

SOBRE AS RELAÇÕES ENTRE NOTA E SÍLABA NA PALAVRA CANTADA

José Roberto do CARMO JR.*

- **RESUMO:** A palavra cantada apresenta frequentemente construções anômalas do ponto de vista prosódico. Tudo indica que tais construções não podem ser descritas exclusivamente pelo ambiente linguístico em que ocorrem. A excepcionalidade dessas construções se manifesta na realização de processos fonológicos em ambientes nos quais tais processos seriam bloqueados e, reciprocamente, em seu bloqueio em ambientes nos quais seria esperada a sua realização. O objetivo do presente trabalho é descrever esse comportamento divergente da palavra cantada com relação à palavra falada. Inicialmente apresentaremos exemplos que ilustram a dimensão e a generalidade do problema. Em seguida argumentaremos que a excepcionalidade constatada pode ser explicada se assumirmos a hipótese de que a relação entre texto e melodia na palavra cantada é determinada por condições de boa formação. Uma dessas condições, que denominamos Pareamento Métrico, estabelece que a relação entre nota e sílaba é sempre bijetiva, ou seja (I) toda e qualquer nota de uma melodia deve ser pareada a uma e apenas uma única sílaba, e (II) toda e qualquer sílaba de uma cadeia deve ser pareada a uma e apenas uma única nota. A observância dessa condição de boa formação explica tanto a realização quanto o bloqueio de processos fonológicos em divergência com a fonologia do português falado.
- **PALAVRAS-CHAVE:** reestruturação silábica; processos fonológicos; melodia; palavra cantada; prosódia.

O problema

Não é raro encontrar na palavra cantada construções prosodicamente anômalas. Irrealizáveis na fala, em quaisquer dos seus registros, essas construções, supostamente, passam despercebidas quando cantadas. Tudo indica que esse fenômeno não ocorre fortuitamente, mas depende de contextos extralinguísticos determinados. Em tais contextos, processos fonológicos que fazem referência à sílaba e ao acento podem ser bloqueados; paralelamente, processos fonológicos via de regra bloqueados na fala

* Universidade Federal do Paraná (UFPR). Curitiba – PR – Brasil. carmojr@ufpr.br. Orcid: 0000-0002-2293-9962.

passam a se realizar no canto. Compare-se, por exemplo, a realização falada (1a) e cantada (1b) de um verso de *Rouxinol* (Gilberto Gil)¹.

(1a) TraTEI da sua Asa

(1b) #TRAtEi da suø Asa

Em (1b)² observa-se, aparentemente, uma retração acentual (traTEI > TRAtEi) e uma degeminação (sua Asa > suAsa). Falamos em retração e degeminação “aparentes” porque, de fato, em nenhum dos dois casos há ambiente para esses processos fonológicos. O ambiente para a retração é a colisão acentual (Abousalh, 1997), fenômeno que se dá entre duas sílabas adjacentes, portadoras de acento primário e pertencentes a palavras de uma mesma frase fonológica. Por exemplo, (caFÉ QUENte > CAfé QUENte) e (JeSUS CRISto > JESus CRISto) apresentam retração acentual. Não é o caso (traTEI > TRAtEi). Do mesmo modo, convém falar em degeminação aparente em (sua Asa > suAsa) porque, como mostra Bisol (2002, p. 66), a degeminação está sujeita a restrição rítmica. A degeminação aplica-se apenas quando a segunda das vogais em contato não é acentuada, do contrário ela é bloqueada. Dado que em (1b) a segunda das vogais porta acento primário, a degeminação deveria ser bloqueada, porém isso não ocorre na interpretação de Gilberto Gil. Em outras palavras, embora em (1b) de fato ocorra o deslocamento do acento em (TRAtEi) e a supressão de uma vogal em (suAsa), não se pode afirmar que se trata do mesmo fenômeno que conhecemos como retração e degeminação na fala natural.

Verifica-se algo similar quando contrastamos a realização falada (2a) e cantada (2b) de um verso de *Valsa brasileira* (Chico Buarque e Edu Lobo), interpretada por Chico Buarque³.

(2a) Como de um filme // A ação que não valeu

(2b) #Como de um filmjação que não valeu

Aqui temos degeminação associada à ditongação (filme, a ação > filme ação > filmjação), sem que se possa determinar a ordem em que tais processos ocorrem. Qualquer que seja essa ordem, no entanto, parece-nos uma construção estranha à fonologia do português. Dada a ocorrência da anástrofe (“como de um filme, a ação que não valeu” *versus* “como a ação de um filme que não valeu”), e dado que a fonologia tem acesso à sintaxe (Abaurre, 1996), faz-se necessária uma pausa entre os versos (um filme//a ação). Conseqüentemente, a ditongação (fil.mja) deveria ser bloqueada, o que não ocorre quando a canção é interpretada por Chico Buarque.

¹ O ROUXINOL. Intérprete: Gilberto Gil. Compositores: G. Gil e J. Mautner. In: REFAZENDA. Intérprete: Gilberto Gil. [S. l.]: Philips Records Brasil, 1975. 1 CD, faixa 9.

² O símbolo # indica uma construção prosódica agramatical

³ VALSA BRASILEIRA. Intérprete: Chico Buarque. Compositores: C. Buarque e E. Lobo. In: CHICO BUARQUE - 1989. Intérprete: Chico Buarque. [S. l.]: RCA Brasil, 1989. 1 LP, faixa 10.

Degeminação (1b) e ditongação (2b) são processos de reestruturação silábica em que ocorre redução da quantidade de sílabas da cadeia. Reestruturação pelo incremento de sílabas é também comum na palavra cantada, embora não se confunda com nenhum processo fonológico conhecido.

Em (3a) e (3b) temos realizações falada e cantada, respectivamente, do seguinte verso de *Asa branca* (Luiz Gonzaga e Humberto Teixeira), interpretada por Luiz Gonzaga⁴.

