

HIERARQUIA PROSÓDICA E HIERARQUIA MELÓDICA NA CANÇÃO *GABRIELA*¹

(Prosodic Hierarchy and Melodic Hierarchy in *Gabriela*)

José Roberto do CARMO JR.

(Departamento de Linguística – FFLCH
Universidade de São Paulo)

Raquel Santana SANTOS

(Departamento de Linguística – FFLCH
Universidade de São Paulo)

Resumo: *O presente artigo propõe uma investigação das relações de interface entre o componente linguístico – prosódico – e o componente musical na palavra cantada. Tomando como objeto a canção Gabriela, de Tom Jobim, investigamos as relações que se estabelecem entre a hierarquia prosódica (Nespor & Vogel, 1986) e a hierarquia melódica (Carmo Jr 2007). Mostramos que no caso de incompatibilidade entre a estrutura prosódica e melódica, os processos fonológicos do português brasileiro de elisão, degeminação, ditongação e retração acentual podem violar seus domínios prosódicos de aplicação, mas ainda assim são organizados de acordo com os domínios melódicos.*

Palavras-chave: *canção; prosódia; processos fonológicos; semiótica.*

Abstract: *In this article we investigate the relation between the linguistic (prosodic) component and the musical grammar in lyrics. Specifically, by analyzing Gabriela, by Tom Jobim, we discuss the relation between the Prosodic Hierarchy (Nespor & Vogel 1986) and the Melodic Hierarchy (Carmo Jr 2007). We show that when there is a mismatch between these two hierarchies, the phonological processes of elision, degemination, diphthongation and stress retraction in Brazilian Portuguese may not comply with the prosodic domain for their application, while still satisfying the restrictions related to the melodic domains.*

Key-words: *popular song; prosody; phonological process, semiotics.*

1. Agradecemos a dois pareceristas anônimos a discussão deste texto. Todos os problemas remanescentes são falta unicamente nossa. O primeiro autor agradece também à FAPESP (processo 2008/06656-7) e o segundo autor agradece à FAPESP (2006/00965-2) e ao CNPq (311041/2006-0) pelos auxílios à pesquisa.

1. INTRODUÇÃO

A teoria semiótica procura explicitar os mecanismos de geração de sentido dos textos, ou seja, ela procura revelar *o que* um texto diz e *como* ele faz isso. No modelo estabelecido por Tatit para a análise semiótica da canção (Tatit 1996), o procedimento básico consiste em verificar a ressonância maior ou menor entre o *conteúdo* linguístico (isotopia temático-figurativa, estrutura narrativa, tensividade) e a *expressão* melódica (duração das notas, tessitura e andamento).

O estudo da canção *Gabriela* que se segue trata o problema do sentido de uma maneira diferente, mas que vem complementar a proposta de Tatit. Investigaremos a compatibilidade entre letra e melodia comparando a expressão linguística (hierarquia prosódica) com a expressão melódica (hierarquia melódica) sem fazer nenhuma referência ao conteúdo do texto. Examinaremos sucintamente como se apresenta a hierarquia prosódica dos 20 versos que compõem a canção (§2) e como as seqüências melódicas associadas a estes versos se estruturam metricamente (§3), para em seguida observar algumas consequências sobre os processos fonológicos decorrentes da justaposição destes dois componentes na canção (§4). Finalizaremos o ensaio mostrando como os resultados aqui obtidos podem contribuir para o desenvolvimento de um modelo semiótico-prosódico da melodia.

2. O COMPONENTE LINGUÍSTICO (PROSÓDICO)

A proposta da teoria prosódica é que processos fonológicos não ocorrem obedecendo apenas a fatores segmentais, mas também são circunscritos a domínios prosódicos. A motivação da proposta de domínios é serem estes os *loci* onde determinados processos fonológicos ocorrem. A proposta mais corrente na literatura assume que, se um domínio existe numa língua qualquer (dado que foram encontrados processos fonológicos que nele ocorrem), ele existe em qualquer língua, mesmo que em uma outra língua ele pareça ser inoperante (cf. Nespor & Vogel 1986). Segundo Nespor & Vogel, as informações prosódicas estão organizadas em sete níveis ou domínios hierárquicos: sílaba (S), pé (F), palavra fonológica (W), grupo clítico (C), frase fonológica (PhP), frase entoacional (I) e enunciado (U). Os níveis acima da palavra fonológica (esta incluída) seriam construídos a partir de noções não-fonológicas, ou seja, levando em conta informações de outros compo-

nentes da gramática (morfologia, sintaxe, semântica). A palavra fonológica e o grupo clítico fariam uso de noções morfo-fonológicas; a frase fonológica usa de noções sintáticas; a frase entoacional usa de informações sintáticas e semânticas; e o enunciado usa de informações semântico-discursivas. Cada domínio pode incluir várias unidades do domínio inferior (no mínimo uma), como é possível observar na Figura (1) abaixo:

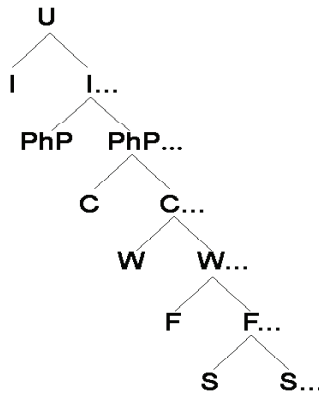


Figura 1: Hierarquia prosódica.

Esta é a proposta geral da hierarquia prosódica, mas assume-se que vários aspectos de sua aplicação são definidos por cada língua particular. Para nós, interessa-nos aqui apenas aspectos particulares do pé em português brasileiro. Segundo Nespor & Vogel, o pé é uma unidade que compreende uma e apenas uma sílaba forte e pode ter de nenhuma a várias sílabas fracas. Portanto, a escolha de quantas sílabas os pés terão e qual a posição da sílaba forte (no início ou no final) é uma escolha de cada língua. Assume-se que o português tem um ritmo binário (no máximo duas sílabas), com a sílaba forte à esquerda (portanto, um troqueu) (cf. (Carvalho 1989), (Collischon 1994), entre outros).²

2. Um dos pareceristas chama a atenção de que uma análise alternativa seria a de assumir o pé n-ário para o português e de que a proposta de Collischon é para explicar o acento secundário. De fato, os estudos sobre o pé no português brasileiro propõem diferentes análises. Estudos que buscam por correlatos acústicos do acento secundário argumentam não encontrar evidências que apontem para uma alternância binária (cf. Arantes & Barbosa 2002, Moraes 2003, Arantes 2005, por exemplo). Por outro lado, estudos que discutem processos fonológicos que ocorrem relacionados ao ritmo apontam para a alternância binária (cf. Abousalh 1997, Abaurre & Svartman 2008, entre outros).

