

SOBRE A GRAMÁTICA DA PALAVRA CANTADA

JOSÉ ROBERTO DO CARMO JR*

RESUMO

O presente ensaio propõe um método de análise para a palavra cantada tomando-a como uma superposição entre duas estruturas hierárquicas, uma prosódica e outra melódica, governada por regras de preferência que determinam o *alinhamento* entre os elementos terminais (nota e sílaba), e a *associação* entre os não-terminais (pé, palavra, grupo clítico, frase fonológica e entoacional de um lado, e pulso, célula e frase melódica de outro). O não-isomorfismo entre as hierarquias prosódica e melódica permite explicar a ocorrência de certos fenômenos anômalos observados na palavra cantada. **Palavras chave:** fonologia; hierarquia melódica; canção.

ABSTRACT

In this essay, we propose a method of analysis for the sung melody based on the assumption that it presents a superposition between two hierarchies, a melodic and prosodic one. The relationship between these hierarchies is governed by rules that (i) establish the *alignment* between the terminals (note and syllable) and (ii) the *association* among non-terminals (foot, word, clitic group, phonological phrase and intonation phrase from one side and beat, cell and melodic phrase from the other). The fact that melody and prosody are both structured but are not isomorphic help us to understand some anomalous phonological phenomena observed in the sung melody.

Key-words: phonology; melodic hierarchy; song.

1. INTRODUÇÃO

A palavra cantada constitui um interessante objeto de análise para o linguista, uma vez que não raro apresenta um comportamento desviante com relação à gramática da palavra falada, notadamente no que diz respeito à fonologia. São comuns os deslocamentos acentuais, os clíticos acentuados, as sílabas epentéticas e o bloqueio a processos fonológicos sem que existam contextos linguísticos adequados. No presente ensaio mostramos como tais fenômenos podem ser

*USP, São Paulo (SP), Brasil. jrcarmojr@gmail.com

*Agradecemos as valiosas sugestões recebidas dos pareceristas anônimos do Caderno de Estudos Linguísticos. Os problemas remanescentes são de nossa inteira responsabilidade. Agradecemos ainda à FAPESP (processo 2008/06656-7).

compreendidos a partir das seguintes assunções: (i) a palavra cantada constitui-se pela superposição entre duas hierarquias, uma melódica e outra prosódica; (ii) dado que tais estruturas não são isomórficas, a cadeia melódica pode ser afetada pela pressão exercida pela estrutura prosódica ou, ao contrário, pode ocorrer um ajuste prosódico que encontra sua causa na estrutura melódica; (iii) o ajuste entre texto e melodia deve satisfazer às condições de boa formação da palavra cantada; e (iv) tais condições de boa formação constituem uma *gramática da palavra cantada*, um conjunto de regras de preferência para o *alinhamento* entre os elementos terminais (nota e sílaba), e a *associação* entre os não-terminais (pé, palavra, grupo clítico, frase fonológica e entoacional de um lado, e pulso, célula e frase melódica de outro).

2. CONDIÇÕES DE BOA FORMAÇÃO

A primeira condição de boa formação da palavra cantada diz respeito ao alinhamento entre os elementos terminais da melodia, as *notas*, e os elementos terminais da cadeia prosódica, as *silabas*.

Princípio de alinhamento dos terminais: *cada elemento terminal da cadeia melódica (nota) deve ser alinhado a um e apenas um constituinte terminal da cadeia prosódica (sílaba) e vice-versa.*

Por trás de um princípio aparentemente banal, esconde-se uma importante diferença entre palavra falada e palavra cantada, mais precisamente entre entoação¹ e melodia. Na entoação da fala, a sílaba pode ser associada a mais de um tom e este pode ser associado a mais de uma sílaba. Na melodia cantada, ao contrário, a relação entre tom e sílaba é sempre biunívoca. Convém esclarecer este ponto. A entoação é um conjunto limitado e não hierarquizado de padrões sonoros, que são encadeados um após o outro moldando-se ao conteúdo semântico, à estrutura sintática e à condição pragmática da enunciação. O inventário da gramática entoacional organiza-se em torno de apenas dois tons, alto (H) e baixo (L), cuja combinatória dá origem a um número limitado de acentos tonais e de tons de fronteira (Pierrehumbert, 1988:22). Em razão dessa limitação – limitação essa essencial para o bom funcionamento da entoação enquanto sistema paralinguístico – a relação entre tom e sílaba pode ser de um para um, de um para muitos ou de muitos para um. Segundo os princípios formulados por Goldsmith (1979: 27), (i) todas as vogais estão associadas a ao menos um tom e (ii) todos os tons estão associados a ao menos uma vogal. É claro que a associação entre um tom e uma vogal é apenas um caso particular decorrente de (i) e (ii).

¹ Nossa análise toma a entoação do ponto de vista fonológico (PierreHumbert, 1988; Ladd, 1996). Um enfoque fonético da questão revelaria certamente outros aspectos do problema que estamos tratando, e por si só exigiria um outro ensaio.

Suponhamos uma interlocução na qual é apresentada a seguinte informação: “O João vai a São Paulo amanhã só pra comprar uma camisa do Ronaldinho”. O interlocutor, surpreso, pode manifestar sua reação à informação por meio de diversas frases, por exemplo, “O João?”, “O João vai a São Paulo amanhã?”, “O João vai a São Paulo amanhã só pra comprar uma camisa do Ronaldinho?”, etc. Embora cada uma destas frases contenha um número diferente de sílabas e de palavras, todas apresentam um perfil entoacional similar, aquele que permite identificar a expressão de surpresa do interlocutor. Ou seja, *a entoação se caracteriza por ser uma cadeia definida de tons associada a uma cadeia indefinida de sílabas*. A elasticidade da entoação faz dela uma espécie de melodia moldável a enunciados de diferentes extensões. Daí a relação entre tom e sílaba poder ser de um para muitos, de muitos para um, ou de um para um.

