminologia específica. Entretanto, mesmo os indivíduos mais leigos em música possuem um conhecimento musical implícito que os permite gostar de uma música, não gostar de outra, se emocionar com um trecho musical, diferenciar estilos musicais e perceber quando uma música foi executada com erros. Embora não tenham sido treinados explicitamente e não sejam capazes de utilizar um jargão musical, os leigos também dispõem dos mecanismos necessários para adquirir a gramática de um idioma

⁶ Como bem observado por Katz e Pesetsky (2011), embora a teoria de Lerdahl e Jackendoff se diga "gerativa", ela é antes "analisadora" do que "gerativa". Isso porque seu sistema de regras não gera melodias, ao contrário da teoria linguística, cujo sistema de regras gera sentenças da língua. O sistema de regras da teoria musical toma qualquer superfície musical e dá a ela uma estrutura de acordo com a maneira como a música tonal é estruturada, mas não é capaz de tomar uma sequência sonora e classificá-la como pertencente ou não ao idioma tonal. Ainda que esse seja um problema relativamente grande a uma teoria que se diz "gerativa", não discutiremos essa questão a fundo, dado que não compromete a discussão presente neste trabalho.

musical e, após expostos a dados desse idioma, tornam-se ouvintes experientes e passam a "entender música", nesse sentido tácito. O ouvinte experiente de um idioma musical não é diferente do falante nativo de uma língua: uma pessoa não alfabetizada, comparável ao leigo em música, tem um conhecimento implícito sobre sua língua, que a torna capaz de entender e produzir as sentenças de sua língua, mesmo que não saiba (explicitamente) o que é um verbo ou um substantivo e não tenha a capacidade de falar metalinguisticamente. Por isso, a teoria de Lerdahl e Jackendoff procura dar conta das intuições musicais de *todos* os ouvintes experientes, sejam eles leigos ou não, assim como a teoria gerativa da linguagem procura dar conta das intuições linguísticas de *todos* os falantes nativos, sejam eles alfabetizados ou não. Por este motivo, doravante, quando falarmos sobre a capacidade musical do ser humano e sobre ouvintes experientes, estaremos nos referindo ao conhecimento musical geral tácito, que é compartilhado por todos os indivíduos que são expostos a estímulos musicais na infância, independentemente de terem recebido treinamento musical explícito ou não (da mesma forma que um falante nativo de uma língua é qualquer indivíduo que foi exposto a estímulos linguísticos dessa língua na infância, independentemente de ter recebido treinamento explícito ou não).

Assim como acontece com a aquisição da linguagem, o entendimento musical tal qual descrito no parágrafo acima é adquirido espontaneamente, sem treinamento explícito, o que também parece demonstrar que existe algum viés cognitivo que possibilita a aquisição da gramática musical após exposição a estímulos externos. Assim, a teoria gerativa da música, através da descrição de estruturas musicais dos idiomas musicais naturais, tem o intuito de buscar o estado inicial da nossa capacidade musical, que pode se manifestar através de *universais musicais*, ou seja, propriedades em comum a todos os idiomas musicais naturais.

1.3 O formalismo da teoria...

Conforme visto na seção anterior, a teoria gerativa da linguagem e a teoria gerativa da música compartilham o mesmo viés teórico e as mesmas perguntas dirigidas aos seus respectivos objetos. Como a teoria musical inspirou-se na teoria linguística, alguns formalismos das duas teorias também são compartilhados, embora, pelo fato de terem sido desenvolvidas independentemente, acabem tendo uma arquitetura geral

bastante distinta. Nas subseções abaixo traçaremos a arquitetura de cada uma das teorias, apontando suas semelhanças e suas diferenças.

1.3.1 ...da linguagem

Uma das contribuições da teoria gerativa para os estudos da linguagem e sua computação é a afirmação de que as palavras dentro de uma sentença não se relacionam linearmente, embora o *output* seja uma sequência linear de sons e palavras; os elementos de uma sentença relacionam-se hierarquicamente. Assim, existem palavras que, embora na cadeia da fala não estejam adjacentes, são interdependentes. Um exemplo clássico de Chomsky (1957, p. 22) é a relação entre as palavras "se" e "então" (do inglês *if* e *then*, respectivamente) e entre as palavras "ou" e "ou" (do inglês *either* e *or*, respectivamente), como vemos em (2a) e (2b) abaixo. Não seria possível trocar, por exemplo, a palavra *então* da sentença (2a) pela palavra *ou*, pois isso geraria o resultado agramatical em (2c).⁷

(2)

- a. *Se* você não for à festa, *então* eu também não vou.
- b. *Ou* você come o bolo, *ou* você toma o sorvete.
- c. **Se* você não for à festa, *ou* eu também não vou.

O componente responsável pela relação estrutural entre termos da oração é a sintaxe, que é o estudo dos princípios e processos que presidem à construção de sentenças em línguas particulares (CHOMSKY, 1957, p. 11). O recurso visual que se consagrou na literatura gerativa para a representação da estrutura das sentenças são os diagramas arbóreos, como mostra a Figura 1.

⁷ A agramaticalidade da sentença em (2c) é representada pelo asterisco que a precede, recurso difundido na literatura linguística.

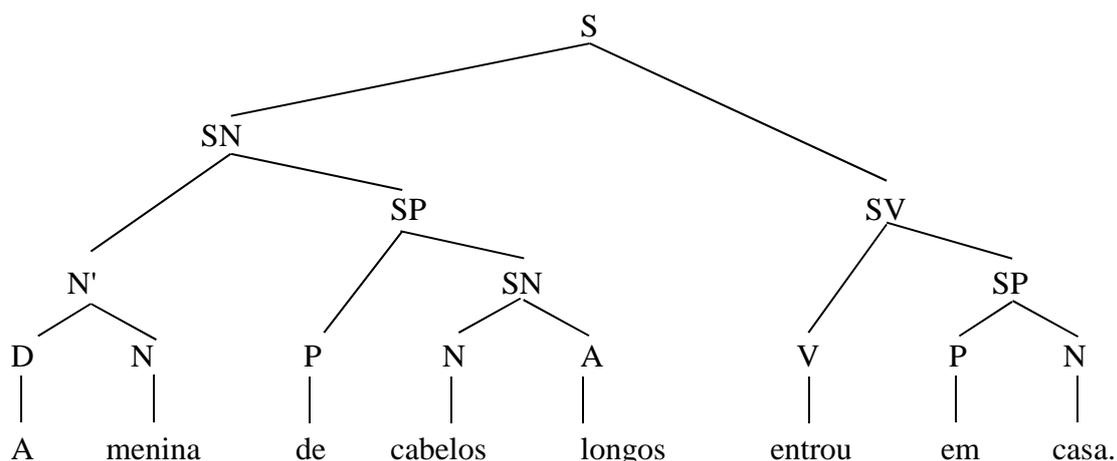


FIGURA 1: DIAGRAMA ARBÓREO SIMPLIFICADO PARA A SENTENÇA "A MENINA DE CABELOS LONGOS ENTROU EM CASA".

No diagrama acima, cada nó recebe um rótulo. À parte os nós terminais, que recebem como rótulo as próprias palavras da sentença, cada galho recebe um rótulo abstrato, que pode ser de três níveis distintos: SX, X' ou X (sendo x uma variável). O nível SX confere a um agrupamento de palavra o *status* de sintagma, que é um agrupamento de palavras que desfruta de certas propriedades na sentença. A expressão "em casa", por exemplo, forma um SP (sintagma preposicional) que tem um comportamento relativamente independente dentro da sentença; podemos, por exemplo, movimentá-lo ao início da sentença em uma construção clivada, formando "Foi em casa que a menina de cabelos longos entrou", o que não poderia ser feito com uma sequência de palavras presente na sentença que não forma um sintagma: "*Foi menina de que a cabelos longos entrou em casa". O nível X' (lê-se *x-barra* ou *x-linha*) é um nível intermediário, em que ainda não se tem um sintagma completo, e tampouco uma única palavra.⁸ O N' mais à esquerda, por exemplo, une "a" com "menina", mas isso ainda não forma um SN (sintagma nominal), pois o SN completo, nesta sentença, é "a menina de cabelos longos", como podemos ver pela clivagem apenas do N': "*Foi a menina que de cabelos longos entrou em casa". Já o nível X é o nível diretamente ligado ao nó terminal, o núcleo que atribui a uma palavra sua categoria. D é "determinante", N é "nome" (substantivo), P é "preposição", V é "verbo", A é "adjetivo" e assim por diante.

O diagrama da Figura 1 é o resultado final de uma série de operações que concatena elementos de dois em dois, que é uma operação denominada *merge*. A cada

⁸ A relação entre ser uma projeção X' e conter mais do que uma, mas não todas as palavras de que o sintagma é formado, não é tão trivial, até porque existem sintagmas formados de uma só palavra. Mantemos essa relação no texto para fins didáticos.

operação de *merge*, um dos dois elementos concatenados é tido como núcleo do complexo formado pela operação. A união, por exemplo, do determinante (D) "a" com o nome (N) "menina", forma um complexo cujo núcleo é "menina"; por isso rotula-se o galho resultante dessa operação de N', e não, por exemplo, de D'. Além da união entre dois elementos terminais, *merge* também pode operar sobre um elemento terminal com um elemento que já é fruto de uma operação anterior de *merge* (como a união da preposição (P) "de" com o sintagma nominal (SN) "cabelos longos") ou ainda de dois elementos não-terminais (como a união do N' "a menina" com o SP "de cabelos longos").

O que esse diagrama nos diz, dentre outras coisas, é que a sequência de palavras "a menina de cabelos longos" forma uma unidade (um *constituente* ou *sintagma*), cujo núcleo é a palavra "menina" (representado pela relação de dominância⁹ entre SN (sintagma nominal) e N (nome)) e que a sequência de palavras "entrou em casa" também forma uma unidade, cujo núcleo é a palavra "entrou" (representado pela relação de dominância entre SV (sintagma verbal) e V (verbo)). Essas duas unidades formadas relacionam-se diretamente. Ou seja, seus núcleos relacionam-se diretamente, embora sejam formadas por mais palavras. Assim, "menina" e "entrou", nessa estrutura, são palavras correlacionadas, embora não estejam linearmente adjacentes. A palavra "cabelos" ou mesmo a expressão "cabelos longos" está linearmente mais próxima do verbo do que a palavra "menina". Ainda assim, a estrutura devidamente nos mostra que é "menina" que se relaciona com o verbo, e não "cabelos" (caso contrário, o verbo estaria no plural, concordando com "cabelos", e não com "menina", o que não acontece). Poderíamos colocar ainda mais material entre o núcleo do sintagma nominal "menina" e o núcleo verbal "entrou" e a relação entre essas duas palavras continua sendo exatamente a mesma, já que ela não depende da adjacência, e sim da relação hierárquica: "a menina de cabelos longos com quem eu estava conversando ontem quando você me ligou entrou em casa".

O componente sintático é aquele que computa a estrutura das sentenças, como a representação vista na Figura 1. Para haver uma estrutura, é necessário haver antes um inventário de palavras ou de morfemas com seus conceitos e suas funções, que serão as unidades que a sintaxe concatenará na estrutura arbórea, formando os sintagmas e as

⁹ A relação de dominância caracteriza-se pelo percurso unicamente descendente de galhos, em que o galho superior domina o galho inferior.

sentenças. A esse inventário de palavras, que é armazenado na memória, dá-se o nome de *léxico*. Quando temos uma sentença formada, há duas informações simultâneas e diferentes que precisam ser computadas: a do som e a do significado. Dito de outro modo, existe uma cadeia de sons (composta de fonemas, sílabas, ritmos, contornos entoacionais) que precisa ser decodificada e essa cadeia sonora apresenta um significado, que também deve ser decodificado. Além do léxico e da sintaxe, portanto, é necessário também haver um componente que compute os sons e um componente que compute o significado. Ao primeiro, dá-se o nome de *componente fonológico* ou *prosódico* (ou ainda *forma fonológica*) e ao segundo, *componente semântico* (ou *forma lógica*). O componente fonológico toma a estrutura sintática como *input*, transforma-a em uma sequência linear e confere a essa sequência linear fonemas, contorno entoacional e ritmo através de regras e princípios que regem esse componente da gramática. Já o componente semântico "lê" a estrutura sintática, dando a ela um significado baseado nos elementos individuais retirados do léxico e na maneira como a sintaxe os concatena. Levando-se tudo isso em conta, a arquitetura da gramática de uma língua tem a forma apresentada na Figura 2.¹⁰

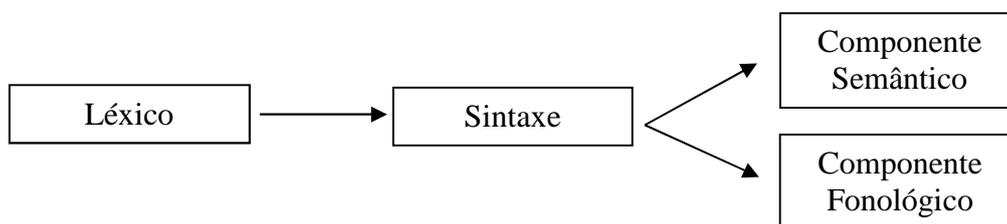


FIGURA 2: ARQUITETURA DA GRAMÁTICA

Acima fizemos referência ao componente fonológico de maneira geral como aquele componente que lida com as propriedades sonoras da linguagem. Numa formulação mais precisa, o componente fonológico pode ser subdividido em suas duas dimensões: a segmental, que lida com os fonemas de uma língua (ver CHOMSKY; HALLE, 1968), e a suprasegmental, que lida com aspectos do som como a entoação, a acentuação e o ritmo, o que é chamado de *prosódia* (ver SELKIRK, 1984; NESPOR; VOGEL, 1986; LIBERMAN, 1975; LIBERMAN; PRINCE, 1977). As teorias rítmicas,

¹⁰ O quadro geral mostrado na Figura 2 é mais condizente com o modelo GB (do inglês *government and binding*) da teoria gerativa, que é o modelo de Chomsky (1981). Nesse modelo, a sintaxe seria ainda subdividida em dois níveis: *estrutura profunda* e *estrutura superficial*. Nos modelos mais atuais, entretanto, como de Chomsky (1995), a sintaxe não chega a ser um nível desse quadro; ela é apenas o "caminho" (uma sucessão de processos derivacionais) para se chegar do léxico até a forma lógica.

na linguagem, costumam ter duas representações distintas, a depender do modelo teórico: o diagrama arbóreo, conforme visto acima para a representação sintática, mas codificando outro tipo de informação (Figura 3), e a grade métrica, presente na Figura 4.

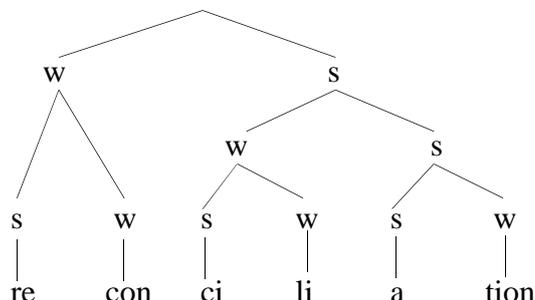


FIGURA 3: DIAGRAMA ARBÓREO PARA REPRESENTAR O RITMO DA PALAVRA INGLESA "RECONCILIATION". FONTE: LIBERMAN; PRINCE, 1977.

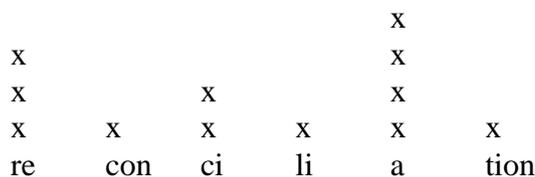


FIGURA 4: GRADE MÉTRICA PARA REPRESENTAR O RITMO DA PALAVRA INGLESA "RECONCILIATION". FONTE: SELKIRK, 1984, p. 66.

Na árvore (Figura 3), os nódulos rotulados *s* (do inglês *strong*) são mais proeminentes ritmicamente do que os nódulos rotulados *w* (do inglês *weak*); na grade métrica (Figura 4), as sílabas com uma coluna maior de *x* são mais proeminentes ritmicamente do que aquelas com uma coluna menor. Assim, embora as representações sejam visualmente diferentes, a informação codificada nos dois diagramas é basicamente a mesma: existe um padrão binário de alternância rítmica entre as sílabas, gerando um padrão forte-fraco-forte-fraco-forte-fraco, sendo a sílaba *-a-* a mais forte dentre todas as sílabas da palavra *reconciliation*. A informação adicional obtida através de uma representação como a da Figura 3 é a informação sintagmática: as duas sílabas iniciais *re-* e *-con-*, após serem ligadas uma à outra, unem-se a *-ciliation*. A unidade *-ciliation*, por sua vez, é um complexo formado por *-cili-* e *-ation*. Embora o exemplo aqui apresentado seja da árvore fonológica de uma única palavra, esses diagramas são também usados para representar a prosódia das sentenças. A informação sintagmática na representação fonológica das sentenças faz interface com a informação sintagmática da sintaxe, que tem um papel na atribuição de proeminências às unidades da sentença. Essa informação sintagmática não é formalmente representada nas grades métricas, embora a

informação sintagmática faça parte da construção da grade, da mesma forma como faz parte da construção da árvore prosódica.

Não apresentaremos neste trabalho a representação para outros aspectos da fonologia, como a entoação e a fonologia segmental, e tampouco falaremos a respeito do componente semântico, já que esses aspectos da gramática não são relevantes para o restante do trabalho.

1.3.2 ...da música tonal

A "redução" é um conceito da análise musical que foi cunhado por Heinrich Schenker no início do século XX. A redução consiste na eliminação gradual de notas de uma superfície musical, até se chegar à sua estrutura subjacente mais básica, irreduzível. A "elaboração", ao contrário, consiste na modificação de uma superfície musical através da inserção de novas notas, mantendo-se a estrutura subjacente, o que mantém a identidade entre a superfície original e a superfície elaborada. Os processos de redução e elaboração são processos intuitivos perceptíveis na música que demonstram a interdependência entre eventos musicais não adjacentes e a natureza hierárquica da representação mental da música.

Embora a obra de Schenker consista na elaboração de uma análise musical que reduz peças musicais do idioma tonal até a sua estrutura mais profunda (que é a mesma para todas, denominada *Ursatz*), Schenker não formula regras reducionistas computáveis, mantendo a análise no plano puramente intuitivo. Lerdahl e Jackendoff (1983), baseando-se no trabalho de Schenker, propõem uma teoria formal que parte do pressuposto de que as notas musicais são ouvidas em termos estritamente hierárquicos. Para representar a hierarquia entre os eventos musicais, Lerdahl e Jackendoff lançam mão da mesma representação gráfica que a teoria gerativa da linguagem, que é o diagrama arbóreo, como o apresentado na Figura 5.

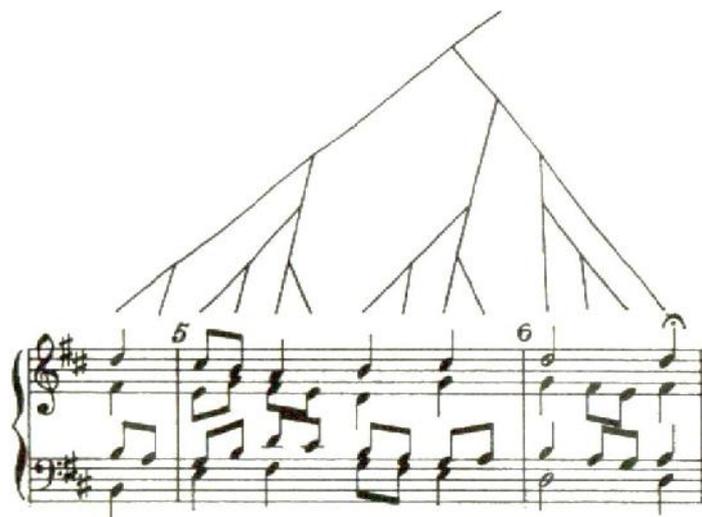


FIGURA 5: DIAGRAMA ARBÓREO PARA OS COMPASSOS 5 E 6 DE "O HAUPT" DE BACH.
 FONTE: LERDAHL; JACKENDOFF, 1983, p. 157.

A árvore da Figura 5, assim como a árvore linguística apresentada na Figura 1, possui diversos nós; no entanto, contrariamente à árvore linguística, esses nós não são rotulados. Da mesma forma como no diagrama linguístico, a união entre eventos na estrutura se dá de dois em dois, sendo que um dos dois deverá ser núcleo, o que, na música, não é representado através do rotulamento dos galhos. Aqui, a indicação de núcleo se faz geometricamente da seguinte forma: a nota cujo galho nasce de um galho maior é elaboração da nota representada pelo galho maior. Dessa forma, no exemplo (a) da Figura 6, há um par de irmãos cujo núcleo é representado pelo galho maior, e o outro galho é uma elaboração à direita do núcleo; no exemplo (b) da Figura 6, há uma nota núcleo – o galho maior – que tem uma elaboração à sua esquerda. Em (c), na Figura 6, temos uma ramificação maior: o maior galho (núcleo da árvore toda) tem uma elaboração à sua direita, e essa elaboração tem, por sua vez, uma elaboração à sua esquerda.



FIGURA 6: TRÊS EXEMPLOS DA MANEIRA COMO SE DÁ A RAMIFICAÇÃO DOS DIAGRAMAS ARBÓREOS DA TEORIA DE LERDAHL E JACKENDOFF (1983).

A informação codificada na estrutura arbórea apresentada na Figura 5 é de que os eventos musicais representados por galhos maiores são eventos mais proeminentes, estáveis e mais importantes estruturalmente, o que significa que, para se fazer uma redução, uma elaboração ou uma variação daquele trecho musical sem que se perca sua identidade, deve-se preferir manter antes esses eventos a outros, hierarquicamente mais baixos. Essas estruturas são designadas pelos autores de *redução do time-span* (doravante TSR, do inglês *time-span reduction*) e são construídas a partir de *regras de boa-formação*, que especificam as descrições estruturais possíveis, e *regras de preferência*, que designam as descrições estruturais, dentre as possíveis, que correspondem à maneira como um ouvinte experiente ouve uma dada peça musical.

Entretanto, a estrutura TSR necessita de outros dois tipos de estrutura como *input*, aos quais as regras de preferência se referirão, que são a *estrutura métrica* e a *estrutura de agrupamento*. A estrutura métrica expressa a intuição descrita na seção 1.1 de que os eventos de uma peça se relacionam a uma alternância regular entre eventos ritmicamente mais proeminentes e menos proeminentes em diversos níveis hierárquicos. A estrutura de agrupamento expressa a segmentação hierárquica da peça em grupos de eventos e grupos de grupos de eventos (ou seja, em sintagmas, para emprestar o termo da linguística) (LERDAHL; JACKENDOFF, 1983, p. 8). A redução do *time-span* expressa a informação das duas estruturas, simultaneamente. A Figura 7 mostra a estrutura métrica para o mesmo trecho presente na Figura 5. Como o trecho da Figura 5 é demasiadamente curto para apresentar uma estrutura de agrupamento complexa, a Figura 8 mostra a estrutura de agrupamento de um outro trecho musical.

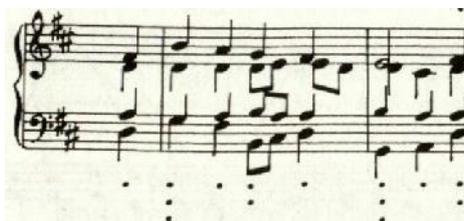


FIGURA 7: ESTRUTURA MÉTRICA PARA OS COMPASSOS 5 E 6 DE "O HAUPT" DE BACH.
 FONTE: LERDAHL; JACKENDOFF, 1983, p. 132.



FIGURA 8: ESTRUTURA DE AGRUPAMENTO PARA OS COMPASSOS INICIAIS DA SINFONIA Nº 40 DE MOZART. FONTE: LERDAHL; JACKENDOFF, 1983, p. 48

Para além das estruturas supracitadas (redução do *time-span*, estrutura métrica e estrutura de agrupamento), Lerdahl e Jackendoff postulam ainda uma quarta estrutura hierárquica, que codifica um quarto tipo de informação estrutural, que é a chamada *redução prolongacional* (doravante PR, do inglês *prolongational reduction*). A PR toma como *input* a TSR e atribui aos eventos musicais uma hierarquia que expressa movimentos harmônicos de tensão e relaxamento, continuidade e progressão. (LERDAHL; JACKENDOFF, 1983, p. 9). Embora a PR seja muito parecida com a TSR, as duas estruturas não são isomórficas e expressam informações diferentes.

Com essa visão geral da gramática musical conforme formulada por Lerdahl e Jackendoff (1983), vemos na Figura 9 a arquitetura geral (e simplificada) da gramática musical.¹¹

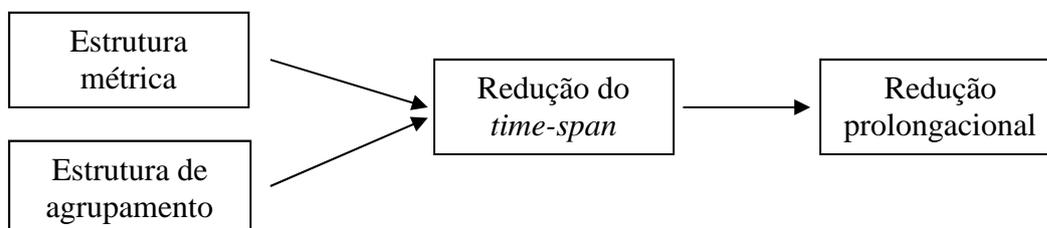


FIGURA 9: ARQUITETURA SIMPLIFICADA DA GRAMÁTICA MUSICAL. FONTE: KATZ; PESETSKY, 2011, p. 13.

1.4 Música, linguagem e sua interdisciplinaridade

Nos anos 80, o cientista Jerry Fodor recuperou uma noção a respeito da arquitetura funcional da mente humana, que é a noção de *modularidade da mente*, uma hipótese que postula que as faculdades humanas são associadas a estruturas

¹¹ Essa arquitetura simplificada apresenta apenas os *componentes estruturais* da teoria de Lerdahl e Jackendoff (1983). Na arquitetura original, os componentes interagem com os tipos de regras, que são, além das regras de boa-formação e das regras de preferência, também as regras transformacionais.

neuroológicas distintas e computacionalmente autônomas. Isso significa que cada faculdade humana está relacionada a um "órgão" cerebral específico para a aquisição e computação dessa faculdade particular. De acordo com essa hipótese, a linguagem estaria relacionada a um módulo cerebral específico à linguagem e a música estaria relacionada a um módulo cerebral específico à música. A hipótese alternativa é de que as faculdades humanas são todas provindas de um mesmo sistema cognitivo geral. Embora se classifique usualmente a hipótese da modularidade da mente como uma hipótese inatista e a hipótese alternativa como uma hipótese behaviorista, é possível uma combinação entre as duas hipóteses, em que se defende que existem propriedades inatas para o desenvolvimento de faculdades humanas (condizente com a hipótese da modularidade da mente), mas que essas propriedades são gerais, e não específicas (condizente com a hipótese alternativa).

Nas seções 1.2.1 e 1.2.2, foi levantada a questão a respeito do interesse das teorias musical e linguística pela busca de universais musicais e universais linguísticos. As implicações da descoberta de universais de uma atividade humana são grandes para o entendimento da mente humana, pois a universalidade de uma propriedade pode significar um viés cognitivo inato para o tratamento do fenômeno em questão; entretanto, não necessariamente são grandes para a discussão a respeito da modularidade da mente, que é uma questão atualmente em pauta no estudo das ciências cognitivas.

Encontrando-se universais de uma atividade humana qualquer, de que forma é possível saber se esses universais são próprios dessa atividade ou se pertencem a um sistema cognitivo geral, podendo, assim, ser encontrados na computação de outras atividades humanas, como a visão, a música e a locomoção? Uma maneira de responder a essa pergunta é comparando-se diferentes atividades humanas, a fim de observar se, por exemplo, uma única operação está disponível para a computação dessas diferentes atividades ou se é específica a apenas uma delas. Assim, o estudo sobre as relações entre música e linguagem, que tem ganhado foco no cenário acadêmico (sobretudo experimentalmente), pode trazer contribuições para a discussão a respeito da modularidade da mente. Nas seções 1.4.1, 1.4.2 e 1.4.3 abaixo, será apresentada, com base em Santana (2012), a descrição de alguns estudos recentes sobre a interface entre música e linguagem.

1.4.1 Aquisição

Dentre os tópicos de pesquisa que a teoria gerativa da linguagem busca explicar, está a maneira como o ser humano adquire a linguagem. As crianças aprendem toda a gramática de sua língua em poucos anos, apenas a partir da exposição ao *input*, sem serem explicitamente treinadas. Que mecanismos cerebrais estão em jogo nessa tarefa aparentemente tão difícil, porém bem sucedida por todas as crianças falantes de todas as línguas existentes, salvo casos patológicos? A mesma questão se coloca para a música: que mecanismos cerebrais estão em jogo quando uma criança adquire a gramática musical sem ser explicitamente treinada, de forma a ser capaz de apreciar música, salvo casos patológicos?

O estudo interdisciplinar da aquisição de música e linguagem pode ser útil na compreensão de questionamentos acerca da modularidade da mente. Qualquer adulto é capaz de diferenciar música de fala, seja por aspectos funcionais ou estruturais. Para um bebê, no entanto, música e linguagem talvez não sejam tão diferentes assim. Afinal, o bebê, ao ser exposto a dados do sistema musical e a dados do sistema linguístico, ainda não sabe quais são as intenções comunicativas desses dados e, portanto, não tem de saída o conhecimento das propriedades mais salientes que nos fazem distinguir música de linguagem. Dessa forma, é possível que música e linguagem compartilhem no cérebro os mesmos mecanismos de aprendizagem, já que são duas atividades que organizam elementos sonoros no tempo.

Um dos mecanismos de aprendizagem que pode ser coincidente entre música e linguagem no cérebro é aquele relativo à aprendizagem da percepção de categorias sonoras. Afinal, tanto música quanto linguagem são formadas de categorias de sons (notas musicais e fonemas, respectivamente), ou seja, de unidades discretas cujas propriedades só se fazem relevantemente perceptíveis quando determinantes para o pertencimento da unidade em questão a uma determinada categoria. Antes mesmo de serem suficientemente expostas a dados próprios de uma língua e a dados próprios de um idioma musical, as crianças já apresentam predisposições para a percepção categorial de *inputs* sonoros. Crianças ainda muito novas não atentam, por exemplo, para diferentes manifestações de um mesmo fonema /a/ e tratam essas diferentes manifestações como pertencentes a uma mesma categoria; já percebem, no entanto, a diferença entre /pa/ e /ba/ como mudança de categoria, embora sejam sons extremamente semelhantes, tendo como única diferença o fato de que, para a consoante

/p/ da primeira sílaba, as pregas vocais não vibram e, para a consoante /b/ da segunda sílaba, as pregas vocais vibram (EIMAS *et al.*, 1971¹², *apud* MCMULLEN; SAFFRAN, 2004). Além disso, crianças japonesas muito novas tratam os sons [r] e [l] como pertencentes a duas categorias distintas, apesar de a língua japonesa e os falantes adultos de japonês não distinguem um som do outro (MCMULLEN; SAFFRAN, 2004). Isso mostra a predisposição das crianças para lidar com categorias sonoras linguísticas. Da mesma forma, crianças já parecem preferir certos intervalos musicais (os consonantes) a outros (os dissonantes) (TRAINOR; HEINMILLER, 1999¹³, *apud* MCMULLEN; SAFFRAN, 2004) e já tratam uma mesma peça tocada em tonalidades diferentes como sendo a mesma, mas tratam uma mesma peça musical tocada em modos diferentes como peças diferentes¹⁴ (TRAINOR *et al.*, 2002¹⁵, *apud* MCMULLEN; SAFFRAN, 2004). Isso mostra, igualmente, uma predisposição das crianças para lidar com categorias sonoras musicais distintas, percebendo antes intervalos que a manifestação desses intervalos em notas distintas. Patel (2008) formula a hipótese de que o mecanismo envolvido na categorização de sons é um único mecanismo e, portanto, é compartilhado pela música e pela linguagem. O autor a chama de *hipótese do mecanismo compartilhado de aprendizagem de categorias sonoras* (SSCLMH: *shared sound category learning mechanism hypothesis*).

Um outro recorte possível no estudo de aquisição da linguagem e da música está nos processos prosódicos, ou seja, em padrões de ritmo, acento, contorno e entoação. De acordo com o estudo de Hirsch-Pasek *et al.* (1987¹⁶, *apud* MCMULLEN; SAFFRAN, 2004), crianças de 7 meses de idade tendem a dar atenção por mais tempo a um discurso em que as pausas coincidem com fronteiras de sintagmas, e escutam com

¹² EIMAS, P.; SIQUELAND, E.; JUSCZYK, P.; VIGORITO, J. (1971) Speech perception in infants. *Science*, 171, p. 303-306.

¹³ TRAINOR, L.; HEINMILLER, B. (1999) The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, 21, p. 77-88.

¹⁴ Músicas executadas em tonalidades diferentes são músicas que exibem exatamente o mesmo padrão intervalar, ou seja, a mesma relação entre as notas, embora se utilizem de notas diferentes. Por exemplo, se é tocada a nota dó e, em seguida, a nota mi, para mudar de tonalidade, deve-se utilizar duas notas cuja distância acústica seja a mesma (como fá e lá, por exemplo). Músicas executadas em modos diferentes são músicas que são construídas sobre escalas diferentes e, portanto, exibem relações diferentes entre as notas, mas mantendo-se os graus da escala. Por exemplo, se é tocada a primeira nota da escala e, em seguida, a terceira nota da escala, para mudar de modo deve-se tomar como base uma nova escala e, em seguida, repetir os graus (no caso, a primeira nota da escala, seguida da terceira nota da escala), independentemente da distância entre esses dois graus em cada uma das escalas.

¹⁵ TRAINOR, L.; WU, L.; TSANG, C. D.; PLANTINGA, J. (2002) Long-term memory for music in infancy. Paper presented at the International Conference on Infant Studies, Toronto.

¹⁶ HIRSH-PASEK, K.; KEMLER NELSON, D.; JUSCZYK, P.; CASSIDY, K. (1987) Clauses are perceptual units for young infants. *Cognition*, 26, p. 269-286.

atenção por menos tempo quando no discurso as pausas são inseridas no meio de uma unidade sintagmática. Adicionalmente, de acordo com o estudo de Jusczuk e Krumhansl (1993¹⁷, *apud* MCMULLEN; SAFFRAN, 2004), crianças dessa mesma idade tendem a ouvir com atenção por mais tempo uma peça musical em que as pausas coincidem com as fronteiras entre frases musicais, e escutam com atenção por menos tempo uma peça musical em que as pausas são inseridas no meio de uma frase. Essas fronteiras costumam, na música e na linguagem, ter as mesmas características prosódicas: um declínio de altura, um prolongamento da nota/vogal final e uma determinada estrutura hierárquica (sintática ou de redução do *time-span*), e as crianças parecem ser sensíveis a esses padrões. É possível, assim, que essas características em comum observadas na música e na linguagem compartilhem um mesmo mecanismo de aquisição no cérebro humano.

Uma terceira maneira de abordar a questão da aquisição da linguagem e da música é a partir da sintaxe. Sendo música e linguagem sistemas de natureza combinatória infinita, regidos por regras que nos permitem distinguir sequências bem formadas de sequências mal formadas, podemos afirmar que esses dois sistemas possuem uma sintaxe. Pesquisas mostram que crianças de apenas 6 meses de idade já têm percepção para padrões sintáticos, demonstrando estranhamento quando expostas a uma sequência sonora cujo padrão sintático foge das sequências a que foram expostas anteriormente. Assim, crianças que foram expostas a algumas sentenças curtas que exibem um padrão como AAB tendem a ouvir mais atentamente sentenças novas que fogem desse padrão (como ABA) do que sentenças novas que obedecem a esse padrão (MARCUS *et al.*, 1999¹⁸, *apud* MCMULLEN; SAFFRAN, 2004). Da parte musical, ainda não há um número suficiente de pesquisas sobre essa questão, mas os dados das pesquisas já existentes mostram que musicalmente as crianças também têm essa percepção, mas ela é adquirida um pouco mais tarde do que a percepção sintática linguística (para visão contrária, ver BRANDT; GEBRIAN; SLEVC, 2012). McMullen e Saffran (2004) apresentam duas hipóteses para explicar esse retardamento na aquisição da sintaxe musical em relação à sintaxe linguística: (i) as crianças são expostas a muito menos dados musicais do que a dados linguísticos; e (ii) os benefícios

¹⁷ JUSCZYK, P. W., & KRUMHANSL, C. L. (1993) Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, 19, p. 627–640.

¹⁸ MARCUS, G.; VIJAYAN, S.; BANDI RAO, S.; VISHTON, P. M. (1999) Rule learning by seven month-old infants. *Science*, 283, p. 77–80.

comunicativos se mostram maiores quando do conhecimento da linguagem do que quando do conhecimento da música.

Pode-se, ainda, pensar na maneira como crianças adquirem a significação e a memória para a música e para a linguagem, além do que foi visto nos parágrafos precedentes. É um campo de estudo que pode auxiliar bastante na solução de questões sobre a modularidade da mente. Ainda que a próxima seção também trate de um campo de estudos que igualmente auxilia na discussão sobre modularidade da mente (patologias relacionadas à língua e à música), a aquisição é provavelmente mais fácil de ser testada empiricamente, por lidar com algo que é comum a todos, e não com casos excepcionais.

1.4.2 Patologias relacionadas à música e à linguagem

Uma evidência para supor que tanto a música quanto a linguagem são módulos específicos e distintos no cérebro são as patologias a elas relacionadas. Afinal, existem patologias que afetam somente a linguagem, e nada mais, e patologias que afetam somente a música, e nada mais. As patologias que afetam o processamento ou o desempenho linguístico são denominadas *afasias* e as patologias que afetam o processamento ou o desempenho musical são denominadas *amusias*.

