Phonetic Assessment of Tone Spreading

Author(s): Sung-A Kim

Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals (1998), pp. 129-138

Phonetic Assessment of Tone Spreading

Sung-A Kim
University of Texas at Austin

1. Introduction. The alignment of fundamental frequency (f0 henceforth) contour with other units in speech is an important issue not only for modeling f0 contour in speech analysis, but also for theoretical understanding of tone and intonation systems. The f0-syllable alignment has been investigated by two groups of researchers from different perspectives. The first line of research has been conducted in tone studies within a framework of theoretical linguistics. Tones can be spread, shifted, or copied, all involving changes of alignment of tones with their host units (Hyman and Schuh 1974, Schuh 1978). Major portions of metrical and autosegmental phonology of tone/accent languages are devoted to various complex tone-syllable alignment issues. The f0-syllable alignment has been formalized by resorting to the notion of 'association lines' in autosegmental phonology (Goldsmith 1979, 1990, among others).

The second line of research has been mainly conducted in the field of phonetics. It concerns physical aspects of alignment of f0 contour with the segmental units in speech. For example, Bruce (1977) found that f0 realizations in Swedish are characterized by the relative stability of certain tonal targets. However, a review of these tone/accent studies displays a general lack of communication between the groups. This paper tries to fill this gap by investigating one of the most common tone alternations: tone spread.

Tone spread is described as a phonological process whereby "a tone moves beyond its original segmental domain to replace or displace the tone of the following syllable or syllables" (Schuh, 1978: 231). In autosegmental phonology, the tone spread is formalized as a process in which two tone-bearing units share the same tone, as schematically shown in figure 1.

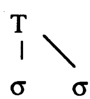


Figure 1. Phonological analysis of tone spreading

Recent works on fundamental frequency realization cast a doubt on the nature of the phonological analysis of tone spreading. It is agreed that the f0 peak corresponding to high tone tends to be delayed until the later part of the tone-bearing unit or the onset of the next syllable (Silverman and Pierrehumbert 1990, Prieto et al 1995, Arvaniti et al 1998, Kim 1998, among others) because of a sluggish cessation of f0 movement (Ohala 1978, Fujisaki 1988). Furthermore, phonetic studies on implementation of tones have shown that only a single peak is found in the context of tone spreading (Xu and Wang 1997, Kim 1998). Findings of the phonetic studies raise a serious question of how tone spread is distinguished from f0 peak delay described above. Will f0 peak delay and tone spreading be isomorphic in nature? This issue has not been addressed before, simply assuming that phonological spreading would be realized by an f0 plateau laid upon two tone-bearing units at the phonetic implementation level (Pierrehumbert and Beckman 1988). This paper attempts to answer this question and proposes a phonetic assessment for the facts previously described as tone spreading in Yoruba and Chichewa.

The remainder of the paper is organized as follows: Facts of the tone alternation in Yoruba and Chichewa are introduced in section 2. A brief sketch of the previous literature on f0 peak realization will be presented in section 3. The experimental method and the result will be addressed in sections 4 and 5, respectively. Implications of this study will be discussed in the conclusion.

2. Tone Alternation in Yoruba and Chichewa. Yoruba has a three-tone system, which is composed of high, mid, and low tones. Among the three tones, mid tone is phonologically inert in the sense that it does not interact with adjacent high or low tones. It is not subject to tone spreading rules (Pulleyblank 1988, Laniran 1992). In comparison, high tone spreads to the next low-toned syllable but not to the mid-toned syllable as shown in (1).

(1) a. /Láyò/	[Layô]	'a personal name for male'
/Débò/	[Debô]	'a personal name for male'
b. /lóko/	[lóko]	'a farm'

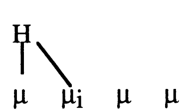
In (1a), high tone on the first syllable spreads to the next low-toned, and the second syllable bears a contour tone. On the other hand, such a contour does not occur where the high tone is followed by a mid tone, as shown in (1b). The tone alternation described above occurs regardless of the position in which the high tone occurs in a phrase.

