UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS CLÁSSICAS E VERNÁCULAS

Cássio Augusto Alves de Andrade Santos

**A sincronização em dois ritmos da canção:
uma observação experimental acerca da fala cantada**

Versão corrigida

São Paulo
2012

CÁSSIO AUGUSTO ALVES DE ANDRADE SANTOS

cassioandradesantos@hotmail.com

**A sincronização em dois ritmos da canção:
uma observação experimental acerca da fala cantada**

Dissertação apresentada ao programa de
pós-graduação em Filologia e Língua Portuguesa
do Departamento de Letras Clássicas e
Vernáculas da Faculdade de Filosofia, Letras e
Ciências Humanas da Universidade de São Paulo
para a obtenção do título de Mestre em Letras.

Versão Corrigida

Área de concentração: Filologia e Língua Portuguesa

Orientadora: Prof. Dr. Beatriz Raposo de Medeiros

São Paulo

2012

Aos meus pais e irmãos
Ao vô Manuel, pelos papagaios e enciclopédias

Agradecimentos

À minha mãe e meu pai, Cida e Ivan, simplesmente por serem minha mãe e meu pai e tudo aquilo que isto implica.

Aos meus irmãos, Filipe e Fábio, companheiros de sempre.

À Bia, por me orientar e pela liberdade que me deu.

À Rute, pelas leituras, idéias, apoios, palavras de motivação e pela amizade.

Aos cantores e cantoras Larissa Smid, Harry Castro, Cristiane Martinez, Cristiano Veder, Thalita Bertolini, Camila de Sá, Beth Aquino e Caíque Guirao.

Ao pessoal do Estúdio Multimeios do Centro de Computação Eletrônica da USP: Flávia Pires Ribeiro, Edson Ramos e os monitores João Caputi, Ana Paula Fiorotto e Camila Luppi.

À prof. Dr. Eliane G. Fischer, da área de alemão do Departamento de Letras Modernas, pela ocasião da iniciação científica.

Ao pessoal da Biblioteca do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da USP: Sandra, Leandro, Conceição, Lucinha, Rosa, Marta, Fran, e Célia; e ao Freitas das Salas Pró-Aluno da FFLCH, pelas ocasiões de estágios na graduação.

Ao amigo e músico Carlos Ferreira, pela ajuda com os estudos para a prova de proficiência, com as partituras e conhecimentos musicais.

Ao Antônio Pessotti, da UNICAMP, pelas ajudas ao longo deste trabalho.

Ao pessoal do Centro Cultural Esplanada, pelo apoio desde a 7º série do ensino fundamental.

Aos professores.

Aos professores do CASD Vestibulares da turma de 2003.

Ao pessoal do Grupo de Estudos em Fonética da USP.

À Lili e à Maria Joana.

À todos que colaboraram para o meu desenvolvimento intelectual.

Resumo

SANTOS, C. A. A. A. **A sincronização em dois ritmos da canção: uma observação experimental acerca da fala cantada.** Dissertação de Mestrado – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2012.

Este trabalho apresenta um estudo acerca da fala cantada com o português brasileiro inserido no contexto da canção popular. Utilizando como base o processo de sincronização entre os pares de quatro duplas de cantores, elaboramos um experimento fatorial com dois fatores, denominados Ritmo da Canção e Ponto de Observação de Sincronização. Com o fator Ritmo da Canção, buscamos observar se a fala cantada inserida no contexto sincopado, no caso o da bossa-nova, é mais difícil de ser realizada quando comparada a mesma inserida no contexto não sincopado, aqui o do rock. Com o fator Ponto de Observação de Sincronização buscamos observar se os sujeitos têm a mesma capacidade de sincronização em inícios e em finais de frases. Ambos os cálculos foram realizados através das medidas das *lags*, diferença temporal entre os Pontos de Observação de Sincronização dos cantores de uma mesma dupla, ou seja, ora separamos as *lags* por canções (fator Ritmo da Canção), ora separamos por ponto (fator Ponto de Observação de Sincronização). Em paralelo, e de forma a complementar o estudo fonético-acústico, observamos o desenvolvimento métrico-acentual das canções, comparando os acentos lingüísticos e os musicais das canções utilizadas. O conjunto de resultados obtidos por meio de cálculos estatísticos e as observações métrico-acentuais indicam que os cantores apresentam maior facilidade em realizar a tarefa pedida – cantar o mais sincrônico possível – no contexto no qual o ritmo de acompanhamento das canções não era sincopado, assim como indica maior facilidade em sincronizar nos finais de sentenças.

Palavras-chave: fonética, ritmo, sincronização, música, síncopa.

Abstract

SANTOS, C. A. A. A. The Entrainment in two rhythms of the song: an experimental observation about sung speech. Dissertation (Master) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2012.

This paper presents a study about the Brazilian Portuguese placed in the context of the song. Using as a base the entrainment process between pairs of four pairs of singers, we designed a factorial experiment with two factors, called Rhythm of the Song and Entrainment Observation Point. With the factor Rhythm of the Song, we observe whether the speech into the context of singing bossa-nova is more difficult to achieve when compared to the same inserted in the context of rock. With the factor Entrainment Observation Point we seek to observe whether individuals have the same ability to sync at the beginning and end of sentences. Both calculations were performed by measuring the lag, time difference between entrainment observation points of the singers of the same pair, i.e., separated the lag by songs (factor Rhythm of the Song), and separated by Entrainment Observation Point (factor Entrainment Observation Point). In parallel, and to complement the acoustic-phonetic study, we observed the development of accentual metric-songs comparing the linguistic and musical accents of the songs used. The set of results obtained by statistical calculation and accentual-metric observations indicates that the singers have it easier to perform the requested task - to sing as synchronous as possible - in the context in which the rhythm accompaniment of songs was not syncopated, so ease as indicated to sync at the end of sentences.

Keywords: phonetics, rhythm, entrainment, music, syncopation.

LISTA DE FIGURAS

Legenda	Página
Figura 1: Fatores experimentais do estudo piloto e seus níveis	14
Figura 2: Gráfico do desvio padrão na condição <i>com metrônomo Gita</i>	15
Figura 3: Gráfico do desvio padrão na condição <i>com cancionista Gita</i>	16
Figura 4: Gráfico do desvio padrão na condição <i>com metrônomo Corcovado</i>	16
Figura 5: Gráfico do desvio padrão médio na condição <i>com cancionista Corcovado</i>	16
Figura 6: Pauta Gestual da sílaba /du/	26
Figura 7: Sequência de quatro compassos com métrica binária	33
Figura 8: Sequência de três compassos com métrica ternária.	33
Figura 9: Sequência de dois compassos com métrica quaternária.	33
Figura 10: Trecho de valsa da opereta O Morcego, de J. Strauss.	33
Figura 11: Partitura para bateria da batida de acompanhamento de uma valsa.	34
Figura 12: Partitura para bateria da batida de acompanhamento de um rock.	34
Figura 13: Exemplo 1 de síncope (Priolli, 1987)	35
Figura 14: Exemplo 2 de síncope (Priolli, 1987)	35
Figura 15: Ritmo de acompanhamento da bossa-nova (Adolfo, 1994)	35
Figura 16: Exemplo 1 de frase musical.	44
Figura 17: Exemplo 2 de frase musical (Chediak, 2004)	45
Figura 18: Exemplo 3 de frase musical (Chediak, 2004)	45
Figura 19: Exemplo 4 de frase musical (Chediak, 2004)	46
Figura 20: Exemplo de segmentação da canção Gita (Chediak, 2004).	47
Figura 21: Exemplo de segmentação da canção Águas de Março (Chediak, 2004).	47
Figura 22: Exemplo de segmentação da canção Aluga-se. As 4 primeiras notas são a introdução.	48
Figura 23: Exemplo de segmentação da canção Corcovado.	48
Figura 24: Ponto Inicial e Ponto Final da primeira sentença da canção Aluga-se, pela cantora LS.	49
Figura 25: Pontos Inicial e Final da cantora LS comparados aos pontos Inicial e Final de HC.	50
Figura 26: Árvores métricas das palavras <i>tatu</i> , <i>livro</i> e <i>lingüística</i>	55
Figura 27: Grades métricas das palavras <i>tatu</i> , <i>livro</i> e <i>lingüística</i>	55
Figura 28: Grade métrica da sentença <i>Às vezes você me pergunta</i>	56
Figura 29: Grade métrica simplificada da sentença <i>às vezes você me pergunta</i>	56
Figura 30: Representação métrica da primeira frase musical de Gita.	56
Figura 31: Representação métrica da primeira sentença lingüística de Gita e	

de sua respectiva frase musical.	57
Figura 32: Representação métrica da primeira sentença lingüística de Corcovado e de sua respectiva frase musical.	57
Figura 33: Distribuição dos dados gerais. A abscissa indica os blocos dos valores das <i>lags</i> em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de <i>lags</i> nos determinados blocos.	78
Figura 34: Exemplo de <i>lag</i> em torno da média. A faixa rosada indica justamente a <i>lag</i> , de 48 milissegundos.	79
Figura 35: Exemplo de sincronia perfeita, ou seja, <i>lag</i> igual a zero, no <i>onset</i> de <i>viu</i> .	79
Figura 36: Exemplo de <i>outlier</i> . As setas indicam os pontos iniciais. A diferença entre eles é de 1087 milissegundos e representa uma perda de deixa por parte de um dos cantores.	80
Figura 37: Gráfico descritivo do fator Ritmo da Canção	81
Figura 38: Distribuição dos dados do nível RNS. A abscissa indica os blocos dos valores das <i>lags</i> em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de <i>lags</i> nos determinados blocos.	82
Figura 39: Distribuição dos dados do nível RS. A abscissa indica os blocos dos valores das <i>lags</i> em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de <i>lags</i> nos determinados blocos.	83
Figura 40: Gráfico descritivo do fator Ponto de Observação de Sincronização	85
Figura 41: distribuição dos dados do nível Ponto inicial. A abscissa indica os blocos dos valores das <i>lags</i> em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de <i>lags</i> nos determinados blocos.	86
Figura 42: Distribuição dos dados do nível Ponto Final. A abscissa indica os blocos dos valores das <i>lags</i> em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de <i>lags</i> nos determinados blocos.	87
Figura 43: Gráfico da existência de interação entre os fatores	89
Figura 44: Resultados do experimento	91

LISTA DE TABELAS

Legenda	Página
Tabela 1: Resultados do estudo piloto (Santos; Medeiros, 2010)	17
Tabela 2: Delineamento do Experimento. Os fatores e seus níveis.	39
Tabela 3: Duplas e respectivos sujeitos	40
Tabela 4: Quantidade de sentenças por canção	76
Tabela 5: Estatística dos dados gerais	77
Tabela 6: Tabela de Percentiis dos dados gerais	77
Tabela 7: Estatística descritiva do fator Ritmo da canção	80
Tabela 8: Tabela de percentiis dos dados do fator Ritmo da canção	84
Tabela 9: Estatística descritiva do fator Ponto de Sincronização	84
Tabela 10: Tabela de percentiis dos dados do fator Ponto de sincronização	88
Tabela 11: Quantidade de sincronias perfeitas	88
Tabela 12: Combinação entre os níveis dos dois fatores	88

Sumário

1. Introdução	11
1.1 Delimitação do objeto de estudo	13
1.2 Hipóteses	13
1.3 Estudo piloto	14
1.3.1 Nota sobre a utilização do GIPC	18
2. Ritmo e conceitos circundantes	19
2.1 Conceito clássico de ritmo lingüístico	20
2.1.1 O ritmo do português brasileiro	21
2.2 Novos conceitos sobre ritmo lingüístico	22
2.3 Conceito de oscilador e sua relação com o processo de sincronização	24
2.4 Fala como movimento	25
2.5 Planejamento cerebral dos movimentos da fala	27
2.6 Acentos e Ritmos	29
2.6.1 Caracterização do acento em português brasileiro	29
2.6.2 Acento e ritmo na música	29
2.6.2.1 Ritmo na música: aspectos objetivos e subjetivos	31
2.6.2.2 Relações entre metro e ritmo e o conceito de síncope	33
3. Metodologia	37
3.1 <i>Corpus</i>	39
3.2. Sujeitos	40
3.3 Preparação para a coleta	41
3.4 Coleta	42
3.5 Segmentação e rotulagem do sinal acústico	43
3.5.1 A frase musical	43
3.5.2 Segmentação das canções	46
3.5.3 Rotulagem e definição de <i>lag</i>	48
3.5.3.1 Rotulagem automatizada: o <i>Dynamic Time Warping</i>	50
3.6 Padronização dos valores das <i>lags</i>	51

4. Observações dos Acentos das Canções	53
Introdução	54
4.1 Observações dos Acentos	58
4.1.1 Aluga-Se	58
4.1.2 Gita	61
4.1.3 Corcovado	65
4.1.4 Águas de Março	66
4.2 Discussão	72
5. Estatística Descritiva	75
5.1 Descrição geral dos dados	76
5.2 Descrição do fator Ritmo da canção	80
5.3 Descrição do fator Ponto de Observação de Sincronização	84
5.4 Sincronias perfeitas	88
5.5 Observações iniciais acerca da interação entre os fatores	88
6. Conclusão	90
7. Referências Bibliográficas	93
8. Apêndice 1 – Segmentação das canções	99
9. Apêndice 2 – Grades métricas das frases musicais	102

1 - Introdução

A palavra ritmo tem origem na Grécia, oriunda da raiz *rheo*, que significava fluir. A idéia de movimento circunda o conceito de ritmo e está presente também em várias outras palavras, como em valsa, do alemão *walzer*, que em português pode ser traduzida por *rolante*. Esta última, em grego, pode ser traduzida por *trochaío*, que por sua vez chegou ao português como troqueu. O termo inglês *rock and roll* também tem sua origem na idéia de movimento e pode ser traduzido para português por *balançar e rolar*. A palavra marcha não fica de fora, sendo originada do latim *marcare*, palavra que descrevia uma passada marcada. Segundo o Dicionário Cravo Albin da Música Popular Brasileira¹ o termo samba também se originou de um movimento: o de umbigada, traduzido por *semba* em quimbundo, língua falada em Angola.

Tradicionalmente, o conceito de ritmo é quase que fadado a estar presente somente no contexto da música e da dança. Mais raramente, ele aparece em outros contextos nos quais sons e/ou movimentos atrelados a uma linha temporal sejam matérias primas, como a fonética e a fonologia, principais áreas da lingüística no estudo sobre o ritmo das línguas. Na música, área em que o conceito de ritmo é mais consolidado quando comparado ao conceito lingüístico, Cooper e Meyer (1960) o definem como o modo ao qual uma ou mais batidas fracas são agrupadas em relação a uma batida forte, definição que retomaremos no item 2.6.2 *Acento e ritmo na música*.

Na lingüística, o ritmo tem sido objeto de estudos desde a primeira metade do século passado, mas as observações e definições a seu respeito sofrem constantes transformações e acarretam em definições, por vezes, distintas. A definição clássica básica de ritmo lingüístico se baseia na idéia de isocronia, dividindo as línguas do mundo em dois grupos: línguas de Ritmo Acentual a línguas de Ritmo Silábico. Como veremos adiante, a isocronia nunca foi encontrada e novas definições sobre ritmo surgiram. Fred Cummins, da University College Dublin, parte de um ponto de vista dinâmico e ubíquo do ritmo (segundo o autor, presente em toda a natureza) para explicar o ritmo lingüístico.

Assim, tomamos como objeto de estudo a fala cantada inserida no contexto da canção, levando a diante as idéias de Cummins e inserindo a fala em outro contexto, tão dinâmico quanto o da fala falada²: o contexto musical.

Baseando-nos no processo de sincronização, apresentaremos neste trabalho uma perspectiva do ritmo de canções em português brasileiro. No *Capítulo 1 – Introdução*

¹ (<http://www.dicionariompb.com.br/samba/dados-artisticos>), acessado em 02/07/2011.

² Aqui se contrapõe *fala falada* de *fala cantada* (Medeiros, 2002)

apresentamos as principais idéias do trabalho, hipóteses e estudo piloto. O *Capítulo 2 – Ritmo e conceitos circundantes* retoma a literatura sobre sincronização, ritmo, acento e síncopa, com os quais buscamos apresentar um ponto de vista dinâmico da *fala cantada* (Raposo de Medeiros, 2002), que norteou esta pesquisa. O *Capítulo 3 – Metodologia* apresenta o desenvolvimento de um experimento que elaboramos a fim de observar o comportamento da fala cantada em dois contextos rítmicos distintos. No *Capítulo 4 – Observações dos Acentos das Canções* observamos os acentos lingüísticos e musicais das sentenças das canções utilizadas, buscando relacionar a presença da síncopa com a coincidência acentual entre as estruturas musical e lingüística. No *Capítulo 5 – Estatística Descritiva* apresentamos os resultados dos dados coletados, seguido do último capítulo, no qual apresentamos as conclusões finais deste trabalho, *Capítulo 6 – Conclusões*.

1.1 Delimitação do objeto de estudo

No contexto musical o ritmo pode se manifestar de diversas maneiras, dependendo do gênero: rock, forró, samba, jazz, reggae, *etc.*. Assim, decidimos por observar a sincronização entre dois cantores de uma mesma dupla (utilizamos quatro duplas) que realizavam a língua portuguesa cantada em dois desses gêneros, tanto quanto diferentes entre si: o rock e a bossa-nova, sendo a principal diferença rítmica entre esses dois gêneros o fato de o ritmo de acompanhamento da bossa-nova se estruturar com base na síncopa.