Os dois tipos de patologia manifestam-se de maneiras diversas, podendo atacar subáreas específicas da linguagem e da música (como perda da capacidade sintática, da capacidade de nomear objetos ou da capacidade de pronúncia, por um lado, e perda da capacidade de reconhecer alturas diferentes, da capacidade rítmica ou da capacidade de cantar, por outro). Ambas podem ser adquiridas ou congênitas e ambas subdividem-se em dano receptivo (incapacidade de interpretação) e dano clínico (incapacidade de desempenho).

Sacks (2007) relata o caso de uma paciente com amusia congênita para tons. Como não consegue discriminar tons, mesmo se ouvir diversas vezes uma mesma melodia, ela jamais será capaz de reconhecê-la quando a escutar novamente, porque, embora ela ouça, não é capaz de processar. Sendo tão surda para reconhecer tons, era de se esperar que ela também tivesse um *déficit* para processar a expressividade da fala, já que expressamos diferentes emoções discursivamente com mudança de contornos entoacionais, o que implica a mudança de altura da voz a partir da frequência

fundamental. Entretanto, o autor afirma não ter detectado nada de anormal nem em sua percepção e nem em sua fala; ao contrário, afirma que a paciente lhe recitou um poema com bastante expressividade.

Apesar dessa aparente independência entre a “surdez musical” e a percepção da entoação da fala, Patel *et al.* (2005), mostrando-se céticos quanto a isso, desenvolveram uma pesquisa com indivíduos surdos para tons para testar se música e linguagem, nesse sentido, de fato não se relacionam. Os estímulos apresentados aos participantes amúsicos foram (i) pares de sentenças que diferem entre si somente pela entoação (como, por exemplo, "Pedro corre." e "Pedro corre?" em português), (ii) pares sequências de tons análogas a cada uma das sentenças de cada um dos pares, imitando completamente a entoação da fala, inclusive reproduzindo os movimentos de deslize (portamento) da F0 e (iii) pares de sequências de tons análogas a cada uma das sentenças de cada um dos pares, em que cada sílaba foi substituída por uma frequência distinta, formando-se, assim, uma sequência de alturas discretas, sendo essa frequência a média entre a frequência máxima e a mínima presentes na sílaba falada em cada enunciado.

Esperava-se, se Sacks (2007) estivesse correto, que o resultado mostrasse nos amúsicos uma dificuldade perceptiva apenas sobre a sequência não-linguística de frequências distintas, mas não sobre as sentenças com entoações diferentes e, tampouco, sobre as sequências não-linguísticas que imitam fielmente a entoação das sentenças em questão. Afinal, a música é composta, sobretudo, por alturas distintas, e não contornos deslizantes, enquanto a entoação da fala é composta de contornos. Não é o que acontece, no entanto. Os amúsicos têm a mesma dificuldade de percepção nos dois tipos de sequências sonoras não-linguísticas, ainda que uma delas reproduza exatamente a entoação que, sem o esvaziamento semântico, há nas sentenças que lhes são compreensíveis em sua completude. Embora os autores não tenham formulado uma hipótese que explicasse esses resultados, os resultados possivelmente mostram que certos elementos da linguagem estão mais próximos da nossa capacidade musical do que se poderia pensar ao olhar somente para as manifestações superficiais de afasias e amusias.

Com relação às afasias, Jentschke *et al.* (2008) investigam a deficiência de processamento sintático musical em crianças que portam algum *distúrbio específico da linguagem*, tendo como grupo controle crianças com desenvolvimento linguístico típico

e adultos. O padrão revelou-se o mesmo em crianças que não portam deficiências linguísticas e em adultos. No entanto, em crianças com distúrbio específico da linguagem, o padrão encontrado no processamento da sintaxe musical foi diferente, refletindo sua dificuldade em processar regularidades sintático-musicais. Isso é compatível com a hipótese de que há uma justaposição cerebral de processamentos sintáticos na linguagem e na música.

Ainda que tenhamos apontado somente para os casos em que amusias revelam de alguma forma algum problema também relacionado à linguagem e para os casos em que afasias revelam de alguma forma algum problema também relacionado à música, há também muitos estudos de deficiências linguísticas e musicais que procuram, a partir da apresentação de casos de amusia sem afasia e de afasia sem amusia, provar o exato contrário, isto é, a independência desses *déficits* (cf. PERETZ; COLTHEART, 2003 e PERETZ, 2008).

1.4.3 Ritmo

Pike (1945) propõe uma classificação bipartida para as línguas do mundo em relação a seu tipo rítmico: existem as línguas do tipo acentual e as línguas do tipo silábico¹⁹. Quando primeiro foram descritas essas duas categorias, a propriedade a partir da qual elas supostamente se diferenciavam era o tipo de isocronia. As línguas do tipo acentual eram tidas como aquelas cujo elemento regularmente recorrente são as vogais tônicas e as línguas do tipo silábico eram tidas como aquelas cujo elemento regularmente recorrente são as sílabas. Hoje, no entanto, há fortes evidências para descartar a isocronia do ritmo prosódico, mas continua-se falando em uma dicotomia entre ritmo silábico e ritmo acentual no que diz respeito à estrutura silábica²⁰: em línguas do tipo silábico, a estrutura da sílaba é menos variada (quase não tem reduções silábicas) do que nas línguas do tipo acentual²¹.

¹⁹ Existem também as línguas moraicas (tendo como exemplo a língua japonesa), que muito se assemelham às línguas silábicas, com exceção do fato de que uma língua moraica pode ter sílabas que medem duas unidades rítmicas, ou seja, há sílabas que equivalem, ritmicamente, a duas sílabas. Esse tipo, no entanto, não entrará na presente discussão.

²⁰ Segundo estudos realizados por Grabe e Low (2002), não é bem uma *dicotomia* a diferença rítmica entre as línguas. Há uma gradação: certas línguas são, por exemplo, mais silábicas do que acentuais, mas não exclusivamente de um tipo ou do outro.

²¹ Essa reanálise foi proposta por Rebecca Dauer (1983, *apud* GRABE; LOW, 2002).

Havendo tal variedade rítmica entre as línguas, há que se perguntar se essa variedade se reflete em diferença rítmica na música das culturas que falam esses diferentes tipos de línguas. Patel e Daniele (2003) conduziram um estudo sobre as possíveis relações entre o ritmo da música de uma cultura e o ritmo da língua falada por essa mesma cultura. Para poder determinar a interferência ou não do ritmo da fala na música, eles compararam uma língua de tipo acentual (a língua inglesa) com uma língua de tipo silábico (a língua francesa) e, em seguida, compararam a música da cultura que fala a língua acentual (a música inglesa) com a música da cultura que fala a língua silábica (a música francesa), para por fim comparar os resultados.

A maneira como foram medidos os padrões rítmicos tanto na música quanto na língua foi através de um algoritmo denominado nPVI (*normalized pairwise variability index*) que mede a duração de unidades adjacentes dentro de uma cadeia sonora – e, em seguida, o padrão de variabilidade entre as durações dentro de uma mesma cadeia – sendo essas unidades adjacentes notas musicais e unidades inter-silábicas na música e na língua, respectivamente. O índice de variabilidade (nPVI) se mostra maior em línguas do tipo acentual, já que essas línguas são aquelas cuja estrutura silábica é mais variada. Sendo assim, a ideia inicial da pesquisa era encontrar, na música das culturas cuja língua é do tipo acentual, um nPVI mais alto do que o nPVI da música das culturas cuja língua é do tipo silábico. Com sucesso, Patel e Daniele conseguiram mostrar empiricamente que, paralelamente à língua, o nPVI da música inglesa é mais alto do que o da música francesa, como se pode ver na Figura 10.

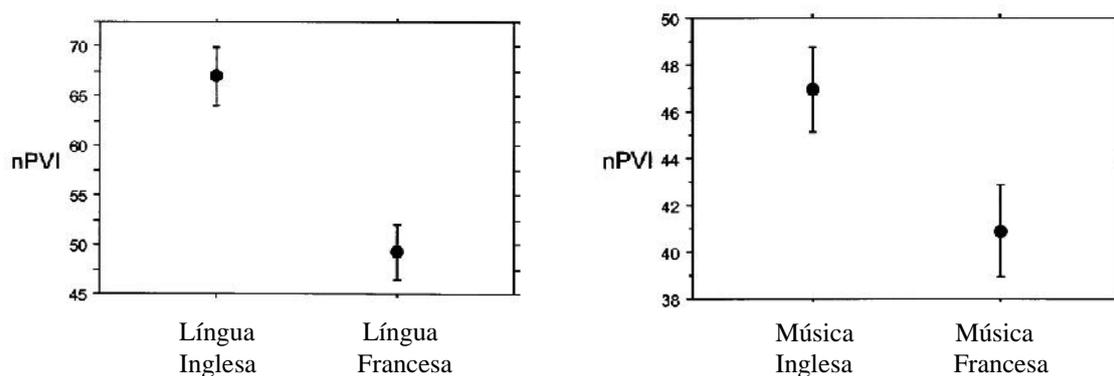


FIGURA 10: GRÁFICOS COM O ÍNDICE DE VARIABILIDADE RÍTMICA DAS LÍNGUAS INGLESA E FRANCESA POR UM LADO, E DAS MÚSICAS INGLESA E FRANCESA POR OUTRO. FONTE: PATEL; DANIELE, 2003, p. B38 e B41.

Quanto às melodias escolhidas para fazer a medição do índice de variabilidade, Patel e Daniele tomaram o cuidado de colher exemplos somente de música instrumental, porque se a música cantada se baseia em palavras, as canções obviamente carregariam as características prosódicas da língua, devido antes à necessidade de ajuste da música ao texto do que por uma influência dos padrões rítmicos da língua nativa do compositor. Além disso, procuraram também deixar de fora peças musicais com influência explícita da música folclórica – por ela ser cantada – já que uma música instrumental sob sua influência supostamente tomaria também seus padrões rítmicos.

A verificação do alinhamento do ritmo da música de uma cultura com o ritmo de sua respectiva língua pode ser explicada pelo fato de a internalização dos padrões rítmicos da fala se dar muito cedo, algo que é verificado em estudos sobre aquisição linguística (RAMUS, 2002a²², *apud* PATEL; DANIELE, 2003a). Tendo internalizado esses padrões na primeira infância, o compositor inconscientemente incorpora à sua música esses mesmos padrões aos quais já está habituado. Embora esse estudo não necessariamente elucide questões sobre a justaposição cerebral de propriedades rítmicas linguísticas e musicais, é um dos muitos exemplos de estudos *empíricos* sobre as relações entre música e linguagem que poderiam se apoiar em estudos *teóricos* sobre as relações entre música e linguagem, conforme propomos no presente trabalho.

1.5 Considerações finais

Neste capítulo, apresentamos em linhas gerais os pressupostos teóricos e o formalismo de uma teoria sobre linguagem e de uma teoria sobre música. Conforme se pôde notar, as duas teorias compartilham os mesmos pressupostos teóricos e, além disso, alguns mecanismos formais são muito semelhantes. Assim, não só as duas teorias colocam a seus respectivos objetos perguntas como "o que um falante/ouvinte *sabe* quando sabe um idioma linguístico/musical?", "de que maneira um falante/ouvinte adquire esse conhecimento?", "que parte desse conhecimento é inata e que parte é adquirida?" e "quão geral ou quão específico é o conhecimento inato?", como buscam responder a essas perguntas através de formalismos semelhantes. A estrutura métrica da música (Figura 7), por exemplo, assemelha-se à grade métrica da linguagem (Figura 4); os diagramas arbóreos da linguagem (Figuras 1 e 3) assemelham-se ao diagrama

²² RAMUS, F. (2002a) Language discrimination by newborns: teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, 2, p. 85–115.

arbóreo da música (Figura 5); a estrutura de agrupamento da música (Figura 8) é uma estrutura que divide a superfície musical em "sintagmas", da mesma forma como faz os rótulos da árvore linguística (Figura 1); ademais, a informação codificada por algumas dessas estruturas é bastante paralela, conforme notam Lerdahl e Jackendoff (1983, cap. 12) e conforme será brevemente discutido na seção 2.3.3.

Neste capítulo, apresentamos também algumas das diversas pesquisas experimentais sobre possíveis relações entre música e linguagem na mente humana, que é uma interdisciplinaridade que vem ganhando cada vez mais espaço (cf. PATEL, 2008). Neste trabalho, buscaremos recriar uma pesquisa empírica existente sobre música e linguagem, tomando-se como base para a construção dos experimentos que serão relatados as teorias descritas nas seções anteriores. O estudo que está sendo tomado como base é de Nevins e Endress (2007) e o fenômeno em observação é verificado apenas na linguagem, embora os autores o busquem experimentalmente também na música, com a finalidade de observar se o fenômeno em questão é restrito à linguagem ou se é compartilhado com a música. O próximo capítulo (capítulo 2) apresenta o trabalho de Nevins e Endress e discute a maneira como os autores conduzem seu estudo, sugerindo novas formas de pesquisar o mesmo tema. O capítulo 3 descreve os experimentos que nós conduzimos com base na discussão presente no capítulo 2 e, por fim, o capítulo 4 discute os resultados a que chegamos com nossos experimentos.

2. DESCRIÇÃO E PROBLEMATIZAÇÃO DE NEVINS E ENDRESS (2007)

2.1 Introdução

Neste capítulo, será primeiramente apresentada uma descrição do trabalho original de Nevins e Endress (2007) que estamos questionando, a fim de fornecer o material necessário para o leitor poder julgar a pertinência ou não de nossa proposta de continuação dos experimentos por eles conduzidos. Em seguida, traremos uma extensa discussão sobre detalhes dos experimentos dos autores, que justificam o fato de este trabalho reformular esses experimentos e trazer à discussão esses novos resultados, assunto dos próximos capítulos (capítulos 3 e 4).

2.2 Descrição dos experimentos de Nevins e Endress (2007)

O objetivo principal de Nevins e Endress (2007) é testar a universalidade de uma hipótese que parte da observação e análise de um tipo particular das chamadas *línguas secretas*. Por este motivo, a seção abaixo (2.2.1) contempla a caracterização desse fenômeno linguístico, em especial do subconjunto de fatos investigado pelos autores. A seção seguinte (2.2.2) ilustra o experimento linguístico conduzido pelos autores para testar sua hipótese a respeito das restrições sobre as possibilidades de manifestações desse fenômeno. Por fim, a seção 2.2.3 ilustra o experimento musical conduzido pelos autores para verificar a (não)especificidade linguística das restrições cognitivas impostas sobre a manipulação feita pelo subconjunto de línguas secretas investigado.

2.2.1 Línguas secretas

Praticamente todas as comunidades do mundo apresentam pelo menos uma manifestação do sistema de manipulação da fala denominado *língua secreta* (que também pode ser denominado *jogo linguístico* ou ainda *língua lúdica*)²³.

²³ No inglês, chama-se a este fenômeno linguístico de *ludling*, termo cunhado por Laycock (1972). *Ludling* vem do latim *ludus* ('jogo') + *lingua* (língua, linguagem, idioma) – significando literalmente 'jogo linguístico', portanto. O motivo pelo qual se utiliza esse termo em inglês, e não o esperado *language games*, é devido ao fato de essa última expressão ter um significado mais abrangente e, por isso, ser associada a códigos linguísticos que não são de interesse para o estudo da fonologia ou mesmo da estrutura linguística (como o código Morse), por ter uma definição antes sociolinguística do que formal.

As línguas secretas são sistemas de codificação da linguagem, com a função social de ocultar de ouvintes não treinados o conteúdo de conversas entre pessoas treinadas na instância particular da língua secreta. As línguas secretas são mais comumente utilizadas por crianças e adolescentes para encobrir o conteúdo de suas conversas de, por exemplo, seus pais ou crianças e adolescentes pertencentes a outros grupos sociais. Como exemplo de língua secreta do português brasileiro, temos a conhecida *língua do pê*²⁴. Assim como a *língua do pê*, todo jogo linguístico parte de uma língua natural (compartilhada por todos os "falantes" do jogo) e, a partir de um sistema morfológico pobre e semanticamente vazio, transforma a estrutura fonológica das palavras "originais", criando, assim, um código, cujo conteúdo não se diferencia do conteúdo da superfície da língua base sem as transformações morfológicas da língua secreta.

Com o advento da fonologia autosssegmental, as línguas secretas, antes deixadas de lado pelo estudo da fonologia, começaram a mostrar seu caráter sistemático, baseado em princípios e formalmente relacionado a fenômenos já conhecidos da linguagem, ainda que, superficialmente, a linguagem padrão e as línguas secretas sejam muito distintas. Além disso, embora antes da fonologia autosssegmental as línguas secretas servissem ao estudo da fonologia apenas como uma evidência a mais para operações fonológicas das línguas das quais as línguas secretas provêm, estudos recentes sobre as línguas secretas revelam mais sobre a capacidade linguística geral do ser humano do que sobre sua língua de origem. Por este motivo, acredita-se que o estudo das línguas secretas pode revelar muito sobre os princípios que subjazem a nossa capacidade linguística (BAGEMIHLE, 1995).

Segundo Bagemihl (1995), as línguas secretas enquadram-se em certas categorias tipológicas, sendo que as quatro grandes categorias são (i) infixação, (ii) inversão, (iii) substituição e (iv) *templating*. As línguas secretas do tipo (i) envolvem a concatenação de um afixo em uma palavra. Esse afixo é geralmente um *infixo*, que se caracteriza por ser um morfema que se localiza dentro da raiz da palavra, dividindo-a em duas partes descontínuas. Como exemplo de língua secreta por infixação temos a *língua do pê*, mencionada no primeiro parágrafo desta seção e explicada na nota 24. As línguas secretas do tipo (ii) modificam a ordem das sílabas das palavras. Como exemplo de

²⁴ A *língua do pê* apresenta algumas variações em sua manifestação, mas em uma de suas variações (possivelmente a mais conhecida), essa língua insere, entre cada sílaba, uma sílaba nova cujo *onset* é o som [p] e cuja rima é copiada da sílaba anterior. Por exemplo, para a sentença "A menina é bonita", teríamos como resultado "Apa mepenipinapa épé boptonipitapa".

língua secreta por inversão, temos a língua secreta do francês *Verlan*, cujo nome se inspira na inversão da expressão francesa *à l'envers*, que significa "ao contrário". Nas línguas secretas do tipo (iii), todas ou quase todas as vogais são substituídas por um ou dois fonemas específicos. Na língua secreta da língua *cuna*, por exemplo, todas as vogais são substituídas pela vogal [i], transformando, assim, uma palavra como "nuka" (que significa "nome") em "niki". No português, há uma instância desse tipo de língua secreta na canção infantil "O sapo não lava o pé", em que há cinco repetições, cada uma delas substituindo-se todas as sílabas da letra original por uma das vogais da língua, exceto as vogais médias abertas (a vogal da palavra "pé" e a vogal da palavra "pó"). Por fim, as línguas secretas do tipo (iv) apresentam um modelo fixo de consoantes e vogais, ao qual o material fonológico será acrescentado. Por exemplo, é possível haver um modelo do tipo *CayCC*, que quer dizer: produza a primeira consoante da palavra, substitua a vogal da palavra pelo ditongo *ay* e, em seguida, reduplique a consoante seguinte da palavra. Para a palavra "bet", por exemplo, teríamos "bayt t"²⁵.

Dentre essas quatro grandes classes, o tipo mais comum é a inversão, também referido como *alterador de precedência* (do inglês *precedence-modifying*). Essa é a classe de línguas secretas que será investigada por Nevins e Endress. Por este motivo, de agora em diante nós nos referiremos unicamente a esse tipo de jogo linguístico.

A inversão silábica em palavras com apenas duas sílabas ($1\ 2\ 2\ 1$ que, aplicada à palavra [pra.to], por exemplo, resulta em [to.pra]) é extremamente comum em todas as manifestações de línguas secretas do tipo alterador de precedência. Entretanto, em palavras com mais do que duas sílabas, encontra-se muita variação nos dados da fala. Por exemplo, para uma palavra com quatro sílabas, diferentes línguas secretas podem revelar as ordens apresentadas em (3a-e):

(3)

- a. $1\ 2\ 3\ 4\ 2\ 3\ 4\ 1$ (Língua Fula)
- b. $1\ 2\ 3\ 4\ 4\ 1\ 2\ 3$ (Tagalog)
- c. $1\ 2\ 3\ 4\ 2\ 1\ 3\ 4$ (Marquesano)
- d. $1\ 2\ 3\ 4\ 1\ 2\ 4\ 3$ (Lucazi)

²⁵ O símbolo \emptyset representa uma vogal neutra, que deve ser inserida foneticamente naquela posição devido a uma incapacidade motora de produzir o som [t] duas vezes, sem que uma vogal seja interposta entre as duas produções de [t].

e. 1 2 3 4 4 3 2 1 (Saramacano)

Para os autores, a variação acima não é caótica, e sim devida a diferentes generalizações que os falantes dão a um *input* limitado. A inversão dissilábica é compatível com diversos comandos logicamente possíveis que se manifestam de forma diferente quando há o aumento do número de sílabas das palavras. Ou seja, ao se depararem com diversos dados com apenas duas sílabas da língua secreta que estão aprendendo, é possível que os falantes imaginem diversas regras capazes de gerar esses dados com duas sílabas invertidas e que funcionariam diferentemente em palavras com mais do que duas sílabas. Por exemplo, se, ao ouvir um dado 1 2 2 1, ou seja, um dado como [pra.to] [to.pra], um falante *x* infere que a regra por trás desse dado é “mova 1 (primeira) para o final”, um falante *y* infere a regra “mova F (final) para o início”, um falante *w* infere a regra “transponha (1, 2)”, um falante *z* infere a regra “transponha (F, F-1)” e um falante *k* infere a regra “inverta a ordem de todas as”, todos eles obterão corretamente a ordem 2 1 para palavras dissilábicas. Entretanto, para palavras com, por exemplo, quatro sílabas, a gramática de cada um desses falantes gerará uma ordem diferente, cada uma delas compatível com as ordens apresentadas em (3a-e). Veja (4a-e) abaixo:

(4)

a. falante *x*: mova 1 (primeira) para o final:

1 2 3 4 2 3 4 1 (ex. [si.na.li.zar] [na.li.zar.si])

b. falante *y*: mova F (final) para o início:

1 2 3 4 4 1 2 3 (ex. [si.na.li.zar] [zar.si.na.li])

c. falante *w*: transponha(1, 2):

1 2 3 4 2 1 3 4 (ex. [si.na.li.zar] [na.si.li.zar])

d. falante *z*: transponha(F, F-1):

1 2 3 4 1 2 4 3 (ex. [si.na.li.zar] [si.na.zar.li])

e. falante *k*: inverta a ordem de todas as :

1 2 3 4 4 3 2 1 (ex. [si.na.li.zar] [zar.li.na.si])

Ainda assim, os exemplos acima mencionados parecem um pouco caóticos, dado que aparentemente é possível existirem milhares de generalizações compatíveis com esses dados, sendo o *input* tão limitado. Entretanto, Bagemihl (1989) relata algumas restrições sobre as possibilidades de inversões silábicas dentro do observado nas diversas línguas secretas que se enquadram na classe alteradora de precedência:

(5)

- a. Nenhuma língua secreta inverte as duas sílabas do meio;
- b. Nenhuma língua secreta move a sílaba final para o meio;
- c. Nenhuma língua secreta permuta cada duas sílabas em uma palavra;
- d. Nenhuma língua secreta permuta pés;²⁶
- e. Nenhuma língua secreta permuta traços suprasegmentais;²⁷
- f. Nenhuma língua secreta cria palíndromos.²⁸

Com essa constatação, surge a seguinte questão: as restrições acima listadas são decorrentes do fato de as línguas secretas terem um número muito pequeno de falantes, o que justifica a ausência desses padrões nos dados ou seriam essas restrições decorrentes de algum princípio cognitivo que leva os falantes a preferirem certos tipos de ordenamentos a outros? Para Nevins e Endress, a segunda opção é a verdadeira. Segundo os autores, existe um viés cognitivo que restringe essas possibilidades. Por isso, as perguntas que os autores se colocam e que serão investigadas por eles são: dentre todos os comandos logicamente possíveis desse tipo de transformação do componente fonológico, quais são utilizados pelos falantes e que fazem parte de um viés cognitivo? Esse viés é linguístico (pertencente à gramática universal (GU)) ou geral (pertencente ao domínio cognitivo geral e, portanto, encontrado também em outras

²⁶ "Pé" é um termo utilizado por alguns fonólogos para designar uma unidade rítmica nas línguas naturais que exibe isocronia, ou seja, em que as sílabas tônicas se apresentam em intervalos regulares dentro de um enunciado. Esse termo provém dos estudos tradicionais da estrutura métrica do verso (da poesia), em que se classificavam muitos padrões regulares de sílabas acentuadas e não-acentuadas. Por exemplo, chama-se "jâmbico" o padrão não-acentuado+acentuado e trocaico o padrão acentuado+não-acentuado. Na hierarquia prosódica, o "pé" é o constituinte prosódico logo acima da sílaba.

²⁷ "Traços suprasegmentais" são elementos fonológicos que se estendem ao longo de uma sequência de segmentos (fonemas) em um enunciado, como a entoação e o ritmo. Os traços suprasegmentais estão em contraste com os traços segmentais, que se referem a "segmentos" da fala, como os fonemas.

²⁸ Um palíndromo é uma frase, palavra, número ou qualquer outra sequência de símbolos cuja ordem invertida é igual à ordem natural. Um exemplo muito popular de palíndromo do português é a frase "Socorram-me, subi no ônibus em Marrocos", cuja sequência de letras (ignorando outros sinais gráficos, como pontuação e acentuação) é igual nas duas direções.

faculdades humanas)? Para responder à primeira questão, os autores conduziram o experimento linguístico que será descrito na próxima seção (2.2.2). Para responder à segunda, os autores conduziram o experimento musical que será descrito na seção 2.2.3.

2.2.2 *Experimento linguístico*

Para o controle dos dados de seu experimento linguístico, Nevins e Endress utilizaram a metodologia de gramáticas artificiais, que consiste em ensinar aos participantes algumas estruturas linguísticas que não estão presentes em sua língua, a fim de verificar quais dessas estruturas eles preferem e quais delas eles rejeitam. Essa metodologia permite constatar o que é possível e o que é impossível de ser aprendido pelos participantes sem treinamento explícito, uma vez que as estruturas preferidas por eles poderiam ser geradas por regras disponíveis à cognição para aprendizagem, enquanto as estruturas rejeitadas poderiam ser geradas por regras cognitivamente impossíveis – e, portanto, interpretadas pelos participantes como estruturas aleatórias, que não seguem nenhum padrão.

O experimento foi executado da seguinte forma: primeiramente, informou-se aos participantes que eles iriam testemunhar um ritual marciano. Neste ritual, um marciano chefe pronuncia uma sentença, à qual um marciano subordinado deve responder apropriadamente. Os participantes também foram informados de que esses dois marcianos dominam o ritual perfeitamente e, em seguida, foi pedido para que os participantes tentassem descobrir como o ritual funciona. Então, foram apresentados 25 estímulos aos participantes, em que uma voz sintetizada (o marciano chefe) pronunciava uma sequência de três sílabas e uma outra voz sintetizada (o marciano subordinado) respondia usando as mesmas sílabas, mas na ordem inversa. Ou seja, em todas as 25 vezes, o marciano chefe dizia uma sentença na forma $s_1 s_2 s_3$ e, em seguida, o marciano subordinado respondia $s_3 s_2 s_1$. Esses estímulos são compatíveis com, pelo menos, quatro generalizações diferentes, apresentadas em (6) abaixo:

(6)

- **Comando 1:** Inverta a ordem de todas as sílabas;
- **Comando 2:** Permute a primeira sílaba com a última sílaba;
- **Comando 3:** Permute a última sílaba com a antepenúltima sílaba;

- **Comando 4:** Permute a cada duas sílabas (*i.e.* j com $j+2$).

Embora o resultado desses quatro comandos seja sempre o mesmo para palavras trissilábicas, para palavras com quatro sílabas, ao contrário, cada um deles gera um ordenamento diferente, como vemos em (7):

(7)

- **Comando 1:** Inverta a ordem de todas as sílabas:

1 2 3 4 4 3 2 1

- **Comando 2:** Permute a primeira sílaba com a última sílaba:

1 2 3 4 4 2 3 1

- **Comando 3:** Permute a última sílaba com a antepenúltima sílaba:

1 2 3 4 1 4 3 2

- **Comando 4:** Permute a cada duas sílabas (*i.e.* j com $j+2$):

1 2 3 4 3 4 1 2

Por isso, após fazer os sujeitos se familiarizarem com os estímulos de inversão trissilábica, os participantes foram informados de que eles iriam testemunhar o ritual com o marciano chefe e com um outro marciano subordinado que, agora, não domina o jogo tão bem. Os informantes deveriam julgar em uma escala de 1 a 9 se a resposta do novo marciano se conformava às regras do ritual. Eles deveriam dar a nota 1 se eles tivessem certeza que a resposta do marciano estava errada, 9 se eles tivessem certeza de que estava correta e 5 se eles não tivessem certeza – podendo o participante escolher valores intermediários, conforme o grau de certeza de seu julgamento. Foi apresentado um total de 20 estímulos em que o marciano chefe pronunciava uma sequência de quatro sílabas; nesses 20 estímulos, a resposta do marciano subordinado apresentava as mesmas sílabas da sequência do marciano chefe, mas distribuía essas respostas em igual proporção pelas quatro diferentes ordens resultantes dos quatro diferentes comandos apresentados em (7) – ou seja, havia cinco dados para cada uma das inversões. Os autores referem-se às inversões resultantes dos comandos 1 e 2 como "naturais", por serem ordenamentos atestados em línguas secretas do tipo alterador de precedência

existentes. As outras duas inversões, resultantes dos comandos 3 e 4, são chamadas de "não naturais" pelos autores, por não serem atestadas na realidade da fala.

O objetivo do experimento era verificar se os participantes optam ou não pelos comandos 3 e 4, dado que são regras não atestadas na tipologia existente sobre línguas secretas. Sua inexistência pode ser (i) devida ao fato de simplesmente não ter sido encontrada nenhuma língua secreta que faça uso dessas regras, sendo, portanto, uma lacuna decorrente meramente de fatores históricos ou (ii) devida a um viés cognitivo que elege 1 e 2 como regras possíveis e 3 e 4 como impossíveis.

Segundo os autores, diversas pesquisas já demonstraram que, em sequências de objetos linguísticos, não só os objetos em si são importantes, mas também sua posição na sequência. Tais pesquisas demonstram que as posições de extremidade (final e inicial) das sequências são posições privilegiadas. Sob a hipótese de que as posições privilegiadas em oposição às não-privilegiadas restringem as possibilidades de operações sobre objetos linguísticos, Nevins e Endress sugerem que as operações de modificação do ordenamento silábico das línguas secretas só podem fazer referência às sílabas de extremidade, e jamais a posições não extremas.²⁹ Isso explicaria, segundo os autores, por que as transformações geradas pelos comandos 1 e 2 são encontradas em línguas secretas existentes e por que as transformações geradas pelos comandos 3 e 4 não são. E mais ainda: prevê que as transformações resultantes dos comandos 3 e 4 *jamais* poderão ser encontradas em manifestações naturais de línguas secretas, por serem impossíveis de serem processadas cognitivamente.

O resultado desse experimento corroborou a hipótese dos autores. Os dados resultantes dos comandos 1 e 2, por um lado, foram significativamente mais bem aceitos do que os dados resultantes dos comandos 3 e 4 ($F(1,11) = 20.43$; $p = 0.0009$) e, ademais, os dados resultantes dos comandos 1 e 2 tiveram uma média significativamente acima de 5, que é o "ponto neutro" ($p = 0.0005$) e os dados resultantes dos comandos 3 e 4 tiveram uma média significativamente abaixo do "ponto neutro" ($p = 0.037$). A média do comando 1, embora maior do que a média do comando 2, não foi *significativamente* maior do que a média do comando 2, o que permite

²⁹ Nevins (2010) estende essas possibilidades ao dizer que, além de fazer referência às sílabas de extremidade, as operações de transposição de sílabas também podem fazer referência a sílabas definidas por uma função relativizada de alguma das sílabas de extremidade; crucialmente, contudo, não a *qualquer* sílaba definida nesses termos, mas somente às sílabas imediatamente precedentes ou subsequentes – caso contrário, o comando 3, que diz "permuta a última sílaba com a antepenúltima sílaba", ou "permuta F com $F-2$ ", não poderia ser considerado "não natural" pelos autores.

enquadrar essas transformações naturais em uma mesma "categoria de aceitação", mas a média entre o comando 3 e o comando 4 foram significativamente diferentes entre si ($p = 0.017$), o que não era previsto pela hipótese inicial do estudo. Os autores atribuíram esta divergência à possível existência de uma estratégia de curto-circuito que dê conta de explicar essa diferença numérica. Esses resultados podem ser visualizados na Figura 11.

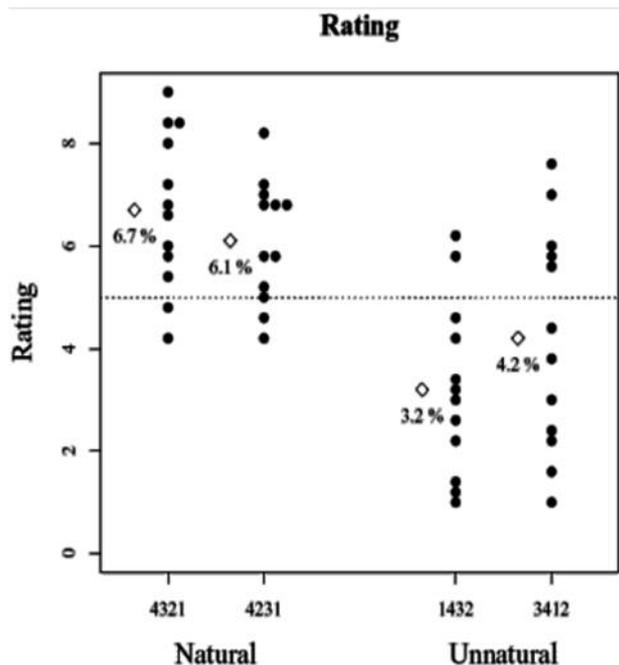


FIGURA 11: JULGAMENTO DOS PARTICIPANTES PARA AS DUAS INVERSÕES NATURAIS (À ESQUERDA) E AS DUAS INVERSÕES NÃO NATURAIS (À DIREITA). FONTE: NEVINS; ENDRESS, 2007, p. 6.

Dada a extensa discussão na literatura linguística sobre a existência de um módulo cerebral puramente linguístico (HAUSER; CHOMSKY; FITCH, 2002; PINKER; JACKENDOFF, 2005), os autores, na apresentação dos resultados desse experimento, discutem se essas transformações são operações gerais de aprendizagem de sequências ou se são específicas a operações de unidades linguísticas. A afirmação de que propriedades linguísticas são *puramente* linguísticas só é possível após a verificação de que elas não estão disponíveis a outras atividades humanas. Assim, no intuito de verificar a hipótese de que a restrição verificada sobre as possibilidades de manipulação do ordenamento de unidades linguísticas é uma restrição puramente

linguística, os autores replicaram o experimento com estímulos musicais, o que será descrito a seguir.

2.2.3 *Experimento musical*

O procedimento do experimento musical conduzido por Nevins e Endress (2007) foi semelhante ao procedimento do experimento linguístico: primeiramente, informou-se aos participantes que eles iriam testemunhar um ritual marciano. Neste ritual, um marciano chefe executa uma melodia, à qual um marciano subordinado deve responder apropriadamente. Os participantes também foram informados de que esses dois marcianos dominam o ritual perfeitamente e, em seguida, foi pedido para que os participantes tentassem descobrir como o ritual funciona. Então, foram apresentados 30 estímulos musicais aos participantes, em que o marciano chefe executava uma melodia de quatro notas musicais e o marciano subordinado respondia a essa melodia em um outro instrumento através da inversão da melodia do índio chefe. O motivo pelo qual se utilizaram melodias com quatro notas musicais, segundo os autores, é que as pessoas não processam os valores absolutos das notas musicais, e sim as relações entre elas. Assim, a unidade que está sendo manipulada no experimento é o intervalo musical formado entre duas notas adjacentes da melodia, e não as notas, formando-se, assim, uma sequência de três intervalos musicais com uma melodia de quatro tons. Ademais, a direção dos intervalos foi também invertida, pois, quando se toca um intervalo de trás para frente, sua direção muda. Ou seja, um intervalo ascendente torna-se descendente e vice-versa. Dessa forma, em todos os 30 estímulos apresentados, o marciano chefe executava uma melodia com a forma $1\ 2\ 3$ e, em seguida, o marciano subordinado respondia $3\ 2\ 1$. Assim como no caso dos estímulos linguísticos, esses estímulos são matematicamente compatíveis com, pelo menos, quatro generalizações diferentes, apresentadas em (8).

(8)

- **Comando 1:** Inverta a ordem de todos os intervalos;
- **Comando 2:** Permute o primeiro intervalo com o último intervalo;
- **Comando 3:** Permute o último intervalo com o antepenúltimo intervalo;
- **Comando 4:** Permute a cada dois intervalos (*i.e.* j com $j+2$).

Exatamente como no caso do experimento linguístico, embora o resultado dessas quatro regras seja sempre o mesmo para melodias com três intervalos, para melodias com quatro intervalos cada um desses comandos gera um ordenamento diferente, como vemos em (9).

(9)

- **Comando 1:** Inverta a ordem de todos os intervalos:

1 2 3 4 4 3 2 1

- **Comando 2:** Permute o primeiro intervalo com o último intervalo:

1 2 3 4 4 2 3 1

- **Comando 3:** Permute o último intervalo com o antepenúltimo intervalo:

1 2 3 4 1 4 3 2

- **Comando 4:** Permute a cada dois intervalos (*i.e.* j com $j+2$):

1 2 3 4 3 4 1 2

Por isso, após os participantes se familiarizarem com os dados de inversão com três intervalos, eles foram informados de que iriam testemunhar o ritual com o marciano chefe e com um outro índio subordinado que, agora, não domina o jogo tão bem. A tarefa era julgar em uma escala de 1 a 9 se a resposta do novo marciano se conformava às regras do ritual. Eles deveriam dar a nota 1 se eles tivessem certeza que a resposta do marciano estava errada, 9 se eles tivessem certeza de que estava correta e 5 se eles não tivessem certeza – podendo o participante escolher valores intermediários, conforme o grau de certeza de seu julgamento. Foram apresentados estímulos em que o marciano chefe pronunciava uma sequência de quatro intervalos (e, conseqüentemente, cinco notas); nesses dados, a resposta do marciano subordinado apresentava os mesmos intervalos da sequência do índio chefe, mas as respostas foram distribuídas em igual proporção pelas quatro diferentes ordens resultantes dos quatro diferentes comandos apresentados em (9). Os autores não divulgaram a quantidade de dados de quatro unidades apresentadas aos informantes no experimento musical.