É preciso ressaltar também que não são apenas as três frases mencionadas que portam a mesma entoação, mas um número indefinido delas, uma vez que a entoação associada ao conteúdo “surpresa” pode recobrir um sem número de manifestações linguísticas. Empregamos os mesmos morfemas entoacionais recorrentemente toda vez que expressamos os mesmos conteúdos. É na qualidade de morfemas entoacionais, portanto, que as entoações constituem uma classe fechada. Elas fazem parte de inventários relativamente estáveis da língua e esta é a razão pela qual não criamos a cada ato da fala novas entoações, mas apenas as adaptamos a novas cadeias de sílabas e palavras. Em suma, a entoação não é *produtiva*. Como mostra Pierrehumbert, na entoação temos uma *gramática de estados finitos*, de modo que os morfemas entoacionais apenas se sucedem linearmente uns aos outros sem constituir hierarquias. Ao contrário da entoação, a melodia é uma classe aberta e apresenta uma estrutura complexa, hierárquica e recursiva. Como discutiremos em detalhe mais adiante, as notas de uma melodia não constituem uma sequência linear, mas são estruturadas em grupos hierarquicamente organizados segundo parâmetros rítmicos, harmônicos e semióticos. Por essa razão, a produtividade da palavra cantada é ilimitada.² Sempre poderemos criar novas melodias e a elas associar textos, desde que respeitada a gramática da palavra cantada.

3. ALINHAMENTO NOTA-SÍLABA

O princípio basilar dessa gramática, princípio que a distingue da gramática entoacional, é a relação biunívoca nota-sílaba. Assim:

Melodia		Entoação			
<i>Sílaba</i>	σ	<i>Sílaba</i>	σ	σ	σ
			/¥	ou	¥/
<i>Nota</i>	N	<i>Tom</i>	T	T	T

² Evidentemente, se considerada em sua realização fonética, a entoação da fala expressiva pode ser muito mais rica que a melodia musical. No entanto, o que estamos comparando aqui não são as realizações fonéticas da cadeia falada e da cadeia melódica, mas as formas que lhes são subjacentes.

Em virtude do princípio de alinhamento, o número de sílabas da cadeia deve coincidir com o número de notas da melodia. É o que ocorre com a imensa maioria das canções, das quais o primeiro verso de *Asa Branca* serve de exemplo³.

<031.01>⁴

<i>Sílaba</i>	quan	doj	ei	a	te	rrar	den	do
<i>Nota</i>	N	N	N	N	N	N	N	N

Ocasionalmente encontramos vestígios da ação do cancionista no sentido de não violar o princípio do alinhamento. No terceiro verso da mesma *Asa Branca*, por exemplo, não há relação biunívoca entre o texto (4 sílabas) e a melodia (5 notas), o que o leva a criar a sílaba paragógica **ei*.

<031.03>

<i>Sílaba</i>	eu	per	gun	tei	eu	per	gun	te	*ei
				/¥	>				
<i>Nota</i>	N	N	N	NN	N	N	N	N	N

Não se pode atribuir a criação desta sílaba a outra causa senão ao princípio de alinhamento dos terminais, mesmo porque o processo se repete alguns versos adiante.

<031.16>

<i>Sílaba</i>	pra	mim	vor	ta	pra	mim	vor	ta	*a
				/¥	>				
<i>Nota</i>	N	N	N	NN	N	N	N	N	N

Se tirarmos todas as consequências do princípio de alinhamento dos terminais, seremos levados a desconsiderar o papel do plano do conteúdo no ajuste texto-melodia no nível da sílaba uma vez que a relação aqui é entre significante verbal (sílaba) e “significante” musical (nota). Consequentemente, as sílabas *ai* e *viu* deveriam ser consideradas apenas como sílabas preenchedoras de posições na linha melódica. Assim:

<031.03>

<i>Sílaba</i>	a	Deus	do	céu	a	Deus	do	céu	*ai
				/¥	>				
<i>Nota</i>	N	N	N	NN	N	N	N	N	N

³ Sempre que possível utilizaremos a representação ortográfica na transcrição das sílabas que compõem o texto. Os casos de ressilabação envolvendo ditongos serão representados pela semiconsoantes *j* e *w*, e os casos de alçamento ou abaixamento vocálico serão transcritos com os símbolos do AFI.

⁴ Todos os segmentos analisados foram extraídos do corpus do projeto “Hierarquia melódica e hierarquia prosódica: uma comparação possível?” (Processo Fapesp 2008/06656-7) e são referenciados entre colchetes angulares. Os três primeiros algarismos indicam o número da canção e os dois últimos o verso correspondente. No caso, <031.01> representa o primeiro verso da canção de número 31.

<031.19>

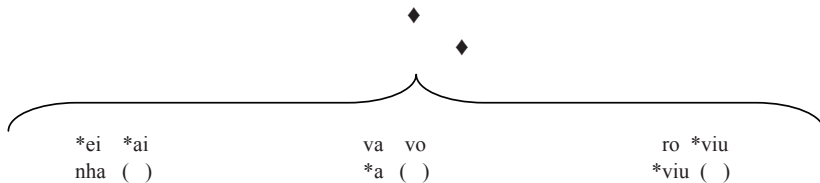
<i>Sílaba</i>	não	cho	re	não	não	cho	re	não	*viu
				/¥	>				
<i>Nota</i>	N	N	NNN	N	N	N	N	N	N

Tal interpretação dos dados impõe-se pelo modo quase mecânico com que o processo ocorre na canção.

