Singing in a Tone Language: Shona

Murray Schellenberg
University of British Columbia

To what extent does speech melody influence the composition of song melody in tone languages? This is a question that has been raised periodically in both linguistics and musicology over the past century but has never been satisfactorily answered. This paper is a preliminary examination of vocal music from Shona that provides further evidence towards the solution of this puzzle through acoustic analysis of sung and spoken melodies of three songs in Shona, a Bantu language spoken in Zimbabwe. This examination shows that, in the songs examined, there is a statistically significant correspondence between the spoken melody and the sung melody but that the locus of parallelism is not predictable.

1. Background

A culture may be presumed to have three basic choices when it comes to representing speech tones in music: linguistic tone can be slavishly followed; it can be abandoned completely; or a position somewhere in the middle can be taken.

Research on the relationship between speech melody and sung melody has been carried out in a variety of languages over a long range of time: Cantonese (Wee, 2007, Yung, 1983), Mandarin (Wong and Diehl, 2002; Stock, 1999; Chao, 1956), Thai (List, 1961), Kalam Kohistani (Baart, 2004), Lushai (Bright, 1957), Navaho (Herzog, 1934), Hausa (Richards, 1972), Ewe (Agawu, 1988 & 1995; Schneider, 1961; Jones, 1959), Igbo (Ekueme, 1974) and Fanti (Ward, 1932). After comparing the melodic transitions in songs with the melodic transitions in the corresponding speech melody, the research has produced a variety of results and conclusions and finds supporters in all three camps.

1.1. Speech melody determines sung melody

There is a frequent assumption in much of the literature that singing and speech are very closely related in a tone language. The most common assumption is that there is a very strong correspondence between the tonal melody of spoken words and the musical melody of those words in song. Ekueme (1974, p. 187), for example, asserts that "the tonal contour of words determines the melodic contour of the music to which the words are sung". Similarly, Schneider (1961, p. 204) writes "if a word is to be grammatically intelligible, the individual syllables cannot be sung arbitrarily high or low. Speech tone and musical tone must be definitely correlated". Devine and Stephens (1994, p. 162) go so far as to say "in tone languages … the correlation of song melody with speech tone is so high that it can generally be expressed in terms of rules."

Richards (1972) conducted an extremely thorough analysis of an extended song in Hausa. Using a variety of statistical measures he found statistical significance in the parallelism between the spoken and sung melodies of the text.

Wong and Diehl (2002) found a very strong correspondence (91.81%) between musical and lexical melodies in an analysis of four contemporary Cantonese songs. There are also arguments for the influence of linguistic tone on sung melodies in Cantonese opera; Yung (1983) found speech tone to have a strong influence on the melodies produced by singers in this particular style of music that requires semi-improvisational performances.

1.2. Speech melody and sung melody are unrelated

Some researchers have argued against any connection between speech melody and sung melody. Ward (1932, p. 709) found that "in at least two well-known Fanti songs there are clear cases in which tune goes against the melody of the words" and, certainly, in the two excerpts he provides there is little correspondence. Bright (1957) provides analysis of two examples in Lushai, a Tibeto-Burman language found in Assam and finds "no particular agreement between them" (p. 27).

More recently, Kofi Agawu has argued against the correspondence of spoken and sung melody. Agawu (1988) presents evidence from eight Ewe songs showing that the pitch patterns of speech and singing do not match. He also points out that while Richards (1972) does find statistical significance in the parallelism in his analysis of the Hausa song,

> calculations for the repeated phrase given in example one show that of the twenty-five speech tones, fourteen correspond to musical direction while eleven do not; in other words 56 percent of the tune is 'correct' from the point of view of the words, while 44 per cent is 'wrong'. But that is a lot of notes to get wrong! (p. 131)

1.3. The middle ground

A third subset of researchers suggests that a culture may take a somewhat ambivalent approach to the connection between spoken melody and sung melody. Two different possibilities have been proposed. Chao (1956) posits a hierarchy of song types in Mandarin: singsong; chanting; recitative; tonal composition; and atonal composition. The singsong end of the hierarchy (children's rhymes; vendor's cries) has the greatest correspondence and atonal composition (his label for contemporary song composition – so called because the song writers "pay no attention to tone" [p. 58]) showing the least correspondence. Rycroft (1979) establishes a very similar hierarchy for Southern African music. At the "close correspondence" end of his scale he lists praise poetry and war chants. At the other end he lists modern church, school and popular music stating that "in all Southern African Black languages there is generally complete disregard for speech tones" for these types of music (p. 313). The idea of variation across song types is supported by List's (1961) study of eight Thai songs with children's recitations and traditional songs showing strong correspondence (79%, 90% and 100% for his three examples) and two popular songs showing the least (59% and 60%). Conversely, however, Herzog (1934) finds greater correspondence in his two examples of songs in Navaho than in his example of a chant.