Assume-se também que as línguas obedeçam ao Princípio de Alternância Rítmica (cf. (Selkirk 1984) e (Takezawa 1981)), segundo o qual, as línguas preferem sentenças o mais eurrítmicas possível. Para isso, ela usa de processos fonológicos (como a retração acentual, a haplologia, as regras de sândi) para otimizar o ritmo de suas sentenças. No entanto, essas regras não ocorrem sem restrições, elas só podem ocorrer dentro de domínios prosódicos específicos. A língua então sofre uma tensão: quer, a todo custo, sentenças ritmicamente ótimas e para isso aplica processos fonológicos; mas os processos têm que obedecer aos domínios prosódicos, o que significa que nem sempre as sentenças terão o melhor ritmo possível.

2.1. Processos fonológicos e hierarquia prosódica

Como dissemos no início desta seção, há alguns processos fonológicos que levam em conta informações prosódicas. Neste artigo, discutiremos alguns processos de sândi externo (especificamente a elisão, a degeminação e a ditongação) e a retração acentual. Por isto, nesta seção, resumimos brevemente as condições de aplicação dos mesmos em português brasileiro.

As regras de degeminação e elisão implicam a diminuição da quantidade de sílabas de um contexto. A degeminação vocálica ocorre entre duas vogais com as mesmas características segmentais, provenientes de diferentes sílabas, que resultam em uma sílaba CV. Segundo Bisol (1996), a degeminação envolve uma fusão das vogais, que pode ser seguida de um encurtamento do segmento gerado. Não há restrições às propriedades segmentais da vogal envolvida no processo. Quanto às propriedades acentuais, o processo pode ocorrer entre duas sílabas fracas, uma sílaba forte e uma fraca, ou uma fraca e uma forte; nunca ocorre entre duas sílabas fortes ou se a sílaba forte carrega proeminência do nível da frase fonológica. Em (1) exemplificamos um caso de possibilidade (por envolver duas sílabas fracas) e em (2) um caso de bloqueio do processo (por envolver duas sílabas fortes):³

Assumindo que a implementação fonética não necessariamente mapeia a estrutura fonológica (em nosso caso, rítmica) subjacente, adotamos aqui esta segunda via de análise.

3. Os exemplos desta seção são de Santos (2007). A sentença no lado esquerdo traz os domínios prosódicos entre parênteses (para sua construção, cf. Nespor & Vogel 1986). No lado direito, a transcrição fonética. O *italico* marca o contexto segmental da regra e as letras em *maiúsculo* indicam as sílabas acentuadas. O asterisco indica inaceitabilidade daquele tipo de produção.

- 1) [eu comprei {um QUAdro boRRÍvel}_{phP}]_I >> ['k^wa.dro.'xi.vew]
- 2) [ele falaRÁ]_{phP} {ALto deMAIS}_{phP}]_I >> * [fa.la'raw.to.de.'majS]

A elisão, em português brasileiro, apaga a primeira vogal em um contexto de encontro vocálico em sílabas de palavras diferentes. Segundo Bisol (2003), a primeira vogal deve ser um /a/ seguido por qualquer outra vogal. Santos (2007) aponta que no dialeto paulista a regra é um pouco mais genérica, ocorrendo sempre que a primeira vogal for [+ posterior] e a segunda vogal diferente da primeira. No que diz respeito às proeminências, Bisol (2003) e Abaurre, Galves & Scarpa (1999) mostram que a elisão ocorre se as duas sílabas forem fracas ou se a primeira sílaba for fraca e a segunda forte (desde que esta não carregue a proeminência de frase fonológica) e não ocorre se a primeira sílaba for forte ou se a segunda sílaba portar a proeminência de frase fonológica. Os exemplos (3) e (4) ilustram um caso de possibilidade e um caso de bloqueio do processo, por conta da proeminência de frase fonológica:

- 3) [eu {masTIgo}_{phP} {ERvas aMARgas}_{phP}]_I >> [mas.'tʃi.'ger.va.za.'mar.gas]
- 4) [eu {masTIgo ERvas}_{phP} {no jarDIM}_{phP}]_I >> * [mas.'tʃi.'ger.vas]

A ditongação, por outro lado, diminui a quantidade de sílabas sem apagar nenhum segmento. Este processo ocorre no contexto de duas vogais, desde que uma delas seja [+ alta]. No que respeita às propriedades acentuais, é também a que ocorre em mais contextos (Bisol 2003): entre duas sílabas fracas, entre uma forte e uma fraca, entre uma fraca e uma forte, entre duas fortes (desde que a segunda sílaba não carregue proeminência de frase entoacional). Os exemplos (5) e (6) ilustram um caso de aplicação e um caso de bloqueio do processo, respectivamente:

- 5) [{o moLEque Arabe}_{phP} {saIU}_{phP}]_I >> [o.mu.'le.'kja.ra.bi.sa.'iw]
- 6) [eu vi {o moLEque Arabe}_{phP}]_I >> *[mu.'le.'kja.ra.bi]

O último processo a ser resumido é a retração acentual, que não modifica nem é regido por propriedades segmentais, levando em conta apenas as propriedades prosódicas das sílabas. Segundo Abousalh (1997), a retração acentual ocorre em português brasileiro quando há um encontro de sílabas

portadoras de acento primário (portanto, provenientes de palavras diferentes). O processo consiste na mudança do primeiro dos acentos para uma sílaba anterior na mesma palavra (assim, é condição necessária para o processo que a palavra seja no mínimo dissílaba).⁴ É o caso ilustrado em (7):

7) Eu comprei um quilo de caFÉ QUENte >> ['ka.fɛ.'kẽːj.tʃi]

Segundo Tenani (2002), as regras de sândi externo são obrigatórias dentro de frases fonológicas, sendo opcionais entre frases fonológicas (e, portanto, entre os domínios prosódicos superiores). A retração acentual, por outro lado, é opcional dentro de frases fonológicas, mas é bloqueada entre domínios acima da frase fonológica (Abousalh 1997).

2.2. Os processos fonológicos na palavra falada

Vejam os alguns casos de processos fonológicos que ocorrem nos versos de *Gabriela*.⁵ O primeiro verso é bem prototípico, com uma sequência que não dá margens nem para mudanças segmentais (não há contextos para regras de sândi, por exemplo) nem suprasegmentais (não há contexto para retração acentual também), como vemos em (8):⁶

8)	(.	U
	(.	X	.)	.	I
	(.	.	.	.x	.)	(.	.	.	X	.)	PbP
	(x	.)	(.	X	.)	(x	.)	(.	X	.)	C
	(x	.)	(x)	(x	.)	(x	.)	(.	X	.)	W
	(x	.)	(x)	(x	.)	(x	.)	(x)	(x	.)	F
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	S
	To	Dos	Os	Di	As	Es	Ta	Sau	Da	De	

4. Há também outras restrições para que a retração acentual aconteça. Santos (2002) mostra que a retração acentual é sensível ao tipo de categoria fonologicamente vazia que pode estar no contexto (isto é, para acontecer entre um verbo e um advérbio, não é possível que haja uma categoria vazia *pro* entre as duas palavras); Sândalo & Truckenbrodt (2002) por sua vez mostram que as frases fonológicas devem ser do mesmo tamanho.