<031.03-19>

<i>Sílaba</i>	eu	per	gun	te	*ei	a	deus	do	céu	*ai
	por	que	ta	ma	nha	ju	dí	a	ção	()
	eu	tja	se	gu	ro	não	cho	re	não	*viu
	kjeu	vol	ta	rei	*viu	meu	co	ra	ção	()
<i>Nota</i>	N	N	N	N	N	N	N	N	N	N

O trecho acima mostra que não se trata apenas de um alinhamento nota-sílaba. Parece que estamos diante de uma estrutura melódica rígida que governa a distribuição das sílabas, uma espécie de molde melódico. Em *Asa Branca*, as sílabas paragógicas ocorrem na primeira, quarta e quinta estrofes, ajustando-se a este molde:



As paragoges silábicas de *Asa Branca* estão longe de representar um caso isolado. Ao contrário, o fenômeno está associado ao princípio de alinhamento de elementos terminais, o qual nada mais faz que adaptar a cadeia de sílabas a uma estrutura ritmo-melódica recorrente, notadamente nas terminações das células melódicas (ver adiante item 6). Fenômeno similar, agora na forma de uma epêntese silábica, ocorre em *pre.gui.i.ça*, *le.i.las* e *te.re.e.zas*, em *Tempo de estio*:

<002.01-03>

<i>Sílaba</i>	que	ro	pre	gui	ça	>	que	ro	pre	gui	*i	ça
---------------	-----	----	-----	-----	----	---	-----	----	-----	-----	----	----

<002.015>

<i>Sílaba</i>	lan	ges	e	lei	las	>	lan	ges	e	le	*i	las
---------------	-----	-----	---	-----	-----	---	-----	-----	---	----	----	-----

<002.017>

<i>Sílaba</i>	cias	e	te	re	zas	>	cias	e	te	re	*e	zas
---------------	------	---	----	----	-----	---	------	---	----	----	----	-----

				/¥								
--	--	--	--	----	--	--	--	--	--	--	--	--

<i>Nota</i>	N	N	NNN	NNN	>	N	N	N	N	N	N	N
-------------	---	---	-----	-----	---	---	---	---	---	---	---	---

Paragoge e epêntese silábica constituem apenas dois dos processos de ajuste de base entre hierarquia prosódica e hierarquia melódica. Processos fonológicos como a ditongação, degeminação, haplologia e elisão são também empregados

com o propósito de alinhar sílaba e nota. É preciso destacar, no entanto, que nem sempre tais processos guiam-se pela fonologia do português brasileiro. Em *Leãozinho*, por exemplo, a diérese *léu > le.u* não pode ser vista como um fenômeno exclusivamente linguístico, mas como um fenômeno melódico-fonológico desencadeado pelo princípio de alinhamento, e que faz parte da gramática da palavra cantada, uma gramática melódico-fonológica.

<043.10>

<i>Sílaba</i>	l	h	e	d	o	u	r	a		p	e		l	e		a	o		l	é		u	
<i>Nota</i>	N	N	N	N	N	N	N	N		N	N	N	N	N	N	N	N	N		N	N	N	N

Até aqui examinamos situações nas quais o número de notas excede o número de sílabas. Mas o contrário também ocorre, sendo até mais frequente. É preciso esclarecer desde já que estamos diante de um fenômeno que apenas aparentemente se confunde com a escansão das sílabas poéticas utilizada na versificação. De fato, Saad Ali afirma que

“Na contagem de sílabas do interior do verso fundem-se frequentemente em sílaba única a terminação vocálica átona e o início vocálico da palavra imediata. Costuma-se dizer que houve “absorção” ou “elisão”, exigida pela técnica versificatória, como se o fenômeno fosse alheio ao falar de todos os dias. Se absorver, elidir, significam eliminar, suprimir ou incorporar uma coisa em outra, não se dará isso senão em caso de serem as vogais em contato iguais ou parecidas. Não há *com q’amor, led’ e cego, porqu’as mãos* senão por hipótese. Ninguém lê verso diferentemente da prosa comum, absorvendo vogais em cada linha. Nem há jeito de dar sumiço a algum dos fonemas em *já a luz, no assento, e os barões*, etc. E, apesar de tudo isso, sem fazermos violência à expressão usual, os versos soam com o número exato de sílabas e o desejado movimento rítmico” (Ali, 1999:23-24).

Nossos dados indicam que o princípio do alinhamento de terminais reestrutura a cadeia silábica nos moldes de uma escansão, com a importante diferença que a “violência à expressão usual” de que fala Ali pode ocorrer. Vejamos este ponto mais detidamente considerando alguns versos da canção *Valsa brasileira*.

<029.04-07>

Eu descartava os dias
Em que não te vi
Como de um filme
A ação que não valeu

Como temos aqui uma anástrofe (*como de um filme, a ação que não valeu* versus *como a ação de um filme que não valeu*), faz-se necessária uma pausa entre o terceiro e quarto versos ...*filme/a ação*, a qual bloqueia a ditongação *fil.me.a > *fil.mja*. Por outro lado, nada impede que ocorra a ditongação em *de.um > djum*.

No entanto, o fator decisivo na realização ou não das ditongações não é linguístico, mas musical. Para não violar o princípio de alinhamento dos terminais, a ditongação provável *de.um > djum* se realiza, e a ditongação improvável *fil.me.a > *fil.mja* também se realiza.

