Tonal Aspects of Languages 2016
24-27 May 2016, Buffalo, New York

Methods of analysis for tonal text-setting. The case study of Fe'Fe' Bamileke

Teresa Proto

Leiden University and Meertens Institute, the Netherlands

t.proto@hum.leidenuniv.nl

Abstract

This paper focuses on the interaction of speech tones and musical pitches in Fe'Fe' Bamileke, a Grassfields Bantu language of Cameroon. Like earlier studies of tonal text-setting, it aims to test the hypothesis that, in song, the melodic contour of tunes matches the tonal contour of the lyrics. The study is based on four songs and was carried out with the help of a native speaker who provided the songs and accurate feedback on the musical settings. His acceptability judgments of musical settings in song #4 were integrated into the study to complement the evidence from song data. Most of the paper is a detailed discussion of the findings, and the methodology employed for measuring the degree of correspondence between speech and song. Some issues with a bearing on the interpretation of the data are addressed, and alternative analyses are proposed. It is argued that all of the suggested analyses have drawbacks and, as revealed through a comparison of the different sets of results, need to be supplemented with qualitative approaches that evaluate the interaction between the phonological and musical systems involved.

Index Terms: speech tone, musical pitch, Fe'Fe' Bamileke, song, tone-melody correspondence

1. Introduction

The expression 'text-setting' denotes the process of aligning textual units to musical units in creating and/or performing vocal music. Native speakers (who are active participants in their singing traditions) have clear intuitions as to how to map a text onto a tune, despite their lack of explicit training. A long-standing question about text-setting in tone languages is: how can speakers of tone languages understand lyrics when they are set to music, if pitch is used to also distinguish meaning? It has been argued that if a word is to be grammatically intelligible, language tone and musical pitch must be somehow correlated ([1]). This idea has been supported or challenged in a number of studies, sometimes yielding conflicting results (e.g. for Ewe songs, [1] found direct correlation, whereas [2] claimed the opposite). In this paper, I will test this hypothesis on a small sample of songs in Fe'Fe' Bamileke, a Bantu language of Cameroon. The methodology employed in the first approach to the corpus was borrowed from [3]; subsequently, the analysis was expanded to introduce qualitative evaluation of tone-tune parallelism.

2. Tones and songs in Fe'Fe'

This work was carried out with the help of a native speaker of Fe'Fe' Bamileke, who has a lifelong experience with the music of his country and happens to be the author of one of the examined songs (song #1). His role mainly consisted in selecting the songs from [4], in annotating the spoken tones of the lyrics, and in providing feedback on the tone-pitch associations found in the songs.

2.1. Tonal inventory

Fe'Fe' Bamileke ([5], [6], [7]) is a Grassfields Bantu language (Bantoid, Benue-Congo, Niger-Congo). Besides being spoken in Western Cameroon, it is taught in schools and seminaries outside the Bafang area, and is especially used as a vehicular language in the religious sphere. This is relevant to our purposes as the songs under examination are church songs.

In citation, Fe'Fe' has four discrete level tones and four contour tones ([7]). The level tones are called Low, Raised-Low, Mid and High; they will be referred to with the following symbols: L, ↑L, M, H, respectively. It should be noted that in isolation very few monosyllabic morphemes carry H tone; moreover, the contrast between L and ↑L is found only before pause [Hyman 1993:84]. Regarding the contour tones, two have an ascending contour (↑LM, ↑LH) and two have a descending contour (ML and HL). ↑LM characterizes nouns, whereas ↑LH is found on adjectives and some grammatical morphemes. ML occurs on a few lexical morphemes derived from verbs (such as 'Mboo', the word for 'God'), while HL characterizes a few grammatical morphemes. Contour tones are less frequent than level tones in the language and, as we will see, they also play a secondary role in this analysis of songs.

In addition to these eight phonetic tones, Fe'Fe' also has floating tones, i.e. tones that are not attached to segmental material, which occur for example in associative constructions, between the head noun and the possessive pronoun.

2.2. Corpus of songs

This study is based on four songs selected from a songbook. My informant was recorded both singing the songs and reading the lyrics (without singing). The resulting audio recordings were transferred to PRAAT, segmented and annotated. For each tune, a music score was created, and for each spoken text the surface tones were manually annotated.

