Singing in Tashlhiyt Berber, a language that allows vowel-less syllables

François Dell

Charles E. Cairn & Eric Raimy, eds., 2011, *Handbook of the Syllable*, 173-193. Leiden: Brill.

Singing in Tashlhiyt Berber, a language that allows vowel-less syllables

1. Introduction[1]

A song is a composite object with two components, a linguistic object, the 'text', and a musical object, the 'melody'. The two objects have structures that are independent of one another, and each can be realized in the absence of the other. An essential feature of singing is that the text and the melody are produced simultaneously by the same machinery, i.e. the mind and the vocal apparatus of the same person. An interesting question is what happens when the text and the melody make conflicting demands on that machinery, as in the case of singing in a tone language. This article deals with singing in Tashlhiyt, a language in which some syllables are entirely made up of voiceless consonants. What happens to the musical pitches associated with such syllables?

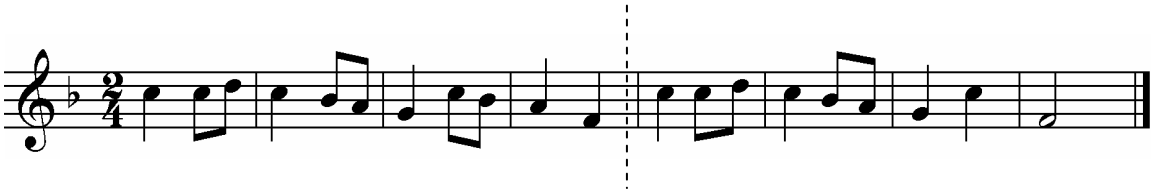
This work is part of a larger endeavour, the search for a theory that would account for text-to-tune alignment in the world's singing traditions. Our discussion of textsetting in Tashlhiyt has at least three implication that are of general interest. One is that text-to-tune alignment may make crucial use of properties of the speech stream that do not play any role in the phonology of the language nor in the composition of metrical verse. Another is that text/melody mappings must be represented at two different levels of abstraction. The third implication is that the representations for text-to-tune alignment in singing differ from those of autosegmental phonology in at least one important respect: they only allow one-to-one correspondences. This difference is related to the fact that musical melodies have an intrinsic temporal structure, whereas sequences of tones do not.

[1] I wish to thank Dianne Bradley for calling my attention to the article by Sundberg and Bauer-Huppmann (2007), Takeki Kamiyama for helping me with the Japanese example in section 5, and Shigeko Shinohara for lending me Anon. (1977) and a CD with recordings of some of the songs gathered there. I also thank John Frampton, John Goldsmith and John Halle for their comments. The shortcomings in this article are my sole responsibility.

2. Some basic assumptions about text-to-tune alignment.

We first present some basic assumptions about text-to-tune alignment in singing. (1) below shows the melody of the French children's song *J'ai lié ma botte*, and (2) shows the text of three stanzas.[2]

(1)



(2) a Au bois voisin l'y a des v i olett e s, de l'aubépine et de l'églantier.

In the woods nearby there are violets, hawthorn and eglantine.

b J'en cueillis tant, j'en avais plein ma hott e . Pour les porter j'ai dû les lier.

I gathered lots of them, so that my basket was full. To carry them I had to tie them up.

c En revenant j'ai rencontré un princ e . Avec mes fleurs je l'ai salué.

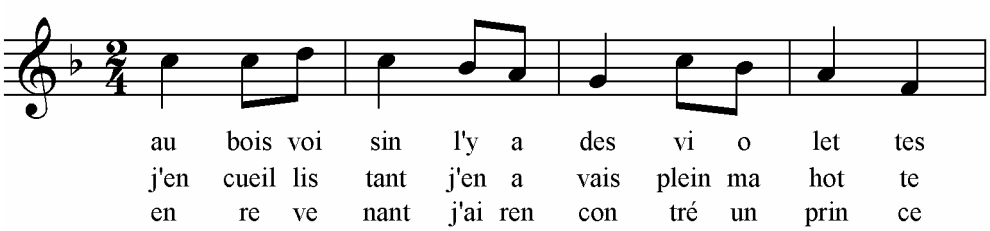
On my way back I met a prince. With my flowers I greeted him.

This song has a strophic form. In songs with a strophic form, the same melody is repeated over and over and each repetition is combined with a different portion of the text. Strophic songs are useful for understanding text-to-tune alignment in singing. They are a natural experiment in which the text varies while the melody is held invariable, or nearly so.

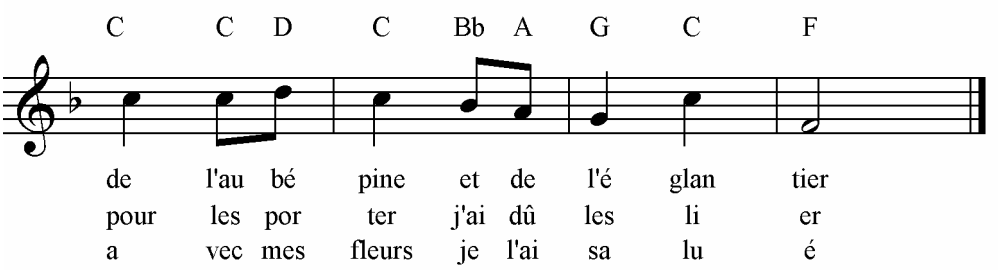
[2] Berthier (1979: 149). The boxed letters will be explained later.

The melody is a sequence of notes and the text is a sequence of syllables. The two sequences are mapped onto one another. The syllables in each of the three chunks of text in (2) are associated with the notes in (1). The result is shown in (3)-(4) below, where the score in (1) has been broken into two halves to fit between the margins of the printed page. (The dashed line in (1) indicates the breaking point.) The capital letters above the notes of (4) have been added as an aid for readers who are not familiar with musical notation.

(3)



(4)



Musical scores like (1) are conventional ways of representing melodies in the Western musical tradition, but when one wants to understand the structure of melodies and their relation to texts in singing, representations like (5) below are more insightful.

(5)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Time grid	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	|		|	|	|		|	|	|		|		|			
Pitch	C		C	D	C		Bb	A	G		C		F			
a	de		l'au	bé	pine		et	de	l'é		glan		tier			
b	pour		les	por	ter		j'ai	dû	les		li		er			
c	a		vec	mes	fleurs		je	l'ai	sa		lu		é			

(5) represents the same three settings as those displayed in (4). The representation of the melody is that part of the diagram which is enclosed in the dashed box. Two kinds of information are necessary in order to characterize a given melody: one must specify a certain sequence of pitches and one must specify the time intervals that separate the attack points of the successive pitch events. For the purposes of this paper, the only difference between (4) and (5) that matters has to do with how durations are represented. In (5) pitches are symbolized by capital letters and durations are represented in terms of a time grid. A time grid is a sequence of positions which represent points evenly spaced in time. Every pitch in the melody is linked to a grid position that represents the point in time when that pitch begins, i.e. its attack point. The time grid in (5) is a sequence of 16 positions, which are numbered for ease of reference. As grid positions are evenly spaced in time, durations are measured by counting grid positions. We see for instance that the duration between the attack of Bb and the attack of the preceding C is twice the duration between the attack of that same Bb and that of the following A.[3] In what follows the terms 'note' and 'pitch event' are used interchangeably to refer to a pitch together with its associated grid position. In (5), grid positions 1, 3 and 5 are the loci of three different pitch events with the same pitch C.