Observamos também este processo em dois diferentes pontos das sentenças das canções utilizadas no *corpus*: Ponto Inicial e Ponto Final, elaborando um experimento (ver capítulo 3) com dois fatores, denominados *Ritmo* (com os níveis *Ritmo Sincopado* (RS) e *Ritmo Não Sincopado* (RNS) e *Ponto Observação de Sincronização* (com os níveis *Ponto Inicial* (PI) e *Ponto Final* (PF)).

1.2 Hipóteses

Colocamos neste trabalho duas hipóteses principais. A primeira delas defende a idéia de que seja mais fácil para os sujeitos sincronizarem suas falas em canções cujo ritmo de acompanhamento não seja sincopado, pois acreditamos que a síncopa, conflito métrico-rítmico que veremos a seguir, tenha direta influência na realização da tarefa, dificultando-a.

A segunda hipótese, baseada em estudos de Cummins (2002, 2003, 2006), defende a idéia de que, independentemente do ritmo, seja mais fácil para os sujeitos se sincronizarem nos Pontos Finais das sentenças em comparação aos Pontos Iniciais.

Estas hipóteses germinaram em searas diferentes: os fatores do experimento realizado. A partir da combinação entre os fatores pudemos fazer emergir uma terceira hipótese, aquela na qual acreditamos que seja mais fácil sincronizar a fala em pontos finais de sentenças não sincopadas.

1.3 Estudo piloto

Elaboramos um experimento piloto (Santos; Medeiros, 2010) que nos permitisse uma observação prévia acerca do processo de sincronização da fala cantada. Observamos esse processo no canto de duas diferentes canções, Gita e Corcovado, realizado por um sujeito que cantava ora acompanhado de um metrônomo ajustado de acordo com o andamento da canção, se guiando pela partitura das canções (Chediak 1994b, Chediak, 2004); ora acompanhado de gravação original das referidas canções. Assim, tínhamos dois fatores experimentais: *com metrônomo* e *com cancionista*, sendo esses divididos em dois níveis, cada uma referente a uma canção, do seguinte modo:

Figura 1: Fatores experimentais do estudo piloto e seus níveis.

Cada canção foi cantada seis vezes, sendo três em cada fator.

Segmentamos as canções em unidades denominadas de Grupo *Inter Perceptual Center*, ou GIPC (Barbosa, 1994, 2000), obtendo 60 unidades de Gita e 60 de Corcovado, que foram numerados em seqüência.

O GIPC é uma unidade que se inicia no *onset* de determinada vogal e se estende até o *onset* da vogal seguinte. Utilizamos tal unidade, pois, segundo Morton, Marcus e Frankish (1976), seria no *onset* das vogais que se encontra o *perceptual center*, ponto em que o ouvinte se ancoraria para perceber a regularidade temporal da fala. No sintagma “um cantinho”, por exemplo (trecho da canção Corcovado), tem-se o primeiro GIPC entre os *onsets* de /u/ e o de /a/, ou seja, tem-se um GIPC formado pelos fones /um k/. O /a/ pertence já ao GIPC seguinte, /ant/.

Hipotetizamos que a condição *com cancionista Corcovado* propiciaria maior variação na duração dos GIPC, e que a menor variação seria na condição *com metrônomo Gita*, pois acreditamos que o metrônomo tende a tornar a interpretação mais “quadrada”, mais regular, assim como a canção não-sincopada. Por outro lado, acreditamos que a canção sincopada tenderia a acrescentar dificuldade à realização da tarefa, independentemente da condição, justamente devido à presença da síncopa, o que acarretaria maior variação nas durações.

Calculamos, então, a média das durações das três realizações de cada GIPC e em seguida, também o desvio padrão dessas médias. No valor do desvio padrão observamos o quanto variou a duração das unidades em cada uma das condições, ou seja, o quanto o sujeito variou na duração de cada segmento de sua fala cantada.

Os gráficos abaixo trazem os valores do desvio padrão de cada GIPC em cada uma das quatro condições experimentais:

Figura 2: Gráfico do desvio padrão na condição *com metrônomo Gita*

Figura 3: Gráfico do desvio padrão na condição *com cancionista Gita*

Figura 4: Gráfico do desvio padrão na condição *com metrônomo Corcovado*

Figura 5: Gráfico do desvio padrão médio na condição *com cancionista Corcovado*

Em seguida calculamos o desvio padrão médio de cada condição, obtendo, portanto, quatro valores:

Tabela 1: Resultados do estudo piloto (Santos; Medeiros, 2010)

Condição	Desvio padrão médio
Com metrônomo-Gita	15,58 ms
Com cancionista- Gita	19,60 ms
Com metrônomo-Corcovado	50,96 ms
Com cancionista-Corcovado	34,76 ms

Observando-se os gráficos, notam-se menores valores de duração na condição *com metrônomo Gita*, como esperávamos, sendo o valor de desvio padrão médio igual a 15,58 ms, como mostra a tabela. Em seguida, observa-se o valor de 19,60 ms para o desvio padrão médio na condição *com cancionista Gita*, confirmando assim menores valores nas condições com a canção não sincopada.

Nas condições com a canção Corcovado, no entanto, pode-se perceber que, ao contrário do que aconteceu na condição com a canção Gita, obtiveram-se maiores valores de duração no canto com metrônomo, e não no canto com a gravação.

Concluimos que o uso de partitura e metrônomo dificultou o exercício de sincronização na canção sincopada. Acreditamos que isso se deva a dois fatores:

1) As batidas do metrônomo são sempre iguais em intensidade, frequência e duração, o que o torna um oscilador-guia sem referências a não ser por marcar o andamento e o pulso. Em outras palavras, o metrônomo não marca, por exemplo, a métrica, essencial para a realização da síncopa.

2) A gravação da canção é um oscilador guia mais eficiente, pois a fala cantada facilita o exercício de entrosamento, já que o sujeito se utiliza antes de tudo da voz do intérprete para se guiar. A presença ainda do violão, que se ouve claramente na gravação, marca a batida tão característica da bossa-nova, o que também facilita a tarefa de entrosamento, já que a batida tem contida em si o ritmo do acompanhamento, que conduz a canção.

1.3.1 Nota sobre a utilização do GIPC

No experimento realizado no presente trabalho, abandonamos a noção de GIPC. O observado por Massini-Cagliari (1992) é o principal motivo que nos levou a tomar tal decisão: em amplo estudo fonético-acústico do acento em português, a autora notou que este é caracterizado, também, por uma maior duração da sílaba como um todo, e não somente da vogal, ou seja, as consoantes também apresentam maiores durações quando em sílabas tônicas, sejam elas sílabas CV, VC, VCC, CCV³, e assim, acreditamos que segmentar uma sílaba CV, por exemplo, em dois GIPC distintos pode acabar por descaracterizar o acento lexical, já que a consoante da sílaba seria pertencente ao GIPC anterior.

³ C= Consoante
V= vogal

2 - Ritmo e conceitos circundantes

2.1 Conceito clássico de ritmo lingüístico

Há décadas tem-se buscado definir e entender o ritmo das línguas. Emprestado da música, este termo define, grosso modo, certa regularidade temporal da fala.

A classificação rítmica mais divulgada foi apresentada por Pike (1945), que dividiu as línguas em *stress-timed rhythm* (ritmo acentual) e *syllable-timed rhythm* (ritmo silábico) a partir de estudo do inglês americano e do espanhol falado nos Estados Unidos, respectivamente. Inspirado na classificação de Lloyd-James (1940, *apud* Barbosa, 2006 ⁴), que separava as línguas em *machine-gun rhythm* e *Morse-code rhythm*, Pike afirma, a partir do conceito de isocronia, que em línguas de ritmo acentual (*Morse-code rhythm*, na classificação de Lloyd-James) as regularidades ocorrem entre sílabas acentuadas, ou seja, o espaço temporal entre as sílabas acentuadas seria o mesmo (isocronismo acentual), enquanto em línguas de ritmo silábico (*machine-gun*, segundo Lloyd-James) a regularidade ocorre entre as sílabas, e assim, todas as sílabas teriam a mesma duração (isocronismo silábico). O próprio autor, no entanto, já afirma que o inglês pode apresentar características de ritmo silábico em determinados contextos, como em alguns estilos de elocução e na fala cantada.

A partir do apresentado por Pike, Abercrombie (1967) classificou o inglês, o russo e o árabe como línguas de ritmo acentual, classificando como de ritmo silábico o francês, o telugu e o iorubá, afirmando que todas as línguas do mundo são faladas com um dos dois tipos de ritmo.

Major (1981), no entanto, afirma serem três os tipos de ritmo lingüístico: além dos dois já citados, o ritmo moraico, tendo como exemplo a língua japonesa, seria aquele no qual as moras teriam igual duração.

Dauer (1983) comparou dados do inglês (ritmo acentual) com dados do grego, do italiano e do espanhol (ritmo silábico) e afirmou que em todas essas línguas a duração entre os acentos, ou seja, o espaço temporal entre um acento e outro, depende da quantidade de sílabas presentes entre eles. A autora concluiu que as diferenças rítmicas ocorrem muito mais devido a diferenças na estrutura da língua do que na regularidade entre os intervalos, colocando em xeque, portanto, a presença da isocronia na fala e as conseqüentes classificações rítmicas consagradas até o momento.

⁴ LLOYD-JAMES, A. *Speech signals in telephony*. Londres: Sir Issac Pitman and Sons, 1940.

Ainda pautados na noção de isocronia, Cagliari e Abaurre (1986) defendem a existência desse fenômeno em um nível psicológico. Tal isocronia psicológica seria baseada em interpretações do ouvido de acordo com as expectativas do ouvinte em relação ao tempo.

Barbosa (2006) cita diversos trabalhos a respeito do ritmo lingüístico (Lea, 1974; Lehisté, 1977; Pointon, 1980; Wenk; Wioland, 1982; Nooteboom, 1991; *apud* Barbosa, 2006⁵) para concluir que o isocronismo absoluto nunca foi encontrado em dados de produção da fala. Para o autor, o motivo dessa busca infrutífera se dá devido ao fato de o ritmo lingüístico ser “construído na inter-relação entre regularidade e estruturação” (p. 61), ou seja, muitas vezes a regularidade da fala é quebrada em detrimento da estrutura da língua.

Cummins defende que na fala não há isocronia e que o ritmo está presente na fala por estar presente em todas as coisas: *Rhythm is a ubiquitous phenomenon found throughout nature, and central to many human activities* (2009b, p. 3).

2.1.1 O ritmo do português brasileiro

Tentativas de classificar o português brasileiro quanto à sua ritmicidade surgiram na década de 1980. Primeiramente, Cagliari (1981) classificou como de ritmo acentual o português falado no Brasil. Em estudo posterior, Cagliari e Abaurre (1986) observaram que alguns falantes possuíam ritmo acentual, enquanto outros possuíam ritmo silábico, mas em ambos os casos, pôde-se observar certa flutuação rítmica. Major (1981, 1985, *apud* Massini-Cagliari, 1992) e Moraes e Leite (1989, *apud* Massini-Cagliari, 1992⁶) também classificaram o português brasileiro como de ritmo acentual. Massini-Cagliari (1992), após minucioso estudo baseado na noção de isocronia, concluiu que o português brasileiro pode ser classificado tanto como língua de ritmo acentual como de ritmo silábico. Por outro lado, a autora notou também que há evidências contra a classificação do português em um desses dois tipos, o que a levou a concluir que a dicotomia acentual/silábico é muito vaga.

⁵ LEA, W. A. *Prosodic aids to speech recognition: IV. A general Strategy for prosodically-guided speech understanding*. St. Paul, Minnessota, EUA, 1974.

LEHISTE, I. Isocrony reconsidered. *Journal of Phonetics*, v. 5, p. 253-263, 1977.

POINTON, G. E. Is Spanish really syllable-timing? *Journal of Phonetics*, v. 8, p. 293-304, 1980.

WENK, B. L.; WIOLAND, F. Is French really syllable-timed? *Journal of Phonetics*, v. 10, p. 193-216, 1982.

NOOTBOOM, S. G. Some observations on the temporal organisation and rhythm of speech. In: *Proceedings of the XIIth International Congress of Phonetic Science*. Aix-en-Provence, França: [sn], 1991. V.1, p. 228-237.

⁶ MAJOR, R. C. Stress and rhythm in Brazilian Portuguese. *Language*, 61(2): 259-282. 1985

MORAES, J. A.; LEITE, Y. Ritmo e velocidade da fala na estratégia do discurso: uma proposta de trabalho. Comunicação apresentada no III Seminário do Projeto “Gramática do Português Falado”, Águas de Lindóia, 1989

2.2 Novos conceitos sobre ritmo lingüístico

Partindo do ponto de vista de que de falar é também movimento, Cummins (2009b) apresenta um novo conceito de ritmo aplicado às línguas, e mais do que isso, aplicado não só às línguas, mas também à música, à dança, ou seja, a qualquer fenômeno ou processo nos quais movimentos, de um modo geral, são observados a partir de um ponto de vista temporal, dinâmico. Para esse autor, ritmo é *an affordance for the entrainment of movement*, uma idéia que necessita de alguns esclarecimentos conceituais para ser entendida.

O conceito de *affordance* (Gibson, 1979; Chemero, 2003; *apud* Cummins 2009b⁷) é oriundo da psicologia ecológica e define “*relations between the abilities of organisms and features of the environment*” (p. 5). Como exemplo, o autor traz a relação de “agarrabilidade” entre um copo e uma pessoa que busca agarrá-lo: essa relação só existe se o copo for “agarrável” (se tiver uma alça ou se não estiver muito quente, por exemplo) e se a pessoa for capaz de agarrá-lo (uma pessoa sem mãos não conseguiria realizar tal ação ou uma pessoa sem luvas não poderia agarrar um copo a determinada temperatura).

Ora, aplicado ao estudo do ritmo das línguas, *affordance* define a relação entre as habilidades de movimento de fala de um sujeito e a estrutura temporal de sua língua (no texto: sinal específico) (Cummins, 2009b).

Entrainment é um conceito trazido da física e das ciências biológicas que lidam com o tempo. É definido como *o processo* no qual dois osciladores interagem entre si de modo a se ajustarem e, eventualmente, entrarem em fase ou período⁸ (Clayton *et. al.*, 2004). Utilizamos a palavra *sincronização* para traduzir este conceito. A *sincronia* perfeita (com fase igual a zero, chamada de *zero phase*) seria uma das conseqüências do *entrainment* e assim distinguimos *sincronização* como um processo; processo esse que pode, como conseqüência, implicar em uma *sincronia* perfeita ou não entre os osciladores envolvidos. Portanto, a sentença *rhythm as an affordance for the entrainment of movement* pode ser traduzida para o português por *ritmo como um possibilitador da sincronização de movimento*.

⁷ GIBSON, J. J.: *The Ecological Approach to Visual Perception*. (Houghton Mifflin, Boston, 1979).

CHEMERO, A. An outline of a theory of affordances. *Ecological Psychology*, **15**(2):181–195, 2003

⁸ [in.train.m'nt] se difere de [en.treyn.ment], ambos grafados da mesma forma e derivados da palavra *train* (trem). O primeiro se refere ao processo acima referido. O segundo é um conceito usado na físico-química e na meteorologia e se refere ao arrastamento de certos gases por uma nuvem de vapor ou pelo arrastamento de um fluído por outro. *Arrastamento* ou *acarretamento* são traduções correntes para o português dessa segunda acepção e por isso usamos a palavra *sincronização*. (Dictionary.com) (<http://dictionary.reference.com/browse/entrainment>, acessado em 10/04/2011)

O processo de sincronização tem sido usado como ferramenta de investigações científicas quando se busca entender alguns aspectos em torno à organização, produção e percepção de estruturas temporais. Trabalhos no âmbito da fisiologia e da cronobiologia⁹, por exemplo, têm observado processos de sincronização de seres vivos com o ciclo temporal do meio em que habitam, ou seja, com fatores externos à própria fisiologia (como o ciclo do Sol ou o ciclo das estações do ano), levando justamente em consideração o fato de ritmos biológicos e fisiológicos serem essenciais para a vida (Aschoff 1979, Strogatz; Stewart 1993, Glass 1996, *apud* Clayton *et al.*¹⁰).

Por outro lado, estudos têm apontado que seres humanos e animais exibem também padrões de organização temporal baseados na sincronização com a própria fisiologia e não somente com padrões externos, processo este chamado de auto-sincronização¹¹ (Port *et. al.*, 1996). Esses padrões internos, geralmente oriundos do ciclo temporal da respiração ou dos batimentos cardíacos, se refletem, por exemplo, nos movimentos dos braços e das pernas ao andar ou ao correr.

Também em pesquisas sobre a fala, tem-se observado como se comporta o processo de sincronização. Port *et. al.* (1996) relatam os resultados de algumas experiências que testam a relação entre inícios de frases repetidas e a posição das sílabas tônicas destas frases em diferentes andamentos. Eles relatam que seres humanos realizam muita ou pouca sincronia na fala, dependendo da oportunidade para fazer a tarefa e sugerem ainda que estudos sobre auto-sincronização podem ser profundamente reveladores sobre a maneira em que o tempo é tratado no sistema nervoso dos animais, indo muito além da coordenação dos próprios membros, como braços e pernas ao andar. Mayberry *et al.* (1998¹²), por exemplo, observaram que pessoas que sofrem de gagueira tendem a parar o movimento das mãos ao gaguejar, o que exemplifica o processo de auto-sincronização entre a fala e um outro oscilador (conceito logo abaixo esclarecido), neste caso, as mãos.