The examined songs fall into two sets: the first three are original Fe'Fe' songs; the fourth is a *contrafactum* based on a pre-existing French tune to which new lyrics in the vernacular have been set. The song in question is the Christmas carol *Les anges dans les campagnes* and it is known as *Gloria* in the Fe'Fe' version.

The songs in the two sets also differ in their overall metrical structure. While the original songs are typically antiphonal, joining two voice parts (soloist and choir) in call and response, the *contrafactum* shows the stanza-refrain alternation commonly found in Western folk music. The refrain consists of the Latin verse 'Gloria in Excelsis Deo'. Each stanza can be analyzed as a quatrain, made up of two couplets, each consisting of a fixed number of syllables. Couplets are sung on the same, repeating melody, which is made up of two musical phrases (A and B) whose boundaries match the boundaries of the lines in the couplet. This is in striking contrast to the original songs, where the line length and the melody may vary significantly from one call to

 http://dx.doi.org/10.21437/TAL.2016-35

another; only the response part is stable across the song. An additional difference between the two groups concerns the type of musical setting they exhibit: while in the original songs each syllable is assigned to an individual note, thus resulting in a syllabic style, the *contrafactum* contains melismas.[i]

3. Methodology

For the analysis, I adopted the method employed in [3] for Shona, also a Bantu language. Within this framework of analysis, speech and song are characterized as melodies made up of sequences of transitions from one pitch to the next. Based on this assumption, it seems possible to determine the degree of correspondence between the two melodies by comparing the transitions of any spoken text with the corresponding sung rendition. A transition is defined as the pitch interval that occurs between a syllable and the next one, in both song and speech.

In order to annotate all the transitions, the audio recordings of the songs and spoken lyrics were transferred to PRAAT, the signals were segmented into syllables and for each syllable the mean F0 value was automatically calculated. After calculation of F0, the transitions from one frequency to the next were examined.[ii] From any given syllable of a text, the fundamental frequency can move in three possible directions, namely UP, DOWN or remain the SAME.[iii] Each transition in both spoken and sung text was thus coded for directionality. Then, the direction of each spoken transition was compared with the corresponding sung transition.[iv] If the direction was the same in both speaking and singing, the transitions were coded as PARALLEL; if one transition moved up and the other moved down, they were coded as OPPOSING; if the movement was not in opposite directions (e.g., if the transition went down in the language but remained the same in singing), the transitions were coded as NON-OPPOSING.

σ	F0 sung	Dir.	F0 spoken	Dir.	Compared
Pαh	192.32		120.54		
sα'	151.9	down	99.49	down	P
mα	170.22	up	95.36	down	O
nshi-	190.37	up	104.31	up	P
nαh	191.77	same	116.93	up	N
mbhi^o	172.94	down	102.53	down	P
Shʉa-	152.2	down	91.72	down	P
pʉ	155.66	up	80.04	down	O

Table 1. *Melody comparison for line 1 of song #1. σ = syllable; Dir. = direction of pitch movement; P = Parallel; O = Opposing; N = Non-opposing*

Table 1 reports the analysis of the transitions in line 1 of song #1. In this specific case, there is a high degree of correspondence between song and speech (4 transitions out of 7 are PARALLEL). However, as we will see in more details in the next section, this is not true for all lines in all four songs. Once again, the *contrafactum* stands out from the original songs in that it displays a pattern of its own.

3.1. Results

Excluding transitions separated by major pauses (notably at the end of lines or half-lines) and transitions involving melismas, the results of the analysis are summarized below (number of occurrences/percentage):[v]

Song	TT	P	O	N
Song #1	72	48/66.6	14/19.4	10/13.8
Song #2	103	68/66.0	24/23.3	11/10.67
Song #3	135	96/71.1	14/10.3	25/18.5
Song #4	168	81/48.2	47/27.9	40/23.8

Table 2. *Transition breakdown for all songs. TT = Total number of Transitions; P = Parallel; O = Opposing; N = Non-opposing*

A high proportion of PARALLEL transitions is found in the first three songs and especially in song #3 (71.1%) and #1 (66.6%); in song #4, however, the percentage rating drops dramatically to 48.2%. The *contrafactum* also contains more OPPOSING transitions (27.9%) than the other three songs averaged (≈18%). The percentage of NON-OPPOSING transitions is lowest in song #2 (10.67%) and highest in song #3 (18.5%). Visual examination of the melody comparison for song #4 (cf. Fig. 1) does not show the similarity of shape that is exhibited by the other songs.