5. Apresentamos a grade total apenas para o primeiro exemplo. Nos exemplos seguintes, apresentamos apenas os domínios relevantes.

6. Como dissemos, interessa-nos aqui apenas o domínio do pé. Chamamos a atenção, no entanto, que cada domínio tem suas próprias regras (universais ou específicas da língua) de estabelecimento da sílaba forte. Por exemplo, em *os dias*, o pé é marcado com sílaba dominante à esquerda (por isso *o* e *di*). No nível do grupo clítico, a sílaba dominante é aquela que satisfaz duas condições: ter sido sílaba dominante nos níveis inferiores e ser a sílaba da palavra lexical. Deste modo, a proeminência fica à esquerda na sequência *o dia*.

O primeiro nível é o da sílaba. Cada sílaba recebe uma marcação. O segundo nível é o do pé. Veja que *os* é marcado como um pé formado de uma única sílaba. Ele não pode associar-se a *todos* ou a *dias* porque o pé é formado no nível da palavra, isto é, dentro de uma palavra, de acordo com a proposta de Nespor & Vogel.⁷ No nível da palavra fonológica, cada palavra é separada em uma unidade, com uma e apenas uma sílaba proeminente. No nível do grupo clítico, palavras funcionais fracas (sempre monossílabos, normalmente determinantes, conjunções) são associadas a uma palavra lexical. Por isso temos *os dias*, mas não *esta saudade*. É que *esta* não é um clítico fonológico. No nível acima, temos as frases fonológicas: núcleos sintáticos com seus complementos. No nível acima, temos a frase entoacional. Finalmente, o nível do enunciado une frases entoacionais que tenham relação semântico-pragmática entre si. É por isso que o verso acima junta-se com *felicidade, cadê você?* formando um único enunciado, em (9).

9)	x)	U
	(.	.	.	x	.)	(.	.	.	x)	I
	(.	.	.	x	.)	(.	.	.	x)	PbP
	(.	.	.	x	.)	(.	x)	(.	x)	C
	(.	.	.	x	.)	(.	x)	(.	x)	W
	(x)	(x)	.)	(x	.)	(x)	(x)	(x)	(x)	Σ
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	S
	Fe	Li	Ci	Da	De	Ca	Dê	Vo	Cê	

Em (9), há dois momentos em que o ritmo da sentença é perturbado. O primeiro é em *felicidade*. Como é possível observar na linha do pé, as duas primeiras sílabas recebem acento de pé, criando um encontro acentual de acentos secundários (o acento de palavra está em *DA*). Segundo Collishmon (1994), em português brasileiro, quando temos este tipo de encontro acentual, podemos escolher entre apagar o acento da primeira ou da segunda sílabas (produzindo ou *FELiCiDAde* ou *feLiCiDAde*).⁸ O segundo momento ocorre em *cadê você*. Veja que aqui todas as sílabas recebem acento de pé,

7. Lembramos que o domínio do pé, proposto por Nespor & Vogel (1986) não obedece às mesmas regras de construção do pé métrico, do qual se derivam os acentos secundários. O domínio proposto por Nespor & Vogel deve ser construído dentro da palavra fonológica (caso contrário, violaria o princípio de construção de constituinte prosódico: “Join into na n-ary branching X^P all X^{P-1} included in a string delimited by the definition of the domain of X^P.” – cf Nespor & Vogel 1986, 7). Outra característica do pé prosódico é que todas as sílabas devem ser mapeadas no nível do pé, o que gera uma grande quantidade de pés degenerados.

8. Veja que em português europeu a escolha é por apagar sempre a segunda sílaba deste tipo de encontro, produzindo *FELiCiDAde*

dado que o pé está sendo construído dentro da palavra. No primeiro caso, havia a possibilidade de apagar o primeiro ou segundo acento; agora não há esta possibilidade porque as sílabas *dê* e *cê* são também portadoras do acento de palavra.⁹ Desse modo, aplicado o apagamento de acento, há duas possibilidades rítmicas de pronúncia deste verso ((10a) e (10b)). Falantes do português brasileiro lendo (e não cantando) este verso tenderão a preferir a primeira opção (10a):¹⁰

10a) . x . x . . x . x 10b) x . . x . . x . x
 Fe lí ci da de ca dê vo cê Fe lí ci da de ca dê vo cê

Nenhuma opção é totalmente perfeita ritmicamente (no sentido de ter uma perfeita alternância entre fracas e fortes), mas a primeira opção tem apenas uma violação (duas fracas seguidas em *de ca*), enquanto a segunda opção viola o ritmo binário duas vezes (em *lici* e *deca*).

Processos fonológicos diferentes ocorrem no verso a seguir:

11)	(.	.	.	X)	(.	x)	(.	X)	<i>PbP</i>
	(x)	(x)	(.	X)	(.	x)	(x)	(x)	<i>C</i>
	(x)	(x)	(.	X)	(.	x)	(x)	(x)	<i>W</i>
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	<i>F</i>
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	<i>S</i>
	Já	Não	Con	Si	Go	Vi	Ver	Sem	E	La	

Neste caso, há dois momentos que perturbam o ritmo da sentença. O primeiro é na sequência *já não consi*, todas as sílabas recebendo acento secundário. A escolha do português será por apagar os acentos em *já* e *con* e criar uma sequência perfeita de sílabas fortes e fracas. O segundo caso é o encontro de acentos entre *ver sem*. Embora a palavra *viver* seja dissílaba e portanto tenha sílabas para a retração acentual, este processo só pode ocorrer

9. Não é possível apagar um acento em um nível sem apagá-lo também nos níveis superiores. Apagar os acentos em *dê* e *cê* acabaria por fazer perder a informação de qual é o acento dessas palavras. Em palavras monossílabas não há problemas de o acento ser apagado porque como só há uma sílaba, o falante não terá dúvidas de onde o acento deveria estar.