The tone alternation in Chichewa, a language with high and low tones, differs from the tone alternation in Yoruba in the sense that it is position-sensitive. In this language, low tones are assumed to be phonologically inert and thus low tones do not play a role in tone patterns. It is a high tone that spreads and deletes. In Chichewa, high tones generally spread forward one syllable as in (2a) and (2b), but not onto or within the disyllabic phrase-final foot as in (2c). High-toned vowels created by tone spread are underscored in (2).

(2) a. tinapátsá mwaná dé^{n}guiili	'We gave the child a basket'
b. chigawé^{n}gá iichi	'this terrorist'
c. chigawée^{n}ga	'terrorist'

Kanerva (1989) formulates the position-sensitive tone spread rule for the data in (2) as nonfinal high tone-spreading as depicted in (3).

(3) Nonfinal doubling:



condition: μ_i is not in phrase-final foot.

According to this rule, a high tone spreads onto the following mora if the mora is not in the phrase-final foot. In (2a) and (2b), there occurs high tone spreading because the target mora is not in the phrase-final foot. On the other hand, the high tone in (2c) does not spread to the following mora since the target mora is included in the phrase-final foot. The same tone alternation is also observed by Mtenje

(1988). The phonological analyses of the tone alternation in Chichewa can be summarized as follows: Tone spreading occurs in pre-penult positions, while it does not in penult positions in Chichewa.

Before moving on to the next section, let us briefly discuss a problem with the spreading analysis given in (3). The problem arises from the special reference to the phrase-final element. According to the phonological analysis, the spreading rule must refer three moras ahead before tone-spreading takes place. It makes this rule typologically peculiar in the sense that local tone spreading is triggered by an element at the end of a phrase and by counting the syllables from the end of the phrase. This pattern leads to a violation of the locality condition in an otherwise well-motivated generalization. Phonological rules are usually assumed to be subject to conditions of locality where the trigger and the target should be structurally adjacent to each other (McCarthy and Prince 1986, Odden 1994). The tone-spreading as described is a violation of this well-motivated constraint.

The summary of tone alternations in the two languages is tabulated in (4). In Yoruba, high tone spreading is assumed to occur in a high-low context (henceforth, HL context), but not in a high-mid context (henceforth, HM context). In Chichewa, high tone spreading is supposed to occur in pre-penult positions but not in penult positions.

Languages	Locality	Phrase Position	So-called Spreading Case
Yoruba	local	Insensitive	HL contexts
Chichewa	non-local	Sensitive: Spreading only in pre-penult positions	Pre-penult positions

(4) Summary of Phonological Analysis of Tone Alternation. I explored an instrumental analysis of the facts previously described as tone spreading, which suggests another description of the facts and another analysis. This ultimately proposes that the position-sensitive tone alternations are not tone spreading, but a by-product of f0 peak alignment.

3. Theoretical Background: F0 Peak Delay. Compared to the unusualness of the phonological analysis in Chichewa given in section 2, the position-sensitive processes are not typologically unusual, from the perspective of the phonetic timing of the f0 peak. A number of studies on the phonetic realization of accent and tone have found that phonetic prominence, specifically pitch prominence, may not always align with the accented or tone-bearing syllable (Steele 1986, Silverman and Pierrehumbert 1990 for English, Prieto et al. 1995 for Mexican Spanish, Liberman et al. 1993 for Igbo, and Liberman 1996 for Yoruba).

