

# Recognizing phrase and utterance as prosodic units in non-tonal dialects of Kammu

Anastasia Karlsson<sup>1</sup>, David House<sup>2</sup> and Damrong Tayanin<sup>1</sup>

<sup>1</sup>Department of Linguistics and Phonetics, Lund University

<sup>2</sup>Speech, Music and Hearing, KTH, Stockholm

## Abstract

*This paper presents a study of prosodic phrasing in a non-tonal dialect of Kammu, a Mon-Khmer language spoken in Northern Laos. Prosodic phrasing is seen as correlated with syntactic and informational structures, and the description is made referring to these two levels. The material investigated comprises sentences of different lengths and syntactic structures, uttered by seven male speakers. It is found that, based on prosodic cues, the distinction between prosodic utterance and prosodic phrase can be made. Prosodic phrase is signaled by a sequence of low + high pitch while the right edge of the prosodic utterance gets low pitch. This low terminal is replaced by a high terminal in expressive speech. The study is performed using elicited speech.*

## Introduction

Kammu, a Mon-Khmer language, has dialects with lexical tones (low and high lexical tone) and dialects with no lexical tones. Tones arose at a late stage of the language's development in connection with loss of the contrast between voiced and voiceless initial consonants in a number of dialects (Svantesson and House, 2006). There are no other phonological differences between toneless and tonal dialects and this makes the different Kammu dialects well suited for studying and comparing the use of phrase intonation.

In this paper we present an investigation of prosodic phrasing in the non-tonal dialects of Kammu. We concentrate on the use of pitch in signaling prosodic grouping. It is assumed that a spoken utterance can be prosodically signaled as one prosodic unit or divided into smaller prosodic units. We do not have any pre-assumptions about the types and numbers of prosodic units in Kammu. Instead we assume that, due to the elicited type of the material, the utterances are read as autonomous utterances, and it is interesting to find out whether the rightmost utterance boundary is signaled prosodically differently from the utterance in-

ternal boundaries. Due to the SVO word order Kammu typically places new information at the right edge of the utterance.

Words in Kammu are monosyllabic or sesquisyllabic (Svantesson and Karlsson, 2004). Sesquisyllabic words consist of one minor and one major syllable. The minor syllable has schwa as its nucleus. Schwa insertion is often absent in casual speech, but appears consistently in some types of singing (Lundström and Svantesson, 2008). There is also a phonological distinction between short and long vowels.

## Method

### Material

Material was collected in Laos in 2007. For the present investigation 16 sentences with different lengths and different syntactic structures were chosen. Kammu lacks a written script and informants were asked to translate the material from Lao to Kammu. Kammu speakers are bilingual with Lao being their second language. This resulted in somewhat different but still compatible versions of the target sentences. The resulting utterances were checked and transcribed by Damrong Tayanin who is a native speaker of Kammu.

Each target sentence was read three times by the speakers. A total of 226 utterances were investigated.

### Subjects

A total of nine speakers, seven men and two women were recorded. Their ages ranged from 14 to 57 years. For the present investigation seven speakers, all men, were chosen. They are labeled as S1, S3, S5, S6, S7, S8 and S9.

### Recording and analysis

The subjects were recorded with a portable Edirol R-09 digital recorder. Five of the speakers were recorded in quiet hotel rooms, S3 was recorded in his native village, and S7 was recorded at his home.

The recordings were analyzed using the Praat program. For each utterance, an  $f_0$  contour was extracted. Main pitch features such as turning points, lows and highs, relations between lows and highs specified by finding the lowest and the highest point in the  $f_0$  contours, and shapes of pitch gestures (fall, rise or level tone) were measured.

The observed tonal features were analyzed by referring to the syntactic and information structure of the utterances. Thus, the division of the sentences into syntactic phrases (NP, VP etc) and types of words (function- or lexical words) were matched to the tonal events observed. The placement of pragmatic focus was unambiguous from the sentence contents.

## Results

### Prosodic utterance

Regarding the signaling of the right utterance edge, the speakers can be divided into two groups. Speakers in the first group (S1, S5, S7, S9) have high pitch gestures utterance finally. Speakers in the second group (S3, S6, S8) have both low and high terminals. In the second group, S3 and S6 have a prevalence for low terminals and S8 has mostly high terminals.