O resultado do experimento musical também mostrou que os informantes julgaram melhor as transformações ditas "naturais" do que as transformações não

naturais ($F(1,12) = 11.96$; $p = 0.006$). Entretanto, diferenças significativas foram encontradas entre o comando 1 e os comandos 2, 3 e 4 ($p = 0.034$, $p = 0.0343$ e $p = 0.038$, respectivamente), enquanto não houve nenhuma outra diferença significativa entre quaisquer dois outros comandos. Assim, o primeiro grupo foi avaliado com os *scores* mais altos por parte dos participantes. A Figura 12 abaixo mostra os resultados:

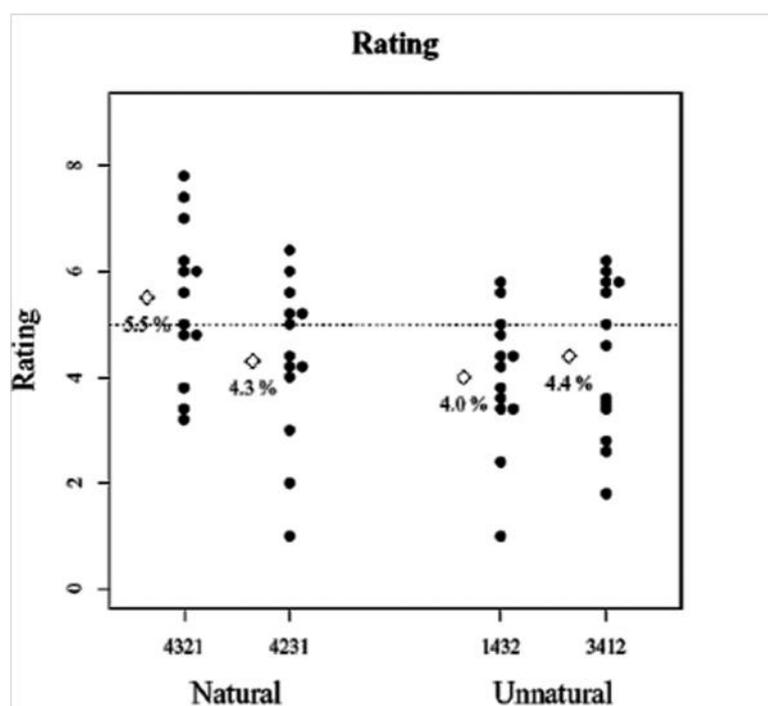


FIGURA 12: JULGAMENTO DOS PARTICIPANTES PARA AS DUAS INVERSÕES NATURAIS (À ESQUERDA) E AS DUAS INVERSÕES NÃO NATURAIS (À DIREITA). FONTE: NEVINS; ENDRESS, 2007, p. 8.

O resultado deste experimento não coincidiu com o resultado do experimento linguístico, o que sugere, para os autores, que a maneira como transformações de sequências musicais são aprendidas é diferente da maneira como as transformações de sequências linguísticas são aprendidas. Isso é um forte indício de que as generalizações relacionadas à inversão de unidades de extremidade – ou, mais amplamente, o privilégio das posições de extremidade – são únicas à computação da linguagem, já que são generalizações verificadas na linguagem e não verificadas, pelo menos, na música. Para explicar a aceitabilidade da inversão total na música, os autores esboçam a hipótese de que o que está em jogo na música para o fenômeno observado são os *contornos melódicos*, e não as unidades individuais das sequências e, tampouco, as unidades localizadas nas extremidades das sequências. Isso porque, com a inversão dos intervalos

e com a inversão da direção dos intervalos, os dados resultantes do comando 1 apresentam um contorno exatamente contrário ao contorno pré-transformação, o que não acontece com os dados resultantes dos comandos 2, 3 e 4.

2.3 Discussão

Neste trabalho, defende-se que, para a realização de um estudo experimental, é necessária uma discussão teórica relativa à área do conhecimento em que o estudo desse fenômeno se encaixa, tal que justifique a condução do experimento da maneira como os condutores do experimento defendem. Conseqüentemente, defende-se que, quando se opta por um estudo que compara duas atividades humanas diferentes, é necessário que se tenha uma discussão teórica a respeito de cada uma dessas atividades isoladamente, bem como uma discussão a respeito de *como* essas duas atividades podem ser comparadas. Isso exige que as teorias usadas para abordar cada uma das atividades tenham a mesma estrutura conceitual. Para o nosso estudo em particular, é necessário que se tenha experiência considerável sobre as teorias linguísticas, musicais e outras que considerem interfaces entre música e linguagem.

Um dos objetivos desta seção é argumentar em favor de uma nova maneira de conduzir os experimentos de Nevins e Endress, levando-se em conta as teorias musical e linguística descritas no capítulo 1, bem como o trabalho de Katz e Pesetsky (2011) – que será descrito na seção 2.3.3 – sobre a interface formal entre música e linguagem. A principal modificação feita nos nossos experimentos em relação aos experimentos originais foi a mudança da unidade musical que os autores tomam para a análise da transformação que está sob observação. Em vez de utilizar os intervalos musicais, utilizamos os valores relativos das notas musicais dentro do sistema tonal-harmônico, noções que serão definidas na próxima seção. Essa modificação justifica-se principalmente quando se leva em conta o tratamento dado por Katz e Pesetsky (2011) ao alinhamento entre as teorias formais da música e da linguagem. Primeiramente, entretanto, é necessário levantar uma discussão musical a respeito dos estímulos musicais.

2.3.1 Uma discussão musical sobre os estímulos musicais

A partir da década de 80, aproximadamente, começou a se desenvolver nos estudos musicais uma tradição de pesquisa experimental em percepção e cognição musical, sob a perspectiva da psicologia cognitiva. Foi nessa época que surgiu a teoria de Lerdahl e Jackendoff (1983) descrita no capítulo anterior (seções 1.2.2 e 1.3.2) e que começou a se realizar uma série de experimentos musicais, com o objetivo de descrever as intuições de ouvintes de um idioma musical particular e encontrar as regras gerais que governam o entendimento da música.

Krumhansl (1990) descreve uma série de experimentos musicais que buscam entender e descrever o funcionamento cognitivo do idioma musical denominado por ela de tonal-harmônico: "tonal" por ser organizado em redor de um centro de referência denominado tônica pela tradição musical e "harmônico" pelo fato de a harmonia (o uso de notas simultâneas) ser importante para o estabelecimento deste quadro. Embora costumemos denominar o idioma musical da música ocidental de "tonal", a grande maioria dos idiomas musicais naturais também são "tonais", no sentido de apresentarem um ponto de ancoragem perceptual, um *centro tonal* ou *tom básico*³⁰, que é o ponto de maior estabilidade na música; o ponto de repouso.³¹ Todos os outros tons de um sistema musical tonal se relacionam a sua tônica devido a seu grau de tensão, ou seja, a seu grau de desvio da estabilidade (LERDAHL; JACKENDOFF, 1983, p. 280), o que é corroborado pelos experimentos descritos por Krumhansl (1990) para o sistema musical ocidental tonal-harmônico. Assim, a literatura expõe o fato de ser muito mais comum nos sistemas musicais das culturas humanas a existência de escalas musicais

³⁰ Alguns autores optam pelo termo "tom básico", para que a expressão "centro tonal", em seu sentido amplo, não seja confundida com o centro tonal próprio do idioma musical tonal-harmônico. Outros sistemas musicais, como o sistema modal, também têm um tom básico/centro tonal, no sentido de que possuem um ponto de ancoragem perceptual.

³¹ Embora tenhamos dito que *a grande maioria* dos idiomas musicais são tonais, Lerdahl e Jackendoff (1983) consideram a propriedade de estabelecimento de um centro tonal como um universal da gramática musical (p. 280), que se define por ser um princípio disponível a todo ouvinte experiente para organizar as superfícies musicais que ouve, independentemente de qual seja o idioma musical no qual ele é experiente (p. 278). Para os autores, o fato de existir uma pequena exceção a esse aparente universal musical (a música pertencente à escola denominada "atonalismo") não elimina a universalidade cognitiva dessa propriedade musical. Assim, os autores defendem que há uma distinção crucial entre os princípios através dos quais uma peça musical é composta e os princípios através dos quais ela é ouvida. Um compositor, por exemplo, pode utilizar princípios estatísticos do movimento das moléculas para determinar suas escolhas composicionais. Isso não significa que o ouvinte se utilizará dos mesmos princípios para ouvir a música. O ouvinte se utilizará, muito provavelmente, dos princípios musicais disponíveis na sua gramática musical tácita, mesmo que a superfície musical pareça violar tais princípios, o que gera um determinado efeito estético. Neste trabalho, também assumiremos a universalidade do tonalismo em seu sentido amplo, não contemplando, portanto, discussões a respeito do atonalismo.

assimétricas, isto é, formadas por sequências de notas em que os intervalos formados entre duas notas adjacentes da escala exibem pelo menos duas distâncias diferentes. O motivo pelo qual se supõe que existe essa assimetria intervalar na maior parte dos idiomas musicais naturais é que tal assimetria torna cada tom dentro da escala único em relação aos padrões intervalares que estabelece com as outras notas da escala (PATEL, 2008, p. 20), o que facilita ao ouvinte a percepção da tônica e de outras funções do idioma musical dado.

Assim, quando inserida em um sistema musical tonal, cada nota adquire um valor único, o que significa que os tons musicais, além de terem seu valor absoluto, têm também um valor relativo, que é justamente o valor que ganham quando inseridos dentro de um sistema musical particular. Desta forma, a justificativa de Nevins e Endress para utilizar a unidade intervalar como unidade de manipulação de seu experimento musical torna-se questionável, já que sua única motivação para a utilização de intervalos musicais foi a oposição entre os intervalos musicais e os valores absolutos das notas: sendo o último não processável pelo ser humano, opta-se pelo primeiro. Tendo agora em mãos duas possibilidades processáveis (os intervalos musicais e o valor relativo das notas musicais), nós defendemos a utilização do valor relativo como unidade de manipulação para o desenho de experimento criado por Nevins e Endress. Entretanto, para que se utilize o valor relativo como unidade de manipulação dos experimentos, os estímulos musicais deverão apresentar uma informação a mais, que é a contextualização tonal, já que as notas musicais só adquirem um valor relativo quando inseridas em um sistema musical tonal. Nos próximos parágrafos, apresentaremos argumentos em favor da composição de estímulos musicais com contexto tonal, em oposição a estímulos que não se apropriam de um contexto tonal – como os criados por Nevins e Endress – o que permitirá a utilização dos valores relativos, e não dos intervalos musicais, como unidade de manipulação do experimento, o que, por sua vez, será motivado na seção 2.3.2.

Para se testar experimentalmente o conhecimento musical tácito de um ouvinte, os estímulos do experimento devem ser construídos de acordo com o sistema musical "nativo" do ouvinte – do mesmo modo que, para se testar o conhecimento linguístico tácito de um falante, deve-se criar os estímulos do experimento de acordo com a sua língua nativa, mesmo que esses estímulos sejam criados artificialmente. Os estímulos linguísticos de Nevins e Endress, por exemplo, embora não façam uso de palavras da

língua inglesa, fazem uso de *logatomas*³² da língua inglesa. Caso contrário, o que estaria sendo testado seria algo como o processamento cadeias de sons, quaisquer que sejam esses sons, e não de linguagem. Entretanto, os estímulos musicais de Nevins e Endress são melodias que não se apropriam de um contexto tonal, ou seja, melodias que não estão inseridas dentro de nenhum sistema musical sobre o qual um "ouvinte nativo" tem intuições.

Ademais, Dowling (1978) reporta um experimento em que se verifica a interdependência entre as dimensões musicais *escala* e *contorno melódico*. Seu experimento utiliza um paradigma de reconhecimento em que melodias são executadas e, em seguida, pede-se para o ouvinte julgar a relação de uma segunda melodia com a primeira melodia executada. A segunda melodia pode se assemelhar à primeira (i) por apresentar o mesmo contorno melódico e a mesma relação intervalar entre as notas, o que resulta numa melodia *exatamente* igual, mas tocada em outra tonalidade; (ii) por apresentar o mesmo contorno melódico, porém não a relação intervalar, mas ainda dentro de um contexto tonal (o que se costuma chamar de "resposta tonal"); e (iii) por apresentar o mesmo contorno melódico, porém não a relação intervalar e tampouco dentro de um contexto tonal. O resultado de seu experimento mostra que os ouvintes tiveram mais facilidade em reconhecer um mesmo contorno quando os dois *inputs* são tonais – ou seja, quando a relação entre a melodia julgada e a original é aquela descrita em (i) ou em (ii), mas não quando a relação é aquela descrita em (iii). Isso demonstra que o contorno melódico isoladamente não é informação suficiente para que os ouvintes julguem duas melodias como melodias semelhantes, e que a construção de estímulos musicais dentro de um sistema tonal tem um papel significativo na compreensão de outras dimensões musicais. Dito de outro modo, o contorno musical, quando é o mesmo para duas melodias, é reconhecido como tal quando as duas melodias são tonais. Isso significa que o contexto tonal faz parte da identidade musical, havendo maior dificuldade de processamento de *inputs* sonoros quando estes são atonais³³.

³² Logatomas são "palavras em potencial" de uma língua. Ou seja, são sequências de sons que obedecem às regras fonológicas da língua, mas que ainda não foram atreladas a um significado. Por exemplo, "bralancho" (lê-se /bra.1ã.fo/), embora não seja uma palavra do português brasileiro, poderia perfeitamente ser, pois todos os sons pertencem ao inventário de sons dessa língua, e a maneira como esses sons se concatenam obedece às regras fonológicas do português. Já "dafpe" (lê-se /'daf.pe/) não poderia ser uma palavra do português brasileiro, pois a sequência de sons /fp/, assim como a utilização do som /f/ como coda silábica, não são características da fonologia desta língua.

³³ O termo "atonal" refere-se à ausência de qualquer centro tonal/tom básico. Quando for necessário nos referirmos à ausência do centro tonal/tom básico *próprio* do idioma tonal-harmônico, utilizaremos o termo "não tonal".

Sabe-se que existe uma tática muito difundida entre músicos para reconhecer intervalos musicais tocados isoladamente: cada músico tem um inventário de melodias de referência, cada uma delas iniciando-se com um intervalo musical diferente. Ao ouvir um intervalo isoladamente (e, portanto, fora de qualquer contexto tonal), o músico busca em seu inventário aquela melodia cujo início copia o intervalo ouvido. Assim, o músico torna-se capaz de identificar o intervalo. Essa tática funciona justamente porque as músicas de referência, por serem músicas e terem um contexto tonal, facilitam o reconhecimento de um intervalo que, de outro modo, não apresenta contexto tonal. Outro exemplo sobre o contexto tonal como facilitador para a audição e entendimento musical dos ouvintes é a leitura à primeira vista de cantores. Ao deparar-se pela primeira vez com uma partitura e ao tentar entoá-la sem jamais tê-la ouvido, um cantor se sai muito melhor nesta tarefa quando a música é tonal do que quando é atonal.

Dada a discussão apresentada acima, acredita-se, no presente estudo, na hipótese de que a utilização de um contexto tonal nos estímulos musicais do modelo de experimento musical criado por Nevins e Endress lidará mais diretamente com a capacidade musical humana. Como a inserção em um sistema musical tonal dá às próprias notas musicais um valor que é processável pelo ser humano, pode-se abrir mão da utilização dos intervalos musicais como unidade de manipulação do ordenamento, e manipular a ordem dos próprios eventos musicais.³⁴ A questão neste momento é: por que abriríamos mão dos intervalos musicais? Dito de outro modo, criando-se melodias tonais, qual objeto deverá sofrer transformações no experimento: os intervalos musicais, conforme fazem os autores, ou os eventos musicais, atrelados a seu valor relativo? Na seção abaixo, será justificada a utilização dos eventos musicais como unidade de manipulação, o que é basicamente motivado por questões de interface entre música e linguagem.

2.3.2 Uma discussão de interface sobre os estímulos musicais

Como o presente trabalho lida com duas atividades cognitivas diferentes, para que seja possível a comparação dos resultados de experimentos que tratam de cada uma

³⁴ Neste trabalho, acataremos a visão de Lerdahl e Jackendoff (1983) de que o estabelecimento de um centro tonal/tom básico é um universal musical, embora exista a escola musical denominada "atonalismo" (ver nota 31). Por este motivo, a criação de estímulos musicais tonais nos experimentos justifica-se pelo fato de os estímulos estarem sendo criados de acordo com a gramática musical do homem, e não violando um de seus princípios.

delas isoladamente, é necessário levantarmos aqui uma discussão a respeito da interface entre essas duas atividades.

A linguagem humana possui um grande nível de complexidade, parcialmente devido à sua organização em várias camadas, ou seja, em unidades complexas de diferentes tipos. Por exemplo, existem princípios que governam a construção de sílabas a partir de fonemas; existem princípios que governam a construção de palavras a partir de morfemas; existem princípios que governam a construção de sintagmas a partir de palavras; existem princípios governando a atribuição de contornos entoacionais a sentenças e assim por diante. Embora essas camadas estejam todas conectadas e exercendo influência uma sobre a outra, são camadas distintas e a existência de tantas camadas é uma das propriedades da linguagem humana que a torna tão complexa. A música também possui diversas camadas diferentes, o que igualmente a torna altamente complexa. Existem princípios governando a construção de escalas a partir de um contínuo de frequência sonora; existem princípios governando a construção de melodias a partir dos sons da escala; existem princípios governando a construção de acordes; existem princípios governando as relações entre escalas diferentes e assim por diante.

Dentre tantos níveis na música e tantos níveis na linguagem, acreditamos que devemos motivar o alinhamento entre a sílaba, que é a unidade sendo manipulada no experimento linguístico, e o valor relativo das notas, que é a unidade por nós proposta como unidade de manipulação do experimento musical. Dito de outro modo, ao comparar música e linguagem criando dois experimentos que manipulem unidades de cada uma delas, deveremos apresentar uma motivação teórica para utilizar, dentre tantos níveis musicais distintos, o nível da nota musical (atrelada a seu valor relativo) para verificar a existência, na música, de fenômenos semelhantes a fenômenos que ocorrem no nível silábico na linguagem, e não em qualquer outro nível – como o *pé*, por exemplo, que é um nível prosódico assim como a sílaba, mas que não sofre operações de manipulação de ordenamento nas línguas secretas atestadas, conforme mostra o item (5d), p. 43, da lista de propriedades não atestadas em línguas secretas do tipo alterador de precedência de Bagemihl (1989)). Afinal, embora tenhamos justificado a construção de estímulos musicais com contexto tonal na seção 2.3.2, ainda não foi apresentado nenhum motivo pelo qual seja melhor a utilização do valor relativo como unidade de manipulação, e não o intervalo musical ou mesmo qualquer outra unidade musical. A justificativa para a utilização dessa unidade musical reside justamente no alinhamento

entre essa unidade musical e a sílaba. Para justificar esse alinhamento, entretanto, precisamos apresentar as linhas gerais do trabalho de Katz e Pesetsky (2011), que mostra as semelhanças e diferenças entre a teoria gerativa da música e a teoria gerativa da linguagem.

2.3.3 Katz e Pesetsky (2011)

A fim de verificar quais propriedades linguísticas e musicais poderiam ser descritas através dos mesmos mecanismos teóricos e motivados pela existência de muitos trabalhos que aproximam música e linguagem em um âmbito neurológico e cognitivo, porém não formal, Katz e Pesetsky (2011) procuram alinhar a teoria gerativa da música de Lerdahl e Jackendoff (1983) e a teoria gerativa da linguagem (em especial os modelos de Chomsky (1995a) e Chomsky (1995b)), partindo da hipótese forte de que a música e a linguagem compartilham a mesma gramática. Em suas palavras:

Todas as diferenças formais entre música e linguagem são consequências das diferenças em seus blocos de construção fundamentais (emparelhamento arbitrário de som e significado no caso da linguagem; notas musicais e combinações de notas musicais no caso da música). Em todos os outros aspectos, linguagem e música são idênticas. (KATZ; PESETSKY, 2011, p. 3, tradução nossa)

Duas grandes questões que desafiam esta hipótese são: (i) por que então música e linguagem são tão diferentes superficialmente e (ii) por que as teorias de cada uma delas também tomaram um formato tão diferente, já que a hipótese dos autores apresentada acima diz que elas são praticamente idênticas? Segundo os autores (p. 2), a razão dessas aparentes diferenças se resume aos dois seguintes motivos:

- Diferenças que a princípio parecem grandes entre música e linguagem são apenas consequências da diferença entre os “blocos de construção” que cada domínio utiliza, o que é uma real (e *única*) diferença prevista pela teoria desses autores. Pelo fato de essa diferença se mostrar na superfície da música e da linguagem, ela acaba obscurecendo semelhanças mais profundas;
- Dentro de suas respectivas áreas, a música e a linguagem são abordadas de maneiras diferentes, o que impossibilitaria, *a priori*, uma comparação entre teorias linguísticas e teorias musicais. Pensando nas teorias que alicerçam o

trabalho de Katz e Pesetsky, o diagnóstico encontrado para explicar as diferenças entre a maneira de abordar seus respectivos objetos repousa sobre (i) o fato de a teoria gerativa da música ter se desenvolvido independentemente da teoria gerativa da linguagem, embora a primeira tenha sido inspirada na última, (ii) o fato de o gerativismo na música ser antes um *parser* gerativo (ou seja, uma "máquina" que busca as leis gerais que definem a classe de possíveis análises de um dado *input* dentro de um idioma específico) do que uma gramática gerativa (ou seja, uma "máquina" que procura as leis gerais que definem a classe de *outputs* possíveis dentro de um idioma específico) e o gerativismo na linguística ser antes uma gramática gerativa do que um *parser* gerativo e (iii) mesmo que não houvesse essa diferença de nível de descrição entre as teorias, toda teoria da música e toda teoria da linguagem até hoje existentes são, obviamente, incompletas, o que pode fazer com que elas pareçam muito diferentes entre si simplesmente porque estão fragmentadas de formas diferentes. Assim, a diferença entre as teorias não é efetivamente representativa da diferença entre essas duas atividades.

Katz e Pesetsky procuram amenizar os resultados indesejáveis para sua hipótese que surgem em consequência das duas observações listadas acima, através de uma tentativa de alinhamento entre as duas teorias. Nesse alinhamento, os autores notam que existem alguns mecanismos postulados pelas duas teorias que, embora diferentes, poderiam receber um mesmo tratamento formal. Um exemplo disso é o chamado *merge interno* ou *movimento* pela teoria linguística e a chamada *retenção da cadência* pela teoria musical, o que será exposto a seguir.

Conforme apresentado na seção 1.3.1, os objetos sintáticos na linguagem se relacionam hierarquicamente, sendo que a união de dois elementos é feita através de uma operação designada *merge* e o resultado dessa união é tal que os dois elementos são linearizados lado a lado na fonologia. Por exemplo, da união de "a" com "menina", obtém-se o resultado "a menina"; a união de "a menina" (que já é resultado de uma operação de *merge*) com "corre" resulta em "a menina corre". Tanto este exemplo como o exemplo apresentado na seção 1.3.1 (ver Figura 1) são exemplos mais simples, em que os elementos sintática e semanticamente correlacionados estão "em seus devidos

lugares". Entretanto, há sentenças em que algum de seus elementos parece ter "se movido", conforme vemos nos exemplos abaixo.

(10)

- a. **Que livro** o menino leu ____ ?
- b. Foi **com eles** que eu conversei ____ ontem.
- c. O leite **que** eu comprei ____ estava azedo.

Nas três sentenças acima, temos três construções diferentes (interrogativa-QU, clivagem e oração relativa, respectivamente) em que é necessário o deslocamento de um elemento (apresentado em negrito) de seu "lugar de origem" (representado pela lacuna). Em todos os casos, o elemento deslocado é complemento do verbo: "o menino [leu algum livro]", "eu [conversei com eles]" e "eu [comprei o leite]". Evidências apontam para o fato de que, no português, deve haver uma operação de *merge* entre o verbo e o seu complemento, de tal forma que o complemento venha imediatamente à direita do verbo na linearização da estrutura, como o próprio fato de falarmos nessa ordem em sentenças assertivas comuns. Nos exemplos apresentados acima, entretanto, não é o que acontece: os complementos "que livro" do verbo "ler", "com eles" do verbo "conversar" e "que" (que é um pronome relativo que representa o sintagma nominal "o leite" dentro da oração relativa) do verbo "comprar" não estão em sua posição canônica. A maneira como a gramática gerativa-transformacional lida com esse tipo de fenômeno é através de uma operação denominada *movimento* ou *merge interno*.³⁵ Essa operação ocorre após todas as operações de *merge* e, conseqüentemente, após a união entre o verbo e o complemento – o que satisfaz as exigências semânticas e sintáticas do verbo – e desloca o complemento à posição requerida pelo tipo de estrutura (ver Figura 13).³⁶

³⁵ Ao subconjunto de modelos gerativistas que postulam a operação de movimento é dado o nome de gramática gerativa-transformacional. Há outros modelos gerativistas que buscam dar conta dos mesmos fenômenos da língua sem lançar mão do movimento, como o modelo GPSG (do inglês *generalized phrase structure grammar*).

³⁶ Algumas versões da teoria falam em *movimento* propriamente dito e outras falam em *cópia e apagamento*.

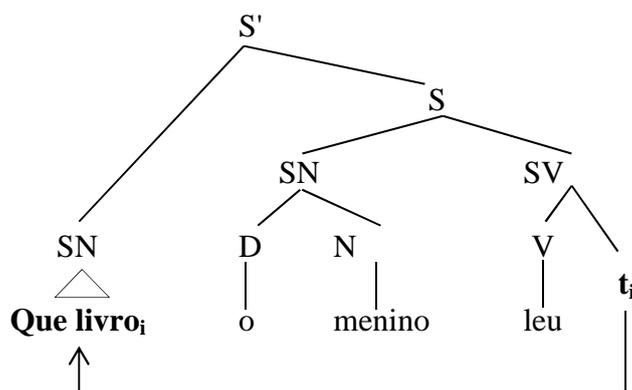


FIGURA 13: DIAGRAMA ARBÓREO DA SENTENÇA "QUE LIVRO O MENINO LEU?", EM QUE A FLECHA REPRESENTA O MOVIMENTO DO SINTAGMA INTERROGATIVO "QUE LIVRO" E O SÍMBOLO t_i REPRESENTA O VESTÍGIO DEIXADO PELO SINTAGMA EM SEU LUGAR DE ORIGEM.

O movimento apresentado na Figura 13 é denominado *movimento de sintagma*, por ser um movimento que envolve um elemento que já é resultado da operação de *merge* e que é um SX, e não um X' e tampouco um X (ver 1.3.1). Entretanto, postula-se também a existência de movimento somente de itens lexicais (de núcleos ou X), sem que o sintagma todo se mova. A esse tipo de movimento dá-se o nome de *movimento de núcleo*.

O movimento de núcleo mais notório na teoria é o movimento que envolve o verbo e a informação de tempo, modo e aspecto (grosso modo, a flexão verbal). Existem evidências para se crer que o verbo e a flexão de tempo em uma sentença sejam dois núcleos distintos, e que sujeito+verbo+objeto são unidos antes de o verbo se unir à flexão de tempo. Uma das evidências para essa afirmação é que é possível tratar da combinação sujeito+verbo+objeto como um sintagma, sem a flexão de tempo, o que é evidenciado pela pronominalização observada na seguinte sentença: "Os republicanos ganharam as eleições em 1986, mas **isso** não vai acontecer de novo em 2014". Nessa sentença, o pronome "isso" se refere ao evento de *os republicanos ganharem as eleições*, o que omite a informação de passado do verbo. Como pronomes são itens lexicais que servem justamente para substituir *sintagmas*, é possível postular a hipótese de que existe um sintagma que é formado pelo verbo e seus argumentos (sujeito e objeto), sem a flexão de tempo. Isso gera um problema na linearização em algumas línguas, inclusive o português, afinal, como podemos ver em diversas sentenças como "O João comprou um carro", o tempo está não só entre o verbo e o complemento, como forma com o verbo uma única palavra. Para dar conta desses dados, postula-se que o

verbo (V) une-se através da operação de *merge* com seu complemento (o objeto direto), formando um V' e, em seguida, o V' une-se com o sujeito, formando um SV. Por último, a informação temporal une-se ao SV, formando um ST (sintagma de tempo ou *tense*). Se linearizássemos o resultado dessas operações de *merge*, o tempo precederia todos os outros elementos da sentença, resultando em algo como "-ou o João compr- um carro". Por isso, após as operações de *merge* e antes do "envio" da estrutura sintática para a fonologia para linearização, ocorrem algumas operações de *merge interno* ou *movimento*, que fazem a ordem dos elementos da sentença ser a mesma observada na sentença falada. Assim, há o movimento à esquerda do núcleo verbal para acoplar-se à flexão de tempo e o movimento do sujeito para uma posição à esquerda, dentro do ST (o movimento do sujeito é um movimento de sintagma sobre o qual não falaremos no presente trabalho). Observe, abaixo, os movimentos presentes na sentença "O João comprou um carro".

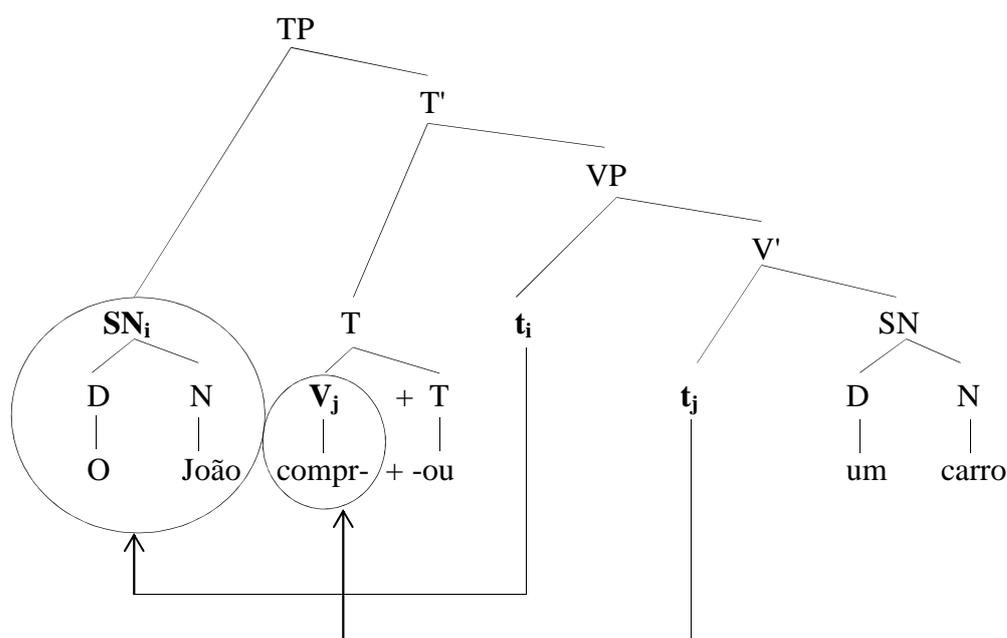


FIGURA 14: ESTRUTURA DA SENTENÇA "O JOÃO COMPROU UM CARRO", COM OS MOVIMENTOS EVIDENCIADOS.

Certamente, coincidir a ordem dos elementos da sentença com a maneira como falamos não é a única motivação para os movimentos postulados pela teoria, caso contrário, o movimento seria um mecanismo extremamente *ad hoc*. Existem fenômenos que evidenciam a existência desse movimento. Observe, por exemplo, as duas sentenças em francês em (11). Na língua francesa, em construções negativas, é possível verificar o

movimento do verbo (sublinhado) para a flexão de tempo (-á) em relação à partícula negativa (em negrito), como é possível observar na sentença (11a). Entretanto, quando não há informação temporal (em construções infinitivas, por exemplo, como (11b) abaixo), é possível verificar que o verbo não se move em relação à partícula negativa:

(11)

a. *La fille n'acheter-á **pas** _____ le livre.* (A menina não comprará o livro.)



b. *Il y a trois raisons pour ne _____ **pas** acheter le livre.* (Há três razões para não comprar o livro.)



Dessa forma, parece existir alguma exigência da flexão de tempo para "receber" o verbo. Quando a flexão é de outra natureza (infinitiva, por exemplo), o verbo não precisa se mover para satisfazer a exigência da flexão e, portanto, permanece em sua posição de origem. Dadas essas considerações a respeito do movimento de núcleo, podemos extrair algumas propriedades características dessa operação.

1. Após o núcleo X de um sintagma SX ter sofrido movimento, X é **pronunciado adjacientemente** ao núcleo de um sintagma que é mais alto na estrutura arbórea.
2. O resto do SX permanece um **sintagma independente** que se comporta exatamente como um sintagma cujo núcleo não se moveu.
3. O movimento é obrigatório. O movimento de V para T satisfaz uma *necessidade* de algum elemento na estrutura. O fato de o movimento do verbo para T ser desencadeado por propriedades do próprio T pode ser visto a partir do fato de que um verbo no infinitivo não se move dessa forma. Assim, a relação T-V envolve alguma **alteração nos traços de T**.
4. O núcleo que sofre movimento acaba ficando **fortemente acoplado** ao núcleo a que foi movido. Os dois núcleos acabam se comportando como uma única palavra para os processos seguintes da gramática. (KATZ; PESETSKY, 2011, p. 40)

Segundo Katz e Pesetsky, o mecanismo da teoria de Lerdahl e Jackendoff denominado *retenção de cadência* apresenta exatamente as mesmas propriedades que

aquelas acima verificadas para o movimento de núcleo na linguagem. As cadências são sequências de eventos musicais que desempenham um papel especial na organização estrutural das peças musicais. As cadências funcionam como sinalizadores que marcam a fronteira direita de unidades estruturais – como frases musicais e seções dentro de uma peça. A teoria musical tradicional reconhece uma variedade de tipos de cadência, classificados através de critérios harmônicos, rítmicos e melódicos. Entretanto, para a presente discussão, contemplaremos apenas a chamada *cadência perfeita*. A cadência perfeita é uma sequência de dois acordes pertencentes à mesma escala, em que V (a dominante) é uma tríade (acorde de três sons) cuja fundamental (nota a partir da qual o acorde se forma) é mais aguda por um intervalo de quinta justa de I (a tônica). Embora seja tradição focar na relação de quinta justa entre as unidades da cadência como propriedade central de uma cadência perfeita, segundo Katz e Pesetsky tais cadências são, em realidade, caracterizadas por um conjunto de propriedades aparentemente não relacionadas. São elas:

1. V deve ser **adjacente** a I;
2. V deve ter um conjunto normal de dependentes sintáticos, linearizados normalmente. Assim, V também é um núcleo de **seu próprio sintagma**. A presença de pelo menos um dependente de V é extremamente comum, o que representaremos como IV (subdominante). Há evidências para o fato de que, na sequência cadencial IV-V-I, IV é dependente de V, e não de I, como:
 - o fato de alguns acordes que funcionam como subdominante nunca aparecerem antes de I, sem a presença de V (como o acorde I_4^6);
 - o fato de, na redução (ver 1.3.2) dessa sequência cadencial, nunca existir um passo em que se reduz V, restando IV-I; apenas os passos IV-V-I, V-I e I;
3. A participação de V em uma cadência é crucial para o **estabelecimento de I como a tônica**;
4. Na PR (ver 1.3.2), V se comporta como se fosse estruturalmente diretamente subordinada a I (ambas hierarquicamente altas), mesmo quando seu nível de proeminência devesse motivar seu ligamento à estrutura em um nível mais baixo. Essa propriedade, particularmente, é aquilo que Lerdahl e Jackendoff chamam de **retenção de cadência**. (KATZ; PESETSKY, 2011, p. 42)

É bastante evidente que as propriedades 1 e 2 da cadência parecem corresponder às propriedades 1 e 2 do movimento de núcleo. Quanto à propriedade 3 tanto do movimento de núcleo como da cadência, embora sua semelhança não seja óbvia, Katz e Pesetsky defendem que o *estabelecimento de tonalidade* na música é equivalente à *alteração de traço* na linguagem. O movimento de um núcleo para outro (muito possivelmente *qualquer* tipo de movimento, em realidade), conforme já visto, acontece devido a exigências feitas pelo núcleo ao qual o outro é movido. Dito de outro modo, o movimento ocorre porque o núcleo superior na estrutura apresenta algum traço que precisa ser satisfeito através do acoplamento de outro núcleo a ele. Da mesma forma, para Katz e Pesetsky, a propriedade de estabelecimento da tonalidade surge devido a um traço em I que precisa ser satisfeito. Assim, V se move para I, atribuindo a I o traço [+tônica]. Como consequência disso, quaisquer nós na estrutura que forem dominados por I com o traço [+tônica] será interpretado como *pertencente à tonalidade de I*.

Por último, a propriedade 4 do movimento de núcleo e da cadência também não parece coincidir, em princípio, mas Katz e Pesetsky defendem que a retenção da cadência corresponde ao acoplamento dos dois núcleos envolvidos. Quando dois elementos adjacentes formam uma cadência, a teoria de Lerdahl e Jackendoff "promove" o elemento V da cadência a uma posição mais proeminente na estrutura TSR, mesmo que fatores rítmicos que servem de *input* para a construção dessa estrutura impelisse V a ocupar uma posição bastante encaixada (hierarquicamente baixa) na estrutura. Com o procedimento da retenção da cadência no processo de redução, permite-se que os dois núcleos que compõem uma cadência se comportem como **um único núcleo**. Em suas palavras:

Na cadência perfeita ou na cadência deceptiva, a dominante, assim como a resolução em I ou VI, deve ser retida para que a redução faça sentido. A sequência de dois eventos resultante age como uma unidade gramatical. Para permitir que tal redução seja bem-formada, a condição para selecionar o núcleo de um *time-span* deve ser enriquecida para permitir que uma sequência de eventos formando uma cadência sirva como núcleo. (LERDAHL; JACKENDOFF, 1983, p. 155-6; tradução nossa)

De acordo com a citação, o mecanismo utilizado para dar conta do acoplamento dos dois eventos que formam uma cadência não é através do movimento de núcleo, que os une, e sim através da melhoria da condição para selecionar um núcleo, de forma a permitir a seleção de *dois* núcleos, quando esses formam uma cadência. Parece claro

que se trata de um mecanismo teórico diferente para lidar com um fenômeno muito semelhante aos fenômenos que envolvem movimento de núcleo na linguagem. Na estrutura arbórea, em vez de essa proximidade entre os dois núcleos ser representada pelo movimento, como é na linguagem, a cadência é representada por uma circunferência conectando V e I na estrutura (ver Figura 15).