Also, it has already been well established that the f0 delay does not occur where the syllable is close to a prosodic phrase edge. Steele (1986) and Silverman and Pierrehumbert (1990) show that two factors, rhyme duration and upcoming prosodic contexts, are the main source of peak location variation in English. That is, when a vowel is lengthened because of slow speech, the f0 peak is correspondingly delayed. Given an identical phrase position, there is a positive correlation between vowel duration and f0 peak delay relative to the vowel onset. In contrast to this, the f0 peak is aligned early in the syllable where the syllable is close to a prosodic edge in English. Prieto et al. (1995) report a similar result for Mexican Spanish. In addition, they observe that the same finding is obtained when f0 peak delay is measured relative to accented syllable onset as well as vowel/rhyme onset. Kim (1998) reveals that the distance from syllable onset to f0 peak highly correlates with the duration of tone-bearing syllables in Chichewa. Of particular interest is the lack of correlation between the syllable duration and distance between syllable offset and f0 peak (offset-to-f0 peak) in Chichewa. This means that offset-to-f0 peak does not systematically vary with any consistency. In other words, the f0 peak simply stays close to the syllable offset. The high correlation between f0 peak relative to the tone-bearing syllable onset and the syllable duration indicates that the f0 peak is moving closely with the syllable offset. Let us call this pattern 'peak delay.' As Ohala (1978) and Fujisaki (1988) argue, the f0 peak delay appears to be due to a function of sluggish cessation of an f0 movement. In other words, f0 is a function of the strain of the muscles such as cricothyroid and sternohyhoid, the mass of the thyroid cartilages and stiffness of the cricothyroid joint. Even if the neural commands for producing a pitch target is issued simultaneously with those for producing the syllable that carries it, the pitch target would be attained more slowly than the segmental targets. Therefore, it is possible that f0 peak is realized on the following syllable.

The peak delay pattern in f0 represents a regular timing relation between the laryngeal gesture that leads to the f0 peak and landmarks in the syllable (syllable onset, syllable offset). Since the groundbreaking work of Bruce (1977) on Swedish, it has been known that at least some f0 patterns are characterized by relative phonetic stability of certain tonal targets. Bruce finds that the local f0 peak corresponding to Swedish word accent is constantly aligned with the segmental material and the f0 value is constant across multiple repetitions of the same utterance by the same speaker. Bruce interprets this to mean that, for the Swedish accentual distinction, "reaching a certain pitch level at a particular point in time is the important thing, not the movement (rise or fall) itself (1977:132)." The same kind of stability is observed elsewhere as well. For instance, Huffman (1993) shows that syllable landmarks are important in the timing of velar gestures. The regular and proportional timing between gestures of different articulators is also founded in work within the framework of task dynamics (Tuller and Kelso 1984 and Browman and Goldstein 1990 among others).

4. Experimental Design. The experiment was designed to compare the f0 alignment of the so-called spreading case to the non-spreading case in each language. Consider the sentences in (5) from the Yoruba corpus used in the experiment.

(5) Corpus in Yoruba

a. Ó	<u>mòómò</u>	lò	ìyen	b.	<u>Màmámia</u>	lè	lò	ìyen.
He	intentionally	use	that one		mother my	is able to use		that one
"He intentionally uses that one."					"My mother is able to use that one."			

The two sentences in (5) are almost identical to each other except for the target words, indicated by underlines. The underscored word in (5a) contains a sequence of a high tone followed by a low tone (henceforth, HL), while the one in (5b) has a sequence of high tone followed by a mid tone (henceforth, HM). Therefore, the target word in (5a) corresponds to the spreading case, whereas the one in (5b) corresponds to the non-spreading case. If there is tone spreading only in (5a), clear differences in the f0 realization between the spreading and non-spreading cases are predicted by the phonological analysis. The predictions made by phonological analysis of tone-spreading are summarized in (6).

	F0 realization	Number of Associated Syllable(s)
Spreading	F0 plateau (f0 peak laid upon two syllables)	2
Non-spreading	no plateau (single f0 peak)	1

(6) Prediction Borned out of Phonological Analysis of Tone Spreading.