<029.06-07>

Sílaba	co	mo	de	um	fil	me	a	a	ção	co	mo	djum	fil	mja	ção
			¥ /			¥ /	>								
Nota	N	N	N	N	N	N		N	N	N	N	N	N	N	N

Dado que a fonologia do PB tem acesso à sintaxe, tal tipo de reestruturação não ocorreria na fala. Como se vê, porém, tudo indica que a gramática melódico-fonológica pode ignorar a sintaxe quando o que está em jogo é a preservação do princípio do alinhamento dos terminais.

Outro exemplo de que a gramática melódico-fonológica é relativamente independente da gramática fonológica pode ser observado numa degeminação anômala que ocorre em *O rouxinol*. Segundo Bisol (1996) a degeminação em PB está sujeita a uma restrição rítmica que limita sua aplicação ao encontro de duas vogais apenas quando a segunda não porta acento primário. Dada essa restrição, a degeminação ocorre em (a) mas não ocorre em (b).

(a) me.NI.na.a.LE.gre > me.NI.na.LE.gre

(b) me.NI.na.AL.ta > *me.NI.NAL.ta

No entanto, na canção *O Rouxinol*, essa restrição é violada e a degeminação se realiza a despeito da presença do acento na segunda vogal (sublinhada) :

<033.7>

Sílaba	tra	tei	da	su	a	<u>A</u> .sa	>	*tra	tei	da	su	<u>A</u> .sa
--------	-----	-----	----	----	---	--------------	---	------	-----	----	----	--------------

Mais uma vez, é o princípio de alinhamento dos terminais que explica esta violação, uma vez que a linha melódica possui seis notas e o texto sete sílabas, determinando, portanto, a degeminação, a despeito da restrição apontada por Bisol. Assim:

<033.07>

Sílaba	tra	tei	da	su	a	a	sa	tra	tei	da	su	a	sa
					¥ /		>>						
Nota	N	N	N	N	N	N		N	N	N	N	N	N

É interessante verificar que, apesar de soar antinatural quando falada, esta degeminação parece passar despercebida quando cantada. Voltaremos a este ponto mais adiante.

4. PRINCÍPIO DE ASSOCIAÇÃO

O princípio de alinhamento faz referência apenas às unidades terminais da melodia, sem qualificá-las. Ele explica porque ocorre a degeminação no verso de *O rouxinol*, mas não explica porque a distribuição de proeminências desse verso é totalmente alterada em relação ao padrão da fala. De fato, Gilberto Gil não canta

tra.TEI.da.SU.a.A.sa mas *TRA.tei.DA.su.A.sa*. Tal distribuição de proeminências pode ser explicada pela segunda condição de boa formação da palavra cantada, que diz respeito não mais ao *alinhamento*, mas à *associação* entre os elementos não-terminais da hierarquia melódica e da hierarquia prosódica.

Princípio de Associação: *Elementos não-terminais da hierarquia prosódica devem ser preferencialmente associados aos elementos não-terminais da hierarquia melódica.*

Quando a cadeia de sílabas é alinhada à cadeia de notas, os núcleos das divisões da melodia (pulso, célula, frase) atraem os núcleos dos constituintes prosódicos (pé, palavra, frase fonológica e frase entoacional), de onde resulta uma associação entre proeminências melódicas e proeminências linguísticas.

5. ASSOCIAÇÃO PULSO-PÉ

A mais básica dessas associações por atração ocorre entre os pulsos da melodia e as sílabas acentuadas do texto. A regra geral é a de que o pulso tem a capacidade de atrair o acento silábico. É o pulso que determina a distribuição do acento, de modo que sílabas a ele associadas são mais proeminentes que sílabas não alinhadas.

O verso de *O Rouxinol* que acabamos de analisar ilustra a relação pulso-acento.

<033.07>

<i>Sílaba</i>	tra	TEI	da	SU	a	A	sa	TRA	tei	DA	su	A	sa
					¥/		>>						
<i>Nota</i>	N	N	N	N	N	N		N	N	N	N	N	N
<i>Pulso</i>								P		P		P	

Embora pareça à primeira vista que o pulso (hierarquia melódica) se relacione ao pé da hierarquia prosódica, isso de fato ocorre apenas acidentalmente. Há uma diferença fundamental entre a métrica verbal e a métrica musical. Na primeira, o pé, unidade rítmica, organiza-se seja em torno da sílaba, seja em torno do acento, de onde a clássica (e ao que tudo indica anacrônica) divisão das línguas do mundo entre aquelas que têm ritmo silábico (o italiano e o espanhol, por exemplo) e aquelas de ritmo acentual (o inglês, por exemplo) (cf. Barbosa, 2006). No entanto, quaisquer que sejam as línguas naturais, temos que:

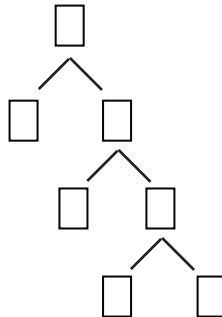
- (i) o inventário rítmico é reduzido; segundo Hayes (1992) existem apenas pés binários e degenerados;
- (ii) as estruturas rítmicas não são recursivas *stricto sensu*;

Na música, ao contrário, as estruturas rítmicas são muito mais complexas. Tal complexidade se deve, em primeiro lugar, ao fato de a gama de valores de duração

e intensidade ser mais ampla, uma vez que a relação entre seus constituintes não pode ser definida apenas em termos de proeminência relativa. Ao contrário, ela precisa ser plenamente especificada a partir de um inventário de dezesseis valores de duração e três valores de intensidade. Assim, se a língua atribui às quatro sílabas da frase fonológica *o.MEU.a.MOR* uma única representação estrutural possível, na qual se alternam valores W e S, na música, devido à riqueza de suas categorias rítmico-melódicas, a alternância W-S é apenas uma dentre inúmeras possibilidades. Por exemplo, na canção *O meu amor* (Chico Buarque) encontramos a sequência S-S-W-S:



Em segundo lugar, a música apresenta uma estrutura rítmica recursiva *stricto sensu*, cuja base é o conceito de divisão musical, expresso através da hierarquia de valores de duração.