In our previous investigation (Karlsson et al., 2007), we found two main types of focal accent in the non-tonal dialects. The material comprised recordings made by Kristina Lindell in the 1970's of one male speaker telling a folktale. Based on the contents, we assumed that the high focal accent was more expressive than the low one. The present material gives more support to our assumption about the different pragmatic load of the focal accents. Thus, we find two different gestures, a falling pitch and a high pitch in the same position in the utterance. This is a default position for focus and both gestures thus function as focal accent. The high focal accent differs from the low one by its expressive load. For instance, S3 is reserved when reading the written material. He uses high terminals only in some cases. On the other hand, in his spontaneous storytelling, S3 is very relaxed and uses high terminals instead of low ones.

A default pattern of a neutral utterance uttered as one prosodic unit is a declining  $f_0$  course with a low (falling) terminal. The low terminal tone signals the right boundary of a prosodic utterance. A prosodic phrase can not have a low terminal (more on this in the next

section), until the phrase can (syntactically and semantically) function as an autonomous utterance.

Speakers with high terminals usually reach the highest  $f_0$  value in the utterance at the final rise. This is typical even for utterances with multiple focal accents, such as listings of objects that a person owns. Figure 1 illustrates variants of the sentence *hmraŋ o dəŋ wɛt, traak o dəŋ wɛt* "A horse I'll buy and a buffalo I'll buy". In the top panel the  $f_0$  course of S3 is presented. The sentence is realised as two prosodic utterances with low terminals. The utterance boundaries are shown with arrows. In the bottom panel, the  $f_0$  course of S1 is presented. This speaker is expressive and has high boundary tones. He uttered the sentences as consisting of four prosodic phrases, each ending with a high tonal gesture (shown with arrows).

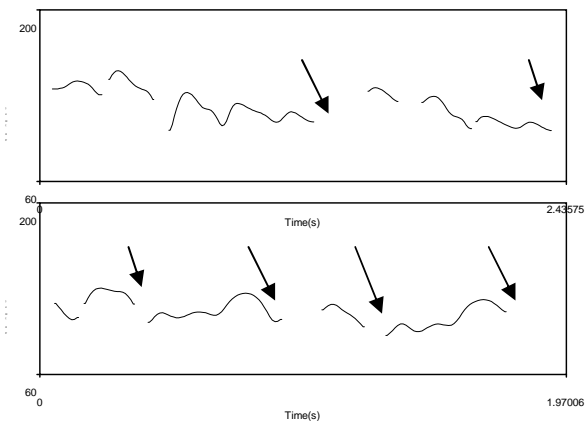


Figure 1. F0 courses of the sentence (its mutual variants) *hmraŋ o dəŋ wɛt, traak o dəŋ wɛt* "A horse I'll buy and a buffalo I'll buy" uttered by S3 (the upper panel) and S1 (the bottom panel).

### Prosodic phrase

When the utterance is divided into smaller prosodic units (named prosodic phrases here), it is signaled by a combination of low + high pitch. The basic prosodic phrase comprises two words with the first word getting low pitch and the second word getting high pitch. These two-word groups overlap with the syntactic grouping in the sense that prosodic grouping does not pose boundaries which are syntactically impossible. The prosodic phrasing pattern low + high can be supposed to reflect the placement of the focal accent at the right edge of a focused unit. This is, however, a subject for future study. The

right boundary of a prosodic phrase is signaled by a high pitch, and the right boundary of a prosodic utterance gets the highest  $f_0$  value in expressive speech. In Figure 2  $f_0$  courses of the utterance əə, o jəə phaən hmraŋ “Yes, I’ll kill the horse” (top panel) and of the utterance o jəə phaən hmraŋ too mee wæt knaay “I’ll kill the horse that you bought”, Speaker 9 (an expressive one). Both utterances are divided into three prosodic phrases; their final high tonal gestures are shown with arrows. The word “horse” is utterance final in the top panel and gets the highest  $f_0$  values. In the bottom panel the same word is in utterance medial position and does not get the highest  $f_0$  values; it is the utterance final word that gets in most cases the highest  $f_0$  value. The high pitch on “horse” in the first case is the utterance-final high, while in the second case it is a phrase-final high.

