

DIFFERENTIAL PERCEPTION OF TONAL CONTOURS THROUGH THE SYLLABLE

David House

Department of Linguistics and Phonetics, Lund University
Helgonabacken 12, S-223 62, Lund, Sweden

ABSTRACT

This paper proposes a model of differential tonal perception in which perception of tonal movement in spoken language is influenced by syllable structure. In the proposed model, tonal movement at syllable onset is perceptually coded differently from movement in the syllable nucleus or in the coda. An F₀ contour through the syllable onset or at the beginning of the nucleus is coded as a level tone while a contour through the nucleus or at the beginning of the coda is coded as a contour tone.

1. INTRODUCTION

A current issue in the description of tonal contours in spoken language involves the percept of level tones and contour tones. When is tonal movement perceived as a level tone and when is it perceived as a contour tone? What are the perceptual constraints acting upon the course of F₀ through the syllable to optimize perception of tonal movement as a sequence of levels or as perceptual movement?

In a previous study [1] it was shown that in synthesized VCVCV sequences where V= [a] and C= [m], tonal movement in vowels is perceptually more salient than in consonants. On the basis of these results and on theories of pitch perception, a model of optimal tonal perception was proposed. In the model, tonal movement early in the vowel (through areas of changing spectral and intensity characteristics) is recoded as level features, while movement later in the vowel (through areas of spectral stability) is coded as contour features.

This can be illustrated by a series of figures adapted from [1]. In Figure 1, a stylized example of changing spectral and intensity characteristics through a CVC syllable is shown. New spectral information is defined as an integration of the amount of spectral change and relative intensity at a particular point in time in the CVC sequence.

The moderate amount of new spectral information at the beginning of the syllable decreases through the relatively stable consonant segment (the syllable onset). At vowel onset (the beginning of the syllable nucleus) a combination of factors such as consonant release, aspiration, increasing intensity, and rapid formant transitions during the beginning of the vowel can contribute to create a maximum of new spectral information shortly after vowel onset. New spectral information decreases during the more spectrally stable vowel

and increases again at the end of the vowel into the following consonant (at the beginning of the syllable coda). However, because intensity decreases from the nucleus through the coda, the amount of new spectral information here is far less than the maximum level at vowel onset.

New spectral information

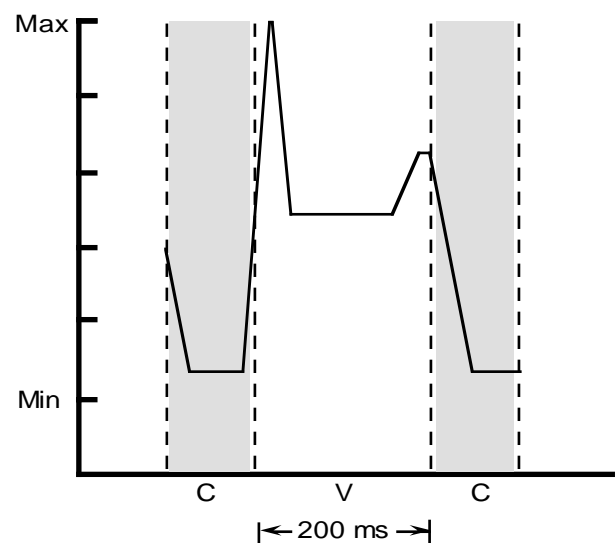


Figure 1: Stylized example of changes in amount of new spectral information in arbitrary units through a CVC syllable.

The model of optimal tonal perception is based on the assumption that there exists a general relationship between signal complexity and pitch sensitivity. As the complexity of the signal increases (reflected in a rising amount of new spectral information) pitch sensitivity decreases. When the perceptual mechanism is maximally loaded with the task of resolving spectral information and rapidly changing intensity, its capacity to resolve fundamental frequency movement is decreased. It follows, therefore, that the optimal timing for a tonal contour to be perceived as a true contour tone (movement feature perception) is within the spectrally stable portion of the vowel. According to the model, this movement must also be synchronized with vowel onset so that the beginning of the fall or rise occurs 30-50 milliseconds into the vowel. This is to enable pitch extraction of a relative pitch frequency to which the perception of pitch movement direction can be calibrated. Finally, the model proposes a duration constraint which

requires a vowel duration greater than 100 milliseconds for optimal movement feature perception.

