

## ON THE RELATIONSHIPS BETWEEN SYLLABLES AND MUSICAL NOTES IN SUNG WORDS IN BRAZILIAN PORTUGUESE

José Roberto do CARMO JR.\*

- **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 behaviour 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 words; prosody.

### The problem

It is not uncommon to find prosodically anomalous constructions in sung words. Though anomalous constructions sound preposterous in any speech register, they are supposedly hardly noticed when sung. It is to be suspected that this phenomenon does not occur fortuitously; rather, it depends on specific extralinguistic contexts. In such contexts, phonological processes referring to syllable and stress may be blocked; conversely, phonological processes usually blocked in speech may occur in singing.

---

<sup>1</sup> Universidade Federal do Paraná (UFPR). Curitiba – PR – Brasil. carmojr@ufpr.br. Orcid: 0000-0002-2293-9962.

Compare, for example, the spoken (1a) and sung (1b) renderings of a line in *Rouxinol* (Gilberto Gil).<sup>1</sup>

(1a) TraTEI da sua Asa

(1b) #TRAtEi da suø Asa

In (1b)<sup>2</sup> we see what is apparently a stress shift (traTEI > TRAtEi) and a degemination (sua Asa > suAsa). We say “apparent” shift and degemination because, in fact, in neither case is there an environment for those phonological processes. The environment for the shift is the stress clash (Abousalh, 1997), a phenomenon that occurs between two adjacent syllables bearing primary stress and belonging to words of the same phonological phrase – for instance, stress shifts occur in [caFÉ QUENte > CAFÉ QUENte] and [JeSUS CRISto > JESus CRISto]. That is not the case in [traTEI > TRAtEi]. Likewise, it is convenient to speak of apparent degemination in [sua Asa > suAsa] because, as Bisol (2002, p. 66) has demonstrated, degemination is subject to rhythmic restrictions. The term degemination applies only when the second of the vowels in contact is not stressed, otherwise it is blocked. Given that in (1b) the second vowel has the primary stress, degemination should be blocked; however, this does not occur in Gil’s delivery. In other words, although (1b) has indeed a stress displacement in [TRAtEi] and the suppression of a vowel in [suAsa], it cannot be said that these are instances of the natural speech phenomena we know as “stress shift” and “degemination”.

Something similar can be seen when we contrast the spoken (2a) and sung (2b) renderings of the lines of *Valsa brasileira* (Chico Buarque & Edu Lobo), performed by Chico Buarque.<sup>3</sup>

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

(2b) #Como de um filmj ação que não valeu

Here we see degemination associated with diphthongization [filme, a ação > filme ação > filmj ação], in such a way that makes it impossible to determine in which order the two processes occur. Whatever the order, however, it looks like a construction outside the realm of Portuguese phonology. Given the occurring anastrophe (“*como de um filme, a ação que não valeu*” versus “*como a ação de um filme que não valeu*”), and given that phonology has access to syntax (Abaurre, 1996), a pause between the lines is necessary (um filme//a ação). As a consequence, the diphthongization (fil.mja) should be blocked, which does not occur in Buarque’s delivery.

---

<sup>1</sup> O ROUXINOL. Performed by: Gilberto Gil. Composed by: G. Gil and J. Mautner. In: REFAZENDA. Performed by: Gilberto Gil. [S.L.]: Philips Records Brasil, 1975. 1 CD, track 9.

<sup>2</sup> The hash symbol # here indicates a prosodically anomalous construction.

<sup>3</sup> VALSA BRASILEIRA. Performed by: Chico Buarque. Composed by: C. Buarque and E. Lobo. In: CHICO Buarque – 1989. Performed by: Chico Buarque. [S.L.]: Philips Records Brasil, 1989. 1 LP, track 10.

Degemination (1b) and diphthongization (2b) are syllable adjustment processes in which a reduction occurs in the number of syllables in the string. Adjustment by increasing the number of syllables is also common in songs, although it is not confused with any known phonological process.

(3a) and (3b) are, respectively, spoken and sung renderings of a line in *Asa branca* (Luiz Gonzaga and Humberto Teixeira), delivered by Luiz Gonzaga.<sup>4</sup>

(3a) Eu per gun tei a Deus do céu ai

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

The difference between (3a) and (3b) is solely the number of syllables, nine and ten respectively. If this increase were the result of diaeresis (per.gun.tej > per.gun.te.i), it would not be unusual, as it is a resource commonly used by songwriters and poets. However, what is observed in (3b) is not a diaeresis (tej > te.i), but an odd “syllabic epenthesis” (tej > te.ej) which preserves the diphthong, creating a syllabic adjustment that has no parallel in spoken language.

