Tones in Tunes

A preliminary look at speech and song melody in Mambila

Bruce Connell

York University

Toronto

bconnell@yorku.ca

Abstract

The phonetic and phonological factors governing the realization of linguistic tone in speech have for some years been the focus of study by number of scholars. While the basic effects of such factors as declination, downdrift, and downstep, among other influences, are reasonably well understood, it is nevertheless fair to say that considerable work yet remains to be done in this area. A number of studies have focussed on the relation between speech melody and song melody to better understand pitch realization. Mambila, a Bantoid language with four level lexical tones, presents an excellent test case for exploring the possible effects of song on linguistic tone. This paper presents an initial view of the interaction between speech melody and song melody in Mambila through comparison of productions in speech and song with respect to overall pitch range used, and the spacing of individual tones. The results of the investigation not only contribute to the debate concerning the relation between speech and song, but also allow for more general conclusions to be drawn concerning pitch realization.

1. Introduction

1.1. Language, music and Wiktor Jassem

It would be accurate, I think, to say that one of Wiktor Jassem's abiding interests has been the music found in language. Going back even, or perhaps especially, to his *Intonation of Conversational English*, which appeared in published form 60 years ago [1], this interest is apparent. In that volume and many other publications that followed, Jassem relied on concepts from music to both inform and convey his understanding of intonation and rhythm. One thinks, for example, of his use of anacruses in describing the structure of English rhythm, and of course both tone and tune when discussing

intonation. In *Intonation of Conversational English*, Jassem seems to eschew the use of instrumental investigative techniques, and offers compelling arguments for use of the ear, still and always the phoneticians most important tool. He went on however, to make a substantial contribution in the area of acoustic phonetics as well as speech synthesis and technology. The present contribution investigates through instrumental means an area which – to my knowledge – Professor Jassem didn't manage to explore, though one suspects it would not have been for lack of interest.

1.1. Research perspectives on the relationship between language and music

There are at least three distinct research perspectives from which a possible relationship between language and music has been investigated. The most general of these has as its thesis the proposition that the two, language and music, are intimately related and in fact must share a common origin with respect to human cultural development. A second view, independent of the first, assumes music can be usefully analysed using linguistic models; Blacking [2] and Feld [3] provide critical views of the appropriateness and value of such work. A third approach examines the overlap or interface between music and language; the relation, for example, between song structure and poetry, on one hand, or the musical characteristics of spoken language on the other hand. While such overlapping might be based in a shared origin of music and language, this need not be the case. The present study follows the third approach without assumptions or claims as to a common origin of these two human facilities.

1.2. Speech and song melody

Tone languages are found in three major geographical areas, the Americas, East and Southeast Asia and Africa, and though research on the relation between language and music has involved languages from each of these areas [4, 5, 6, 7, 8, 9], the total number of studies is relatively few. This is particularly true when one takes into account the range of tone systems found across these regions. Among studies that have been done, the main focus has been to investigate the relation between speech melody and song melody: e.g. to explore to what extent which, if either, dominates the other; typically either a very strong correlation or a weak to no correlation has been found to exist. Little work has been done to investigate pitch realization characteristics of lexical tones in song melody; four such studies are [8], [9], [10] and [11].

Studies investigating the correspondence between speech melody and song melody in tone languages have employed different methods, but their essence has always been to determine whether pitch movement in the song melody correlates directionally with that of the corresponding spoken text; e.g. to the extent that a H-H-L-L-H sequence in a spoken text is realized with a similar F0 contour, and matched syllable by syllable in the sung version, then to that extent correspondence is said to exist. Typically it is phonemic tone that has been considered, abstracted away from various possible coarticulatory or downtrend effects. Thus, one of the few studies that has taken such factors into account [8], was able to show that the previously reported tone–tune correspondence of 53.4% in a Hausa song [13], was in fact closer to 100%. For studies that have found a strong correspondence between spoken and sung texts,, the claim has been advanced that tune is determined by tone, and for those with a weak

correspondence, the converse has been posited; the realization of lexical tone is influenced by song melody.