10. As discussões sobre as possíveis escolhas dos falantes foram feitas levando-se em conta as possibilidades lógicas de modificações da estrutura a partir do julgamento de informantes (através de análise de ouvira). Não era objetivo deste trabalho uma análise quantitativa destes julgamentos e por isso os dados não foram codificados quantitativa ou sócio-linguisticamente. Logo, é possível que haja casos em que falantes de diferentes regiões escolham diferentes estratégias de resolução destas perturbações (ou mesmo não escolham nenhuma estratégia).

dentro da frase fonológica (cf. Abousalh 1997). Como o encontro acentual ocorre entre frases fonológicas, a retração acentual não ocorrerá. Ao invés disso, a escolha mais comum do falante de português será de apagar o acento em *sem*. O resultado terá apenas uma violação, como ilustrado em (12):

- (12) . x . x . . x . x .
 já não con si go vi ver sem e la

Os processos apontados acima alteram a estrutura rítmica *sem*, no entanto, alterar a sequência segmental. Nos versos abaixo (13), apontamos para processos que alteram a estrutura rítmica alterando também a quantidade de sílabas do enunciado.

- | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------------|
| 13) | (. | . | . | X | .) | (. | X | .) | (. | . | x) | <i>PbP</i> |
| | (x) | (. | . | X | .) | (. | X | .) | (. | . | x) | <i>C</i> |
| | (x) | (x) | (. | X | .) | (x) | (x | .) | (x) | (. | x) | <i>W</i> |
| | (x) | (x) | (x) | (x | .) | (x) | (x | .) | (x) | (x) | (x) | <i>F</i> |
| | (x) | (x) | (x) | (x) | (x) | (x) | (x) | (x) | (x) | (x) | (x) | <i>S</i> |
| | Eu | Te | Per | Do | O | Com | Rai | Va | E | A | Mor | |

Em (13), há vários momentos que perturbam a sentença. Os dois primeiros (*eu te perdôo* e *com rai*) são dos tipos discutidos acima. O último é *raiva e amor*. Se levássemos em conta apenas o nível do pé, a otimização dessa sequência seria através do apagamento do acento em *a*, criando um ritmo perfeitamente binário *RAIva E aMOR*. No entanto, o *e* é um clítico fonológico que não ‘suporta’ um acento. Assim, a sequência ficaria com muitas sílabas átonas seguidas (o que também é indesejável na otimização rítmica): *RAIva e aMOR*. O que os falantes de português brasileiro fazem é, então, elidir uma sílaba, ficando com a sequência em (14):

- 14) X . . x
 rai vi a mor

Veja que esta sequência ainda não é perfeita, mas resulta da tensão da língua entre otimizar ritmicamente a sentença (o que privilegiaria a sequência *RAIva E aMOR*) e manter informações linguísticas importantes para outros componentes da língua (como a morfologia, que diz que a conjunção não pode portar o acento). Como é possível observar, a língua prefere privilegiar informações que são importantes em outros componentes.

Um outro exemplo de processos que alteram a quantidade de sílabas é a degeminação, como acontece em (15):

15)	(x .)	(x)	(x .)	(x .)	(x)	(x .)	(x)	(x .)	(x .)	W		
	(x .)	(x)	(x .)	(x .)	(x)	(x .)	(x)	(x .)	(x .)	F		
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	S		
	Mo	Lha	A	Tu	A	Bo	Ca	Na	Mi	Nha	Bo	Ca

na sequência *molha a tu*, ao apagar o acento de *a*, cria-se uma sequência de sílabas desacentuadas. O falante, então, une as duas sílabas criando uma sequência ótima (16):

16)	x .	x .	x .	x .	x .	x .					
	mo	lha	tu	a	bo	ca	na	mi	nha	bo	ca

Mas nem sempre a aplicação dos processos fonológicos ocorre para otimizar a estrutura rítmica das sentenças. Finalizamos esta subseção analisando dois casos de aplicação de processos fonológicos que não otimizam a estrutura rítmica.

17)	(x .)	(. x .)	(x .)	(x .)	(x .)	(x .)	(x)	W		
	(x .)	(x)	(x .)	(x .)	(x .)	(x .)	(x)	F		
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	S		
	Vol	Ta	Ban	Di	Da	Ma	Ta	E	Ssa	Dor

No exemplo em (17), o único problema com o ritmo está em *bandida*. A solução, como já vimos, é o apagamento do acento da primeira sílaba. Isto cria uma sequência de duas sílabas desacentuadas, mas esta sequência é permitida em português brasileiro (veja que as palavras proparoxítonas são todas com essa sequência). No entanto, o falante de português brasileiro vai aplicar elisão na sequência *ta e*, criando um encontro acentual (18):

18)	x . .	x .	x x .	x					
	vol	ta	ban	di	da	ma	te	ssa	dor

não há razões claras de porque o falante aplica os processos fonológicos nesses casos, mas ele o faz. O caso mais radical de apagamento é o que ocorre no verso a seguir:

19)	(x)	(x .)	(x .)	(x)	(x)	(. x .)	(. x .)	W					
	(x)	(x .)	(x .)	(x)	(x)	(x)	(x .)	F					
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	S					
	A	Mi	Nha	Ca	Sa	É	Um	Es	Cu	Ro	De	Ser	To

em (19), a sequência *é um es* está em encontro acentual. Através apenas do apagamento de acentos, não há como transformar esta sequência numa sequência rítmica ótima (veja em 20a o apagamento do acento em *es* e em 20b o apagamento em *um*):

- 20a) . x . x . x x . x . . x . 20b) . x . x . x . x x . . x .
 a mi nha ca sa é um es cu ro de ser to a mi nha ca sa é um es cu ro de ser to

Mas há ainda uma opção mais radical, o apagamento do acento em três sílabas com sua posterior ditongação, como em (21):

- 21) . x . x x . x . . x .
 a mi nha ca séw es cu ro de ser to

embora alguns falantes acabem também por apagar o segundo acento do encontro acentual, não é sempre que isso ocorre:

- 22) . x . x . . x . . x .
 a mi nha ca sew es cu ro de ser to

Assim, nesta seção, vimos como frequentemente os processos fonológicos de sândi externo são violados na canção *Gabriela*, quando consideramos seus versos produzidos de forma falada. A questão que se levanta é se estes processos ocorrem obedecendo a algum outro domínio.

3. O COMPONENTE MELÓDICO

Nesta seção discutiremos como a métrica melódica interfere na métrica do componente linguístico. Apresentamos a proposta de análise melódica de Carmo Jr (2007) e a aplicação desta proposta para os versos de *Gabriela*, mostrando como a métrica melódica favorece ou bloqueia certos processos fonológicos e determina a realização de deslocamentos acentuais.

