Prosodic features of Kammu tonal and non-tonal dialects: read and spontaneous speech

Karlsson, Anastasia

Published in:
Linguistic Circle for the Study of Eastern Eurasian Languages

2011

The 18th Meeting of the Linguistic Circle for the Study of Eastern Eurasian Languages
Aoyama Gakuin University, Tôkyô, Japan
20 February, 2011

Prosodic Features of Kammu Tonal and Non-tonal Dialects: Read and Spontaneous Speech

Anastasia Mukhanova KARLSSON
Center for Languages and Literature, Lund University

1 Introduction

A problem which is often encountered in work on intonation in tone languages is the difficulty in separating tone from intonation, which may lead to uncertainty about what in the intonation contour is the contribution of lexical tones and what is utterance intonation. Kammu is a Mon-Khmer language. It is spoken by some 600,000 people mainly in Northern Laos, but also in adjacent areas of Vietnam and Thailand. There are more than 1,000 speakers in the southern part of Yunnan province in China. One of the main dialects of this language is a tone language of the "East Asian" type with (high or low) tone on each syllable, while the other main dialect lacks lexical tones.

The origin of the tones of the tonal dialect is due to the development of high pitch in vowels following a voiceless consonant and low pitch in vowels following a voiced consonant, and the subsequent merger of voiceless and voiced consonants into the unmarked member of the pair, voiceless for stops and voiced for sonorants. Thus, *puuc* 'to undress' became *púuc* (high tone) in the tonal dialect and *buuc* 'wine' became *pùuc* (low tone). The non-tonal dialect kept the original forms unchanged. Other differences, phonological, morphological or syntactic, between the dialects are marginal, and speakers of different dialects understand each other without difficulty (Svantesson 1983, Svantesson & House 2006).

Kammu gives us a rare opportunity to use a novel approach to the problem of describing intonation in tone languages by analyzing a language that has developed lexical tones rather recently, from the point of view of language history. We can investigate the tonal contours in nearly identical sentences produced by speakers of dialects with and without tones.

Our working hypothesis is that the existence of lexical tones in a language imposes restrictions on the possibility to express important communicative functions, such as focussing and phrasing, tonally.

2 Data collection and speech material

Data collection was carried out during a recording trip to Laos (November 2007) and to Thailand (March 2008). A total of 24 speakers (11 women and 13 men ranging in age from 14 to 72) were recorded using a portable Edirol R-09 digital recorder and a lapel microphone. Ten are speakers of non-tonal dialects and 14 are speakers of tonal dialects. The recordings comprise about 15 hours of speech. For acoustic measurements, the *Praat* program was used.

The speech material was comprised of 47 read sentences and was composed in order to control for lexical tone, to elicit focus in different positions and to elicit phrasing. Kammu speakers are bilingual with Lao or Thai being their second language. Kammu lacks script and informants were asked to translate the material from Lao or Thai to Kammu. This sometimes resulted in slightly different but still compatible elicitations of the target sentences.

The non-scripted material comprises elicitation of focus using pictures, spontaneous account of growing and preparing rice, and of three series of drawings, comprising five pictures each.

The recorded material was checked, transcribed and translated into English by Damrong Tayanin, who is a native speaker of Kammu.

3 Prosodic units in Kammu: read speech

An utterance is always marked on its right edge by a rising-falling gesture. The gesture is truncated when word ends with a voiceless segment and only the rise is realised. Longer utterances are divided into several prosodic phrases marked by high boundary tone on their right edges. The utterance final phrase (word) gets highest f0 values. Phrasing is illustrated in Figures 1 and 2. The two sentences *ʔòʔ ʔàh tráak, ʔàh sɨ́aŋ, ʔàh hjíar* (tonal dialect)*/ʔoʔ ʔah traak, ʔah sɨaŋ, ʔah hʔiar* (non-tonal dialect) “I have a buffalo, a pig and a chicken” and *ʔòʔ ʔàh hmràŋ, ʔàh mɛ̀ɛw, ʔàh prùul* (tonal dialect)*/ʔoʔ ʔah hmraŋ, ʔah mɛɛw, ʔah bruul* (non-tonal dialect) “I have a horse, a cat and a badger” are divided in three phrases with high right boundary tones and the highest rise in utterance final position.