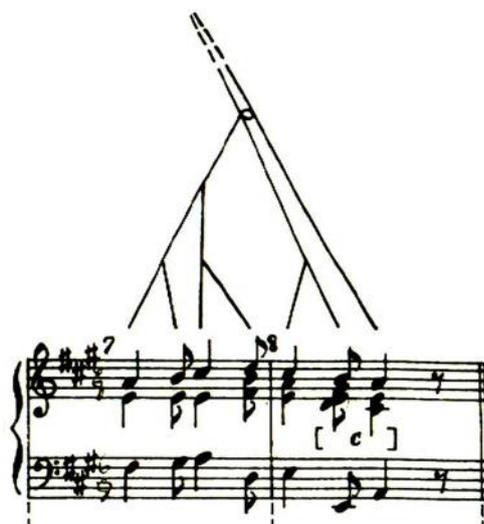


FIGURA 15: REPRESENTAÇÃO DA CADÊNCIA NA ESTRUTURA TSR. FONTE: LERDAHL; JACKENDOFF, 1983, p. 156.

Para além da observação de que o movimento de núcleo e a retenção da cadência são mecanismos bastante semelhantes e que poderiam ser tratados da mesma forma pelas duas teorias, há também outra propriedade em comum entre as teorias que é posta em evidência por Katz e Pesetsky e que é particularmente importante para a discussão proposta no presente trabalho. Katz e Pesetsky notam que tanto a teoria musical quanto a teoria linguística postulam dois componentes diferentes: um que codifica relações estruturais entre elementos que não estão necessariamente adjacentes e outro que invoca noções de proeminência relativa. O primeiro deles, na linguagem, é o componente sintático e, na música, a redução prolongacional (PR); e o segundo deles, na linguagem, é o componente fonológico (prosódico) e, na música, a redução do *time-span* (TSR) (ver seções 1.3.1 e 1.3.2). Os próprios Lerdahl e Jackendoff (1983, capítulo 12), chamam a atenção para o fato de que a estrutura prosódica e a TSR compartilham muitas propriedades formais. Ademais, os dois pares de estruturas (sintaxe e prosódia, por um lado, e TSR e PR, por outro) têm relações muito fortes, embora não sejam isomórficas; há um mapeamento de uma para a outra que mantém algumas relações estruturais e

modifica outras. Assim, dada a hipótese de identidade, Katz e Pesetsky afirmam que a TSR e a estrutura prosódica são uma mesma estrutura, construídas a partir dos mesmos princípios – apesar de cada uma delas estar lidando com unidades de natureza diferente – e que a PR e a estrutura sintática são a mesma estrutura, também construídas a partir dos mesmos princípios.

A sílaba, que é o nível que sofre as operações do experimento de Nevins e Endress, é um nível linguístico pertencente à hierarquia prosódica e, conforme algumas teorias prosódicas propõem (SELKIRK, 1984; NESPOR; VOGEL, 1986), é o nível mais baixo dessa hierarquia. Assim, parece razoável a lógica segundo a qual, para saber quais unidades musicais estão sujeitas a sofrer as mesmas operações que a sílaba, basta encontrarmos o menor constituinte da TSR, que é a estrutura musical alinhada à estrutura prosódica da linguagem por Katz e Pesetsky. A TSR hierarquiza objetos musicais e agrupamentos de objetos musicais (e jamais intervalos entre objetos musicais), sendo o mais baixo na hierarquia os eventos musicais únicos, ou seja, as notas musicais. Como a grande maioria dos seres humanos não processa o valor absoluto das notas musicais, conforme já apontado, o que nos resta para ser tomado como o menor constituinte da TSR é o valor relativo de cada um dos eventos musicais. O intervalo musical, dada essa discussão, deixa de ser uma opção para se alinhar com a sílaba.

2.3.4 *Por que um alinhamento bottom-up?*

Acabamos de defender que se deve alinhar sílaba e evento musical pelo fato de essas duas unidades serem **o constituinte mais baixo da mesma hierarquia**. A caracterização de um constituinte como sendo *o mais baixo* dentro de uma hierarquia carrega implicitamente a ideia de que o alinhamento que estamos fazendo entre as duas hierarquias (TSR, para a música, e prosódia, para a linguagem) é posicional. Isso quer dizer que se um constituinte musical *x* e um constituinte linguístico *y* ocupam a posição das estruturas TSR e prosódica, respectivamente, esses dois constituintes são correspondentes e estão, portanto, sujeitos à aplicação das mesmas regras da gramática, dada a hipótese de identidade de Katz e Pesetsky. Suponhamos, entretanto, que, por algum motivo relativo às propriedades dos “blocos de construção” da música, o número de constituintes da hierarquia prosódica da música seja menor do que o número de

constituintes da hierarquia prosódica da linguagem. Se assim for e se de fato o alinhamento for feito de acordo com a posição do constituinte dentro da hierarquia, temos duas possibilidades de direção de alinhamento que gerarão alinhamentos diferentes: *bottom-up* e *top-down*. Se fizermos um alinhamento *bottom-up*, o constituinte mais baixo da hierarquia prosódica musical vai coincidir com o constituinte mais baixo da hierarquia prosódica linguística (ver Figura 16a), o que não entra em conflito com a nossa suposição de que a sílaba e o evento musical sejam correspondentes. Porém, se fizermos um alinhamento *top-down*, é o constituinte mais alto das duas hierarquias que coincidirão (ver FIGURA 16b), e não o mais baixo, dada a diferença em número de constituintes. Com esse segundo tipo de alinhamento, a sílaba e o evento musical não seriam correspondentes. Para que possamos manter o alinhamento proposto e desejado, precisaríamos ter uma justificativa para que sua direção seja *bottom-up* e não *top-down*.

a.	b.	c.
CL5	CL1 CM1	CL CM
CL4 CM4	CL2 CM2	CL CM
CL3 CM3	CL3 CM3	CL
CL2 CM2	CL4 CM4	CL CM
CL1 CM1	CL5	CL CM
Alinhamento <i>bottom-up</i>	Alinhamento <i>top-down</i>	Alinhamento sem direção

FIGURA 16: POSSIBILIDADES DE ALINHAMENTO ENTRE OS CONSTITUINTES DE DUAS ESTRUTURAS DIFERENTES. *CL* CORRESPONDE A CONSTITUINTE LINGUÍSTICO E *CM* CORRESPONDE A CONSTITUINTE MUSICAL.

Entretanto, é possível postular ainda uma terceira possibilidade de alinhamento entre as estruturas, que não faz referência à sua posição, e sim às propriedades inerentes dos próprios constituintes. Isso quer dizer que se um constituinte musical x e um constituinte linguístico y têm a propriedade intrínseca ϕ , esses dois constituintes são correspondentes. Dessa forma, a “falha” existente devido ao diferente número de constituintes não se encontraria necessariamente nas extremidades, mas poderia se encontrar entre dois constituintes em qualquer posição da hierarquia (ver FIGURA 16c). Se assim for, mesmo com um número distinto de constituintes, o constituinte mais baixo da TSR pode ser correspondente ao constituinte mais baixo da estrutura prosódica e o constituinte mais alto da TSR ser igualmente correspondente ao constituinte mais alto da estrutura prosódica, sem gerar um paradoxo. Ou seja, a posição do constituinte

dentro da hierarquia seria parcialmente irrelevante (*parcialmente*, porque não se esperaria encontrar correspondências cruzadas entre constituintes, o que violaria relações de dominância) e o que seria relevante seriam as propriedades desses constituintes. Considerando-se esse tipo de alinhamento, que é mais interessante do que o alinhamento puramente posicional, para manter a correspondência entre a sílaba e o evento musical, precisaríamos encontrar ao menos uma propriedade em comum entre essas unidades. A próxima seção é voltada à descrição de uma propriedade que afirmamos ser comum entre as duas unidades que estamos alinhando.

2.3.4.1 A atribuição de estrutura métrica

A estrutura métrica é uma estrutura que codifica a alternância entre pulsos fortes e fracos. Para as teorias musicais e linguísticas que postulam a existência do componente métrico na gramática, os mecanismos formais utilizados para representá-lo costumam ser os mesmos (com a exceção irrelevante da representação gráfica que, na música, são pontos e, na linguagem, são *x*'s), conforme vemos nas Figuras 17 e 18.



FIGURA 17: REPRESENTAÇÃO MÉTRICA MUSICAL. FONTE: LERDAHL; JACKENDOFF, 1983, p. 98.

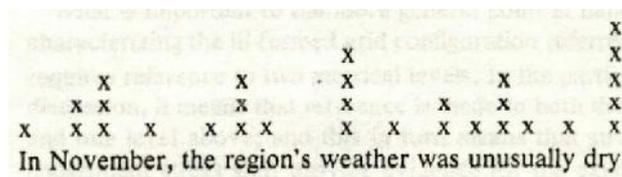


FIGURA 18: REPRESENTAÇÃO MÉTRICA LINGUÍSTICA. FONTE: SELKIRK, 1984, p. 46.

A primeira regra de boa-formação da estrutura métrica da teoria de Lerdahl e Jackendoff (1983) diz que *todo ponto de ataque deve ser associado com um beat no nível mais baixo da estrutura métrica* (p. 69). O *beat* é cada um desses pontos/*x*'s

alinhados às notas e às sílabas que vemos nas Figuras 17 e 18. A intuição por trás da representação de *beats* através de pontos é o fato de que os *beats* não têm duração (p. 18). O ponto de ataque é o momento exato em que se inicia o som de alguma nota musical. Como *todo* ponto de ataque deve ser associado com um *beat*, toda nota deverá ser associada com um *beat*, pois, supondo que uma nota está soando, no exato momento em que se muda de nota, mesmo sem que haja nenhum momento de silêncio entre as duas notas, necessariamente um *beat* deverá ser alinhado com a nova nota. Ou seja, sempre que há uma nota há um *beat*. Uma nota está sempre, necessariamente, associada a um *beat*.

A primeira regra de construção da grade métrica de Selkirk (1984) exige que se *alinhe apenas um demibeat com toda sílaba* (p. 57). Ou seja, toda sílaba deve estar associada a pelo menos um ponto na estrutura métrica, chamado pela autora de *demibeat*. Na teoria prosódica em questão, *demibeat* é o nome dado aos pontos presentes no primeiro nível da estrutura métrica (esse a que todas as sílabas devem estar alinhadas), sendo que os pontos em todos os outros níveis são chamados de *beat*. Lerdahl e Jackendoff não fazem essa distinção terminológica, o que não influencia no fato de que, para as teorias em questão, todas as notas e todas as sílabas devem estar metricamente representadas nas estruturas.

Além das semelhanças acima apontadas, tanto na teoria musical de Lerdahl e Jackendoff como teoria prosódica de Selkirk, às vezes pode haver *beats* que não estão alinhados com nenhum evento na estrutura superficial. Afinal, as regras das teorias preveem apenas que toda sílaba/nota deve estar alinhada com um *beat*, mas não que todo *beat* deva estar alinhado com uma sílaba/nota. Na teoria de Selkirk, tais *beats* são denominados "posições silenciosas na grade" (do inglês *silent grid positions*) (cap. 6). Na teoria de Lerdahl e Jackendoff não há um tratamento especial para esses *beats* silenciosos.

Dada essa propriedade em comum entre a nota musical e a sílaba, o alinhamento entre esses dois constituintes se justifica mesmo se houver, na linguagem, constituintes mais baixos ainda do que a sílaba, conforme propõem teorias diferentes das teorias de Selkirk (1984) e Nespor e Vogel (1986). Afinal, embora consideradas pelas autoras como as menores unidades visíveis para operações prosódicas, as sílabas apresentam uma estrutura interna de segmentos, e os segmentos são formados por traços distintivos que se relacionam de maneira ordenada, conforme defende a fonologia autosssegmental

(GOLDSMITH, 1976). É possível que essas unidades menores sejam também visíveis na estrutura prosódica, o que impossibilitaria o alinhamento da sílaba com a nota musical em um alinhamento *bottom-up*, mas que não impossibilitaria esse alinhamento se levado em consideração o alinhamento não-posicional, ou seja, se levado em consideração o fato de que as unidades apresentam propriedades em comum (no caso, a atribuição de estrutura métrica).

Dada a discussão levantada nesta seção, a fim de verificar a hipótese levantada por Nevins e Endress a respeito da especificidade à linguagem ou não do fenômeno observado em seu experimento linguístico, optamos por reconduzir o experimento musical utilizando como unidade de manipulação os valores relativos das notas musicais, e não os intervalos musicais, como é feito pelos autores.

O leitor possivelmente está se perguntando: mas se os resultados de Nevins e Endress foram estatisticamente relevantes, apontando, inclusive, para uma hipótese plausível sobre o contorno melódico (ver final da seção 2.2.3), isso não torna a recondução desses experimentos desnecessária? A próxima seção apresenta uma hipótese capaz de demonstrar de duas formas distintas o motivo pelo qual Nevins e Endress chegaram a um resultado estatisticamente relevante.

2.3.5 Os dois vieses de Nevins e Endress

Quando analisados em detalhes, os dados musicais criados por Nevins e Endress para seu experimento em realidade não possuem a ambiguidade relatada pelos autores relacionada às possibilidades de análise dos dados de julgamento. Segundo eles, os dados de familiarização, que são aqueles com apenas três unidades musicais intervalares, são compatíveis com pelo menos quatro comandos, já listados em (8) na seção 2.2.3, p. 48, e repetidos em (12) abaixo.

(12)

- **Comando 1:** Inverta a ordem de todos os intervalos;
- **Comando 2:** Permute o primeiro intervalo com o último intervalo;
- **Comando 3:** Permute o último intervalo com o antepenúltimo intervalos;
- **Comando 4:** Permute a cada dois intervalos (*i.e.* j com $j+2$).

Entretanto, em decorrência de um detalhe da formação dos dados, essa ambiguidade desaparece de forma dupla, o que será explicado nas seções 2.3.5.2 e 2.3.5.3. Antes dessa explicação, porém, é necessário fazer algumas considerações sobre a maneira como foram criados os dados do experimento de Nevins e Endress, o que veremos a seguir.

2.3.5.1 A falsa inversão

Nas palavras de Nevins e Endress, os estímulos musicais foram criados da seguinte maneira:

A razão lógica para usar melodias de quatro tons em vez de sequências de três itens como no Experimento 1 foi que os participantes geralmente codificam intervalos entre tons, e não suas alturas absolutas; em termos de intervalos, entretanto, nós usamos novamente sequências de três itens. [...] **Adicionalmente, como os intervalos são invertidos quando tocados ao contrário (uma oitava ascendente, por exemplo, se torna uma oitava descendente), os intervalos também foram invertidos.** (Nevins e Endress, p. 7, tradução nossa, grifo nosso)

O fato de Nevins e Endress utilizarem o intervalo musical como unidade de manipulação torna o intervalo uma unidade atômica, por assim dizer. Dessa forma, as notas de que são compostos os intervalos cristalizam-se no momento em que o intervalo é formado. Com isso, inverter a direção de cada um dos intervalos ao mexer em seu ordenamento, conforme os autores dizem no trecho grifado da citação acima, torna-se sem sentido. Explicamos:

Se tomarmos as notas musicais como unidades atômicas, ao juntarmos duas notas em uma sequência linear, como *dó3 mi3*, essas duas notas formam entre si, "epifenomenicamente", um intervalo. Neste caso particular, o intervalo é uma terça maior *ascendente*. Se invertermos a ordem dessas duas notas, *mi3 dó3*, a classificação intervalar se mantém, exceto pela sua direção, ou seja, o intervalo continuará sendo uma terça maior, mas *descendente*. Da mesma forma, se tomarmos uma sequência linear de três notas, como *ré3 fá3 sol3*, entre cada duas notas forma-se um intervalo: *ré3* com *fá3* forma um intervalo de terça menor ascendente e *fá3* com *sol3* forma um intervalo de segunda maior ascendente. Ao inverter completamente a ordem das notas, isto é, se de *ré3 fá3 sol3* obtivermos *sol3 fá3 ré*, nós também estamos, não coincidentemente,

invertendo a ordem dos intervalos e, além disso, a direção desses intervalos: de *sol3* para *fá3* temos uma segunda maior descendente e de *fá3* para *ré3*, temos uma terça menor descendente. Veja a Figura 19 abaixo:

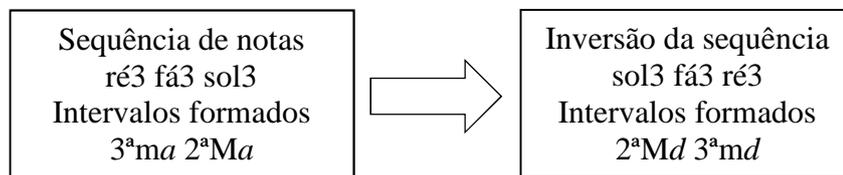


FIGURA 19: RESULTADO DO COMANDO "INVERTA A ORDEM DE TODAS AS NOTAS".

Ou seja, sempre que é feita a inversão total de uma sequência linear de **notas musicais**, é obtida, "epifenomenicamente", a inversão de ordem dos intervalos formados entre essas notas com a mudança da direção desses intervalos. Essa aparente coincidência ocorre porque a inversão total não muda a relação de adjacência entre as notas, e o intervalo musical se caracteriza justamente pela relação de adjacência entre as notas. Entretanto, como invertemos a ordem dessas notas, o intervalo é, obviamente, realizado ao contrário, ou seja, na direção oposta.

Por outro lado, se tomarmos um intervalo musical como unidade atômica, supõe-se que sua direção está cristalizada dentro dele e que o epifenômeno desta vez serão as notas que surgirão ao inverter a ordem dos intervalos (lembrando que a direção do intervalo muda ao inverter a ordem das *notas* de que o intervalo é composto). Por exemplo, se tomarmos a sequência linear de intervalos *3ªma 2ªMa* e escolher aleatoriamente uma nota inicial (neste caso escolherei a nota *ré3*), nós teremos a sequência de notas *ré3 fá3 sol3*. Agora, se a ordem desses **intervalos** for invertida, obteremos a ordem *2ªma 3ªMa* e, de novo, escolhendo aleatoriamente uma nota inicial (neste caso escolheremos, a bem da simetria, a última nota da sequência: *sol3*), obteremos a sequência de notas *sol3 lá3 dó#4*. O epifenômeno, desta vez, são as notas obtidas através da inversão de ordem dos intervalos. Veja na Figura 20 abaixo:

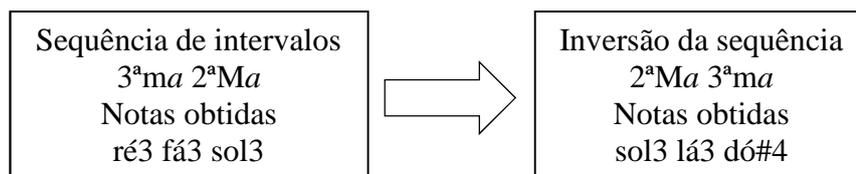


FIGURA 20: RESULTADO DO COMANDO "INVERTA A ORDEM DOS INTERVALOS".

Dada essa constatação, poderíamos parafrasear a citação de Nevins e Endress que diz que "como os intervalos são invertidos quando tocados ao contrário (uma oitava ascendente, por exemplo, se torna uma oitava descendente), os intervalos também foram invertidos" da seguinte forma: como os intervalos são invertidos quando **as notas são invertidas** (uma oitava ascendente [dó3 dó4], por exemplo, se torna uma oitava descendente [dó4 dó3]), os intervalos também foram invertidos. Assim vê-se que o motivo pelo qual Nevins e Endress inverteram a direção dos intervalos é um falso motivo, dado que suas unidades musicais atômicas são os próprios intervalos, e não as notas musicais. Note que o exemplo utilizado pelos autores no trecho acima é composto de apenas um intervalo (uma oitava ascendente tornando-se uma oitava descendente), o que é irrelevante para o fenômeno de inversão da ordem de intervalos, já que é ilógico inverter a ordem dos intervalos de uma sequência composta de um único intervalo.

Para fazer um paralelo com a linguagem, a criação dos dados musicais de Nevins e Endress é comparável à seguinte descrição de uma possível criação de dados para o experimento linguístico: "como as sílabas são invertidas quando os sons são falados ao contrário (a sílaba [pas], por exemplo, se torna [sap]), as sílabas também foram invertidas". Claramente, os dados não devem ser tratados dessa forma.

De que maneira tal falha desambigua os dados?

2.3.5.2 A não-ambiguidade dos dados

Conforme dito acima, a maneira como os dados de Nevins e Endress foram criados elimina a ambiguidade dos dados de duas maneiras. Nesta seção, explicaremos uma das maneiras e, na seção que se segue, mostraremos uma segunda maneira de eliminar a ambiguidade que se revela mesmo quando eliminamos nossos pressupostos e mantemos a argumentação de Nevins e Endress.

O fato de a ordem dos intervalos e suas respectivas direções serem completamente invertidos cria, em realidade, dados em que as **notas musicais** são invertidas, conforme visto na seção anterior, eliminando, portanto, a hipótese de que as unidades sendo manipuladas são os intervalos – ou talvez não eliminando, mas tornando essa hipótese mais complexa e, portanto, menos preferível, já que seriam dois os comandos: inverta a ordem dos intervalos e inverta a direção dos intervalos. Assim, os dados de familiarização do experimento apontam para a hipótese de que a ordem das *notas* é o que se está invertendo³⁷. Nos dados de julgamento, o mesmo ocorre para aqueles que são resultantes do primeiro comando ("inverta a ordem de todos os intervalos" que em realidade deveria ser "inverta a ordem de todas as notas"), mas os dados resultantes dos outros comandos sequer chegam perto. Isso porque, ao criar os dados da maneira como os autores criaram, nem mesmo as notas utilizadas permanecem as mesmas, independentemente de qual nota da sequência original utilizamos para iniciar a sequência transformada, conforme mostra a Tabela 1.

³⁷ Certamente, os dados de familiarização também apontam para a hipótese de que é o *contorno melódico* que está sendo invertido. Não contemplaremos essa hipótese na discussão aqui proposta.

Input	Intervalos					Notas					
	1	2	3	4		dó3	ré3	sib2	mib2	solb2	
	2 ^a Ma	3 ^a Md	5 ^a Jd	3 ^a ma							
Comando 2	4	2	3	1		Nota 1	dó3	lá2	fá2	sib1	láb1
						Nota 2	ré3	si2	sol2	dó2	sib1
						Nota 3	sib2	sol2	mib2	láb1	solb1
	3 ^a md	3 ^a Md	5 ^a Jd	2 ^a Md		Nota 4	mib2	dó2	láb1	réb1	dób1
						Nota 5	solb2	mib2	dób2	solb1	fáb1
Comando 3	1	4	3	2		Nota 1	dó3	ré3	si2	mi2	sol#2
						Nota 2	ré3	mi3	dó#3	fá#2	lá#2
						Nota 3	sib2	dó3	lá2	ré2	fá#2
	2 ^a Ma	3 ^a md	5 ^a Jd	3 ^a Ma		Nota 4	mib2	fá2	ré2	sol1	si1
						Nota 5	solb2	láb2	fá2	sib1	ré2
Comando 4	3	4	1	2		Nota 1	dó3	sol3	mi3	ré3	fá#3
						Nota 2	ré3	lá3	fá#3	mi3	sol#3
						Nota 3	sib2	fá3	ré3	dó3	mi3
	5 ^a Ja	3 ^a md	2 ^a M d	3 ^a Ma		Nota 4	mib2	sib2	sol2	fá2	lá2
						Nota 5	solb2	réb3	sib2	láb2	dó3

TABELA 1: RESULTADO DO COMANDO "INVERTA A ORDEM DE TODOS OS INTERVALOS", INICIANO-SE O *OUTPUT* COM CADA UMA DAS NOTAS PRESENTES NO *INPUT*.

Assim, supondo que a interpretação dos dados de familiarização pelos participantes seja sobre a manipulação de notas, e não de intervalos, esses dados de familiarização *não criam nenhuma* ambiguidade que seja desambiguizada com os dados de julgamento criados pelos autores. Dito de outro modo, se é que há alguma ambiguidade nos dados de familiarização, não é a mesma ambiguidade de que falam os autores, já que, de acordo com a hipótese em que se baseia esta seção, os dados de

familiarização foram construídos de tal forma que só é possível para os ouvintes interpretarem-nos como uma inversão de ordem de *notas musicais*, e não de *intervalos musicais*, e já que os dados de julgamento, tendo sido construídos com base em uma ambiguidade intervalar não processada pelos ouvintes, são absolutamente caóticos à luz da interpretação de que a transformação dos dados de familiarização resulta da inversão das notas musicais – com exceção daqueles resultantes do comando 1, que se mantêm compatíveis com a hipótese da inversão de notas. Por isso, ao serem expostos aos dados de julgamento após terem ouvido os de familiarização, os informantes do experimento de Nevins e Endress julgaram melhores os dados resultantes do primeiro comando, dada a completa desconexão entre os dados de familiarização e os dados resultantes dos comandos 2, 3 e 4.

2.3.5.3 *E se a falsa inversão não for falsa?*

Nas duas seções anteriores, argumentou-se que a ambiguidade que Nevins e Endress afirmam existir em seus dados em realidade não existe, pois a maneira como os autores criam os dados musicais possibilita uma interpretação não contemplada por eles, a saber, que as unidades manipuladas em seus dados são as notas e não os intervalos, que é possivelmente a estratégia utilizada pelos informantes, dado que a interpretação relacionada aos intervalos é cognitivamente mais complexa por necessitar de duas operações, e não apenas de uma. Entretanto, podemos supor que, por algum motivo, a mudança de direção dos intervalos não é uma operação diferente da inversão do ordenamento dos intervalos, e sim, como defendido por eles, uma operação que se segue da mudança de ordenamento. Embora não se tenha encontrado motivos para defender a fusão das operações mudança de direção e mudança de ordenamento, suponhamos, a bem da argumentação, que a maneira como os autores criaram os dados é de fato a melhor maneira para criar dados com manipulação do ordenamento de intervalos musicais. Ou seja, suponhamos que mudança de ordem e mudança de direção de intervalos vêm em um só pacote.

Mesmo com a suposição feita no parágrafo acima, os dados de familiarização Nevins e Endress não são ambíguos. Nos dados de familiarização, a ordem e a direção de *todos* os intervalos são invertidas. Como são apenas três intervalos, o intervalo central obviamente acaba ficando exatamente no mesmo lugar após a inversão de ordem

de todos os intervalos. Logo, pela sua posição, o intervalo central pode, ambigualmente, ter ou não ter sofrido alteração de ordem, que é justamente o que causa a ambiguidade observada pelos autores. Assim, numa sequência de três unidades [1 2 3], ao obter a inversão [3 2 1], nós podemos tanto estar lidando com a inversão de *todas* as unidades, como com a inversão da primeira com a última unidade, o que, em uma sequência de quatro unidades [1 2 3 4], não é ambíguo, pois para cada um dos tipos de inversão obtemos um resultado diferente: [4 3 2 1] e [4 2 3 1], respectivamente.

Embora pela sua posição o intervalo central cause ambiguidade na maneira como interpretamos a generalização por trás das inversões, o fato de a direção do intervalo mudar quando o intervalo sofre mudança de ordem desambiguiza a interpretação: quando o intervalo central tem sua direção invertida, ele sofreu alteração de ordem e, quando o intervalo central não tem sua direção invertida, ele permanece na mesma posição de origem. Assim, numa sequência de três unidades com direção ascendente [1 2 3], se fizermos a inversão de acordo com o comando "inverta a ordem de todos os intervalos", obteremos o resultado [3 2 1] e, se fizermos a inversão de acordo com o comando "inverta o primeiro com o último intervalo", obteremos o resultado [3 2 1]. Como o intervalo carrega uma informação intrínseca sobre a sua mudança ou não de ordem (a mudança em sua direção), ele próprio já nos mostra se sofreu alteração ou não.

A ausência de ambiguidade explicitada no parágrafo acima corresponde à falta de ambiguidade entre sequências de intervalos resultantes dos comandos 1 e 2. Vejamos como ficariam os resultados das sequências de três intervalos após aplicados os comandos 3 e 4. Tomando como *input* a mesma sequência do parágrafo anterior [1 2 3], a sequência resultante do comando 3 "inverta o último com o antepenúltimo intervalo" seria [3 2 1] e a sequência resultante do comando 4 "inverta a cada duas notas" seria [3 2 1]. Observando todos esses resultados, fica claro que não existe ambiguidade entre o comando 1, por um lado, e os comandos 2, 3 e 4, por outro. Porém, entre os comandos 2, 3 e 4 a mesma ambiguidade existe, já que o intervalo central mantém-se inalterado. Essa ambiguidade entre apenas os três últimos comandos pode ser desambiguada quando lidamos com sequências de três intervalos não só pela ordem como também pela direção: fica claro exatamente quais são os intervalos que estão "em movimento". Para a sequência [1 2 3 4], obtemos [4 2 3 1], [1 4 3 2] e [3 4 1 2] para os comandos 2, 3 e 4, respectivamente.

Os estímulos de familiarização criados por Nevins e Endress são todos do tipo [1 2 3] [3 2 1]³⁸. Isso quer dizer que os dados de familiarização só são compatíveis com o comando 1, já que todos os intervalos têm suas direções invertidas na resposta ao *input*. Dessa forma, temos um segundo possível motivo pelo qual os resultados de Nevins e Endress mostram que os informantes preferem apenas os dados resultantes do primeiro comando: os dados de familiarização são logicamente compatíveis *apenas* com o primeiro comando.

Dada a discussão presente nesta seção, nós replicamos o experimento musical de Nevins e Endress, com modificações – como a mudança de unidade musical em manipulação que foi motivada neste capítulo – que serão apresentadas e descritas no próximo capítulo. O próximo capítulo, portanto, volta-se para a descrição dos experimentos por nós conduzidos, com base nas teorias e discussões apresentadas.

³⁸ Ao dizer que todos os dados são desse tipo, não queremos dizer que os dados de *input* são necessariamente formados apenas de intervalos ascendentes e que as respostas são necessariamente formadas apenas de intervalos descendentes, e sim que quaisquer que sejam as direções dos intervalos nos dados *input*, os dados de resposta terão a direção oposta para todos os intervalos.

3. EXPERIMENTOS

3.1 Introdução

Dada a discussão presente no capítulo anterior, a proposta deste trabalho é reconduzir o experimento musical de Nevins e Endress (2007), sobretudo para observar se os resultados serão diferentes se criarmos estímulos com contexto tonal e substituirmos o intervalo musical pela nota musical como unidade a sofrer transformação nos estímulos.

Adicionalmente, contudo, defendemos também a recondução do experimento linguístico, dado que os participantes de nossos experimentos são falantes nativos do português brasileiro, e não do inglês, como é o caso do experimento original.³⁹ Embora, obviamente, a suposição seja de que os resultados do experimento linguístico não difiram de língua para língua, visto que não versa sobre propriedades específicas da língua nativa dos participantes, e sim sobre sua capacidade linguística geral de lidar com uma gramática artificial, há ainda outras duas razões que motivam a recondução também do experimento linguístico: (i) como os experimentos musical e linguístico se utilizam da mesma metodologia e como foram feitas algumas mudanças metodológicas no experimento musical original, a comparação entre os resultados do experimento linguístico e o experimento musical seria mais refinada se o experimento linguístico repetisse as mudanças metodológicas sofridas pelo experimento musical e (ii) também em vista de um maior grau de refinamento na comparação dos resultados dos experimentos linguístico e musical, a recondução do experimento linguístico permite a utilização da mesma população nos experimentos.

Devido à extrema dificuldade evidenciada no percurso deste trabalho de se controlarem todas as variáveis musicais em jogo nos estímulos, mostrou-se necessária a condução não de um, mas quatro experimentos musicais, cada um deles com um tratamento específico dos estímulos. Assim, temos um total de cinco experimentos (quatro musicais e um linguístico), cuja diferença essencial está nos estímulos propriamente ditos. O desenho básico dos experimentos é o mesmo para todos,

³⁹ É uma suposição nossa o fato de o experimento ter sido conduzido com sujeitos falantes nativos de inglês, pelo fato de o experimento, o artigo e os autores serem de língua inglesa. Em momento algum os autores explicitam isso no texto.

conforme mostra a seção 3.2, com apontamentos sobre as diferenças metodológicas em relação ao desenho dos experimentos de Nevins e Endress. A descrição em separado de cada experimento, assim como seus resultados, será apresentada na seção 3.3. A discussão geral desses resultados é feita no próximo capítulo.

3.2 Diretrizes dos experimentos

Para todos os experimentos, foi simulada uma situação em que o participante testemunha, através de uma gravação, um ritual indígena. Neste ponto já se apresenta uma pequena diferença entre o desenho dos experimentos de Nevins e Endress e o nosso: no primeiro, o que foi testemunhado pelos participantes foi um ritual marciano, e não indígena. Pode parecer uma mudança insignificante, mas nós acreditamos que o fato de os participantes terem em mente o fato de estarem lidando com seres humanos, e não com seres extraterrestres desconhecidos, facilita seu desempenho na tarefa proposta pelo experimento, pois não tentarão buscar nos estímulos algum padrão extraordinário.

No ritual indígena testemunhado, primeiramente o participante ouviu 30 instâncias do ritual em que um índio chefe ou pronuncia uma sentença – para o caso do experimento linguístico – ou toca uma melodia em um instrumento musical – para o caso dos experimentos musicais – e um índio subordinado responde-a apropriadamente. Nessa primeira etapa do experimento, que é puramente observacional, os dois índios têm bastante experiência na realização do ritual. As sentenças/melodias pronunciadas pelo índio chefe são formadas de três unidades ($1\ 2\ 3$) e as respostas pronunciadas pelo índio subordinado são as mesmas unidades das sentenças/melodias do índio chefe, mas com a ordem invertida ($3\ 2\ 1$). Exatamente como nos experimentos de Nevins e Endress, esses estímulos, que apresentam a forma $1\ 2\ 3\ 3\ 2\ 1$, são compatíveis com, pelo menos, as quatro generalizações apresentadas em (13):

(13)

- **Comando 1:** Inverta a ordem de todas as unidades;
- **Comando 2:** Permute a primeira unidade com a última unidade;
- **Comando 3:** Permute a última unidade com a antepenúltima unidade;
- **Comando 4:** Permute a cada duas unidades (*i.e.* j com $j+2$).

Para saber qual(is) generalização(ões) os participantes apreenderam dos estímulos do ritual que ouviram, em seguida eles foram informados de que ouviriam um novo ritual – dessa vez com um índio subordinado com menos experiência no ritual – e que eles deveriam avaliar quão bem o índio está respondendo seu chefe, numa escala Likert de 1 a 5. Neste ponto, é possível verificar mais duas diferenças entre os experimentos originais e os nossos. A primeira delas é que, nos experimentos originais, os autores utilizaram uma escala de 1 a 9. Lá, os participantes foram informados que 1 correspondia a ruim, 9 a bom, 5 a neutro e os números entre 1 e 5 e entre 5 e 9 correspondiam a uma gradação de aceitabilidade. Em nossos experimentos, cada número foi rotulado: 1 é ruim, 2 é provavelmente ruim, 3 é neutro (nem bom nem ruim), 4 é provavelmente bom e 5 é bom. Nós acreditamos que é mais fácil para o ouvinte julgar dentro de uma escala com menos níveis, pois há uma dificuldade de precisão quando somos convidados a julgar nossa intuição dentro de uma escala com muitas opções entre os opostos. A segunda diferença é que, em vez de ter sido dito aos participantes que o índio cujas respostas eles deveriam julgar não domina o ritual muito bem, foi dito aos participantes que o índio cujas respostas eles deveriam julgar é *menos experiente* no ritual. Com isso, buscamos amenizar o viés de que o índio sendo julgado *deve* ter respostas ruins.

Nessa segunda etapa do experimento, os participantes ouviram 20 estímulos em que o índio chefe pronuncia (ou toca) uma sequência de quatro unidades (1 2 3 4) e o índio subordinado, dessa vez, responde com 5 diferentes tipos de inversões, 4 das quais são resultantes das generalizações listadas acima: 4 3 2 1 para o comando 1; 4 2 3 1 para o comando 2; 1 4 3 2 para o comando 3; 3 4 1 2 para o comando 4. A quinta "inversão", não presente nos experimentos de Nevins e Endress, é um distrator que inverte a última unidade com a primeira unidade, mas que modifica a natureza das unidades centrais, resultando, assim, em 4 x y 1. A função dos distratores é observar se os participantes estão atentando *apenas* para as unidades presentes nas extremidades, independentemente das unidades centrais, ou se a identidade de *todas* as unidades também está sendo observada, o que se justifica principalmente para o primeiro experimento musical, o que será discutido na subseção 3.3.1.1.

A hipótese investigada em todos os experimentos é a mesma hipótese investigada por Nevins e Endress: considerando-se a posição privilegiada das extremidades na linguagem, os participantes deveriam dar melhores notas aos estímulos

resultantes dos comandos 1 e 2 e notas mais baixas aos estímulos resultantes dos comandos 3 e 4 para o experimento linguístico. Ademais, considerando-se a hipótese de identidade entre música e linguagem de Katz e Pesetsky (2011) apresentada na seção 2.3.3, os experimentos musicais devem apresentar o mesmo resultado que o experimento linguístico. Quanto aos distratores, espera-se que também sejam julgados ruins por mudar a identidade das unidades que não se encontram nas extremidades.