3.1. *Compatibilidade entre fala falada e cantada*

Vimos na seção anterior que os processos fonológicos procuram dotar as sentenças de uma estrutura rítmica regular, própria de cada língua e

que, no caso do português brasileiro, seria uma sequência de pés binários. Como a realização de tais processos é circunscrita a domínios prosódicos, muitas vezes não se atinge um padrão rítmico ótimo.¹¹ A sequência dos dois primeiros versos que compõem a primeira estrofe de *Gabriela* é um bom exemplo dessa limitação, como ilustrado em (23):

- (23) x . . x . x . . x .
 To dos os dí as es ta sau da de

 . x . x . . x . x
 Fe li ci da de ca de vo cê

Nenhum dos versos apresenta uma estruturação rítmica regular identificável. A questão que se apresenta é que esse fato vai de encontro ao efeito de sentido fortemente rítmico criado quando esses versos são cantados pela voz de Tom Jobim¹². O verso 32, por exemplo, é cantado como (24):¹³

- (24) x . . x . . x . . x . .
 Já não com si % go vi ver sem e % leu

esse efeito rítmico se dá através do alongamento (%) de *si* e *e*, além da retração acentual em *viver*. No entanto, prosodicamente (linguisticamente) este acento não poderia ser retraído, porque ele ocorre entre frases fonológicas. Não existe tampouco uma explicação de porque o processo fonológico de alongamento seja aplicado nas sílabas *si* e *e*. Portanto, estamos diante de um processo que afeta a cadeia da expressão e que parece ter sua origem na melodia e não na fala.

3.2. Hierarquia melódica

Há duas maneiras de se tentar explicar porque os processos fonológicos discutidos acima violam a estrutura linguístico-prosódica. Uma é dizer que,

11. Ótimo aqui, está sendo usado como a maneira perfeita, ideal. O termo é usado nos estudos sobre ritmo e métrica para caracterizar a alternância regular entre fortes e fracas. E como chama a atenção um dos pareceristas, “um ótimo absoluto não é operacional, nem encontrado na prática”.

12. Jobim, Tom. 1987. Gabriela. Intérprete: Tom Jobim. In: *Passarim*. São Paulo: PolyGram. 1 CD. Faixa 11.

13. Empregamos o símbolo % para referir uma posição na grade melódica que se concretiza seja como um alongamento da sílaba, seja como uma pausa.

na música, eles ocorrem aleatoriamente, sem serem circunscritos de maneira alguma, outra é dizer que eles obedecem a um outro tipo de hierarquia. Ao final desta seção teremos apontado razões de porque defendemos a segunda hipótese neste ensaio.

Carmo Jr. (2007) propõe uma hierarquia melódica composta de 6 níveis: período (U), frase (I), célula (C), pé (Σ), nota (σ), cronema (χ), como se observa na Figura 2:

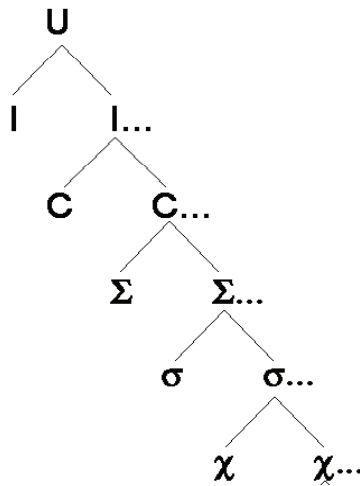


Figura 2: Hierarquia melódica

Vejam, ainda que rapidamente, as características de cada um desses níveis. A unidade elementar de uma cadeia melódica é o *cronema* (χ), que pode se realizar seja como uma nota, seja como um prolongamento de nota, seja como uma pausa. Isto quer dizer que esta unidade pode se realizar como um *silêncio significativa*. Os *Cronemas* são unidades de duração – daí a motivação para o termo – capazes de distinguir diferentes enunciados melódicos¹⁴.

A *nota* musical (σ) é o constituinte imediatamente superior ao *cronema*. Uma *nota* é constituída por um ou mais cronemas. Pode ser segmentada

14. O termo *chroneme* foi criado por Jones (1962:115) para indicar o emprego da duração com função distintiva em algumas línguas. Aqui utilizamos o termo *cronema* para referir uma unidade de duração capaz de distinguir diferentes segmentos melódicos.

com base no traço *duração* [\pm longo] e pode ser analisada com base nos traços *intensidade* [\pm forte] e *altura* [\pm agudo].

O constituinte que domina a *nota é o pé* (Σ), que tem a *intensidade* como traço pertinente. O *pé* melódico é definido como uma cadeia de duas ou mais notas uma das quais tem o traço [+forte] enquanto a(s) restante(s) têm o traço [-forte]. Veja que, diferentemente do pé fonológico (cf. seção 2), o pé melódico obrigatoriamente apresenta uma nota com o traço [-forte].

A *célula* (C) é a unidade que domina o pé. Construída com informações de *duração* [\pm longo] e de *intensidade* [\pm forte], a célula apresenta as seguintes características: a) presença *obrigatória* de um núcleo de sonoridade (*crusa*) com os traços [+forte] e [+longo]; e b) presença *facultativa* de adjacências (*anacrusa*) ou (*metacrusa*) com os traços [-forte] e/ou [-longo].

O nível que se segue imediatamente à *célula* é a *frase* (I), que faz uso de traços entoativos como ascendência *vs.* descendência de tom. Uma frase constitui uma unidade aproximada àquilo que se pode cantar em um só fôlego (Schoenberg 1996:29). Portanto, o fim de uma frase melódica é o ponto onde se inserem pausas numa cadeia melódica. Essa característica nos leva a associar a frase melódica à frase entoacional (Nespor & Vogel 1986:188).

Por fim, o *Período* (U) compõe-se de uma ou mais frases e apresenta uma função harmônica implícita chamada *cadência*. Uma cadência tem o traço [\pm perfectivo], que cria o efeito de sentido de conclusão (cadência perfeita) ou não-conclusão (cadência imperfeita) num período melódico.

3.3. A estrutura métrica de *Gabriela*

Vamos agora empregar alguns destes conceitos na descrição da estrutura métrica da melodia de *Gabriela*, procurando equacionar a questão da compatibilidade entre letra e melodia dessa canção. A Figura (3) representa o primeiro *período* da canção (doravante U). A *cadência*, traço característico de U, aparece no final do período com o valor [-perfectivo], sob as cifras harmônicas “B7” e “E”.