In Figure 2, the CVC frame from Figure 1 is shown along with three falling tonal contours differing only in timing. The earliest contour would receive the level feature category *low* (L) since the movement placement is through the area of maximum new spectral information resulting in a low level through most of the syllable nucleus. The percept of this tonal contour is thus a jump from a previous high level to a low level (cf. [2]).

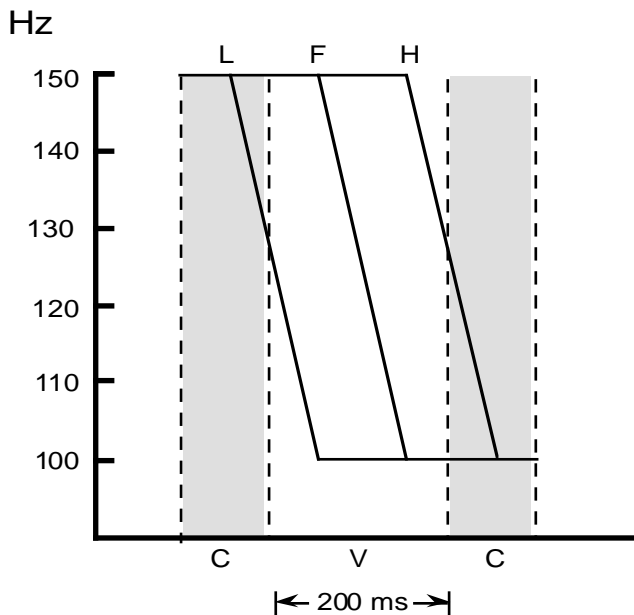


Figure 2: Stylized example of three falling tonal contours differing in timing through a CVC syllable. According to the model of optimal tonal perception, the three contours are categorized as *low* (L), *falling* (F) and *high* (H).

The middle contour in Figure 2 would receive the contour feature category *falling* (F) since the movement placement is through the area of spectral stability in the vowel. The percept of this tonal contour is thus a gliding, falling tone through the syllable nucleus.

The final contour in Figure 2 would receive the level feature category *high* (H) since the movement placement is through an area of spectral change at the beginning of the syllable coda. The percept of this tonal contour is thus a high level in the syllable nucleus.

Recent findings, however, give reason to question the percept of the final contour as a high level. In production data for the tone described as falling in Standard Thai [3,4,5,6], the falling gesture is typically produced late in the vowel or at the beginning of the sonorant coda of the syllable and is thus not restricted to the area of spectral stability of the vowel. In another study [7] it was shown that a silent interval influences tonal perception in the preceding sonorant consonant, i.e. the syllable coda, resulting in increased pitch salience immediately

before silence. In light of these findings, the optimal tonal perception model is reevaluated concerning the falling tonal contour at the beginning of the syllable coda.

2. EVIDENCE FROM THAI

Syllable onset in Thai may consist of simple consonants or consonant clusters such as [kl] or [kw]. In the case of onsets consisting of simple consonants the pattern of spectral and intensity change through a Thai CVC syllable is similar to that described in Figure 1. In the case of onsets consisting of consonant clusters a more complicated picture is necessary to describe the changing spectral information through the CCVC syllable. Figure 3 is adapted from production data presented in [6] showing a CCVC syllable consisting of a voiceless stop, a liquid, a vowel and a nasal consonant. The amount of new spectral information is traced in the same way as in Figure 1.