Blacking [5], in studying Venda, found in most genres of Venda song a near perfect correlation between the two text types, sufficient to assert "song differs from speech chiefly because its words are arranged in a strict metrical framework … Patterns of speech tone affect melodies not merely for the sake of intelligibility, but because the Venda seem to feel an aesthetic impulse to pay attention to speech-tone" (p. 166). In this regard, Blacking draws comparison with composers working in Western music traditions who draw inspiration for their melodies from those of speech: one may cite examples ranging from Elgar to Schönberg and from Stravinsky to Wagner; or even Spanish folk melodies.

Blacking initially gives the impression that song melody in Venda follows very closely that of speech, but this is not true of every musical genre, as he later cites at least one in which speech melody is ignored entirely, the song melody in this case being a chant.

A simple match up of Hs and Ls, however, is at best a basic measure only of the influence tone might (or might not) have on tune, and vice versa. A more sophisticated examination would take into account the various factors mentioned above, such as overall pitch range and tone spacing.

1.3. General hypotheses

Three basic hypotheses can be proposed, each of which are almost certainly too simple; but they do provide initial direction for non-musicological investigation. First, and this appears to be the prevailing view, song melody is constrained by the need to maintain the lexical meaning of the lyrics. However given that all music melody is built around notions of scale and harmonic progression, a second hypothesis would be that lexical tones in song are constrained by the melody of the song; they are realized to fit the scale, its intervals and the harmonic progression used in the music of the culture in question. A third hypothesis combines these two: song melody in a tone language will conform to the demands of both the linguistic systems and the musical systems of the language.

2. Mambila background

Mambila is a Bantoid language spoken in the Nigeria–Cameroon borderland. It has a complex dialect situation, with some 20 recognizably different lects grouped into two dialect clusters, East and West Mambila. Mutual intelligibility doesn't exist between clusters and is restricted within clusters. One of the differences distinguishing the two groupings is tone inventory; EM has three level tones and WM four. The lect which has been subject to the greatest amount of investigation with respect to tone is Ba-Mambila, a WM lect. Ba is also the object of study in the work reported here.

2.1. Pitch realization in Mambila

The tone system of Ba Mambila is relatively uncommon: it has four level lexical tones, which also combine to form several contours; tone is also used to encode several grammatical and discourse functions, the result being that the tone space of Mambila is

unusually 'crowded'. (In what follows, the four tones are numbered T1–T4, from high to low. In transcription High (T1) is marked with an acute accent, Upper Mid (T2) with a macron, Lower Mid (T3) is unmarked and Low (T4) with a grave accent.) Mambila tone, therefore is of interest from several perspectives: phonetic, phonological, morphsyntactic and, given its genetic affinities, its historical development. Several experimental investigations into pitch realization in Mambila have been reported [14, 15, 16, 17, 18, 19]. The principal findings of these studies are as follows:

- downstep is not found in the language [14, 19]
- tone coarticulation, e.g. higher tones following lower tones are lowered, exists and is corrected within one or two syllables [14, 19];
- declination and final lowering only occur consistently with the lowest tone [16]
- IF0 is nullified [15];
- utterance length seems to play no role in the scaling of utterance initial/final tones [17];
- register manipulation, e.g. expansion/raising for polar questions, does not occur [17, 18]

The primary conclusion to be drawn from these findings, which are consistent in the direction of constraining pitch realization, is that F0 values for individual tones in Mambila are fairly narrowly specified, in order to maintain phonological tone contrasts. An obvious question regarding song in Mambila, then, is: is pitch realization similarly constrained in Mambila song? It is of interest to see to what extent tones are variable in Mambila song melody.

2.2. Mambila Drinking Songs

There have been no enthnomusicological studies done of Mambila music, a lacuna which obviously constrains the present study. Only one of several Mambila song genres is explored here, drinking songs. Other genres include laments, nursery songs, and religious songs, and the latter includes many songs which are borrowed from European culture. The drinking song was chosen for this work because it is an indigenous form of music, on one hand, and on the other because of its particular characteristics. Mambila drinking songs are normally sung *a cappella* and solo and while drinking, for example during a *tantine*, and typically after several drinking bowls of mbè, the local beer made from maize. Two or more people may sing together, or alternating phrases. The structure of the songs is flexible; words may vary slightly, and phrases or lines may be repeated several times before moving on to the next. The songs also appear to be generation specific; those sung by the speaker who participated in the work reported here were not known other speakers who participated in the research, though they did recognize his songs as coming from an earlier time or generation. Such songs reportedly may also be borrowed in whole or in part from village to village. They are often, perhaps usually, about specific people; Elizabet, in the song examined here, is the mother of a neighbour of the singer. The song was composed as an attempt to improve her spirits following the death of her husband.