We now turn to the correspondence between the melody and the text, which is represented by double bars in (5). Each syllable is linked to a note. As a consequence of being linked to a note, every syllable is linked to a grid position. This point bears emphasizing. In the view of textsetting that we are advocating, the melody and the text do not play symmetrical roles as far as timing is concerned. The timing of the speech flow is melody-driven, meaning that the

[3] Our time grids are simplified versions of the objects commonly known as metrical grids, for a detailed presentation of which see Lerdahl and Jackendoff (1983). Metrical grids are multi-tiered representations in which each grid position has two attributes, a location in time and a metrical strength. As metrical strength does not play any role in this paper, we restrict our attention to timing, which can be represented on a single tier. For discussions of textsetting that make use of full-blown metrical grids, see e.g. Halle and Lerdahl (1993), Hayes and Kaun (1996), Dell and Halle (2009).

relationship between the text and the time grid is mediated by the melody. Because it is mapped onto the melody, the text has a time structure which is in part dictated by that of the melody. The melody provides a rigid temporal frame to which the text must adapt. In French and in Tashlhiyt this adaptation is effected primarily by stretching the vowels. In our view the basic conditions that govern textsetting in these two languages are the following:

(6) BASIC CONDITIONS ON TEXT-TO-TUNE ALIGNMENT:

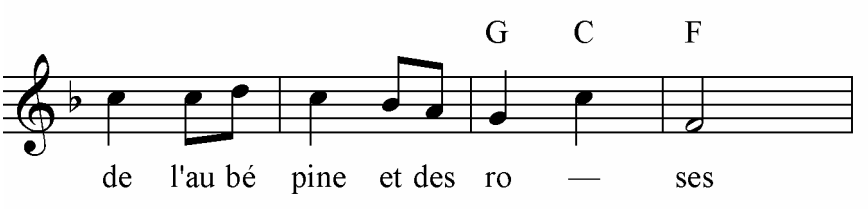
a Every syllable must be associated with one and only one note.

b Two syllables may not be associated with the same note.

These conditions prohibit many-one correspondences between the two sequences, an important difference with autosegmental representations in phonology, which allow correspondences between the tonal tier and the segmental tier that are not one-one. Lack of space prevents us from justifying (6) here. Let us just point out one implication that will be relevant when we discuss textsetting in Tashlhiyt.

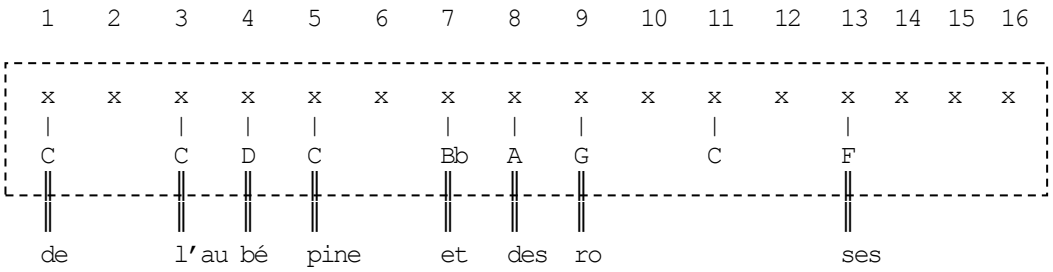
In (1) the melody is comprised of 20 notes and in (2) each stanza can be parsed as a sequence of 20 syllables, and so in (3)-(4) each syllable carries just one note, but this is not always the case; a syllable may carry several notes in succession. For instance, if at the end of (2)a the phrase *de l'églantier* (4 syllables) is replaced by *des roses* (3 syllables), the changed text can still be set to the melody so as to abide by the conventions of traditional French singing. This is shown in (7) below, which only depicts the second half of the modified stanza (cf. (4)a).

(7)



In (7) the first syllable of *roses* is drawn over two notes. The change from pitch G to pitch C occurs approximately at the halfway point in the production of the vowel [o]. If we use a time grid instead of a conventional musical score to represent the facts of (7), we get (8) below, in which the C pitch is not linked to *ro-*:

(8)



Note that although (6) requires every syllable to be linked to a note, it does not require the converse, and so (6) is not contravened by the C pitch at position 11, that does not have an associated syllable. In (8) the syllable *ro-* serves as a carrier for two pitch events in succession: the note G associated with *ro-*, and the unassociated note C that follows it. There is no need to draw an association line between C and *ro-*, for the fact that the final portion of the vowel [o] serves as a carrier for the note C is already implicit in (8). To see why this is so, consider how representations like (8) are implemented phonetically.

We stated earlier that when a pitch is linked to a grid position, that position represents the inception of the pitch event. This implies that in the implementation of the melody in (8), the pitch change from G to C must occur at position 11, regardless of where position 11 happens

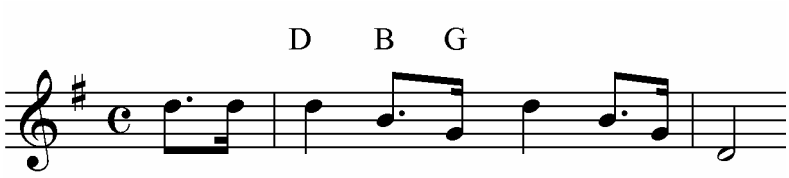
to be located in relation to the vowels and consonants in the text. We assume furthermore that when a syllable and a note are linked, the beginning of the syllable's nucleus is simultaneous with the note's attack (Sundberg and Bauer-Huppmann, 2007). In (8), if the attack of the note F coincides with the beginning of the vowel in the syllable *–ses*, the preceding note C can only be realized on the vowel of *ro-*.[4]

Configurations in which successive notes are carried by the same syllable are called melismas. In (8) the syllable *ro-* is said to be melismatic. Melismas occur when a melody has more notes than there are syllables in the associated text. In such circumstances, some notes must be left unassociated in order to ensure that text-to-tune alignment is to meet condition (6)a.

Melismas only have a limited freedom of occurrence, but in general their distribution is not predictable, and consequently their location in a song must be memorized. A case in point is the fifth line in every stanza of the *Marseillaise*, in which the melody is made up of nine notes while the text has only eight syllables. Consider for instance the first stanza, where the text of the fifth line is *L'étendard sanglant est levé* '(Tyranny's) blood-soaked standard has been raised'. (9)a and (9)a' below are alternate settings for this line. ((9)b,b' will become relevant later.)