New spectral information

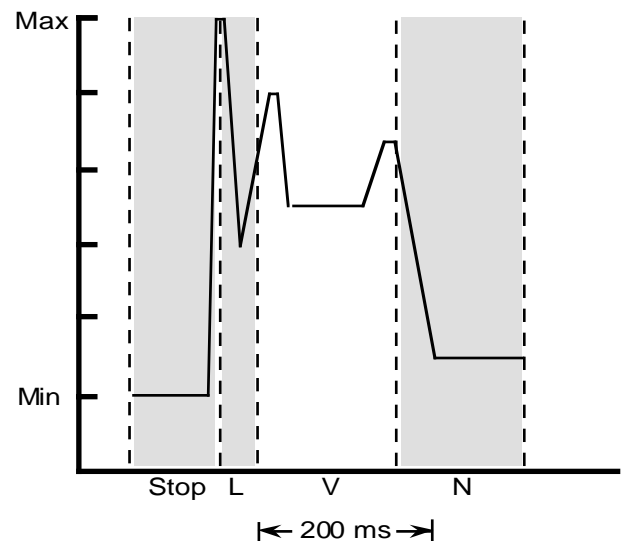


Figure 3: Stylized example of changes in amount of new spectral information in arbitrary units through a CCVC syllable, based on Thai production data.

Here, the new information is at a minimum during the voiceless occlusion for the stop and then rises rapidly to a maximum at stop release. New spectral information decreases during the short, stable portion of the liquid and then rises again at vowel onset. From this point through the syllable rhyme, the new information contour is the same as in Figure 1.

By examining production data using Thai words with and without consonant clusters at syllable onset, House and Svantesson [6] tested the hypothesis that the different syllable configurations would influence the timing of the tonal gesture for the two Thai tones described as low and falling (cf. [3,4,5]). Both tones have a falling gesture when preceded by a high tone: the low tone falls early in the vowel and the falling tone falls late in the vowel. If vowel onset functions as a synchronization

point for tonal timing, the falling contours would be delayed in the CCVC (cluster condition) compared to the CVC syllables when measured and time normalized from onset of voicing. If, on the other hand, the maximum peak of new spectral information functions as the synchronization point, there would be no delay in the contours. Production results in [6] indicate the latter situation to be the case. This is illustrated in Figures 4 and 5. Figure 4 shows the two falling contours in the cluster context, while Figure 5 shows the contours in the CVC context.

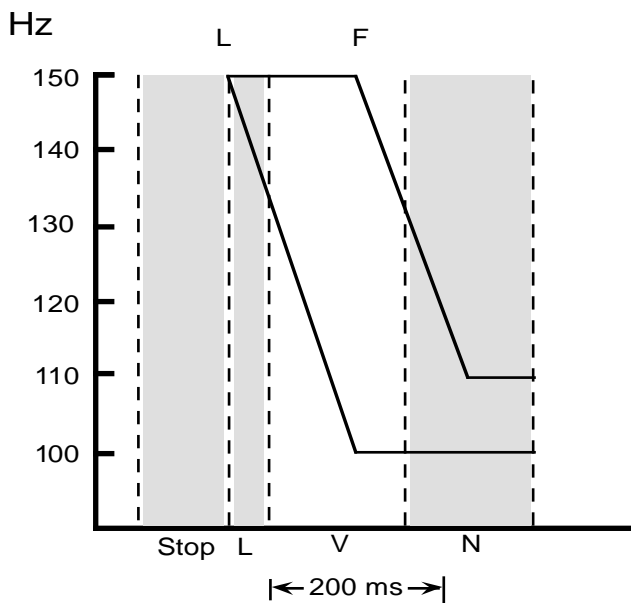


Figure 4: Stylized example of two falling tonal contours differing in timing through a CCVC Thai syllable. According to the revised model of optimal tonal perception, the two contours are categorized as *low* (L), and *falling* (F).

In Figure 4, the early contour falls through two areas of new spectral information, one at stop release (onset of voicing) and one at vowel onset. According to the optimal tonal perception model, this contour is perceived as a *low* (L) in the vowel. This low also extends through the syllable coda. The second contour presents an interesting problem for the model. As it falls through the area of spectral change coinciding with the beginning of the syllable coda, the model would predict this contour to be perceived as *high* (H) in the vowel. This contour has traditionally been described as falling (cf. [3,4]) although Tumtavitikul [5] recently proposed describing the tone as a succession of level tones, H+L. In view of the fact that the amount of new spectral information is less at the beginning of the coda than at the beginning of the nucleus, it may, however, be possible that a falling contour here is perceptually coded as a true contour tone receiving the contour feature category *falling* (F).

Thus a crucial difference between the low tone and the falling tone in Thai is one of timing. Although there is also a difference in final F0 with the low tone ending somewhat lower

as indicated in Figures 4 and 5, the timing difference between the two tones is maximized by the falling tone's late fall in the syllable.