(3a) Eu per gun tei a Deus do céu ai

(3b) #Eu per gun te ej a Deus do céu ai

A diferença entre (3a) e (3b) decorre unicamente da quantidade de sílabas, nove e dez respectivamente. Caso esse incremento resultasse de diérese (per.gun.tej > per.gun.te.i), nada haveria de estranhar, por se tratar do recurso habitual do cancionista e do poeta. Porém o que se observa em (3b) não é uma diérese (tej > te.i), mas uma estranha “epêntese silábica” (tej > te.ej) que preserva o ditongo, ou seja, uma reestruturação silábica que não encontra paralelo na língua falada. Os casos genuínos de diérese, por outro lado, não deixam de apresentar um interesse particular.

Assim, em (4a) e (4b) temos as realizações falada e cantada de um verso de *Tempo de estio* (Caetano Veloso)⁵.

(4a) O Rio es tá chei o de So lan ges e Lej las

(4b) O Rio es tá chei o de So lan ges e Le i las

Aqui, a diérese em (4b) parece bastante clara (Lej.las > Le.i.las). Porém é interessante comparar (4a/b) com os versos homólogos (5a/b), que ocupam idêntica posição na estrofe seguinte.

(5a) A nas e Ma ri nas e Lu cias e Te re zas

(5b) #A nas e Ma ri nas e Lu cias e Te re e zas

Assim como em (3b), em (5b) também temos inserção de uma sílaba epentética (Te.re.zas > Te.re.e.zas). Dado o paralelismo estrutural entre (Te.re.zas > Te.re.e.zas) e (Lej.las > Le.i.las), somos levados a nos questionar sobre o elemento desencadeador da diérese que aparentemente ocorre em (4b). Voltaremos a esse ponto mais adiante.

Sumarizando o problema descrito até aqui, a palavra cantada, assim como a palavra falada, está sujeita a processos de reestruturação silábica, seja por redução ou por incremento de material fonético. Na fala, tais processos de reestruturação são

⁴ ASA BRANCA. Intérprete: Luiz Gonzaga. Compositores: L. Gonzaga e H. Teixeira. In: NOVA HISTÓRIA DA MÚSICA POPULAR BRASILEIRA – 11. Intérprete: Luiz Gonzaga. [S. l.]: RCA Brasil, 1977. 1 LP, faixa 3, lado B.

⁵ TEMPO DE ESTIO. Intérprete: Caetano Veloso. Compositor: C.Veloso. In: MUITO - DENTRO DA ESTRELA AZULADA. Intérprete: Caetano Veloso. [S. l.]: Polygram/Phillips Brasil, 1978. 1 LP, faixa 2.

balizados por ambientes linguísticos determinados e pelo estilo de enunciação (mais ou menos formal, maior ou menor velocidade da fala). Diferentemente do que ocorre na fala, muitas das reestruturações silábicas observadas na palavra cantada não se deixam explicar apenas pelo ambiente linguístico. De fato, as construções que acabamos de observar parecem-nos estranhas do ponto de vista da fonologia do português falado. Como o objetivo deste trabalho é descrever tais construções e propor uma hipótese para as ocorrências, apresentaremos inicialmente a hipótese das condições de boa formação da palavra cantada para, então, discutirmos, à luz de tal hipótese, alguns dos processos que fazem referência à sílaba e ao acento, e à maneira como estes se realizam na palavra cantada.

Condições de boa formação da palavra cantada

Até o momento, temos feito uma descrição contrastiva de construções faladas e cantadas, como se essas fossem comparáveis entre si. Um exame mais acurado da natureza da palavra cantada mostra que esse procedimento não se sustenta. Não se pode tomar a palavra cantada como se esta fosse constituída, como a palavra falada, pela simples interação entre uma cadeia segmental e uma cadeia prosódica. Vejamos essa questão detidamente.

Do ponto de vista estritamente linguístico, a fala se constitui pela sobreposição de duas cadeias, segmental (segmentos e respectiva organização em sílabas) e suprasegmental ou prosódica (acento, entoação e ritmo). A análise fonológica faz referência a uma, a outra, ou à interação entre essas duas cadeias. A palavra cantada, diferentemente, é dotada de uma melodia que não se confunde com a entoação da fala, um ritmo que não se confunde com o ritmo da fala e uma acentuação que não se confunde com a acentuação primária e secundária da fala natural. Portanto, qualquer tentativa de análise da palavra cantada tem que considerar o fator “melodia”, um elemento extralinguístico aparentemente capaz de ativar ou bloquear processos fonológicos.

Diante desse quadro, duas abordagens metodológicas para descrever a palavra cantada se apresentam, ambas teoricamente plausíveis:

- (I) na primeira, mais afim com nossa intuição do que seja o canto, a palavra cantada pode ser considerada como a fala acompanhada de uma melodia musical. Nessa perspectiva, a palavra cantada seria uma variedade da palavra falada constituída, não por duas, mas por três cadeias sobrepostas: segmental, prosódica e melódica.
- (II) na segunda, a cadeia melódica na palavra cantada não coexiste com a cadeia prosódica, mas a substitui. Nesse ponto de vista, a palavra cantada seria uma variedade da palavra falada constituída por duas cadeias sobrepostas: segmental e melódica.

Adotaremos a segunda abordagem no presente trabalho. Trata-se de uma opção estritamente metodológica, teoricamente robusta e econômica. Essa hipótese emerge da observação de que melodia e prosódia compartilham o mesmo material fonético, ou seja, ambas se resolvem nos mesmos traços de altura, duração e intensidade. Nessa perspectiva, a melodia musical, de um lado, e a prosódia (ritmo, entoação e acento), de outro, seriam apenas modos diferentes de organizar e estruturar o mesmo material fonético, constituído por traços de altura, duração e intensidade, o que se torna particularmente claro quando consideramos as semelhanças e diferenças entre entoação e melodia.