A genuine case of diaeresis, on the other hand, is not without interest. For instance, in (4a) and (4b) are, respectively, spoken and sung renderings of a line from *Tempo de estio* (Caetano Veloso).<sup>5</sup>

(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

Here, the diaeresis em (4b) is quite obvious (Lej.las > Le.i.las). However, it is interesting to compare (4a/b) with an homologous line which resides in the same position in the next stanza:

(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

Just like (3b), (5b) has the insertion of an epenthetic syllable (Te.re.zas > Te.re.e.zas). Given the homology between [Te.re.zas > Te.re.e.zas] and [Lej.las > Le.i.las], we are led to question what factor produces the apparent diaeresis occurring in (4b). We shall return to this point later.

To summarise the problem described so far, both song and speech are subject to syllabic adjustment processes, either by reduction or by increase of phonetic material. In speech, such adjustment processes are governed by the linguistic environments in which they occur and by the enunciation style (e.g., more or less formality, higher or lower

<sup>4</sup> ASA BRANCA. Performed by: Luiz Gonzaga. Composed by: L. Gonzaga and H. Teixeira. In: NOVA HISTÓRIA DA MÚSICA POPULAR BRASILEIRA, Vol. 11. Performed by: Luiz Gonzaga [S. I.]: RCA Brasil, 1977. 1 LP, track 3.

<sup>5</sup> TEMPO DE ESTIO. Performed by: Caetano Veloso. Composed by: C. Veloso. In: MUITO – DENTRO DA ESTRELA AZULADA. Performed by: Caetano Veloso [S. I.]: Polygram/Phillips Brasil, 1978. 1 LP, track 2.

speech rate). In contrast, much of the syllabic adjustment observed in songs cannot be explained by reference to the linguistic environment. Indeed, the constructions we have just observed seem to be at variance from spoken Portuguese phonology. Since the aim of this paper is to describe such constructions and propose a hypothesis for why they occur, we shall initially present the hypothesis of the conditions for well-formed sung words and then use that hypothesis as a basis to discuss some of the processes that refer directly to the syllable and to the stress, and the way these are actualised in sung words.

### **Conditions for well-formed sung words**

We have so far been finding contrasts in descriptions of spoken and sung constructions, as if these were comparable with each other. However, if we examine the nature of sung words, we find that this procedure does not hold. One cannot take sung words as if they were made up of a simple interaction between segmental strings and prosodic strings, like spoken words are. Let us look at this issue in more detail.

From a strictly linguistic viewpoint, speech is created by coupling two strings: a segmental one (segments and their respective organisation into syllables), and a suprasegmental or prosodic one (stress, intonation and rhythm). Phonological analysis refers to one or the other string, or else to the interaction between them. Sung words, on the other hand, are combined with a melody which is never confused with the intonation of speech, a rhythm which is never confused with the rhythm of speech, and stress patterns which are never confused with the primary or secondary stress patterns of natural speech. Therefore, any attempt to analyse sung words has to factor in their “melody”, an extralinguistic element apparently capable of activating or blocking certain phonological processes.

In this scenario, two methodological approaches to describe sung words present themselves, and both are theoretically plausible:

- (I) The first approach is in line with our intuition of what singing is, in which the sung word is regarded as speech with the addition of a musical melody. In this approach, sung words are a variety of spoken words which comprises not two but three overlapping strings: segmental, prosodic and melodic.
- (II) The second approach considers that the melodic string in sung words does not coexist with the prosodic string of those words; rather the prosodic string is replaced by the melodic string. In this approach, sung words are a variety of spoken words which comprises only two overlapping strings: segmental and melodic.

We shall adopt the second approach in this paper. It is a strictly methodological choice which is theoretically robust and economical. This approach emerges from the observation that melody and prosody share the same phonetic material, that is, both

resolve into attributes of pitch, duration and intensity. In this perspective, musical melody on the one hand and prosody (rhythm, intonation and stress) on the other are just different ways of organising and structuring the same phonetic material, which comprises those attributes of pitch, duration and intensity. This is made particularly clear when we consider the similarities and differences between intonation and melody.