⁹ Ramo da biologia dedicado ao estudo de processos fisiológicos dos seres vivos relacionados à regularidade temporal.

¹⁰ ASCHOFF, J. Circadian rhythms: General features and endocrinological aspects. *In*: Krieger (Ed.), *Endocrine rhythms*. New York: Raven Press. 1979

STROGATZ, S. H.; STEWART, I. Coupled oscillators and biological synchronization. *Scientific American*, 269(6), 68-75. 1993

GLASS, L. Synchronization and rhythmic processes in physiology. *Nature*, 410 (6825), 277-284. 1996

¹¹ Na literatura internacional, *self-entrainment*.

¹² MAYBERRY, R.I.; JAKES, J.; DEDE, G. (1998). What stuttering reveals about the development of the gesture- speech relationship. *In* J.M. IVERSON; GOLDIN-MEADOW, S. (Eds.), *The Nature and Function of Gestures in Children's Communication. New Directions for Child Development*, no.79. San Francisco: Jossey-Bass.

No contexto da fala, ainda, nota-se o processo de auto-sincronização quando o falante coordena os movimentos dos músculos e órgãos envolvidos na tarefa de falar, como o arredondamento dos lábios coordenado ao movimento da língua na realização da sílaba /du/, por exemplo.

Cummins (2002, 2003, 2006, 2009a, 2009b) realizou diversos trabalhos utilizando o processo de sincronização na fala como ferramenta para estudar o ritmo como fenômeno ubíquo. Em experimento (Cummins, 2002), o autor mediu a sincronização entre dois falantes instruídos a ler um texto juntos e comparou aos resultados obtidos quando um desses falantes lia o mesmo texto acompanhando uma gravação realizada pelo parceiro da primeira tarefa. O autor concluiu que a segunda tarefa é mais difícil de ser realizada (medianas de 30 ms na primeira tarefa contra 56 ms na segunda). Cummins (2003) observou também a sincronização em três diferentes pontos das sentenças de um texto – inicial, medial e final – e notou menor sincronia no ponto inicial, não observando diferenças entre a sincronia nos pontos finais comparados aos mediais. Observou-se também que a prática não ajuda a melhorar a realização da tarefa, ou seja, que os falantes têm uma grande capacidade de sincronização, mesmo sem prática. Nesse trabalho, o autor notou, ainda, a importância do contato visual entre os falantes, afirmando que a tarefa é mais fácil de ser realizada com a manutenção deste tipo de contato e que os sujeitos fazem, quase que exclusivamente, uso da visão periférica ao invés da visão foveal (a visão que projeta a imagem do objeto focalizado). Finalmente, ainda concluindo o observado no estudo de 2002, Cummins (2003) notou que a boa realização da tarefa de sincronização depende da colaboração entre os falantes, ou seja, na fala sincronizada não há um líder, os dois falantes interagem de forma mútua.

2.3 Conceito de oscilador e sua relação com o processo de sincronização

O conceito de oscilador é trazido da física e se refere a qualquer sistema que apresente movimento harmônico de oscilação, que, dentre outras palavras, pode ser descrito como o movimento de vai-e-vem constante em torno de uma posição central, como o pêndulo de um relógio ou o movimento de translação da Terra.

Já os osciladores acoplados são aqueles que interagem entre si, em um mesmo processo de sincronização. As pernas no ato de andar ou correr, duas pessoas cantando ou falando juntas são bons exemplos de osciladores acoplados.

De acordo com Clayton *et. al.* (2004), há dois componentes básicos envolvidos nos processos de sincronização:

- 1) *É necessário haver dois ou mais osciladores **autônomos***, ou seja, as oscilações não podem ser conseqüências da interação. Elas devem ser processos ativos independentes que requerem uma fonte interna de energia. Fenômenos de ressonância, por exemplo, não são considerados sincronização, pois o oscilador que ressoa é dependente do outro, fonte da energia. Cessada a fonte, cessa-se o ressoar.
- 2) *Os osciladores precisam interagir*, ou seja, um oscilador deve **influenciar** o outro e vice-versa. A quantificação da interação ainda é um dos desafios das pesquisas sobre sincronização, mas sabe-se que na grande maioria dos casos, os sistemas interagem de forma fraca. A grande força de uma interação pode fazer com que os osciladores percam a individualidade de seus movimentos e passem a agir como um sistema único.

2.4 Fala como movimento

Conforme exposto anteriormente, Fred Cummins parte de um ponto de vista dinâmico para apresentar suas idéias e desenvolver estudos acerca do ritmo, com foco em aspectos temporais da fala. Ver a fala como movimento parece algo óbvio, porém trata-se de uma idéia que só ganhou reconhecimento lingüístico a partir de Liberman e Mattingly (1985), que apresentaram o conceito de gesto fonético, uma unidade lingüística surgida de questionamentos sobre a percepção da fala.

A *Teoria motora da percepção da fala* (Lieberman; Mattingly, 1985) sustenta a tese de que a unidade fonética não são os traços, os conhecidos ‘labial’, ‘palatal’, ‘fricativo’, *etc.*; como afirmam as teorias estruturalistas tradicionais, mas o gesto fonético:

[...] the objects of speech perception are the intended **phonetic gestures** of the speaker, represented in the brain as invariant motor commands that call for movements of the articulators through certain linguistically significant configurations. These gestural commands are the physical reality underlying the traditional phonetic notions. (p. 2)

Essa unidade lingüística é fruto da observação acústica da fala: percebeu-se que os traços se sobrepunham temporal e espacialmente, influenciando-se mutuamente em sua realização, o que vai contra a tradicional idéia, de que os fones eram isolados uns dos outros e não se influenciavam. O sinal acústico observado em uma consoante como /d/, por exemplo, varia consideravelmente, pois ela pode vir seguida de /i/, vogal não-arredondada, ou de /u/, vogal arredondada. O conseqüente arredondamento da consoante é visível no sinal acústico, pois o gesto de /u/ se sobrepõe ao de /d/, havendo, portanto, uma coarticulação. A pauta gestual, trazida abaixo, é o modo de representação do gesto. Cada barra corresponde a um gesto e a sobreposição destas barras mostra a sobreposição destes gestos na fala. Assim:

Figura 6: Pauta Gestual da sílaba /du/

Não se trata, o gesto articulatório, de uma materialização em movimentos efetivos dos articuladores, mas de comandos invariantes para a implementação desses movimentos em tempo real no trato vocal (Albano, 2002).

Os estímulos responsáveis pela percepção da fala são de ordem motora e prescindem de processamentos intermediários e, assim, os autores acreditam que perceber a fala é perceber um padrão específico de gestos pretendidos.

Um interessante efeito conhecido a respeito da percepção da fala é o chamado Efeito McGurk, descoberto acidentalmente pelo pesquisador Harry McGurk e seu assistente (McGurk, MacDonald; 1976). Eles dispuseram imagens de falantes realizando a sílaba /ga/, dublado com o som da sílaba /ba/, ou seja, os ouvintes viam na imagem um trato vocal dizendo /ga/, mas o som apresentado era /ba/. Curiosamente, o que os sujeitos alegavam ouvir e ver era /da/, ou seja, o /ba/ ouvido era influenciado pela imagem vista.

Na presente pesquisa, não realizamos uma análise que envolvesse diretamente o gesto articulatório como unidade lingüística, mas nos baseamos no conceito de tal unidade, principalmente, quando segmentávamos os dados (ver em *Capítulo 2: Metodologia*), pois a sobreposição de gestos foi fator decisivo na colocação dos Pontos de Observação de Sincronização.

Além disso, acreditamos que atribuir a uma unidade lingüística o caráter da dinamicidade, ou na contramão deste discurso, inserir nas ciências lingüísticas uma unidade que parta de um ponto de vista dinâmico e cognitivo sobre a fala é algo essencial para o desenvolvimento de um estudo do ritmo da mesma e por isso, mesmo que de forma indireta, tomamos o gesto articulatório e não os traços como unidade para a observação.

2.5 Planejamento cerebral dos movimentos da fala

Como buscamos mostrar, tanto o gesto articulatório quanto a sincronização estão diretamente ligados à busca pelo entendimento de processos cognitivos relacionados à fala. As questões que norteiam esta pesquisa também se pautam na observação de tais processos cognitivos que, no nosso caso, são referentes a um objeto específico: a fala cantada, termo usado sistematicamente em Raposo de Medeiros (2002) para diferenciar dois tipos de fala.

No entanto, vistas puramente como movimento, ou seja, considerando-se apenas questões de ordem físico-fisiológica, pressupusemos que *fala falada* e *fala cantada*, em princípio, não se diferem, simplesmente por operarem com o mesmo instrumento: o trato vocal. Em outras palavras, a fisiologia do instrumento utilizado para a realização dessas “duas falas” é a mesma, o que nos leva a acreditar que o controle neural de seus movimentos também podem ser o mesmo, ou muito próximos entre si.

Guenther (2003) apresenta alguns quadros de referência (*reference frames*) envolvidos no planejamento dos movimentos da fala que são imprescindíveis para um delineamento dos processos neurais em torno de sua realização. São seis esses quadros:

1) *quadro de referência do comprimento muscular*: referências pautadas no comprimento e velocidades dos músculos. Ilustram a capacidade do sistema motor humano por usar diferentes configurações musculares para a realização de um mesmo fonema sobre diferentes condições.

2) *quadro de referência do articulador*: corresponde aos principais graus de independência dos movimentos dos articuladores. Controla, por exemplo, o envolvimento de vários músculos na realização de um único movimento.

3) *quadro de referência tátil*: descreve os estados dos receptores de pressão na superfície dos articuladores da fala. Registra, por exemplo, a pressão entre língua e alvéolos na realização de um /t/.

4) *quadro de referência de constrição*: descreve localizações e graus das principais constrições, como as do corpo e da ponta da língua e a constrição labial.

5) *quadro de referência acústica*: descreve as principais propriedades do sinal acústico produzido pelo trato vocal, como frequências, amplitudes e larguras de banda.

6) *quadro de referência de percepção auditiva*: sistema que traduz o sinal acústico ao sistema nervoso.

Pode-se notar que esses quadros são de grande importância quando se trata de produção e percepção da fala, pois descrevem como a fala é produzida e processada pelo cérebro. No contexto da tarefa de sincronização, produção e percepção devem andar de mãos dadas, pois é necessário a um dado sujeito perceber a fala realizada pelo parceiro ao mesmo instante em que produz a sua fala de forma adequada.

O quadro de referência do articulador (2), o quadro de referência acústica (5) e o quadro de referência de percepção auditiva (6), a nosso ver, estão diretamente ligados à tarefa de sincronização entre os dois cantores do experimento aqui realizado. O primeiro deles (2) justamente por ser o regente dos movimentos dos articuladores do trato. O quadro de referência acústica (5) por descrever, por exemplo, as durações das sílabas ou gestos; e o quadro de referência de percepção auditiva (6) por traduzir o sinal acústico ao sistema nervoso. Na tarefa de sincronização, este quadro é de grande importância, pois é aquele que faz com que o ouvinte se ouça e ouça também o parceiro de dupla, ou seja, além de traduzir o próprio sinal do falante, permitindo, por exemplo, que o cantor se auto-avale, observando se “entrou” ou não no momento correto, este quadro é aquele que permite ao cantor ouvir o outro.

Guenther (2003) chama a atenção ainda para um conceito que permeia todos esses quadros: o de *equivalência motora*. Esta seria a capacidade de usar diferentes movimentos para se obter o mesmo resultado sobre diferentes condições. Acreditamos que esse quadro seja aquele que permita a existência de um elo entre fala cantada e fala falada, permitindo aos cantores a realização da fala no contexto da canção, pois, por mais que se diferenciem fala falada e fala cantada, os fones utilizados nestes dois contextos são os mesmos.

2.6 Acentos e Ritmos

O conceito de acento está na base dos conceitos de métrica e de ritmo, tanto na música como na lingüística. Em cada um desses universos, tal conceito recebe uma definição específica, mas ambas são pautadas na idéia de proeminência. Veremos abaixo como se caracteriza o acento nesses dois universos.

2.6.1 Caracterização do acento em português brasileiro

O acento em português brasileiro foi definido pelas gramáticas tradicionais como uma maior força expiratória na produção da fala (Luft, 1978; Bechara, 2009).

Delgado Martins¹³ (1986, *apud* Massini-Cagliari, 1992) aponta como principais correlatos do acento do português de Portugal a duração e a intensidade em oxítonas e proparoxítonas, sendo as paroxítonas não marcadas acusticamente. A variante paulista do português também foi estudada. Fernandes (1976, *apud* Massini Cagliari, 1992¹⁴) concluiu que nesse caso a intensidade é o que marca o acento lexical, seguido da frequência e por último da duração.

Atualmente, tem-se amplamente divulgado e aceito o apresentado por Massini-Cagliari (1992), que estudou o acento em português também em busca do ritmo da fala. A autora aponta como principal correlato do acento em português brasileiro a duração, seguida da intensidade e, por último, da qualidade vocálica, sendo ainda o acento uma característica da sílaba como um todo, e não somente da vogal.

2.6.2 Acento e ritmo na música

Para entender o conceito de ritmo e acento na música, apresentamos também os conceitos de pulso e de metro.

Pulso, no contexto musical, é, de acordo com Cooper e Meyer (1960), um estímulo sonoro dentro de uma série regular de tempo. Como cada uma das batidas de um relógio ou de um metrônomo, o pulso demarca unidades iguais no contínuo temporal e, embora

¹³ DELGADO MARTINS, M. R. *Sept sur la perception*. Lisboa: Instituto Nacional de Investigação Científica, 1986. Tese de doutoramento.

¹⁴ FERNANDES, N. *Contrin]buições para uma análise instrumental da acentuação e intonação (sic) do português*. São Paulo: USP, (dissertação de mestrado), 1976.

geralmente estabelecido e apoiado por estímulos objetivos (o som realizado), a sensação de pulso pode existir subjetivamente (p. 3). Por estarem organizados em intervalos regulares de tempo, esses pulsos, mesmo depois de cessados, tendem a causar expectativa na mente e na musculatura daqueles que os ouvem (Cooper; Meyer, 1960). Em outras palavras, o pulso não é algo que dependa de algum instrumento para existir, ele está presente em toda a música, mas não é necessariamente tocado por nenhum instrumentista específico. É uma batida abstrata, a qual serve de referência para as execuções de cada músico.

A definição de acento apresentada pelos autores diz que um acento é um estímulo (em uma série de estímulos) marcado pela consciência, de alguma forma (Cooper; Meyer, 1960, p. 8). A diferença entre batidas acentuadas e não acentuadas reside no fato de as batidas acentuadas serem o ponto-foco, o núcleo do ritmo, em torno da qual batidas não acentuadas são agrupadas e em relação ao qual eles são ouvidos. Os autores afirmam que não há como qualificar, do ponto de vista acústico, as características dos acentos musicais. Eles podem recair tanto em notas longas como em breves, tanto nas suaves quanto nas intensas, tanto em notas graves quanto em notas agudas, e com frequência de ocorrência regular ou irregular, sendo assim um conceito axiomático (Cooper; Meyer, 1960).

Metro é a medida da regularidade de pulsos acentuados. Três fatores influenciam para determinar se um pulso é ou não acentuado: a intensidade com que esse pulso é tocado, a duração desse pulso e a sua localização (mais grave ou mais agudo). Pulsos acentuados tendem a ter maior duração e maior intensidade em relação aos pulsos não acentuados. Ambos, quando inseridos no contexto métrico, são chamados de batida (*beat*) sendo adjetivados de "forte" os acentuados em oposição às batidas "fracas", as não acentuadas. Embora o metro tenda a ser regular, irregularidades podem ocorrer sem destruir a sensação de organização métrica, pois essas irregularidades tendem a ser temporárias. Na música ocidental tomou-se por convenção, por questões editoriais das partituras, representar o acento métrico sempre no início dos compassos, imediatamente após a barra de compasso.

Finalmente, ritmo é definido pelos mesmos autores (Cooper; Meyer, 1960) como sendo o modo como uma ou mais batidas fracas são agrupadas em relação a uma batida forte. Tais autores afirmam serem cinco os ritmos básicos: o iambo, o anapesto, o troqueu, o dátilo e o anfíbraco. Curioso notar que eles usam termos usados pela métrica poética clássica. Tal fato se explica, de acordo com os autores, com o argumento de que na Grécia Antiga, local e época onde surgiram os primórdios das teorias musicais empregadas hoje no ocidente, a música era usada exclusivamente como acompanhamento para poesias.

Percebe-se que Cooper e Meyer distinguem o acento métrico do acento rítmico, embora, de modo geral, eles sejam coincidentes, levando-nos a pensar que são o mesmo fenômeno. Segundo os mesmos, ainda, a regularidade métrica pode ocorrer sem que haja, necessariamente, a existência de um ritmo; assim como este pode ocorrer sem que haja a necessidade da ocorrência de regularidade métrica.