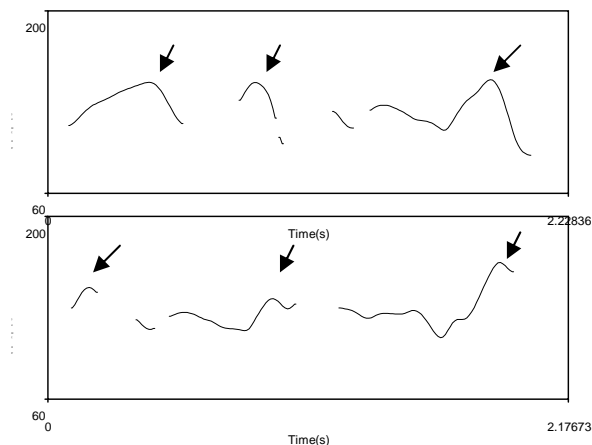


Figure 2.  $f_0$  courses of the utterances əə, o jəə phaən hmraŋ “Yes, I’ll kill the horse” (top panel) and of the utterance o jəə phaən hmraŋ too mee wæt knaay “I’ll kill the horse that you bought”, S9.

The basic phrasal pattern low + high can be modified due to a number of factors. Thus, when more than two words are included in a prosodic phrase, the non-final words get low pitch and the last one is marked with high pitch.

One-word phrases occur. They get only high pitch, and there are no prosodic phrases with only low pitch. One-word phrases are typically the utterance-initial word, thus the pronominal o “I” is often phrased as a one-word phrase. Adverbials, placed utterance initially and syntactically being self-sufficient units, are always marked by a high pitch and phrased as a one-word phrase. Figure 3 illustrates the sentence sgii gəə waar “Today it is sunny” uttered as sgii waar by S1. Both words get high pitch and the

utterance is by virtue of this divided into two prosodic phrases.

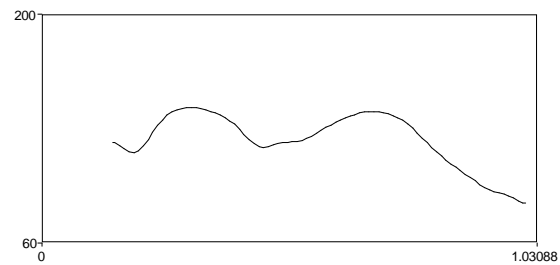


Figure 3.  $f_0$  courses of the utterance sgii waar “Today it is sunny” uttered by S1.

Clusters of two high tones within the same prosodic phrase occur but only at utterance final position. It happens in cases when the last two words are pragmatically highlighted. The pattern low + high is the preferred one, and no more than two high tones are found in the same prosodic phrase.

## Final word

In elicited Kammu speech the division into smaller and larger prosodic units, the prosodic phrase and the prosodic utterance, is clearly signaled by prosodic means. The next step is to test this description for spontaneous speech. In a previous study (Karlsson et al., 2007) we could not give any prosodic cues to discriminate between the prosodic phrase and the prosodic utterance and chose to operate only with the prosodic phrase. Given that a prosodic phrase cannot have a low tone finally, the previous material could be reanalyzed.

## Acknowledgements

This work has been carried out within the research project, “Separating intonation from tone” (SIFT), supported by The Bank of Sweden Tercentenary Foundation.

## References

- Karlsson A.M., House D., Svantesson J-O. and Tayanin D. (2007). Prosodic phrasing in tonal and non-tonal dialects of Kammu. Proceedings of the 16th International congress of phonetic sciences, 1309-1312.
- Lundström, Håkan & Svantesson, Jan-Olof. (2008). *Hrl̥i̯ singing and word-tones in Kammu*. Working Papers 53: 117–131.

Lund University: Dept of Linguistics and  
Phonetics.

Svantesson, J-O., & House, D. (2006). Tone  
production, tone perception and Kammu  
tonogenesis. *Phonology*, 23, 309-333.

Svantesson, J-O & Karlsson, A.M. (2004). Mi-  
nor Syllable Tones in Kammu. *Proceedings  
of International Symposium on Tonal As-  
pects of Languages, Beijing 28-30 March,*  
177-180.