Wee (2007) offers a different interpretation of the middle ground. He proposes that speech melody and musical melody correspond in positions of metrical prominence; that is, that syllables on the most prominent beats of a bar will match for lexical tone and melody (taken as a pitch relative to both preceding and following pitches). He supports this claim with an analysis of 10 randomly chosen folk songs in Mandarin; finding conformity to be 97.2% (649 syllables out of a total 668). Richards (1972) makes a somewhat similar observation in Hausa. He notes that much of the song is based on two "melodic sentences" (p. 140) which he categorizes as A and B. These two sentences appear with variations throughout the song. He does not make it clear in his transcription just what the relationship between these two musical phrases is but he shows that the vast majority of parallelism between lexical tone and song melody occurs at the beginning of the A-type melodic sentences. The general claim of this interpretation is that lexical tone is only marked at certain, strategic points in the music and the rest of the tones are assumed from context. It is interesting to note that, as early as 1932, Ward (1932, p. 710), while rejecting "the invariable agreement of tones and melody" suggests that "nothing is more likely than that there is some more subtle principle of coordination between words and music".

2. Methodology

One subject was recorded singing three different songs in Shona. The recordings were analysed acoustically and the results examined statistically.

The subject was a 28 year old, female, native-speaker of the Zezuru dialect of Shona (S12) living in Vancouver, BC. She had been living in Canada for approximately 3 years. Prior to that, she had been living in Zimbabwe. She still spoke Shona regularly at home with her husband and son.

Recording took place at the Integrated Speech Research Laboratory at the University of British Columbia. Three songs were recorded: a biblical hymn (*Jesu Idombo*), the national anthem of Zimbabwe (*Simudzai mureza wedu weZimbabwe*) and a traditional song (*Nyii dzaibva*). The composer and date of the hymn are not known but the subject characterized it as a very traditional hymn she had known since her childhood. It is an original Shona hymn, not a translation from another language. The national anthem was written by Solomon Mutswairo and Fred Lecture Changundega who won a competition held in 1994 to choose the national anthem. The traditional song is part of a larger narrative. It was recorded in context and later extracted for analysis. The subject was also recorded reading the texts of all three songs.

The audio recordings of all sessions were transferred to PRAAT for acoustic analysis. The signals were segmented into syllables and the mean fundamental frequency was calculated for each syllable. Mean F0 was used to capture any changes in F0 over the duration of the syllable; if the singer modifies the song melody in some way to reflect the speech melody, mean F0 would be expected to capture the changes.

Once the F0 measures were calculated, the transitions from one frequency to the next were examined. Transitions are relative, not absolute measures. From any given syllable, there are three possible directions in which the fundamental frequency can move: up, down or remain the same (see Figure 1).

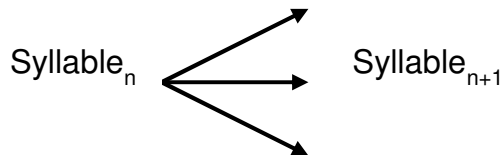$$\text{Syllable}_n \quad\text{Syllable}_{n+1}$$

Figure 1: the three possible directions of a transition

A sequence of transitions creates a melody. In this paper, *melody* is used to refer only to the transitions. In musicology, melody is usually considered to include both transitions and rhythm but in this study, rhythm has been excluded. This is not meant to deny its importance; the division is made only for the simplification of analysis.