[4] An additional assumption is that it is vocalic segments that undergo most of the durational adjustments involved in text-to-tune alignment (Scotto di Carlo 2007): in (8) the initial consonant of *–ses* cannot be stretched so as to take up the time interval between positions 11 and 13, and so it is the vowel of *ro-* that must be drawn out.

(9)



a	l'é ten dart — san glant est le vé	STANZA 1
a’	l'é ten dart san — glant est le vé	
b	ces fers dès — long temps pré pa rés	STANZA 2
b’	ces fers dès long — temps pré pa rés	

Settings (9)a and (9)a’ are both well-attested. The choice between them is not a matter of free variation. Some speakers are used to one setting, others to the other. To memorize a song, then, it is not enough to memorize the melody and the text independently of one another; the location of melismas must also be memorized. Diagrams like (5) and (8) represent the knowledge that a person has when one says that that person knows a song. These representations are of a rather abstract nature and we call them ‘underlying’ to set them apart from less abstract ones that reflect the correspondences between texts and melodies that are observed in actual song performances. As we shall see in our discussion of Tashlhiyt, there are cases where an underlying melisma does not have a melisma as its surface reflex, and others in which a melisma occurs in the surface at a point where none is present in the underlying representation.

A notable feature of traditional singing in French and in Tashlhiyt is that melismas recur at the same place in every stanza. Returning to our example from the first stanza of the *Marseillaise,* the corresponding line in the second stanza, also a sequence of eight syllables, is *Ces fers dès longtemps préparés* ‘These long-prepared fetters’. Someone who sings the line in the first stanza with a D-B melisma on the third syllable will do the same with this line (cf. (9)a,b), whereas someone who sings the line in the first stanza as in (9)a’ will sing this line as in (9)b’, with a B-G melisma on *long*-. The level of representation at which this parallelism

across stanzas is enforced is the underlying level.[5] At the surface level, the parallelism is sometimes perturbed locally, as we shall see below.

3. Syllables and meter in Tashlhiyt Berber

This discussion of text-to-tune alignment in Tashlhiyt draws heavily on work done in collaboration with Mohamed Elmedlaoui, currently at the Institut Universitaire de la Recherche Scientifique in Rabat, see especially DE (2002, 2008).[6] Tashlhiyt (henceforth TB) is a Berber language of Western Morocco spoken by around five million people. TB speakers have their own distinctive musical tradition, which differs from that of the other Berber languages and from that of Moroccan Arabic. Due to the deep changes that Moroccan society is undergoing, TB singing is in a phase of rapid evolution, and we will only be concerned here with the traditional singing style, what TB speakers call *amarg aqqdim* 'old poetry'. *Amarg aqqdim* is still very much alive in the countryside. Up till very recently, in TB-speaking villages everyone was fluent in it. There is no explicit teaching of the conventions which regulate the structure of TB verse or its relation to music. Children simply acquire these conventions by listening to songs and singing along with others. TB has no instituted writing system. Most composers of TB verse are illiterate in any language.[7]

(10) below is a list of the simplex phonemes of TB. Except for /a/, /i/ and /u/, every phoneme in the table has a geminate counterpart not listed there.[8]

---

[5] The uniform distribution of melismas across stanzas is a consequence of 'positional parallelism', on which see Dell and Halle (2009).

[6] References to that work use the abbreviation DE (= Dell and Elmedlaoui).

[7] On music and poetry in TB-speaking areas, see e.g. Schuyler (1979a,b), Lortat-Jacob (1980), Rovsing Olsen (1997), DE (2008). Our data on performance in songs come from commercial recordings on tape cassettes by professional artists. DE (2008) contains a CD with audio recordings of nine songs discussed in the text. These recordings are divided into line-size portions labeled for easy reference to the line-by-line parses in the text.

[8] 'š', 'ž' and 'x' respectively stand for IPA 'ʃ', 'ʒ' and 'χ'.

(10) THE PHONEMES OF TB

	t	$t^ʕ$			k	k^w	q	q^w					
b	d	$d^ʕ$			g	g^w							
f	s	$s^ʕ$	š	$š^ʕ$			x	x^w	ħ				
			ž	$ž^ʕ$			ʁ	$ʁ^w$	ʕ	ɦ			
m	n	$n^ʕ$											
w	l	$l^ʕ$	r	$r^ʕ$	y								
u					i				a				

TB allows long consonant clusters, but syllable structure is rather simple. As a result, some syllables contain only consonants, with even voiceless stops acting as syllable nuclei. Here I follow the analysis of TB syllabification in verse that is given in DE (2002), where readers can find detailed references to other work on TB phonology. That analysis is summarized in (11).

(11) THE CENTRAL TENETS OF SYLLABIFICATION IN TB VERSE (A and B are inviolable; they take precedence over C when they conflict with it):

A Syllables must satisfy the syllable template in (12)a below, where O, R, N and C respectively stand for Onset, Rime, Nucleus and Coda. Except for /a/ which can only be a nucleus, any simplex segment can be an onset, a nucleus or a coda. Only geminates can act as complex codas.

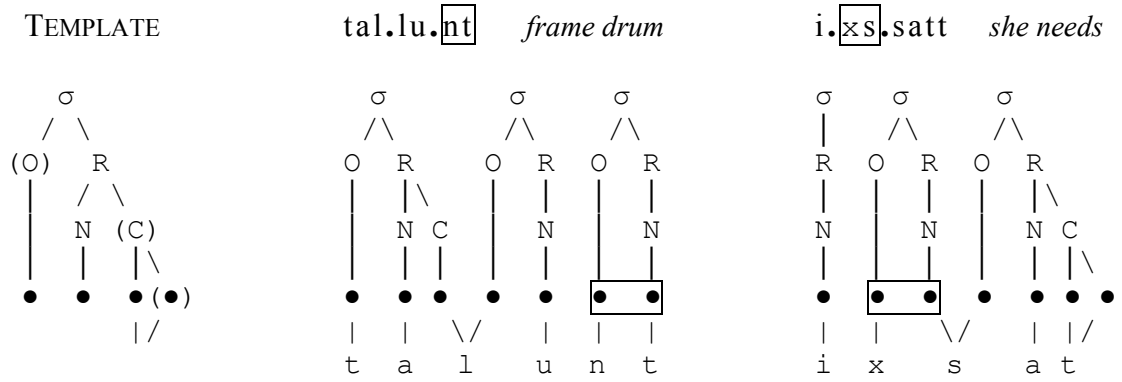
B Except at the beginning of lines, syllables must have an onset (no hiatus allowed).

C Sonority peaks are syllable nuclei.

(12)b and (12)c are given by way of exemplification. A geminate is represented as a single bundle of distinctive features associated with two skeletal positions, which are symbolized there as '●', to keep them apart from grid positions, symbolized as 'x'. (While grid positions

are part of the make-up of melodies, skeletal positions belong to phonological representations.) Note the boxed syllables, whose nuclei are voiceless obstruents, /t/ in (12)b and /s/ in (12)c.