2.6.2.1 Ritmo: aspectos objetivos e subjetivos

Especificamente na área da música, Mário de Andrade (1995) afirma que a dificuldade da aquisição do conceito de ritmo se dá devido ao fato de o tempo, sua ‘matéria prima’, ser uma entidade abstrata. Explica-se: uma organização temporal prática teria sido criada pelo homem para facilitar seu cotidiano. Seria aquela baseada no Sol, no dia e na noite e nas fases da Lua. Outra organização temporal, não mais pautada no exterior, mas sim no interior do sujeito, seria aquela originada justamente da natureza fisiológica de movimentos repetitivos, como principalmente os da circulação do fluído sanguíneo e do coração. Para auxiliar a sua manifestação vital físico-psíquica, o homem teria organizado o tempo naquilo que se dá o nome de ritmo. Para Mário de Andrade ainda, o ritmo realizado concreta ou mentalmente se organizaria por meio de duas unidades: uma temporal-fisiológica, chamada de unidade artística do tempo rítmico; e outra afetivo-espiritual: o acento. Esse acento seria de caráter afetivo por ser um ponto de referência escolhido pelo ser psíquico sem que tenha uma razão determinante para essa escolha.

De acordo com Zamprona (2007), assim como a melodia, a harmonia, o timbre e seus movimentos sintáticos, o ritmo é capaz de mover o ouvinte, de transformá-lo, de induzi-lo. Essa indução ocorreria através de dois diferentes fatores: o fisiológico, através da duração, e o psicológico, através da intensidade: “Pela duração, o ritmo penetra em nossa vida fisiológica e pela intensidade, em nossa vida psicológica, (...) induzindo esquemas de movimento e mobilizando formas de comportamento” (p. 44). O ritmo seria ainda o canal pelo qual o indivíduo aprenderia a perceber e a viver a passagem do tempo. Um bebê em gestação já percebe a periodicidade dos sons do meio em que vive, onde o seu coração e o da mãe são marcadores temporais presentes durante todo o período e variáveis devido a diversos fatores como temperatura, pressão arterial, humor, etc. (p. 44). Todo esse processo indutivo se daria de forma forçosa, ou seja, o ouvinte seria um ser passivo na relação, não podendo “tapar o ouvido” diante de seus efeitos (p. 45).

A partir da observação através de experimentos na área de percepção psicológica do tempo, Fraisse (1982), por sua vez, também define dois modos de ritmização: a subjativa e a objetiva. A ritmização subjativa é aquela que faz com que ouvintes tendam a agrupar seqüências de sons idênticos, geralmente em grupos de 2 ou 4. A ritmização objetiva, por sua vez, é o processo estimulado por uma diferença na seqüência dos sons: um dos elementos pode ser mais longo, mais forte, mais alto ou mais baixo em frequência. Nesses casos ainda, pausas são julgadas pelo ouvinte como marcadores de final de grupo e os sons mais longos e mais intensos são ouvidos como acentuados.

Os conceitos apresentados por estes três autores (Zampranha, Andrade e Fraisse) andam em paralelo e, de certo modo, se complementam. Nota-se que os três distinguem, de maneira particular, dois modos de ritmo ou de ritmização: um psicológico, subjativo e afetivo, e outro fisiológico e objetivo. O agrupamento exposto por Fraisse (1982), fruto da ritmização subjativa, se relaciona com o exposto por Andrade (1995) a respeito do acento (ritmo afetivo espiritual), pois é através da acentuação que o ouvinte demarca o agrupamento de sons idênticos (Krumhansl, 2006). O apresentado por Zampranha (2007), por sua vez, também, se relaciona com o exposto por Fraisse por distinguir a ritmização fisiológica, que se aproxima do que Fraisse chamou de ritmização objetiva, da ritmização psicológica, que se aproxima do que o autor francês chamou de ritmização subjativa.

Repp (2004), em estudo dos acentos métrico, dinâmico e rítmico, também a partir do processo de sincronização, afirma que o acento métrico é, ao contrário dos outros dois, puramente subjetivo.

Assim, podemos concluir que o acento métrico se diferencia do acento rítmico por este último ser um fenômeno expresso na obra musical, algo audível; enquanto o acento métrico seria um fenômeno fruto da estrutura de organização de uma peça musical, o que permite que o mesmo nem sempre esteja expresso, audível, mas presente nos alicerces da linha temporal de uma peça musical, e mais do que isso, presente na expectativa do ouvinte.

2.6.2.2 Relações entre metro e ritmo e o conceito de síncopa

A organização métrica tradicional afirma serem três os tipos de métrica musical básica (Priolli, 1987): a métrica binária, com uma batida acentuada a cada duas batidas; a métrica ternária, com uma batida acentuada a cada três batidas; e a métrica quaternária, com uma

batida acentuada a cada quatro batidas, sendo a terceira batida desta considerada “meio acentuada”, ou seja, um intermédio entre tônica e átona¹⁵. As figuras abaixo trazem exemplos básicos desses compassos, nas quais colocamos todas as notas têm a mesma duração. As batidas acentuadas são, por convenção, as imediatamente a direita das barras que demarcam os compassos:

Figura 7: Sequência de quatro compassos com métrica binária

Figura 8: Sequência de três compassos com métrica ternária.

Figura 9: Sequência de dois compassos com métrica quaternária.

Como vimos na seção anterior, o ritmo é definido como o agrupamento de algumas batidas não acentuadas em torno de uma batida acentuada. Esta batida acentuada geralmente coincide com a batida acentuada pela métrica, ou seja, geralmente o acento de um agrupamento ritmo se encontra no ataque dos compassos, marcando também a métrica. É o caso da valsa. No exemplo abaixo, pode-se notar na pauta inferior o ritmo de acompanhamento da valsa executado pelo piano:

Figura 10: Trecho de valsa da opereta O Morcego, de J. Strauss (sem referência).

¹⁵ Nota-se que as convenções adotadas estão ligadas ao observado por Fraisse (1982), que percebeu serem mais frequentes os agrupamentos de 2 ou de 4 unidades.

Tocado na bateria, esse acompanhamento teria a forma trazida abaixo. A nota redonda corresponde ao bumbo, batida forte, e os losangos, às batidas fracas, tocadas pela caixa. Os x indicam batidas realizadas no prato de condução, ou seja, indicam os pulsos:

Figura 11: Partitura para bateria da batida de acompanhamento de uma valsa.

O rock é outro exemplo de ritmo que coincide com a métrica, nesse caso, quaternária:

Figura 12: Partitura para bateria da batida de acompanhamento de um rock.

Essas batidas de acompanhamento expostas acima se encaixam em diversas obras dos referidos gêneros. A primeira delas se encaixa em qualquer peça composta no gênero valsa, como as peças de Strauss, Chopin, Villa-Lobos, Chiquinha Gonzaga e até as “valsinhas” e “mazurquinhas” de Luiz Gonzaga. A segunda delas é a batida original do rock e se encaixa perfeitamente, do modo com transcrita, em muitas peças de Elvis Presley, Little Richard, Beatles, *etc.*, e também em várias peças do rock brasileiro, dentre as quais citamos “Rock das Aranha” e “Aluga-se”, de Raul Seixas.

Outros ritmos de acompanhamento, tanto de valsa, como de rock, se desenvolveram a partir das apresentadas acima, mas sempre mantendo a coincidência entre as acentuações métrica e rítmica. Outros exemplos de ritmos que apresentam este tipo de coincidência são a marcha, o *blues*, a maioria dos sub-gêneros do rock – *heavy metal*, *punk*, *hard rock*, *hardcore* – o *rap*, dentre outros.

No entanto, em muitos ritmos de acompanhamento pode ocorrer um conflito acentual entre os acentos métrico e rítmico, conflito esse denominado síncope, ou seja, eles não coincidem. Priolli (1987) define síncope (segundo a autora, síncope) como o prolongamento de uma nota executada em tempo fraco até um tempo forte, prolongamento este que causa a sensação de deslocamento do padrão de acentuação métrica. Na figura 13, a síncope ocorre

no adiantamento do acento do segundo compasso, ou seja, no prolongamento da segunda nota do primeiro compasso até o primeiro tempo do compasso seguinte:

Figura 13: Exemplo 1 de síncopa (Priolli, 1987)

A síncopa pode também ocorrer da seguinte forma, em quartos de tempo:

Figura 14: Exemplo 2 de síncopa (Priolli, 1987)

Ela pode ocorrer tanto na linha melódica da voz (como ocorre, inclusive, no rock (Temperley, 2001), por exemplo), como também no ritmo de acompanhamento (como no samba, no *jazz*), como na linha do baixo (como no *reggae*). No ritmo de acompanhamento da bossa-nova ocorre justamente esse conflito, a síncopa, essencial para a caracterização deste ritmo como tal:

Figura 15: Ritmo de acompanhamento da bossa-nova (Adolfo, 1994)

Na figura acima: a síncopa ocorre na ultima batida de cada compasso (representadas pelas figuras pontuadas), que se inicia no tempo fraco do terceiro tempo, se prolongando por todo o quarto tempo¹⁶. Esse ritmo de acompanhamento se encaixa em qualquer canção do gênero bossa-nova, como Corcovado, Águas de Março, Garota de Ipanema, *etc.*.

Hipotetizamos, assim como defendem muitos músicos (cantores ou instrumentistas), que canções com ritmo de acompanhamento sincopado são mais difíceis de cantar,

¹⁶ A bossa nova é comumente descrita como de métrica 2/2. Utilizamos a 4/4 para facilitar a visualização dos tempos fortes e fracos, já que ambas as métricas são correspondentes, ou seja, se estruturam da mesma maneira. Fizemos assim em todo este trabalho.

justamente devido à presença da síncopa. Podemos assim dizer que a fala cantada é mais difícil de ser executada quando é acompanhada de ritmo sincopado. De modo a colocar isto a prova, ou, de confirmar de maneira científica esta hipótese, resolvemos usar a sincronização como ferramenta para observar o comportamento da fala cantada ora em um contexto com ritmo de acompanhamento sincopado e ora em um contexto com ritmo de acompanhamento não sincopado, medindo a sincronia entre os cantores de quatro duplas diferentes.

O capítulo a seguir descreve, então, um experimento que desenvolvemos com a fala cantada nos dois contextos rítmicos acima citados.

3 - Metodologia

No contexto da fala cantada, a articulação do trato vocal sofre algumas modificações justamente devido a uma outra estrutura que também sustenta a canção que não a lingüística, a estrutura musical. Os mesmos fones usados na fala falada são transportados ao universo da música, universo dominado por frequências pré-determinadas (as notas musicais) e por relações de tempo (duração e ritmo) que acabam por caracterizar a fala como fala cantada.

A sincronização da fala falada já tem sido objeto de pesquisas e a motivação para tal está em questões cognitivas acerca da língua inserida na linha do tempo. Por muito tempo, como vimos, buscou-se relacionar a fala na linha do tempo com “ritmos” determinados: ritmo acentual e ritmo silábico, através da equívoca idéia de isocronia. A literatura a respeito classifica as línguas como se classificam os ritmos musicais, separando-as em dicotomias que, de fato, nunca foram comprovadas.

No entanto, muito se tem a observar, ainda, sobre o comportamento das línguas na linha do tempo, principalmente no que se refere a questões de ordem cognitiva. Inserir a fala no contexto musical é transportá-la para uma estrutura temporal bem mais rígida que a da fala falada, pois o tempo na música, como vimos, é o que move uma peça e o que, justamente, ajuda músicos de uma mesma orquestra a tocarem “a mesma coisa ao mesmo tempo”.

Assim, a sincronização como ferramenta de observação nos permite justamente olhar para a fala inserida no contexto temporal, o que, por sua vez, abre portas para observações cognitivas acerca da relação entre língua, tempo e ritmização, questão tão controversa há anos.

Assim, desenvolvemos um experimento no qual pudéssemos observar o processo de sincronização na fala cantada em dois contextos rítmico-musicais: o rock e a bossa-nova. Observamos esse processo na fala cantada realizada por quatro duplas distintas: a partir de análises fonético-acústicas medimos ao longo do canto de quatro canções (duas bossas e dois rocks) alguns pontos de sincronização entre os cantores de uma mesma dupla, que cantaram as canções em uníssono.

O ritmo das canções e os pontos de sincronização formam, cada um, um fator de nosso experimento, definido, portanto, como um experimento fatorial com dois fatores (Calegare, 2001), denominados justamente de Fator A: Ritmo da Canção e Fator B: Ponto de Observação de Sincronização.

Cada um desses fatores apresenta dois níveis. O fator A apresenta o nível Ritmo Sincopado (RS) e o nível Ritmo Não-Sincopado (RNS); enquanto o fator B apresenta os

níveis Ponto Inicial (PI) e Ponto Final (PF). A tabela abaixo ilustra esse delineamento experimental:

Tabela 2: Delineamento do Experimento. Os fatores e seus níveis.

Níveis	Fatores			
	Ritmo da Canção		Ponto de Observação da Sincronização	
	RS	RNS	PI	PF

Detalhamos, a seguir, a elaboração desse experimento.

3.1 *Corpus*

Observamos a canção brasileira inserida em dois contextos rítmico-musicais diferentes: aquele no qual o ritmo de acompanhamento (aquilo que é realizado por um pandeiro, uma bateria ou a mão direita de um violonista em um acompanhamento) é sincopado, e aquele outro onde o ritmo de acompanhamento não é sincopado. Colocamos nosso olhar sobre a síncopa justamente por ela ser um conflito envolvendo aquilo que sustenta toda a estrutura rítmico-musical da canção: o acento musical (Cooper; Meyer, 1960).

Quanto à relação do acento da fala com o ritmo musical, há observações qualitativas no Capítulo 3. No presente capítulo, apresentamos os métodos utilizados para que pudéssemos aferir a sincronização no canto.

Selecionamos quatro canções populares compostas por cancionistas brasileiros contemporâneos para integrar nosso *corpus*. Duas são do gênero bossa-nova e duas do gênero rock. São elas, respectivamente, *Águas de Março* e *Corcovado*, ambas de Antônio Carlos Jobim; *Aluga-se*, de Raul Seixas e Cláudio Roberto e *Gita*, de Raul Seixas e Paulo Coelho. Essas canções foram separadas em dois grupos: de um lado, colocamos as bossas-novas, que, por definição, possuem o rítmico de acompanhamento sincopado; e do outro colocamos os *rocks*, que por definição têm o ritmo de acompanhamento não sincopado.

Escolhemos essas canções devido a alguns fatores elucidados a seguir.

Primeiramente, buscamos canções que fossem comercialmente conhecidas, afim de que fossem já familiares aos cantores que serviram de sujeitos. Em segundo lugar, precisávamos de canções que se enquadrassem em nosso *design* experimental: andamento

médio (entre 80 e 120 batidas por minuto), com sentenças de aproximadamente igual extensão (tanto na comparação entre as canções quanto na comparação entre as sentenças de uma mesma canção) e, principalmente, sem síncopas (ou com o mínimo delas) na linha melódica da voz para o caso da condição RNS. Finalmente, buscamos canções cujas partituras já tivessem sido publicadas em algum meio.

Quanto à partitura da canção *Aluga-se*, não encontramos edição da mesma já publicada. Assim, tivemos que elaborá-la através do *software Encore 4.5.3*. Para isso, nos baseamos na gravação da canção lançada em álbum pelo cancionista (Seixas, 1980).

3.2. Sujeitos

Para a coleta dos dados, contamos com 4 diferentes duplas de sujeitos¹⁷: Dupla 1 (Cantores LS & HC), Dupla 2 (CM & CV), Dupla 3 (CS & TB) e Dupla 4 (CQ & BE), como na tabela a seguir:

Tabela 3: Duplas e respectivos sujeitos

Dupla	1	2	3	4
Sujeitos	LS & HC	CM & CV	CS & TB	CQ & BE

Todas as duplas eram formadas por cantores e cantoras profissionais com idades entre 22 e 32 anos e com mais de 3 anos de experiência com canto popular. Todos se apresentam em bares e casas de show de São Paulo, estudaram ou estudam canto popular formalmente e tinham já contato tanto com rock nacional como com bossa-nova. Aqueles que formavam uma mesma dupla já se conheciam previamente e já haviam tido a experiência de “cantar juntos” em outros momentos, mesmo que de maneira não-profissional.

Em nosso desenho experimental, uma dupla é um sistema de dois osciladores acoplados, no qual cada um dos sujeitos é um oscilador. Eles realizam movimentos, os movimentos próprios da fala cantada, a partir de fontes de energia interna próprias (o sistema respiratório) e tais movimentos são ao mesmo tempo influenciadores e influenciados pelos movimentos do outro oscilador do sistema: o outro cantor; corroborando, assim, com os conceitos de oscilador e de osciladores acoplados mostrados no *Capítulo 1*.

¹⁷ Oferecemos R\$ 50,00 a cada cantor para a realização da tarefa.

O ritmo que observamos aqui é o “elo” que permite a comunicação motora entre os dois sujeitos, permitindo-os que se movimentem de forma sincrônica, ou a mais sincrônica possível. Um ritmo não-sincopado, portanto aquele que não entra em conflito com a estrutura métrica básica, é um “elo” mais simples, que facilita a coordenação dos movimentos justamente pela ausência da síncopa, do conflito. Em comparação a este, o ritmo sincopado é um elo de maior complexidade, justamente devido à presença da síncopa, que acreditamos tornar tal elo mais frágil, mais fácil de ser rompido.