Dizer que a divisão rítmica na música é recursiva significa dizer que é sempre possível encaixar dois ou mais valores de duração no intervalo antes ocupado por um único valor. Ou seja, um valor rítmico pode ser expandido ou condensado indefinidamente. Na palavra, ao contrário, aos valores rítmicos (os pés) são unidades fixas⁵.

Tais características combinadas – riqueza da gama de valores e recursividade – criam um abismo entre música e linguagem. Apesar disso, há uma tendência natural a associar o pulso melódico ao acento silábico, seja ele primário ou secundário. Dizemos que esta tendência é natural apenas para frizar que, sendo a música dotada de pulsos, e sendo a cadeia linguística dotada de proeminências, é natural que estas saliências se atraiam. Como vimos, o princípio de associação diz respeito sobretudo a uma *tendência à atração* entre elementos musicais e elementos verbais.

⁵ Não custa repetir que a perspectiva adotada aqui é a fonológica e não fonética.

Dado que o pulso atrai o acento, ele tem o poder de reestruturar o pé prosódico, como acabamos de ver em *O rouxinol*. De maneira semelhante, em *Tempo de estio*, o ritmo ternário transforma-se em binário pela associação com o pulso *QUE.ro.co.MER* > *que.RO.co.MER*.

<002.01-02>

<i>Sílaba</i>	que	RO	co	MER		que	RO	ma	MAR
					*	*			
<i>Nota</i>	N	N	N	N		N	N	N	N
<i>Pulso</i>			P		P		P		P

Em *Carinhoso*, ao contrário, é o ritmo binário que se transforma em ternário pela mesma razão: *meu.CO.ra.ÇÃO* > *MEU.co.ra.ÇÃO*.

<009.01-02>

<i>Sílaba</i>	MEU	co	ra	ÇÃO						NÃO	sei	por	QUE
					*	*	*	*	*	*	*		
<i>Nota</i>	N	N	N	N								N	N
<i>Pulso</i>	P			P			P		P			P	

O fato de o pulso atrair o acento explica a maioria dos deslocamentos acentuais, tão comuns na canção popular. Do ponto de vista da interface entre música e linguagem, é interessante observar o quanto tais deslocamentos afetam a gramática fonológica do português. Vejamos alguns casos.

Sabe-se que os clíticos são partículas átonas sintaticamente independentes (plano do conteúdo) mas fonologicamente dependentes da palavra (plano da expressão). Por conta desta dependência o clítico nunca é acentuado. Como a hierarquia melódica parece não “enxergar” nem a fonologia, nem a sintaxe, nem o discurso, o clítico pode sim ser acentuado na palavra cantada. É o que ocorre de maneira quase didática em *Xica da Silva*. No refrão desta canção o pulso impõe uma divisão binária onde esperaríamos uma divisão ternária *XI.ca.da.SIL* > *XI.ca.DA.%.SIL*. Tal reajuste leva à acentuação do clítico.

<027.04>

<i>Sílaba</i>	XI	ca	da	SIL		XI	ca	DA	SIL
					>				*
<i>Nota</i>	N	N	N	N		N	N	N	N
<i>Pulso</i>						P		P	P

A presença do pulso também determina processos de ressilabação. Em *Morro Dois Irmãos*, a elisão não ocorre porque o alinhamento de proeminências ao pulso torna tônica a vogal final da primeira palavra prosódica *NOU.tra* > *nou*. *TRA*, e a elisão apenas ocorre se ambas as vogais forem átonas (Bisol, 1992). Consequentemente, a elisão dá lugar a uma ditongação, que admite a vogal acentuada.

<004.06>

<i>Sílaba</i>	jo	que	se	RÁ	nou	traj	xis	TÊN
	*							
<i>Nota</i>	N	N	N	N	N	N	N	N
<i>Pulso</i>	P			P				P

A ação do pulso parece poder afetar também a qualidade da vogal da sílaba a ele associada. No bordão *Olê Peixe*, a vogal final postônica, reduzida, cede lugar a uma vogal plenamente especificada *PEI.xi* > *pei.XE*.

<020.01-02>

<i>Sílaba</i>	o	LE	pei	XE	o	LE	pei	XE
			*	*			*	*
<i>Nota</i>	N	N	N	N	N	N	N	N
<i>Pulso</i>		P	P	P	P	P	P	P

Já no bordão *Olô Porco*, é a vogal posterior que recebe o acento, tornando-se plenamente especificada *POR.cv* > *por.CO*.