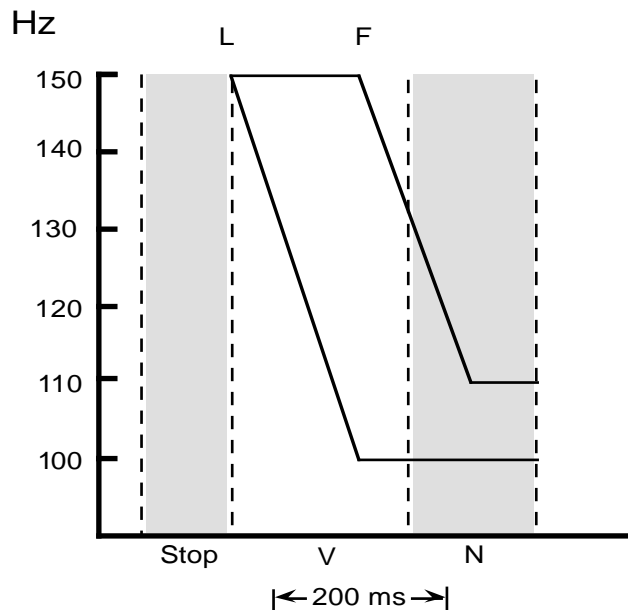


Figure 5: Stylized example of two falling tonal contours differing in timing through a CVC Thai syllable. According to the revised model of optimal tonal perception, the two contours are categorized as *low* (L), and *falling* (F).

3. PERCEPTION OF TONAL MOVEMENT BEFORE SILENCE

Further evidence of an increase in tonal movement perception in a sonorant consonant at a syllable coda comes from a study in which the influence of a silent interval on the preceding tonal contour was investigated [7]. In the study, listeners were asked to match a falling tonal movement in a vowel with a falling tonal movement in a vowel and the following sonorant consonant as illustrated in Figure 6. If listeners matched the contours based mostly on the falling movement in the vowel only, they would be expected to match the fall gradient resulting in the contour represented by the solid line in Figure 6 being matched with the contour represented by the lower dashed line. If listeners use information in both vowel and consonant, they would be expected to match the consonant final F0 resulting in the contour represented by the solid line in Figure 6 being matched with the contour represented by the upper dashed line. Two conditions were tested, one in which a CVCVC sequence served as the context for the tonal contour as in Figure 6, and the other in which a silent interval followed the second consonant producing a CVC context for the tonal contour as in Figure 7.

The results indicated that listeners used information in the consonant to a greater extent when the consonant was syllable final (before silence) as shown in Figure 7 than when

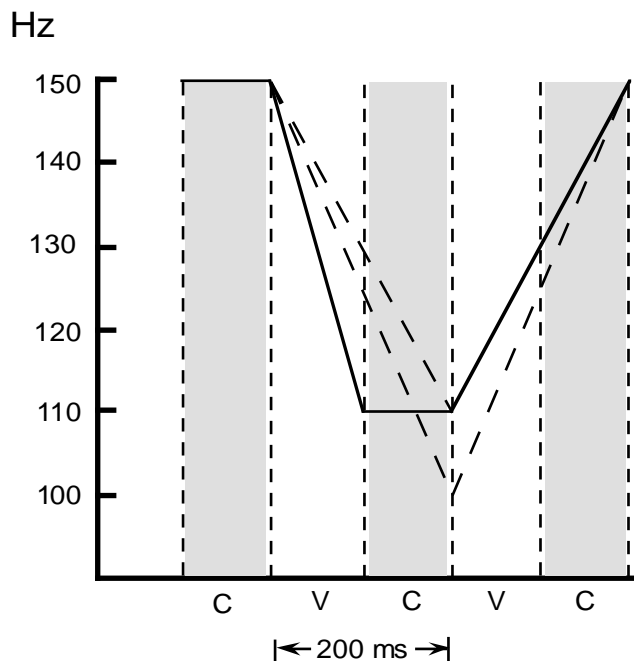


Figure 6: Stylized contours of stimuli used in [7]. Listeners had the task of matching the vowel only fall (solid line) with the VC-fall (dotted line) in a pause and a no-pause condition after the second consonant. In the no-pause condition shown in this figure, listeners matched the solid line to the lower dashed line demonstrating closer matching of the fall gradient in the first vowel.