3.3 Preparação para a coleta

Fornecemos previamente aos cantores, cerca de uma semana antes da gravação, as partituras, as letras e uma cópia da gravação original de cada canção. Nenhum deles, no entanto, demonstrou necessidade, vontade ou capacidade de fazer uso das partituras, acabando por utilizar, para treino, apenas as cópias das letras e do áudio que fornecemos (ver a seguir). Para a gravação as duplas utilizaram apenas as cópias das letras.

Todas essas canções possuem várias gravações e interpretações, como as versões da banda Titãs e da banda Camisa de Vênus da canção *Aluga-se*, as de *Gita* por Maria Bethânia, Rita Lee e pela dupla sertaneja Milionário & José Rico, e a de *Corcovado* do álbum *Elis & Tom*, entre outras.

Fornecemos, então, a gravação de uma interpretação consagrada de cada canção como referência para treino prévio e para evitar desacordos no ato da coleta, como diferenças de andamento¹⁸ e de interpretação. De *Gita*, fornecemos a versão do álbum *Gita* (Seixas, 1974) e de *Aluga-se* a do álbum *Abre-te Sésamo* (Seixas, 1980), ambos do compositor da canção, Raul Seixas. De *Corcovado* fornecemos a versão contida no álbum *O Amor, o Sorriso e a Flor* (Oliveira, 1960), de João Gilberto, álbum no qual o compositor da canção, Tom Jobim, trabalhou como arranjador, diretor musical e pianista. De *Águas de Março*, finalmente, fornecemos a interpretação do álbum *Elis & Tom* (Costa; Jobim, 1974), em que o compositor fez duetos vocais com a cantora Elis Regina, além de ter tocado violão e piano.

Essas versões indicadas são as primeiras versões de estúdio das canções selecionadas, com exceção de *Águas de Março*, lançada no compacto *O Tom de Antônio Carlos Jobim e o*

¹⁸ Não controlamos o andamento no ato da gravação, pois um metrônomo seria um componente a mais em nosso design experimental, o que não era nossa intenção. Ele acabaria sendo um guia, ao qual os cantores teriam que se adaptar. Não era essa nossa intenção.

tal João Bosco - *Disco de Bolso*, dois anos antes (1972, portanto) do lançamento do álbum *Elis & Tom* e parte integrante do semanário *O Pasquim*. Além de não alcançar sucesso comercial (o álbum continha apenas mais outra canção: *Agnus Sei*, de João Bosco, estreante no cenário musical), nessa gravação de *Águas de Março* o andamento é mais rápido do que o das gravações posteriores e podem-se ouvir, claramente, trechos nos quais Tom Jobim teve dificuldade de cantar. Por isso indicamos a versão de 1974, comercialmente mais conhecida e aceita hoje em dia¹⁹.

Essas versões são, ainda, as que serviram de referência para a elaboração das quatro partituras aqui utilizadas, ou seja, das partituras elaboradas por Chediak (1994a, 1994b), Chediak (2004) e da elaborada por nós.

3.4 Coleta

O *set* de gravação foi disposto de forma que os cantores ficassem a cerca de 3 metros um do outro e no mesmo ambiente de gravação²⁰, ou seja, com contato visual e sonoro direto. Assim, impedia-se através da distância que a voz de um cantor fosse captada pelo microfone do outro, o que inviabilizaria as análises acústicas. Dessa forma também, se permitia aos cantores a possibilidade de se comunicarem gestual e visualmente durante a realização da tarefa, o que fizeram todos eles.

Os microfones utilizados eram específicos para gravação de voz - *Shure SM58* - foram ligados a canais separados e colocados em pedestais próprios. Com a utilização dos pedestais, possibilitávamos aos cantores que se comunicassem gestualmente com as mãos.

Utilizou-se o *software Soundforge*, da marca *Sony*, para coleta/digitalização do áudio, salvos em formato *wav*.

Como tarefa, pedimos às duplas que cantassem as canções selecionadas em uníssono, “o mais sincrônico possível”, *a capella* e na tonalidade e ordem que preferissem. Cada canção foi cantada 3 vezes e a essas vezes precederam algumas a fim de adaptação ao ambiente do estúdio, alongamento e aquecimento vocais, sendo que as três repetições de cada canção foram coletadas em sequência. Passamos a gravar quando as duplas se declararam prontas para tal. De acordo com Cummins (2003), a partir da terceira repetição, não há

¹⁹ Outra versão desta canção foi lançada em 1973 no álbum *Matita Perê*, de Tom Jobim. Não utilizamos esta também por não ser tão conhecida.

²⁰ Utilizamos o Estúdio Multimeios do Centro de Computação Eletrônica da USP. Tal ambiente é próprio para gravações de áudio.

diferenças significativas na tarefa de sincronização e por isso pedimos aos sujeitos que cantassem as canções algumas vezes antes da gravação e passamos a gravar, pelo menos, a partir da terceira.

Gravamos uma dupla por dia, assim, uma dupla não tinha contato com o trabalho das outras.

Do que foi exposto até aqui, desde a descrição dos sujeitos, até a da coleta, reiteramos que buscamos montar duplas homogêneas, com sujeitos que realizassem a tarefa com igual nível de exigência e de desenvoltura. Não trabalhamos com músicos amadores²¹. Nenhum dos cantores era interprete só de rock ou só de bossa, todos passaram pelo estudo formal do canto popular, todos puderam cantar as canções na tonalidade de sua tessitura, todos já conheciam as canções e todos eles foram instruídos da mesma forma.

3.5 Segmentação e rotulagem do sinal acústico

Segmentamos então as canções em sentenças, definidas a partir das *frases musicais* (Schönberg, 1996) cuja definição vê-se a seguir. Em seguida, rotulamos essas sentenças com a marcação de dois pontos pelos quais observamos a sincronização entre a dupla: um ponto colocado em algum dos primeiros fones de cada sentença, ao qual chamamos de *Ponto Inicial*; e outro ponto colocado em algum dos últimos fones de cada sentença, chamado de *Ponto Final*. Esses pontos formam também um fator experimental, denominado *Ponto de Sincronização*, cujos níveis são denominados justamente de *Ponto Inicial (PI)* e *Ponto Final (PF)*.

Apresentamos, a seguir, a metodologia desses dois processos: segmentação e rotulagem.

3.5.1 A frase musical

Schönberg (1996) define frase musical justamente a partir de exemplos da fala cantada: “o termo frase significa, do ponto de vista da estrutura, uma unidade aproximada àquilo que se pode cantar em um só fôlego. (...) Seu final sugere uma forma de pontuação, tal como a vírgula.” (p. 29).

²¹ Uma das duplas que convidamos não satisfaz nossos critérios, fato que nos obrigou a descartar os dados.

Ammer (1987) define frase como a divisão de uma melodia que pode ser comparada a uma linha de um poema.

Ambos teóricos, da área da música, utilizam exemplos ligados à língua para conceituar *frase musical*. Ao sugerir a pontuação como critério, Schönberg utiliza-se, implicitamente, do conceito lingüístico de *sintagma*, já que uma vírgula ou um ponto sempre ocupam as fronteiras dos sintagmas sintáticos. Ammer também se utiliza de uma comparação baseada em um “fragmento” da língua, mesmo que inserida no contexto literário: o verso. Schönberg, ainda, utiliza-se especificamente, de algo que, se inserido no contexto da lingüística, seria por natureza objeto da fonética: o fôlego.

O compositor afirma que a frase é a menor unidade estrutural de uma peça musical, formada por notas que se unem *com certa completude de sentido e adaptáveis com outras unidades similares*. Como exemplo, traz uma frase do primeiro movimento do Quarteto de Cordas, *opus* 18/1, de Beethoven:

Figura 16: Exemplo 1 de frase musical.

O final de uma frase musical é delimitado por alguma alteração em relação ao padrão que compõe a própria frase, como “redução rítmica, relaxamento melódico determinado por uma queda de frequência, uso de intervalos menores e de um menor número de notas, ou por qualquer outra *forma adequada de diferenciação*” (Schönberg, 1996, p. 30, grifo nosso).

No exemplo acima trazido, tem-se uma frase formada pela sequência de sete notas, a saber: Fá, Fá, Sol, Fá, Mi, Fá e Dó. O final dessa frase é marcado pela nota Dó, que quando comparada ao restante da frase, diferencia-se devido à queda de frequência, já que é a nota mais grave da sequência. Esse Dó, ainda, é o quinto grau da escala de Fá maior, a tonalidade da peça. O quinto grau de uma escala diatônica qualquer tem por uma das funções indicar tensão no discurso harmônico e o direcionamento à tônica (Schönberg, 1999), como reticências ou um ponto de indagação. Assim, no exemplo apresentado, o Dó caracteriza-se como diferenciador não apenas do ponto de vista rítmico, por durar mais, mas também como um diferenciador a partir do ponto de vista melódico-harmônico.

Com um exemplo trazido de nosso *corpus*, mostramos que na canção *Gita*, a primeira frase tem o final demarcado por uma nota mais longa quando comparada às outras:

Figura 17: Exemplo 2 de frase musical (Chediak, 2004)

Pode-se notar que a última nota da seqüência, Fá#, vale 2 tempos de uma tercina (2/3 de 2 tempos), enquanto que as outras notas da seqüência valem 1, 1 tempo de tercina (1/3 de 2 tempos) ou 1/2 tempo²². Essa nota de 2 tempos de tercina portanto, é a nota que apresenta a diferenciação, por ser mais longa em relação ao restante da frase, e por isso demarca o seu final. Do ponto de vista melódico-harmônico, o Fá# é, na tonalidade E (o desse trecho da canção *Gita*) o segundo grau da escala, que pode sugerir a completude de uma frase, mas indica que o texto musical ainda vai continuar.

Na canção *Águas de Março*, as formas de diferenciação que delimitam as frases são também notas longas, que nesse caso são justamente síncopas²³. Nessa canção quase não há pausas e o que delimita o fim da frase são somente essas síncopas, aqui circuladas em azul. Essas síncopas ainda são as notas mais graves das frases, Si:

Figura 18: Exemplo 3 de frase musical (Chediak, 2004)

A presença de pausa em muitos casos pode definir o final de uma frase (Krumhansl, 2006) por representar justamente a ausência de som e, assim, aquilo que seria o esgotamento do fôlego (Schönberg, 1996). No entanto, a ausência ou presença de pausa não implica em final ou não da frase. Como afirma Schönberg (1996.), o que delimita a frase musical é aquilo que apresenta diferenciação na frase.

²² No apêndice X podem-se consultar as figuras musicais de acordo com seus valores de tempo e os valores de frequência (de altura) das notas.

²³ Nesse caso, síncopas na linha melódica da voz, e não na estrutura rítmica básica. Síncopas na linha melódica são também frequentemente encontradas no rock.

Na seqüência da canção *Gita*, logo na segunda frase musical, pode-se notar a presença de uma pausa - aqui circulada em verde - que não representa o final da frase em questão:

Figura 191: Exemplo 4 de frase musical (Chediak, 2004)

Essa pausa não determina o fim de uma frase, já que as duas frases resultantes seriam duas frases sem sentido melódico completo, principalmente se comparadas às outras frases musicais da canção. Portanto, o final dessa frase, assim como a que a precede, é marcado por uma nota mais longa, nesse caso, o Sol#4 com valor de 2 tempos de tercina.

3.5.2 Segmentação das canções

Baseamo-nos em *songbooks* elaborados por Chediak (1994a, 1994b) e Chediak (2004) para a realização da segmentação em sentenças. A partitura de *Aluga-se*, como já dissemos, foi elaborada por nós²⁴. Esses *songbooks* são muito utilizados por conservatórios e universidades onde se estuda música popular e por músicos de um modo geral, devido ao rigor e a fidelidade às canções.

Tendo as frases musicais como critério para a segmentação, obtivemos 44 sentenças da canção *Gita* e 20 da canção *Aluga-se*, somando 64 na condição RNS. Da canção *Corcovado* obtivemos 12 sentenças e de *Águas de Março*, 51, somando 63 na condição RS.

No entanto, algumas dessas frases não se encaixavam em nosso *design*, ora por serem muito curtas, ora por não apresentarem fones oclusivos ou fricativos. Essas frases foram excluídas das análises. Outro motivo que nos levou a retirar algumas outras sentenças foi o fato de querermos a mesma quantidade de sentenças nas duas condições, a fim de simplificar as análises estatísticas. Assim, conseguimos obter 55 sentenças em cada condição.

A figura abaixo traz exemplos da segmentação em frases nos sete primeiros compassos da canção *Gita*. As pausas circuladas em azul, portanto, são exemplos das que foram usadas para segmentar a canção em sentenças, ou seja, a partir da pausa contamos uma

²⁴ Agradecemos ao músico e amigo Carlos Ferreira pela revisão da partitura.

nova sentença. A pausa em verde é um exemplo daquelas que não usamos como critério para a segmentação. As sentenças estão sublinhadas em dourado:

Figura 20: Exemplo de segmentação da canção Gita (Chediak, 2004).

A Figura 21 traz exemplos da segmentação na canção Águas de Março. As síncopas nesse caso, como dissemos, delimitam as frases musicais e, portanto, o que chamamos de sentença em nosso experimento:

Figura 212: Exemplo de segmentação da canção Águas de Março (Chediak, 2004).

A Figura 22 mostra a segmentação da canção Aluga-se. Nessa canção, as pausas são o que delimitam as sentenças:

A so - lu - ção pro nos - so po - vo_eu vou dar
Os es - tran - gei - ro _eu sei que _e - les - vão gos - tar

Ne - gó - cio bom as - sim nin - guém nun - ca viu
Tem o A - tlân - ti - co tem vis - ta pro mar

Figura 22: Exemplo de segmentação da canção Aluga-se. As 4 primeiras notas são a introdução.

E finalmente a Figura 23 traz exemplos da segmentação da canção Corcovado. A partitura publicada dessa canção não traz a letra subscrita. A apresentada a seguir foi transcrita por nós através do *Encore*, baseada na publicada por Chediak (1994b). As frases apresentadas são delimitadas por notas longas precedidas por pausas. A terceira frase musical apresenta já uma variação das duas primeiras frases, se estendendo por mais dois compassos:

Um can ti nho _um vi o lão E sse _a mor u ma can ção

Pra fa zer fe liz a quem se a ma

Figura 23: Exemplo de segmentação da canção Corcovado.

A segmentação completa das quatro canções pode ser vista no Apêndice 1.

3.5.3 Rotulagem e definição de *lag*

Em cada uma das sentenças obtidas, colocamos dois pontos pelos quais observamos a sincronização entre os cantores. Esses pontos, um no início (Ponto Inicial) e outro no final (Ponto Final), foram colocados principalmente a partir de critérios fonéticos, mas também a partir de critérios musicais. Priorizamos a colocação desses pontos no ataque de determinados

fonos fricativos e oclusivos por terem a borda esquerda mais facilmente distinguível de vogais e sonorantes) cujo *onset* coincidia com o ataque de notas musicais. Não conseguimos fazer assim com todos os pontos, devido à natureza das canções, fato que nos levou a colocar alguns deles no *onset* de outros fonemas que não oclusivos ou fricativos (mas que mesmo assim tinham o início bem definidos – como vogais em início) e cujos inícios coincidiam ainda com o ataque de notas musicais. Foram poucos esses casos.

Primamos em colocar os pontos em fonos assim caracterizados, pois necessitávamos de uma precisão da ordem de milissegundos para medir a sincronização entre os pontos. Colocar um ponto de observação de sincronização, por exemplo, entre duas vogais, seria tarefa difícil, já que esses gestos vocálicos podem se sobrepor de tal forma que seria difícil identificar o *onset* do segundo fone, onde seria colocado o ponto.

Utilizamos o *software Praat 5.2.22* (Boersma; Weenink, 2011) para a colocação desses pontos, através das ferramentas *TextGrid*. Usamos o *TextGrid* para colocar os pontos no sinal acústico dos dois cantores de cada dupla.

Colocamos os pontos no sinal dos dois cantores, um de cada vez, pois é assim que o *software* permite. A figura abaixo mostra os pontos colocados no sinal acústico da cantora LS, além de mostrar o espectrograma (na faixa central) e o sinal de onda (na faixa superior):

Figura 24: Ponto Inicial e Ponto Final da primeira sentença da canção Aluga-se, pela cantora LS.

Nessa sentença, o Ponto Inicial foi colocado no *onset* da sentença, no fone [a], e o Ponto Final no *onset* do fone [d] da palavra “dar”.

Colocados os pontos no sinal acústico dos dois cantores, sobreposamos os dois rótulos, podendo assim, calcular a diferença entre os pontos de observação de sincronização, pois o *Praat* nos fornece a posição numérica dos pontos no eixo do tempo:

Figura 25: Pontos Inicial e Final da cantora LS comparados aos pontos Inicial e Final de HC.²⁵

A figura acima mostra, portanto, quatro pontos: dois iniciais e dois finais, da primeira sentença da canção *Aluga-se* cantada pela dupla LS & HC. A linha 1 mostra os pontos colocados no sinal acústico da cantora LS e a linha 2, os pontos colocados no sinal acústico do cantor HC. A sincronização entre os cantores pode ser vista, portanto, nas *lags*: a diferença temporal (representada na figura através da diferença espacial) entre os pontos respectivos entre si, como nos exemplos circutados em amarelo (*Pontos Iniciais*) e em verde (*Pontos Finais*). Essas *lags* são, de fato, aquilo que tomamos nesse experimento como **dado**, pois foram elas que medimos e que usamos para embasar nossa discussão.