<019.01-02>

<i>Sílaba</i>	o	LE	por	CO	o	LE	por	CO
			*	*			*	*
<i>Nota</i>	N	N	N	N	N	N	N	N
<i>Pulso</i>		P	P	P	P	P	P	P

Por fim, a canção infantil *Atirei o pau no gato* apresenta como que um síntese dos efeitos da associação pulso-acento. No grupo clítico *do.BE.ro*, reestruturado em *DO.be.RRO*, temos (i) a acentuação do clítico *do* > *DO*; (ii) o alçamento da vogal média-baixa *be* > *be*; e (iii) o abaixamento da vogal posterior *ro* > *ro*. Estes dois últimos processos, aliás, estariam em acordo com a fonologia do PB que prevê a neutralização das vogais postônicas. Segundo Câmara Jr.(1970) e Wetzels (1992), a partir de um sistema de sete vogais tônicas, ocorrem sucessivas simplificações em subsistemas dentro de determinados contextos (pretônico, postônico e postônico final) e dentro de certos domínios (palavra prosódica, pé e fronteira de palavra). A primeira neutralização, que diz respeito ao alçamento das médias ($\varrho > o$) e ($\varepsilon > e$), tem uma motivação morfológica, como se pode observar nas alternâncias em *porta* > *porteiro* e *belo* > *beleza*, por exemplo. Na palavra cantada, ao contrário, o alçamento tem motivação musical, uma vez que não há nenhuma transformação morfológica ou qualquer outro fator linguístico envolvido no processo.

No caso, ao transformar o pé troqueu em iambo *Be.ro* > *be.RO*, o deslocamento acentual impede a neutralização da átona final ($*o > u$) e, ao mesmo tempo, promove o alçamento da vogal média ($\varepsilon > e$).

<018.06-07>

<i>Sílaba</i>	DO	be	RRO	DO	be	RRO	kju	GA	to	DEU
				*						
<i>Nota</i>	N	N	N	N	N	N	N	N	N	N
<i>Pulso</i>	P	P	P	P		P	P			

6. ASSOCIAÇÃO CÉLULA-FRASE FONOLÓGICA

A relação pulso-acento, fundada em critérios exclusivamente rítmicos, explica apenas parcialmente a distribuição de proeminências na canção. Assim, no primeiro verso de *Cajuina*, as sílabas *MOS* e *TI* claramente se destacam das demais.

<001.01>

<i>Sílaba</i>	e	XIS	tir	<u>MOS</u>	a	QUE	se	RA	que	SE	des	<u>TI</u>	na
													* *
<i>Nota</i>	N	N	N	N	N	N	N	N	N	N	N	N	N
<i>Pulso</i>		P		P		P		P		P		P	P

O deslocamento de acento em *e.xis.TIR.mos* > *e.xis.tir.MOS* ocorre porque a sílaba *MOS* se associa a um pulso, e, como acabamos de ver, este atrai o acento. Mas, pode-se perguntar, por que o acento desloca-se para a direita, *e.xis.tir.MOS*, e não para a esquerda, *e.XIS.tir:mos*, se ambas as sílabas estão igualmente associadas ao pulso? Em princípio, não há direcionalidade no deslocamento acentual, de modo que *XIS* e *MOS* deveriam ser igualmente acentuadas. A resposta é que a sílaba *MOS* destaca-se das outras doze sílabas do verso porque está associada ao núcleo de uma célula melódica.

Convém explicar o que seja uma célula melódica. As notas de uma melodia tendem a organizar-se em grupos chamados células melódicas (Bas, 1978; Schoenberg, 1996). Os núcleos das células melódicas são as notas que recebem o acento harmônico, ou seja, as notas que coincidem com a posição nas quais ocorrem alterações harmônicas ao longo da melodia. Assim como a melodia é ritmicamente demarcada pelos pulsos – os quais percebemos intuitiva e naturalmente quando batemos os pés ou as mãos acompanhando uma melodia –, ela é também demarcada harmonicamente pelos núcleos das células, que são percebidos como os lugares onde, entre outras coisas, o violonista troca de acordes e o contrabaixista toca a nota mais longa da linha do baixo. A percepção destes pontos de demarcação da cadeia melódica é uma competência básica do músico, mesmo o iniciante e amador, muito mais dependente da exposição ao estímulo musical que de um treinamento escolar sistematizado. O que se denomina “tocar de ouvido” nada mais é que a competência para reconhecer, na cadeia melódica, os pontos em que ocorrem os ataques dos acordes (os pulsos, P) e os pontos em que ocorrem as mudanças de acordes (os núcleos das células, C).

Na melodia da canção popular, a célula melódica é uma unidade harmônica (há apenas um acorde para cada célula), rítmica (o núcleo da célula coincide com

a cabeça de um compasso musical) e melódica (muitas vezes a célula tem um perfil melódico que se repete ao longo da melodia). Por essa razão, a nota-sílaba associada ao núcleo da célula é proeminente em relação às demais. Neste verso de *Cajuína*, as duas células contêm os acordes de Dó menor e Fá menor. Estes acordes instauram uma direcionalidade harmônica mínima, condição essencial para a constituição da frase melódica.

<001.01>

<i>Sílaba</i>	e	XIS	tir	MOS	a	QUE	se	RA	que	SE	des	TI	na
<i>Nota</i>	N	N	N	N	N	N	N	N	N	N	N	N	N
<i>Pulso</i>		P		P		P		P		P		P	
<i>Cel</i>				C								C	

Uma observação ainda que rápida do verso de *Cajuína* já revela uma questão teórica interessante sobre o status da célula enquanto constituinte da hierarquia melódica. Se assumimos que a célula é um constituinte melódico – assim como a palavra prosódica e a frase fonológica são constituintes prosódicos –, então devemos explicitar como tal constituinte se forma e quais são seus limites. No entanto, tudo indica que os limites da célula somente podem ser estabelecidos em condições muito particulares. No exemplo de *Cajuína* é possível afirmar a existência dos núcleos de duas células – porque sabemos quais são os acentos harmônicos – mas não se pode dizer onde termina a primeira delas, onde principia a segunda, e mesmo se estes limites coincidem ou não.