Although intonation is commonly referred to as “the melody of speech” (Waugh, 1980; Bolinger, 1989), there is a fundamental difference between intonation and melody and, of course, between speech and singing. The unit of intonation is the tone, that is, the fundamental frequency (f0) at the core of every syllable. Speech intonation is organised around just two tones levels, high (H) and low (L), which get combined in several ways and thus give rise to a limited number of pitch stresses and boundary tones (Pierrehumbert, 1980).<sup>6</sup> Due to this limitation — which, by the way, is essential for the functioning of intonation as a paralinguistic system —, the relationship between tone and vowel (the nucleus of the syllable) is not bijective, that is, the relationship between tone and vowel can be one to one, one to many, or many to one. In Goldsmith’s formulation (1976, p. 27): (I) all vowels are associated with at least one tone; and (II) all tones are associated with at least one vowel. The association between a single tone and a single vowel is only a particular case arising from (I) and (II), as shown in (6).

(6)

Tone	T T	T	T
	\ /		/ \
Syllable	σ	σ	σ σ

In contrast to melody, intonation has an “elastic” nature, which allows it to mould itself to any linguistic string. Accordingly, intonations constitute relatively stable language inventories, which is why we do not create new intonations with each act of speech, but, thanks to their elasticity, we just apply them to new strings of syllables. In short, intonation is not productive.

Another important feature of intonation is that it is not recursive. In Pierrehumbert’s hypothesis (1980, p. 29), intonation constitutes a grammar of finite states in that tonal morphemes merely string themselves in a sequence without constructing a hierarchy, and mould themselves to the semantic content, the syntactic structure and the pragmatic condition of the enunciation.

Melody is of a different nature. Firstly, the melody unit is not the tone, but the note (N) and, unlike the case of intonation, the relationship between note and syllable is bijective, that is, it is a one-to-one relationship, as shown in (7).

---

<sup>6</sup> At this point, as we try to emphasise the differences between intonation and melody more than their similarities, it seems to us inconsequential to adopt Pierrehumbert’s (1980) proposals as our model of intonation.

(7)	#		#
Note	N N	N	N
	\ /		/ \
Syllable	$\sigma$	$\sigma$	$\sigma \sigma$

To assert that the relationship between note (N) and syllable ( $\sigma$ ) in sung words is necessarily bijective is the same as saying that it is subject to a restriction. In sung words, it is not possible to associate a single note with more than one syllable, just as it is not possible to associate a single syllable with more than one note. Given the universality of this restriction –that is, given the fact that every sung melody has a one-to-one correspondence between its notes and syllables, without exception–, it can be said that this correspondence constitutes one of the conditions for well-formed sung words, a condition we have called *Metrical Pairing*.

- (8) *Metrical Pairing* - In sung words, each terminal element of the melodic string (N, note) must be paired to one and only one terminal element of the syllabic string ( $\sigma$ , syllable) and vice versa.

### Adjustment rules

We have now reached a position from which we can formulate a hypothesis to describe the occurrence of anomalous prosodic constructions in sung words, some examples of which we have seen above. Since one of the conditions for the well-formedness of sung words is *Metrical Pairing*, it is inviolable. When notes and syllables threaten to dissociate, a necessary adjustment between the melodic string and the syllable string occurs involving either a deletion or insertion of syllables and/or notes. There are four possible scenarios for this adjustment, which correspond to four rules:

- (I) syllable deletion rule — delete  $\sigma$ ;
- (II) syllable insertion rule — insert  $\sigma$ ;
- (III) note deletion rule — delete N;
- (IV) note insertion rule — insert N.

In this paper, we shall focus mainly on the first two scenarios as they are the only ones that have any bearing on linguistic analysis. Once rule (I) “delete  $\sigma$ ” or (II) “insert  $\sigma$ ” is applied, the syllable string is adjusted, becoming subject to the usual phonological processes. Let us look more closely at each of those cases.

- (9) delete  $\sigma$  ( $\sigma \rightarrow \emptyset / N_n < \sigma_n$ )

If the number of syllables in a string is greater than the number of notes in the melody, the supernumerary syllables are deleted so that (8) is satisfied.

The application of “delete  $\sigma$ ” explains the apparent degemination in (1b). Indeed, the line has seven syllables, but the melody has only six notes (Fig. 1).