Podemos também visualizar na Figura (3) as duas *frases*, I1 e I2 que dividem U. Outra informação aí contida é a estrutura das *células* (C) de U. É a estrutura de C que mais nos interessa de imediato. Retomemos a estrutura rítmica do primeiro verso de *Gabriela* em (25):

- 25) x . . x . x . . x . . x . x . . x . x
 To dos os di as es ta sau da de fe li ci da de ca de vo cê

Ora, quando cantado, o segmento (25) exhibe uma outra estrutura métrica, a saber:

- 26)

x . . x . .	x . . x . .	x . . x . .	x . . x . .
To dos os di %	as es ta sau da %	de fe li ci da %	de ca de vo ce % %

É fácil perceber que a divisão rítmica do pé (x . .) e da célula (x . . x . .), reiterados ao longo de todo o trecho, são os responsáveis pela unidade rítmica do verso cantado. Outro dado importante presente em (26) é o papel desempenhado pelos cronemas no equacionamento métrico da canção. Eles não apenas se realizam como prolongamentos de notas, como em (*di* (%*as*) e (*da* (%*de*)), como também como silêncios musicalmente significantes, como é o caso do segmento final de I1 (*cê* (%)(%)). O primeiro destes cronemas prolonga a duração da sílaba *ce*, mas o segundo apenas insere uma pausa para que o ritmo da melodia como um todo não se quebre. Caso contrário, na juntura entre I1 e I2 teríamos uma sobreposição entre o trecho final de I1 e o inicial de I2, como ilustra (27):

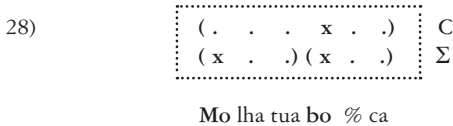
- 27)

x . . x . .	x . . x . .	x . . x . .	x . . x . .
fe li ci da %	de ca de vo ce %	Já não com si %	go vi ver sem e % la
		fim de I1	início de I2

Cantada dentro desse esquema métrico, o ritmo ternário da canção se quebra. É quando dizemos que o cantor “atravessa” o ritmo da melodia.

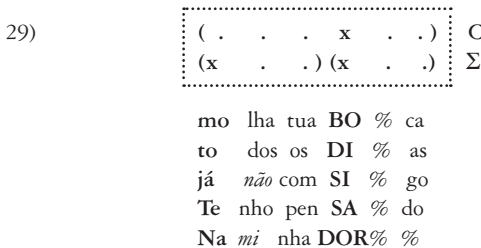
A partir dos segmentos (26) e (27) podemos visualizar o mecanismo de construção dos pés e das células melódicas em *Gabriela*. Vimos que o domínio do pé é construído com informações de *intensidade* [\pm forte] enquanto o domínio da célula é construído com informações de *intensidade* [\pm forte] e *duração* [\pm longo]. Vimos também que a célula, composta de

dois pés, se define pela presença *obrigatória* de um núcleo de sonoridade (*crusa*) com os traços [+forte] e [+longo]. Assim, em (28) as sílabas *mo* e *bo*, recebem ambas o traço [+forte], na medida em que são os cabeças dos dois pés que compõem a célula.



No entanto, apenas *bo* tem o traço [+longo], uma vez que sua duração é alongada pela adjacência com o cronema. Portanto, *mo* tem os traços [+forte] e [-longo] enquanto *bo* tem os traços [+forte] e [+longo], o que faz desta última sílaba o núcleo de sonoridade da célula (*crusa*), conforme a definição dada.

A partir disso, pode-se afirmar que é na estrutura da célula melódica que encontramos as razões pelas quais as sílabas acentuadas são alongadas em determinados pontos e em outros não. Na fala, o alongamento ocorre nas sílabas portadoras de acento (primário ou secundário) ou nas sílabas finais (demarcando os limites dos domínios prosódicos). Mas quando a fala é recoberta pela melodia parece ocorrer um fenômeno interessante: todo alongamento se dá sobre sílabas portadoras de acento, mas nem toda sílaba portadora de acento é alongada. Em (29) vemos que as sílabas que sofrem alongamento sempre coincidem com as *crusas* das células, e que as sílabas acentuadas que não coincidem com a *crusa* (marcadas em *itálico*) nunca são alongadas.



É por isso que o compositor não pode acrescentar um tempo em *mi* e em *não*. Essa constatação parece confirmar a existência de uma hierarquia melódica que interage com a hierarquia prosódica e que determina a realização ou não de alguns processos fonológicos, como veremos a seguir.

4. A TENSÃO MELODIA/PROSÓDIA

Nesta seção, discutimos a tensão provocada entre a superposição de uma estrutura linguística e de uma estrutura musical. Como vimos na seção 2, o ritmo é definido como uma alternância entre sílabas fracas e fortes, definindo então, um ritmo binário. Especificamente para o português, também se defende que este ritmo binário é resultado de vários pés binários com a sílaba forte à esquerda (portanto, um troqueu). Na seção 3 vimos que *Gabriela* caracteriza-se por um ritmo musical ternário (três sílabas com a sílaba forte também à esquerda). Desse modo, instaura-se a tensão: privilegia-se o ritmo linguístico ou o musical, no caso de conflitos entre os dois?

Interessantemente, na maior parte da música o conflito não se instaura. A escolha do formato prosódico das palavras é cuidadosa. Não há palavras proparoxítonas, que teriam o formato do pé musical e as palavras oxítonas são apenas três: *você, carnaval e amor*.

Mas então, como, com palavras paroxítonas (portanto, privilegiando o ritmo linguístico), o compositor conseguiu manter o ritmo musical ternário? A resposta está no uso abundante de clíticos fonológicos e palavras trissílabas paroxítonas e em algumas poucas violações do ritmo linguístico.

No que se refere ao primeiro caso, vejamos alguns exemplos:

30) (x . . .) (x . . .) (x . . .) (x . . .)
che ga mais per to meu rai o de sol % %

como é possível observar em (30), é o uso de *mais, meu e de* que, associados às palavras anteriores, criam sequências ternárias. Mas, para isso ocorrer, o domínio em que o ritmo está sendo construído muda. Como vimos, Nespor & Vogel propõem que o pé se constrói dentro da palavra. Aqui, o pé está sendo construído dentro da sentença.¹⁵

Em (31) observamos que as palavras monossílabas também são usadas para criar pés ternários:¹⁶

15. É por esta razão que alguns pesquisadores propõem uma distinção entre o pé da prosódia – construído dentro da palavra e *locus* de processos fonológicos – do pé métrico – local de otimização rítmica. Cf. Fudge (1999).

16. Observe-se que o compositor acrescenta dois tempos (% %) ao final da sequência para fechar a métrica melódica, no caso, para criar quatro pés de três tempos cada.

- 31) (x . . .)(x . . .)(x . . .)(x . . .)
 mais um pa lha ço no teu car na val % %

Como é possível observar, em (30) *mais* sofre apagamento de acento, enquanto que em (31) ele tem o acento mantido e carrega o núcleo do pé ternário.