Dado que a célula melódica é um conjunto de notas das quais uma única destaca-se das demais por receber o acento harmônico, ela teria como contraparte linguística a palavra prosódica, o grupo clítico ou a frase fonológica, que são cadeias de sílabas dotadas de uma única proeminência. De fato, a célula é uma espécie de *locução melódica*, ou seja, uma unidade musical que encontra seus limites na capacidade articulatória do gesto musical, seja este gesto vocal ou instrumental. Daí a tendência da célula melódica a associar-se a estes constituintes prosódicos. Por exemplo:

<045.01>

<i>Palavra</i>	[ω]
<i>Sílaba</i>	DRÃO
<i>Nota</i>	N
<i>Pulso</i>	P
<i>Cel</i>	C

<009.01>

<i>Gr. Clítico</i>	[C]
<i>Sílaba</i>	MEU	co	ra	ÇÃO	
<i>Nota</i>	N	N	N	N	
<i>Pulso</i>	P			P	
<i>Cel</i>				C	

<014.01>

<i>FrasFon</i>	[PhP]
<i>Sílaba</i>	te	nho	pen	SA	do
					*
<i>Nota</i>	N	N	N	N	N
<i>Pulso</i>	P			P	
<i>Cel</i>				C	

Vimos que a relação pulso-acento é de ordem exclusivamente rítmica. Trata-se de uma relação entre o plano da expressão musical (o pulso) e o plano da expressão verbal (o acento primário ou secundário). A relação da célula com o texto, por outro lado, parece envolver informação de ordem morfológica (portanto, plano do conteúdo). Os núcleos das células atraem os núcleos das palavras lexicais, ou seja, nomes, verbos, adjetivos e advérbios. No primeiro verso de *Cajuína* temos três palavras lexicais (existir, ser, destinar) que formam três frases fonológicas:

<001.01>

<i>FrasFon</i>	[PhP]	[PhP]	[PhP]
<i>Sílaba</i>	e	XIS	tir	MOS	a	QUE	se	RA	que	SE	des	TI	na
													* *
<i>Nota</i>	N	N	N	N	N	N	N	N	N	N	N	N	N
<i>Pulso</i>		P		P		P		P		P	P	P	P
<i>Cel</i>				C									C

Assim, a regra básica para a construção da célula melódica diz que seu núcleo deve ser preferencialmente associado ao núcleo de uma palavra lexical. Como temos duas células e três núcleos lexicais, um destes é descartado. O mesmo ocorre no décimo verso de *Anunciação*, onde temos três palavras lexicais (anunciar, sino, catedral).

<010.10>

<i>FrasFon</i>	[PhP]	[PhP]	[PhP]
<i>Sílaba</i>	eu	TjA	nun	CI	o	NOS	si	NOS	das	CA	te	DRAIS		
<i>Nota</i>	N	N	N	N	N	N	N	N	N	N	N	N	N	
<i>Pulso</i>		P		P		P		P		P		P		
<i>Cel</i>				C									C	

7. FRASE MELÓDICA-FRASE ENTOACIONAL

Uma frase musical elementar é constituída por duas células melódicas contíguas. A análise de melodias de canções populares mostra que o encadeamento de duas células melódicas cria uma espécie de predicação melódico-harmônica, responsável pelo efeito de sentido de frase musical.

Uma breve digressão é necessária para minimizar a estranheza que a expressão “predicação melódico-harmônica” possa causar ao leitor não especialista em

▲ *Gita* (Raul Seixas & Paulo Coelho):

<047.03>

FrasEnt	[IP]
FrasFon	[PhP]	[PhP]	[PhP]
Sílaba	as	VE	zes	vo	CÊ	me	per	GUN	ta	
Nota	N	N	N	N	N	N	N	N	N	N
Pulso		P			P				P	
Cel		C							C	
FrasMel									F	

▲ *Travessia* (Milton Nascimento & Fernando Brandt):

<057.01>

FrasEnt	[IP]
FrasFon	[PhP]	[PhP]	[PhP]	
Sílaba	quan	do	VO	cê	foi	em	BO	ra		
				*		*		*		
Nota	N	N	N	N	N	N	N	N	N	N
Pulso			P		P				P	
Cel			C						C	
FrasMel									F	

Observe-se que o núcleo de uma frase melódica coincide com a célula mais à direita e usualmente é seguido de pausa. Muitas vezes, este núcleo atrai a sílaba mais proeminente de uma frase entoacional. Nesses casos (como nos exemplos acima), o limite da frase melódica é determinado pelo alinhamento célula-pulso-nota, o qual por sua vez coincide com a sílaba mais proeminente do verso, pois é simultaneamente o núcleo da palavra prosódica, frase fonológica (PhP) e frase entoacional (IP). Desse modo, o núcleo da frase melódica coincide com a coluna mais alta da grade estrutural. Daí a naturalidade com a qual a frase melódica se associa à frase entoacional, que por sua vez tende a servir de base para a construção do verso na canção popular.