Embora seja usual referir-se à entoação como a “melodia da fala” (Waugh, 1980; Bolinger, 1989), há uma diferença fundamental entre ambas e, por extensão, entre fala e canto. A unidade da entoação é o tom, ou seja, a frequência fundamental (f_0) presente no núcleo de toda sílaba. A entoação da fala organiza-se em torno de apenas dois tons, alto (H) e baixo (L), cuja combinatória dá origem a uma quantidade limitada de acentos tonais e de tons de fronteira (Pierrehumbert, 1980)⁶. Em razão dessa limitação — essencial para o funcionamento da entoação enquanto sistema paralinguístico — a relação entre tom e vogal (o núcleo da sílaba) não é bijetiva, ou seja, a relação entre tom e vogal pode ser de um para um, de um para muitos, ou de muitos para um. Na formulação de Goldsmith (1976, p. 27): (i) todas as vogais estão associadas a ao menos um tom; e (ii) todos os tons estão associados a ao menos uma vogal. A associação entre um único tom e uma única vogal é apenas um caso particular decorrente de (i) e (ii), como mostra (6).

(6)

Tom	T T	T	T
	\ /		/ \
Sílaba	σ	σ	$\sigma \sigma$

Comparada com a melodia, a entoação tem uma natureza “elástica”, que lhe permite moldar-se a qualquer cadeia linguística. Assim, as entoações constituem inventários relativamente estáveis da língua, razão pela qual não criamos novas entoações a cada ato da fala, mas, graças a sua elasticidade, apenas as adaptamos a novas cadeias de sílabas. Em suma, a entoação não é produtiva.

Outra característica importante da entoação é o fato de não ser recursiva. Conforme hipótese formulada por Pierrehumbert (1980, p. 29), a entoação é uma gramática de estados finitos, de modo que os morfemas entoacionais apenas se sucedem linearmente sem constituir hierarquias, moldando-se ao conteúdo semântico, à estrutura sintática e à condição pragmática da enunciação.

⁶ Neste momento, em que procuramos enfatizar mais as diferenças que as semelhanças entre entoação e melodia, parecemos indiferente tomar como modelo de entoação as propostas de Pierrehumbert (1980).

A melodia tem natureza diversa. Em primeiro lugar, a unidade da melodia não é o tom, mas a nota (N) e, diferentemente do que ocorre com a entoação, a relação entre nota e sílaba é bijetiva, ou seja, é uma relação de um para um, como mostra (7).

(7)	*	*
Nota	N N	N
	\ /	
Sílaba	σ	σ

Afirmar que na palavra cantada a relação entre nota (N) e sílaba (σ) é necessariamente bijetiva, é o mesmo que afirmar que ela está sujeita a uma restrição. Na palavra cantada, não é possível associar uma única nota a mais de uma sílaba, como também não é possível associar uma única sílaba a mais de uma nota. Dada a universalidade dessa restrição, ou seja, dado que em toda melodia cantada a relação entre nota (constituente terminal da melodia) e sílaba (constituente terminal do texto) é de um para um, sem exceção, pode-se afirmar que essa relação constitui uma das condições de boa formação da palavra cantada, a qual denominamos Pareamento Métrico.

(8) Pareamento Métrico.

Na palavra cantada, cada terminal da cadeia melódica (N, nota) deve ser pareado a um e apenas um terminal da cadeia silábica (σ , sílaba) e vice-versa.

Regras de reestruturação

Estamos agora em condições de formular uma hipótese para descrever a ocorrência de construções prosódicas anômalas na palavra cantada, das quais acabamos de ver alguns exemplos. Sendo uma das condições de boa formação da palavra cantada, o Pareamento Métrico é inviolável. Na eventualidade de uma desassociação entre notas e sílabas, dá-se necessariamente um ajustamento entre a cadeia melódica e a cadeia silábica na forma de apagamento ou de inserção de sílabas e/ou notas. São quatro os cenários possíveis para esse ajustamento, os quais correspondem a quatro regras:

- (i) regra do apagamento da sílaba — apague σ ;
- (ii) regra da inserção da sílaba — insira σ ;
- (iii) regra do apagamento da nota — apague N;
- (iv) regra da inserção da nota — insira N.

Neste artigo nos deteremos sobretudo nos dois primeiros casos, pois são os únicos pertinentes para análise linguística. Uma vez aplicadas as regras apague σ e insira σ , a

cadeia silábica pode ser reestruturada, sujeitando-se aos processos fonológicos usuais. Vejamos cada um desses casos.

(9) apague σ ($\sigma \rightarrow \emptyset / N_n < \sigma_n$)

Se a quantidade de sílabas de uma cadeia é maior que a quantidade de notas da melodia, apagam-se as sílabas para satisfazer (8).

A aplicação de apague σ explica a aparente degeminação em (1b). De fato, o verso correspondente é constituído por sete sílabas, mas a melodia por apenas seis notas, como se pode ver na figura 1.

Figura 1 – excerto de *Rouxinol* (Gilberto Gil)



Fonte: Elaboração própria.

A não associação entre terminais resolve-se pela aplicação de apague σ , de modo a satisfazer (8). Assim:

(1a)

Nota	N	N	N	N	N	N
					/ \	
Sílaba	Tra	tei	da	su	a a	sa

apague σ ($\sigma \rightarrow \emptyset / N_n < \sigma_n$)

(1b)

Nota	N	N	N	N	N	N
Sílaba	Tra	tei	da	su	a	sa

Nessa perspectiva o apagamento da sílaba não resulta de um processo fonológico *stricto sensu* (no caso, a degeminação), mas de um processo melódico-fonológico específico da palavra cantada. Pode-se objetar que (8) seria igualmente satisfeito sem que fosse necessária a degeminação, mas apenas ditongando-se o hiato (su.a > swa), como pode ser visto em (1c).