The *contrafactum* was singled out by my informant as an example of "bad" text-setting. Internal and external evidence suggests that the lyricist was not a native speaker. Unlike all the other songs in the songbook, in fact, the text of the *contrafactum* is anonymous; furthermore syntactic peculiarities, such as the aberrant syntactic order in the sequence 'nkwe yoh' (trans. *whole our,*) and the lack of parallelism between spoken and sung melody are evidence in favour of this hypothesis.

3.2. Notes on the methodology

In this section, I wish to comment on the methodology that yielded the results in Table 2. In particular, I will focus on the issue of setting parameters that capture significant movements in F0 and will propose alternative analyses of the transitions in my corpus. Consider, for example, the analysis of the first line of song #1 as reported in Table 1. The last transition in the sung melody (involving the syllables *Shʉα-* and *-pʉ*) was coded as going UP on the basis of the parameters set for the analysis (mean F0 values for these syllables were 152.20Hz and 155.66Hz, respectively).

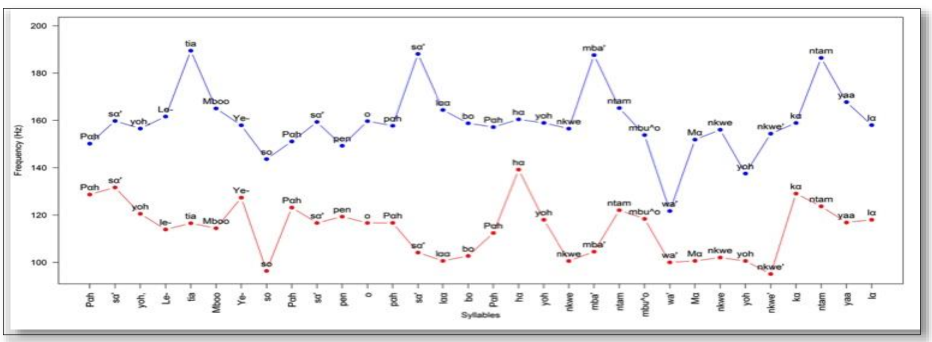

Figure 1: *Melody comparison for stanza 7 of song #4. Upper line = sung melody; lower line = spoken melody*

However, when transcribed in music notation (Fig. 2), the two syllables are assigned the same musical pitch, namely E flat. Based on this evidence, the sung transition in question should therefore be coded as being the SAME.

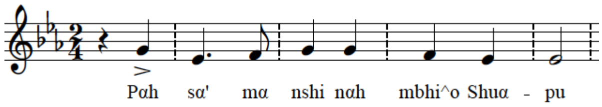

Figure 2: Music score for line 1 of song #1

Consider again the line in Table 1. The second and third syllables (*sɑ′* and *mɑ*) both carry a L tone when pronounced in isolation. However, in speech, the second tone is realized at a lower pitch than the first. In Fe'Fe', like other Grassfields languages, downstepping low tones are responsible for register shifts. Based on the frequency values assigned to each syllable in the spoken rendition (99.49Hz and 95.36Hz, respectively), the transition from *sɑ′* to *mɑ* would be coded as going DOWN. The corresponding transition in the sung melody is coded as going UP. In the comparison, this would count as an OPPOSING transition. However, if we had used the underlying tones instead of the surface tones as a measure, the transition would count as NON-OPPOSING. In the same line, conflicting results would be found for the last transition. The recorded mean F0 for the syllables in *Shuɑpu* are 91.72Hz and 80.04Hz, respectively, although underlyingly the word has a LL pattern. Comparing the direction of pitch in the spoken and sung renditions of *Shuɑpu*, we would get an OPPOSING transition on the basis of the frequency values; however the same transition would count as PARALLEL if we relied on the abstract, underlying categories (tone and musical pitch) alone. If, furthermore, we compared the frequency value of the spoken transition with the (abstract) musical pitch value of the sung transition, we would obtain yet another result, namely a NON-OPPOSING transition. This divergence in the potential results points to an intrinsic weakness of the methodology applied here. In fact, if the acoustic measure of the fundamental frequency fails to capture significant movements in F0, namely pitch changes in the sung melody, this may result in veritable 'errors' in the assessment of the transitions, as in the case of *Shuɑpu*. In a number of cases, it would predict parallel/opposing movement where *underlyingly* the movement is in NON-OPPOSING directions.