**Figure 1** – fragment of *Rouxinol* (Gilberto Gil)



**Source:** Author’s elaboration

The dissociation between terminal elements is resolved by the application of “delete  $\sigma$ ” so that (8) is satisfied. Thus:

(1a)

Note	N	N	N	N	N	N
					/ \	
Syllable	Tra	tei	da	su	a a	as

delete  $\sigma$  ( $\sigma \rightarrow \emptyset / N_n < \sigma_n$ )

(1b)

Note	N	N	N	N	N	N
Syllable	Tra	tei	da	su	a	as

Seen this way, the deletion of the syllable does not result from a *stricto sensu* phonological process (in this case, degemination), but from a melodic-phonological process that is unique to sung words. It may be objected that degemination is not strictly necessary and that (8) would be equally satisfied by simply diphthongizing the hiatus (su.a > swa), as can be seen in (1c).

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

Nevertheless, our analysis looks at some specific recordings, which in this case is Gilberto Gil’s rendering of the song. In this rendering, the performer sings (1b) and not (1a) or (1c).

Let us now see how the “delete  $\sigma$ ” rule can explain what happens in (2a) and (2b), where a diphthongization is observed that would not occur in spoken language (film, a ação > filme ação > filmjação). Again, it seems that the decisive factor in whether or not a phonological process is performed comes from outside the realm of linguistics. We see that the thirteen syllables in (2a) are sung on a melody of ten notes (Fig. 2).

**Figure 2** – fragment of *Valsa Brasileira* (Chico Buarque & Edu Lobo)



**Source:** Author’s elaboration

In order to satisfy (8), “delete  $\sigma$ ” is applied:

(2a)

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

delete  $\sigma$  ( $\sigma \rightarrow \emptyset / N_n < \sigma_n$ )

(2b)

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

Note that diphthongization (de um > djum) presents no problem, as it already occurs in natural speech and is therefore not in the scope of this paper. The diphthongization + degemination in (filme, a ação > filmjação) does not sound like a prosodically acceptable construction in spoken Portuguese. It can only be explained as the result of the association between the segmental string and the melodic string, an association that has (8) as a principle of well-formedness.

Metrical Pairing states that the number of notes in a song must be equal to the number of syllables sung. We have seen how this principle is satisfied by deleting syllables when their number is greater than the number of notes in the string ( $N_n < \sigma_n$ ) by applying “delete  $\sigma$ ” ( $\sigma \rightarrow \emptyset / N_n < \sigma_n$ ); and when, conversely, the number of notes is greater than the number of syllables ( $N_n > \sigma_n$ ), “insert  $\sigma$ ” ( $\emptyset \rightarrow \sigma / N_n > \sigma_n$ ) is applied.

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



When the number of notes exceeds the number of syllables in a string, syllables are inserted so that (8) is satisfied. The application of this rule can only be confirmed indirectly. Thus, simply comparing (3a) and (3b) does not make it clear that the syllable ‘ej’ was inserted in order not to violate (8). At first hearing, the insertion sounds merely like the result of the songwriter’s poetic freedom. However, we have seen that Metrical Pairing determines a bijective note/syllable relationship, implying the impossibility of singing a note without a syllable associated with it, whatever it may be. This is what would happen anyway in (6a), because the number of notes in the melody (10) exceeds the number of syllables in the line (9).

(3a)

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

In (3a), there is a clear violation of (8). Our hypothesis is that (3a) is adjusted into (3b) so that (8) is satisfied.

(3b)

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

The hypothesis is corroborated by comparing the song’s homologous lines that align themselves perfectly to the melody in (3a). Thus:

	N	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

Several adjustments are present here; but they are all predictable from speech patterns, except the insertion of the syllable ‘ej’, which emerges from applying Metrical Pairing.

The same argument is relevant to (4a), where a diaeresis apparently occurs in (Lej.las > Le.i.las). Here, too, the number of notes exceeds that of syllables. The melody has 14 notes, while the line has 13 syllables.

(4a)

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	Lej	las		

By applying “insert  $\sigma$ ”, (4a) is adjusted into (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	

At first hearing, we could describe the adjustment (Lej.las > Le.i.las) as a simple diaeresis emerging from the constrictions of poetic metre, bearing no necessary relationship with the song’s melody. However, given the melody’s structural rigour, the insertion via diaeresis in (4b) is repeated in (4c), now as epenthesis.