	N	N	N	N	N	N
(1b)	Tra	tei	da	su	a	sa
(1c)	Tra	tei	da	swa	a	sa

No entanto, nossa análise tem por objeto fonogramas específicos, no presente caso, a interpretação da canção feita por Gilberto Gil indicada mais acima. Nesse fonograma, o intérprete canta (1b) e não (1a) ou (1c).

Vejamos agora como a regra apague σ pode explicar o que ocorre em (2a) e (2b), em que se observa uma ditongação que não se realizaria em condições normais (filme, a ação > filme ação > filmjação). Novamente, parece que o fator decisivo na realização ou não de um processo fonológico é extralinguístico. Vê-se que (2a) é constituído por treze sílabas ao passo que a melodia correspondente é constituída por dez notas, como se observa na figura 2.

Figura 2 – Excerto de *Valsa brasileira* (Chico Buarque e Edu Lobo)



Fonte: Elaboração própria

Para que (8) não seja violado, aplica-se apague σ . Assim:

(2a)

N	N	N	N	N	N	N	N	N	N	N
		/ \		/ \						
co	mo	de um	fil	me a	a	ção	que	não	va	leu

apague σ ($\sigma \rightarrow \emptyset / N_n < \sigma_n$)

(2b)

N	N	N	N	N	N	N	N	N	N	N
co	mo	dewm	fil	mja	a	ção	que	não	va	leu

Observe-se que a ditongação (de um > djum) não apresenta problema algum, pois já ocorre na fala e, por isso, não é considerada em nosso argumento. Já a ditongação + degeminação em (filme, a ação > filmjação) não nos parece ser construção prosodicamente aceitável no português falado. Portanto, somente pode ser produto da associação entre cadeia segmental e cadeia melódica, associação esta que tem (8) como princípio de boa formação.

Pelo Pareamento Métrico, a quantidade de notas deve ser igual à de sílabas. Já vimos como essa condição é satisfeita pelo apagamento de sílabas quando a quantidade destas é superior à quantidade de notas da cadeia ($N_n < \sigma_n$) pela aplicação de apague σ ($\sigma \rightarrow \emptyset / N_n < \sigma_n$). Quando, inversamente, a quantidade de notas é superior à quantidade de sílabas ($N_n > \sigma_n$), aplica-se insira σ ($\emptyset \rightarrow \sigma / N_n > \sigma_n$).

(10) insira σ . ($\emptyset \rightarrow \sigma / N_n > \sigma_n$)

Se a quantidade de notas excede a quantidade de sílabas de uma cadeia, inserem-se sílabas para satisfazer (8). Essa regra somente pode ser verificada indiretamente. Assim, pela simples comparação entre (3a) e (3b), não se pode afirmar categoricamente que a sílaba *ej* tenha sido inserida para não violar (8). Em uma primeira análise, a inserção poderia ser meramente fruto da liberdade poética do cancionista. Porém vimos que o Pareamento Métrico determina a relação biunívoca nota/sílaba, o que implica a impossibilidade de se cantar uma nota sem que lhe seja associada uma sílaba, qualquer que seja ela. É o que ocorreria em (6a), dado que a quantidade de notas da melodia (dez), excede a quantidade de sílabas do verso (nove).

(3a)

N	N	N	N	N	N	N	N	N	N
			\	/					
eu	per	gun	te	j	a	Deus	do	céu	ai

Há em (3a) clara violação de (8). Nossa hipótese é que (3a) reestrutura-se em (3b), satisfazendo (8).

(3b)

N	N	N	N	N	N	N	N	N	N
eu	per	gun	te	ej	a	Deus	do	céu	ai

Essa hipótese é corroborada pela comparação com versos homólogos da canção que se alinham à mesma melodia de (3a). Assim:

	N	N	N	N	N	N	N	N	N	
(3b)	Eu	per	gun	te	ej	a	Deus	do	céu	ai
(3c)	por	fal	ta	d'á	gua	per	di	meu	ga	do
(3d)	En	ton	cew	di	sse	a	deus	Ro	si	nha
(3e)	Es	pe	rwa	chu	va	ca	ir	de	no	vo
(3f)	Eu	tja	sse	gu	ro	não	cho	re	não	viu

Observam-se aqui diversas reestruturações, todas previsíveis na fala, com exceção da inserção da sílaba ej, motivada, conforme o que acabamos de expor, pelo Pareamento Métrico.

O mesmo argumento se aplica a (4a), onde aparentemente ocorre uma diérese em (Lej.las > Le.i.las). Aqui também a quantidade de notas excede a de sílabas. A melodia é constituída por quatorze notas, ao passo que o verso é constituído por treze sílabas.

(4a)

N	N	N	N	N	N	N	N	N	N	N	N	N	N	
												\	/	
O	Rio	es	tá	chei	o	de	So	lan	ges	e	Lej	las		

Pela aplicação de insira σ , (4a) reestrutura-se em (4b)

(4b)

N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
O	Rio	es	tá	chei	o	de	So	lan	ges	e	Le	i	las	

Em princípio poderíamos descrever a reestruturação (Lej.las > Le.i.las) como simples diérese motivada pela métrica poética, sem relação alguma com a melodia da canção. No entanto, dada a rigidez estrutural que caracteriza a melodia, a inserção via diérese em (4b) se repete em (4c), agora como epêntese.