Each transition was coded for directionality: up, down or the same (if the F0 of the second syllable was within 1.5 Hz of the first syllable, that transition was coded as being the same) and the direction of each corresponding transition in the sung and spoken melody was compared. Corresponding transitions between two given syllables in the text were coded as parallel if the direction was the same in both sung and spoken melodies; as opposing if one transition went up and the other went down; or as non-opposing if the directions were mismatched but not in opposite directions. For example, if one transition went up but the other remained the same, this was coded as a non-opposing transition. In cases where one syllable corresponded to two notes in the song melody both frequencies were recorded for the song melody and the corresponding frequency in the spoken melody was recorded (and plotted in the graphs) in line with the first note of the song melody, leaving a gap for the second note. This was counted as a single transition. When comparing directionality in these cases, if *either* of the two transitions of the song melody matched the direction of the spoken melody, the transition was counted as parallel. Occasionally, phrase final syllables in the spoken melody were completely devoiced. Consequently, these had no F0 and were excluded; these also show up as a gap in the graph. The number of parallel transitions is expressed as a fractional proportion of the total number of transitions and also as a percentage. Statistical
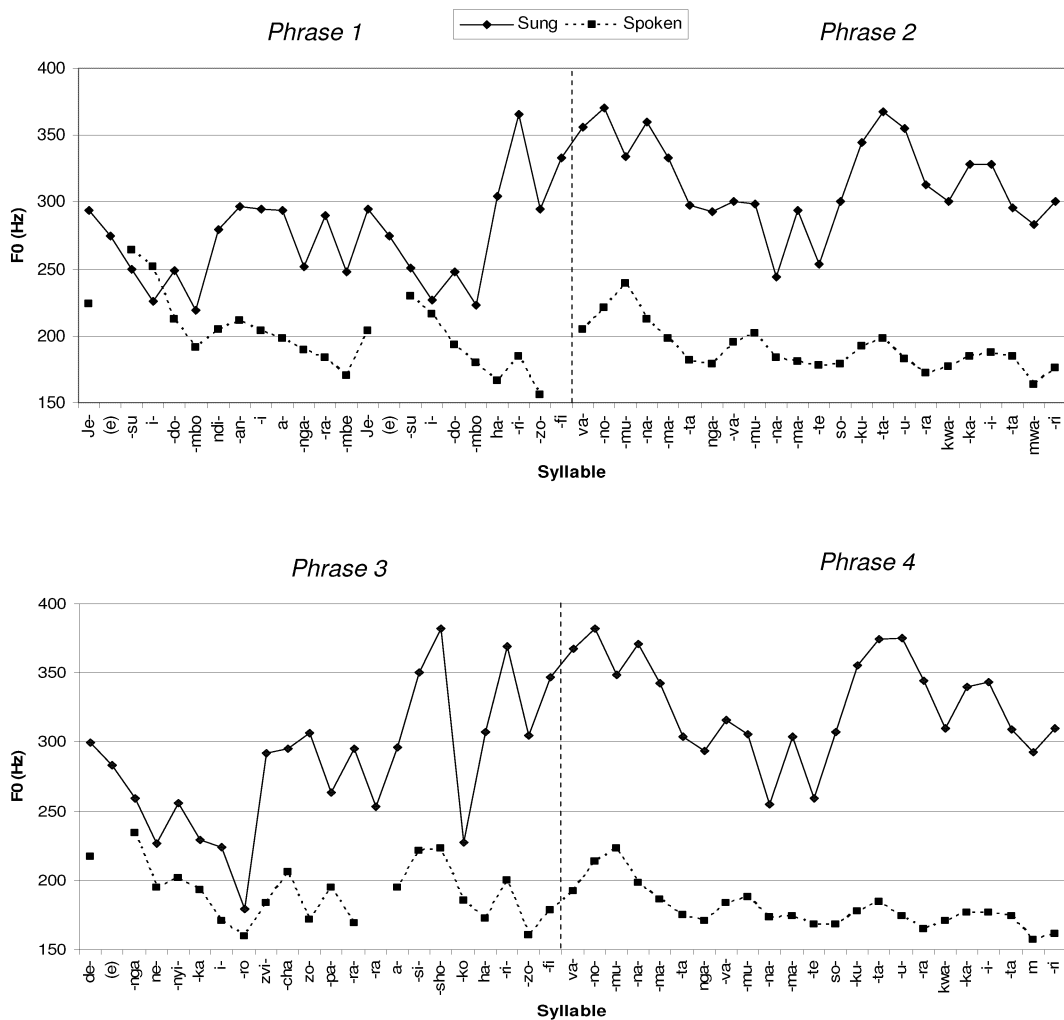
Table 1: Melody comparison for the biblical hymn *Jesu Idombo*

significance was calculated with a chi-square test; the chance of random correspondence (33.33%) was calculated mathematically.[1]

3. Results
3.1. Song One – Biblical Hymn (Jesu Idombo)

The results for the first song (*Jesu Idombo*) are shown in Table 1. Transitions across phrase boundaries were excluded. In the song *Jesu idombo*, there were eighty-one transitions. Fifty two transitions were parallel (52/81; 64.2%) showing statistical significance (p ≤ 0.001). Of the remaining transitions, 21 went in opposing directions and 8 differed in a non-opposing manner.