(12) a b c



(13) below shows how the phoneme sequence of a line of verse is divided into syllables according to the analysis summarized in (11).

(13) iqqand akk asiʁ, nsbr issaʕtnk nzzritnt[9]

iq qan dak ka si ʁn sb ris saʕ tn knzz ri tn <t

I must take you (as you are), accept this episode with you and live through it.

The lyrics of TB songs have a metrical structure. TB songs have a strophic form and in most songs all the lines are built on the same meter. The basic principles of TB versification were discovered by Hassan Jouad, see Jouad (1983, 1995). TB meters are based on a distinction between heavy (H) and light (L) syllables. When two lines follow the same meter, they have the same number of syllables, with H and L syllables arranged in the same order. In (14) below the upper half of the table displays the syllabic parses of five lines in a song whose

[9] Line 30 in Song 7 in DE (2008). The angled bracket indicates a line-final segment that is extrasyllabic; see DE (2008) on line-final extrasyllabicity.

meter requires that every line have 13 syllables, with the 2nd, the 9th and the 11th syllables H and all the other syllables L.[10]

(14)	1	2	3	4	5	6	7	8	9	10	11	12	13	
	L	H	L	L	L	L	L	L	H	L	H	L	L	<
1a	a	daʁ	su	lu	ri	ri	na	ra	ʁitt	mu	nad	wa	ya	<d
2a	i-	qan	da-	ka	si	ʁn	sb	ri-	saʕ	tn	knzz	ri	tn	<t
3a	n	šrk	di	wn	ma	du	ra-	tu	nar	ki	ʁngg	za	ka	<l
4a	i	ninn	si	kn	ga-	sa	ħt	ky	yin	tr	mit	gi	tn	<ʁ
5a	ha-	nur	da	sg	ga	su	la	si	nad	la-	ħlx	ba	rn	<k
1b	a	daʁ	su	lu	ri	ri	na	ra	ʁitt	mu	nadə̲	wa	ya	<d
2b	i-	qan	da-	ka	si	ʁnə	səb	ri-	saʕ	tnə	kənzzə̲	ri	tn	<t
3b	nə	šrək	di	wnə	ma	du	ra-	tu	nar	ki	ʁnəggə̲	za	ka	<l
4b	i	ninn	si	knə	ga-	sa	ħət	kəy	yin	trə	mitə̲	gi	tnə	<ʁ
5b	ha-	nur	da	səg	ga	su	la	si	nad	la-	ħləxə̲	ba	rnə	<k

In TB verse syllabic weight is assigned as follows: syllables without a coda are L and syllables with a coda are H, except that when a coda is the first half of a geminate the syllable where it occurs can count as either L or H. In (14) and hereafter such 'ambiguous' codas are represented by hyphens.[11]

The mastery of verse forms like that in (14) is not limited to a small group of people with a special training. In the countryside all TB speakers have an intuitive grasp of these meters. We are now ready to discuss how melodies and texts are put together in TB songs.

[10] Lines 36, 30, 26, 28 and 16 in Song 7 in DE (2008). The lower half of the table will be explained below.
[11] The parse in row 2a is that in (13) rewritten so as to follow this convention.

4. Schwa as a carrier for pitch

In addition to the underlying vowels /a, i, u/, TB has a vowel-like sound that we call schwa. The status of schwa in the phonology of TB is discussed at length in our earlier work, where it is argued that schwas are nothing more than transitions between adjacent consonants.[12] TB schwas do not play any role in the phonology of the language (e.g. in syllable structure) nor in versification. In text-to-tune alignment, however, schwas acts as carriers for pitch, exactly like *bona fide* vowels, as we shall now see.

The transcriptions in the lower half of table (14) indicate how the five lines in the upper half actually sound in a recording of the song. Syllables with schwas are boxed for the sake of conspicuousness. Even in singing, schwa does not occur in all syllables with consonantal nuclei (see syllable 13 in row 2b in (14)), but its distribution is the following in those syllables where it does occur:

(15) Schwa is always adjacent to the nucleus. (a) If the nucleus belongs to a geminate, schwa precedes it. Otherwise, (b) schwa precedes the nucleus if it is an obstruent, and (c) schwa follows it if it is a sonorant.

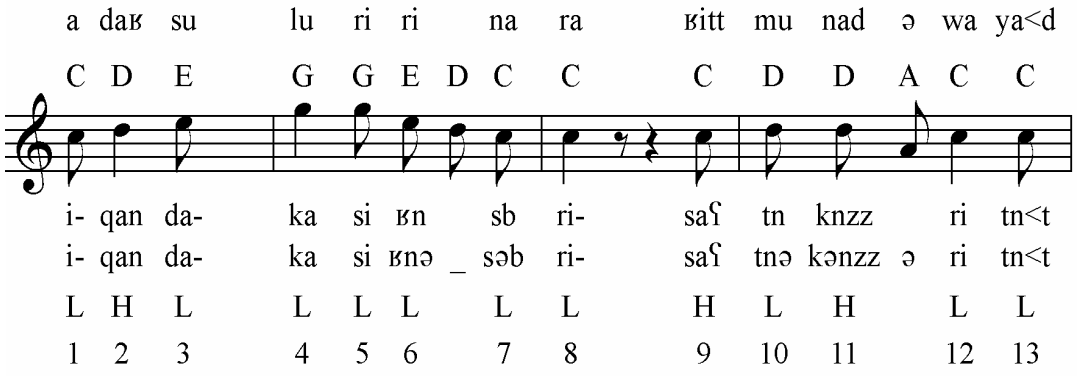
(15) does not cover the schwas which are underlined in column 11 in table (14); these represent a special case that will be discussed below.[13] Otherwise the schwas in the lower half of (14) all obey (15), except for syllable 11 in line 2b, which contradicts (15)c. (15)c is only a first approximation; see DE (2008) for details.

To illustrate the role of schwa in singing, let us consider (16) below, which shows how the first two lines in (14) are set to music. Line 1 is aligned above the score, and line 2 is aligned

[12] See Hall (2006) on other languages with 'intrusive vowels'.
[13] See diagram (17) and the text around it.

under it. The center of interest is actually line 2; line 1, in which all the nuclei are vowels, is only given to provide a baseline.