3. Experimental work

3.1. Method

A comparison of tones in song and speech was done from two different perspectives. First, an examination of the two text types to ascertain to what extent tones in the song match those in the spoken version of the text. This followed, in broad outline if not every detail, the procedure used in such studies as [5, 8, 9]. Second, a series of measurements were done on the extracted F0 from comparable portions of the two text types; these include measurements of overall pitch range, average F0, F0 values for individual tones, and differences between adjacent tones. Results from the spoken text were compared with the sung version and with results of measurements done on this speaker's productions in previous work.[1]

3.1.1 The singer

The subject for this study, SM, is a native speaker of Ba-Mambila; he is also fluent in three other languages, Kwanja, Fulfulde, and French. He was in his early 50s at the time the recordings were made, and has no known speech problems or abnormalities. He has participated in a number of other studies on Mambila tone conducted by the present author. Pitch realization characteristics of this speaker from the earlier studies are presented below. Two other speakers, one male and one female, also participated in the study, though singing songs of different genres; results from these speakers will be reported at a later date.

3.1.2. The song

At least three repetitions of each of three different Mambila drinking songs of this speaker were recorded. The lyrics of the songs were then written out and recorded as spoken text. Analysis of only one song is presented here. As mentioned above, Mambila drinking songs are flexible; words may vary slightly, and phrases or lines may be repeated several times before moving on to the next. (An analogy might be drawn with the jazz singing of Van Morrison) In the rendition of the particular song examined here, lines were repeated, parts of lines sometimes spliced together, and words either deliberately omitted or absent through syncopation. As a result, the sung version varies in places from the spoken version.

3.1.3. The recording

Recordings were made in the village of Somié, Adamawa Province, Cameroon, using a stereo digital recorder (Sony TCD D-100) and an Audiotechnica Pro-80 headset microphone. The audio signal was collected in one channel, while an electroglottograph signal (Laryngograph Field Unit) was fed into the other channel. Pitch extraction and measurements were done using the Macquirer signal processing package (SciconRD). It should be remarked that the speaker was sober for the recordings; this bears mentioning since the usual situation for singing such songs is in a state of (increasing) inebriation. Consequently the situation was to some extent

---

[1] This latter comparison was done simply to ascertain that he was, in terms of pitch realization, behaving consistently. There were no significant differences between his recordings on this occasion and the previous ones.

artificial, and SM was apparently a bit reticent; it is not known to what extent his singing would have been different had he been in circumstances usual to the song.