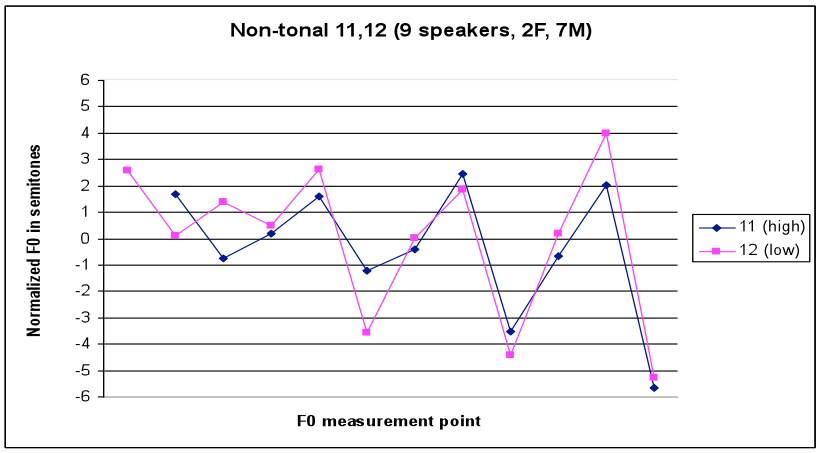

Figure 1. *Illustration of phrasing. Normalised F0 measurement points from nine speakers of the non-tonal dialect. Lexical tone in parentheses refers to the tonal dialect.*

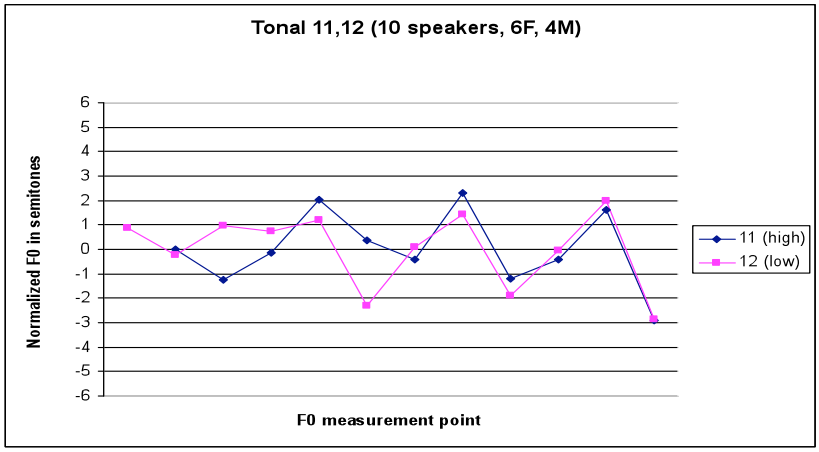

Figure 2. *Illustration of phrasing. Normalised F0 measurement points from ten speakers of the tonal dialect. Lexical tone is indicated in parentheses.*

Thus we find phonetic evidence for two prosodic units in Kammu, prosodic phrase and prosodic utterance. The right boundary tone of the prosodic phrase is called phrase boundary and the rightmost boundary tone of the prosodic utterance is called sentence accent. At narrow focus (i.e. at pragmatic highlighting of a word within an utterance), focus is realised by strengthening phrase boundary tones. The sentence accent is an obligatory marker of the utterance. In our controlled material (Karlsson et al. 2010) we used utterances of the type "This is [noun + adjective]" and changed the place of focus within the noun phrase. Speakers of both dialects have a pronounced gesture on the utterance final word, both with final and non-final focus. The final rise is larger when focus is on the final word, see Fig.3.

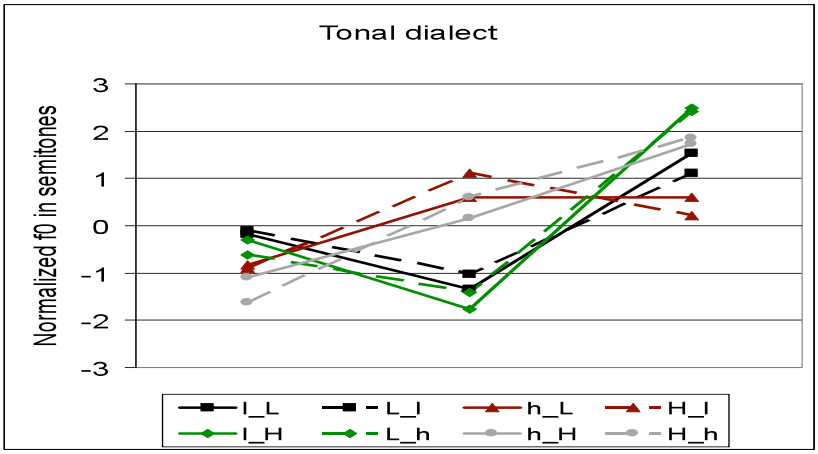

Figure 3. *Normalised F0 measurement points for the tonal dialect. Measurement points are F0 maxima in the three last words (copula + [noun+adjective] of sentences of type "This is [noun+adjective]. Lexical tone and focal status are indicated for the last two words. l=low lexical tone, h=high tone, L/H=focus.*