A crucial difference between spreading and non-spreading cases is whether there is a tone target in the syllable next to the underlying high tone-bearing syllable. As shown at (6), the so-called spreading case is characterized by the existence of an additional tone target. In order to test the prediction born out of the phonological analysis of tone spreading, I compared the f0 peak alignment in the two target words in (5). If there is a tone spread in pre-penult positions, then we should find f0 peak to be timed with the syllable following the high tone-bearing syllable, while the f0 peak is timed with the high tone-bearing syllable itself if there is no tone spread.

The same reasoning applies to the tone spreading case in Chichewa. For Chichewa, the sentence in (7) was used in the experiment. Notice that the sentence has four high tones. Among those, the third high tone and the fourth high tone underlined are the main concern here. It is the fourth high tone that is located in the phrase-final foot, while the third high tone is not. For the sake of convenience, let us call the third and the fourth high tones the pre-penult and penult tones respectively. In order to minimize segmentally induced perturbation on f0, target words indicated by italics are composed of sonorants, except the *k* in *kun^{y}en^{y}a* . Also notice that the pre-penult and the penult tone-bearing vowels are identical (i.e., *e*) to control for the intrinsic f0 of vowels.

(7) Mló^{n}da *áma yenéra kun^y én^y a.*
The watchman must to goof off
"The watchman must goof off."

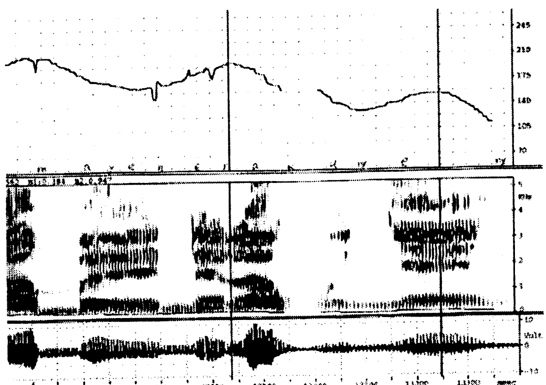
If there is a spreading only in pre-penult positions in Chichewa, then the f0 alignment in the pre-penult positions should differ from the one in penult positions. If there is no tone spread, f0 is expected to be determined by reference to the tone-bearing syllable regardless of the positions.

In the experiment, a male Chichewa speaker and a female Yoruba speaker participated at the recording. The Chichewa speaker and the Yoruba speaker uttered the sentence in (5) and (7) respectively. To induce a broad range of f0 values and syllable duration, the speaker was asked to vary loudness and speech rate. There were three conditions with respect to loudness as used in Liberman et al. (1993). A total of 576 tokens (1 speaker x 2 sentence types x 3 loudness levels x 2 speech rates x 2 target words x 24 repetitions=576) was obtained for Chichewa. A total of 120 tokens was obtained for Yoruba. The speech signal recorded on a digital audio tape-recorder was digitized at a sampling rate of 22 kHz and segmented from waveform and spectrogram display, using Sound Scope, a Macintosh-based sound analysis by GW Instruments.

The following measurements were taken for the two target words in spreading and non-spreading cases in each language.

(8) Measurement Points

a. The onset and offset of the tone-bearing syllable,
b. the onset and offset of the tone-bearing rhyme,
c. the onset and offset of the syllable following the tone-bearing syllable,
d. the onset and offset of the rhyme in the syllable following the tone-bearing syllable,
e. the f0 peak corresponding to high tones in both spreading and non-spreading cases.



m a y e n e r a k u n^y e n^y

Figure 2. A pitch track of pre-penultimate and penultimate high tones in Chichewa.

Figure 2 shows the waveform display, f0 contour, and spectrogram corresponding to the target words *amayenera* and *kun^{y}en^{y}a*. Segmental transcription is given at the bottom of Figure 2.

The following null hypotheses are tested in the study. If those null hypotheses are rejected, then it can be said that tone spreading analysis is confirmed in each language.

(9) The Hypotheses:

a. In Yoruba, the temporal locations of the f0 peak are determined by reference to high tone-bearing syllable in both HL and HM contexts.
b. In Chichewa, the temporal locations of the f0 peak are determined by reference to the high tone-bearing syllable in both pre-penult and penult positions.