In the table below, I report the results we would have obtained from the analysis of each song if we had used musical pitch values (instead of frequency values) to code the direction of the sung transitions.

Analysis 2	TT	P	O	N
Song #1	72	41/56.9	7/9.7	24/33.3
Song #2	103	55/53.3	19/18.4	29/28,1
Song #3	135	90/66.6	12/8.8	33/24.4
Song #4	168	64/38.0	32/19.0	72/42.8

Table 3. *Results of Analysis 2*

For songs #1-3, the values in the PARALLEL column are close to Table 2. Unlike in Table 2, however, NON-OPPOSING transitions consistently outnumber the OPPOSING ones (average score ≈28%). In song #4, NON-OPPOSING transitions also outnumber the PARALLEL ones. In most stanzas, in fact, PARALLEL movement is found in a smaller number of cases than NON-PARALLEL movement.

Most of the research conducted on tonal text-setting is based on tone (and pitch) categories rather than acoustic measures ([8], [9], [10]). In those studies, the direction of pitch change between adjacent syllables was tracked and compared on the basis of underlying tones (in speech) and musical pitches (in song). In Table 4, I report the results we would obtain for the transitions in my corpus using this approach.

Analysis 3	TT	P	O	N	X
Song #1	72	56/82.3	0	12/17.6	4
Song #2	103	55/60.4	12/13.1	25/27.4	12
Song #3	135	105/84.0	4/3.2	16/12.8	10
Song #4	168	53/37.8	27/19.2	60/42.8	28

Table 4. *Analysis 3. X = unclassified tokens*

The first notable difference concerns the presence of unclassified tokens (X column). These are transitions whose direction could not be determined because they involved contour tones. In analyses of tonal text-setting in Cantonese ([8]) and Mandarin ([10]), contour tones were classified according to the ending or starting pitch level, respectively. Since either classification would have been arbitrary if applied to Fe'Fe', transitions involving contour tones were excluded from the comparison.[vi]

Leaving aside these instances, the general trend is a consistent fall in the number of OPPOSING transitions, which turns out to be absolute for song #1 (0 occurrences). In consequence, PARALLEL (82.3%) and NON-OPPOSING (17.6%) transitions score higher than under Analysis 1 for Song #1. This difference is mainly due to the fact that in the first analysis the transitions in the response are attributed different frequency values across the (three) repetitions, and are thus coded differently, while under Analysis 3 all (three) repeats are coded in exactly the same way. For song #2, it can be seen that while the percentage of PARALLEL transitions is above 60%, like in Table 2, the ratings for the other two classes are closer to those in Table 3. At present, I have no explanation for this difference. In song #3 a noticeable increase of PARALLEL transitions is observed (84.0%), to the detriment of both other classes. The main source of this increase appears to be the coding of spoken sequences of two or more successive Ls. For instance, a word like *nusipe*, which carries a L tone on each syllable, gives rise to two DOWN-transitions under Analysis 1 and 2 (presumably due to the effect of downstep), but two SAME-transitions under Analysis 3. The corresponding transitions in the song melody are coded as being the SAME. In the melody comparison, this results in both transitions being coded as PARALLEL under Analysis 3 (but not in the other analyses). For song #4, the high proportion of unclassified tokens (28) makes it particularly difficult to evaluate the results.[vii] What is relevant is that, like in Table 3, NON-OPPOSING transitions (42.8%) outnumber PARALLEL transitions (37.8%), and percentually their score is comparatively much higher than in the other songs. Again, song #4 displays a pattern of its own.