	N	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	

It seems to us that the most defensible hypothesis to explain (4b) and (4c) is that, whether by diaeresis or epenthesis, the rule “insert  $\sigma$ ” is applied in both cases in order to prevent the violation of the Metrical Pairing principle. It remains for us to present the rules for note insertion or deletion, whose interest is more musicological than linguistic. The insertion or deletion of syllables manifests itself as a phonological process whose ultimate motivation, as we have seen, is musical. The insertion or deletion of notes, on the contrary, presents itself as a melodic “process” which originates in the text. In the previous case, the melody imposes a structure on the string of syllables; in the present case, the semantic-syntactic-phonological integrity of the text is to be preserved, which forces the melody to be adapted to the text.

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

If the number of syllables in a string is smaller than that of notes in the melody string, notes are deleted so that (8) is satisfied. The stanza (12 a–d) with the first four initial lines of Raul Seixas’ *Gîtã*, in his own rendering<sup>7</sup>, illustrates rule (11) and also the difference between poems with rigorous metre and scansion, and song lyrics, whose metre is usually more flexible.

<sup>7</sup> GÎTÃ. Performed by: Raul Seixas. Composed by: R. Seixas and P. Coelho. In: GÎTÃ. Performed by: Raul Seixas. [S. l.]: Philips/Universal Music Brasil, 1974. 1 LP, track 12.

	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

Firstly, while (12 a, c & d) have nine syllables and nine notes, (12b) has eight syllables and eight notes. All lines are in accordance with the Metrical Pairing principle (8). This sort of disparity in the number of syllables in the lines making up a stanza is fairly common in popular songs. In this respect, song lyrics are different from traditional poetry and popular poetry, both of which are strongly linked to metre. When a lyricist takes the liberty of decreasing or increasing the number of syllables in a given line — whatever their reason for doing so — the number of notes in the melody must necessarily and correspondingly increase or decrease. This is what happens in (12b), by applying “delete N”. Thus:

(12b<sub>1</sub>)

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

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

(12b<sub>2</sub>)

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

A similar process occurs when notes are inserted, as in the initial lines of Caetano Veloso’s *Sampa*<sup>8</sup> (13a–c).

(13a–c)

		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	dwew	che	guei	por	a	qui	eu	na	den	ten	di

<sup>8</sup> SAMPÁ. Performed by: Caetano Veloso. Composed by: C. Veloso. In: MUITO — DENTRO DA ESTRELA AZULADA. Performed by: Caetano Veloso. [S. l.]: Polygram/Phillips Brasil, 1978. 1 LP, track 7.





inserted, which will determine syllable boundaries. For instance, in a passage from Handel’s *Messiah*,<sup>10</sup> the syllable ‘born’ is associated with a string of 57 notes.

**Figure 4** – Fragment of *Messiah* (G.F. Händel)



Source: Musescore<sup>11</sup>.

As it happens, it is not humanly possible to sing this sequence without introducing not only one, but several pauses for breathing. How could we assert then that it is a single syllable?

Finally, there is the argument based on metre in support of the bijective, one-note-to-one-syllable correspondence hypothesis. A song melody is nearly always made up of a limited set of musical phrases; often, a single phrase is repeated with variations. Just like the lines of a poem, the musical phrase establishes a metric pattern which is filled in by different texts, that is, by different strings of syllables. The lines in *Asa branca* are a good example of this:

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	çã	o

In one of the author Luiz Gonzaga’s renderings of the song<sup>12</sup>, he sings those melodic phrases with a single syllable, ‘mm’.

mm	mm	mm	mm	mm	mm	mm	mm
mm	mm	mm	mm	mm	mm	mm	

<sup>10</sup> FOR unto us a child is born. Performed by: Academy of St Martin in the Field. Composer: G. Handel. In: MESSIAH – HWV 56 1.12. Performed by: Academy of St Martin in the Field. [S. l.]: Decca Music Group Ltd., 2002. 1 CD, track 8.

<sup>11</sup> Available at: <https://musescore.com/user/181766/scores/2144106>. Access on: 25 out. 2023.

<sup>12</sup> Available at: <https://www.youtube.com/watch?v=zsFSHg2hxbc>. Access on: Oct. 25<sup>th</sup>, 2023.