---

[1] As a confirmation, four random chains of 49 transitions were generated and the same analysis was carried out on all six possible combinations of random chains. The mean percentage of correspondence of all six random chains was 32.98% with a range of 26.5% - 38.8%.
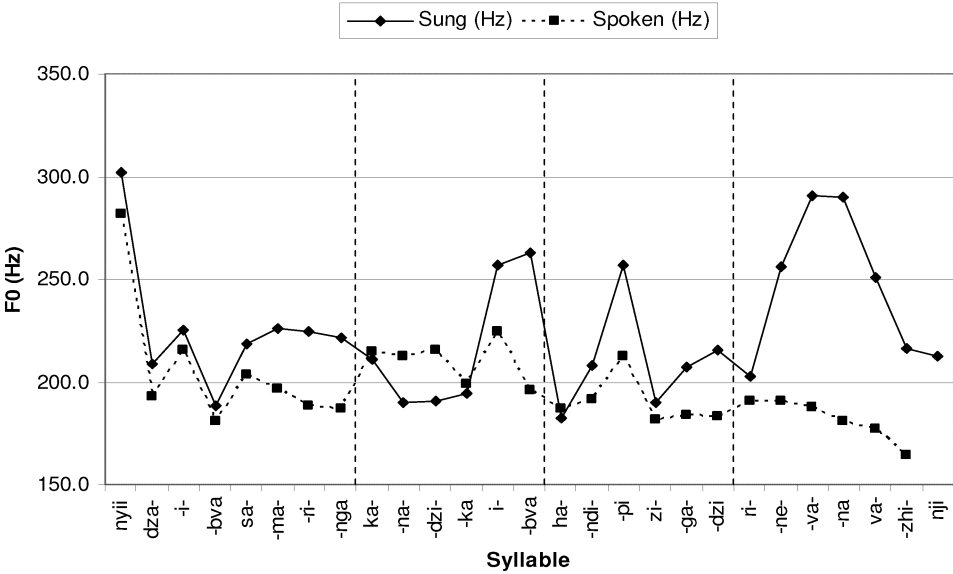
Table 2: Melody comparison for the traditional song *Nyii Dzaibva*

3.2. Song Two – Traditional Song (Nyii Dzaibva)

The analysis of the traditional song followed the same procedure as that for the biblical hymn and the results are shown in table two. In this analysis several text disparities were found between the sung and spoken versions: the repeats in the sung version did not match the repeats in the spoken version. To deal with this disparity, all repeats were omitted and only the first occurrence of each phrase was analysed. This gave 23 transitions, 13 of which were parallel (56.5%) which is statistically significant ($p \leq 0.025$). Six transitions were opposing and 4 were non-opposing.

3.3. Song Three – National Anthem (Simudzai mureza wedu weZimbabwe)

The national anthem consists of three verses with five phrases each. Each verse has different words with the exception of the final phrase which is identical in all three verses. The sung melody is, of course, the same for all three verses. The results are shown in table four.

The verse-by-verse breakdown of transitions is shown in table three. The number of parallel transitions for the whole song (75/161; 46.6%) is statistically significant ($p \leq 0.01$) but the numbers for each individual verse are not.

Verse	Parallel	Opposing	Non-opposing	Total
1	22	21	6	49
2	26	21	10	57
3	27	18	10	55
Whole song	75	60	26	161

Table 3: Transition breakdown for national anthem (*Simudzai mureza wedu weZimbabwe*)