In order to test these hypotheses, I compared models of the temporal relation using multiple regression. The relevant variables used in the multiple regression are given in (10):

(10) Variables

a. Peak delay: Temporal distance between f0 peak and the onset of the tone-bearing syllable.
b. T-Syllable duration: Duration of the tone-bearing syllable.
c. N-syllable duration: Duration of the syllable following the tone-bearing syllable.

F0 peak relative to the onset of the tone-bearing syllable in (10a) is the dependent variable and position and the syllable duration in (10b) and (10c) are the independent variables in the multiple regression. Penult position and pre-penult position are coded as 1 and 0 respectively.

5. Results and Discussion.

5.1. Yoruba. First, let us discuss the results of Yoruba. With regard to the f0 peak in pitch tracks, only a single f0 peak is observed in both HM and HL contexts. There was no f0 plateau laid upon the two syllables in general. The f0 peak is delayed until the syllable following the high tone-bearing syllable. As a result, the longer the high-toned syllable, the longer the f0 peak is delayed. At first glance, it appears that there is no difference in f0 timing between the HL context (i.e., spreading case) and HM context (i.e., non-spreading cases) with regard to the f0 peak alignment. A close examination, however, reveals that the two contexts differ from each other in the syllable with which the f0 peak has a constant relation. Consider the table in (11) where the results of simple regression models in HM context are summarized. The one in (11a) indicates the regression model where the peak is assumed to be timed with the tone-bearing syllable itself, while the one in (11b) represents the regression model where the peak is supposed to be timed with the syllable next to the tone-bearing syllable. We can compare models using relative measures of goodness of fit. One of such measures is the Pearson R^2 values, which is indicated in the second column in the table. For example, the R^2 value of 0.684 indicates that the model accounts for 68% of the variation in the data.

a. Peak = 0.049+ 0.676*Tsyll	R^2: .684	HM context (non-spread)
b. Peak = 0.001+0.022*Nsyll	R^2: .002	HM context (non-spread)

(11) Simple regression models for the prediction of f0 peak location

In the table in (11), higher R^2 value in (11a) indicates that f0 peak is timed with the tone-bearing syllable in non-spreading context. The extremely low R^2 value in (11b) shows that the f0 peak has almost no relation with the syllable following the tone-bearing syllable.

In the spreading context, however, the opposite result is obtained, as shown in (12). The equation in (12b) shows the result of the case where the f0 peak is assumed to be timed with the next syllable in the spreading context. This equation model has higher R^2 value than the one in (12a). This means that the temporal location of f0 peak is best predicted by reference to the syllable next to the tone-bearing syllable in spreading cases.

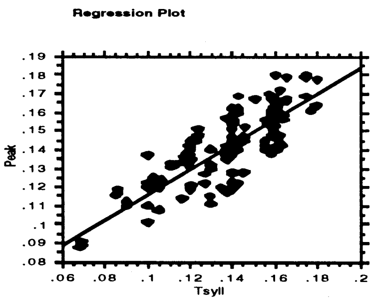
a. Peak = 0.134+0.901*Tsyll	R^2: .592	HL context (spread)
b. Peak =-0.027+0.823*Nsyll	R^2: .621	HL context (spread)

(12) Simple regression models for the prediction of f0 peak location

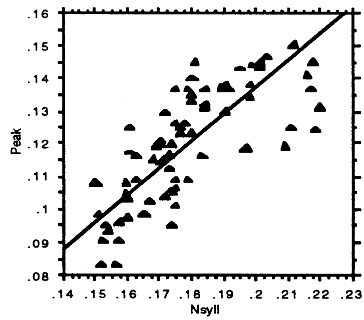
If we compare the tables in (11) and (12), it is clear that there is an asymmetry between HM and HL contexts. In a non-spreading context, f0 peak has a constant relation with the high tone-bearing syllable itself. On the other hand, in a spreading context, it is constantly timed with the syllable next to the tone-bearing syllable. The equation models in (11a) and (12b) are graphically shown in (13a) and (13b).