5.2.3.1 Rotulagem automatizada: o *Dynamic Time Warping*

Buscamos automatizar o processo de rotulagem através da ferramenta *Dynamic Time Warping* (DTW), um algoritmo que permite a comparação entre séries temporais distintas: rotulado manualmente o sinal acústico de um dos cantores de determinada dupla, o DTW faz a inserção dos rótulos (no nosso caso, os pontos de sincronização) no sinal do segundo cantor, a partir de comparações entre os sinais e do reconhecimento dos rótulos já colocados manualmente.

A eficiência dessa ferramenta se mostrou satisfatória enquanto rotulávamos a canção *Corcovado*, o que nos levou a querer divulgar o uso de tal ferramenta em artigo publicado (Santos; Medeiros, 2011).

A rotulagem do DTW, no entanto, necessita sempre de revisão e, em muitas vezes, de posteriores ajustes manuais dos rótulos, ou seja, muitas vezes o ponto (*inicial* ou *final*) não foi colocado no ataque do fone que determinamos. Notamos, ao decorrer da utilização da

²⁵ Sobrepostos os rótulos dos dois cantores, o *Praat* só apresenta o espectrograma de um deles. Na figura, vê-se o do cantor HC.

ferramenta, quando passamos a rotular as outras canções, que os rótulos colocados automaticamente estavam muito distantes do ideal. Os ajustes posteriores passaram a ser tarefa mais trabalhosa que a colocação manual dos rótulos nos dois sinais e assim, deixamos de utilizar o *DTW* em certo momento da rotulagem, passando a realizar tal tarefa manualmente. Acreditamos que tal fato se deva à grande extensão de alguns de nossos arquivos de áudio: entre 120 e 240 segundos, o que pode ter sobrecarregado o algoritmo.

3.6 Padronização dos valores das *lags*

Como as canções foram compostas em diferentes andamentos e como não controlamos os mesmos no ato da coleta, as gravações coletadas para essa pesquisa foram realizadas, portanto, em diferentes andamentos. Encontramos diferenças de andamento também quando observamos uma mesma canção cantada por duplas diferentes, ou até mesmo, quando observávamos as mesmas canção e dupla, mas em repetições distintas. Dessa forma, as *lags* não puderam ser comparadas diretamente, com seus valores brutos, pois pressupusemos que, relativamente, quanto maior o andamento, menor o valor da *lag*.

Assim, utilizamos dos princípios matemáticos *razão e proporção* para padronizar os dados, igualando proporcionalmente as razões entre *lag* e andamento reais com as razões entre *lag* e andamento ideais. Estipulamos como andamento ideal o valor de 120 batidas por minuto (bpm), o que equivale a *allegretto*. Nesse andamento, as mínimas devem valer entre 500 ms e 600 ms e as colcheias, portanto, entre 250 ms e 300ms. De acordo com Fraisse (1982), o valor de 600 ms é o mais comum na duração de tarefas espontâneas na realização de pulsos periódicos, e por isso determinamos esses valores.

Para esse processo, calculamos a média de duração, em cada repetição, de algumas sílabas aleatórias que correspondiam a colcheias em tempo fraco²⁶ e então calculamos os valores das *lags* proporcionalmente aos valores dessas notas. Convencionamos que todas as colcheias devessem valer 300 ms e assim conseguimos obter valores proporcionais de *lags* através do cálculo dessas proporções.

A fórmula a seguir representa esse cálculo:

$$X = \frac{\lambda \cdot 300}{\alpha}$$

²⁶ A colcheia (♩) é a figura musical que representa uma nota com valor de ½ tempo.

na qual λ representa o valor real de *lag* e α representa o andamento real. O valor obtido X representa o valor ideal de *lag* a partir do valor estipulado ideal das colcheias, 300 ms. Assim, se tínhamos uma *lag* real de 20 ms, a 90 bpm, teríamos:

$$\frac{20.300}{90} = 67$$

Ou seja, 20 ms de *lag* a 90 bpm correspondem a 67 ms de *lag* a 300 bpm.

4. Observações dos Acentos das Canções na Relação Fala- Canto

Introdução

Como vimos no *Capítulo 1 – Ritmo e conceitos circundantes*, o acento musical é protagonista na organização das estruturas métrica e rítmica na música. Vimos que em canções não sincopadas os acentos métricos e rítmicos são coincidentes, soando como se fossem o mesmo, mas independentes quando observados a partir do ponto de vista da estrutura que organizam. Essa coincidência acentual não ocorre nas canções sincopadas, colocando-se, assim, as estruturas métrica e rítmica em conflito, fazendo esta última parecer andar sempre “um passo a frente” da estrutura métrica, devido ao adiantamento, ou melhor, à sensação de adiantamento do acento: a síncopa.

Neste capítulo, realizamos uma observação das estruturas musical e lingüística das canções utilizadas tomando como ponto de partida justamente a acentuação. Mostraremos em quais casos há coincidência entre o acento musical e o acento lexical. Acreditamos que a relação entre essas estruturas acentuais, musical e lingüística, tenha relação com a tarefa de sincronização realizada no experimento: supusemos que, quanto mais coincidente as estruturas lingüística e musical, mais fácil é a realização da tarefa.

A seguir, apresentamos a indicação dos acentos nas sentenças lingüísticas, comparando-as aos acentos de suas respectivas frases musicais. A representação aqui utilizada é baseada na proposta pela Fonologia Métrica (Hayes, 1981, 1989, 1995), baseada por sua vez na Fonologia Autossegmental (Goldsmith, 1976a, 1976b). A representação dos acentos musicais é baseada em Lerdhal e Jackendoff (1983).

A Fonologia Autossegmental considera a existência de dois ou mais níveis de cadeias paralelas de segmentos na representação lingüística, abandonando a idéia de representações fonológicas em uma única cadeia (representação linear). A existência de mais de um nível acarreta uma determinada organização hierárquica da estrutura fonológica, hierarquização que ajuda a entender e a representar, entre outros fenômenos lingüísticos, o acento. Liberman e Prince (1977) afirmam que o fato de muitas línguas apresentarem múltiplos graus de acentos – acento primário, secundário, terciário, *etc.*, é uma evidência da organização acentual em níveis hierárquicos.

Utilizando-se, primeiramente, de árvores e, num segundo momento histórico, também de grades, a representação proposta pela Fonologia Métrica é baseada na Fonologia Autossegmental por pressupor justamente a existência de diferentes níveis hierárquicos. Na teoria de Hayes (1981, 1989), a métrica é sempre pautada na binaridade forte/fraco, sendo

considerada mais forte a sílaba que for acentuada em todos os níveis hierárquicos. Assim, as palavras *tatu*, *livro* e *lingüística* seriam representadas da seguinte maneira, sendo *s* para *forte* e *w* para *fraco*²⁷:

Figura 26: Árvores métricas das palavras *tatu*, *livro* e *lingüística*

ou:

Figura 27: Grades métricas das palavras *tatu*, *livro* e *lingüística*

Reconhece-se o acento principal observando-se os ramos da árvore: a sílaba dominada apenas com nós rotulados *S* é a portadora do acento principal. Na representação por grade, o acento é representado pela sílaba com mais marcações *X*, sendo que cada linha de *X* representa um nível hierárquico.

Trazemos a seguir a grade métrica da sentença *às vezes você me pergunta*, de *Gita*. As sílabas com dois *X* são as portadoras do acento:

Figura 28: Grade métrica da sentença *Às vezes você me pergunta*

²⁷ Utiliza-se, em inglês, as letras *s* e *w* para, respectivamente, forte (*strong*) e fraco (*weak*).

Hayes (1989) apresenta uma grade simplificada, excluindo da representação o primeiro nível e indicando as sílabas átonas com um ponto:

Às vezes você me pergunta
· X · · X · · X ·

Figura 29: Grade métrica simplificada da sentença *às vezes você me pergunta*

Por sua vez, a estrutura métrica de uma frase musical, também pautada na relação *forte/fraco*, pode ser representada através de grades métricas, já que essa binaridade também se faz presente na música e está relacionada aos conceitos de ritmo e de métrica musical. Na representação musical de origem européia, através da pauta, o compasso tem o acento principal no ataque de seu primeiro tempo e o acento secundário no ataque do terceiro tempo, sendo átonos os segundo e quarto tempos, como explicamos no Capítulo 1.

Lerdhal e Jackendoff (1983) apresentam uma nova proposta de representação de uma peça musical, tanto no que se refere ao ritmo, quanto ao que se refere à estrutura melódica; baseada nas intuições musicais do ser humano e em estruturas hierárquicas. Assim, baseamos-nos nesses autores para a representação da métrica musical a partir de grades. A figura abaixo traz a grade métrica da frase musical correspondente à sentença lingüística “às vezes você me pergunta”, apresentada nas figuras 31 e 32. Os principais acentos, portanto os que se encontram nos primeiros tempos dos compassos, são os indicados com a maior quantidade de X:

Figura 30: Representação métrica da primeira frase musical de Gita.

Comparando os acentos da sentença lingüística com os da frase musical apresentada acima, pode-se perceber, portanto, que são coincidentes, ou seja, as sílabas *ve* de *vezes* e a

sílabas *gun* de *pergunta*, as sílabas tônicas, coincidem com os acentos da frase musical, caindo no primeiro tempo dos compassos:

Figura 31: Representação métrica da primeira sentença lingüística de Gita e de sua respectiva frase musical.

A não coincidência acentual entre música e língua pode ocorrer de duas formas. A figura abaixo ilustra esses casos. No primeiro deles, mantém-se a acentuação no primeiro tempo do compasso, mas canta-se nesse tempo uma sílaba átona, como ocorre com o *vi*, de violão. O segundo caso ocorre quando uma sílaba tônica é cantada em tempo fraco, como ocorre com *lão* de violão:

Figura 32: Representação métrica da primeira sentença lingüística de Corcovado e de sua respectiva frase musical.

Este segundo caso é muito freqüente, pois a estrutura lingüística dispõe, a cada compasso, de maior quantidade de acentos que a estrutura musical. Não levamos esses casos em consideração nas observações a seguir.

4.1 Observações dos Acentos

Apresentamos a grade métrica simplificada da sentença lingüística, representando com x minúsculo os acentos e com pontos as sílabas átonas. Colocamos em letras maiúsculas as coincidências entre acentos lingüístico e musical, marcando também em maiúscula o x da grade métrica lingüística nesses casos.

Acima de cada sentença da letra, trazemos o trecho da partitura correspondente, das quais não apresentamos a grade métrica, já que os acentos métrico-musicais, por definição, sempre caem na primeira nota de um compasso, sendo grafados imediatamente a direita das barras de compasso. Utilizamos a primeira sentença da canção *Aluga-se* como exemplo:

A so lu ção pro nosso PO vo_eu vou dar
 . . . x . x . X . . x x

Assim, a grade métrica abaixo da sentença lingüística apresenta as sílabas tônicas marcadas com x. A sílaba coincidente ao acento musical, *po*, é grafada em maiúscula. As sublinhas indicam sílabas cantadas na mesma nota.

4.1.1 Aluga-se

1.

a so lu ção pro nosso PO vo_eu vou dar
 . . . x . x . X . . x x

2.

ne gó cio bom as sim nin GUÉM nun ca viu
 . x . x . x . X x . x

3.

tá tudo pronto_a qui é SÓ vir pe gar
 x x . x . . x x X x . x

4.

5.

6.

7.

8.

9.

10.

11.

12.

13.

e o dó lar de les pa ga_o NOSso mingau (I)
 . . x . x . x . . x . . x

14.

nós não va mos pa gar NA da (II)
 x x x . . x X .

15.

nós não va mos pa gar NA da_é tu do *FREE* (II-A)
 x x x . . x X . x x . X

16.

tá na ho ra_a go ra_é *FREE* (III)
 x . x . . x . x X

17.

va mo_em bo ra dar lu gar (III-A)
 x . . x . x . x

18.

pois es se_i mó vel tá pra_a lu GAR (IV)
 x x . . x . x . . . X

4.1.2 Gita

1.

às VÊ zes vo cê me perGUNta (I)
 . X . . x . . X .

2.

porque_é que_eu sou tão ca LA do (I)
 . x x . . x x . . X .

3.

não FA lo de_a mor quase NAda (I)
 . X . . . x x . X .

4.

nem FI co sorrindo_ao teu LAdo (I)
 . X . . x . . . X .

5.

vo CÊ pensa_em mim toda HO ra (II)
 . X x . . . x . X .

6.

me COme, me cospe e me DEIxa (II)
 . X . . x . . . X .

7.

tal VEZ você não em TENDa (II)
 . X . x . . X .

8.

mas HO je eu vou lhe mos TRAR (II)
 . X . . x . . X

9.

eu SOU a luz das es TRE las (III)
 . X . x . . X .

10.

eu SOU a cor do lu AR (III)
 . X . x . . X

11.

eu SOU as coisas da VI da (III)
 . X . x . . X .

12.

eu SOU o medo de a MAR (III)
 . X . x . x . X

13.

eu SOU o medo do FRA co (III)
 . X . x . . X .

14.

a FORça da i ma gi na ÇÃO (III)
 . X X

15.

o BLEfe do jo ga DOR (III)
 . X X

16.

eu SOU o seu sa cri FÍ cio (III)
 . X X .

17.

a PLAcA de contra MÃO (III)
 . X . . x . X

18.

o SANgue no o lhar do vam PI ro (III)
 . X . . . x . . X .

19.

eu SOU a ve la que a CEN de (III)
 . X . x . . . X .

20.

eu SOU a luz que se a PA ga (III)
 . X . x . . . X .

21.

eu SOU a beira do a BISmo (III)
 . X . x . . . X .

22.

por QUE você me per GUNta (I)
 . X . x . . X .

23.

perGUNtas não vão lhe mos TRAR (I)
 . X . . x . . X

24.

do FOgo da água_e do AR (I)
 . X . . x . . X

25.

voCÊ me tem todo DI a (II)
 . X . x x . X .

26.

mas não SA be se_é bom ou ru IM (II)

. . X . . x x . . X

27.

mas SAI ba que_eu_estou em vo CÊ (II)

. X x . . X

28.

mas vo CÊ não_es tá em MIM (II)

. . X . . x . .

29.

das TELhas eu sou o te LHAdo (III)

. X . . x . . X .

30.

a PESca do pesca - DOR (III)

. X X

31.

dos SO nhos eu sou o_a Mor (III)

. X . . x . . X

32.

eu SOU a dona de CAsa (III)

. X . x . . X .

33.

nos PEgue pa gues do MUNdo (III)

. X . x . . X .

34.

eu SOU a mão do car RASco (III)

. X X .

35.

36.

37.

4.1.3 Corcovado

1.

2.

3.

4.

5.

6.

da ja ne la VÊ - se_o cor co VA do_o Reden TOR que lin do
 . . x . X . . . X . . . X . x .

7.

quero_a vi da SEMpre_as sim (I)
 x . . x . X . . x

8.

com você perto de mim a TÉ o a pa GAR da velha CHA ma (I-C)
 . . x x . . . X . . . X . x . X .

9.

e eu que e ra TRIS te (II)
 . . . x . X .

10.

descrente deste MUN do (II-A)
 . x . x . X .

11.

ao encontrar vo CÊ eu conhe CI (II-B)
 . . . X . x . . . X .

12.

o que_é fe li ci DA de meu a MOR (II-C)
 . . x . . . X . . . X .

4.1.4 Águas de Março

1.

é pau, é PEdra é o fim do caminho (III)
 X x x X . x . x . . x .

2.

3.

4.

5.

6.

7.

8.

9.

10.

11.

12.

13.

14.

15.

16.

17.

18.

festa da cumeeira (I)

x x .

19.

é a chuva chovendo (I)

x . x . . x .

20.

das águas de Março (I)

. x . . x .

21.

é_o pé é_o chão, é a marcha estradeira (III)

x . x x . . x . x . . . x .

22.

passa rinho na mão (I)

. . x . . x

23.

pedra de_atiradeira (I)

x x .

24.

uma ave no céu (I)

x . x . . x

25.

uma ave no chão (I)

x . x . . x

26.

é um regato_é uma fonte (I)

x . . x . x x . x .

27.

é o fundo do poço (I)

x . x . . x .

28.

é o fim do caminho (I)

x . x . . x .

29.

no rosto o desgosto (I)

x x . . . x .

30.

é um estrepe, é um PREgo, é uma ponta, é um ponto (III)

x . . x . . . X . x x . x . x . x .

31.

é um pingo pingando (I)

x . x . . x .

32.

é uma com ta_é um conto (II)

x x . x . x . x .

33.

é_o ti jolo chegando (I)

x . . x . . x .

34.

35.

36.

37.

38.

39.

40.

41.

42.

43.