At present, I have no explanation for the overlap between the results in Table 3 and 4, nor for the observed difference between the first three songs, on the one hand, and the *contrafactum* on the other. The point I wish to make here, however, is that fairly different (even conflicting) results are achieved under the three analyses, and that each approach has its pros and cons. What Analysis 1 captures best is the gradient, cumulative effects found in both speech and singing.

In fact, the tonal variability found in speech, due to phonetic and phonological processes such as automatic and non-automatic downstep, is taken into account in the coding of spoken transitions. On the other hand, however, in a number of cases frequency variations turn out not to be significant, i.e. those detected within the same musical intervals. This constitutes the most serious shortcoming of Analysis 1 and an important source of 'error'. A further source may lie upstream in the calculation of the syllables' fundamental frequency. In Fe'Fe', as in other Bantu languages, certain consonants tend to raise or lower the tone of the syllable. This is the case, for instance, with nasal prefixes, which carry an underlying low tone ([8:135]). Since consonant perturbations were not excluded from the measurement of the mean F0 value of the syllables, it may well be that the recorded values do not reflect the actual frequencies. It would therefore be advisable in future to use the mean frequency values of vowels instead of syllables.

Analysis 2 appears to provide stronger results, thanks to the correct coding of transitions found within the same musical intervals; however it is not free of drawbacks, namely as regards the tone perturbations mentioned for Analysis 1.

Finally, Analysis 3, while less exposed to errors, being based on abstract categories such as phonological tones and musical notes, fails to yield any satisfactory result, due to the high proportion of unclassified tokens. If this approach is used, it would be crucial to establish a criterion for the inclusion of transitions involving contour tones and melismas. A tentative proposal, for contour tones, consists in using highly recurrent words bearing a contour tone, such as *Mboo* (the word for 'God'), as a diagnostic to find out which pitch level (starting or ending) is relevant in spoken transitions. Alternatively, the choice of either classification could be substantiated through examination of the tone rules (particularly assimilation rules, such as absorption) that are applied to contour tones in the language; or on the basis of morphological categories, given that contours starting with a ↑L are observed in lexemes and contours ending with a L are observed in grammatical morphemes. Finding substantiation for the classification of contour tones in Fe'Fe' songs is the next step on the way to achieving a better understanding of how speech tones interact with musical pitches in this language.

4. Discussion

It seems clear from the discussion in 3.2 that fairly different results can be obtained from the same corpus of songs, depending on whether transitions are analyzed as purely acoustic objects, or whether they are also evaluated on the basis of the available knowledge of the phonological and musical systems involved. An additional source of knowledge is formed by the intuitions of native speakers about musical settings in singing idioms with which they are familiar. In the case of the *contrafactum* for instance, it was revealing to check native speakers' judgments of musical settings as a way of complementing the data analysis.[viii]

When first examining song #4, my informant claimed that the whole setting was ill-formed. However, when I asked him to rank each transition in the song on a five-point scale ranging from acceptable to unacceptable, he actually evaluated 103 transitions as fully acceptable.[ix] This accounts for 74% of the total number of transitions in the song (net of unclassified tokens), thus covering a much higher proportion than just the total number of PARALLEL transitions (under any of the analyses proposed above). This fact suggests, on the one hand, that the system internalized by the native speakers of Fe'Fe' does not require all transitions to be parallel. On the other hand, it implies that some kind of restriction must also operate on NON-OPPOSING transitions.

Looking at the transitions judged to be unacceptable, I noticed that in many cases they involved a sequence of non-L (H or M) + L associated with a level melody (i.e. two notes of the same pitch). One example is given in Fig. 3 where the adjacent syllables *Mα pαh* carry the tones H and L respectively, and are both set to the musical note E.