(16) 1

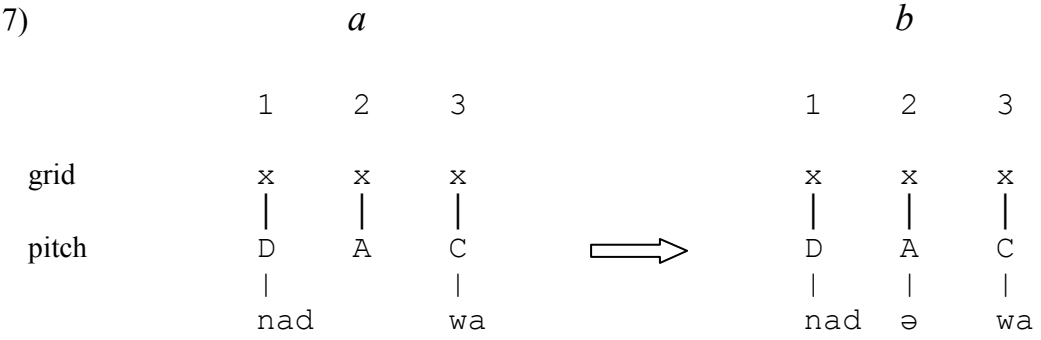


2a	i-	qan	da-	ka	si	ʁn	sb	ri-	saʕ	tn	knzz	ri	tn<t
2b	i-	qan	da-	ka	si	ʁnə	_	səb	ri-	saʕ	tnə	kənzz	ə	ri	tn<t
	L	H	L	L	L	L	L	L	H	L	H	L	L
	1	2	3	4	5	6	7	8	9	10	11	12	13

(16)-2a shows the underlying correspondence between syllables and pitch events, and (16)-2b shows how this correspondence is implemented phonetically. Let us list the syllables with consonantal nuclei (nuclei are underlined): /ʁn/ (6), /sb/ (7), /tn/ (10), /knzz/ (11), /tn/ (13). In the last syllable, which is not realized with a schwa, the nuclear consonant /n/ serves as a carrier for the associated C pitch. The four other syllables are pronounced with schwas and it is those schwas, rather than the nuclear consonants, that serve as carriers for pitch. Take for instance syllable 6, which must bear the pitch sequence E-D. In line 1, the articulation of vowels and consonants in the text is timed so that the attack of E is heard to coincide with the beginning of the nucleus /i/. In line 2, on the other hand, where the nucleus of syllable 6 is /n/, the melisma is not realized on that nucleus but on the following schwa, whose articulation is maintained as long as required by the combined durations of notes E and D. The principle which governs text-to-tune alignment at the phonetic level seems to be the following: when a syllable is linked to a note, the note must begin at the same point as that portion of the syllabic rime which is highest on the sonority scale. The same principle holds trivially for French, where it is always the nucleus which has the highest sonority within a syllabic rime.

Schwa occurs in a wider range of contexts in singing than in everyday speech. We have for instance encountered occurrences of schwa between two voiceless consonants,[14] a context where schwa never seems to occur in everyday speech.[15] Other occurrences of a kind found only in singing are those after syllable 11 in (16), which serve as carriers for a note that does not have any associated syllable at the underlying level. Like syllable 6, syllable 11 in (16) is underlyingly melismatic, meaning that at the underlying level it is followed by an unlinked note, and one would expect it to be melismatic on the surface as well, as is the case for syllable 6. No melisma is heard on syllable 11, however. What happens instead is depicted in (17) below, which represents the situation in (16)-1. (17)a depicts the underlying correspondence between the melody and the text, and (17)b shows how the correspondence is implemented phonetically.

(17)



At the underlying level, the note D associated with /nad/ is followed by a note A which does not have any syllable linked to it. One would expect the vowel in /nad/ to take up most of the time interval between positions 1 and 3 so as to serve as a carrier for notes D and A in succession: the portion of [a] between positions 1 and 2 would be sung to pitch D, and the portion between positions 2 and 3 would be sung to pitch A. Instead, the coda of /nad/ is

[14] Syllable 7 in (14)-4b is a case in point.
[15] On syllables that are phonetically voiceless, see our previous work and also Ridouane (2003, 2008).

released into a schwa that begins at position 2 and serves as a carrier for the A pitch. A schwa is heard after syllable 11 in every line of the song; these schwas are underlined in row 11 of table (14). What is special about them is that they occur only in environments that meet certain requirements on metrical structure and on the associated melody: the preceding syllable must be an H linked to a note immediately followed by a note not associated with any syllable; see DE (2008) for more details.

At this point it is instructive to compare TB schwas with the paraphonological vowels that occur in French singing. I adopt the term 'paraphonology' from Kiparsky (1977) to designate phonological patterns that arise from conventions specific to poetry. French singing allows pronunciations that are not licit in the colloquial language. Take for instance *violettes* in (2)a. In conversational Parisian French this word's only licit pronunciation is [vjɔlɛt], which is bisyllabic, but as sung in (3)a it is pronounced quadrisyllabic: [vi.jɔ.lɛ.tə]. The word has in addition two trisyllabic pronunciations that are licit in singing but not in the colloquial language, [vi.jɔ.lɛt] and [vjɔ.lɛ.tə]. *Violettes* illustrates the two main paraphonological phenomena that occur in French singing: the blocking of prevocalic gliding and the occurrence of schwas unacceptable in other styles of delivery.[16] In (2) paraphonological vowels are enclosed in boxes.

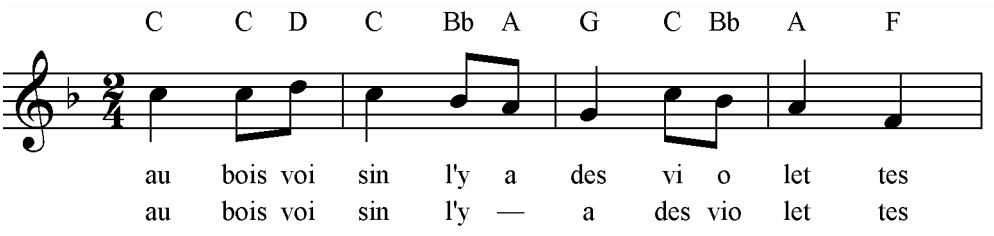
There is an important difference between paraphonological vowels in French singing and schwa in TB: TB schwa is metrically irrelevant. It is absent from the representations which are inputs to syllabification and to the computation of syllabic weight; see e.g. the representations in the upper half of table (14). Its only relevance is to text-to-tune alignment below the syllable level. This is in marked contrast with French, where paraphonology

[16] The paraphonology of singing is not the same as that of classical French poetry. For a summary of the facts about schwa in French singing, see Dell (1989).

provides alternative syllable counts not available in everyday speech, as the following example illustrates.

Let us return to the word *violettes* in (2)a. Instead of singing it as [vi.jɔ.lɛ.tə] as in (3)a, if one sings it as [vjɔ.lɛ.tə], the utterance to be mapped onto the melody in (3) has one syllable less, and consequently one note must be left without an associated syllable. One felicitous setting is depicted below in (18)a', with the setting (3)a reproduced as (18)a for the sake of comparison.