8. PARA ALÉM DA FRASE

A frase melódica parece ser a unidade superior da hierarquia melódica. Ou seja, ela é a mais extensa cadeia de notas a constituir dependências internas (Hjelmslev, 1971:109). Por essa razão, seu estudo merece um papel de destaque dentro da teoria semiótica da canção e, de maneira mais ampla, dentro dos estudos sobre as relações entre música e linguagem (Carmo Jr, 2007). O fato de a frase ter tal papel dentro da hierarquia melódica não significa que a melodia não possua estruturas superiores, mas apenas que tais estruturas se comportam muito mais como formas cristalizadas pelo uso, padrões recorrentes de construção melódica que se servem da combinação de elementos musicais de modo mais ou menos livre. Trata-se de formas melódicas fixas, assim como as formas fixas da versificação, que nada mais são que combinações padronizadas de certo número de versos. A forma *soneto*, por exemplo, não constitui uma estrutura, no sentido forte da

palavra, mas apenas uma forma canônica da poesia lírica que se fixou pelo uso. Assim também na melodia.

Embora o número de tais combinações seja muito grande, ele certamente não é ilimitado. Uma taxonomia da melodia popular baseada em tais esquemas poderia nos ajudar a entender melhor como certas combinações estão associadas a certos gêneros e principalmente a certos estilos tensivos. Uma tal taxonomia, no entanto, foge aos propósitos do presente ensaio.

9. CONCLUSÃO

O problema que está na origem deste ensaio diz respeito à possibilidade ou não de nos servirmos de modelos linguísticos como um recurso auxiliar na análise de melodia musical. Até que ponto um olhar linguístico sobre a melodia musical poderia auxiliar o musicólogo a aprimorar seus métodos e ferramentas de análise? E, por que não, em que medida esse tipo de pesquisa interdisciplinar poderia contribuir para nosso conhecimento da língua (cf. Massini-Cagliari, 2008 e 2009; Gilbers & Schreuder, 2002). Acreditamos que o problema aqui proposto representa apenas uma das faces de um outro, mais amplo, acerca da abrangência dos modelos linguísticos e da natureza semiótica da música e das línguas naturais.

Nossa investigação teve como ponto de partida a busca de uma hipotética hierarquia melódica, que suspeitávamos ser similar à hierarquia prosódica das línguas naturais. Nossos resultados mostram que a melodia exhibe sim uma estrutura hierárquica, porém ela se diferencia da hierarquia prosódica na medida que não possui constituintes claramente delimitados. Em função disso, e diferentemente dos domínios prosódicos, as estruturas melódicas pedem uma representação em grade, que não faz referência aos limites das unidades. Os principais resultados deste ensaio podem agora ser sumarizados.

- ♣ A hierarquia melódica é constituída de quatro níveis, *nota*, *pulso*, *célula* e *frase*;
- ♣ Ela estabelece diferentes relações com os contituíntes prosódicos, o *alinhamento* entre nota e sílaba, e a *associação* entre as unidade superiores.
- ♣ Dado o não-isomorfismo entre as duas hierarquias, são frequentes os ajustes entre texto e melodia na palavra cantada. Tais ajustes são perceptíveis na forma da realização de processos fonológicos estranhos à fonologia do PB ou na forma de bloqueios onde os processos seriam esperados;
- ♣ A partir das propriedades da hierarquia melódica tornam-se mais claras as diferenças entre entoação e melodia. A entoação é uma gramática de estados finitos, com um inventário limitado de morfemas entoacionais. A entoação é uma classe fechada. A melodia, ao contrário, possui um inventário de sons muito mais rico, e uma estrutura recursiva, de onde provém sua produtividade ilimitada. A melodia constitui uma classe aberta.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALI, M.S. (1999). *Versificação portuguesa*. São Paulo: Edusp.
- BARBOSA, P. (2006). *Incursões em torno do ritmo da fala*. Campinas: Pontes/Fapesp.
- BAS, J. (1972). *Tratado de la Forma Musical*. Buenos Aires, Ricordi. 6a ed.
- BISOL, L. (1992). “Sândi vocálico externo”. in: *Gramática do português falado*, Campinas: Editora da UNICAMP, vol.2, p. 21-38.
- BISOL, L. (1996). “O sândi e a ressilabação”. in: *Letras de Hoje*, v.31, n.2, p.159-168.
- CÂMARA, J.M. (1970). *Dicionário de Filologia e Gramática*. Petrópolis: Vozes.
- CARMO JR. (2007). *Melodia & Prosódia: um modelo para a interface música-fala com base no estudo comparado do aparelho fonador e dos instrumentos musicais reais e virtuais*. Tese de doutorado, São Paulo: FFLCH/USP.
- GILBERS, D & SCHREUDER, M.J. (2002). “Language and music in Optimality Theory”. In: *Rutgers Optimality Archive* (<http://roa.rutgers.edu>), ROA # 517-0103.
- GOLDSMITH, J. (1979). *Autosegmental phonology*. New York: Garland Press.
- HAYES, B. (1995). *Metrical Stress Theory*. Chicago: The University of Chicago Press.
- HJELMSLEV, L. (1975). *Prolegômenos a uma teoria da linguagem*. São Paulo: Perspectiva.
- LADD, D.R. (1996). *Intonational Phonology*. Cambridge University Press, Cambridge.
- MASSINI-CAGLIARI, G. (2008). “Das cadências musicais para o ritmo lingüístico: uma análise do ritmo lingüístico do português arcaico, a partir da notação musical das Cantigas de Santa Maria.”. In: *Revista da ABRALIN*, v. 7, p. 9-26.
- PIERREHUMBERT, J. (1988). *The phonology and phonetics of English intonation*. PhD thesis, MIT. Indiana University Linguistics Club.
- SCHOENBERG. (1996). *Fundamentos da composição musical*. São Paulo:EDUSP.
- WETZELS, L. (1992). “Midi vowel neutralization in Brazilian Portuguese”. Coimbra: Phonology Workshop.