	N	N	N	N	N	N	N	N	N	N	N	N	N	N
(4b)	O	Rio	es	tá	chei	o	de	So	lan	ges	e	Le	i	las
(4c)	A	nas	e	Ma	ri	nas	e	Lu	cias	e	Te	re	e	zas

Parece-nos que a hipótese mais defensável é que, seja pela diérese (4b), seja pela epêntese (4c), em ambos os casos temos a aplicação da regra insira σ , a qual impede a violação do Pareamento Métrico. Resta-nos apresentar as regras de inserção e apagamento de notas, cujo interesse é muito mais musicológico que linguístico, porque a inserção ou apagamento de sílabas se manifesta como um processo fonológico cuja motivação última, como acabamos de ver, é musical. Já a inserção ou apagamento de notas, ao contrário, se apresenta como um “processo” melódico que tem origem no texto. No primeiro caso, a melodia impõe sua estrutura à cadeia silábica; no segundo, a integridade semântico-sintático-fonológica do texto é preservada, forçando a adaptação da melodia ao texto.

(11) apague N ($N \rightarrow \emptyset / \sigma_n < N_n$)

Se a quantidade de sílabas de uma cadeia é menor que a quantidade de notas da melodia, apagam-se as notas para satisfazer (8).

A estrofe (12 a–d) com os quatro versos iniciais de *Gîtá*, interpretada por Raul Seixas⁷, ilustra a regra (11), ao mesmo tempo em que nos permite destacar a diferença entre um poema rigidamente estruturado pela métrica, e uma canção, cuja métrica é muito mais flexível.

	N	N	(N)	N	N	N	N	N	N
(12a)	As	ve	zes	vo	cê	me	per	gun	ta
(12b)	Por	kje	∅	kjew	sou	tão	ca	la	do
(12c)	Não	fa	lo	dea	mor	qua	se	na	da
(12d)	Nem	fi	co	so	rin	dwaw	teu	la	do

Em primeiro lugar, observamos que, enquanto (12 a, c e d) possuem nove sílabas e nove notas, (12b) possui oito sílabas e oito notas, logo, em consonância com o Pareamento Métrico (8). Essa disparidade entre a quantidade de sílabas dos versos que compõem uma estrofe é bastante comum na canção popular, e nesse aspecto a letra de uma canção se distingue da poesia tradicional e da poesia popular, ambas fortemente atreladas à métrica. Quando o letrista toma a liberdade de diminuir ou aumentar a quantidade de sílabas de determinado verso — não importa a razão para que o faça — a quantidade de notas da melodia tem que ser necessária e correspondentemente aumentada ou diminuída. É o que verificamos em (12b), pela aplicação de apague N. Assim:

(12b₁)

	N	N	N	N	N	N	N	N
Por	kje	∅	kjew	sou	tão	ca	la	do

apague N ($N \rightarrow \emptyset / \sigma_n < N_n$)

(12b₂)

	N	N	N	N	N	N	N
Por	kje	kjew	sou	tão	ca	la	do

⁷ GÎTÁ. Intérprete: Raul Seixas. Compositor: R. Seixas e P. Coelho. In: GÎTÁ. Intérprete: Raul Seixas. [S. l.]: Philips/Universal Music Brasil, 1974. 1 LP, faixa 12.

Processo semelhante ocorre quando do acréscimo de notas, como mostrado nos versos iniciais de *Sampa*, interpretada por Caetano Veloso⁸ (13a–c).

(13a–c)

	N	N	N	N	N	N	N	N	N	N	N	N	N
(a)	AL	gu	ma	coi	sa	con	te	ce	no	meu	co	ra	ção
(b)	que	SÓ	quan	do	cru	zaj	pi	ran	gja	ve	ni	da	são joão
(c)	é	que	QUAN	dweu	che	guei	por	a	qui	eu	na	den	ten di

Aqui vemos um incremento na anacruse que antecede o primeiro tempo forte (em negrito) de cada verso. Diferentemente do poema, em que o verso se organiza em torno da quantidade de sílabas, contadas a partir da primeira delas, a canção tem como ponto de referência o tempo forte do primeiro compasso da frase musical. Anacruse é todo material que precede o tempo forte, podendo, em princípio, ter qualquer quantidade de notas/sílabas. Assim, (13c) é reestruturado como (13d) pela aplicação da regra insira N.

(14) insira N ($\emptyset \rightarrow N / \sigma_n > N_n$)

Se a quantidade de sílabas de uma cadeia é maior que a quantidade de notas da melodia, inserem-se notas para satisfazer (8).

(13c)

	N	N	N	N	N	N	N	N	N	N	N	N	N
é	que	QUAN	dweu	che	guei	por	a	qui	eu	na	den	ten	di

insira N ($\emptyset \rightarrow N / \sigma_n > N_n$)

(13d)

N	N	N	N	N	N	N	N	N	N	N	N	N	N
é	que	QUAN	dweu	che	guei	por	a	qui	eu	na	den	ten	di

Podemos concluir preliminarmente, a partir dos dados até aqui mostrados, que existem contextos nos quais alguns componentes da gramática tornam-se invisíveis para a melodia. Em outros termos, dado que a palavra cantada se constitui pela superposição de dois componentes, verbal e musical, e dado que cada um desses apresenta gramática própria, pode ocorrer conflito entre estruturas verbais e melódicas, de onde resulta

⁸ SAMPA. Intérprete: Caetano Veloso. Compositor: C. Veloso. *In*: MUITO - DENTRO DA ESTRELA AZULADA. Intérprete: Caetano Veloso. [S. l.]: Polygram/Phillips Brasil, 1978. 1 LP, faixa 7.

algun tipo de ajustamento de um ou de outro lado. Se as estruturas verbais impõem-se sobre a melodia, esta se adapta inserindo ou suprimindo notas na cadeia, de modo a não violar o Pareamento Métrico. Se, ao contrário, a estrutura melódica impõem-se sobre o texto — e esse é o caso mais interessante do ponto de vista linguístico — são as sílabas que agora podem ser suprimidas ou criadas. Nesses casos, a palavra cantada parece poder violar processos fonológicos usualmente observados na variedade falada. Em outros termos, tudo indica que certas construções são aceitas quando cantadas, mas não o são quando faladas. Assim, os dados observados até aqui nos levam a indagar por que certas construções soam antinaturais quando faladas (isto é, são fonologicamente anômalas), mas passam despercebidas quando cantadas.