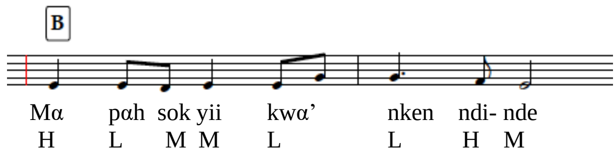

Figure 3: *Setting of line 2d, song #4*

Here, the transition counts as NON-OPPOSING under Analysis 2, and yet it is judged unacceptable. Assessing the validity of any conjecture about the setting of consecutive non-L and L tones in Fe'Fe' songs would require expanding the corpus and conducting more tests. This will be the subject of future research. The important point to underline here is that level musical melody is not compatible with *any* linguistic tone sequence. Furthermore, instances like Fig. 3 demonstrate that the ill-formedness of musical settings is not directly linked to the occurrence of OPPOSING transitions. In fact, transitions can be NON-OPPOSING and yet violate some kind of well-formedness constraint. We should therefore study the nature of NON-OPPOSING transitions in more depth in order to gain some insight into how text-setting works in this language. It might be the case that the inventory of NON-OPPOSING transitions allowed in songs is restricted to only a few types, or that these transitions are not freely distributed across the various metrical positions (e.g. at line beginning/end, on upbeats/downbeats etc.).

One potential explanation for the ill-formedness of the setting at the beginning of the line in Fig. 3 is found in a generalization postulated for Igbo songs. Regarding the shape of melodies in Igbo, [11: 339] states that it is "a terraced one, which, in general, starts at the highest point and gently works its way down to the lowest point". This tendency is impressionistically confirmed by observation of the outline of tunes #1-3. Significantly, this downtrend is not realized in either melodic phrase in the *contrafactum*. The same author further argues that when the initial syllable in a phrase bears a H tone, the melody starts with the highest note and "although the melody may at several points during the course of the song rise to this initial note, never it will go higher than this note within that musical phrase" ([11: 340]). This principle appears to be violated in Fig. 3, where the melody reaches its peak on the G at the end of the first measure, in correspondence with a L-toned syllable. In fact, not only is the L-toned syllable *pαh* at the beginning of the line assigned the same musical pitch as the preceding H-toned *Mα*, but the L-toned syllables *kwα'* and *nken* in the middle are also sung on the highest note of the musical phrase. This results in a failure to realize the downtrend described above. It is probably this failure that makes all the transitions in Fig. 3 (except the last one) range from unacceptable to hardly acceptable, and accounts for the first impression of complete ill-formedness elicited in my informant. An important consequence to be drawn in light of the methodology used in this study is that, in certain contexts, the pitch direction of individual transitions may play only a minor role: the well-formedness of the settings may depend less on the direction of each transition *per se* and more on some constraint placed on the distribution of high and low pitches across the line.

Other authors have underlined the importance of studying transitions in relation to their context. [12] has proposed that transitions should be studied in sequences of three, instead of two, syllables/notes (in his words, *trigrams* instead of *bigrams*), at least in some contexts. In fact, different kinds of three-note musical sequences may allow different tonal sequences. Applied to the first three syllables of the line in Fig. 3, this would mean that the musical sequence SAME-DOWN (E-E-D) would not allow the tonal sequence DOWN-UP (H-L-M). The source of the unacceptability of the first transition in Fig. 3 would actually lie in the next transition. This interpretation is sustained by the observation that in songs #1-3, sequences of H-L are indeed found on level musical melodies (although elsewhere within the line).

Finally, it should be mentioned that for some melodic movements it is unclear whether they are musically or linguistically motivated. [5] noted that when singing plainchant (= on a Latin text), the Bamileke tend to modify the melody locally. More particularly, if they come across a sequence of two major seconds, they tend to change the second interval into a minor third. [5] relates this idiosyncrasy to the fact that the minor third is the most common interval found in the language, both in ascending and descending order, being roughly the one that occurs between ↑L and M. This is clearly an area of study that deserves caution and needs to be researched carefully.

5. Conclusions

This study is a preliminary examination of the correspondence between speech melody and sung melody in Fe’Fe’ songs. Three different analyses were proposed, based on the same corpus of four songs. All of the analyses revealed the following pattern: songs #1-3 cluster together showing a preference for the melodic contour of tunes to match the tonal contour of the corresponding lyrics, whereas in song #4, which is a *contrafactum* based on a French tune, the tone-tune parallelism is less strong.