3.3 Metodologia dos experimentos musicais

Conforme dito no início deste capítulo, foi necessária a condução de quatro experimentos musicais, pelo fato de existirem muitas variáveis em jogo, o que tornaria muito difícil a interpretação dos resultados de apenas um experimento, lidando com essas variáveis de apenas uma maneira. A ideia por trás da realização dos quatro experimentos foi retirar gradativamente informações dos estímulos, a fim de se verificar quanto essas informações estavam influenciando os resultados. Os estímulos dos experimentos 1 e 2 (E1 e E2) apresentam a informação de domínio de aplicação e de contexto tonal; os estímulos do experimento 3 (E3) não apresentam a informação de domínio de aplicação, mas apresentam a informação de contexto tonal, e os estímulos do experimento 4 (E4) não apresentam nenhuma dessas informações.

E1 e E2, embora apresentem as mesmas informações, abarcam duas instâncias diferentes da variável *domínio de aplicação*, que serão definidas no parágrafo que se segue e na seção 3.3.1.1. O motivo pelo qual esta variável foi inserida nos estímulos é o fato de que, na música, não é possível dissociar a estrutura prosódica (TSR) da estrutura sintática (PR) na criação de estímulos pertencentes a um idioma musical particular, visto que a menor unidade da estrutura prosódica coincide com a menor unidade da estrutura sintática (a nota musical). Ou seja, ao olharmos os diagramas arbóreos da TSR e da PR para uma mesma peça musical, nas duas estruturas todas as notas musicais são necessariamente núcleos no nível mais baixo dessas estruturas. Isso significa que, quando se tem uma estrutura TSR mínima, tem-se automaticamente uma estrutura PR mínima. A dissociação dessas estruturas, ao contrário, é possível na linguagem, já que a menor unidade prosódica da linguagem (a sílaba) é uma unidade dissociada da estrutura sintática, por ser uma unidade puramente fonológica, sem significado, enquanto a estrutura sintática lida com unidades semânticas ou funcionais. Ao olharmos um diagrama arbóreo da estrutura sintática, vemos que alguns núcleos são complexos

silábicos. Portanto, um núcleo sintático no nível mais baixo não necessariamente corresponde a uma sílaba, o que possibilita que tenhamos uma estrutura prosódica mínima (com, digamos, três sílabas), sem ter ainda um núcleo sintático. Costuma-se referir ao que estamos chamando de "dissociação entre estrutura sintática e prosódica" como "dupla articulação" (termo cunhado pelo linguista André Martinet). A dupla articulação da linguagem refere-se à propriedade que a linguagem tem de se organizar estruturalmente em dois níveis distintos: no nível mais alto (a primeira articulação) a linguagem é vista em termos de combinações de unidades com significado (morfemas, palavras e sintagmas, por exemplo) e no nível mais baixo (a segunda articulação) a linguagem é vista em termos de combinações de unidades desprovidas de sentido (fonemas, sílabas e pés, por exemplo). Assim, o fenômeno em questão que, na linguagem, é puramente prosódico, ou seja, que manipula unidades da segunda articulação, sem colocar questões para a primeira articulação, deve, na música, ser analisado dentro de sua estrutura sintática, o que acrescenta variáveis a serem controladas na música, como o domínio de aplicação.⁴⁰

São poucos os domínios conhecidos na música. Em análise musical, costuma-se falar sobre domínios complexos como motivo, frase, tema e seção, sendo cada um deles formado por um complexo do complexo anterior (ou seja, existe uma noção de hierarquia e de constituência): uma frase pode ter mais de um motivo, um tema pode ter mais de uma frase e uma seção pode ter mais de um tema. A definição de cada uma dessas unidades musicais complexas na tradição musical é bastante difícil, pouco científica e no geral bastante intuitiva. Dessas, a única a que foi dado um tratamento formal através dos mecanismos desenvolvidos por Lerdahl e Jackendoff (1983) em sua teoria é a frase musical ou, nos termos desses autores, a *estrutura prolongacional normativa*. Os estímulos dos dois primeiros experimentos foram montados em cima desse domínio, com diferenças a serem especificadas nas seções 3.3.1.1 e 3.3.2.1. O terceiro experimento utilizou unidades randômicas do idioma tonal, ou seja, estímulos

⁴⁰ Andrew Nevins (comunicação pessoal) justifica a condução de seu experimento musical com estímulos atonais através do *insight* de que a música, assim como a linguagem, também apresenta uma dupla articulação: o nível que combina elementos com significado seria aquele que aparece quando da construção de uma melodia tonal ou construída a partir de princípios combinatórios de um idioma musical particular e o nível que combina elementos sem significado estaria em jogo quando da utilização das unidades musicais (notas musicais) sem o comprometimento com princípios combinatórios de um idioma musical particular. Embora o *insight* seja muito interessante, trabalhamos aqui com a hipótese de que, pelo fato de a TSR e a PR terem as mesmas unidades mínimas, a dupla articulação não existe na música, pois, embora existam duas estruturas que codificam informações semelhantes àsquelas das duas articulações na linguagem, a unidade mínima da primeira articulação é, na música, a mesma que a unidade mínima da segunda articulação, o que impossibilita sua dissociação.

construídos em cima do sistema tonal, porém sem controle sobre sua estrutura (por isso apresenta contexto tonal sem apresentar domínio de aplicação). Por último, o quarto experimento não foi construído em cima de nenhum idioma musical (não tendo, assim, nem contexto tonal e nem domínio de aplicação).

Como o contorno melódico é uma variável impossível de se isolar, dado que toda melodia tem também um contorno melódico, nós procuramos distribuir igualmente entre todos os estímulos quatro contornos possíveis para a melodia do índio chefe: ascendente, descendente, ascendente-descendente e descendente-ascendente. Para todos os experimentos musicais, procurou-se ter aproximadamente oito estímulos de cada tipo de contorno para os estímulos de familiarização – como se totaliza 30 estímulos de familiarização, é impossível que haja exatamente a mesma quantidade para os quatro contornos. Quanto aos estímulos de julgamento, que são subdivididos em cinco condições, cada uma delas com quatro estímulos, os contornos foram distribuídos igualmente entre os estímulos de uma única condição. Assim, cada estímulo de uma condição tem um contorno diferente, dado que são quatro contornos e quatro estímulos por condição.

Um último comentário geral a respeito dos estímulos musicais é que o idioma dentro do qual eles foram criados é o idioma tonal-harmônico (ver 2.3.1), que é o idioma a que nós, da cultura ocidental, mais estamos habituados. O idioma tonal-harmônico apresenta dois padrões escalares, denominados *modos*⁴¹. Os dois modos do idioma tonal-harmônico são os modos *maior* e *menor*. A fim de diminuir o número de variáveis, nossos estímulos são criados apenas em cima do modo maior.

3.3.1 Experimento 1

3.3.1.1 Estímulos

Os estímulos do Experimento 1 (E1) foram criados de forma que a melodia do índio chefe juntamente com a resposta do índio subordinado formasse uma única frase musical, do tipo antecedente-consequente, em que o antecedente é a melodia do chefe e,

⁴¹ Na música, é possível utilizar um mesmo inventário de notas musicais para construir diferentes escalas musicais. Por exemplo, as notas dó, ré, mi, fá, sol, lá e si, nesta ordem, formam a escala de Dó Maior; entretanto, essas mesmas notas, na ordem ré, mi, fá, sol, lá, si e dó, formam a escala de Ré Dórico. "Dórico" e "Maior" são os *modos* dessas escalas, ou seja, são as formas como se definem as relações hierárquicas entre as notas a partir de um mesmo conjunto de notas.

através da inversão, a melodia do subordinado torna-se o conseqüente. Uma frase musical é um trecho de música "gramaticalmente completo" que tem certa independência, por poder ser, por exemplo, apresentada isoladamente a um ouvinte e ser considerada completa, sem a impressão de suspensão ou de incompletude.

Para Lerdahl e Jackendoff, para se ter uma frase musical completa e bem-formada, é preciso ter uma estrutura mínima, denominada por eles *estrutura prolongacional normativa*. A estrutura prolongacional normativa requer um início estrutural estável que se encaminhe para a tensão (formando um antecedente) seguida de uma cadência com preparação, que é um caminho inverso, de tensão para estabilidade (formando um conseqüente). Dentro do idioma tonal, a tônica (1º grau da escala) representa o maior ponto de estabilidade e a dominante (5º grau da escala) é onde geralmente culmina um trecho de tensão. É por isso que a cadência mais utilizada por este idioma consiste na sequência dominante-tônica, denominada *cadência perfeita*. Dessa forma, um antecedente típico pode apresentar a forma 'tônica ... dominante' e um conseqüente típico, a forma 'dominante ... tônica'.

Por este motivo, os estímulos de E1 foram formados de modo que a melodia do chefe partisse de um centro de estabilidade e culminasse numa unidade de tensão, caminho que seria invertido na resposta do índio subordinado. Os estímulos com três unidades tinham todos a forma $I \times V \quad V \times I$, sendo I a tônica (1º grau), V a dominante (5º grau) e x qualquer outra função do mesmo campo harmônico (2º, 3º, 4º ou 6º graus) com a exceção do 7º grau⁴². Embora as combinações possíveis entre as funções no formato $I \times V$ sejam poucas (apenas quatro, para ser exata: I ii V, I iii V, I IV V e I vi V), cada uma dessas funções pode ter pelo menos três notas melódicas, o que aumenta substancialmente as possibilidades de combinações. Abaixo, explicamos de que maneira se comporta essa combinação função-melodia.

O sistema tonal baseia-se em escalas de sete sons, sendo que cada um desses sons tem um comportamento específico dentro da escala. Por isso, a cada um deles dá-se tradicionalmente um nome diferente: tônica, supertônica, medianta, subdominante, dominante, superdominante e sensível. Como existe um ordenamento intrínseco desses sons em uma escala – já que uma escala, inclusive, é definida como um conjunto de

⁴² O sétimo grau não foi utilizado, pois o acorde formado sobre esse grau é o único da escala que apresenta um intervalo dissonante (uma quinta diminuta). Por causa dessa dissonância, o acorde do sétimo grau é um acorde pouco utilizado no idioma tonal-harmônico, sendo muito mais utilizado como função do quinto grau.

notas *ordenadas* de acordo com sua frequência fundamental – pode-se referir a cada um desses sons pela ordem em que aparece na escala: 1º grau, 2º grau, 3º grau, etc. Na escala de Dó Maior, por exemplo, temos as notas dó, ré, mi, fá, sol, lá, si, nesta ordem. Assim, dó é o 1º grau e é chamado de tônica; ré é o 2º grau e é chamado de supertônica; mi é o 3º grau e é chamado de medianta; fá é o 4º grau e é chamado de subdominante; sol é o 5º grau e é chamado de dominante; lá é o 6º grau e é chamado de superdominante e, por fim, si é o 7º grau e é chamado de sensível. Como o idioma tonal-harmônico é um idioma musical, obviamente, harmônico, em que a combinação de notas musicais simultâneas forma uma dimensão da música tanto quanto a combinação sucessiva de notas musicais, sobre cada um desses graus de uma escala pode-se construir uma tríade, ou seja, um acorde de três sons simultâneos. Obedecendo as regras de formação de acordes do idioma tonal-harmônico (basicamente a sobreposição de intervalos de terça) e utilizando apenas as notas pertencentes à escala, o acorde construído sobre o primeiro grau da escala de Dó Maior é composto pelas notas dó, mi e sol (formando a tríade de dó maior); o acorde construído sobre o segundo grau da escala é composto pelas notas ré, fá e lá (formando a tríade de ré menor); o acorde construído sobre o terceiro grau da escala é composto pelas notas mi, sol e si (formando a tríade de mi menor); o acorde construído sobre o quarto grau da escala é composto pelas notas fá, lá e do (formando a tríade de fá maior); o acorde construído sobre o quinto grau da escala é composto pelas notas sol, si e ré (formando a tríade de sol maior); o acorde construído sobre o sexto grau da escala é composto pelas notas lá, dó e mi (formando a tríade de lá menor); o acorde construído sobre o sétimo grau da escala é composto pelas notas si, ré e fá (formando a tríade de si diminuto). Seguindo a tradição, doravante nos referiremos ao *acorde* montado em cima de um grau da escala fazendo uso de algarismos romanos e nos referiremos à nota isoladamente por meio de algarismos arábicos. Assim, na escala de Dó Maior, a nota dó é o grau 1 e o acorde de dó é rotulado I.

Esclarecida a utilização dos algarismos romanos, ao dizer que os estímulos de três unidades têm a forma I x V V x I, entende-se, no presente trabalho, que a primeira unidade da melodia do índio chefe está sobre o acorde formado em cima do primeiro grau de uma escala, podendo a nota da melodia ser qualquer uma das notas que pertencem a esse acorde. Assim, a primeira unidade de um estímulo em Dó Maior poderá ter como nota melódica o dó, o mi ou o sol. Essas possibilidades serão

representadas como 1(I), 3(I) e 5(I), respectivamente – o algarismo arábico representando a nota da melodia e o algarismo romano representando o acorde sobre o qual a nota melódica é sobreposta. Já a terceira unidade de um estímulo em Dó Maior poderá ter como nota melódica sol, si ou ré, o que será representado como 5(V), 7(V) e 2(V), respectivamente. O evento central pode ser qualquer outra função: ii, iii, IV ou vi, cada uma delas podendo apresentar qualquer das três notas melódicas. Segue abaixo, na Tabela 2, um exemplo de um estímulo de três unidades. Todos os outros estímulos se apresentam na tabela 16 (Apêndice A, p. 131).

Melodia índio chefe			Melodia índio subordinado		
<i>1</i>	<i>2</i>	<i>3</i>	<i>3</i>	<i>2</i>	<i>1</i>
1(I)	3(vi)	5(V)	5(V)	3(vi)	1(I)

TABELA 2: REFERENTE A UM ESTÍMULO DE FAMILIARIZAÇÃO DE E1.

Os estímulos com quatro unidades foram criados exatamente da mesma forma, mas com duas unidades centrais – I x y V – com as quatro possíveis inversões – V y x I, V x y I, I V y x e y V I x. Abaixo, na tabela 3, segue um exemplo de estímulo com quatro unidades. Este exemplo apresenta somente a inversão total (1234 4321). A lista completa de estímulos é apresentada na tabela 17 (Apêndice A, p. 132).

Melodia índio chefe				Melodia índio subordinado			
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>3</i>	<i>2</i>	<i>1</i>
5(I)	3(vi)	2(ii)	7(V)	7(V)	2(ii)	3(vi)	5(I)

TABELA 3: REFERENTE A UM ESTÍMULO DE JULGAMENTO DE E1.

A lógica por trás dessa maneira de construir os estímulos é o não comprometimento da estrutura da frase musical ao se modificar o ordenamento das unidades de que é formada. Se a inversão das unidades de uma sequência melódica mantém uma estrutura prolongacional minimamente bem-formada, supõe-se ser mais fácil para o ouvinte o entendimento dessa inversão. Entretanto, pode parecer um viés

indesejado deste experimento o tratamento dos estímulos desta forma; afinal, se o que queremos observar é a maneira como o ouvinte generaliza a inversão de ordem de unidades numa cadeia linear, dar a essa cadeia uma estrutura pode enviesar os ouvintes para os resultados esperados, fazendo-os olhar para a estrutura, e não para o ordenamento, principalmente pelo fato de que a estrutura só se mantém bem-formada nas duas primeiras inversões (4321 e 4231) – que são as inversões que mantêm a relação das extremidades – e não nas duas últimas – que são de fato as inversões que, espera-se, seriam rejeitadas pelo ouvinte. Dessa forma, este experimento, mostrando um resultado que está em consonância com nossa hipótese, seria, em realidade, compatível com pelo menos duas interpretações: (i) a de que os ouvintes estavam atentos ao ordenamento das unidades musicais e que sua preferência pelas inversões resultantes dos comandos 1 e 2 é devida à importância das posições de extremidade ou (ii) os ouvintes estavam atentos à estrutura frasal das melodias e que sua preferência pelas inversões resultantes dos comandos 1 e 2 é devida à manutenção das funções I e V nas extremidades.

Nós admitimos que esse possa de fato ser um viés de nosso experimento, o que impossibilita uma interpretação única dos resultados. É justamente por este motivo que criamos outros experimentos: para averiguar o papel da variável relacionada à estrutura das melodias e para averiguar se a inserção da informação relativa ao domínio de aplicação facilita ou não o julgamento dos estímulos. A ideia dos quatro experimentos musicais é justamente criar uma "escala", na qual o primeiro experimento teria mais informações em princípio facilitadoras e os experimentos subsequentes tirariam passo a passo essas informações, para ver o comportamento do ouvinte sobre esse mínimo e máximo de informações dadas.

3.3.1.2 Participantes

Reproduzindo o tamanho da amostra dos experimentos de Nevins e Endress, este experimento foi rodado com 12 participantes, todos brasileiros. O grau de escolaridade de todos os participantes é superior incompleto ou completo (a amostra é composta de estudantes da Universidade Federal do Paraná, podendo esses estudantes ser de graduação ou de pós-graduação). A idade varia entre 18 e 39 anos.

3.3.1.3 Equipamento

Os estímulos deste experimento foram criados através de combinação e randomização das unidades no programa Excel, versão 2007, e, em seguida, escolhidos atentamente dentre as combinações geradas, de forma a impedir a repetição de uma mesma unidade em um estímulo e de forma a obedecer a distribuição de contornos descrita no fim da seção 3.3. Para preservar a estrutura musical dos estímulos e minimizar a interferência entre diferentes estímulos, a cada um deles foi atribuída randomicamente uma tonalidade. Os estímulos, em seguida, foram transformados em partitura no *software* NoteWorthy Composer 2 e gravados em formato WAV através da reprodução da partitura oferecida pelo próprio programa. Assim, pôde-se controlar intensidade, duração e timbre. Todas as unidades de todos os estímulos foram gravadas com a mesma intensidade.

Para distinguir o índio chefe do índio subordinado nos estímulos, foram utilizados timbres diferentes para cada um deles: o timbre de ocarina oferecido pelo programa para a melodia do índio chefe e o timbre de flauta oferecido pelo programa para a melodia do índio subordinado. Ritmicamente, os estímulos não são marcados; todas as unidades são equidistantes, exceto entre a melodia do chefe e a melodia do subordinado, onde se inseriu uma pausa, para que a distinção entre eles ficasse duplamente marcada: pelo timbre diferente e pelo silêncio que os separa. Assim, os participantes seriam capazes de distinguir um do outro.

O *software* utilizado para a realização do experimento foi o PRAAT, versão 5.2.01. Os participantes realizaram o experimento no Laboratório de Fonética da UFPR, que é um ambiente acusticamente isolado, o que impede que ruídos externos interfiram no seu desempenho. Os participantes ouviram os estímulos através de um fone Audio-Technica, modelo ATH-A500, e as instruções foram apresentadas por escrito, em tela de computador.

3.3.1.4 Procedimento

Primeiramente, pediu-se aos participantes que preenchessem uma ficha que solicitava informações a respeito de sua idade, local de nascimento, escolarização e nível de conhecimento musical (ver ficha no Apêndice B, p. 141). Não se esperava controlar a população através dessa ficha; a população de cuja amostra colhemos os

dados é formada por brasileiros estudantes da UFPR. O intuito desta ficha é colher informações sobre os participantes que possam ser verificadas durante a análise qualitativa dos dados.

Em seguida, os participantes foram verbalmente instruídos a ler os comandos na tela de um computador e clicar com o *mouse* no centro da tela toda vez que necessário. As primeiras instruções por escrito dadas aos participantes são as que vemos na Figura 21. Como este experimento apresenta contexto tonal, antes de cada estímulo foi apresentado ao ouvinte um sinal que estabelece o contexto tonal dentro do qual o estímulo estará inserido (o "sinal de iniciação" das instruções). Esse sinal é uma cadência perfeita (ver segundo parágrafo da seção 3.3.1.1) dentro da tonalidade do estímulo, com que se esperou eliminar possíveis ambiguidades dos estímulos.

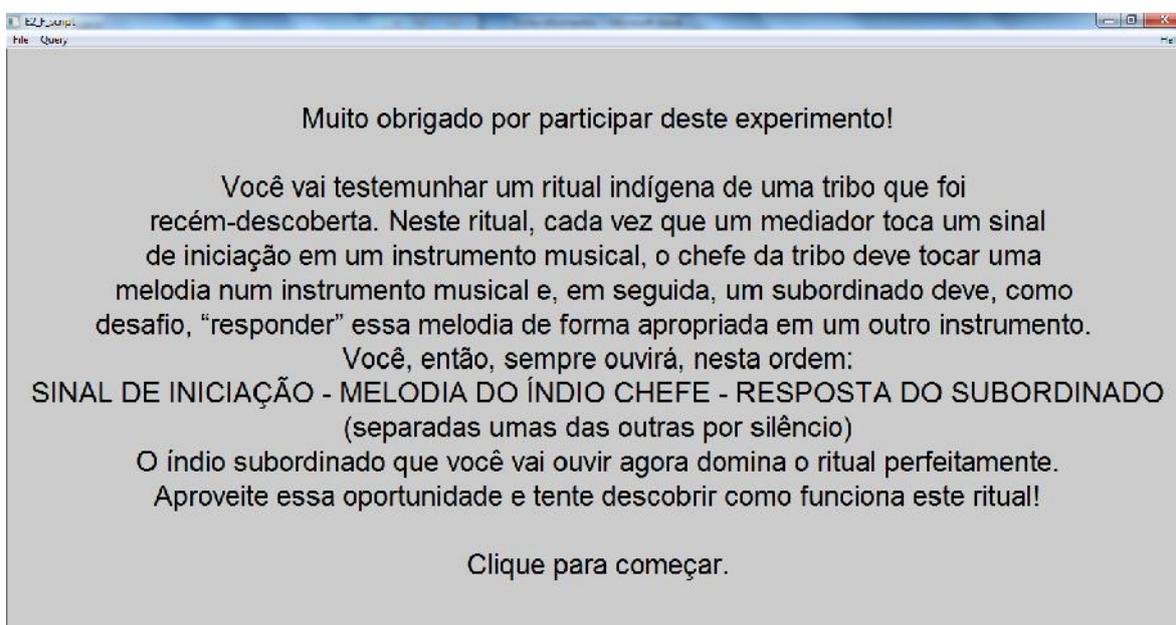


FIGURA 21: TELA DE INSTRUÇÕES DO EXPERIMENTO 1.

Em seguida, os participantes ouviram 15 estímulos de familiarização, que são os estímulos ambíguos, compostos de apenas três unidades, e, após essas 15 escutas, os participantes depararam-se com a instrução presente na Figura 22:

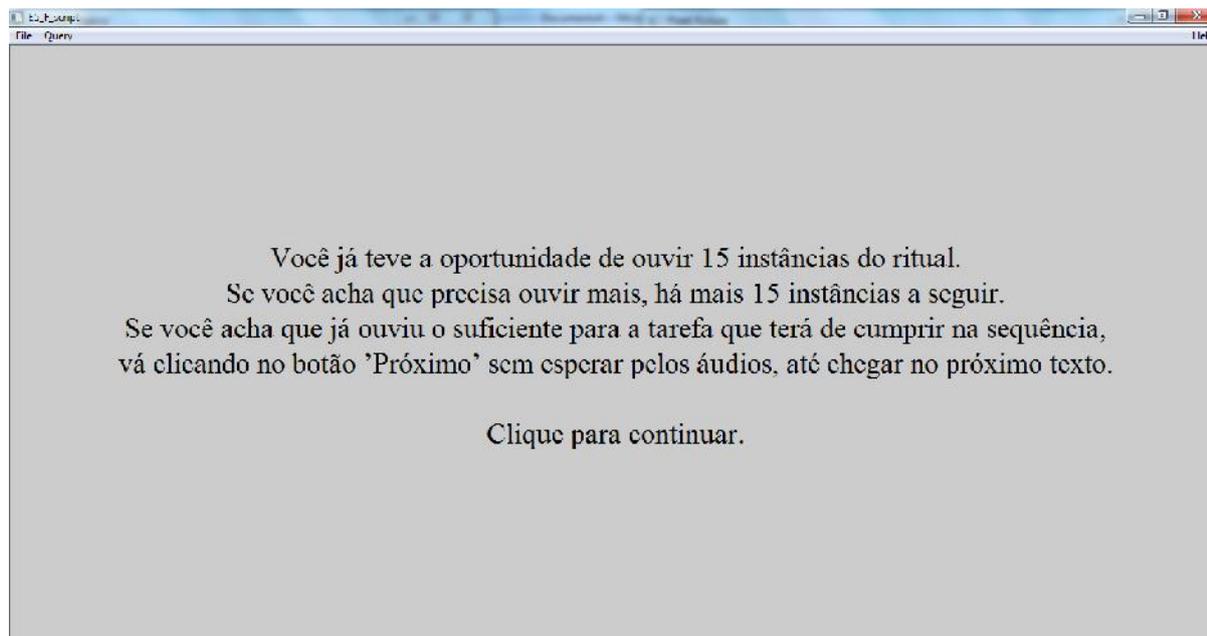


FIGURA 22: INSTRUÇÕES APÓS EXPOSIÇÃO A 15 ESTÍMULOS DE FAMILIARIZAÇÃO.

O motivo pelo qual foi dada ao participante essa escolha é porque trinta estímulos seguidos podem deixá-lo muito cansado para efetuar a tarefa que ainda lhe será dada. Além disso, supôs-se que aqueles que ainda não tivessem encontrado um padrão nos estímulos ouviriam os próximos estímulos e aqueles que já tivessem chegado a alguma generalização sobre os estímulos poderiam passar para frente. Após escutarem ou pularem os 15 estímulos de familiarização seguintes, os ouvintes chegaram às duas seguintes telas:

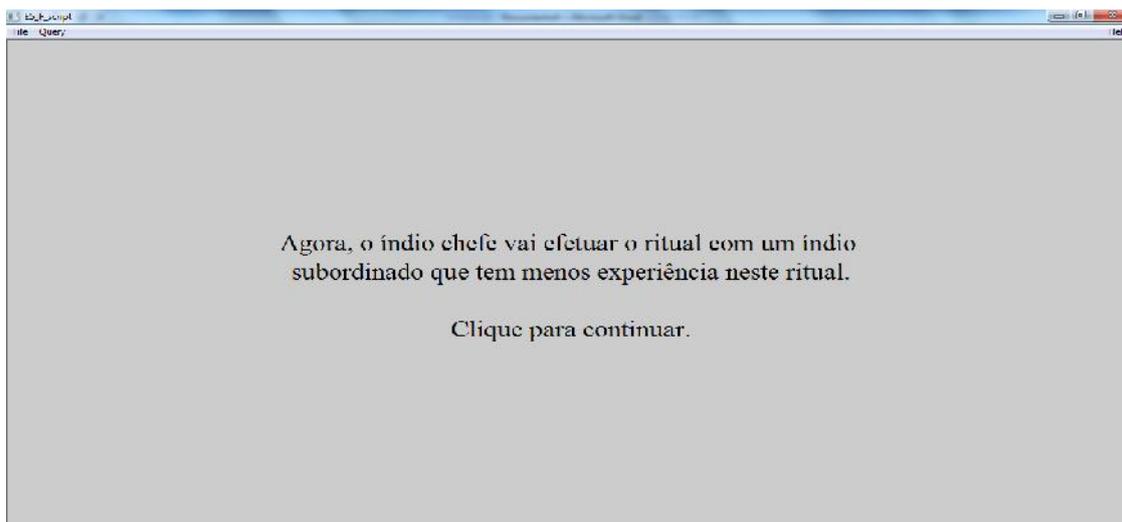


FIGURA 23: TELA DE INSTRUÇÕES 1 APÓS EXPOSIÇÃO A ESTÍMULOS DE FAMILIARIZAÇÃO.

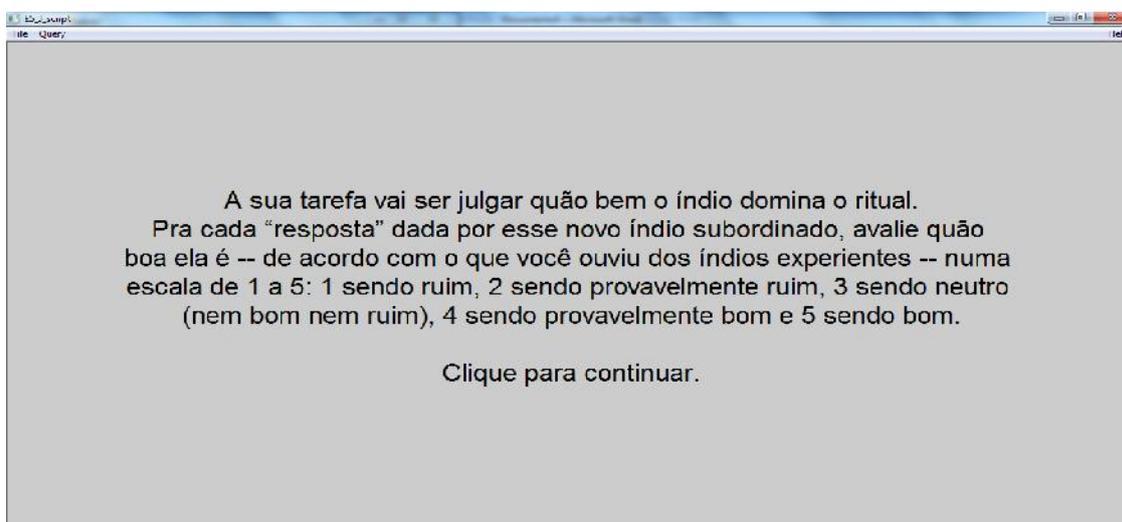


FIGURA 24: TELA DE INSTRUÇÕES 2 APÓS EXPOSIÇÃO A ESTÍMULOS DE FAMILIARIZAÇÃO.

Em seguida, os participantes ouviram 20 estímulos não-ambíguos para julgamento, sendo eles divididos igualmente para cada um dos quatro comandos, com mais quatro distratores (os estímulos foram randomizados pelo próprio programa PRAAT, tendo cada participante ouvido uma ordem diferente). A imagem na tela do computador durante o julgamento dos participantes é:

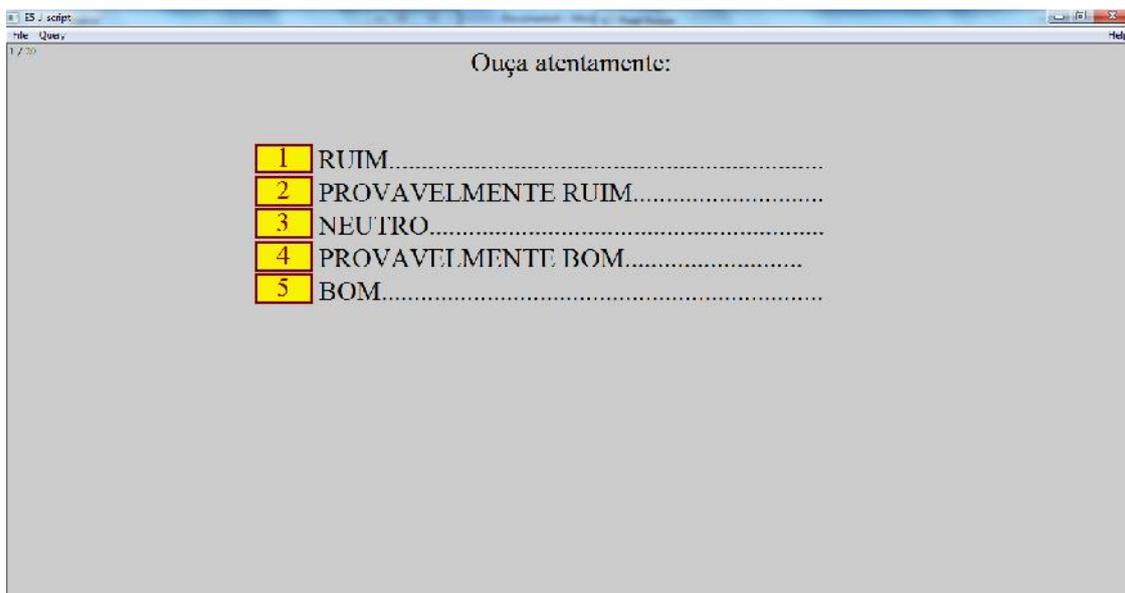


FIGURA 25: TELA PRESENTE DURANTE A ESCUTA DOS ESTÍMULOS DE JULGAMENTO.

Por fim, aparecia à vista do participante uma tela indicando o fim do experimento e agradecendo-o por ter participado.

3.3.1.5 Resultado

A análise estatística deste e dos outros experimentos foi realizada pelas pesquisadoras Fernanda Bühner Pizzato e Andressa Avendaño Forbellone, através do programa de assessoria estatística oferecido pelo Laboratório de Estatística Aplicada da UFPR (LEA). O relatório completo de sua análise encontra-se no Anexo (p. 142). O teste estatístico considerado mais apropriado para o tipo de dado utilizado em nossos experimentos foi o teste não-paramétrico Friedman. Com este teste, avaliou-se a relação entre as condições duas a duas, ou seja, buscou-se a existência de diferenças significativas entre cada duas das cinco inversões dos dados de julgamento do experimento. O resultado a que se chegou foi que, ao nível de 5% de significância, este experimento não apresenta diferenças estatisticamente significativas entre as condições.

3.3.2 Experimento 2

3.3.2.1 Estímulos

O Experimento 2 (E2), assim como E1, também manipula frases musicais, mas, ao contrário de E1, a melodia do índio chefe forma uma única frase e a melodia do índio subordinado forma também uma única frase, diferente da do índio chefe. Assim, o domínio "frase musical", em E1, abarca tanto a "pergunta" quanto a "resposta" de um mesmo estímulo mas, em E2, a "pergunta" forma um domínio e a "resposta" forma outro.

É um pouco impróprio afirmar, a bem da verdade, que os estímulos de E2 formam uma *frase musical*, dado que a definição de Lerdahl e Jackendoff para frase musical exige que uma única frase possua, no mínimo, cinco elementos: um começo estrutural, um movimento de tensão, uma preparação para a cadência e uma cadência (que em geral é formada de dois elementos). Como os estímulos de familiarização possuem três elementos e os estímulos de julgamento possuem quatro elementos, nenhum deles, a rigor, forma uma frase musical. O que opõe os estímulos deste experimento aos estímulos de E1 é a "independência" da melodia do índio chefe em relação à melodia do índio subordinado, pois não há um caminho de tensão na "pergunta" do chefe que exija um caminho de resolução na "resposta" do subordinado; a melodia do índio chefe tem a forma I ... I, assim como a melodia do índio subordinado; ou seja, o caminho de tensão e de resolução já é satisfeito no enunciado de qualquer um dos índios isoladamente. Como tipicamente uma frase musical inicia e finaliza com a tônica, estamos chamando essas melodias curtas de "frase musical" por apresentarem essa mesma característica, embora, conforme acabamos de dizer, sejam demasiadamente curtas para poderem ser consideradas frases musicais segundo a definição de Lerdahl e Jackendoff.

Exatamente como em E1, vamos nos referir às unidades dos estímulos pelo grau a que pertence a nota melódica fazendo uso de algarismos romanos e pelo acorde a que essa nota melódica é sobreposta fazendo uso de algarismos arábicos. A primeira e a última unidade dos estímulos pertencem à função tônica, o que faz com que o acorde seja o mesmo. Embora isso diminua as possibilidades de combinação, pois faz com que uma das extremidades só tenha duas opções disponíveis das três opções de notas melódicas do acorde a que pertence, visto que a outra extremidade já está utilizando uma das opções, isso não interferiu negativamente na criação dos estímulos. Abaixo,

nas tabelas 4 e 5, segue um exemplo de estímulo de familiarização e um exemplo de estímulo de julgamento de E2, respectivamente. As tabelas completas referentes a esses estímulos encontram-se nas tabelas 18 e 19 (Apêndice A, p. 133-4).

Melodia índio chefe (123)			Melodia índio subordinado (321)		
<i>1</i>	<i>2</i>	<i>3</i>	<i>3</i>	<i>2</i>	<i>1</i>
1(I)	6(ii)	5(I)	5(I)	6(ii)	1(I)

TABELA 4: REFERENTE A UM ESTÍMULO DE FAMILIARIZAÇÃO DE E2.

Melodia índio chefe (1234)				Melodia índio subordinado (4321)			
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>3</i>	<i>2</i>	<i>1</i>
1(I)	7(V)	6(ii)	5(I)	5(I)	6(ii)	7(V)	1(I)

TABELA 5: REFERENTE A UM ESTÍMULO DE JULGAMENTO DE E2.

3.3.2.2 Participantes

O número e perfil dos participantes é igual ao do Experimento 1. Ver seção 3.3.1.2, p. 88.

3.3.2.3 Equipamento

O equipamento utilizado em nada se diferencia do equipamento utilizado para a realização do Experimento 1. Ver, portanto, seção 3.3.1.3, página 87.

3.3.2.4 Procedimento

Assim como o perfil dos participantes e o equipamento utilizado, o procedimento deste experimento reproduz o procedimento do Experimento 1. Ver seção 3.3.1.4, página 87.

3.3.2.5 Resultado

O teste estatístico utilizado neste experimento foi o teste não paramétrico Friedman. Com este teste, avaliou-se a relação entre as condições duas a duas, ou seja, buscou-se a existência de diferenças significativas entre cada duas das cinco inversões dos estímulos de julgamento do experimento. O resultado a que se chegou foi que, ao nível de 5% de significância, este experimento apresenta diferenças estatisticamente significativas entre as condições ($p = 0.04$). A Tabela 6 mostra quais condições são significativamente diferentes, através das letras apresentadas na segunda coluna: as condições que são significativamente diferentes não possuem nenhuma letra em comum; quando duas condições possuem letras em comum, não são significativamente diferentes. A terceira coluna mostra a soma dos *ranks* de cada condição. O resultado nos mostra que as duas inversões consideradas "naturais" são significativamente mais bem aceitas do que os distratores, não havendo distinção significativa entre as quatro inversões e tampouco entre as inversões não-naturais e os distratores.

Resultado Experimento 2		
Condições	Grupo	Soma dos <i>ranks</i>
C1 (4321)	A	43
C2 (4231)	A	43,5
C3 (1432)	AB	34,5
C4 (3412)	AB	34,5
DIS (4xy1)	B	24,5

TABELA 6: RESULTADO DO EXPERIMENTO 2.