O uso de trissílabas paroxítonas também ajuda na construção de pés ternários porque a primeira sílaba passa a, metricamente, pertencer ao pé anterior. Em (32) marcamos, acima da sequência segmental, o ritmo musical (Σ); abaixo, marcamos o domínio das palavras (W):

- 32) (x . . .)(x . . .) Σ
 quan ta ca cha % ça
 (x . . .)(. x . . .) W

Finalmente, a maneira mais radical de fazer com que a sequência rítmica linguística se conforme a sequência musical diz respeito a aplicações de processos fonológicos que violam os domínios prosódicos. Em (10), vimos que o verso *Felicidade, cadê você* poderia ter duas sequências rítmicas ótimas. Normalmente, o falante de português brasileiro opta pela opção de acentuar *li*, enfatizando o ritmo binário; no entanto, a segunda opção obedece mais ao ritmo ternário, e é ela a escolhida na canção. Porém, ambas as estruturas apresentam um pé binário no final *dê vo*. O que ocorre então é a mudança de acento de *dê* para *ca*, mesmo não estando em situação de encontro acentual (portanto, violando as propriedades de aplicação do processo).¹⁷ Tal processo gera a seguinte estrutura, em (33):

- 33) x . . x . x . . x
 fe li ci da de ca dê vo cê

Embora a troca pareça inócua, pois continuamos tendo 2 pés ternários e 1 pé binário, o compositor agora alonga a sílaba tônica *da*, e com isso consegue os três pés ternários, como ilustra (34):

- 34) x . . x . . x . . x
 fe li ci da de % ca dê vo cê

17. O mesmo processo de mudança de acento entre frases fonológicas ocorre no verso *Já não consigo viver sem ela*.

O segundo caso é o que se segue em (35):

35)	(x)	(.)	(.)	(x)	(x)	(.)	(x)	(x)	(.)	(.)	(x)	C
	(x)	(x)	(.)	(x)	(x)	(.)	(x)	(x)	(.)	(x)	(x)	W
	(x)	(x)	(x)	(.)	(x)	(.)	(x)	(x)	(.)	(x)	(x)	F
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	S
	Mas	Com	Vo	Cê	E	La	É	Chei	A	De	Sol	

o ritmo binário ocorre através do apagamento do acento da primeira sílaba e da elisão e apagamento de acento da sequência *la é*, como representado em (36):

36)	.	x	.	x	x	.	x	.	.	x	
	mas com vo cê e lé chei a de sol										

linguisticamente, o choque acentual entre *cê* e *e* não pode ser desfeito porque estão em frases fonológicas diferentes. Não há processos fonológicos que possam ser aplicados de modo a desfazer o pé ternário final.

Musicalmente, o que o compositor faz é que aparece em (37):

37)	x	.	.	x	.	.	x	.	.	x	
	mas com vo cê e lé chei a de sol										

o autor mantém o acento da primeira palavra e apaga o da segunda. Mais interessantemente, ele apaga o acento do encontro acentual. No entanto, ao fazer isso, ele apaga a informação de que a palavra *ela* tem um acento. Veja que as duas sílabas passam a ser fracas.¹⁸

A fonologia argumenta que os processos fonológicos (mudança de acento, apagamento de sílabas) que discutimos aqui são opcionais e por isso muitas vezes imprevisíveis. No entanto, do ponto de vista da melodia, é possível determinar quando os processos vão ou não ocorrer. Terminamos a seção 3 com a análise dos versos *molha a tua boca, na minha boca; a tua boca é meu doce é meu sal*. Dois processos fonológicos podem, em princípio, ocorrer aqui: a degeminação de *molha a* e a ditongação, em dois ambientes, de *tua*. No entanto, o compositor só ditonga no primeiro caso. Se prosodicamente não há explicação de porque aplicar em um caso e não em outro, é perfeitamente possível explicar (e prever os contextos de aplicação) pela dominância

18. O mesmo ocorre com a sequência *mata essa dor*.

que a métrica melódica exerce sobre a métrica prosódica. Prosodicamente, as duas instâncias de *tua* podem ocorrer, mas é só no momento em que tais palavras são situadas em um ambiente melódico que os processos podem finalmente ser aplicados ou não, como temos em (38):

38) x . . x . . x . . x . . x . . x . . x . . x
 mo lha twa bo % ca na mi nha bo ca % a tu a bo kE meu do cE meu sal

Finalmente, um último exemplo de violação de regras linguísticas é o de (39):

39)	(x	.	(x)	(x	.	(x)	(x	.	(x)		W
	(x	.	(x)	(x	.	(x)	(x	.	(x)		F
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)		S
	Quan	Ta	Ca	Cha	Ça	Na	Mi	Nha	Dor		

Em (39), há dois encontros de acento: *cacha* e *na mi*. No primeiro caso, o compositor apaga o acento da primeira sílaba. No segundo, no entanto, o compositor apaga o segundo acento. Ao fazer isso, ele viola duas propriedades linguísticas: deixa uma palavra dissílaba sem acento e o faz em detrimento a um clítico fonológico, como ilustrado em (40):

40) x . . x . x . . x
 quan ta ca cha ça na mi nha dor

Mas isso não torna a sequência ótima. Para que isso ocorra o compositor tem que acrescentar um tempo na sílaba tônica, como em (41):

41) x . . x . . x . . x . .
 quan ta ca cha % ça na mi nha dor % %

No entanto, o que ocorreria se ele tivesse optado por apagar o primeiro acento?

42) x . . x . . x . x
 quan ta ca cha ça na mi nha dor

Neste caso, representado em (42), ele teria novamente uma sequência de dois ritmos ternários e um binário. A solução seria alongar a sílaba *mi*, criando um pé ternário. Então por que esta segunda alternativa não é possível?

A resposta está no fato de que o autor não pode simplesmente aplicar os processos fonológicos. Estes processos devem obedecer à hierarquia melódica. Como vimos no início da seção 3, o alongamento só pode ocorrer na sílaba acentuada que coincide com a crusa. É o que ocorre neste caso. O pé *minba* é formado por uma sílaba *nba* fraca (que não pode ser alongada) e por uma sílaba *mi* que, embora acentuada, não coincide com a crusa no domínio melódico C.

5. CONCLUSÃO

Ao comparar a estrutura métrica dos componentes linguístico e melódico da canção *Gabriela*, mostramos que, quando se instaura uma tensão, os processos fonológicos passam a obedecer à estrutura melódica em detrimento da estrutura prosódica. Isto é, mostramos que o componente melódico domina o componente linguístico, ou seja, impõe sua estrutura à cadeia linguística. Em *Gabriela*, a realização ou não de processos segmentais, como a ditongação e a elisão, e suprasegmentais, como a retração acentual, é determinada, em última instância, pela estrutura dos pés e das células melódicas.

Estes resultados são sem dúvida interessantes para o fonólogo e para o musicólogo, que podem, a partir de outros estudos semelhantes, afinar sua teoria prosódica e sua técnica de análise musical. Mas a hierarquia melódica que estabelecemos para a análise de *Gabriela* tem também uma pertinência semiótica. Ela nos fornece, ainda que obliquamente, alguns indícios do que se pode entender por “sentido” de um texto musical, independentemente de qualquer referência ao conteúdo linguístico.