Most of the paper is devoted to a discussion of the criteria and methods employed to measure the degree of correspondence between transitions in singing and speech. The remainder is dedicated to showing that the system of rules underlying text-setting in Fe’Fe’ cannot be elucidated satisfactorily through a simple quantitative analysis of transitions. It is necessary to formulate generalizations that capture qualitative differences within and across the sets of (parallel/non-parallel) transitions.

6. Acknowledgments

This research was initially supported by grants from the Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS) and the EURIAS Fellowship Program. It was conducted within the framework of the project “Knowledge and Culture”, funded by the Netherlands Organization for Scientific Research (NWO). I am extremely grateful to EMM for his valuable collaboration, and to Paolo Bravi for providing useful PRAAT scripts.

7. References

[1] Schneider, M. “Tone and tune in West African music”, Ethnomusicology, 5(3): 204-215, 1961.

[2] Agawu, V. K., “Tone and Tune: The Evidence for Northern Ewe Music”, Journal of the International African Institute, 58:127-146, 1988.

[3] Schellenberg, M., “Singing in a Tone Language: Shona”, Select. Proc. 39th Annu. Conf.. African Linguistics, 137–144, Somerville, 2009.

[4] A typewritten book in use at the diocese of Nkongsamba and containing (only) the texts of the songs.

[5] Stoll, R. P., *Tonétique des langues bantu et semi-bantu du Cameroun*, Yaounde, 1955.

[6] Ngangoum, F. B., *Le bamiléké des fe'fe' : grammaire descriptive usuelle*, Abbaye Sainte-Marie de la Pierre-Qui-Vire, 1970.

[7] Hyman, L. M., *A Phonological Study of Fe'Fe'-Bamileke*, PhD. dissertation, Dept. Ling., UCLA, Los Angeles, 1972.

[8] Wong, P. C. M. and Diehl, R. L., “How can the lyrics of a song in a tone language be understood?”, Psychology of Music, 30(2): 202-209, 2002.

[9] Ladd, D. R., “Singing in tone languages: An introduction to the kinds of things we might expect to find”, Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften, 4:13–28, 2014.

[10] Wee, L. H., “Unraveling the relation between Mandarin tones and musical melody”, Journal of Chinese Linguistics, 35:128-143, 2007.

[11] Ekwueme, L. E. N., “Linguistic Determinants of Some Igbo Musical Properties”, Journal of African Studies, 3:335-353, 1974.

[12] Ladd Ladd, D. R., “Tone-melody text-setting constraints in tone language singing”. Presented at ‘Textsetting constraints in tone languages’, NIAS, Wassenaar, 2016.

[i] In song #4, a melisma always involves singing a single syllable of text while moving between two different notes in succession. Melismas occur in the same position as in the original French tune. The ornamentation occasionally found in songs #2 and #3 is of a different kind than the melismas in song #4.

[ii] For the spoken text, the transitions annotated in PRAAT were also compared with the manual annotations provided by the informant: discrepancies were checked and discussed.

[iii] Following [3], the transition was coded as being the SAME if the frequency value of the second syllable was within 1.5 Hz of the first syllable.

[iv] Only the direction of each transition was taken into account, not the magnitude of the F0 movement; a transition from L to M was therefore coded in the same way as one from L to H, despite the fact that the magnitude of the F0 rise might be very different.

[v] In the case of melismas, more than one transition occurs between the syllable involved in the melisma and the following one in the song melody. In [3], two transitions occurring on the same syllable in the song melody were counted as parallel if *either* transition matched the direction of the spoken melody. In the present study, this criterion was not adopted and the 28 such transitions found in song #4 were not included in the comparison.

[vi] Each occurrence of a contour tone resulted in two transitions not being classified (both the preceding and the following one). Under Analysis 1 and 2, syllables bearing contour tones are assigned a frequency value, in the same way as syllables bearing level tones.

[vii] Transitions involving simultaneously a contour tone and a melisma were counted (for exclusion) only once.

[viii] This approach has been employed successfully in studies of non-tonal text-setting.

[ix] The scale consists of: “unacceptable”, “between hardly acceptable and unacceptable”, “hardly acceptable”, “between acceptable and hardly acceptable”, “acceptable”.