**Figure 5** – fragment of *Asa Branca* (Luiz Gonzaga e Humberto Teixeira)

(1) hum hum hum hum hum hum hum hum  
 (2) quan do lhei a ter rar den do

**Source:** Author’s elaboration

It does not seem possible to commute a string of eight syllables (2) with another of a single syllable (1). Therefore, the simplest hypothesis is to consider that this string contains not one syllable, but eight repetitions of the same syllable “mm”. Examples like this abound in popular music.

Taken together, all of the arguments above point to a bijective correspondence between note and syllable as the most straightforward hypothesis to explain some oddities of sung words. Moreover, the bijective note/syllable correspondence hypothesis does not contradict the fact observed in countless renderings in which the singer freely varies the  $f_0$  of notes, so that a single vowel is associated with two or more pitch values. Singers like Ed Motta and Aretha Franklin produced instances of the melismatic style. However, we must make a distinction between performers who explore the pitch continuum from those who treat pitch as a set of discrete tones.

In a way, the melody has a “phonological” dimension consisting of a discrete inventory of notes, and a “phonetic” dimension consisting of the pitch continuum (and, of course, duration and intensity). Here, we have described aspects of the phonological dimension of the melody.

### Rigour, and ordering of rules

We must now turn our attention to two problems resulting from what we have just expounded. The first problem concerns the lack of rigour between input and output when the adjustment rules are applied. The second one concerns the hierarchy established between, on the one hand, the phonological processes determined only by linguistic environments and the conditions of enunciation and, on the other, the melodic-phonological processes regulated by the melody. Though these problems go well beyond the scope of this paper and demand one entirely dedicated to them, we shall present their general contours.

About the first problem, as we have seen, the number of notes in (3a) exceeds that of syllables, and the rule “insert  $\sigma$ ” is applied, resulting in (3b).

(3a)

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

insert  $\sigma$  ( $\emptyset \rightarrow \sigma / N_n > \sigma_n$ )

(3b)

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

As it happens, Metrical Pairing has a broad application, and so has the environment of the rules derived from it. The principle determines only that the number of syllables must be equal to that of notes, but does not specify how this equivalence is to be achieved. In other words, when a rule is given, many outputs are possible from the same input. Thus, (3b) does not necessarily follow from applying “insert  $\sigma$ ” to (3a). On the contrary, several adjustments are possible which satisfy Metrical Pairing, such as (3g), (3h) and (3i) below.

	N	N	N	N	N	N	N	N	N	N
(3b)	eu	per	gun	te	<b>ej</b>	a	Deus	do	céu	ai (syllable epenthesis)
(3g)	<b>e</b>	<b>u</b>	per	gun	tej	a	Deus	do	céu	ai (diaeresis: eu > e u)
(3h)	eu	per	gun	tej	a	<b>De</b>	<b>us</b>	do	céu	ai (diaeresis: Dews > De us)
(3i)	eu	per	gun	tej	a	Deus	do	<b>cé</b>	<b>u</b>	ai (diaeresis: céw > cé u)

Why is only (3b) actualised? We have no means to answer this question categorically at this point. Note only that in (3g–i), the diaeresis occurs on monosyllabic prosodic words (eu, Deus, céu), which undermines intelligibility to a certain degree. This does not occur with the epenthesis in (3b). Another factor that may halt the actualisation of those constructions concerns the melodic metre. The diaeresis in (3g–i) not only create new syllables but shift their stresses to the vowels that replace the original semivowels (ew > e U), (cew > ce U), (dews > de US), which also undermines intelligibility. Thus, although Metrical Pairing mentions only the syllable and the note, the description of the adjustments requires reference to other factors, such as metre (of the melody), primary and secondary stresses, and the prosodic domain on which the adjustment occurs (word, foot, phonological phrase), at least.