Discussão

A hipótese apresentada neste trabalho defende a existência de certas condições de boa formação que regulam a relação entre texto e melodia na palavra cantada. A mais básica dessas condições estipula que a relação entre nota e sílaba é bijetiva, o que temos chamado de Pareamento Métrico. Ocorre que esse princípio vai de encontro a outro, tacitamente aceito na literatura musicológica, segundo o qual é sempre possível associar a uma única sílaba mais de uma nota musical. Melodias assim construídas são chamadas “melismáticas” (Hartong, 2007, p.160). Em vista disso, discutiremos agora alguns argumentos a favor da tese da relação bijetiva nota/sílaba.

Do ponto de vista fonológico — e independentemente do modelo de sílaba que adotemos —, devemos admitir que não é possível associar mais de uma nota a uma sílaba completa (ataque + núcleo + coda). Se duas notas são associadas a uma tal sílaba, a primeira necessariamente será constituída pelo ataque + núcleo e a segunda pelo núcleo + coda. Além disso, se mais de duas notas são associadas a uma única sílaba, todas as intermediárias entre a inicial e final serão associadas apenas ao núcleo. Por exemplo, em *Não quero dinheiro*⁹, Tim Maia canta uma cadeia de cinco notas, que são grafadas na partitura sobre uma única sílaba: (a) MOR (1).

Figura 3 – Excerto de *Não quero dinheiro* (Tim Maia)

(1) eu te a mo eu tja do ro meu a MOR

(2) eu te a mo eu tja do ro meu a MO O O O OR

Fonte: Elaboração própria.

⁹ NÃO QUERO DINHEIRO. Intérprete: Tim Maia. Compositor: T.Maia. In: TIM MAIA - 1971. Intérprete: Tim Maia. [S. l.]: Polydor/Polygram Brasil, 1971. 1 LP, faixa 2.

Parece-nos claro que essa cadeia de notas segmenta a sílaba MOR em três partes, “MO”, “O” e “OR”. A primeira nota está associada à sílaba MO, ou seja, ataque + núcleo; as três notas seguintes estão associadas à sílaba O, ou seja, ao núcleo; por fim, a última nota está associada à sílaba OR, ou seja, núcleo + coda. Portanto, a transcrição mais adequada dessa melodia seria (2), não (1). Essa representação é mais consentânea à realidade acústica da cadeia, uma vez que a coda não é audível na primeira sílaba, nem o ataque na última.

Outro argumento, agora fonético-articulatório, é que uma cadeia de notas associada a uma única vogal não pode ser indefinidamente longa. Se assim for, uma pausa será necessariamente inserida, determinando o limite da sílaba. Por exemplo, numa passagem do *Messias*¹⁰ de Händel, a sílaba *born* é associada a uma cadeia de 57 notas.

Figura 4 – *Messiah* (G.F. Händel)



Fonte: Musescore¹¹.

Ocorre que não é possível cantar essa sequência sem introduzir não uma, mas várias pausas. Como afirmar então que temos aí uma única sílaba?

Por fim, há um argumento métrico em apoio à hipótese da associação biunívoca nota/sílaba. A melodia de uma canção é constituída quase sempre de um conjunto limitado de frases musicais, muitas vezes por uma única frase que se repete com variações. Assim como o verso de um poema, a frase musical estabelece um padrão métrico, que é preenchido por diferentes textos, ou seja, por diferentes cadeias de sílabas. Os versos de *Asa branca* são um bom exemplo disso:

Quan	doj	ej	a	te	rar	den	do
Qual	fo	guei	ra	de	são	joão	
Que	bra	sei	ro	que	for	na	lha
Ne	nhum	pé	de	pran	ta	ção	

¹⁰ FOR unto us a child is born. Intérprete: Academy of St Martin in the Field. Compositor: G. Handel. In: MESSIAH – HWV 56 1.12. Intérprete: Academy of St Martin in the Field. [S. L.]: Decca Music Group Ltd., 2002. 1 CD, faixa 8.

¹¹ Disponível em: <https://musescore.com/user/181766/scores/2144106>. Acesso em: 25 out. 2023.

Numa das interpretações dessa canção¹², Luiz Gonzaga canta essas frases melódicas com uma única sílaba, “hum”.

hum hum hum hum hum hum hum hum
| | | | | | | |
hum hum hum hum hum hum hum

Figura 5 – excerto de *Asa Branca* (Luiz Gonzaga e Humberto Teixeira)



Fonte: Elaboração própria.

Não nos parece possível comutar uma cadeia de oito sílabas (2) por outra de uma única sílaba (1). A hipótese mais simples, portanto, é considerar que essa cadeia contém não uma sílaba, mas oito repetições da mesma sílaba “hum”. Exemplos como este são abundantes na música popular.

Por todas essas razões, parece-nos que a associação biunívoca entre nota e sílaba é a hipótese mais simples para explicar alguns fatos estranhos que ocorrem na palavra cantada. Ademais, a hipótese da relação bijetiva nota/sílaba não contradiz o fato observado em inúmeras interpretações nas quais o cantor varia livremente a f_0 da nota, de modo que uma única vogal associa-se a dois ou mais valores de altura. Cantores como Ed Motta e Aretha Franklin são representantes desse estilo melismático. Porém, é necessário distinguir tais casos em que o intérprete explora o *continuum* da altura daqueles que tratam a altura como um conjunto de tons discretos.

De certo modo, a melodia tem uma dimensão “fonológica” constituída por um inventário discreto de notas, e uma dimensão “fonética”, constituída pelo contínuo da altura (e da duração e intensidade, claro). Aqui, procuramos descrever a dimensão fonológica da música.