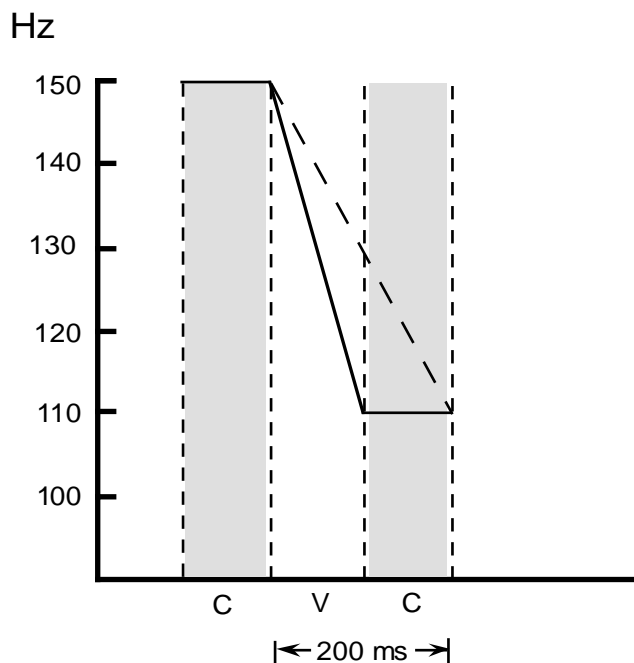


Figure 7: Stylized example of contours matched by listeners in the CVC context (pause condition) showing closer matching of the syllable-final F0.

the consonant was followed by a vowel as in Figure 6. If we assume that listeners perceive the CVCVC configuration as two syllables, then the natural syllable boundary would be before the second consonant. If the second consonant in the no-pause context comprises a syllable onset, then the results can be interpreted as greater tonal movement perception at the end of the syllable than at the onset of a new syllable.

4. CONCLUSIONS

The evidence presented in this paper leads to a reevaluation of the model of optimum tonal movement perception. The revision takes into account syllable structure. The area of maximum new spectral and intensity change occurring typically between syllable onset and syllable nucleus appears to be a crucial point for the timing of tonal movement. Movement through this area will be recoded as tonal levels as indicated in the earlier model. However, movement through the beginning of the syllable coda can be perceived as movement per se and described using movement features. This is consistent with the syllable rhyme as the domain of tone for example in Chinese as postulated in e.g. [8]. This revision will be put to further tests as perceptual evidence from more tonal languages becomes available.

5. REFERENCES

1. House, D. *Tonal Perception in Speech*, Lund University Press, Lund, 1990.
2. Hermes, D. *Timing of pitch movements and accentuation of syllables*, Manuscript no. 1143, Institute for Perception Research, Eindhoven, 1996.
3. Abramson, A. S. *The vowels and tones of standard Thai: acoustical measurements and experiments*, Indiana University Research Center in Anthropology, Folklore, and Linguistics (Publication Twenty), Bloomington, IN. 1962. Also Part II, *International Journal of American Linguistics* 28:2.
4. Gandour, J., Siripong Potisuk, Suvit Ponglorpisit and Sumalee Dechongkit. "Inter- and intraspeaker variability in fundamental frequency of Thai tones", *Speech Communication* 10, 355-372, 1991.
5. Tumtavitikul, A. "Tonal movements in Thai", *Proc. ICPhS 95*, Stockholm, 118-121, 1995.
6. House, D. and Svantesson, J-O. "Tonal timing and vowel onset characteristics in Thai", *Proc. Pan-Asiatic Linguistics, 4th International Symposium*, Bangkok, 104-113, 1996.
7. House, D. "Perception of prepausal tonal contours: Implications for automatic stylization of intonation," *Proceedings of Eurospeech '95*, 949-952, Madrid, 1995.
8. Howie, J.M. "On the domain of tone in Mandarin, Some Acoustical Evidence," *Phonetica* 30, 129-148, 1974.