Verse 1

Verse 2

Verse 3

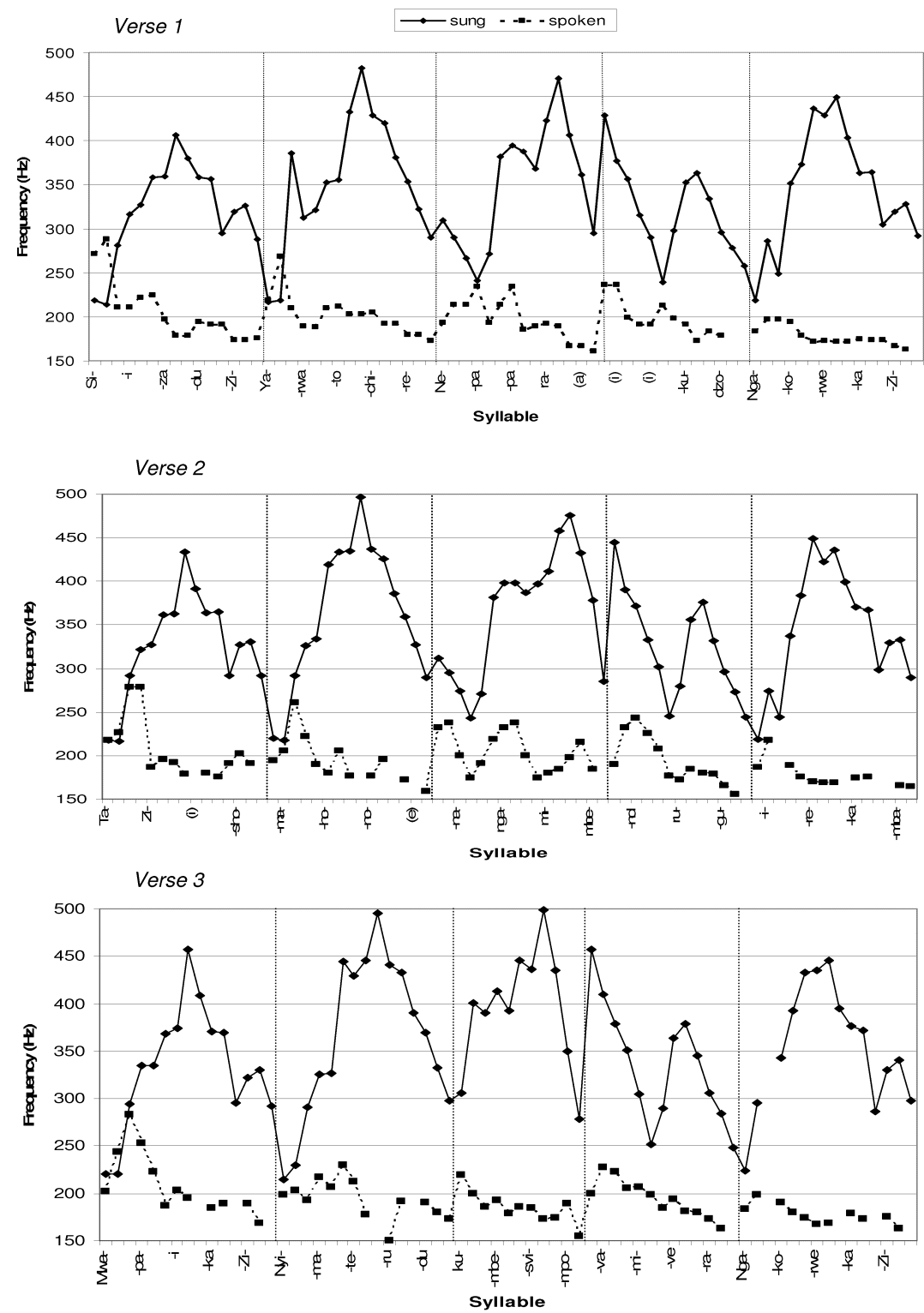Table 4: Melody comparison for the national anthem (*Simudzai mureza wedu weZimbabwe*)

4. Discussion and Conclusions

All three songs examined showed a statistically significantly number of parallel transitions; all three showed a greater proportion of parallel transitions than would be expected purely by chance. Even in the case of the national anthem where the same sung melody corresponds to three different spoken melodies, the overall pattern for the song shows a statistically significant level of correspondence. This suggests that sung melodies in Shona are influenced by the spoken melodies to which they correspond.

Locus, however, does not show any indication of predictability. Scatterplots showed no correlation between position and parallelism, nor was there any significance when examination was limited to the opening sequences of phrases (the first four syllables). There is a visually remarkable parallelism in the opening phrase of the traditional song (table 3) but caution must be used when evaluating this particular song. After recording the spoken version of this text, the subject asked if she could re-record it as she felt the first version was "too much like singing".[2] This particular text is very well known to the subject, but only in its sung version. This may be a case of the sung melody influencing the spoken melody.

A closer examination of non-parallel transitions provides some interesting results: statistical evaluations of the opposing transitions do not show significance but those for the non-opposing transitions in both the hymn and the national anthem do ($p \leq 0.001$ in both cases). This would seem to indicate that Shona does not treat melodic parallelism the same way that Hausa does, for example. In Hausa, non-opposing transitions are acceptable but opposing transitions are not (McHugh, p.c.).