4.2 Discussão

Pode-se observar que, na canção Gita, todos os tempos fortes são cantados com sílabas tônicas, sendo o exemplo de maior coincidência entre as estruturas musical e lingüística. A canção Aluga-se também apresenta, em sua maioria, coincidências, mas apresenta já alguns conflitos. Nas sentenças 7, 8 e 9 tais conflitos ocorrem devido à presença de síncopas na melodia, todas cantadas com sílabas tônicas. Trazemos abaixo a sentença 8:

Na sentença 10 de Aluga-se ocorre também um conflito entre música e língua, pois o primeiro tempo do segundo compasso é cantado com sílaba átona:

Na canção Corcovado, observamos 3 casos de conflito, sendo um deles devido a presença de síncopa, sentença 3, na qual a sílaba tônica de *ama* é cantada com síncopa:

pra fa zer Fe LIZ a quem se a ma (I-A)

. . x . X . . . x .

Os outros dois casos de conflito nesta canção ocorrem devido à ocorrência de sílaba átona no primeiro tempo do compasso:

um canti nho_um vi o lã o (I)

. . x x

e:

com você perto de mim a TÉ o a pa GAR da velha CHA ma (I-C)

. . x x . . . X . . . x . X .

No primeiro dos exemplos acima, como trouxemos na introdução do capítulo, o primeiro tempo do segundo compasso é cantado com sílaba átona, *vi*, de violão, assim como no segundo exemplo, no qual a sílaba átona de *perto* é cantada em tempo forte.

Na canção Águas de Março encontramos o maior número de conflitos, 40, sendo apenas 3 casos coincidentes. Os conflitos ocorrem devido à presença de síncopa, que, além de deslocarem o acento, obrigam o cantor a cantar duas sílabas, uma átona e uma tônica, justamente no tempo sincopado, como ocorre com as sílabas *zi* e *nho*, de sozinho; ou *vi* e *dro*, de vidro:

é um pouco sozinho (I)

x . x . . x .

é um ca co de vidro (II)

x . x . . x .

Buscamos mostrar neste capítulo alguns conflitos possíveis entre língua e música. Como colocamos no início do capítulo, acreditamos que a presença destes conflitos tenha relação com a tarefa de sincronização, deixando-a mais difícil de ser realizada. Não tratamos essas observações acentuais a partir da estatística por não ser aqui o caso, já que nossa intenção foi apresentar, de um ponto de vista qualitativo, a relação entre acento musical e lingüístico. Assim, contamos, no total, com 43 conflitos nas canções sincopadas e 4 conflitos nas canções não sincopadas.

O tratamento estatístico-descritivo dos dados coletados na pesquisa são apresentados no próximo capítulo.

5 - Estatística Descritiva

Apresentamos neste capítulo a estatística descritiva da amostra de dados que coletamos no experimento, ou seja, das *lags*²⁸.

Para cada um dos dois níveis de cada fator utilizamos 55 sentenças. Exemplificamos com o fator Ritmo da Canção:

Tabela 4: Quantidade de sentenças por canção

Nível <i>Ritmo Não-Sincopado</i>	Nível <i>Ritmo Sincopado</i>
Aluga-se: 18 sentenças	Águas de Março: 43
Gita: 37 sentenças	Corcovado: 12

Multiplicando-se esse valor, 55, pelo número de repetições, 3, pela quantidade de duplas, 4, e pela quantidade de pontos de observação de sincronização em cada sentença, 2, obtêm-se 1320 *lags* para cada nível do fator Ritmo da Canção. Multiplicando-se esse resultado pela quantidade de níveis do fator, 2, obtêm-se 2640 *lags* no total. O mesmo cálculo pode ser feito com o fator Ponto de Observação de Sincronização:

$$55 \times 3 \times 4 \times 2 = 1320 \times 2 = 2640$$

5.1 Descrição geral dos dados

Inicialmente, trazemos a estatística descritiva geral dos dados. Buscamos assim apresentar o comportamento geral das *lags*, resultado da tarefa de sincronização, comportamento este que servirá de base para a comparação entre os fatores e seus níveis.

A tabela a seguir traz, portanto, a média e o desvio-padrão de todos os dados, além da mediana e do valor máximo de *lag* encontrado, todos em milissegundos:

²⁸ Para retomar o conceito de lag, ver item 3.53. Rotulagem, às páginas 47-49.

Tabela 5: Estatística dos dados gerais

Média	Desvio Padrão	Mediana	Máximo
50	72	32	1687,5

Observa-se que o valor médio de *lag* é de 50 ms e o desvio-padrão é de 72 ms. Essa média representa o valor médio de *lag* para as quatro canções, calculada usando-se tanto *lags* obtidas nos pontos iniciais quanto nos finais. O mesmo foi feito para as outras medidas. Observamos ainda que a maior *lag* foi de 1687,56 ms.

Pode-se observar também o valor 32 ms para a mediana, valor notadamente mais baixo que a média. Ao trazer a mediana dos dados, buscamos apresentar um valor menos influenciado por *lags* muito longas, como o máximo apresentado acima, já que a mediana indica a *lag* que divide o conjunto de dados em duas metades com igual quantidade de *lags*. Valores muito altos de *lags* também influenciam de forma visível no desvio-padrão, outra razão pela qual trazemos a mediana. Assim, tem-se 1320 *lags* abaixo de 31,86 ms e outras 1320 acima de 32 ms.

Colocamos como um dos objetivos deste capítulo apresentar a distribuição dos dados, tanto dos dados gerais quanto dos separados por fatores, a fim de apresentar como a fala cantada se comporta na tarefa de sincronização. A tabela de percentis e o gráfico de frequências a seguir ilustram a distribuição dos dados gerais:

Tabela 6: Tabela de Percentis dos dados gerais

10 ° Percentil	25° Percentil	50° Percentil	75° Percentil	90° Percentil
5,7	14,5	32	64,5	110,7

A tabela indica que 10% dos dados (264 *lags*) apresentam valor menor ou igual a 5,7 milissegundos; 25% (1° quartil; 660 *lags*) apresentam valor menor ou igual a 14,5 ms; 50% (2° quartil, 1320 *lags*) apresentam valor menor ou igual a 32, como já indicou a mediana; 75% (3° quartil, 1980 *lags*) apresentam valor menor ou igual a 64,5 milissegundos e, Finalmente, que 90% dos dados (2376 *lags*) apresentam valor igual ou abaixo dos 110,7

milissegundos²⁹. O gráfico de frequências traz a distribuição de ocorrências das *lags*, em blocos de 6 milissegundos. Excluímos da contagem no gráfico as *lags* com valores maiores que 264 milissegundos, pois a apresentação delas no gráfico inviabilizaria a visualização do eixo y. Deste gráfico foram excluídas 24 *lags*, distribuídas entre 265 ms e 1687, a maior coletada:

Figura 33: Distribuição dos dados gerais. A abscissa indica os blocos dos valores das *lags* em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de *lags* nos determinados blocos.

²⁹ Nos 25% citados, estão inclusos os primeiros 10%. Nos 50% citados estão inclusos os 25% e os 10%, e assim por diante.

Observando-se o gráfico da figura 33 juntamente com a Tabela 6 pode-se ver que, apesar de 90% dos dados serem *lags* com valores abaixo de 110 ms, ainda conta-se uma grande quantidade (aproximadamente 220 *lags*, cerca de 8% dos dados) com valores superiores a 110 ms, razão pela qual não tomamos *lags* superiores a esse valor como *outliers*. A grande quantidade e a natureza desses dados analisados é também outro motivo pelo qual não tomamos essas *lags* como *outliers*. Obtivemos, em ambos os fatores, valores mais altos do que o máximo apresentado. Estes foram excluídos por serem momentos de hesitação ou de “perda da deixa” na realização da tarefa, momentos nos quais não houve interação entre os osciladores envolvidos. Esses foram considerados *outliers* e excluídos dos cálculos estatísticos.

Trazemos abaixo um exemplo do sinal acústico de cada tipo de *lag*: um exemplo de lag em torno da média, um de sincronia perfeita e um de *outlier*:

Figura 34: Exemplo de *lag* em torno da média. A faixa rosada indica justamente a *lag*, de 48 milissegundos.

Figura 35: Exemplo de sincronia perfeita, ou seja, *lag* igual a zero, no *onset* de *viu*.

Figura 36: Exemplo de *outlier*. As setas indicam os pontos iniciais. A diferença entre eles é de 1087 milissegundos e representa uma perda de deixa por parte de um dos cantores.

5.2 Descrição do fator Ritmo da canção

Em nosso experimento, o fator Ritmo da Canção apresenta dois níveis: Ritmo Sincopado (RS) e Ritmo Não-Sincopado (RNS). A Tabela 7 traz a estatística descritiva desses dois níveis:

Tabela 7: Estatística descritiva do fator Ritmo da Canção

	Média	Desvio Padrão	Mediana	Máximo
RS	55	82,5	34	1687,56
RNS	45	59	29,5	1027,18

Pode-se observar que as medidas apresentam valores maiores no nível ritmo sincopado (RS) quando comparadas ao nível ritmo não-sincopado (RNS). Enquanto no nível RNS a média das lags é de 45 ms, no nível RS a média é de 55 ms, o que indica maior dificuldade na realização da tarefa de sincronização no nível RS, indo de acordo com nossa primeira hipótese.

Os valores de desvio padrão são 82,5 e 59 para os níveis RS e RNS, respectivamente. Essa diferença entre os níveis indica maior amplitude de variação em torno da média no nível RS em relação ao nível RNS. Em outras palavras, a menor variação na realização da tarefa indica que o processo de sincronização é mais estável no nível RNS, resultado próximo ao

que observamos em nosso estudo piloto (Santos; Medeiros, 2010). Com tais valores ainda, baseados em suas respectivas médias, chega-se a amplitudes máximas em torno da média de 137,5 ms para RS e 104,5 ms para RNS³⁰.

O gráfico abaixo ilustra essas medidas (média, desvio-padrão e mediana) e a comparação com os valores referentes aos dados gerais:

Figura 37: Gráfico descritivo do fator Ritmo da Canção

Observa-se que a média, o desvio-padrão e a mediana no nível RS são superiores aos valores dessas mesmas medidas no nível RNS e também aos valores dos dados gerais.

A tabela de percentiis e os gráficos de frequência abaixo ilustram a distribuição desses dados:

³⁰ Amplitudes em torno da média a partir do desvio-padrão geralmente são representadas como intervalos. No nosso caso, representaríamos da seguinte forma: [137,54; -27,68] para S e [104,53; -13,79] para NS. Não representamos assim, pois não existem *lags* negativas em nosso e experimento, o que pressuporia um líder na realização da tarefa, descaracterizando os cantores como osciladores acoplados.

Figura 38: Distribuição dos dados do nível RNS. A abscissa indica os blocos dos valores das *lags* em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de *lags* nos determinados blocos.

Figura 39: Distribuição dos dados do nível RS. A abscissa indica os blocos dos valores das *lags* em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de *lags* nos determinados blocos.

Tabela 8: Tabela de percentiis dos dados do fator Ritmo da Canção

	10° Percentil	25° Percentil	50° Percentil	75° Percentil	90° Percentil
RNS	5	13	29,5	61	102
RS	7	16	34	68	118,5

Observando-se os gráficos das figuras 38 e 39 e a Tabela 8 nota-se que a distribuição dos dados nos dois níveis, RS e RNS, apresentam padrão semelhante, descendente. Na tabela 6 observa-se que, para todos os percentiis apresentados, os valores do nível RS são superiores ao do nível RNS.

Nesta tabela, 10% dos dados representam 132 *lags*, 25 % representam 330 *lags*, 50% representam 660 *lags*, 75% representam 990 *lags* e 90% representam 1188 *lags*.

5.3 Descrição do fator Ponto de Observação de Sincronização

O fator Ponto de Observação de Sincronização também apresenta dois fatores, denominados ponto inicial e ponto final. A Tabela 9 traz a média e o desvio padrão desses níveis:

Tabela 9: Estatística descritiva do fator Ponto de Sincronização

	Média	Desvio Padrão	Mediana	Máximo
PI	63,5	84,7	45	1687,5
PF	36,5	53	24	1027

Observa-se médias com valores de 63,5 e 36,5 no nível PI e PF, respectivamente, o que vai de acordo com nossa segunda hipótese, já que a média nos pontos iniciais é maior que a dos pontos finais, indicando maior dificuldade de realizar a tarefa de sincronização nos pontos iniciais das sentenças. O desvio-padrão também apresenta valor maior para o nível PI, 84,7, contra 53 para o nível PF, mostrando maior estabilidade de sincronização no nível PF. A partir dos desvios-padrão obtidos, pode-se chegar à amplitude de variação em torno da média de 148,5 para PI e 106 para PF. O gráfico abaixo ilustra a apresentação desses dados:

Figura 40: Gráfico descritivo do fator Ponto de Observação de Sincronização

Para ilustrar a distribuição dos dados, trazemos os gráficos de frequências dos dois níveis do fator em questão e a tabela de percentis.

Figura 41: distribuição dos dados do nível Ponto inicial. A abscissa indica os blocos dos valores das *lags* em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de *lags* nos determinados blocos.

Figura 42: Distribuição dos dados do nível Ponto Final. A abscissa indica os blocos dos valores das *lags* em milissegundos, agrupadas a cada 6 milissegundos; a ordenada indica a frequência de ocorrências de *lags* nos determinados blocos.

Tabela 10: Tabela de percentis dos dados do fator Ponto de sincronização

	10° Percentil	25° Percentil	50° Percentil	75° Percentil	90° Percentil
PI	8,5	21	45	84	132,25
PF	4,5	11,5	23,5	45	73,7

5.4 Sincronias perfeitas

Contabilizamos, também, as quantidades de sincronias perfeitas, ou seja, de lags iguais a zero. A tabela abaixo indica esse cálculo:

Tabela 11: Quantidade de sincronias perfeitas

Total	RS	RNS	PI	PF
65	26	39	28	37

Nota-se que, indo de acordo com nossas primeira e segunda hipóteses, encontramos maior quantidade de sincronias perfeitas nos níveis RNS e PF, indicando maior facilidade de realização da tarefa nesses níveis, já que pedimos aos sujeitos que cantassem o “mais sincrônico possível”.

5.5 Observações iniciais acerca da interação entre os fatores

Como se observou nos gráficos de frequência apresentados, as amostras colhidas para o experimento não apresentam distribuição normal, apresentando distribuição exponencial decrescente, o que torna mais complexa a realização de uma análise de variância (ANOVA), motivo pelo qual não apresentamos aqui tal análise.

Com a intenção ainda de apresentar os resultados das combinações entre os fatores e seus níveis, trazemos a tabela a seguir:

Tabela 12: Combinação entre os níveis dos dois fatores

		Ritmo	
		RS	RNS
Ponto	Inicial	69,8	57,3
	Final	40	33,5

Confirmando nossas hipóteses, nota-se que a maior média é a encontrada na combinação dos níveis RS e PI, e que a menor delas é a referente à combinação dos níveis RNS e PF. Esses resultados indicam maior facilidade de sincronização nos pontos finais das canções não sincopadas e maior dificuldade nos pontos iniciais das canções sincopadas.

A existência de interação entre os fatores pode ser observada no gráfico abaixo:

Figura 43: Gráfico da existência de interação entre os fatores

O fato de as linhas azul e vermelha não se cruzarem indica a não existência de interação entre os fatores (Calegare, 2001), ou seja, os níveis de um fator se comportam de maneira semelhante quando combinados aos níveis do outro fator³¹.

O efeito principal do fator Ritmo da Canção é 9,5 e do fator Ponto Observação de Sincronização é -26,8, ou seja, do nível RNS para o RS a resposta aumenta em 9,5 ms, enquanto do nível PI para PF a resposta diminui em 26,8 ms.

³¹ A existência de interação entre os fatores poderia, por exemplo, indicar médias menores no nível S combinado ao nível PI quando comparado ao PF, indicando diferença de comportamento do ritmo sincopado entre os níveis Inicial e Final.

6 - Conclusão

A natureza do objeto tratado neste trabalho nos exigiu, mas também podemos dizer que nos permitiu, observar a fala de um ponto de vista dinâmico. A fala cantada, realizada de acordo com critérios não só lingüísticos, mas também musicais, permitiu uma observação rítmica da língua ainda muito pouco presente na ciência, observação essa que traz sua contribuição para o estudo do ritmo das línguas e também para o estudo de alguns aspectos cognitivos tanto da fala quanto do canto.

Buscamos mostrar como a estrutura musical interfere na realização da fala cantada, colocando a mesma ora inserida em um contexto no qual o ritmo de acompanhamento fosse sincopado, o da bossa-nova, e ora em um contexto no qual ritmo de acompanhamento não fosse sincopado, o do rock. Em paralelo, estudamos o comportamento do processo de sincronização em dois pontos diferentes das sentenças das canções, buscando observar se, ou como, a estrutura musical dá igual suporte à sincronização nos inícios e nos finais de sentença.

Observando os resultados do experimento, podemos concluir que, ritmicamente, a fala cantada inserida no contexto do rock é mais fácil de ser cantada quando comparada à fala cantada inserida no contexto da bossa-nova. Todos os cálculos estatísticos realizados corroboram com essa nossa primeira hipótese, ou seja, média, desvio-padrão e mediana apresentam valores mais elevados no nível RS do que no nível RNS, como pode-se ver na Figura 44:

Figura 44: Resultados do experimento

Em outras palavras, podemos concluir que o conflito entre as estruturas métrica e rítmica, a síncopa, acaba por dificultar a realização da tarefa de sincronização. Acreditamos que tal dificuldade se deva ao fato de o ouvinte se confundir ou perder, de certo modo, as referências métricas e rítmicas da estrutura musical: os acentos.