(18)



In (18)a', to compensate for the gap due to the monosyllabic pronunciation of *vio-*, the words *a* and *des* move over to the right, which leaves the note A unassociated, and so the syllable *l'y* is prolonged so as to serve as a carrier for it, in addition to the preceding note Bb. In the case of TB, an analogue of this example would be one in which the suppression of a schwa would necessitate a change in the mapping of two syllables before that schwa. TB has indeed paraphonological processes which provide alternate syllable counts for an expression, but none of these involves schwa.[17]

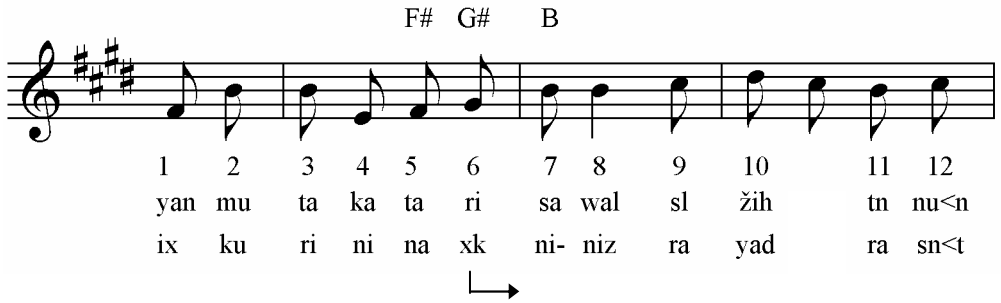
TB is not the only language on record in which textsetting makes use of phonetic material that is not metrically relevant. Khmer inserts nasal consonants to serve as pitch carriers, but these consonants do not play any role in versification. (see Jacob 1966).

[17] /a/ or /wa/ may be inserted at certain line-internal syntactic breaks to round off the meter (DE 2008: 63), and certain morphemes have poetic variants.

5. Derived melismas

In all the examples discussed up to this point, syllables with consonantal nuclei can serve as carriers for pitch because glottal vibration is involved in their production: pitch is realized on a schwa adjacent to the nucleus, or if there is no schwa, pitch is realized on the nucleus itself, a sonorant, as in syllable 13 in (16)-2b. Let us now consider syllables whose rimes are produced without glottal vibrations, i.e. syllables in which the nucleus is a voiceless consonant and there is no abutting schwa. This situation is dealt with in two ways in TB singing. One way is simply not to sing the note associated with the syllable in question. The unrealized note is presumably reconstructed by the listeners' minds.[18] The other way is what we call 'rightward shift': the note is realized on the preceding syllable, which becomes melismatic. Rightward shift is illustrated in line b in (19) below; there is no shift in line a, which is given for the sake of comparison.[19]

(19)

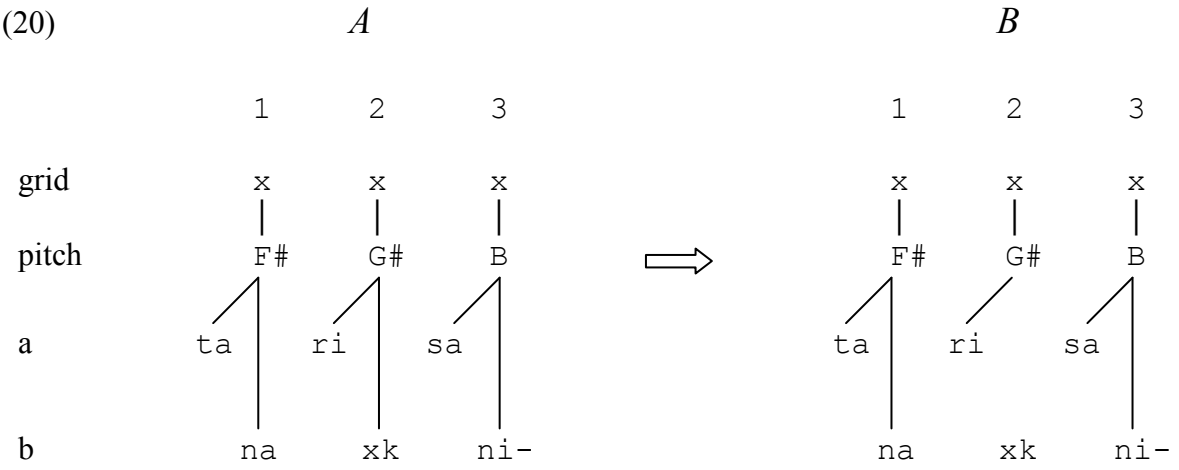


(19) shows how syllables and notes are aligned at the underlying level. The nucleus of syllable 6 in line b is a voiceless stop and there is no schwa adjacent to it. The arrow is meant to suggest what is actually heard in the recording: syllable 6 is shifted to the right and its associated note G# is realized on the preceding syllable. Our analysis of the situation is that

[18] See DE (2008) for details.
[19] Lines a and b in (19) are the respective parses of lines 9 and 2 in Song 2 in DE (2008). In (19)b the onset of syllable 6 is an underlying /ʁ/ devoiced under the influence of the following consonant. That consonant is in fact /k^{w}/, but the right superscript 'w' that represents rounding is omitted for typographic convenience.

rightward shift is an epiphenomenon of delinking. This analysis is diagrammed in (20) below, where, line a is again given for the sake of comparison:

(20)



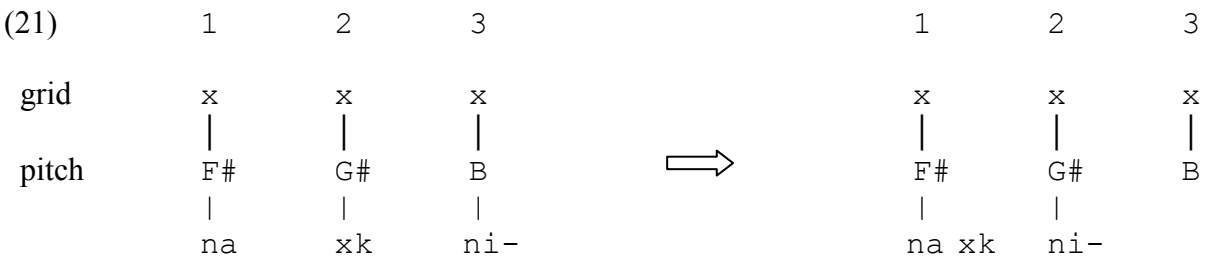
(20)A depicts the mapping between the text and the melody at the underlying level and (20)B depicts it at the surface level. The left and right diagrams are identical for line a. For line b, on the other hand, the syllable /xk/, which is linked to G# in the underlying representation, is no longer linked in the surface representation. As a consequence of this delinking, the surface representation of text-to-tune alignment does not make any demands of its own on how the time interval between grid positions 1 and 3 is to be divided up between the segments in the sequence /axkn/, i.e. in the portion of the speech stream that extends from the beginning of [a] in /na/ to the beginning of [i] in /ni-/. Since consonants can be prolonged only to a limited extent and voiceless segments cannot bear pitch anyway, the vowel [a] takes up most of the time interval and the attack of G# occurs during the production of [a], whence an F#-G# melisma on /na/.