The hierarchies proposed by Chao (1956) and Rycroft (1979) suggest that there should be less correspondence as the musical style moves towards more "contemporary" (*viz.* Western-influenced) composition. The results for the songs studied here seem to contradict that idea. Both the hymn and the national anthem are written in a more "modern" musical idiom yet they both show higher rates of parallelism than the very traditional story-song. Language change may account for a loss of parallelism in a very old song over time but there still remains the fact that both of the more recently composed songs contain a significantly high number of parallel transitions. The composers of these songs seem to have given at least some attention to the spoken melody of the words.

This acoustic analysis of three songs in Shona shows that the song melodies and spoken melodies share a significant number of parallel transitions but that there is not a transition-by-transition correspondence. The Shona evidence contradicts the strongest forms of the "words must determine music" theories but also disagrees with both the hierarchies of song type and the idea that parallel transitions occur at strategic locations.

It may be that a transition-by-transition analysis is too narrow. Visual examination of the melody comparison graphs indicates an overall similarity of shape which does not extend to the very narrow syllable-by-syllable comparison used here and elsewhere but which may be enough to satisfy linguistic needs. The next step may need to be to step back and look at the bigger picture.

References

Agawu, Kofi. (1988). Tone and tune: The evidence for northern Ewe music. *Africa, 58*(2), 127-146.

Agawu, Kofi. (1995). *African rhythm: A northern Ewe perspective*. Cambridge: Cambridge University Press.

Baart, Joan L. G. (2004). *Tone and song in Kalam Kohistani (Pakistan)*. Retrieved Oct. 7, 2007, from http://www.geocities.com/kcs_kalam/tskk.pdf

Chao, Yuen Ren. (1956). Tone, intonation, singsong, chanting, recitative, tonal composition and atonal composition in Chinese. In M. Halle, H.G. Lunt, H. McLean and C.H. Van Schooneveld (Ed.), *For Roman Jacobson: Essays on the occasion of his sixtieth birthday, 11th October 1956* (pp. 52-59). The Hague: Monton & Co.

Devine, A. M., and Laurence D. Stephens. (1994). *The Prosody of Greek Speech*. New York: Oxford University Press.

---

[2] The second version, which the subject felt was better, was used in the analysis.

Ekwueme, Lazarus N. (1974). Linguistic determinants of some Igbo musical properties. *Journal of African Studies, 1*(3), 335-353.

Herzog, George. (1934). Speech-melody and primitive music. *Musical Quarterly, 20*, 452-466.

Jones, A. M. (1959). *Studies in African music*. London: Oxford University Press.

List, George. (1961). Speech melody and song melody in central Thailand. *Ethnomusicology, 5*(1), 16--32.

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

Rycroft, David K. (1979). The relationships between speech-tone and melody in Southern African music. In J. P. Malan (Ed.), *South African music encyclopedia, vol. 2* (pp. 301-314). Capetown: Oxford University Press.

Schneider, Marius. (1950). La relation entre la melodie et le langue dans la musique Chinoise. *Anuario Musical, 5*, 62-69.

Schneider, Marius. (1961). Tone and tune in West African music. *Ethnomusicology, 5*(3), 204-215.

Stock, Jonathan P. J. (1999). A reassessment of the relationship between text, melody and aria structure in Beijing opera. *Journal of Musicological Research, 18*, 183-206.

Ward, W. E. (1932). Music of the Gold Coast. *The Musical Times, 73*(1047), 707-710.

Wee, L. H. (2007). Unraveling the relation between Mandarin tones and musical melody. *Journal of Chinese Linguistics, 35*(1), 128-143.

Wong, Patrick C. M., & Diehl, Randy L. (2002). How can the lyrics of a song in a tone language be understood? *Psychology of Music, 30*(2), 202-209.

Yung, Bell. (1983). Creative process in Cantonese opera I: The role of linguistic tones. *Ethnomusicology, 27*, 29-47.

Selected Proceedings of the 39th Annual Conference on African Linguistics: Linguistic Research and Languages in Africa

edited by Akinloye Ojo and Lioba Moshi

Cascadilla Proceedings Project   Somerville, MA   2009

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Schellenberg, Murray. 2009. Singing in a Tone Language: Shona. In *Selected Proceedings of the 39th Annual Conference on African Linguistics*, ed. Akinloye Ojo and Lioba Moshi, 137-144. Somerville, MA: Cascadilla Proceedings Project.

or:

Schellenberg, Murray. 2009. Singing in a Tone Language: Shona. In *Selected Proceedings of the 39th Annual Conference on African Linguistics*, ed. Akinloye Ojo and Lioba Moshi, 137-144. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2193.