Rigidez e ordenação das regras

É preciso mencionar agora dois problemas decorrentes do que acabamos de expor. O primeiro diz respeito à ausência de rigidez entre entrada e saída das regras de reestruturação. O segundo diz respeito à ordem hierárquica que se estabelece entre, de um lado, os processos fonológicos determinados apenas pelo ambiente linguístico e pelas condições da enunciação e, de outro, os processos melódico-fonológicos determinados pela melodia. Esses problemas extrapolam os limites do presente ensaio e demandam

¹² Disponível em: <https://www.youtube.com/watch?v=zsFSHg2hxbc>. Acesso em: 25 out. 2023.

um texto integralmente a eles dedicado, mas não podemos deixar de apresentar seus contornos gerais.

Já vimos que em (3a) a quantidade de notas excede a quantidade de sílabas, o que leva à aplicação de insira σ , do que resulta (3b).

(3a)

N	N	N	N	N	N	N	N	N	N
			\	/					
eu	per	gun	tei	a	Deus	do	céu	ai	

insira σ ($\emptyset \rightarrow \sigma / N_n > \sigma_n$)

(3b)

N	N	N	N	N	N	N	N	N	N
eu	per	gun	te	ej	a	Deus	do	céu	ai

Ocorre que o Pareamento Métrico é muito amplo, assim como o ambiente das regras dele derivadas. O princípio determina apenas que a quantidade de sílabas deve ser igual à quantidade de notas, sem especificar como esta igualdade deve ser alcançada. Em outras palavras, dada uma regra, é possível que muitas possam ser as saídas a partir de uma mesma entrada. Assim, (3b) não decorre necessariamente da aplicação de insira σ a (3a), ao contrário, diversas reestruturações são possíveis, como, por exemplo (3g), (3h) e (3i), que satisfazem o Pareamento Métrico.

N	N	N	N	N	N	N	N	N	N	
(3b)	eu	per	gun	te	ej	a	Deus	do	céu	ai (epêntese silábica)
(3g)	e	u	per	gun	tej	a	Deus	do	céu	ai (diérese eu > e u)
(3h)	eu	per	gun	tej	a	De	us	do	céu	ai (diérese Dews > De us)
(3i)	eu	per	gun	tej	a	Deus	do	cé	u	ai (diérese em céw > cé u)

Por que apenas (3b) se realiza? Não temos condição de responder categoricamente a esta questão no momento. Observe-se apenas que em (3g-i), a diérese incide sobre palavras prosódicas monossilábicas (eu, Deus, céu) e compromete em certo grau a compreensibilidade do texto, o que não ocorre com a epêntese em (3b). Outro fator que talvez bloqueie essas construções diz respeito à métrica da melodia. As diéreses (3g-i) não apenas criam novas sílabas, mas deslocam os respectivos acentos para a vogal que substitui a semivogal originária (ew > e U), (cew > ce U), (dews > de US), o que também compromete a compreensibilidade do texto. Assim, embora o Pareamento Métrico faça referência apenas à sílaba e à nota, a descrição das reestruturações requer referência a outros fatores, como métrica (da melodia), acento primário e secundário, e

ao domínio prosódico sobre o qual incide a reestruturação (palavra, pé, frase fonológica), pelo menos.

O segundo problema decorrente da amplitude do Pareamento Métrico diz respeito à ordenação entre regras fonológicas e regras “melódico-fonológicas” na palavra cantada. Vimos que tanto insira σ quanto apague σ são regras específicas da palavra cantada. Tais regras coocorrem com as regras de inserção e apagamento da fonologia do português. Assim, na transformação (2a) > (2b) temos ditongação (de um > djum) e degeminação (a ação > ação) em consonância com a fonologia do português brasileiro. O que causa problema é apenas a ditongação (#filme a > filmja), pelas razões já adiantadas. Dado que nosso *corpus* é constituído pelos fonogramas e pelo texto escrito fornecido no encarte de cada disco, nossa análise a eles tem se restringido. Porém devemos lembrar que o texto escrito é uma transcrição da fala conforme as regras de ortografia e que, por isso mesmo, não registra elisões, degeminações, ditongações e outros processos fonológicos eventualmente presentes. Desse modo, com o intuito de registrar apenas as reestruturações resultantes do Pareamento Métrico, convém adotar como critério metodológico uma etapa intermediária, na qual todos os possíveis processos de sândi externo sejam realizados. Assim, por exemplo, reformularíamos a transformação (2a > 2b) em (2a > 2a' > 2b).

```
(2a)  co mo de um fil me a a ção que não va leu
      | | \ / | | \ / | | | | | | (processos fonológicos)
(2a') co mo dewm fil me a ção que não va leu
      (de um > dewn) (a ação > ação)

(2a') co mo dewm fil me a ção que não va leu
      | | | | \ / | | | | | | (processos melódico-fonológicos)
(2b)  co mo dewm fil mja ção que não va leu
      | | | | | | | | | | |
      N N N N N N N N N N
      # (filme a > filmja)
```

Essa diretriz é necessária para, de um lado, depurarmos a análise, separando processos fonológicos de processos “melódico-fonológicos”. De fato, na análise da palavra cantada, interessa determinar e descrever o mais precisamente possível o ambiente extralinguístico (melódico, rítmico, harmônico, etc.) que determina a realização ou o bloqueio de um processo fonológico.

Paralelamente, essa análise mostra quão imprecisa é a terminologia musical que distingue melodia silábica, “*a song in which each syllable has but one note*”¹³ de melodia melismática, “*melody in which more than one tone is sung to a syllable*”¹⁴ (Hartong,

¹³ Original: “melodia cantada em que cada sílaba do texto corresponde a uma única nota”.