(13) Regression equations in non-spreading and spreading cases in Yoruba

a. HM context



b. HL context



To sum up, Yoruba results appear to support the phonological analysis of tone spreading as they display a difference in f0 timing between spreading and non-spreading cases.

5.2. Chichewa. In the previous section, Yoruba results show that there is a clear difference in f0 timing between spreading and non-spreading cases. In comparison, Chichewa results show that there is no difference in f0 timing between spreading and non-spreading cases. F0 peak is timed with the high tone-bearing syllable in both cases in Chichewa. The table in (14) shows two regression models in pre-penult positions whereas the one in (15) contains two regression models in penult positions. It should be noted that regression equations in (14a) and (15a) have higher R^2 values than the ones in (14b) and (15b).

a. Peak =.003 + 1.355 * Tsyll	R^2: .775	pre-penult (i.e. spreading)
b. Peak = .028 + .968 * NSyll	R^2 .381	pre-penult (i.e. spreading)

(14) Simple regression models for the prediction of f0 peak location

a. Peak =-.027 + .637 * Tsyll	R^2: .616	penult (i.e. non-spreading)
b. Peak = .063 + .293 * NSyll	R^2: .217	penult (i.e. non-spreading)

(15) Simple regression models for the prediction of f0 peak location

If there is tone spreading only in penult positions, then it is predicted that the f0 peak in pre-penult positions has the most constant relations with the syllable next to the high tone-bearing syllable, as we found in Yoruba cases. Notice that the R^2 value in the N-syllable based model (i.e., equations in 14b and 15b) is much lower than the one in the T-syllable based model (i.e., equations in 14a and 15a). When the peak delay is plotted against the duration of the high tone-bearing syllable as in figure 3, quite high R^2 values (0.775 in pre-penult positions and 0.616 in penult positions) were obtained in both pre-penult and penult position. Therefore, the tone spreading analysis is rejected in Chichewa.

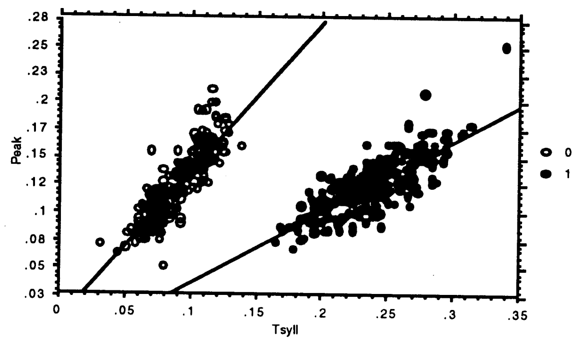


Figure 3. Peak delay as a function of duration of tone-bearing syllable in Chichewa.

The findings of this study are important both empirically and theoretically. Empirically, they provide instrumental data about the little-studied f0 peak delay phenomenon in African tone languages. Theoretically, the experimental results for Chichewa indicate that the tone alternation in the language deserves a phonetic account rather than a phonological analysis. The phonetic account does not require that Chichewa be an exception to phonological constraint on locality. More importantly, the contrast between tone alternations in the two languages provides an empirical ground to tease apart phonetic implementation of tone (e.g. tone alternation in Chichewa) from true phonological rules of tone (e.g. tone alternation in Yoruba).

Acknowledgments

I would like to thank Larry Hyman, Björn Lindblom, Scott Myers, Tony Woodbury, and Yi Xu for helpful discussions and comments. I also thank Amalia Arvaniti for sending her papers. Any error is mine.