It is important to understand that rightward shift only concerns the relative timing of a syllable and its associated note; the temporal structure of the melody remains unaffected. At the surface level, the attack of G# is located at midpoint in the time interval between the

attack of F# and that of B, and this holds for line b as well as for line a or any other line sung to the tune in (19).

In the TB songs we have examined there are numerous cases where a syllable is shifted in time and its associated note is realized on an adjacent syllable. In none of these instances does the shifted syllable contain a full vowel or a schwa. We take this as evidence that rightward shift aims at keeping the melody invariable regardless of the phonetic material mapped onto it, i.e. that rightward shift seeks to ensure that every note is phonetically realized and that it begins at its appointed position on the time grid.

Another remarkable property of the shift is that it is always to the right. Our explanation for this fact is that rightward shift is the smallest change needed to keep the melody invariable. Shifting the syllable to the left would also achieve this aim, but only at an additional cost. Returning to example (20), let us see what would happen if /xk/ were shifted to the left so that its associated note was realized on the following syllable. Such a leftward shift is depicted below in (21).



In the output of (21) the syllable /ni-/ bears a melisma G#-B whose first note is the G# underlyingly associated with /xk/. The melody is identical on both sides of the arrow, but whereas the only change involved in the rightward shift in (20) is the delinking of /xk/, the leftward shift in (21) requires in addition that /ni-/ be delinked from B and reassociated with G#.

We noted earlier that the representations we are using for text-to-tune alignment differ from those of autosegmental phonology in that they do not allow many-one associations. In particular, we hold that linking a syllable with several notes is not the appropriate manner of representing melismas; see condition (6)a. Rightward shift provides us with one piece of evidence in favor of this claim. (22) below shows how rightward shift and leftward shift would be characterized using standard autosegmental representations. (The delinking of /xk/ from G#, which occurs both in right- and leftward shift, is not represented.)

(22) /xk/ shifts... to the right (20) to the left (21)

In standard autosegmental terms, the two shifts are mirror images of one another and there does not seem to be any reason why one should be preferred over the other. The picture given by (22) is deceptive, of course, in that it leaves out the time grid, i.e. it glosses over a fundamental difference between singing and the phonology of tone, which is that musical melodies have an intrinsic temporal structure whereas sequences of tones do not. If one is to look for a phonological analogue to rightward shift, it is compensatory lengthening that should come to mind, rather than the formation of contour tones.

Rightward shift also occurs in Japanese, where it is affects obstruent moras.[20] In Japanese singing the basic linguistic units involved in text-to-tune alignment are moras rather than syllables. Based on the discussions in Vance (1987: 67-70) and in Hayes (2008), and also on our examination of 25 songs in Anonymous (1967), we can state with confidence that to fit

[20] The phenomenon is briefly noted in Vance (1987: 67).

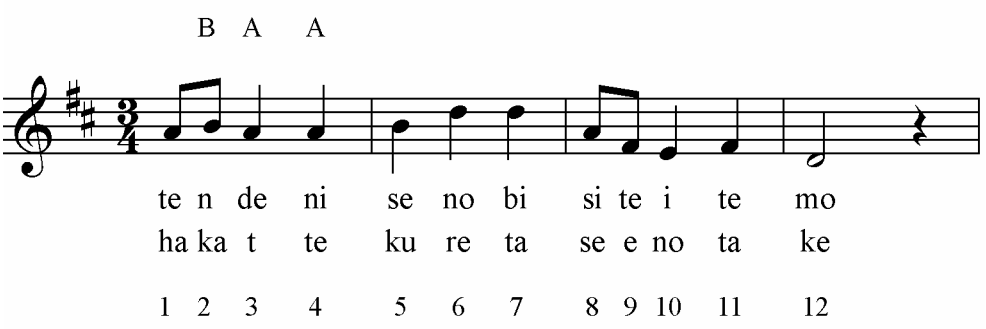
the facts of Japanese, our basic condition (6)a must be reworded by replacing 'syllable' with 'mora'.[21] The second mora of a Japanese syllable can be a short vowel, the final part of a long vowel, a nasal consonant or the first half of a geminate obstruent, the second half being the onset of the next syllable. The obstruent moras of Japanese present text-to-tune alignment with a challenge that is analogous to that of nuclear obstruents not flanked by a schwa in TB.

To illustrate, let us consider two homologous lines from *See kurabe* ('Comparing heights'), a well-known children's song. The text of these lines is given in (23). Each line is a sequence of twelve moras which is in a one-one correspondence with a twelve note melody, as shown in (24). [22]

(23) a tendeni senobi siteitemo

b hakatte kureta seeno take

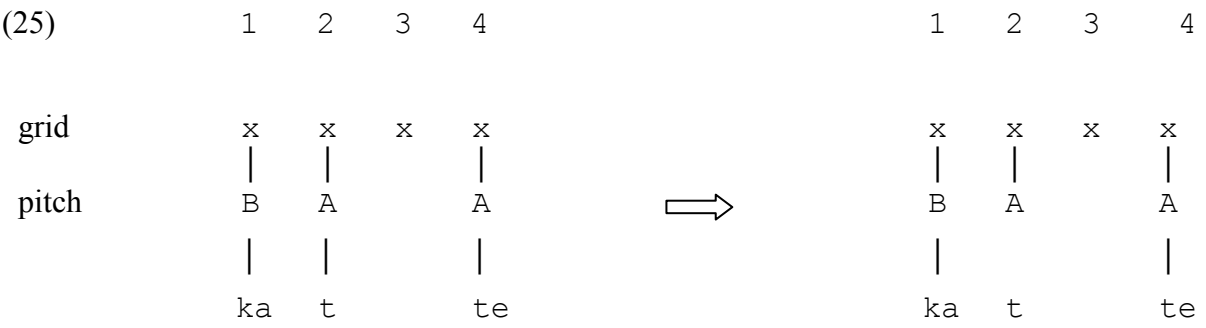
(24)



Like /de/ in line a, the second mora of /kat/ in line b is linked to the third pitch event (A), but whereas in line a the vowel of /de/ serves as a carrier for A, in line b it is the preceding [a] that bears that pitch. The underlying alignment in line b and its surface realization are depicted below in (25).

[21] It is not clear at this point whether condition (6)b must be changed in a similar manner. In any case syllables seem to play a role in Japanese textsetting, in addition to moras. Hayes (2008) notes that in the setting of bimoraic syllables there is a strong tendency for the position of the first mora to be metrically stronger than that of the second mora.

[22] See Anon. (1967: 116-117). The meaning of line a is 'Although they all stand on tiptoe to make themselves taller'; that of line b is 'The height that he measured (for me)'. A recording of this song can be heard on the internet at the end of the file *http://masaring.blogzine.jp/music/dohyohunno.mp3*.