4. Results

Due to the quasi-improvised nature of the singing, comparison of the two text types for tone correspondences was not as straightforward as reported in the studies referred to earlier. First, as mentioned, there were a number of places where the spoken and sung texts were simply not comparable.
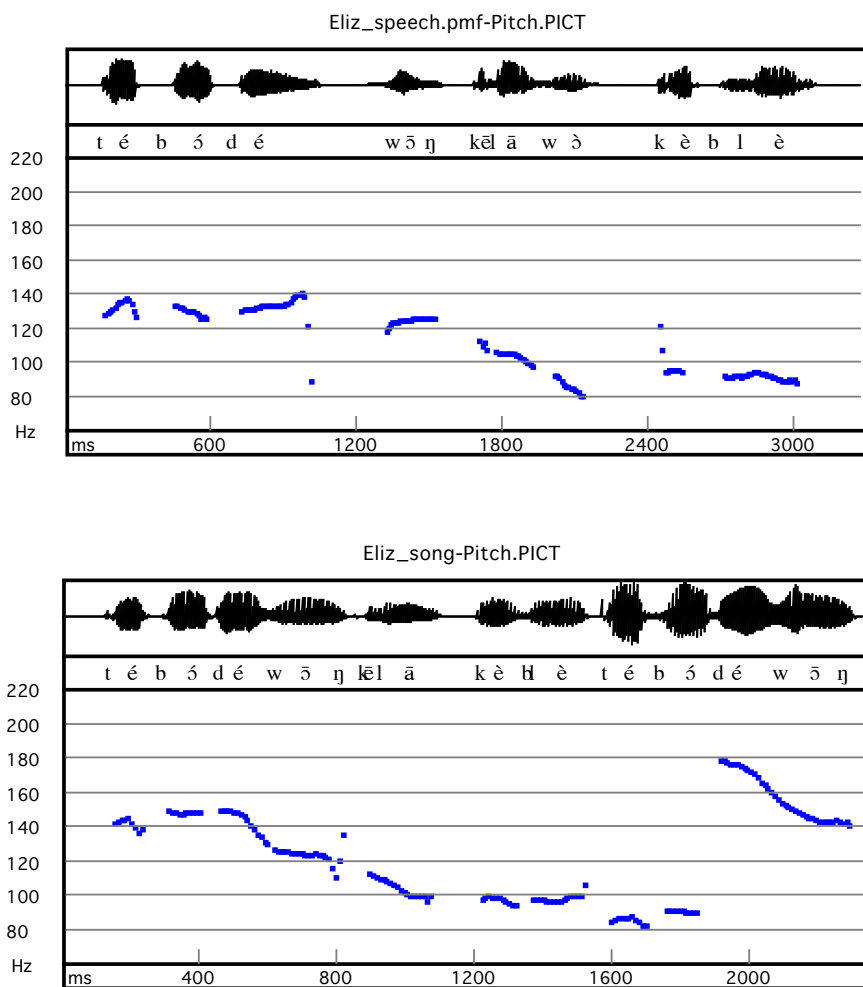

Figure 1. *Pitch traces of a line from a Mambila drinking song; the upper panel reflects spoken text, the lower panel sung text. The lower panel shows a partial repetition of the line (speaker/singer, SM).*

In addition, in some repetitions of lines the match up gave a 100% correspondence, while in other repetitions of the same line, the match was less than perfect. On a few occasions, it was the first repetition of a line that didn't correspond, and one gained the

impression that the line was repeated in order to get it right. On other occasions when tones in the song melody deviated from their spoken counterparts, it appeared the singer was perhaps 'stretching out'. These deviations aside, for at least one repetition of each line, there was a perfect or near perfect correlation of tones with the spoken text. Figure 1 shows pitch tracings for a single line from the two text types and illustrates the strength of the correspondence. The lower panel shows the sung version, and carries through to the first part of a second repetition of this line. In this second repetition, the first two syllables, which carry high tones (T1) are realized on a pitch lower than the preceding Ls. Jones [12] observes a similar occurrence in his study of Ewe, and suggests the singer to be following a musical tendency; for the singer this low pitch is underlyingly equivalent to T1.

4.2. Pitch range comparisons

4.2.1. Overall pitch range

First impressions suggest a greatly expanded pitch range is available to this speaker when singing. Further comparison reveals that the average overall range used is virtually identical in the two text types. Values for the spoken text are given in Table 1 and those for the sung text in Table 2.

Table 1. *Pitch values in Hz for tones in spoken text, with range of each tone given in semitones.*

Tone	Max	Min	Mean (SD)	N	Range (ST)
1	136	107	117.9 (7.2)	23	4.1
2	124	92	103.6 (6.7)	30	5.2
3	100	94	96.4 (2.3)	5	1.1
4	107	78	90.8 (6.7)	40	5.5

Table 2. *Pitch values in Hz for tones in sung text, with range of each tone given in semitones.*

Tone	Max	Min	Mean (SD)	N	Range (ST)
1	180	118	145.6 (15.5)	115	7.3
2	153	95	122.2 (14.9)	99	8.3
4	140	82	112.7 (12.3)	99	9.3

The overall range is larger for sung text, both in absolute terms (Hz) and in semitones: spoken text: 9.6 semitones, sung text: 13.8 semitones. But, the averaged range in semitones across the two text types is virtually identical (see below, Table 3), with spoken text at 4.5 semitones and the sung text: at 4.4 semitones.