Buscamos mostrar ainda que nas canções sincopadas há ainda um segundo tipo de conflito, não mais entre duas estruturas musicais paralelas, mas entre música e língua. Observando as coincidências entre os acentos métricos e os acentos lingüísticos, lexicais, vimos que nas canções sincopadas quase não houve coincidências, sendo toda a canção Águas de Março estruturada em cima deste conflito entre língua e música, conflito este quase não presente nas canções do gênero rock.

Vendo o ritmo de acompanhamento de uma canção como guia para a realização da tarefa, já que ele está presente do início ao fim das canções, de forma contínua, seja em momentos de canto ou de silêncio, buscamos também observar como a fala cantada se guia ritmicamente a cada nova sentença a ser realizada. Em outras palavras, buscamos ver se a estrutura musical guiava da mesma forma a sincronização nos inícios e nos fins das sentenças. Hipotetizamos, apesar de o ritmo de acompanhamento de uma canção ser algo contínuo ao longo de toda a sua execução, que os inícios de sentenças seriam mais difíceis de serem sincronizados quando comparados aos finais das mesmas, hipótese esta inspirada nos estudos de Fred Cummins a respeito da fala falada. Nossa amostra nos permite concluir que realmente há diferenças de sincronização entre os pontos iniciais e os pontos finais.

Acreditamos que música e fala (língua) sejam estruturas com igual “força” para guiar os sujeitos na realização da tarefa de sincronização, e podemos, assim, dizer que nos pontos iniciais houve menor sincronização devido ao silêncio da fala entre uma sentença e outra, o que obrigava os sujeitos a iniciar um novo processo de sincronização a cada sentença, pois o silêncio na fala rompe o elo entre os cantores.

Assim, podemos concluir que houve na realização da tarefa dois processos de sincronização andando em paralelo: um deles colocava os sujeitos num universo em que as canções eram vistas como um todo e outro no qual as canções eram fragmentadas em sentenças, no qual cada sentença era tomada como uma tarefa distinta, percebida como um universo distinto. A música é o guia no primeiro e a língua é o guia no segundo.

7- Referências Bibliográficas

ABERCROMBIE, D. **Elementals of General Phonetics**. Edimburgo: Edinburgh University Press, 1967

ADOLFO, A. **Harmonia e estilos para teclado**. Rio de Janeiro : Lumiar, 1994

ALBANO, E. C. **O Gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro**. Campinas-SP : Mercado de Letras, ALB São Paulo: FAPESP, 2001

AMMER, C. **The Harper Dictionary Of Music**. Harper e Row, 1987

ANDRADE, M. **Introdução a estética musical**. São Paulo : Ucitec, 1995

BECHARA, E. **Moderna Gramática Portuguesa**. Rio de Janeiro, RJ : Lucerna : Nova Fronteira, 2009

BARBOSA, P. A. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on segmental duration generation. **Cadernos de Estudos Lingüísticos**, 31: 33-53. 1994

BARBOSA, P. A. (2000) “Syllable-Timing in brazilian portuguese: uma crítica a Roy Major” **D.E.L.T.A. Vol. 16, N.º 2**, 2000 (369-402)

BARBOSA, P. A. É possível integrar o discreto e o contínuo em um modelo de produção da fala? **Cadernos de Estudos Lingüísticos**, 40, 29-38. 2001

BARBOSA, P. A. **Incursões em torno do ritmo da fala**. Campinas: Pontes Editores, São Paulo : FAPESP, 2006

BOERSMA, P; WEENINK, D. **Doing phonetics by computer**. Versão 5.2.22. Software livre baixado em 31 de Maio de 2011 de <http://www.praat.org/>.

BROWMAN, C. P.; GOLDSTEIN, L. M. Towards an articulatory phonology.

Phonology Yearbook, v. 3, p. 219-252, 1996

CAGLIARI, L. C. **Elementos de Fonética do Português Brasileiro**. Campinas: UNICAMP (Tese de Livre-Docência), 1981.

CAGLIARI, L. C.; ABAURRE, M. B. M. **Elementos para uma investigação instrumental das relações entre padrões rítmicos e processos fonológicos no português brasileiro**. *Cadernos de Estudos Linguísticos 10*. Campinas: UNICAMP/IEL, 1986.

CALEGARE, A. J. A. **Introdução ao delineamento de experimentos**. São Paulo : Edgard Blücher, 2001

CHEDIAK, A. **Songbook Tom Jobim Vol. II**. Rio de Janeiro: Lumiar Editora, 1994a

CHEDIAK, A. **Songbook Tom Jobim Vol. III**. Rio de Janeiro: Lumiar Editora, 1994b

CHEDIAK, J. **As 101 melhores canções do século XX**. Seleção Almir Chediak. Rio de Janeiro: Lumiar Editora, 2004

CLAYTON, M; SAGER, R.; WILL, U. **In time with the music: The concept of entrainment and its significance for ethnomusicology**. ESEM CounterPoint, Vol.1, 2004

COOPER, G.; MEYER L. B. **The Rhythmic Structure of Music**. Chicago: The University of Chicago Press, 1960

COOPLAND, A. **Como ouvir e entender a música**. Rio de Janeiro: Artenova, 1974

COSTA, E. R. C.; JOBIM, A. C. **Elis e Tom**. Phonogram, 1974. CD

CUMMINS, F. On synchronous speech. **Acoustic Research Letters Online**, 3(1):7–11, 2002

_____, F. Practice and performance in speech produced synchronously.

Journal of Phonetics, 31(2):139–148, 2003

_____, F. Measuring Synchronization Among Speakers Reading Together. **Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics**, 28-30. Atenas, Grécia, 2006

_____, F. Rhythm as Entrainment: The Case of Synchronous Speech. **Journal of Phonetics**, 37(1):16-28. Pré-impressão. Elsevier, 2009a.

_____, F. Rhythm as an affordance for the entrainment of movement. **Phonetica**, 66(1–2):15–28, 2009b

DAUER, R. M. Stress-timing end syllable-timing reanalyzed. **Journal of Phonetics**, 11:51-62, 1983

FERRIS, J. **Music: the art of listening**. New York, McGraw-Hill Companies, 2003

FRAISSE, P. *Rhythm and tempo*. Em D. Deutsch (Org.), **The Psychology of music** (pp. 149-180). Nova York : Academic Press, 1982

GOLDSMITH, J. **Autosegmental Phonology** Tese (Doutorado) 1976a

GOLDSMITH, 1976b fonologia autosegmental

GUENTHER, F. Neural control of speech movements. Em: **SCHILLER, N.; MEYER, A., Phonetics and phonology in language comprehension and production: differences and similarities**. Berlim : Mouton de Gruyter, 2003 Nova York.

HAYES, B. *A Metrical Theory of Stress Rules*, Garland Press, New York. 1980

_____, B. **The prosodic hierarchy in meter** in Paul Kiparsky and Gilbert Youmans, eds., *Rhythm and Meter* , Academic Press, Orlando, FL, pp. 201-260.

HAYES, B. *Metrical stress theory: principles and case studies*, Chicago : University of Chicago Press, 1995

HASKINS LABORATORIES "**The McGurk Effect: Hearing lips and seeing voices**"

Visitado em 2 de Outubro de 2011. <http://www.haskins.yale.edu/featured/heads/mcgurk.html>

HOEL, P. G. **Estatística Elementar**. São Paulo : Atlas, 1979

KELSO, J. A. S. **Dynamic Patterns**. MIT Press, Cambridge, MA, 1995

KRUMHANS, C. L. *Ritmo e altura na cognição musical*. In: ILARI, B. S. (org.) **Em busca da mente musical: ensaios sobre os processos cognitivos em música – da percepção a produção**. Curitiba: Editora UFPR, 2006.

LERDHAL, F.; JACKENDOFF, R. *A Generative theory of tonal music*. Cambridge, Mass: MIT Press, 1983

LIBERMAN, A. M.; MATTINGLY, I. G. The motor theory of speech perception revised. **Cognition**, 21. 1-36, 1985

LIBERMAN, M.. PRINCE, A. **On Stress and linguistic rhythm**. 1977

LUFT, P. C. *Moderna Gramática Brasileira*. Porto Alegre : Globo, 1978.

MAJOR, R. C. Stress-timing in Brazilian Portuguese. **Journal of Phonetics**, v. 9 n°3: 343-352, 1981

MASSINI-CAGLIARI, G. **Acento e ritmo**. São Paulo : Contexto, 1992 (Coleção repensando a língua portuguesa)

MCGURK H., MACDONALD J. (1976). "Hearing lips and seeing voices." *Nature* **264** (5588): 746–8.

MAYBERRY, R.I., JAQUES, J., AND DEDE, G. What stuttering reveals about the development of the gesture-speech relationship. Em J.M. Iverson and S. Goldin-Meadow (Org.), *The Nature and Function of Gestures in Children's Communication*. **New Directions for Child Development**, n.79. São Francisco: Jossey-Bass. 1998

OLIVEIRA, J. G. P. P. **O Amor, o sorriso e a flor**. Odeon, 1960. CD

PESSOTTI, A. C. S. **Procedimento semi-automático de segmentação fonética via Algoritmo de Alinhamento Temporal Dinâmico (DTW)**, 2010. (produção técnica sem patente)

PIKE, K. **Phonemics: a technique for reducing languages to writing**. Ann Arbor: The University of Michigan Press. 1971.

PRIOLLI, M. L. M. **Princípios básicos da música para a juventude**. 1º Volume. Rio de Janeiro : Casa Oliveira de Músicas Ltda. 1987

PORT, R.F.; TAJIMA, K.; CUMMINS, F. Self-Entrainment in Animal Behavior and Human Speech. **Online Proceedings of the 1996 Midwest Artificial Intelligence and Cognitive Science Conference**. 1996

RAPOSO DE MEDEIROS, B. R. **Descrição comparativa de aspectos fonético-acústicos selecionados da fala e do canto em português brasileiro**. Tese. Campinas : UNICAMP, 2002

REPP, B. Effects of three kinds of accent on synchronization with simple rhythms. Em: **Proceedings of the 8th international conference on music perception and cognition**. Evaston, 2004

SANTOS, C. A. A.; MEDEIROS, B. R. Interações entre ritmo linguístico e ritmo musical na canção. Em : **Anais do VI SIMCAM, Simpósio de Cognição e Artes Musicais**. Rio de Janeiro, 2010.

SANTOS, C. A. A.; MEDEIROS, B. R; PESSOTTI, A. C. S.. Para medir a sincronização de dois cantores: o caso da bossa-nova. Em: **Anais do VII SIMCAM, Simpósio de Cognição e Artes Musicais**. Brasília, 2010.

SCHÖNBERG, A. **Fundamentos da Composição Musical**. São Paulo : Edusp, 1996

SCHÖNBERG, A. **Harmonia**. São Paulo : Editora Unesp, 1999

SEIXAS, R. S. **Gita**. Phillips, 1974. *Compact Disc* (CD)

SEIXAS, R. S. **Abra-te Sésamo**. CBS. 1980 *Compact Disc* (CD)

TATIT, L. **O cancionista: composição de canções no Brasil** São Paulo : Edusp, 1996

_____ **O Século da canção** Cotia : Ateliê Editorial, 2004.

TEMPERLEY, D. **The cognition of basic musical structures**. Cambridge, Mass. : MIT Press, 2001.

ZAMPRONHA, M. L. S. **Da música, seus usos e recursos**. 2. ed. rev. e ampliada. São Paulo : Editora UNESP, 2007

8 – Apêndice 1 – Segmentação das canções

Apresentamos a seguir a segmentação completa das canções. Trazemos apenas as sentenças utilizadas na pesquisa.

Os asteriscos indicam a localização dos pontos de observação de sincronização.

Aluga-se

A solução pro nosso povo eu vou dá
 *Negócio bom assim ninguém nunca *viu
 *Tá tudo pronto aqui é só vim *pegar
 *A solução é alugar o *Brasil!...
 *Nós não vamo paga *nada
 *Nós não vamo paga nada é *tudo free!
 *Tá na hora ago*ra é free
 *Vamo embora *dar lugar
 *Esse imóvel tá *prá alugar
 *Os estrangeiros
 Eu sei que eles vão gos*tar
 *Tem o Atlântico tem vista pro *mar
 *A Amazônia é o jardim do *quintal
 *E o dólar dele paga o nosso *mingau...
 *Nós não vamo paga *nada
 *Nós não vamo paga nada é *tudo free!
 *Tá na hora ago*ra é free
 *Vamo embora *dar lugar
 *esse imóvel está *pra alugar

Gita

*Às vezes você me per*gunta
 *Por que é que eu sou tão cala*do,
 *Não falo de amor qua*se nada,
 *Nem fico sorrindo ao *teu lado.
 *Você pensa em mim *toda hora.
 *Me come, me cospe, me *deixa.
 *Talvez você não en*tenda,
 *Mas hoje eu vou lhe mos*trar.
 *Eu sou a luz das es*trelas;
 *Eu sou a cor *do luar;
 *Eu sou as coisas da *vida;
 *Eu sou o medo *de amar.
 *Eu sou o medo *do fraco;
 *A força da imagina*ção;
 *O blefe do joga*dor;
 *Eu sou o seu sa*crifício;
 *A placa de *contra-mão;
 *O sangue no olhar do *vampiro
 *Eu sou a vela que a*cende;
 *Eu sou a luz que se a*paga;

*Eu sou a beira do a*bismo;
 *Por que você me *pergunta?
 *Perguntas não vão lhe *mostrar
 *Do fogo, da água e *do ar!
 *Você me tem todo *dia,
 *Mas não sabe se é bom ou *ruim.
 *Mas saiba que eu estou em *você,
 *Mas você não es*tá em mim.
 *Das telhas eu sou o *telhado;
 *A pesca do *pescador;
 *Dos sonhos eu *sou o amor
 *Eu sou a dona de *casa
 *Nos pegue pagues *do mundo;
 *Eu sou a mão do *carrasco;
 *Eu sou a mosca da *sopa
 *E o dente do tu*barão;
 *Eu sou os olhos do *cego

Corcovado

*Um cantinho e um *violão
 *Este amor, uma can*ção
 *Pra fazer feliz a *quem se ama
 *Muita calma pra pen*sar
 *E ter tempo pra so*nhar
 *Da janela vê-se o Corcovado o Redentor que *lindo
 *Quero a vida sempre a*ssim
 * com você perto de mim até o apagar da velha *chama
 *E eu que era *triste
 *Descrente deste *mundo
 *Ao encontrar você eu *conheci
 *O que é felicidade *meu amor

Águas de Março

*É pau, é pedra, é o fim *do caminho
 *É um resto *de toco
 *é um pouco so*zinho
 *É um caco *de vidro,
 *é a vida, é o *sol
 *É a noite, é a *morte
 *é o laço, é o an*zol
 *É peroba do *campo,
 é o *nó da ma*deira
 *Caingá, candeia, é o Matita *Pereira
 *É madeira *de vento,
 *tombo da ri*banceira
 *É o mistério *profundo,
 *é o queira ou não *queira
 *É o vento ven*tando,

*é o fim da la*deira
 *É a viga, é o *vão,
 *festa da cu*meeira
 *É o pé, é o chão, é a marcha estra*deira
 *Passarinho na *mão,
 *pedra de atira*deira
 *uma ave no *céu,
 é u*ma ave no *chão
 *É o fundo *do poço,
 * é o fim do *caminho
 *No rosto o des*gosto
 *É um estrepe, é um prego, é uma ponta, é um *ponto
 *É um pingo *pingando,
 *é uma conta, é um *conto
 *é o tijolo che*gando
 *É a lenha, é o *dia,
 *é o fim da *picada
 *É a garrafa de *cana,
 *o estilhaço na es*trada
 *É um passo, é uma ponte, é um sapo, é uma *rã
 *É um resto de *mato,
 *na luz *da manhã
 *São as águas de março fechando o *verão
 *É a promessa de vida no teu *coração
 *É uma cobra, é um *pau,
 *é João, é *José
 *É um espinho *na mão,
 *é um corte no *pé
 *São as águas de março fechando o *verão,
 *É a promessa de vida no teu cora*ção

9 - Apêndice 2 – Grades métricas das frases musicais

Aluga-se

The musical notation for 'Aluga-se' is presented in six systems, each with a staff and corresponding guitar fret numbers (X) below it. The first system is in 4/4 time, and the others are in 2/4 time.

System 1 (4/4): Notes: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11.

System 2 (2/4): Notes: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7, 8.

System 3 (2/4): Notes: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11.

System 4 (2/4): Notes: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7.

System 5 (2/4): Notes: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7.

System 6 (2/4): Notes: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7, 8, 9.

Gita

The musical notation for 'Gita' is presented in one system, in 4/4 time. The notes are: quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter, quarter. Fret numbers: 1, 2, 3, 4, 5, 6, 7, 8, 9.

Corcovado

1 2 3 4 5 6 7

X X X X X X X
X X X
X X X

1 2 3 4 5 6 7

X X X X X X X
X X X
X X X

1 2 3 4 5 6 7 8 9 10

X X X X X X X X X X
X X X
X X X

X X X X X X X
X X X
X X X

Águas de Março

1 2 3 4 5 6

X X X X X X
X X X
X

1 2 3 5 6 7

X X X X X X
X X X

1 2 3 4 5 6 7 8 9 10

X X X X X X X X
X X X X X X
X X X X X X