3.3.3 Experimento 3

3.3.3.1 Estímulos

Os estímulos do Experimento 3 (E3) não apresentam nenhum tipo de domínio musicalmente reconhecido. Assim, não há nenhuma restrição sobre quais funções se encontram nas extremidades e quais funções se encontram fora delas. Não havendo a necessidade de forçar o ouvinte a interpretar cada uma das unidades dos estímulos como

uma ou outra função, estes estímulos não requerem um acompanhamento harmônico, mas apenas uma melodia não harmonizada.⁴³ A maneira como os estímulos foram criados foi através da randomização dos sete graus da escala em sequências de três e quatro notas. Embora esses estímulos pareçam, à primeira vista, muito desorganizados, talvez até a ponto de se tornarem atonais, nós garantimos que isso não acontecesse através de dois procedimentos: (i) as notas utilizadas nos estímulos são um subconjunto de sete das doze notas que compõem o inventário de sons da música tonal ocidental, escolhidas de tal modo que este subconjunto forme uma escala maior, o que faz com que os estímulos só apresentem notas de uma única escala e (ii) o sinal previamente apresentado aos ouvintes estabelece o contexto tonal dentro do qual o estímulo deverá ser interpretado, o que elimina possíveis ambiguidades dos estímulos (como ambiguidades modais).

Dessa forma, os estímulos de E3 não têm uma estrutura previamente estabelecida, contrariamente aos estímulos de E1 e E2. Por um lado, a ausência de estrutura pode ser negativa pela impossibilidade, na música, de dissociar a estrutura prosódica da estrutura sintática (ver 3.3.1). Não havendo nada controlando a estrutura, com exceção da utilização de eventos pertencentes a uma única escala do sistema tonal, as melodias podem soar demasiadamente fragmentadas aos ouvintes. Por outro lado, a ausência de estrutura é positiva por não gerar um viés estrutural para os ouvintes, como acontece com E1 e E2, o que diminui as chances de os ouvintes se ancorarem na boa-formação estrutural para julgar os estímulos, desviando positivamente sua atenção para o ordenamento das funções. É positivo, ainda, por eliminar a necessidade de utilizarmos nos estímulos a dimensão harmônica da música, o que enxuga os estímulos de informações em excesso.

Como nestes estímulos não há o estabelecimento prévio de função tonal a cada um dos eventos musicais, a representação das unidades será apenas através dos algarismos arábicos correspondentes ao grau da escala sendo utilizado. Abaixo, seguem um exemplo de estímulo de familiarização (Tabela 7) e um exemplo de estímulo de

⁴³ Como a harmonia é uma propriedade fundamental do idioma tonal-harmônico, mesmo uma melodia não harmonizada carrega implicitamente uma harmonização. Para melodias simples, que utilizam notas pertencentes a apenas um campo harmônico (a apenas uma escala), a harmonia implícita é em geral baseada apenas nos graus tônica (I), subdominante (IV) e dominante (V), que são as funções mais importantes do idioma tonal-harmônico. Outras harmonizações são, obviamente, possíveis, dado que uma mesma melodia pode ser harmonizada de diversas formas.

juízo (Tabela 8) de E3. As tabelas 20 e 21 (no Apêndice A, p. 135-6) apresentam todos os estímulos utilizados neste experimento.

Melodia índio chefe (123)			Melodia índio subordinado (321)		
<i>1</i>	<i>2</i>	<i>3</i>	<i>3</i>	<i>2</i>	<i>1</i>
<i>2</i>	<i>7</i>	<i>1</i>	<i>1</i>	<i>7</i>	<i>2</i>

TABELA 7: REFERENTE A UM ESTÍMULO DE FAMILIARIZAÇÃO DE E3.

Melodia índio chefe (1234)				Melodia índio subordinado (4321)			
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>3</i>	<i>2</i>	<i>1</i>
<i>6</i>	<i>5</i>	<i>4</i>	<i>3</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>

TABELA 8: REFERENTE A UM ESTÍMULO DE JUÍZO DE E3.

3.3.3.2 Participantes

O número e perfil dos participantes é igual ao do experimento 1. Ver seção 3.3.1.2, p. 88.

3.3.2.3 Equipamento

O equipamento utilizado em nada se diferencia do equipamento utilizado para a realização do experimento 1. Ver, portanto, seção 3.3.1.3, p. 89.

3.3.3.4 Procedimento

Assim como o perfil dos participantes e o equipamento utilizado, o procedimento deste experimento reproduz o procedimento do experimento 1. Ver seção 3.3.1.4.

3.3.3.5 Resultado

O teste estatístico utilizado neste experimento foi o teste não paramétrico Friedman. Com este teste, avaliou-se a relação entre as condições duas a duas, ou seja, buscou-se a existência de diferenças significativas entre cada duas das cinco inversões dos estímulos de julgamento do experimento. O resultado a que se chegou foi que, ao nível de 5% de significância, este experimento não apresenta diferenças estatisticamente significativas entre as condições.

3.3.4 Experimento 4

3.3.4.1 Estímulos

Os estímulos do Experimento 4 (E4) não apresentam nem domínio musical e nem contexto tonal. Assim, além de os estímulos terem sido construídos de maneira melódica, sem acompanhamento harmônico, como os estímulos de E3, eles foram criados com a utilização de *quaisquer* das doze notas do inventário de notas do sistema musical ocidental e foram apresentados aos ouvintes sem o sinal inicial apresentado em E1, E2 e E3 que estabelece o contexto tonal do estímulo. A ideia por trás da realização deste último experimento musical é replicar o experimento de Nevins e Endress mantendo todos os seus aspectos, com exceção da unidade musical sendo manipulada; no experimento de Nevins e Endress, a inversão era feita sobre os intervalos e, em E4, a inversão foi feita sobre os valores absolutos das notas musicais. Não esperávamos um resultado estatisticamente relevante para este experimento, visto que a grande maioria dos ouvintes⁴⁴ não codifica os valores absolutos das notas musicais. Este experimento teve o objetivo único de cobrir a gradação de informações retiradas de um experimento para outro.

Como neste experimento não é mais possível nos referirmos às unidades dos estímulos pelo grau da escala a que pertencem, já que os estímulos não mais se baseiam em escalas, os estímulos foram criados através da randomização das próprias notas musicais em sequências de três e quatro notas musicais. Ademais, como a única

⁴⁴ Existe um pequeno número de pessoas capazes de codificar o valor absoluto das notas musicais. Essas pessoas são classificadas como possuidoras de um "ouvido absoluto". Aqueles que têm ouvido absoluto são capazes de reter a informação da frequência dos sons em sua memória de longo prazo, o que os possibilita reconhecer e classificar uma nota musical quando produzida sem nenhuma referência prévia ou contexto musical.

informação musical que restou para os ouvintes ancorarem seu julgamento dos estímulos foi, além do próprio valor absoluto das notas, os intervalos formados entre essas notas, o controle sobre os estímulos foi feito através do controle dos intervalos formados, para que não houvesse estímulos com intervalos repetidos, embora com notas diferentes, pois, para um ouvinte que não possui ouvido absoluto, esses estímulos seriam interpretados exatamente da mesma forma. Por isso, às tabelas de estímulos foi acrescida a coluna referente aos intervalos formados pelas notas adjacentes da melodia executada pelo índio chefe. Seguem, abaixo, as tabelas com exemplos de estímulos de três (Tabela 9) e de quatro (Tabela 10) unidades. As tabelas completas são as tabelas 22 e 23 (Apêndice A, p. 137-8).

Melodia índio chefe (123)			Melodia índio subordinado (321)			Intervalos índio chefe	
<i>1</i>	<i>2</i>	<i>3</i>	<i>3</i>	<i>2</i>	<i>1</i>	Intervalo 1	Intervalos 2
fá	ré#	sib	sib	ré#	fá	2M D	4J D

TABELA 9: REFERENTE A UM ESTÍMULO DE FAMILIARIZAÇÃO DE E4.

Melodia índio chefe (1234)				Melodia índio subordinado (4321)				Intervalos índio chefe		
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>3</i>	<i>2</i>	<i>1</i>	Intervalo 1	Intervalo 2	Intervalo 3
ré#	si	fá	mi	mi	fá	si	ré#	3M D	4aum D	2m D

TABELA 10: REFERENTE A UM ESTÍMULO DE JULGAMENTO DE E4.

3.3.3.2 Participantes

O número e perfil dos participantes é igual ao do experimento 1. Ver seção 3.3.1.2, p. 88.

3.3.3.3 Equipamento

O equipamento utilizado em nada se diferencia do equipamento utilizado para a realização do experimento 1. Ver, portanto, seção 3.3.1.3, página 87.

3.3.3.4 Procedimento

Como os estímulos do experimento 4 não são tonais e, portanto, não necessitam de um contexto tonal precedente à apresentação dos estímulos, o procedimento deste experimento se diferencia ligeiramente do procedimento dos experimentos anteriores, pois não há o sinal de iniciação (que é a contextualização tonal) nos estímulos e nas instruções deste experimento. Embora a mudança seja pequena entre os experimentos 1, 2 e 3 e o experimento 4 (apenas a primeira tela de instruções é diferente), nesta seção será descrito todo o procedimento do experimento 4.

Primeiramente, pediu-se aos participantes que preenchessem uma ficha que solicitava informações a respeito de sua idade, local de nascimento, escolarização e nível de conhecimento musical (ver ficha no APÊNDICE B, página 139). Não se esperava controlar a população através dessa ficha; a população de cuja amostra colhemos os dados é formada por brasileiros estudantes da UFPR. O intuito desta ficha é colher informações sobre os participantes que possam ser verificadas durante a análise qualitativa dos dados.

Em seguida, os participantes foram instruídos a ler os comandos na tela de um computador e clicar com o *mouse* no centro da tela toda vez que necessário. As instruções dadas aos participantes são as que vemos na figura 26.

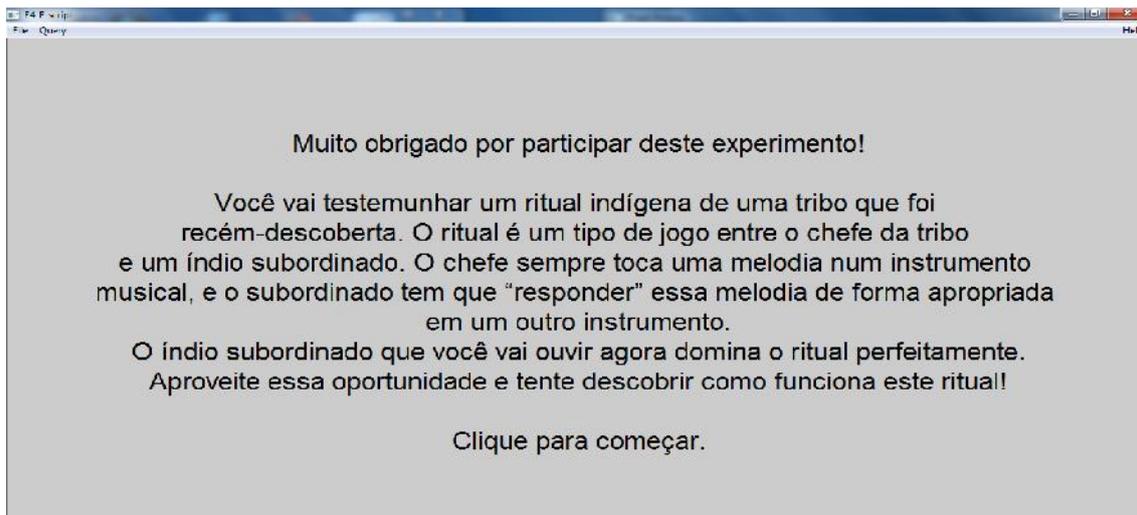


FIGURA 26: TELA DE INSTRUÇÕES DO EXPERIMENTO 4.

Em seguida, os participantes ouviram 15 estímulos de familiarização, que são os estímulos ambíguos, compostos de apenas três unidades, e, após essas 15 escutas, os participantes depararam-se com a seguinte instrução:

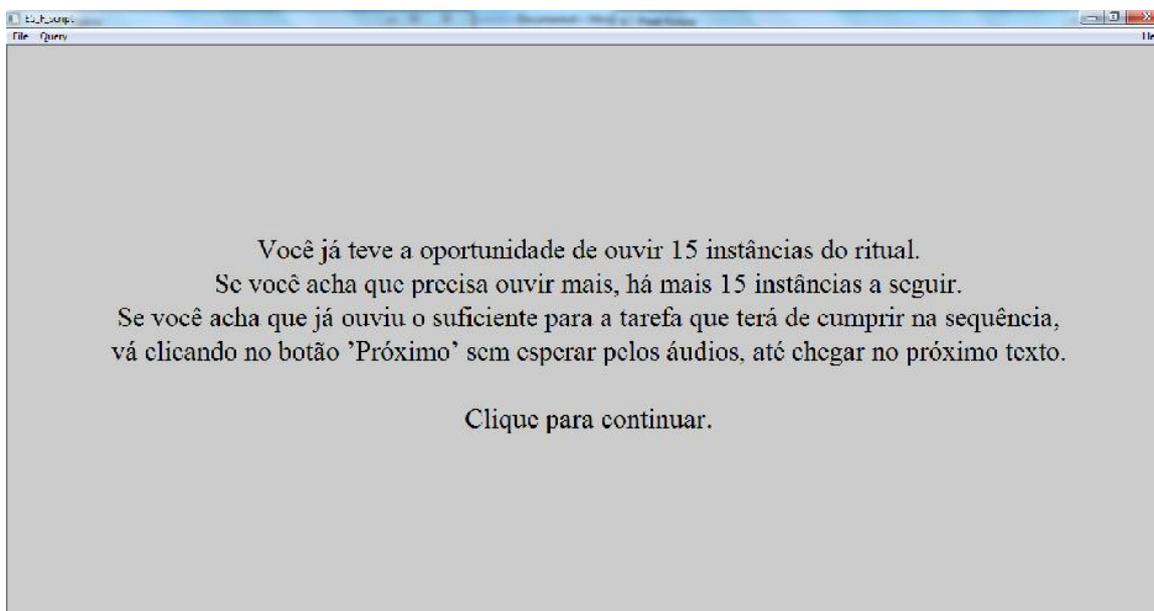


FIGURA 27: INSTRUÇÕES APÓS EXPOSIÇÃO A 15 ESTÍMULOS DE FAMILIARIZAÇÃO.

O motivo pelo qual foi dada ao participante essa escolha é porque trinta estímulos seguidos podem deixá-lo muito cansado para efetuar a tarefa que ainda lhe será dada. Além disso, supôs-se que aqueles que ainda não tivessem encontrado um

padrão nos estímulos ouviriam os próximos estímulos e aqueles que já tivessem chegado a alguma generalização sobre os estímulos poderiam passar para frente. Após escutarem ou pularem os 15 estímulos de familiarização seguintes, os ouvintes chegaram às duas seguintes telas:

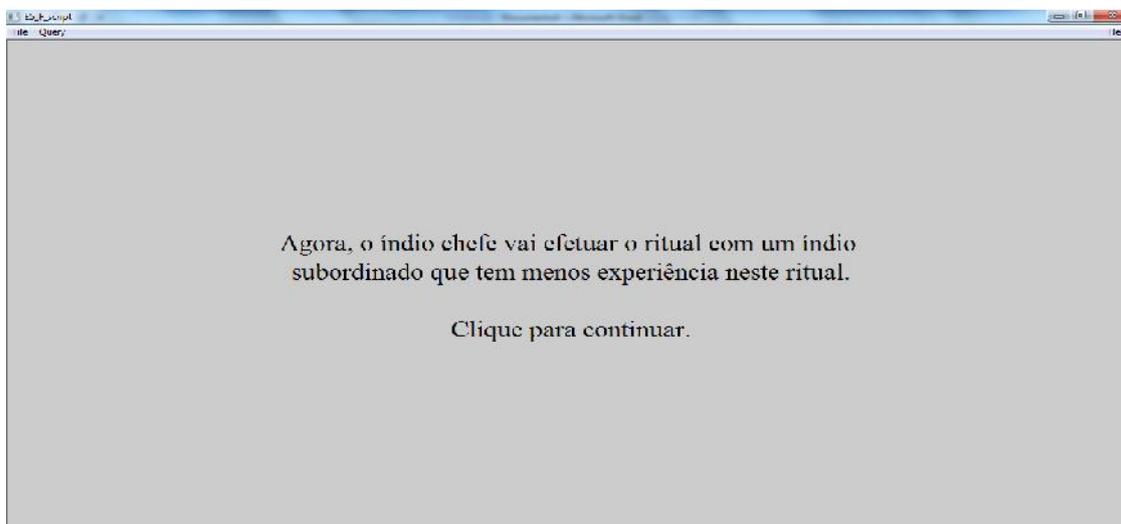


FIGURA 28: TELA DE INSTRUÇÕES 1 APÓS EXPOSIÇÃO A ESTÍMULOS DE FAMILIARIZAÇÃO NO EXPERIMENTO 4.

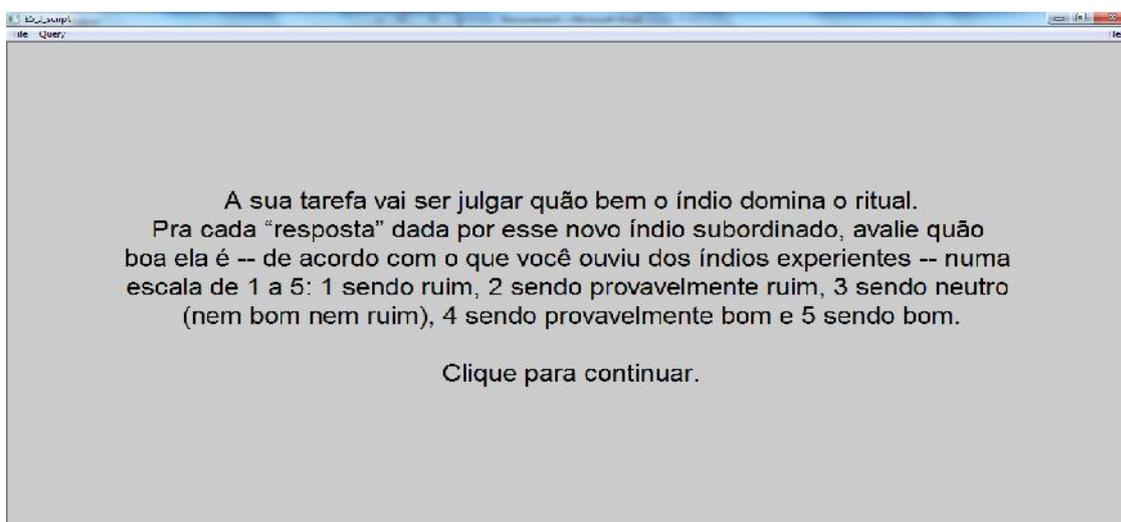


FIGURA 29: TELA DE INSTRUÇÕES 2 APÓS EXPOSIÇÃO A ESTÍMULOS DE FAMILIARIZAÇÃO NO EXPERIMENTO 4.

Em seguida, os participantes ouviram 20 estímulos não-ambíguos para julgamento, divididos igualmente para cada um dos quatro comandos, com mais quatro distratores (os estímulos foram randomizados pelo próprio programa Praat, tendo cada

participante ouvido uma ordem diferente). A imagem na tela do computador durante o julgamento dos participantes é:

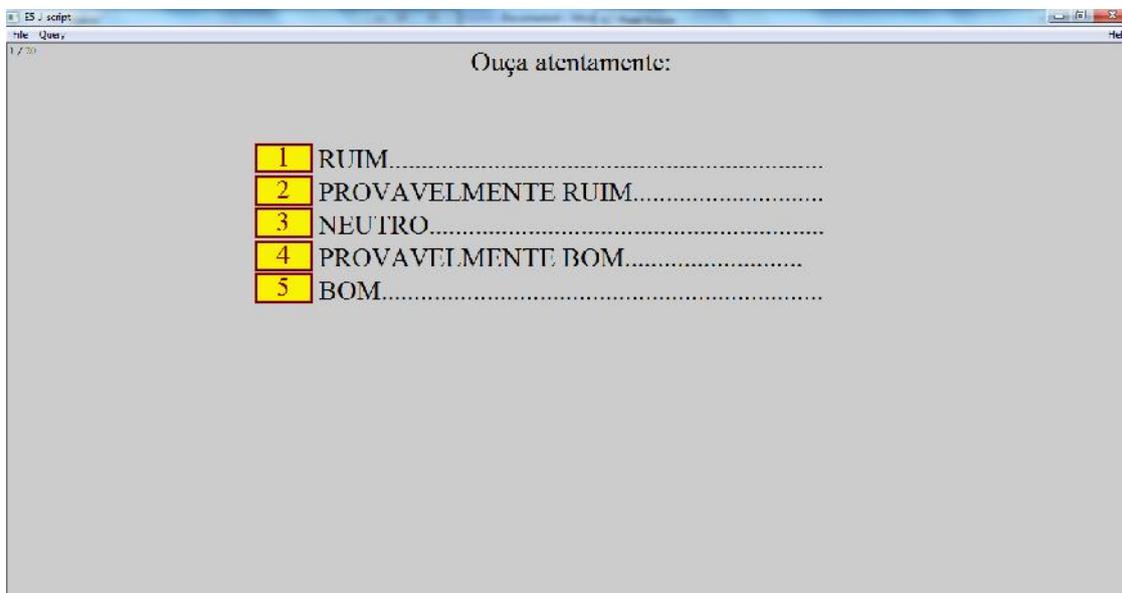


FIGURA 30: TELA PRESENTE DURANTE A ESCUTA DOS ESTÍMULOS DE JULGAMENTO.

Por fim, aparecia à vista do participante uma tela indicando o fim do experimento e agradecendo-lhe por ter participado.

3.3.3.5 Resultado

O teste estatístico utilizado neste experimento foi o teste não paramétrico Friedman. Com este teste, avaliou-se a relação entre as condições duas a duas, ou seja, buscou-se a existência de diferenças significativas entre cada duas das cinco inversões dos estímulos de julgamento do experimento. O resultado a que se chegou foi que, ao nível de 5% de significância, este experimento apresenta diferenças estatisticamente significativas entre as condições ($p = 0.04$). A Tabela 11 mostra quais condições são significativamente diferentes, através das letras apresentadas na segunda coluna: as condições que são significativamente diferentes não possuem nenhuma letra em comum; quando duas condições possuem letras em comum, não são significativamente diferentes. A terceira coluna mostra a soma dos *ranks* de cada condição. O resultado nos mostra que a primeira inversão (inversão total) foi significativamente mais bem aceita

pelos participantes do que as outras três inversões. Os distratores, entretanto, não apresentam diferença significativa em relação a nenhuma das quatro inversões.

Resultado Experimento 4		
Condição	Grupo	Soma dos <i>ranks</i>
C1 (4321)	A	48
C2 (4231)	B	34,5
C3 (1432)	B	30,5
C4 (3412)	B	30
DIS (4xy1)	AB	37

TABELA 11: RESULTADO DO EXPERIMENTO 4.

3.4 Metodologia do Experimento Linguístico

3.4.1 Experimento 5

3.4.1.1 Estímulos

Como os informantes que participariam do experimento linguístico (E5) eram falantes nativos do português brasileiro, os estímulos linguísticos foram criados de modo a obedecer a fonotática da língua, mas sem coincidir com nenhuma palavra já existente. Dito de outro modo, os estímulos são logatomas do português brasileiro com três e com quatro sílabas, para familiarização e julgamento, respectivamente.

Todos os logatomas tanto do índio chefe como do índio subordinado foram acentuados na penúltima sílaba, pelo fato de o português ser formado majoritariamente por palavras paroxítonas.⁴⁵ Todas as unidades prosódicas mínimas dos estímulos (i.e. todas as sílabas) foram compostas de modo a obedecer a estrutura silábica do tipo tido como o mais comum, que é a sílaba consoante-vogal (CV). O motivo pelo qual se utilizaram sílabas apenas com essa estrutura é que esse tipo silábico é considerado o tipo

⁴⁵ A acentuação nas línguas secretas não é matéria trivial e mereceria uma discussão à parte. As palavras transformadas pela operação das línguas secretas do tipo modificador de precedência não necessariamente copiam a acentuação da palavra que serve de *input* para a operação. Tampouco a sílaba acentuada necessariamente carrega consigo a informação de acentuação no processo de transformação. A acentuação às vezes pode ser completamente diferente da acentuação da palavra original, e isso mereceria um estudo à parte. Entretanto, mantemos o padrão acentual das palavras não invertidas nas palavras invertidas pelo fato de Nevins e Endress tê-lo feito em seu experimento (todas as palavras, inclusive as inversões, eram acentuadas na primeira sílaba).

silábico *default* pela teoria fonológica devido a diversos fatores, incluindo o fato de que a sílaba CV é o único tipo silábico comum a todas as línguas, sendo o *único* tipo silábico em línguas como o havaiano. Além disso, utilizar mais do que uma estrutura silábica pode influenciar no julgamento do participante – esta seria uma variável a mais sem controle.

Como os estímulos foram compostos de logatomas com suas respectivas inversões silábicas (que, por sua vez, também formam logatomas), alguns fonemas da língua não foram usados, devido a seu comportamento assimétrico quanto à distribuição fonotática, o que pode tornar um dado de inversão inaceitável por fatores diferentes daqueles que estamos investigando. Assim, as palatais nasal e lateral (os sons representados na escrita pelas letras 'nh' e 'lh', respectivamente), bem como o *tap* (o som da letra 'r' na palavra 'cara'), foram fonemas descartados, por jamais ocorrerem em início de palavra no português brasileiro. Todos os segmentos nasais foram descartados: as consoantes nasais ([m] e [n], além da nasal já mencionada acima) foram descartadas pelo fato de exercerem influência sobre a vogal [a], em especial nos contextos em que a vogal [a] é tônica e a nasal é o *onset* da sílaba que a segue; e as vogais nasais (o som [ã] em 'banda' e o som [] em 'lenda', por exemplo) foram descartadas devido à controvérsia existente em sua representação, ou seja, se uma sílaba formada por uma consoante e uma vogal nasal tem coda (análise bifonêmica) ou não (análise monofonêmica) (ver MENDES (2008, p. 19-22)). As vogais médias abertas [ɛ, ɔ] (representadas na escrita por 'é' e 'ó'), embora não descartadas, foram evitadas por terem sua distribuição limitada à posição tônica nos dialetos do sul do Brasil; elas só foram utilizadas quando a sílaba em que se encontravam estivesse na posição tônica tanto na enunciação do índio chefe quanto na inversão enunciada pelo índio subordinado.

Após descartar todos esses sons da língua, permanecemos com o seguinte inventário de sons: consoantes [p, b, t, d, k, g, f, v, s, z, ʃ, ʒ, l] e vogais [i, e, (ɛ), a, (ɔ), o, u]. Todas as combinações do tipo CV possíveis dão um total de 91 sílabas. Para gerar os estímulos a partir das 91 sílabas a que chegamos, foram feitas diversas combinações triplas e quádruplas dessas sílabas, das quais 50 (30 triplas e 20 quádruplas) foram selecionadas manualmente, para que pudéssemos evitar a repetição da mesma sílaba em um único logatoma e também evitar o surgimento de sílabas átonas com vogais abertas. Além disso, a seleção manual buscou evitar problemas relacionados à neutralização dos

pares [e, i] e [o, u] em posição pós-tônica. Ou seja, quando criados dois logatomas que só se diferenciam por um desses pares de vogais na posição pós-tônica, como [te.ˈka.bo] e [te.ˈka.bu], os participantes julgariam tais estímulos como o mesmo logatoma, devido à neutralização própria dessa posição.

Nas tabelas 12 e 13 abaixo, estão um exemplo de estímulo de familiarização e um exemplo de estímulo de julgamento, respectivamente. As tabelas 24 e 25 (Apêndice A, p. 140-1) apresentam todos os estímulos deste experimento.

Sentença índio chefe (123)			Sentença índio subordinado (321)		
<i>1</i>	<i>2</i>	<i>3</i>	<i>3</i>	<i>2</i>	<i>1</i>
[ze]	[ˈda]	[bo]	[bo]	[ˈda]	[ze]

TABELA 12: REFERENTE A UM ESTÍMULO DE FAMILIARIZAÇÃO DE E5.

Sentença índio chefe (1234)				Sentença índio subordinado (4321)			
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>3</i>	<i>2</i>	<i>1</i>
[fe]	[ga]	[ˈsi]	[de]	[de]	[si]	[ˈga]	[fe]

TABELA 13: REFERENTE A UM ESTÍMULO DE JULGAMENTO DE E5.

3.4.1.2 Participantes

O número e perfil dos participantes é igual ao do experimento 1. Ver seção 3.3.1.2, p. 88.

3.4.1.3 Equipamentos

As unidades dos estímulos (ou seja, as 91 sílabas CV formadas a partir do inventário de sons listado na p. 106) foram combinadas e randomizadas pelo programa Excel, versão 2007 e, em seguida, foram selecionadas manualmente, para controle dos estímulos. Contrariamente aos estímulos linguísticos de Nevins e Endress, que foram criados em um programa de fala sintética, os nossos estímulos foram gravados por dois sujeitos do sexo masculino, idade 24 e 25 anos. O primeiro gravou os logatomas

pronunciados pelo índio chefe e o segundo gravou as inversões pronunciadas pelo índio subordinado. As gravações foram realizadas no Laboratório de Fonética da UFPR, um ambiente acusticamente isolado, e foram monitoradas, para que os participantes pudessem repetir a gravação dos logatomas sempre que necessário. O microfone utilizado nas gravações foi o Sennheiser E838. O *software* utilizado para a gravação foi o Audacity 2.0.3, que gera arquivos em formato WAV, com taxa de amostragem do sinal a 44100Hz. Em seguida, as gravações selecionadas foram manipuladas no programa PRAAT, versão 5.3.52, para que apresentassem exatamente o mesmo contorno entoacional descendente. Entre a "pergunta" do índio chefe e a "resposta" do índio subordinado, foram inseridos 1,5 segundos de silêncio.

Da mesma forma como nos experimentos E1 a E4, o *software* utilizado para a realização deste experimento foi o PRAAT 5.2.01. Os participantes realizaram o experimento no Laboratório de Fonética da UFPR, para impedir que ruídos externos interferissem no seu desempenho. Os participantes ouviram os estímulos através de um fone Audio-Technica, modelo ATH-A500, e as instruções foram apresentadas por escrito, em tela de computador.

3.4.1.4 Procedimento

Como os estímulos do experimento 5 são estímulos linguísticos, o procedimento deste experimento se diferencia um pouco do procedimento dos experimentos anteriores, pois, além de não haver nele o sinal de iniciação (da mesma forma como o experimento 4) nem nos estímulos e nem nas instruções, as instruções fazem menção a *sentenças* pronunciadas pelos índios, e não a *melodias* executadas por eles. Embora a mudança seja pequena entre os experimentos 1, 2, 3 e 4 e o experimento 5, nesta seção será descrito todo o procedimento do experimento 5.

Primeiramente, pediu-se aos participantes que preenchessem uma ficha que solicitava informações a respeito de sua idade, local de nascimento, escolarização e nível de conhecimento musical (ver ficha no Apêndice B, p. 141). Não se esperava controlar a população através dessa ficha; a população de cuja amostra colhemos os dados é formada por brasileiros estudantes da UFPR. O intuito desta ficha é colher informações sobre os participantes que possam ser verificadas durante a análise qualitativa dos dados.

Em seguida, os participantes foram instruídos a ler os comandos na tela de um computador e clicar com o *mouse* no centro da tela toda vez que necessário. As instruções dadas aos participantes são as que vemos na figura 31.

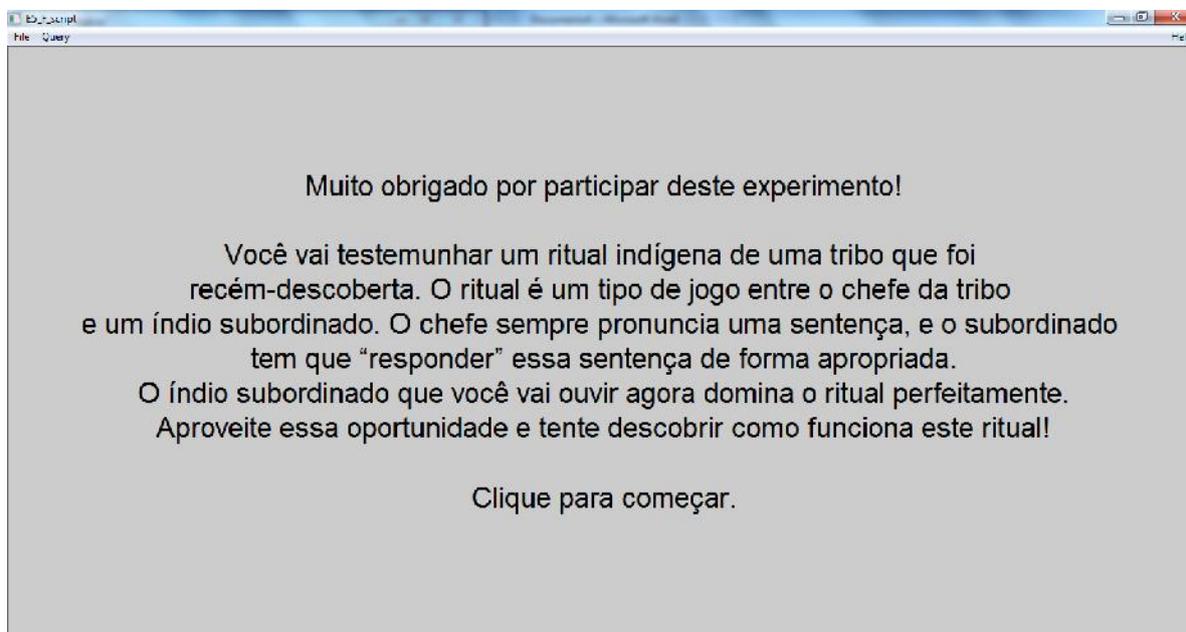


FIGURA 31: TELA DE INSTRUÇÕES DO EXPERIMENTO 5.

Em seguida, os participantes ouviram 15 estímulos de familiarização, que são os estímulos ambíguos, compostos de apenas três unidades, e, após essas 15 escutas, os participantes depararam-se com a seguinte instrução:

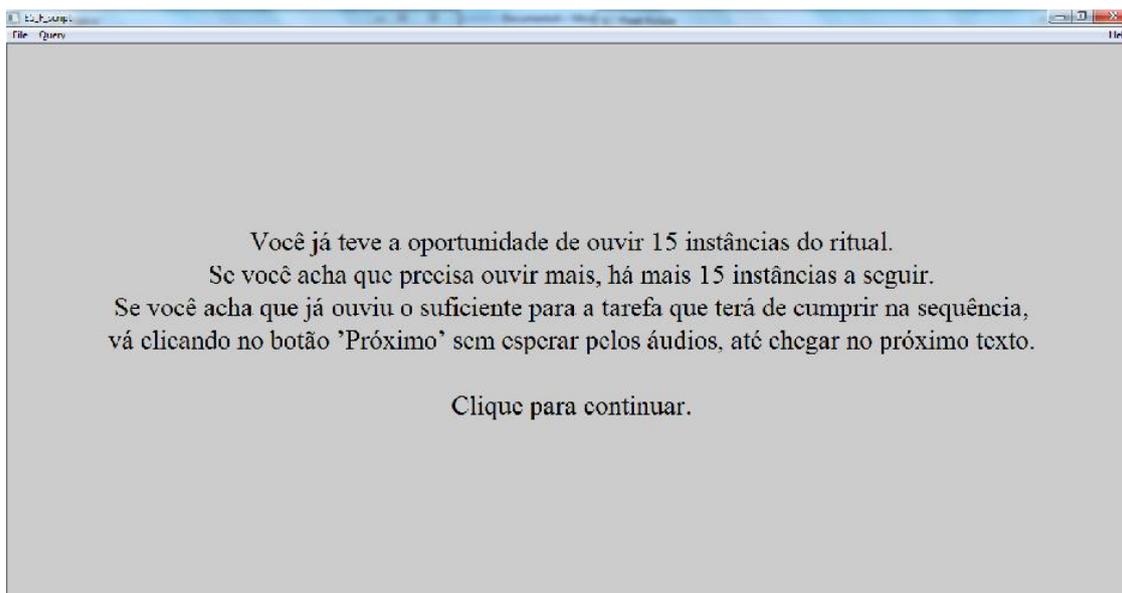


FIGURA 32: INSTRUÇÕES APÓS EXPOSIÇÃO A 15 ESTÍMULOS DE FAMILIARIZAÇÃO.

O motivo pelo qual foi dada ao participante essa escolha é porque trinta estímulos seguidos podem deixá-lo muito cansado para efetuar a tarefa que ainda lhe será dada. Além disso, supôs-se que aqueles que ainda não tivessem encontrado um padrão nos estímulos ouviriam os próximos estímulos e, aqueles que já tivessem chegado a alguma generalização sobre os estímulos poderiam passar para frente. Após escutarem ou pularem os 15 estímulos de familiarização seguintes, os ouvintes chegaram às duas seguintes telas:

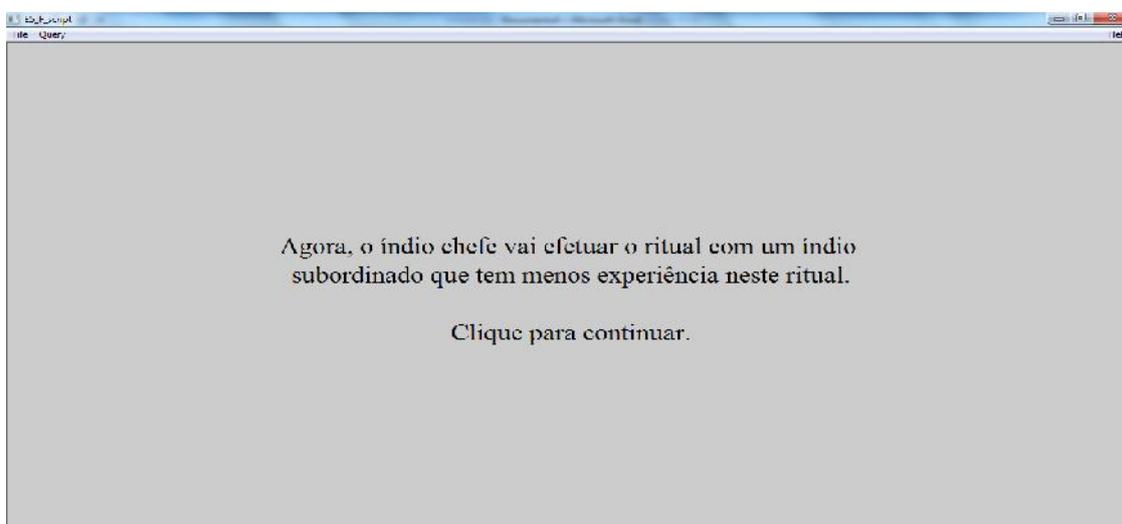


FIGURA 33: TELA DE INSTRUÇÕES 1 APÓS EXPOSIÇÃO A ESTÍMULOS DE FAMILIARIZAÇÃO.