4.2.2. Differences between tones

The differences between speech and song for adjacent tones is similar in both modes, but is expanded in song for T1–T2 and contracted for T2–T4, as shown in Table 3.

Table 3: Differences between mean tone values in spoken and sung text in semitones.

	Speech	*Song*
Tonal Step	Δ ST	Δ ST
T1–T2	2.2	3.0
T1–T3	3.5	
T1–T4	4.5	4.4
T2–T3	1.2	
T2–T4	2.3	1.4
T3–T4	1.0	

It is not clear how T3 would fit in, given the narrow difference between T2 and T4. T3 is the least frequently occurring tone in Mambila, and often seems differentiated from T4 simply by the latter having a fall. (NB, some T3 tokens present in the spoken version dropped out via changes in the sung text.)

5. Discussion

5.1. Pitch realization in speech vs song

While investigations into pitch realization in Mambila speech indicate a system that is fairly constrained, greater overall latitude for pitch is apparent in song. The key difference between the two, speech and song, is that a greater overall pitch range available in song; this is achieved through raising the upper limit of the range. Speech realized in lower portion of available range; the range expands and shifts upward slightly in song. The overall relation between tones, however, remains largely the same, though status of tone 3 unclear. There is little or no evidence of declination in the song melody, i.e. declination is similar across the two genres, though tone 2 is somewhat anomalous. Finally, and impressionistically, tone coarticulation effects appear similar in the two genres.

So, it can be seen that both speech and music contribute to song melody [4]. Since the relation between tones is effectively maintained, it is clear the importance of maintaining lexical contrasts is an important factor in determining song melody. Nevertheless, musical factors also play a role, for example in the initial low realisations of T1. More generally song, at least this genre in Mambila, is like another speech mode; it is comparable to speech; the difference in effect between speech and song here appears similar to that found in experiments where speaking volume is adjusted [20].

5.2. Reservations and further directions

It should be emphasized, in closing, that the results presented here are from a restricted investigation, based on only one speaker and one song, of one genre. Further work both with this speaker and others will reveal the validity of these results. Additional work investigating other music genres in Mambila should also prove revealing, and it may be noted that impressionistic analysis of a Christian hymn shows

the opposite effect: lexical tone is subservient to the melody of the hymn. The key to understanding the difference between the two genres may lie in the fact that one, that studied here, is indigenous, while the other is borrowed.

Finally, one factor that calls for caution in the present work is the uncertain influence of grammatical tone. Some of the anomalies found may in fact be due to incomplete accounting of grammatical tone, and a more detailed study would be able to take this into account. for example some phrase could apparently be interpreted in either of two.

References

[1] Jassem, Wiktor (1952). The Intonation of Conversational English (Educated southern British). Wroclaw: Nakladem Wroclawskiego Towarzystewa Naukowego.

[2] Blacking, John (1973) *How musical is man? (Venda Music)*. Seattle: University of Washington Press.

[3] Feld, Steven. (1974) Linguistic models in ethnomusicology. *Ethnomusicology*. 18.2: 197–217.

[4] Agawu, V. Kofi (1988) Tone and tune: the evidence for Northern Ewe music. *Africa 58*.2: 127–146.

[5] Blacking, John (1967) *Venda children's songs: a study in ethnomusicological analysis*. Johannesburg: Witwatersrand University Press.

[6] Chao, Yuen Ren (1956). Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese. In Halle, Morris, H. G. Lunt, H. McLean & C. H. Van Schooneveld (eds.) *For Roman Jakobson: Essays on the occasion of his sixtieth birthday 11th October 1956*. The Hague: Mouton, 52–59.

[7] Herzog, George (1934). Speech melody and primitive music. Musical Quarterly 20, 452–466.

[8] Leben, William (1985) Correspondence between linguistic tone and musical melody. In Goyvaerts, Didier (Ed.). *African Linguistics: Essays in memory of M.W.K. Semikenke*. Amsterdam: John Benjamins. Pp. 335–343.

[9] List, George (1961). Speech melody and song melody in Central Thailand. *Ethnomusicology 5*.1: 16–32.