The second problem arising from the wide breadth of Metrical Pairing concerns the ordering of phonological rules and “melodic-phonological” rules in sung words. As we have seen, both “insert  $\sigma$ ” and “delete  $\sigma$ ” are rules specific to sung words. Such rules occur simultaneously with the insertion and deletion rules of Portuguese phonology. Thus, in the adjustment (2a) > (2b) we have diphthongization (de um > djum) and degemination (a ação > açã) in line with the phonology of Brazilian Portuguese. What causes problems is only the diphthongization (#filme a > filmja), for the reasons already mentioned. As our corpus comprises the recordings and the written text provided on the insert booklet of each disc, our analysis has been



delimited to them. Nevertheless, we must remind ourselves that the written text is transcribed speech that follows spelling rules and that, for this very reason, does not record elisions, degeminations, diphthongizations and other phonological processes that may be present. Thus, in order to register only the adjustments resulting from Metrical Pairing, it is convenient to adopt an intermediate stage as a methodological criterion in which all possible external sandhi changes occur. So, for example, we would reformulate the adjustment (2a > 2b) as (2a > 2a' > 2b).

```
(2a)  co mo de um fil me a a ção que não va leu
      | | \ / | | \ / | | | | | | (phonological processes)
(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
      | | | | \ / | | | | | | (melodic-phonological processes)
(2b)  co mo dewm fil mja ção que não va leu
      | | | | | | | | | | |
      N N N N N N N N N N
      # (filme a > filmja)
```

This gimmick is useful as it allows us, on the one hand, to untangle the analysis by separating phonological processes from “melodic-phonological” ones. Indeed, it is essential that the analysis of sung words should determine and describe as precisely as possible the extralinguistic (melodic, rhythmic, harmonic, etc.) environment that determines the actualisation or blockage of a phonological process.

At the same time, this analysis demonstrates how imprecise musical terminology is when it comes to telling syllabic melody — “a song in which each syllable has but one note” — from melismatic melody — “a melody in which more than one tone is sung to a syllable” — (Hartong, 2007, p. 160). As we have shown, all melodies are syllabic, without exception, and the distinction between syllabic and melismatic melodies is about the written syllable, not the syllable actually sung.

## Conclusion

To summarise:

- sung words may present prosodically anomalous constructions;
- such constructions, which occur as exceptional syllabic adjustments, are not accepted in speech, but go unnoticed when sung;

- such constructions can be adequately described by regarding sung words as an interaction between two strings, a syllabic and a melodic one, whose metrical elements are respectively the syllable and the note;
- the interaction between the strings is governed by Metrical Pairing, which establishes a bijective relationship between note and syllable;
- the inviolability of this condition is revealed in the systematic way in which phonological processes of syllable adjustment are blocked when they should be actualised, or actualised when they should be blocked;
- although superficially manifested as a phonological process, the syllabic adjustment observed in sung words must be regarded as a melodic-phonological process which characterises the interaction between the syllabic string and the musical melody;
- the environment where these adjustments occur is not linguistic (phonological, syntactic, morphological), but musical.

Metrical Pairing expresses the most fundamental of the conditions for the well-formedness of sung words. For sung words to be well formed, it is absolutely necessary that the number of terminal elements in the syllabic string should be equal to the number of terminal elements in the melodic string. As this condition does not allow exceptions and supersedes any other condition of a linguistic nature, it is not uncommon for there to occur resyllabifications that are strange to Portuguese and which, when spoken, sound like prosodically ungrammatical constructions. We have seen that these constructions can be described by the rules “delete  $\sigma$ ” and “insert  $\sigma$ ”.

Metrical Pairing is the most basic condition for the well-formedness of sung words because its domain is restricted to the correspondence between notes and unspecified syllables. Evidently, the interactions between melody and speech go well beyond this. Syllables carry primary or secondary stress, can be unstressed or stressed monosyllables, or even the core of phonological phrases. In turn, a musical note can fall on a strong or a weak beat, it may or may not be the core of a rhythmic group, it may or may not have a harmonic role, and so on. In other words, in addition to the bijective relationship established between note and syllable, these terminal elements interact with elements which may be hierarchically superior to them. More than that, these hierarchies seem to interact with each other. A description of their interactions goes far beyond the scope of this paper; they are mentioned here only as a reminder that Metrical Pairing and the rules derived from it can be used to explain only a small set of the linguistic-melodic phenomena observed in sung words.

- *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.*

## REFERENCES

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

ABOUSALH, E. **Resolução de choques de acento no português brasileiro**: elementos para uma reflexão sobre a interface sintaxe-fonologia. 1997. 157f. Masters (Dissertation) - 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. Thesis (Doctorate in Philosophy) – Department of Linguistics and Philosophy, Massachusetts Institute of Technology, 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. Thesis (Doctorate in Philosophy) – Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, 1980.

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

Received March 8, 2021

Approved on October 13, 2022