References

ARVANITI, R. LADD and INEKE MENNEN. 1998. Stability of tone alignment: The case of Greek prenuclear accents. To appear in Journal of Phonetics, 26.1

BROWMAN, CATHERINE and LOUIS GOLDSTEIN. 1990. Tiers in articulatory phonology, with some implications for casual speech. Papers in laboratory phonology I, ed. by John Kingston and Mary Beckman, 341-376. Cambridge, MA: Cambridge University Press

BRUCE, GOASTA. 1977. Swedish word accents in sentence perspective. Lund: Gleerup.

FUJISAKI, HIORTA. 1988. A note on the physiological and physical basis for the phrase and accent components in the voice fundamental contour. Vocal physiology: Voice production, mechanisms and functions, ed. by Osamu Fujimura, 347-355. New York: Raven Press.

GOLDSMITH, JOHN. 1979. Autosegmental phonology. New York: Garland.

_____________. 1990. Autosegmental and metrical phonology. Oxford: Blackwell Publishers.

HUFFMAN, MARY. 1993. Phonetic patterns of nasalization and implications for feature specification. Phonetics and phonology 5: Nasals, nasalization and the velum, ed. by Mary Huffman and R. Krakow, 303-327. New York: Academic Press.

HYMAN, LARRY M, and RUSSELL G. SCHUH. 1974. Universals of tone rules: Evidence from west Africa. Linguistic Inquiry 5.81-115.

KANERVA, JONNI. 1989. Focus and phrasing in Chichewa phonology. Stanford, CA: Stanford dissertation.

KIM, SUNG-A 1998. Position-sensitive tone alternation in Chichewa: Phonological analysis vs. phonetic timing. To appear in the proceedings of the 34th Chicago Linguistic Society.

LANIRAN, YETUNDE. 1992. Intonation in tone languages: The phonetic implementation of tones in Yoruba Ithaca, NY: Cornell University dissertation

LIBERMAN, MARK. 1996. Tone, accent and relative prominence. Talk given at the University of Texas at Austin.

_______; J. Schults; Soonhyun Hong; and Vincent Okeke. 1993. The phonetic interpretation of tone in Igbo. Phonetica 50.147-160.

MCCARTHY, JOHN, and ALAN PRINCE. 1986. Prosodic morphology. Amherst and Waltham, MA: University of Massachusetts, and Brandeis University, MS.

MTENJE, AL. 1988. On tone and tone transfer in Chichewa reduplication. Linguistics 26.125-155.

ODDEN, DAVID. 1994. Adjacency paramenters in phonology. Language 70.2.289-330.

OHALA, JOHN J. 1978. The production of tone. Tone: A linguistic survey, ed. by Victoria Fromkin. 5-39. New York: Academic Press.

PIERREHUMBERT, JANET and MARY BECKMAN. 1988. Japanese tone structure. Cambridge, MA: MIT Press.

PRIETO, PILAR; JAN VAN SANTEN; and JULIA HIRSCHBERG. 1995. Tonal alignment patterns in Spanish. Journal of Phonetics 23.429-451.

PULLEYBLANK. DOUGLAS. 1988. Vocalic underspecification in Yoruba Linguistic Inquiry 19.233-270.

SCHUCH, RUSSELL G. 1978. Tone rules. Tone: A linguistic survey, ed. by Victoria Fromkin. 221-256. New York: Academic Press.

SILVERMAN, KIM, and JANET PIERREHUMBERT. 1990. The timing of prenuclear high accents in English. Papers in laboratory phonology I, ed. by John Kingston and Mary Beckman, 72-106. Cambridge, MA: Cambridge University Press.

STEELE, S.A. 1986. Nuclear accent f0 peak location: Effects of vowel, rate, and the number of syllables following. Journal of Acoustical Society of America Supplement 1.80; s 51.

TULLER, BETTY, and SCOTT KELSO. 1984. The timing of articulatory gestures: Evidence for relational invariants. Journal of the Acoustical Society of America 76.1030-1076.

XU,YI , and EMILY WANG. 1997. Intonation contours decomposed: From bottom up. Evanston, IL: Northwestern University, MS.