¹⁴ Original: “melodia em que mais de uma nota é associada uma única sílaba”.

2007, p. 160). Como mostramos, todas as melodias são silábicas, sem exceção, e a distinção entre melodias silábicas e melismáticas diz respeito à sílaba escrita e não à sílaba efetivamente cantada.

Conclusão

Sumarizando o que vimos até aqui:

- a palavra cantada pode apresentar construções prosodicamente anômalas;
- tais construções, que ocorrem como reestruturações silábicas excepcionais, não são aceitas na fala, mas passam despercebidas quando cantadas;
- essas construções podem ser adequadamente descritas se considerarmos a palavra cantada como a interação de duas cadeias, silábica e melódica, cujos terminais são respectivamente a sílaba e a nota;
- a interação entre essas cadeias é regida pelo Pareamento Métrico, que estabelece a relação bijetiva entre nota e sílaba;
- a inviolabilidade dessa condição de boa formação revela-se na maneira sistemática pela qual processos fonológicos de reestruturação silábica são bloqueados (quando deveriam se realizar) ou se realizam (quando deveriam ser bloqueados);
- as reestruturações silábicas observadas na palavra cantada, embora manifestem-se superficialmente como processos fonológicos, devem ser consideradas processos melódico-fonológicos, característicos da interação entre a cadeia silábica e a melodia musical;
- o ambiente para a ocorrência dessas reestruturações não é linguístico (fonológico, sintático, morfológico), mas musical.

O Pareamento Métrico expressa a mais fundamental das condições de boa formação da palavra cantada. A condição *sine qua non* para que a palavra cantada seja bem formada é que a quantidade de terminais da cadeia silábica seja igual à quantidade de terminais da cadeia melódica. Como essa condição não admite exceções e se sobrepõe a qualquer outra condição de natureza linguística, não é raro que ocorram ressilabificações estranhas ao português, as quais, quando faladas, soam como construções prosodicamente agramaticais. Vimos que essas construções podem ser descritas pelas regras apague σ e insira σ .

O Pareamento Métrico é a mais básica condição de boa formação da palavra cantada porque seu domínio restringe-se à relação entre notas e sílabas não especificadas. Evidentemente, a interação entre música e fala vai além disso. Sílabas são portadoras de acento primário ou secundário, podem ser monossílabos átonos ou tônicos, ou ainda, núcleos de frases fonológicas. Por sua vez, sobre a nota incide um tempo forte

ou fraco, ela pode ou não ser o núcleo de um grupo rítmico, pode ou não desempenhar uma função harmônica, e assim por diante. Em outras palavras, para além da relação bijetiva estabelecida entre nota e sílaba, esses terminais interagem com outros elementos que lhes são hierarquicamente superiores. Mais do que isso, essas hierarquias parecem interagir entre si. Uma descrição dessa interação ultrapassa em muito os limites deste ensaio e é aqui mencionada apenas para destacarmos que o Pareamento Métrico e as regras dele derivadas conseguem explicar apenas uma pequena parte dos fenômenos linguístico-melódicos observados na palavra cantada.

CARMO JR., J. On the relationships between syllables and musical notes in sung words in Brazilian Portuguese. *Alfa*, São Paulo, v.68, 2024.

- *ABSTRACT: Sung words often present prosodically ungrammatical structures. Such structures cannot be described exclusively on the basis of the linguistic environment in which they occur. Ungrammaticality manifests itself in the actualisation of phonological processes in environments in which such processes would be blocked and, conversely, in its blockage in environments in which their actualisation would be expected. The aim of this paper is to describe this behavior of sung words that contrasts with that of spoken words. We shall initially introduce examples that illustrate the extent of the problem as well as its ubiquity. Next, we shall argue that the observed ungrammaticality can be explained on the hypothesis that the correspondence between text and melody in sung words is regulated by well-formed conditions. One of these conditions — which we call Metrical Pairing — establishes that the relationship between note and syllable is always bijective, that is (I) each and every note in a melody must be assigned to one and only one syllable, and (II) each and every syllable in a string must be paired with one and only one note. The observance of this condition for well-formedness of sung words explains both the actualisation and the blockage of phonological processes that diverge from the phonology of spoken Portuguese.*
- *KEYWORDS: resyllabification; phonological process; melody; sung music; prosody.*

REFERÊNCIAS

ABAURRE, M. B. M. Acento frasal e processos fonológicos segmentais. **Letras de Hoje**, Porto Alegre, v. 31, n. 2, 1996. Disponível em: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/15591>. Acesso em: 2 out. 2023.

ABOUSALH, E. **Resolução de choques de acento no português brasileiro**: elementos para uma reflexão sobre a interface sintaxe-fonologia. 1997. 157f. Dissertação (Mestrado) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1997.

BISOL, L. Sândi externo: o processo e a variação. *In*: KATO, M. A. (org.). **Gramática do português falado**, Campinas: Ed. da UNICAMP, 2002. p. 53–97.

BOLINGER, D. **Intonation and its Uses**: Melody in Grammar and Discourse. Stanford: Stanford University Press, 1989.

GOLDSMITH, J. **Autosegmental Phonology**. 1976. Tese (Doutorado em Filosofia) – Departamento do Linguística e Filosofia, Instituto de Tecnologia de Massachussets, Cambridge, 1976.

HARTONG, J. L. **Musical Terms Worldwide**: A Companion for the Musical Explorer. The Hague: Semar Publishers, 2007.

PIERREHUMBERT, J. **The phonology and phonetics of English intonation**. 1980. Tese (Doutorado em Filosofia) – Departamento do Linguística e Filosofia, Instituto de Tecnologia de Massachussets, Cambridge, 1980.

WAUGH, L. R. **The melody of language**: intonation and prosody. Baltimore: University Park Press, 1980.

Recebido em 8 de março de 2021

Aprovado em 13 de outubro de 2022