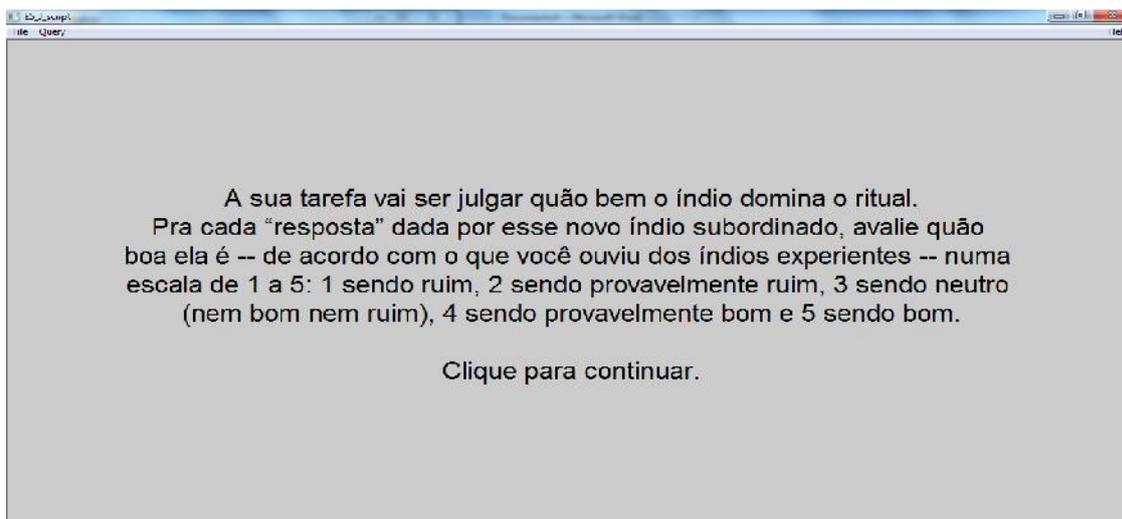


FIGURA 34: TELA DE INSTRUÇÕES 2 APÓS EXPOSIÇÃO A ESTÍMULOS DE FAMILIARIZAÇÃO.

Em seguida, os participantes ouviram 20 estímulos não-ambíguos para julgamento, sendo eles divididos igualmente para cada um dos quatro comandos, com mais quatro distratores (os estímulos foram randomizados pelo próprio programa Praat, tendo cada participante ouvido uma ordem diferente). A imagem na tela do computador durante o julgamento dos participantes é:

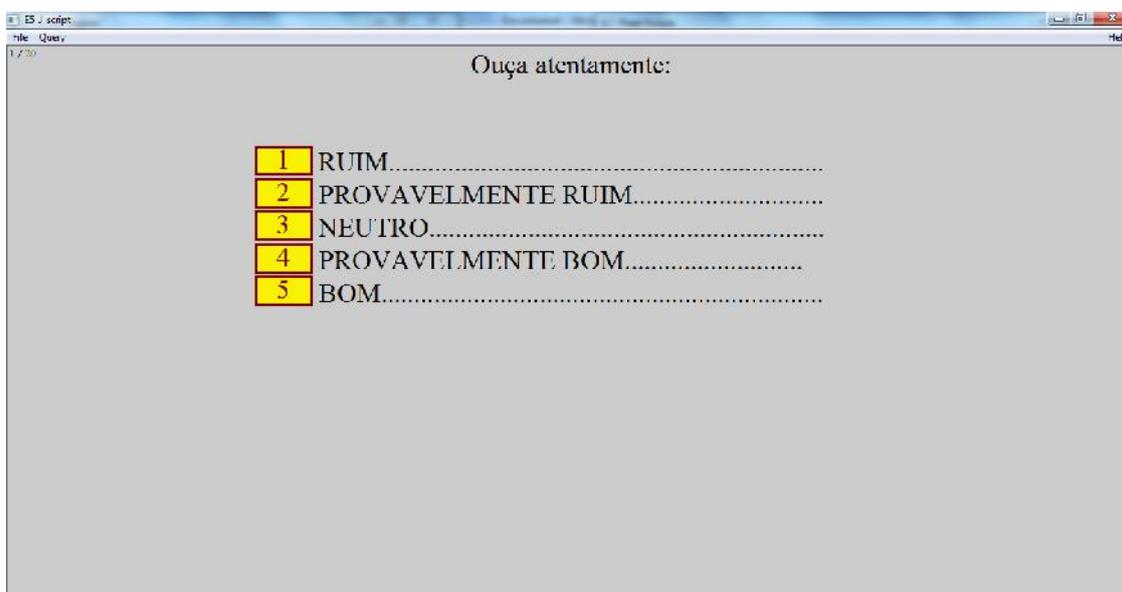


FIGURA 35: TELA PRESENTE DURANTE A ESCUTA DOS ESTÍMULOS DE JULGAMENTO.

Por fim, aparecia à vista do participante uma tela indicando o fim do experimento e agradecendo-lhe por ter participado.

3.4.1.5 Resultado

O teste estatístico utilizado neste experimento foi o teste não paramétrico Friedman. Com este teste, avaliou-se a relação entre as condições duas a duas, ou seja, buscou-se a existência de diferenças significativas entre cada duas das cinco inversões dos estímulos de julgamento do experimento. O resultado a que se chegou foi que, ao nível de 5% de significância, este experimento apresenta diferenças estatisticamente significativas entre as condições ($p < 0.001$). A Tabela 14 mostra quais condições são significativamente diferentes, através das letras apresentadas na segunda coluna: as condições que são significativamente diferentes não possuem nenhuma letra em comum; quando duas condições possuem letras em comum, não são significativamente diferentes. A terceira coluna mostra a soma dos *ranks* de cada condição. O resultado nos mostra que as duas primeiras inversões (as consideradas "naturais" por Nevins e Endress) são significativamente mais bem aceitas do que todas as outras, incluindo os distratores. Entretanto, há também uma diferença significativa entre a aceitabilidade da primeira inversão e da segunda inversão (sendo a primeira, que é a inversão total, mais bem aceita). Ademais, a condição 4 é significativamente diferente dos distratores, mas a condição 3 não é significativamente diferente nem da condição 4 e nem dos distratores.

Resultado Experimento 5		
Condição	Grupo	Soma dos ranks
C1 (4321)	A	55,5
C2 (4231)	B	46
C3 (1432)	CD	28
C4 (3412)	C	30,5
DIS (4xy1)	D	20

TABELA 14: RESULTADO DO EXPERIMENTO 5.

No próximo capítulo teremos a oportunidade de discutir todos esses resultados.

4. DISCUSSÃO

4.1 Introdução

No capítulo anterior, foi mostrado o resultado quantitativo de cada um dos cinco experimentos isoladamente (seções 3.3.1.5, 3.3.2.5, 3.3.3.5, 3.3.4.5 e 3.3.5.5). Neste capítulo, serão contemplados os resultados em conjunto e, ademais, ofereceremos algumas discussões na tentativa de explicar os resultados a que se chegou com os experimentos conduzidos.

Conforme explicitado no capítulo anterior, dentre os cinco experimentos conduzidos, apenas três deles mostraram resultados estatisticamente significativos, ao nível de 5% de significância. São eles (i) o Experimento 2 (E2, em que $p = 0.04$), que é o experimento musical com contextualização tonal e em que o domínio de aplicação da transformação foi a frase musical (com a função *tônica* nas extremidades), (ii) o Experimento 4 (E4, em que $p = 0.04$), que é o experimento musical com estímulos que não apresentam nem contexto tonal e nem domínio de aplicação e (iii) o Experimento 5 (E5, em que $p < 0.001$), que é o experimento linguístico. Para o Experimento 1 (E1) e o Experimento 3 (E3), a análise mostrou que a variável observada não estava afetando significativamente o julgamento dos participantes. As diferenças verificadas entre as condições de E2, E4 e E5 podem ser vistas na Tabela 15.

Condições	Experimentos					
	E2		E4		E5	
	Grupo	Soma dos ranks	Grupo	Soma dos ranks	Grupo	Soma dos ranks
C1 (4321)	A	43	A	48	A	55,5
C2 (4231)	A	43,5	B	34,5	B	46
C3 (1432)	AB	34,5	B	30,5	CD	28
C4 (3412)	AB	34,5	B	30	C	30,5
DIS (4xy1)	B	24,5	AB	37	D	20

TABELA 15: RESULTADOS DOS EXPERIMENTOS QUE APRESENTARAM RESULTADO ESTATISTICAMENTE SIGNIFICATIVO (E2, E4 E E5).

As condições de um mesmo experimento que apresentam pelo menos uma letra em comum no grupo não são significativamente diferentes; as que apresentam letras diferentes são significativamente diferentes, sendo que a soma dos *ranks* mostra para que lado aponta essa diferença (ou seja, quais condições são mais bem-aceitas e quais são menos bem-aceitas). Como podemos ver, em E2 as duas primeiras condições são significativamente mais bem-aceitas do que os distratores, mas, embora mais bem-aceitas do que as condições 3 e 4, essa diferença não é significativa. Em E4, a primeira condição é significativamente mais bem-aceita do que as outras três condições, mas não *significativamente* mais bem-aceita do que os distratores, embora mais bem-aceita. Em E5, vemos diferenças significativas entre diversas condições diferentes. As duas primeiras condições, conforme previsto pela hipótese, são mais bem-aceitas do que todas as outras condições. Entretanto, há também uma diferença significativa entre a primeira e a segunda condição, sendo a primeira mais bem-aceita do que a segunda.

4.2 Análise qualitativa dos resultados

Embora apenas dois dos quatro experimentos musicais tenham tido resultados estatisticamente relevantes, ao analisar as respostas individuais dos participantes, de um

modo geral não pareceu que eles tivessem chegado a alguma generalização a respeito do ordenamento das unidades musicais. Não pudemos encontrar nenhuma variável não controlada que estivesse influenciando na inconsistência dos resultados individuais dos participantes. Tampouco os contornos pareceram influenciar nos julgamentos. Seria esperado, por exemplo, que os estímulos de C1 (inversão total) de E1 cujas inversões do índio subordinado são descendentes (ou seja, em que o índio chefe apresenta uma melodia ascendente), fossem mais bem-aceitos do que os estímulos em que o índio chefe produz uma melodia descendente e o subordinado produz uma melodia ascendente. Isso porque, em uma frase musical do tipo antecedente-consequente na música, o antecedente costuma finalizar em um contorno ascendente e o consequente costuma finalizar em um contorno descendente. Entretanto, tirada a mediana dos dados agrupados de acordo com condição e contorno, obtivemos exatamente o mesmo resultado para o julgamento dos dados do primeiro tipo pertencentes à C1 de E1 (med = 4) e dos dados do segundo tipo pertencentes à C1 de E1 (med = 4). Esperava-se, ainda, que fosse mais visível a distinção de julgamento relacionado ao ordenamento das unidades musicais entre os estímulos cujos contornos são simples (apenas ascendente ou apenas descendente), em oposição aos estímulos cujos contornos são mistos (ascendente-descendente e descendente-ascendente). Entretanto, a análise qualitativa dos resultados não evidenciou nada a esse respeito.

O que a análise qualitativa foi capaz de nos mostrar é o fato de que os participantes músicos foram mais capazes de chegar a uma generalização relacionada ao ordenamento das unidades musicais do que os participantes não músicos. Conforme mencionado na descrição do procedimento dos experimentos, foi dada aos participantes uma ficha para que preenchessem certas informações requeridas. Uma dessas informações era sobre o grau de conhecimento musical: leigo, amador com menos de cinco anos de experiência, amador com mais do que cinco anos de experiência e profissional. Não houve nenhum músico profissional dentre os informantes que participaram dos experimentos musicais, mas houve um total de 4 participantes amadores com mais do que cinco anos de experiência. Desses quatro, três claramente chegaram à generalização de que o que estava em jogo nos dados era a inversão total: em um intervalo de 1 a 5, que foram os valores utilizados para os participantes julgarem a aceitabilidade, a mediana de cada uma das condições para esse grupo de três participantes musicalmente treinados foi: med(C1) = 5, med(C2) = 2, med(C3) = 1,

$\text{med}(C4) = 1,5$, $\text{med}(\text{Dis}) = 1,5$ (apresentamos apenas a mediana, pois não é possível realizar um teste estatístico confiável sobre dados de tão poucos participantes).

Quanto ao experimento linguístico, a análise qualitativa mostrou que a escolha dos distratores – em que são mantidas as sílabas de extremidade, porém com ordem invertida e em que as unidades centrais têm sua identidade modificada – foi uma escolha ruim, pois confundiu pelo menos um participante. Como a ordem em que os estímulos eram apresentados aos participantes era randomizada pelo próprio programa PRAAT, para este participante em questão ocorreu que os dois primeiros dados de julgamento apresentados foram justamente dois distratores. Com esse *input*, o participante passou a aceitar quaisquer estímulos resultantes das quatro inversões, o que mostra que, mesmo que o participante tivesse apreendido uma generalização que diz respeito à *ordem* das sílabas, após esses dois estímulos iniciais sua generalização se converteu para a *identidade* das sílabas: mantendo-se a identidade delas, independentemente da ordem em que aparecem, o estímulo corresponde àqueles ouvidos na etapa de familiarização. Dada a observação de que os distratores influenciaram negativamente o julgamento dos dados linguísticos, é possível também que a mesma influência negativa tenha sido gerada pelos distratores dos experimentos musicais. Entretanto, não foi possível constatar a influência ou não dos distratores através da análise da consistência de julgamento intraindividual nos experimentos musicais.

4.3 Discussão

4.3.1 Sobre música

Os resultados a que chegamos com os quatro experimentos musicais conduzidos foram parcialmente inesperados. Considerando-se a hipótese de que os participantes atentam para a ordem das unidades de uma melodia musical, como os experimentos musicais foram montados de forma a haver uma gradação na retirada de informações dos dados de E1 a E4, esperaríamos das duas uma: ou que (i) os experimentos E1 e E2 (ou apenas um deles) apresentariam um resultado estatisticamente relevante por terem estímulos com mais informações, o que facilitaria o julgamento do ouvinte, por serem informações facilitadoras ou que (ii) os experimentos E3 e E4 (ou apenas E4) apresentariam um resultado estatisticamente relevante por terem estímulos com menos

informação, o que favorece que o ouvinte preste atenção somente na ordenação das unidades, pois o excesso de informações de E1 e E2 desvia a atenção do ouvinte para variáveis que não estão sendo controladas. Entretanto, o que aconteceu foi que E2 e E4 se mostraram estatisticamente relevantes, mas não E1 e E3.

A primeira hipótese formulada após a verificação dos resultados inesperados foi a de que, em princípio, todos os experimentos deveriam ter dado o resultado de que os ouvintes, ao primeiramente se depararem com estímulos musicais de três unidades e com suas respectivas inversões e em seguida serem solicitados a julgar estímulos com quatro unidades com inversões resultantes das quatro generalizações em questão, preferem a inversão total a qualquer outra das inversões. Primeiramente porque, mesmo quando não estatisticamente relevantes, as medianas dos dados nos mostram que a inversão total sempre apresenta um resultado mais alto do que qualquer outra inversão. Em segundo lugar porque, ao isolar os dados dos músicos com mais experiência, foi esse o resultado que a análise individual de cada um deles mostrou, mesmo tendo cada um participado de um experimento diferente. E, em terceiro lugar, porque antes de conduzirmos os experimentos aqui relatados com o público geral, foi feito um experimento com um grupo controle de três participantes com ouvido absoluto, e seus resultados foram categóricos: 100% de aceitação da primeira inversão e 100% de rejeição das outras inversões. Dadas essas motivações para resultados que corroborem nossa hipótese, por que os resultados foram aparentemente caóticos? Em E4, por exemplo, embora tenhamos chegado a uma diferença significativa entre a inversão total e todas as outras inversões – tendo a primeira recebido um maior grau de aceitação por parte dos participantes – os distratores não se diferenciam significativamente de nenhuma das inversões. Isso significa que a diferença no julgamento da inversão total e dos distratores é estatisticamente irrelevante, o que é contrário às nossas expectativas.

Embora tenhamos conduzido o experimento em especial com participantes sem treinamento explícito em música, sob o argumento já reforçado pela literatura de que mesmo quem não tem treinamento explícito em música *entende* música em um sentido tácito, conforme explicitado na seção 1.2.2, acreditamos que a ausência de treinamento musical tenha sido extremamente relevante para o desempenho dos participantes na tarefa proposta em nossos experimentos. Por menor que seja, o treinamento explícito sobre alguma atividade humana possibilita ao participante organizar um mínimo de seu conhecimento tácito em sua consciência; possibilita ao participante refletir a respeito

dessa atividade; possibilita ao participante reconhecer nomeadamente padrões e generalizações dessa atividade. Dado que uma tarefa como a proposta por nossos experimentos exige uma reflexão consciente a respeito do que está acontecendo com os estímulos, os participantes leigos, por não serem capazes de acessar esse conhecimento tácito em um nível consciente no momento de condução do experimento, não conseguiram atingir o mesmo nível de desempenho dos participantes com treinamento explícito em música. Não é novidade que existe essa diferença de desempenho entre participantes músicos e participantes não músicos. Locke e Kellar (1973), por exemplo, ao testar a percepção musical categorial em um grupo de músicos e um grupo de não músicos, chegam ao resultado de que não músicos não mostram evidência clara de percepção categorial. Para Sloboda (1985, p. 27), o fato de músicos treinados terem aprendido nomes de notas musicais e de acordes pode ser um fator que impulse a categorização de sons musicais por parte dos indivíduos.

Os informantes que participaram do experimento linguístico, ao contrário, receberam desde a infância treinamento suficiente, com a alfabetização e escolarização, para serem capazes de compreender o fenômeno em questão em um nível consciente. Para obter uma comparação mais fiel entre os resultados dos experimentos musicais, por um lado, e do experimento linguístico, por outro, o conhecimento explícito recebido pelos informantes a respeito da atividade sendo avaliada deve coincidir ao máximo, o que indica que seria ideal comparar participantes com treinamento musical para os experimentos musicais com participantes escolarizados para os experimentos linguísticos. Uma outra possibilidade seria comparar participantes sem treinamento musical e participantes analfabetos. Entretanto, lidar com analfabetos pode comprometer os resultados, pelo fato de os analfabetos apresentarem problemas gerais de compreensão e interpretação de texto, o que possivelmente faria com que eles sequer entendessem as instruções para realização do experimento.

Há ainda uma segunda hipótese que contempla parte dos resultados dos experimentos musicais. A hipótese é que a diferença fundamental entre os Experimentos 1, 2 e 3 reside na maneira como foi manipulado o domínio de aplicação, sendo que apenas o Experimento 2 apresentou resultados significativos. Abaixo exploramos o significado dessa variável para os resultados.

Na linguagem, o domínio dentro do qual se aplica a operação de manipulação do ordenamento das sílabas nas línguas secretas de inversão é a palavra. Ou seja, se fosse

aplicada a inversão total dentro de uma sentença, como "A menina é bonita", o resultado seria "A nanime é tanibo", e não, por exemplo, "Tanibo é nanime a". Afinal, a ordem das sílabas *dentro de cada palavra* é invertida, mas a ordem das palavras, dos sintagmas e dos fonemas de que as sílabas são formadas permanece a mesma. Por este motivo é que se diz que o domínio ao qual se aplica essa operação é a palavra.

Entretanto, no experimento linguístico que conduzimos, não se manipula nenhum domínio específico, embora o domínio no qual a operação se aplica na realidade da fala seja o domínio da palavra. Isso ocorre porque o experimento lida com logatomas, que são esvaziados de sentido. Logo, embora os logatomas sejam palavras fonológicas⁴⁶, essas "palavras" pronunciadas pelos índios poderiam ser, em princípio, uma frase ou um sintagma qualquer, devido ao esvaziamento de significado. Por exemplo, para o caso dos logatomas com três sílabas, é possível que a primeira sílaba seja um pronome clítico pré-verbal, a sílaba tônica seja um verbo e a última sílaba seja um outro pronome clítico, mas pós-verbal, estrutura que se assemelharia à sentença "Eu vi-te", por exemplo. O motivo pelo qual as transformações feitas nos estímulos do experimento linguístico não necessariamente lidam com palavras reside na propriedade da linguagem de apresentar dupla articulação (ver 3.3). Sendo possível desatrelar som de significado, é possível criar estruturas puramente fonológicas que não apresentam uma estrutura sintático-semântica e que, portanto, são sintaticamente ambíguas.⁴⁷

A música, por outro lado, conforme já foi discutido, não apresenta dupla articulação – ou, se a apresenta, é impossível desatrelar as duas articulações na superfície, devido à coincidência das unidades mínimas das duas articulações – e muito embora tenhamos apresentado uma extensa discussão a respeito do alinhamento da sílaba, que é a unidade sendo manipulada, com determinada unidade musical (no caso, as próprias notas musicais), não foi apresentada nenhuma discussão a respeito do

⁴⁶ Uma palavra fonológica é uma unidade que apresenta independência fonológica, determinada pela sua grade acentual. Os clíticos, por exemplo, como o artigo 'a' e o pronome 'me', embora apresentem certa independência sintática, fonologicamente são dependentes das palavras que os circundam. Assim, o sintagma "a menina", embora apresente pelo menos dois núcleos sintáticos (o artigo 'a' e o nome 'menina'), fonologicamente constitui uma única palavra.

⁴⁷ Dada esta constatação, torna-se necessário discutir o uso da palavra "sentença" nas instruções do experimento linguístico. As instruções diziam que "o chefe sempre pronuncia uma **sentença**, e o subordinado tem que "responder" essa **sentença** de forma apropriada". Para aproximar a situação criada no experimento à realidade em que opera o tipo de transformação observada no experimento, seria ideal que as instruções dissessem que "o chefe sempre pronuncia uma **palavra**, e o subordinado tem que "responder" essa **palavra** de forma apropriada". O motivo pelo qual as instruções foram formuladas da primeira forma, e não da segunda, é que o experimento original de Nevins e Endress utiliza a palavra *sentence* nas instruções. Como o objetivo aqui era recriar o experimento modificando-se os estímulos, as instruções foram reproduzidas semelhantemente àquelas do experimento em que nos baseamos.

alinhamento da palavra, que é a unidade dentro da qual a manipulação ocorre, com determinada unidade musical. Dada a falta de dupla articulação na música, esse alinhamento deveria ser discutido e motivado, já que a estrutura fonológica da música não seria sintaticamente ambígua no sentido exposto no parágrafo acima. Por este motivo, dos quatro experimentos musicais, três deles foram criados de modo a se diferenciar entre si basicamente nessa variável: na falta de uma hipótese para o alinhamento entre a palavra e determinada unidade musical, optamos pela condução de experimentos que manipulassem essa mesma variável de formas distintas: o Experimento 1 manipula antecedente-consequente, o Experimento 2 manipula frases musicais e o Experimento 3 não apresenta nenhum domínio reconhecido.

Dada a discussão apresentada nos três parágrafos acima, o motivo pelo qual acreditamos que o Experimento 1 tenha falhado em apresentar resultados significativos é que a maneira como os estímulos foram criados não foi tomando-se um domínio, transformando-o e gerando uma outra instância do mesmo domínio (por exemplo, tomando-se uma palavra, empregando a operação e gerando outra palavra), e sim tomando-se um domínio (o antecedente de uma frase musical), transformando-o e gerando *outro* domínio (o consequente de uma frase musical). Desse modo, o experimento falha em recriar nos estímulos musicais o mesmo tipo de operação que está sendo verificada no experimento linguístico, o que em si já invalida a comparação entre os dois experimentos (o Experimento 1 e o Experimento 5, que é o linguístico). À parte a irrelevância do Experimento 1 para o fenômeno observado neste trabalho, um estudo diferente deveria ser feito para entender o motivo pelo qual os resultados desse experimento não foram significativos.

Já o Experimento 3 simplesmente não manipula nenhum domínio específico. Entretanto, pelo fato de as estruturas sintática e prosódica na música (PR e TSR, respectivamente) não poderem ser dissociadas, a ausência de domínio, de uma unidade completa, de um nexos sintático, essa falta pode ter influenciado negativamente no julgamento dos participantes e, por isso, o resultado do Experimento 3 se mostrou estatisticamente não significativo.

O Experimento 2, por outro lado, toma uma frase musical como *input* da operação e, após aplicar a esse *input* a operação que modifica a ordem de suas unidades, gera como *output* uma sequência que é do mesmo tipo que o *input*, ou seja, gera também uma frase musical. Essa transformação é paralela àquela verificada no

fenômeno linguístico, em que se toma uma palavra, transforma-a e gera-se outra palavra. É por esse motivo que acreditamos que o Experimento 2 tenha apresentado um resultado significativo. Entretanto, a diferença significativa está entre as quatro inversões, por um lado, e os distratores, por outro. Embora isso pareça invalidar a hipótese em discussão – já que apenas as duas primeiras inversões geram um *output* bem formado, por ser o único que mantém a tônica (que é a unidade mais propensa a ocupar as extremidades de uma frase musical) nas extremidades – se observarmos os *ranks* na tabela que apresenta o resultado desse experimento (Tabela 6, p. 96), vemos que há uma tendência dos participantes em preferir as duas primeiras inversões às outras. Como o experimento foi rodado com apenas 12 participantes, é possível que uma nova aplicação, desta vez com mais sujeitos, defina melhor a significância ou não dessa tendência apresentada pelos participantes do Experimento 2. Se os resultados de uma nova aplicação do Experimento 2 forem condizentes com a tendência à preferência das duas primeiras inversões às outras, teríamos aí um forte paralelo com o resultado do experimento linguístico.

4.3.2 *Sobre linguagem*

Nosso experimento linguístico também mostrou um resultado diferente do resultado do experimento linguístico de Nevins e Endress. Embora tenha havido, sim, um maior grau de aceitação das inversões resultantes dos comandos 1 e 2 em relação às outras, a diferença entre a aceitabilidade do comando 1 e a aceitabilidade do comando 2 também se mostrou, em um nível de 5% de significância, relevante, contrariamente aos resultados de Nevins e Endress, que não mostrou relevância na diferença de aceitação entre essas duas inversões. Houve, ainda, uma diferença significativa entre a aceitabilidade do comando 4 e a aceitabilidade dos distratores. Houve, assim, uma graduação significativa de aceitabilidade entre quatro das inversões: C1-C2-C4-Dis (o comando 3 não se mostrou significativamente diferente nem de C4 e nem de Dis), o que mostra que esses resultados não são categóricos, e sim graduais.

4.3.3 *Sobre música e linguagem*

O que esses resultados nos dizem sobre a relação entre nossa capacidade musical e nossa capacidade linguística? Supondo que a primeira hipótese apresentada na

discussão sobre os resultados musicais (seção 4.3.1) esteja correta, existe uma relação nos resultados de todos os experimentos, que é a aceitação indiscutível da inversão total – resultante do comando 1. Deixando de lado o fato de que os resultados não apontam para uma coincidência total entre a música e a linguagem, dado que, na linguagem, o comando 2 foi bem-aceito e na música não (salva a discussão sobre o Experimento 2), ainda assim não se pode ter uma conclusão definitiva sobre as relações cognitivas entre música e linguagem com apenas a investigação deste fenômeno específico.

Existem, pelo menos, três interpretações para as implicações desses resultados na cognição como um todo: (i) a inversão total é um mecanismo disponível para nossa capacidade cognitiva geral e, por isso, está disponível igualmente para a linguagem e a música, independentemente do fato de música e linguagem serem ou não módulos cerebrais específicos; a verificação desta hipótese poderá ser feita quando testada a disponibilidade desse tipo de inversão para outros domínios cognitivos, (ii) música e linguagem compartilham a mesma gramática, conforme a hipótese de Katz e Pesetsky, o que pode ser verificado pela semelhança dos resultados no que diz respeito à inversão total e no que diz respeito às inversões resultantes dos comandos 1 e 2 para o Experimento 2, particularmente; essa hipótese poderá ser verificada com a investigação de outras propriedades em comum exibidas por essas duas atividades, bem como pela investigação dessa e de outras propriedades em outros domínios cognitivos, para que se possa descartar a hipótese da existência de um único módulo cognitivo geral e (iii) música e linguagem são módulos distintos no cérebro humano, e a existência de uma ou outra propriedade comum a essas duas atividades provém da existência de mecanismos *domain-bound* (ENDRESS; NESPOR; MEHLER, 2009, p. 348), que são mecanismos similares a mecanismos de domínio específico, mas que são compartilhados por pelo menos dois domínios e, embora esses mecanismos possam ter propriedades muito similares, cada um deles opera em apenas um domínio específico, como se cada domínio tivesse uma "cópia" de um mesmo mecanismo, sendo que cada cópia opera somente naquele domínio no qual está; a diferença entre mecanismos *domain-bound* e mecanismos de domínio geral, portanto, é que cada um desses mecanismos é implementado de tal forma que só opera em um domínio particular.

Com relação à aceitabilidade dos dados resultantes do comando 2, se a segunda hipótese apresentada na seção 4.3.1 for corroborada a partir da aplicação de uma nova instância do Experimento 2, é muito possível que se mostre aí um paralelo entre a

música e a linguagem. Entretanto, a maior parte dos outros experimentos musicais mostrou uma maior tendência antes à aceitabilidade apenas do comando 1 do que à aceitabilidade dos comandos 1 e 2. Se esse for o caso, é possível afirmar que música e linguagem diferem nesse ponto. A questão que se coloca sobre tal diferença é: essa é uma diferença da *gramática* ou uma diferença de superfície? Ou seja, o aparecimento dessa diferença na maneira como os dados dos experimentos foram julgados pelos participantes é resultado da ausência desse tipo de operação na gramática musical – o que estabelece uma diferença entre a gramática musical e a gramática linguística – ou é resultado da impossibilidade de essa operação ser realizada na superfície musical devido a alguma propriedade dos "blocos de construção" da música? O fato de a menor unidade sintática na música coincidir com a menor unidade prosódica, o que não acontece com a linguagem, pode estar interferindo na capacidade da música de fazer uso de um mecanismo que, de outra forma, estaria disponível a ela. Isso porque, como podemos ver na implementação dessa operação na linguagem natural, essa operação manipula *somente* unidades fonológicas, sem interferir nas relações sintáticas e semânticas. Se as unidades fonológicas são coincidentes com as unidades sintáticas na música, seria impossível manipulá-las sem interferir nas relações sintáticas. Assim, mesmo com esse resultado discrepante em mãos, é possível manter a hipótese de Katz e Pesetsky de que a gramática da música e da linguagem são iguais. Entretanto, é igualmente possível manter a hipótese de que as gramáticas são diferentes. Ou seja, experimentalmente é impossível verificar qualquer uma dessas hipóteses, visto que superficialmente não será possível verificar o comportamento da estrutura prosódica sem a interferência da estrutura sintática na música. Com isso, podemos afirmar que a hipótese de identidade formulada por Katz e Pesetsky é impossível de ser verificada experimentalmente através da comparação entre propriedades puramente fonológicas e segmentais da linguagem e propriedades musicais, devido à inexistência de segmentos musicais que sejam "vistos" pela TSR, mas não pela PR. A hipótese de Katz e Pesetsky deve, então, ser descartada? Não necessariamente. É ainda possível testar essa hipótese através da comparação entre fenômenos *sintáticos* na música e na linguagem; mas quando se dissocia o conteúdo sintático da superfície fonológica, aí torna-se impossível uma comparação com a estrutura musical, que não permite tal dissociação.

4.4 Considerações finais

Durante a realização deste trabalho, a tentativa de alinhar música e linguagem experimentalmente mostrou ser um desafio bastante grande, mesmo tendo motivado o proposto alinhamento com o trabalho de Katz e Pesetsky (2011), em que já se propõe o alinhamento formal entre música e linguagem, o que, em princípio, facilitaria o trabalho de alinhar experimentalmente música e linguagem. O principal motivo pelo qual este trabalho revelou tamanho desafio é que o fenômeno sendo observado, comparado e controlado experimentalmente é verificado apenas na realidade linguística, porém não na realidade musical. Vale deixar claro que essa diferença verificada na realidade dos dois domínios não é suficiente para se descartar o alinhamento. Afinal, pode ser que a inexistência desse fenômeno de inversão de unidades da cadeia musical seja puramente acidental, seja (i) pelo fato de que a inversão, servindo como mecanismo de codificação, não teria aplicabilidade funcional à música, já que a música não apresenta conteúdo semântico proposicional e, assim, não haveria motivos para codificá-la, ou ainda (ii) por um simples erro de amostragem porque, por exemplo, não foram descritas à exaustão todas as manifestações musicais de todas as culturas existentes do mundo.

Independentemente de qual é o motivo pelo qual esse fenômeno não é observado na realidade musical, tal inexistência torna extremamente difícil a criação de dados em que se simula a existência do fenômeno em melodias musicais, pois as possibilidades de criação de dados são muitas, compatíveis com diversas hipóteses, mesmo após serem restringidas através da utilização do alinhamento formal de Katz e Pesetsky no alinhamento experimental, o que é claramente demonstrado pela necessidade que tivemos de realizar não um, mas quatro experimentos musicais. Isso mostra que a tentativa de verificar a especificidade de um fenômeno linguístico é uma tarefa nada trivial, independentemente de qual seja a propriedade cuja especificidade se quer mostrar e de qual seja o domínio cognitivo com o qual se compara a linguagem.

Os experimentos por nós conduzidos foram capazes de revelar que a complexidade com a qual se deve tratar o alinhamento entre música e linguagem é de veras grande, pois mostraram que os cuidados que se devem tomar são maiores do que os evidentes. Tornou-se claro, por exemplo, que um novo conjunto de experimentos deverá ser realizado em pesquisas futuras, desta vez tomando-se apenas músicos profissionais como participantes. Afinal, conforme apontado na seção 4.3.1, a exposição que qualquer pessoa escolarizada teve à reflexão metalinguística é muito maior do que a

exposição à reflexão metamusical que uma pessoa escolarizada sem treinamento musical teve. Por isso, a fim de comparar a capacidade linguística de uma pessoa escolarizada com a capacidade musical de uma pessoa por meio de experimentos que pedem julgamento de aceitabilidade de dados linguísticos e musicais, a pessoa sujeita ao experimento musical deve ser, também, musicalmente treinada.

Além disso, a realização deste trabalho nos conduziu à hipótese de que, pelo menos no que diz respeito a fenômenos puramente fonológicos, a hipótese de Katz e Pesetsky é impossível de ser experimentalmente averiguada, devido à impossibilidade de dissociação das estruturas TSR e PR na música. Por este motivo, o resultado deste trabalho nos leva a propor que, em pesquisas futuras, seja estabelecida uma comparação entre fenômenos sintáticos na música e na linguagem, já que fenômenos sintáticos pressupõem a coexistência de fenômenos "semânticos" (ou "interpretativos", para ampliar o termo à música) e também, quando apresentados através de *inputs* auditivos, de fenômenos fonológicos, por motivos óbvios. Uma possibilidade, por exemplo, seria averiguar experimentalmente o fenômeno de movimento de núcleo – ou *merge* interno – que é um fenômeno postulado pela teoria gerativa da linguagem e um fenômeno que, segundo Katz e Pesetsky (2011, p. 37-57), também é abarcado pela teoria gerativa musical, mas sob o nome de *retenção da cadência* por Lerdahl e Jackendoff (1983) (ver 2.3.3). Assim, a hipótese de Katz e Pesetsky poderia ser aplicada experimentalmente e, ademais, a verificação da hipótese seria mais facilmente controlada, já que os experimentos estariam verificando o comportamento de fenômenos dos quais as teorias musical e linguística já dão conta isoladamente, o que facilitaria a criação dos dados, por estarmos lidando com um fenômeno verificado em ambos os domínios cognitivos. Dessa forma, não se buscaria a existência ou não de um fenômeno linguístico na música, e sim a coincidência ou não de mecanismos cognitivos para processar um fenômeno linguístico (movimento de núcleo) e um fenômeno musical (cadência).

REFERÊNCIAS

- BAGEMIHL, B. (1989) The crossing constraint and 'backwards languages'. *Natural language and linguistic theory* 7, p. 481–529.
- BAGEMIHL, B. (1995) Language games and related areas. In: GOLDSMITH, J. A. (ed.) *The handbook of phonological theory*, p. 697-712. Oxford: Blackwell.
- BERNSTEIN, L. (1976) *The unanswered question: six talks at Harvard*. Cambridge, Mass: Harvard University Press.
- BRANDT, A. K.; GEBRIAN, M.; SLEVC, L.R. (2012) Music and early language acquisition. *Frontiers in psychology*, p. 3-327.
- CATCHPOLE, C. K.; SLATER, P. J. B. (1995) *Bird song*. Cambridge University Press.
- CHOMSKY, N. (1957) *Syntactic structures*. The Hague/Paris: Mouton.
- CHOMSKY, N. (1981) *Lectures on government and binding: the Pisa lectures*. Holland: Foris Publications. 7ª edição. Berlin and New York: Mouton de Gruyter, 1993.
- CHOMSKY, N. (1995a) Bare phrase structure. In: WEBELHUTH, G. (ed.) *Government binding theory and the minimalist program*, p. 383-439. Oxford: Oxford University Press.
- CHOMSKY, N. (1995b) *The minimalist program*. Cambridge, MA: The MIT Press.
- CHOMSKY, N.; HALLE, M. (1968) *The sound pattern of English*. New York: Harper & Row, Publishers.
- CHOMSKY, N.; LASNIK, H. (1993) The theory of principles and parameters. In: von STECHOW, J.; JACOBS A.; STERNEFELD, W.; VENNEMANN, T. (eds.). *Syntax: an international handbook of contemporary research*. Berlin: De Gruyter.
- COOK, N. (1987) *A guide to musical analysis*. New York: George Baziller.