Sabemos intuitivamente que *uma melodia não é uma sequência qualquer de notas musicais*. Se assim fosse, uma criança de dois anos que martela notas ao piano estaria criando uma melodia. Em outras palavras, se a expressão melódica não tivesse um *sentido* em si mesma, mal poderíamos distingui-la de uma sequência estocástica de notas. Logo, não teríamos como identificá-la ou como memorizá-la. Mas nossa experiência nos mostra o oposto disso. Cada um de nós conhece “de cor” milhares de melodias diferentes, fato que, por si só, exige que estas sequências de sons tenham algum tipo de organização interna independente de qualquer outro componente semiótico, como o componente linguístico, por exemplo.

A questão está em precisar quais são os elementos que caracterizam estas sequências melódicas e que as distinguem das sequências não-melódicas. Com base no presente estudo podemos propor que a sequência melódica – e apenas ela – é estruturada como uma hierarquia composta de níveis. Esta hierarquia de níveis parece ser o mecanismo gerador do fio melódico, sendo responsável pela coesão (Hjelmslev 1975:72) que caracteriza toda melodia, criando o efeito de sentido de uma totalidade orgânica. Como vimos, essa coesão é forte a ponto de impor-se sobre domínios extra-melódicos, como o domínio prosódico: o ritmo linguístico não é capaz de quebrar o ritmo melódico, mas este pode subverter aquele. Mais que isso, essa hierarquia pressupõe que por trás de uma sequência melódica exista algum tipo de “inteligência” musical, numa palavra, de um enunciador musical que diz algo através da música.

Embora estudos mais detalhados precisem ser realizados para esclarecer alguns aspectos da questão – qual a generalidade desta hierarquia? Estarão todos os níveis sempre manifestados em qualquer melodia? – existem razões para crermos que a hierarquia de níveis é o substrato de um conjunto muito extenso de melodias, podendo ser tomada como um dos critérios de distinção entre o melódico e o não-melódico.

Recebido em fevereiro de 2010

Aprovado em maio de 2010

E-mails: carmojr@usp.br

raquelss@usp.br

REFERÊNCIAS BIBLIOGRÁFICAS

- ABAURRE, Maria Bernadete; SVARTMAN, Flaviani Romani Fernandes. 2008. Secondary stress, vowel reduction and rhythmic implementation in Brazilian Portuguese. In: BISOL, Leda; BRESANCINI, Cláudia Regina. (Orgs.). *Contemporary Phonology in Brazil*. Newcastle: Cambridge Scholars Publishing, p. 54-83.
- ABAURRE, Maria Bernadete, GALVES, Charlotte & SCARPA, Ester. 1999. A interface fonologia-sintaxe. Evidências do português brasileiro para uma hipótese top-down na aquisição da linguagem. In: SCARPA, E. (org) *Estudos de Prosódia*, Campinas: Ed. da Unicamp, p. 285-323.

- ABOUSALH, Elaine Ferreira. 1997. *Resolução de choques de acento no Português Brasileiro: elementos para uma reflexão sobre a interface fonologia-sintaxe*. (dissertação de mestrado). Campinas: UNICAMP.
- ARANTES, Pablo. 2005. *Acentuação secundária em português brasileiro: fundamentos de produção e percepção do ritmo da fala*. (ms). Campinas: UNICAMP.
- ARANTES, Pablo & BARBOSA, Plínio Almeida. 2002. Acentuação secundária em português brasileiro à luz do modelo dinâmico do ritmo: um estudo piloto. In: *Anais do I Congresso Internacional de Fonética e Fonologia/ VII Congresso Nacional de Fonética e Fonologia*. Belo Horizonte.
- BISOL, Leda. 1996. Sândi externo: o processo e a variação. In: KATO, M. (org.) *Gramática do Português Falado*. Campinas: UNICAMP. (5):55-96.
- BISOL, Leda. 2003. Sandhi in Brazilian Portuguese. In: *Probus* 15(2): 177-200.
- CARMO JR, José Roberto do. 2007. *Melodia & Prosódia: um modelo para a interface música/fala com base na análise comparada do aparelho fonador e dos instrumentos musicais reais e virtuais*. (tese de doutorado). São Paulo: FFLCH/USP.
- CARVALHO, Joaquim Brandão. 1989. Phonological conditions on Portuguese clitic placement: on syntactic evidence for stress and rhythmical patterns. *Linguistics*, vol.27: 405-436.
- COLLISCHON, Gisela. 1994. Acento Secundário em Português. *Letras de Hoje*, vol.29 (4): 43-53.
- FUDGE, Erik. 1999. Words and feet. *Journal of Linguistics*, vol.35: 273-296.
- HJELMSLEV, Louis. 1975. *Résumé of a theory of language*. Madison: The University of Wisconsin Press.
- JONES, Daniel. 1962. *The Phoneme: Its Nature and Use*. Cambridge: Cambridge University Press.
- JOBIM, Tom. 1987. Gabriela. Intérprete: Tom Jobim. In: *Passarim*. São Paulo: PolyGram, 1 CD. Faixa 11.
- MORAES, João Antonio. 2003. Secondary stress in Brazilian Portuguese: perceptual and acoustical evidence. *Proceedings of the 15th International Congress of Phonetic Sciences*, p. 2063-2066.
- NESPOR, Marina & VOGEL, Irene. 1986. *Prosodic Phonology*, Dordrecht: Foris Publications.
- SÂNDALO, Filomena & TRUCKENBRODT, Hubert. 2002. Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics*. Vol. 42, p. 285-310.

- SANTOS, Raquel Santana. 2002. Categorias sintáticas vazias e retração de acento em português brasileiro. *DELTA – Revista de Documentação de Estudos em Linguística Teórica e Aplicada*. Vol. 18 (1), p. 67-86.
- SANTOS, Raquel Santana. 2007. *A aquisição prosódica do português brasileiro de 1 a 3 anos*. (Tese de Livre-docência) São Paulo: FFLCH-USP.
- SCHOENBERG, Arnold. 1996. *Fundamentos da composição musical*. São Paulo: EDUSP.
- SELKIRK, Elisabeth. 1984. *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge, Massachusetts: The MIT Press.
- TAKEZAWA, Koichi. 1981. Rhythm Rule in Metrical Theory. *Linguistic Analysis*, vol.8 (1):1-14, Elsevier North Holland, Inc.
- TATIT, Luiz. 1996. *Semiótica da canção: melodia e letra*. São Paulo: Escuta.
- TENANI, Luciane Ester. 2002. *Domínios prosódicos no Português do Brasil: implicações para a prosódia e para a aplicação de regras fonológicas* (tese de doutorado). Campinas: UNICAMP.