[10] Wong, Patrick C. M. & Randy L. Diehl (2002). How can the lyrics of a song in a tone language be understood? *Psychology of Music 30*, 202–209.

[11] Gibbon, Dafydd, Firmin Ahoua & Adjépolé Kouamé (2011). Modelling speech-song relations: an exploratory study of pitch contours, tones and prosodic domains in Anyi. *Proceedings ICPhS 2011*. Hongkong.

[12] Jones, A. M. (1959) *Studies in African Music*. London: Oxford University Press.

[13] Richards, P. (1972) A quantitative analysis of the relationship between language tone and melody in a Hausa song. African Language Studies XIII: 137-161.

[14] Connell, B. (2002). Downdrift, downstep and declination. In Gut, Ulrike & Dafydd Gibbon (Eds.) *Typology of African Prosodic Systems*. Bielefeld: University of Bielefeld, 3–12.

[15] Connell, Bruce (2002) Tone languages and the universality of IF0. *Journal of Phonetics, 30*: 101–129.

[16] Connell, Bruce (2003) Pitch realization and the four tones of Mambila. In Kaji, Shigeki (Ed.) *Cross-Linguistic Studies of Tonal Phenomena: Historical development, phonetics of tone, and descriptive studies*. Tokyo: Research Institute for the Language and Cultures of Asia and Africa. Pp. 181–197.

[17] Connell, Bruce (2004) Tone, utterance length and F0 Scaling. In Bel, Bernard and Marlien, Isabelle, (Eds.), *International Symposium on Tonal Aspects of Languages: Emphasis on tone languages. Proceedings*. Beijing: Institute of Linguistics, Chinese Academy of Social Sciences. Pp. 41-43.

[18] Connell, Bruce (2005) Pitch realization of questions and statements in Mambila. In A. Akinlabi (Ed.), *Proceedings of WOCAL3*. Köln: Köppe.

[19] Connell, B. (2011). Downstep. In Van Oostendorp, Mark, Colin J. Ewen, Elizabeth Hume & Keren Rice (Eds.) *The Blackwell Companion to Phonology*. Malden, MA & Oxford: Wiley-Blackwell, 824–847.

[20] Liberman, Mark, Michael J. Schultz, Soonhyun Hong, Vincent Okeke (1993) The phonetic interpretation of tone in Igbo. *Phonetica 50*: 147–160.

Appendix: Elizabet, transcription

The six lines of the song are transcribed below. For each, the top row provides a broad transcription, including lexical tone, the second row give the tone number and the third row the gloss. As described above in §3.1.2, there was on occasion some variation between spoken and sung forms; here below, line 3 reflects one such variation that occurred in the spoken and sung versions used for analysis, where 3a is the spoken text and 3b the sung text.

1	èlízàbét	té	bɔ́	dé	wɔ̄ŋ	kēlā	wɔ̀	kèp	lè
	4141	1	1	1	2	2-2	4	4	4
	Elizabet	*NEG*	*do*	*rest*	*world*	*pass*	*2S*	*side*	*in*

2	wɔ̀	lé	jògò	nāá̄	hî	èlízàbét	ma
	4	1	4-4	212	14	4141	3
	2S	*PAST*	*before*	*PAST*	*where*	*Elizabet*	*EMP*

3a	mî	dere	môn
	1	3	14
	1S	*stay*	*thus*

3b	ndē	ndé	rē	môn	ā̄
	2	1	2	14	24
	FUT	*go*	*EMP*	*thus*	

4	té	bɔ́	dé	bɔ́	hāá	wɔ̀	sùàgà	nāgā	wā
	1	1	1	1	21	4	4-4	2-2	2
	NEG	*3P*	*POSS*	*3P*	*give*	*2S*	*sua*	*oath/lick*	*Q*

5	wɔ̀	jògò	nāá̄	rē	hî	wā	mbɪ́	yɛ̄

	4	44	212	2	14	24	21	24
	2S	before	be PAST	EMP	where	Q	friend	2SPOSS
6	kè	mé	njōlō	té	mé	bē	kèmà	wɔ́yè
	4	1	22	1	1	2	44	14
	look	with	eye	NEG	with	hand	touch	--