COOPER, G.; MEYER, L. B. (1960) *The rhythmic structure of music*. Chicago: University of Chicago Press.

CROSS, I. (2001). Review of *The Origins of Music*. *Music Perception*, v. 18, p. 513–521.

de ARMELLADA, C. (1999) *Gramatica de la Lingua Pemon*. Caracas: Universidad Catolica Andres Bello.

DOWLING, W. J. (1978). Scale and contour: two components of a theory of memory for melodies. *Psychological review*, v. 85, p. 341-354.

ENDRESS, A. D.; NESPOR, M.; MEHLER, J. (2009) Perceptual and memory constraints on language acquisition. *Trends in cognitive science*, 13(8), p. 348-353.

FITCH, W. T.; HAUSER, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303, p. 377–380

GOLDSMITH, J. A. (1976) *Autosegmental phonology*. Tese de doutorado MIT. New York: Garland Press, 1979.

GRABE, E.; LOW, E. L. (2002) Durational variability in speech and the rhythm class hypothesis. In: GUSSENHOVEN, C.; WARNER, N. *Laboratory phonology*. Berlin: Mouton de Gruyter, v. 7, p. 515–546.

HAUSER, M.; CHOMSKY, N.; FITCH, T. 2002. The faculty of language: what is it, who has it, and how did it evolve. *Science* 198. p. 1569-79

HAUSER, M. D.; MCDERMOTT, J. (2003). The evolution of the music faculty: Comparative perspectives. *Nature Neuroscience*, 6, p. 663–668.

HOWE, M. J.; DAVIDSON, J. W.; SLOBODA, J. A. (1998). Innate talents: reality or myth? *Behav Brain Sci*, 21, 399-407; discussão 407-42.

JENTSCHKE, S.; KOELSCH, S.; SALLAT, S.; FRIEDERICI, A. D. (2008) “Children with specific language impairment also show impairment of music-syntactic processing”. *Journal of Cognitive Neuroscience*, v. 20, p. 1940-1951.

KATZ, J.; PESETSKY, D. (2011) *The identity thesis for language and music*. (não-publicado). <http://ling.auf.net/lingBuzz/000959>. Data de acesso: 4-04-2012.

KRUMHANSL, C. L. (1990) *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

LAYCOCK, D. (1972). Towards a typology of ludlings, or play-languages. *Linguistic Communications* 6, p. 61–113.

LERDAHL, F.; JACKENDOFF, R. (1983) *A generative theory of tonal music*. Cambridge; London: The Mit Press.

LEVITIN, D. (2008). *The world in six songs: how the musical brain created human nature*. New York: Penguin.

LIBERMAN, M. (1975) *The Intonational System of English*. Tese de doutorado MIT, Cambridge, Massachusetts.

LIBERMAN, M.; PRINCE, A. (1977) On stress and linguistic rhythm. *Linguistic inquiry*, v. 8, p. 249-336.

LOCKE, S.; KELLAR, L. A. (1973) Categorical perception in a non-linguistic mode. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, v. 9(4), p. 355-369.

MANN, T. (1947) *Doutor Fausto*. Trad. Herbert Caro. Rio de Janeiro: Ed. Nova Fronteira, 2000.

McMULLEN, E.; SAFFRAN, J. R. (2004) Music and language: a developmental comparison. *Music Perception*, v. 21, n° 3; University of California Press.

MEDEIROS, B. R. Em busca do som perdido: o que há entre a lingüística e a música. In: ILARI, B. S. (Org.). *Em busca da mente musical – ensaios sobre os processos cognitivos da música – da percepção à produção*. Curitiba: Editora UFPR, 2006.

MENDES, G. (2008) Análise acústica das vogais nasais. In: 56o. Seminário do GEL, 2008, São José do Rio Preto. Programação do 56.o Seminário do GEL, 2008.

NESPOR, M.; VOGEL, I. (1986) *Prosodic Phonology*. Dordrecht: Foris Publications.

NETTL, B. (2000) An Ethnomusicologist contemplates universals in musical sound and musical culture. In: WALLIN, N. L., MERKER, B. e BROWN, S. (Eds.) *The origins of Music*, p. 463-472. Cambridge, MA: MIT Press.

NEVINS, A.; ENDRESS, A. (2007) The edge of order: analytic biases in ludlings. *Harvard Working Papers in Linguistics*, v. 12, p. 43-53.

NEVINS, A. (2010) Two case studies in phonological universals: a view from artificial grammars. *Biolinguistics 4.2-3*, p. 218-233.

PATEL, A. D.; DANIELE, J. R. (2003) An empirical comparison of rhythm in language and music. *Cognition*, v. 87, p. B35-B45.

PATEL, A. D.; FOXTON, J. M.; GRIFFITHS, T. D. (2005) Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition*, v. 59, p. 310-313

PATEL, A. D. (2003) Language, music, syntax and the brain. *Nature Neuroscience*, v. 6, n° 7, p. 674-681.

PATEL, A. D. (2008) *Music, Language, and the Brain*. New York: Oxford University Press.

PEPPERBERG, I. M. (2000). *The Alex studies: Cognitive and communicative abilities of Grey parrots*. Cambridge, MA: Harvard University Press

PERETZ, I. (2008) Music, language and modularity in action. In: REBUSCHAT, P., ROHRMEIER, M., HAWKINS, J. & CROSS, I. (eds.) *Language and music as cognitive systems*. p. 57-75. Oxford, New York: Oxford University Press; 2008.

PERETZ, I.; COLTHEART, M. (2003) Modularity of music processing. *Nature Neuroscience*, v. 6, n° 7.

PIKE, K. (1945) *The Intonation of American English*. Ann Arbor: University of Michigan Press.

PINKER, S. (1987) The bootstrapping problem in language acquisition. In: MACWHINNEY, B. (Ed.) *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

PINKER, S. (2002). *The blank slate: The modern denial of human nature*. New York: Penguin.

PINKER, S.; JACKENDOFF, R. (2005). The faculty of language: What's special about it? *Cognition*, 95, p. 201–236

SACKS, O. *Alucinações musicais: relatos sobre a música e o cérebro*. São Paulo: Companhia das Letras; 2007.

SANTANA, B. P. (2010) Os padrões que ouvimos: uma introdução à interface música e linguagem. Trabalho de conclusão de curso, UFPR.

SELKIRK, E. *Phonology and syntax: the relation between sound and structure*. Cambridge, London: The MIT Press, 1984.

SLOBODA, J. A. (1985) *The musical mind. The cognitive psychology of music*. Oxford: Clarendon Press.

TOMASELLO, M. (2003). On the different origins of symbols and grammar. In: CHRISTIANSEN, M.; KIRBY, S. (Eds.) *Language Evolution: The States of the Art*. Oxford University Press.

von FRISCH, K. (1953) *The dancing bees*. Harcourt, Brace & World, Inc.

APÊNDICE A – TABELAS DE DADOS

DADO	Melodia chefe			Melodia subordinado			Direção chefe	Tonalidade
E1_F_D1	1(I)	3(vi)	5(V)	5(V)	3(vi)	1(I)	cima	si
E1_F_D2	1(I)	4(ii)	5(V)	5(V)	4(ii)	1(I)	cima	sol \flat
E1_F_D3	1(I)	2(ii)	4(V)	4(V)	2(ii)	1(I)	cima	ré \flat
E1_F_D4	3(I)	6(IV)	7(V)	7(V)	6(IV)	3(I)	cima	fá
E1_F_D5	3(I)	4(IV)	5(V)	5(V)	4(IV)	3(I)	cima	lá \flat
E1_F_D6	3(I)	5(iii)	7(V)	7(V)	5(iii)	3(I)	cima	ré
E1_F_D7	5(I)	6(IV)	7(V)	7(V)	6(IV)	5(I)	cima	mi
E1_F_D8	5(I)	7(iii)	2(V)	2(V)	7(iii)	5(I)	cima	dó#
E1_F_D9	1(I)	5(iii)	4(V)	4(V)	5(iii)	1(I)	baixo	dó
E1_F_D10	1(I)	6(IV)	5(V)	5(V)	6(IV)	1(I)	baixo	dó#
E1_F_D11	3(I)	6(IV)	5(V)	5(V)	6(IV)	3(I)	baixo	sol \flat
E1_F_D12	3(I)	1(vi)	7(V)	7(V)	1(vi)	3(I)	baixo	mi \flat
E1_F_D13	3(I)	7(iii)	5(V)	5(V)	7(iii)	3(I)	baixo	lá \flat
E1_F_D14	5(I)	4(IV)	2(V)	2(V)	4(IV)	5(I)	baixo	ré
E1_F_D15	3(I)	2(ii)	7(V)	7(V)	2(ii)	3(I)	baixo	dó
E1_F_D16	1(I)	2(ii)	7(V)	7(V)	2(ii)	1(I)	cima-baixo	dó \flat (si)
E1_F_D17	1(I)	4(ii)	2(V)	2(V)	4(ii)	1(I)	cima-baixo	mi \flat
E1_F_D18	3(I)	5(iii)	2(V)	2(V)	5(iii)	3(I)	cima-baixo	mi
E1_F_D19	5(I)	6(IV)	2(V)	2(V)	6(IV)	5(I)	cima-baixo	fá
E1_F_D20	5(I)	6(IV)	4(V)	4(V)	6(IV)	5(I)	cima-baixo	dó
E1_F_D21	5(I)	7(iii)	4(V)	4(V)	7(iii)	5(I)	cima-baixo	si \flat
E1_F_D22	5(I)	1(IV)	7(V)	7(V)	1(IV)	5(I)	cima-baixo	fá#
E1_F_D23	1(I)	5(iii)	2(V)	2(V)	5(iii)	1(I)	baixo-cima	lá
E1_F_D24	1(I)	5(iii)	7(V)	7(V)	5(iii)	1(I)	baixo-cima	dó
E1_F_D25	1(I)	7(iii)	2(V)	2(V)	7(iii)	1(I)	baixo-cima	si
E1_F_D26	3(I)	1(vi)	4(V)	4(V)	1(vi)	3(I)	baixo-cima	fá#
E1_F_D27	3(I)	7(iii)	2(V)	2(V)	7(iii)	3(I)	baixo-cima	sol
E1_F_D28	5(I)	2(ii)	4(V)	4(V)	2(ii)	5(I)	baixo-cima	dó \flat (si)
E1_F_D29	5(I)	3(vi)	7(V)	7(V)	3(vi)	5(I)	baixo-cima	ré \flat
E1_F_D30	5(I)	4(IV)	7(V)	7(V)	4(IV)	5(I)	baixo-cima	lá

TABELA 16: DADOS DE FAMILIARIZAÇÃO DE E1.

DADO	Melodia chefe				Melodia subordinado				Condição	Direção chefe	Tonalidade
E1_C1_D1	5(I)	3(vi)	2(ii)	7(V)	7(V)	2(ii)	3(vi)	5(I)	C1	baixo	mi
E1_C1_D2	1(I)	6(IV)	7(iii)	2(V)	2(V)	7(iii)	6(IV)	1(I)	C1	baixo-cima	réb
E1_C1_D3	3(I)	4(IV)	5(iii)	7(V)	7(V)	5(iii)	4(IV)	3(I)	C1	cima	dó#
E1_C1_D4	1(I)	3(vi)	6(IV)	4(V)	4(V)	6(IV)	3(vi)	1(I)	C1	cima-baixo	dó
E1_C2_D1	5(I)	4(IV)	3(vi)	2(V)	2(V)	4(IV)	3(vi)	5(I)	C2	baixo	láb
E1_C2_D2	3(I)	7(iii)	1(IV)	2(V)	2(V)	7(iii)	1(IV)	3(I)	C2	baixo-cima	ré
E1_C2_D3	1(I)	2(ii)	3(vi)	5(V)	5(V)	2(ii)	3(vi)	1(I)	C2	cima	lá
E1_C2_D4	1(I)	3(vi)	6(IV)	4(V)	4(V)	3(vi)	6(IV)	1(I)	C2	cima-baixo	dób (si)
E1_C3_D1	3(I)	2(ii)	1(IV)	7(V)	3(I)	7(V)	1(IV)	2(ii)	C3	baixo	mib
E1_C3_D2	1(I)	5(iii)	6(IV)	2(V)	1(I)	2(V)	6(IV)	5(iii)	C3	baixo-cima	si
E1_C3_D3	1(I)	3(vi)	4(IV)	5(V)	1(I)	5(V)	4(IV)	3(vi)	C3	cima	solb
E1_C3_D4	5(I)	2(ii)	1(vi)	7(V)	5(I)	7(V)	1(vi)	2(ii)	C3	cima-baixo	fá
E1_C4_D1	5(I)	4(IV)	3(vi)	7(V)	3(vi)	7(V)	5(I)	4(IV)	C4	baixo	sol
E1_C4_D2	1(I)	6(IV)	7(iii)	2(V)	7(iii)	2(V)	1(I)	6(IV)	C4	baixo-cima	dó
E1_C4_D3	5(I)	7(iii)	1(IV)	2(V)	1(IV)	2(V)	5(I)	7(iii)	C4	cima	sib
E1_C4_D4	3(I)	5(iii)	2(ii)	7(V)	2(ii)	7(V)	3(I)	5(iii)	C4	cima-baixo	fá#
E1_Dis_D1	1(I)	6(IV)	3(vi)	2(V)	2(V)	5(iii)	4(IV)	1(I)	Dis	baixo	solb
E1_Dis_D2	5(I)	2(ii)	1(IV)	4(V)	4(V)	3(vi)	6(IV)	5(I)	Dis	baixo-cima	mib
E1_Dis_D3	3(I)	5(iii)	6(IV)	7(V)	7(V)	4(IV)	2(ii)	3(I)	Dis	cima	fá
E1_Dis_D4	3(I)	6(vi)	4(IV)	2(V)	2(V)	1(vi)	5(iii)	3(I)	Dis	cima-baixo	si

TABELA 17: DADOS DE JULGAMENTO DE E1.

DADO	Melodia chefe			Melodia subordinado			Direção chefe	Tonalidade
E2_F_D1	1(I)	6(ii)	5(I)	5(I)	6(ii)	1(I)	baixo	lá b
E2_F_D2	1(I)	7(iii)	5(I)	5(I)	7(iii)	1(I)	baixo	dó
E2_F_D3	3(I)	2(ii)	5(I)	5(I)	2(ii)	3(I)	baixo	fá#
E2_F_D4	3(I)	2(V)	1(I)	1(I)	2(V)	3(I)	baixo	dó
E2_F_D5	5(I)	3(iii)	1(I)	1(I)	3(iii)	5(I)	baixo	ré
E2_F_D6	5(I)	4(ii)	3(I)	3(I)	4(ii)	5(I)	baixo	sol
E2_F_D7	1(I)	6(IV)	3(I)	3(I)	6(IV)	1(I)	baixo-cima	sol b
E2_F_D8	1(I)	7(V)	3(I)	3(I)	7(V)	1(I)	baixo-cima	si
E2_F_D9	3(I)	1(vi)	5(I)	5(I)	1(vi)	3(I)	baixo-cima	ré b
E2_F_D10	5(I)	4(IV)	1(I)	1(I)	4(IV)	5(I)	baixo-cima	fá#
E2_F_D11	1(I)	4(V)	5(I)	5(I)	4(V)	1(I)	baixo-cima	si
E2_F_D12	5(I)	2(ii)	3(I)	3(I)	2(ii)	5(I)	baixo-cima	dó#
E2_F_D13	1(I)	2(V)	3(I)	3(I)	2(V)	1(I)	cima	mi b
E2_F_D14	1(I)	2(ii)	5(I)	5(I)	2(ii)	1(I)	cima	ré
E2_F_D15	5(I)	7(iii)	1(I)	1(I)	7(iii)	5(I)	cima	mi b
E2_F_D16	3(I)	4(ii)	5(I)	5(I)	4(ii)	3(I)	cima	sol b
E2_F_D17	3(I)	4(IV)	1(I)	1(I)	4(IV)	3(I)	cima	si b
E2_F_D18	3(I)	5(V)	1(I)	1(I)	5(V)	3(I)	cima	si b
E2_F_D19	3(I)	6(vi)	1(I)	1(I)	6(vi)	3(I)	cima	dó b
E2_F_D20	5(I)	6(ii)	1(I)	1(I)	6(ii)	5(I)	cima	dó b
E2_F_D21	3(I)	7(V)	1(I)	1(I)	7(V)	3(I)	cima	dó#
E2_F_D22	5(I)	1(IV)	3(I)	3(I)	1(IV)	5(I)	cima	lá b
E2_F_D23	5(I)	7(iii)	3(I)	3(I)	7(iii)	5(I)	cima	mi
E2_F_D24	1(I)	3(vi)	5(I)	5(I)	3(vi)	1(I)	cima-baixo	lá
E2_F_D25	1(I)	4(V)	3(I)	3(I)	4(V)	1(I)	cima-baixo	ré b
E2_F_D26	1(I)	5(iii)	3(I)	3(I)	5(iii)	1(I)	cima-baixo	mi
E2_F_D27	3(I)	6(IV)	5(I)	5(I)	6(IV)	3(I)	cima-baixo	lá
E2_F_D28	3(I)	7(V)	5(I)	5(I)	7(V)	3(I)	cima-baixo	fá#
E2_F_D29	5(I)	2(V)	1(I)	1(I)	2(V)	5(I)	cima-baixo	sol
E2_F_D30	5(I)	6(vi)	3(I)	3(I)	6(vi)	5(I)	cima-baixo	fá#

TABELA 18: DADOS DE FAMILIARIZAÇÃO DE E2.

DADO	Melodia chefe				Melodia subordinado				Direção chefe	Tonalidade
E2_C1_D1	1(I)	7(V)	6(ii)	5(I)	5(I)	6(ii)	7(V)	1(I)	baixo	<i>réb</i>
E2_C1_D2	3(I)	2(V)	7(iii)	1(I)	1(I)	7(iii)	2(V)	3(I)	baixo-cima	lá
E2_C1_D3	5(I)	6(IV)	7(V)	1(I)	1(I)	7(V)	6(IV)	5(I)	cima	ré
E2_C1_D4	3(I)	4(IV)	2(V)	1(I)	1(I)	2(V)	4(IV)	3(I)	cima-baixo	<i>mib</i>
E2_C2_D1	3(I)	2(ii)	7(V)	5(I)	5(I)	2(ii)	7(V)	3(I)	baixo	do
E2_C2_D2	1(I)	5(iii)	7(V)	3(I)	3(I)	5(iii)	7(V)	1(I)	baixo-cima	si
E2_C2_D3	5(I)	6(ii)	7(V)	3(I)	3(I)	6(ii)	7(V)	5(I)	cima	mi
E2_C2_D4	1(I)	2(V)	4(ii)	3(I)	3(I)	2(V)	4(ii)	1(I)	cima-baixo	<i>fá#</i>
E2_C3_D1	5(I)	4(IV)	2(V)	1(I)	5(I)	1(I)	2(V)	4(IV)	baixo	<i>fá#</i>
E2_C3_D2	1(I)	6(IV)	4(ii)	5(I)	1(I)	5(I)	4(ii)	6(IV)	baixo-cima	ré
E2_C3_D3	3(I)	5(V)	7(iii)	1(I)	3(I)	1(I)	7(iii)	5(V)	cima	<i>solb</i>
E2_C3_D4	1(I)	4(ii)	5(V)	3(I)	1(I)	3(I)	5(V)	4(ii)	cima-baixo	<i>sib</i>
E2_C4_D1	1(I)	6(ii)	5(V)	3(I)	5(V)	3(I)	1(I)	6(ii)	baixo	si
E2_C4_D2	5(I)	2(V)	1(IV)	3(I)	1(IV)	3(I)	5(I)	2(V)	baixo-cima	<i>láb</i>
E2_C4_D3	1(I)	2(V)	4(IV)	5(I)	4(IV)	5(I)	1(I)	2(V)	cima	sol
E2_C4_D4	3(I)	4(vi)	6(IV)	5(I)	6(IV)	5(I)	3(I)	4(vi)	cima-baixo	<i>fá</i>
E2_Dis_D1	3(I)	2(V)	1(IV)	5(I)	5(I)	4(IV)	6(ii)	3(I)	baixo	<i>dób</i>
E2_Dis_D2	5(I)	4(ii)	7(V)	1(I)	1(I)	2(V)	3(vi)	5(I)	baixo-cima	<i>dó#</i>
E2_Dis_D3	1(I)	2(V)	3(vi)	5(I)	5(I)	4(ii)	7(V)	1(I)	cima	mi
E2_Dis_D4	5(I)	6(IV)	2(V)	1(I)	1(I)	3(vi)	4(ii)	5(I)	cima-baixo	lá

TABELA 19: DADOS DE JULGAMENTO DE E2.

DADO	Melodia chefe			Melodia subordinado			Direção chefe	Tonalidade
E3_F_D1	2	7	1	1	7	2	baixo-cima	dób (si)
E3_F_D2	6	7	2	2	7	6	cima	solb
E3_F_D3	3	7	1	1	7	3	cima	láb
E3_F_D4	3	7	2	2	7	3	baixo-cima	ré
E3_F_D5	2	5	4	4	5	2	cima-baixo	dó#
E3_F_D6	1	2	7	7	2	1	cima-baixo	réb
E3_F_D7	5	3	2	2	3	5	baixo	dó
E3_F_D8	4	1	7	7	1	4	baixo	solb
E3_F_D9	2	4	7	7	4	2	cima-baixo	ré
E3_F_D10	7	5	1	1	5	7	baixo	mib
E3_F_D11	5	6	4	4	6	5	cima-baixo	fá#
E3_F_D12	5	7	1	1	7	5	cima	réb
E3_F_D13	1	4	3	3	4	1	cima-baixo	fá#
E3_F_D14	4	2	7	7	2	4	baixo	fá
E3_F_D15	7	4	5	5	4	7	baixo-cima	mi
E3_F_D16	3	2	5	5	2	3	baixo-cima	mi
E3_F_D17	7	1	6	6	1	7	cima-baixo	dó
E3_F_D18	5	4	1	1	4	5	baixo	dó
E3_F_D19	7	4	6	6	4	7	baixo-cima	mib
E3_F_D20	6	7	4	4	7	6	cima-baixo	sol
E3_F_D21	7	5	2	2	5	7	baixo-cima	dó
E3_F_D22	4	6	7	7	6	4	cima	fá
E3_F_D23	2	6	7	7	6	2	cima	lá
E3_F_D24	3	4	6	6	4	3	cima	si
E3_F_D25	6	1	3	3	1	6	cima	dó#
E3_F_D26	5	2	6	6	2	5	baixo	sib
E3_F_D27	2	1	4	4	1	2	baixo-cima	lá
E3_F_D28	6	5	7	7	5	6	baixo-cima	si
E3_F_D29	3	4	5	5	4	3	cima	dób (si)
E3_F_D30	4	5	3	3	5	4	cima-baixo	láb

TABELA 20: DADOS DE FAMILIARIZAÇÃO DE E3.

DADOS	Melodia chefe				Melodia subordinado				Direção chefe	Tonalidade
E3_C1_D1	6	5	4	3	3	4	5	6	baixo	<i>mib</i>
E3_C1_D2	4	2	7	1	1	7	2	4	baixo-cima	sol
E3_C1_D3	4	7	1	3	3	1	7	4	cima	<i>fá</i>
E3_C1_D4	2	3	1	4	4	1	3	2	cima-baixo	<i>dób</i>
E3_C2_D1	5	4	3	7	7	4	3	5	baixo	mi
E3_C2_D2	6	2	5	7	7	2	5	6	baixo-cima	ré
E3_C2_D3	2	4	5	6	6	4	5	2	cima	<i>sib</i>
E3_C2_D4	4	5	7	6	6	5	7	4	cima-baixo	mi
E3_C3_D1	2	7	6	3	2	3	6	7	baixo	<i>mib</i>
E3_C3_D2	3	6	5	7	3	7	5	6	baixo-cima	dó
E3_C3_D3	1	4	5	6	1	6	5	4	cima	dó#
E3_C3_D4	6	1	4	3	6	3	4	1	cima-baixo	<i>sib</i>
E3_C4_D1	7	5	4	1	4	1	7	5	baixo	lá
E3_C4_D2	4	3	7	1	7	1	4	3	baixo-cima	si
E3_C4_D3	1	2	4	5	4	5	1	2	cima	<i>fá#</i>
E3_C4_D4	1	5	7	6	7	6	1	5	cima-baixo	ré
E3_Dis_D1	1	7	5	2	2	3	6	1	baixo	<i>réb</i>
E3_Dis_D2	7	5	2	6	6	1	3	7	baixo-cima	<i>láb</i>
E3_Dis_D3	5	6	1	2	2	7	4	5	cima	<i>solb</i>
E3_Dis_D4	3	5	7	4	4	6	2	3	cima-baixo	lá

TABELA 21: DADOS DE JULGAMENTO DE E3.

Dados	Melodia chefe			Melodia subordinado			Direção chefe	Intervalos	
E4_F_D1	fá	ré#	sib	sib	ré#	fá	baixo	2Md	4Jd
E4_F_D2	sol	fá#	mi	mi	fá#	sol	baixo	2md	2Md
E4_F_D3	dó	fá	mi	mi	fá	dó	baixo	5Jd	2md
E4_F_D4	dó	lá#	solb	solb	lá#	dó	baixo	2Md	3Md
E4_F_D5	ré	dó#	solb	solb	dó#	ré	baixo	2Md	5Jd
E4_F_D6	lá	fá#	réb	réb	fá#	lá	baixo	3md	4Jd
E4_F_D7	fá	ré#	ré	ré	ré#	fá	baixo	2Md	2md
E4_F_D8	ré	lá#	mib	mib	lá#	ré	baixo-cima	3m-b	4Ja
E4_F_D9	mi	dó	solb	solb	dó	mi	baixo-cima	3M-b	5dimã
E4_F_D10	ré	lá#	si	si	lá#	ré	baixo-cima	3md	2ma
E4_F_D11	lá	fá	solb	solb	fá	lá	baixo-cima	3Md	2ma
E4_F_D12	dó	lá#	réb	réb	lá#	dó	baixo-cima	2Md	3ma
E4_F_D13	si	láb	dó	dó	láb	si	baixo-cima	3md	3Ma
E4_F_D14	si	lá#	mi	mi	lá#	si	baixo-cima	2md	4auma
E4_F_D15	mi	fá#	sib	sib	fá#	mi	cima	2Ma	3Ma
E4_F_D16	sol	sol#	si	si	sol#	sol	cima	2ma	3ma
E4_F_D17	fá	lá#	réb	réb	lá#	fá	cima	4Ja	3ma
E4_F_D18	si	fá	solb	solb	fá	si	cima	5dimã	2ma
E4_F_D19	lá	dó	dó#	dó#	dó	lá	cima	3ma	2ma
E4_F_D20	ré	sol#	sib	sib	sol#	ré	cima	4auma	2Ma
E4_F_D21	si	dó	solb	solb	dó	si	cima	2ma	5dimã
E4_F_D22	ré	fá#	láb	láb	fá#	ré	cima	3Ma	2Ma
E4_F_D23	si	dó#	sib	sib	dó#	si	cima-baixo	2Ma	3md
E4_F_D24	ré	ré#	réb	réb	ré#	ré	cima-baixo	2ma	2Md
E4_F_D25	lá	dó	sib	sib	dó	lá	cima-baixo	3ma	2Md
E4_F_D26	dó	dó#	láb	láb	dó#	dó	cima-baixo	2ma	4Jd
E4_F_D27	ré	fá	réb	réb	fá	ré	cima-baixo	3ma	3Md
E4_F_D28	fá	sol#	ré	ré	sol#	fá	cima-baixo	3ma	4aumã
E4_F_D29	si	dó#	láb	láb	dó#	si	cima-baixo	2Ma	4Jd
E4_F_D30	dó	fá#	si	si	fá#	dó	cima-baixo	4auma	5Jd

TABELA 22: DADOS DE FAMILIARIZAÇÃO DE E4.

DADO	Melodia chefe				Melodia subordinado				Direção chefe	Intervalos Pergunta		
E4_C1_D1	ré#	si	fá	mi	mi	fá	si	ré#	baixo	3Md	4aum <i>d</i>	2md
E4_C1_D2	fá#	dó	mib	sol	sol	mib	dó	fá#	baixo-cima	4dim <i>d</i>	3ma	3Ma
E4_C1_D3	sib	si	mi	sol#	sol#	mi	si	sib	cima	2ma	4Ja	3Ma
E4_C1_D4	lá	mi	fá	ré	ré	fá	mi	lá	cima-baixo	5Ja	2ma	3md
E4_C2_D1	ré#	dó	si	lá	lá	dó	si	ré#	baixo	3md	2md	2Md
E4_C2_D2	réb	dó	sib	ré	ré	dó	sib	réb	baixo-cima	2md	2Md	3Ma
E4_C2_D3	dó	mib	sol#	lá#	lá#	mib	sol#	dó	cima	3ma	4Ja	2Ma
E4_C2_D4	si	ré#	fá#	fá	fá	ré#	fá#	si	cima-baixo	3Ma	3ma	2md
E4_C3_D1	ré	dó#	lá#	sol#	ré	sol#	lá#	dó#	baixo	2md	3md	2Md
E4_C3_D2	si	sib	ré	lá <i>b</i>	si	lá <i>b</i>	ré	sib	baixo-cima	2md	3Ma	5dim <i>a</i>
E4_C3_D3	mi	sol	sol#	dó#	mi	dó#	sol#	sol	cima	3ma	2ma	4Ja
E4_C3_D4	dó	ré	mib	si	dó	si	mib	ré	cima-baixo	2Ma	2ma	3Md
E4_C4_D1	mib	réb	lá	sol#	lá	sol#	mib	réb	baixo	2Md	3Md	2md
E4_C4_D2	sol	fá#	si	réb	si	réb	sol	fá#	baixo-cima	2md	4Ja	2Ma
E4_C4_D3	mib	mi	sib	dó	sib	dó	mib	mi	cima	2ma	5dim <i>a</i>	2Ma
E4_C4_D4	sol <i>b</i>	dó	lá <i>b</i>	fá	lá <i>b</i>	fá	sol <i>b</i>	dó	cima-baixo	4aum <i>a</i>	3Md	3md
E4_Dis_D1	si	lá <i>b</i>	sol	mib	mib	réb	fá#	si	baixo	3md	2md	3Md
E4_Dis_D2	fá	si	sol <i>b</i>	lá#	lá#	dó	dó#	fá	baixo-cima	4dim <i>d</i>	4J <i>d</i>	3Ma
E4_Dis_D3	lá <i>b</i>	lá#	si	mi	mi	dó	mib	lá <i>b</i>	cima	2Ma	2ma	4aum <i>a</i>
E4_Dis_D4	ré	mi	mib	lá	lá	dó#	fá	ré	cima-baixo	2Ma	2md	5dim <i>d</i>

TABELA 23: DADOS DE JULGAMENTO DE E4.

DADO	Frases chefe			Frases subordinado		
E5_F_D1	ze	da	bo	bo	da	ze
E5_F_D2	fo	ja	ve	ve	ja	fo
E5_F_D3	si	ba	ve	ve	ba	si
E5_F_D4	vo	fa	de	de	fa	vo
E5_F_D5	be	ki	vo	vo	ki	be
E5_F_D6	zo	bi	ga	ga	bi	zo
E5_F_D7	fa	zu	de	de	zu	fa
E5_F_D8	ki	sa	fe	fe	sa	ki
E5_F_D9	bi	da	bo	bo	da	bi
E5_F_D10	fo	zu	ki	ki	zu	fo
E5_F_D11	va	3u	di	di	3u	va
E5_F_D12	pa	3ε	si	si	3ε	pa
E5_F_D13	ta	sɔ	zi	zi	sɔ	ta
E5_F_D14	po	zi	ga	ga	zi	po
E5_F_D15	be	pu	go	go	pu	be
E5_F_D16	ba	sε	bi	bi	sε	ba
E5_F_D17	ve	zu	po	po	zu	ve
E5_F_D18	ba	ku	ti	ti	ku	ba
E5_F_D19	je	vi	fi	fi	va	je
E5_F_D20	se	di	vo	vo	di	se
E5_F_D21	ta	vε	po	po	vε	ta
E5_F_D22	du	3i	ge	ge	3i	du
E5_F_D23	pe	fɔ	ta	ta	fɔ	pe
E5_F_D24	gu	vɔ	de	de	vɔ	gu
E5_F_D25	ka	sε	do	do	sε	ka
E5_F_D26	zi	bu	3a	3a	bu	zi
E5_F_D27	te	ku	zo	zo	ku	te
E5_F_D28	fi	3ɔ	pe	pe	3ɔ	fi
E5_F_D29	3o	ka	be	be	ka	3o
E5_F_D30	fe	dɔ	ka	ka	dɔ	fe

TABELA 24: DADOS DE FAMILIARIZAÇÃO DE E5.

Dados	Frases chefe				Frases subordinado			
E5_C1_D1	fe	ga	si	de	de	si	ga	fe
E5_C1_D2	jo	bi	fa	pe	pe	fa	bi	jo
E5_C1_D3	bu	vi	ba	ke	ke	ba	vi	bu
E5_C1_D4	fu	ze	vi	do	do	vi	ze	fu
E5_C2_D1	ji	fo	bε	za	za	fo	bε	ji
E5_C2_D2	pe	ba	ti	ko	ko	ba	ti	pe
E5_C2_D3	do	fe	zi	pa	pa	fe	zi	do
E5_C2_D4	ga	do	ti	fe	fe	do	ti	ga
E5_C3_D1	ta	do	bi	fe	fe	ta	do	bi
E5_C3_D2	vu	ge	fε	zo	zo	vu	ge	fε
E5_C3_D3	du	ti	gɔ	sa	sa	du	ti	gɔ
E5_C3_D4	bi	ju	za	ve	ve	bi	ju	za
E5_C4_D1	ba	di	do	fe	do	fe	ba	di
E5_C4_D2	vu	pe	ti	ca	ti	ca	vu	pe
E5_C4_D3	pu	go	ti	da	ti	da	pu	go
E5_C4_D4	fi	be	3u	ta	3u	ta	fi	be
E5_Dis_D1	ta	ku	bi	po	po	le	fu	ta
E5_Dis_D2	gu	vi	ba	ke	ke	ta	pi	gu
E5_Dis_D3	pe	ga	3i	to	to	fu	fε	pe
E5_Dis_D4	su	ta	ki	ve	ve	do	fε	su

TABELA 25: DADOS DE JULGAMENTO DE E5.

APÊNDICE B – FICHA INFORMATIVA DOS PARTICIPANTES

- 1- Idade: _____
- 2- Local de nascimento: _____
- 3- Escolarização:
 - Superior Incompleto
 - Superior Completo
- 4- Nível de conhecimento musical
 - Leigo
 - Amador pouco experiente (até 5 anos de estudo)
 - Amador experiente (acima de 5 anos de estudo)
 - Profissional

ANEXO – RELATÓRIO ESTATÍSTICO

Universidade Federal do Paraná - Departamento de Estatística

Laboratório de Estatística Aplicada

Relatório de Assessoria Estatística

Pesquisadora: Beatriz

Solicitação: Análise Estatística

Método: A análise estatística dos dados consistiu em realizar o teste não-paramétrico de Friedman (Ferreira, A.M. 2013), com a finalidade de avaliar se existem diferenças significativas entre as condições/tratamentos (C1, C2, C3, C4 e DIS). Optou-se pelo teste de Friedman pelo fato de amostra ter um tamanho reduzido e o experimento ter sido realizado em blocos (sujeitos). Entretanto, para cada combinação blocoXtratamento tínhamos 4 repetições. Para a realização do teste de Friedman necessita-se de apenas um resultado para cada combinação blocoXtratamento, assim utilizamos ou o valor mais frequente (moda) ou no caso de não existir um valor mais frequente, utilizamos a média das 4 repetições. A análise foi realizada usando o software R (R Development Core Team, 2013) e o Excel.

Resultados:

Experimento	Teste de		Interpretação
	Friedman	P-valor	
E1	3,58	0,46	Os julgamentos são estatisticamente iguais para as 5 "condições" (C1, C2, C3, C4 e DIS)
E2	9,82	0,04	Os julgamentos não são estatisticamente iguais para as 5 "condições" (C1, C2, C3, C4 e DIS)
E3	2,6	0,62	Os julgamentos são estatisticamente iguais para as 5 "condições" (C1, C2, C3, C4 e DIS)
E4	9,98	0,04	Os julgamentos não são estatisticamente iguais para as 5 "condições" (C1, C2, C3, C4 e DIS)
E5	33,04	<0,001	Os julgamentos não são estatisticamente iguais para as 5 "condições" (C1, C2, C3, C4 e DIS)

Como observado na tabela acima, apenas para os experimentos E2, E4 e E5 houve diferença significativa entre duas ou mais condições.

A fim de verificar entre quais condições há diferença, realizou-se um teste de comparação múltiplas de Friedman.

Condições	Experimentos					
	E2		E4		E5	
	Grupo	Soma dos ranks	Grupo	Soma dos ranks	Grupo	Soma dos ranks
C1	A	43	A	48	A	55,5
C2	A	43,5	B	34,5	B	46
C3	AB	34,5	B	30,5	CD	28
C4	AB	34,5	B	30	C	30,5
DIS	B	24,5	AB	37	D	20

Condições com mesma letra indicam que, ao nível de 5% de significância, não há diferença significativa entre as condições. Condições com letras diferentes indicam que, ao nível de 5% de significância as condições são estatisticamente diferentes.

Experimento E2:

Ao nível de 5% de significância a condição DIS difere de C1 e C2.

Experimento E4:

Ao nível de 5% de significância a condição C1 difere de C2, C3 e C4.

Experimento E5:

Ao nível de 5% de significância as condições C1 e C2 diferem das demais condições e as condições C4 e DIS diferem entre si.

Bibliografia:

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

FERREIRA, Armando Matheus. **Testes não paramétricos**. 2010. Apostila – Escola Superior Agrária.