Given that in /kat/ the second mora is underlyingly linked to a pitch event at position 2, the closure of the geminate /tt/ would be expected to begin at position 2, which would require that closure to take up the whole time interval between positions 2 and 4. What happens instead is that the /t/ mora delinks from A and the preceding [a] is prolonged to serve as a carrier for the note A. [a] is sung to pitch B over the first third of its duration, and to pitch A over the rest.

Melismas due to obstruent moras are always borne by the preceding mora, i.e. like that of syllable shift in TB, the direction of mora shift in Japanese is always to the right. For someone not looking beyond the facts of Japanese, rightward shift looks at first blush like evidence pointing to the relevance of syllables in Japanese singing. For Japanese, a plausible explanation of the direction of the shift could be that shifting the obstruent to the right, rather than to the left, allows the associated note to stay within the confines of the same syllable: in the input of (25), for instance, B and the following A are both linked to moras that belong to the same syllable /kat/, and in the output they are both borne by the same syllable, as they are both realized on the same vowel. However, if Japanese mora shift and TB syllable shift are the same phenomenon, as we think they are, the explanation for the direction of Japanese mora shift is the same as that proposed earlier for the direction of TB syllable shift, which has nothing to do with tautosyllabicity.

6. Concluding remarks

In the preceding discussion, text-to-tune alignment is represented at two levels of abstraction. In the underlying representations the objects that are put into correspondence are *structural units* in their respective domains: notes in the melody, syllables (TB, French) or moras (Japanese) in the text. In the surface representations, on the other hand, text-to-tune alignment is a correspondence between *points* in the melody and *points* in the speech stream. The points in the melody are the attacks of notes. Note attacks are linked to the left edges of portions of the speech stream best suited to bear pitch: full vowels or schwas when they are available, and sonorant consonants otherwise.[23] Rightward shift in TB and in Japanese is evidence in favor of this framework, which does not allow many-to-one associations, unlike autosegmental representations in phonology.[24]

We began this paper by bringing up situations of phonetic conflict between the melody and the text. Note that such conflicts are not a logical necessity. They only arise because the textual entities that are associated with musical notes at the underlying level are authentic linguistic units (syllables, moras), and in some languages not all such units are suitable for carrying pitch. One can conceive of alternative textsetting conventions which would make conflicts impossible altogether. Imagine for instance that in condition (6), 'syllable' is replaced by 'pitch span', a pitch span being defined as a maximal sequence of adjacent voiced sonorants. Presumably, singing idioms that abide by this modified version of (6) do not exist, and the reason why is that pitch spans as defined above do not play a role in the phonology of any language.

[23] See Gordon (2001) for some discussion of the fact that certain segment types are better suited than others to carry pitch.

[24] The data discussed in this paper are drawn from three singing idioms that are very much alike. It remains to be seen how this framework can accommodate singing idioms of a rather different type, such as free rhythm singing in TB (see DE 2008) or the idioms used for chanting Psalms in Latin and in English that are discussed in an autosegmental framework in Chen (1983, 1984).

References

Anon., 1977: *Nihon dōyō meika 110 kyokushū; 110 Famous Japanese Children's Songs*. Tokyo: ZEN-ON Music Co.

Berthier, Jean-Edel, 1979: *Mille chants. Anthologie du chant populaire*, tome 3. Paris: Presses de l'Ile-de-France.

Chen, Matthew, 1983: 'Toward a Grammar of Singing: Tune-Text Association in Gregorian Chant', *Music Perception* 1, 84-122.

Chen, Matthew, 1984: 'The interfacing of language and music', *Language and Communication* 4, 159-194.

Dell, François, 1989: 'Concordances rythmiques entre la musique et les paroles dans le chant', in Marc Dominicy, ed., *Le souci des apparences*, 121-36. Brussels: Éditions de l'Université de Bruxelles.

Dell, François, and Mohamed Elmedlaoui, 2002: *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer.

Dell, François, and Mohamed Elmedlaoui, 2008 : *Poetic Meter and Musical Form in Tashlhiyt Berber Songs*. (Enclosed audio CD) Cologne: Rüdiger Köppe Verlag.

Dell, François, and John Halle, 2009: 'Comparing musical textsetting in French and in English songs', in Jean-Louis Aroui and Andy Arleo, eds., *Towards a Typology of Poetic Forms*. Amsterdam : Benjamins.

Gordon, Matthew, 2001: 'A typology of contour tone restrictions', *Studies in Language* 25, 423-462.

Hall, Nancy, 2006: 'Cross-linguistic patterns of vowel intrusion', *Phonology* 23, 387-429.

Halle, John, and Fred Lerdahl, 1993: 'A generative textsetting model', *Current Musicology*, 3-23.

Hayes, Bruce, 2008: 'Two Japanese Children's Songs', revised version of ASUCLA Lecture Notes, Winter 1995. Los Angeles: UCLA.

Hayes, Bruce, and Abigail Kaun, 1996: 'The role of phonological phrasing in sung and chanted verse', *Linguistic Review* 13, 243-303.

Jacob, Judith, 1966: 'Some features of Khmer versification', in C. E. Bazell, J. C. Catford, M. A. K. Halliday, and R. H. Robins, eds., *In Memory of J. R Firth*, 227-41. London: Longman.

Jouad, Hassan, 1983: *Les éléments de la versification en berbère marocain tamazight et tachelhit*. Doctorat de Troisième Cycle, Université Paris 3.

Jouad, Hassan, 1995: *Le calcul inconscient de l'improvisation*. Paris and Louvain: Peeters.

Kiparsky, Paul, 1977: 'The rhythmic structure of English verse', *Linguistic Inquiry* 8, 189-247.

Lerdahl, Fred, and Ray Jackendoff, 1983: *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.

Lortat-Jacob, Bernard, 1980: *Musique et fêtes au Haut-Atlas*. Paris: Mouton et École des Hautes Études en Sciences Sociales.

Ridouane, Rachid, 2003: *Suites de consonnes en berbère: phonétique et phonologie*. Doctoral dissertation, Université Paris 3.

Ridouane, Rachid, 2008: 'Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber', *Phonology* 25, 321-359.

Rovsing Olsen, Miriam, 1997: *Chants et danses de l'Atlas*. Cité de la Musique / Actes Sud.

Schuyler, Philip D., 1979a: *A Repertory of Ideas: the Music of the Rways, Berber Professional Musicians from Southwestern Morocco*. Doctoral dissertation, University of Washington.

Schuyler, Philip D., 1979b: 'Rways and Ahwash: opposing tendencies in Moroccan Berber music and society', *The World of Music* 21-(1).

Scotto di Carlo, Nicole, 2007: 'Speech-rate and tempo effects on the organization of spoken and sung syllables', *XVIth International Congress of Phonetic Sciences*.

Sundberg, Johan and Julia Bauer-Huppmann, 2007: 'When does a sung tone start?'. *Journal of Voice* 21, 285-293.

Vance, Timothy J., 1987: *An Introduction to Japanese Phonology*. Albany: SUNY Press.