The rising sentence accent is pervasive in tonal dialects until it jeopardises the identity of lexical tones. The pattern is broken in favour of lexical tones. When the tone combination within the noun phrase in "This is [noun + adjective]" is [l + h], [l + l], [h + h] the final rise is realised at both final and non-final focus. The greatest focus effect is found at final focus in [l + h] condition. The [h + l] condition is in direct conflict with the rising sentence accent and the pattern is broken in favour of the lexical tones.

The intonational structure described is found for both non-tonal and tonal speakers. In tonal dialect this can be observed when intonation does not conflict with lexical tones, i.e. when we have identical lexical tones, see Fig.4. Unfortunately, such utterances are very rare in our material.

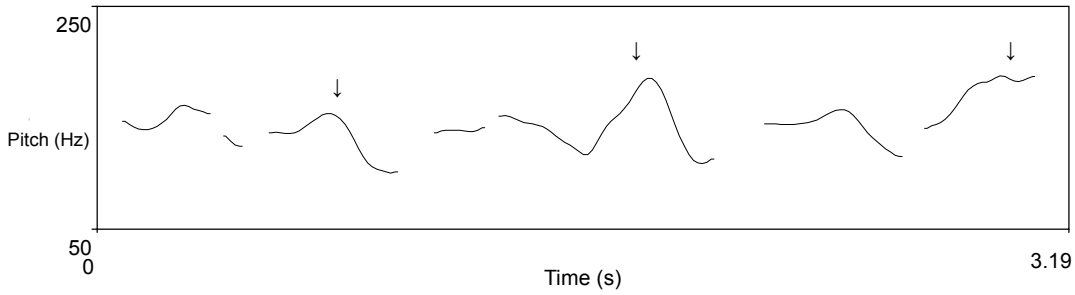

Figure 4. *Phrasing pattern in spontaneous speech, a male tonal speaker. The utterance is [wàaj híc kàaj kùum] [wàaj kùum kàaj rùŋ] [wàaj rùŋ kàaj pə̀ʔ] "After pound then winnow after winnow then steam after steam then eat". All words except the second one have low lexical tone. The utterance is prosodically divided in three prosodic phrases (indicated by the square brackets in the transcription above) with final high phrase tones and utterance final high sentence accent, all shown by an arrow.*

4 Spontaneous speech

4.1 Prosodic structuring of spontaneous speech: background

As investigation of read speech indicates, the tonal and non-tonal Kammu dialects exhibit the same underlying intonational system. The read speech also provides a clear picture about the functions of the identified units. The next stage is to test how these features are applied in spontaneous speech. Spontaneous speech is challenging because of high variation between speakers and poor possibilities for controlled studies. One part of our recordings consists of spontaneous account of rice growing, from the beginning of field works until the rice is on the table as a meal, and eaten. All speakers are well acquainted with this process and their descriptions are very similar. Thus, at least on the lexical level we got fairly homogenous spoken texts.

Discourse has an internal structure: information is chunked into larger units with some internal coherence, "topics" (Swerts 1994). This structuring is achieved by morpho-syntactic and prosodic means. Lehiste (1975) and Bruce (1982) among others have shown that spoken paragraphs have a characteristic prosodic supra-structure as a result of demarcating topics, feed-back seeking, and turn taking. Bruce proposes that an utterance that breaks the prosodic pattern in some way is the strongest cue for a new topic. I follow his way of analysing discourse: the morpho-syntactic and the prosodic structures are described separately and then matched to each other. This is to avoid seeking for direct relations between the two structures.

4.2 Informational structuring of rice-stories

The descriptions of the growing and preparing of rice contains the following events: seeking for a place to make a field, burning the field, sowing of rice, weeding, harvest, putting rice into a barn, drying, pounding, winnowing, soaking, steaming, (putting rice on the table), eating.

Structuring of new and old information is achieved in the same way by all speakers: new information is placed at the end of the utterance, it is then repeated in the next utterance and is followed by new information. The informational structuring is [anchor + new_1] [old new_1 + new $_2$] [old new $_2$ + new $_3$]. The new information becomes an anchor point (old information) in the next utterance. There are thus a lot of repeated words in the speakers' monologues. An example is (only key words are included): "Before there is rice we have to clear the field. After clearing we burn the field[1]. After burning we sow." See also the text in Fig. 4.

[1] Kammu people practice slash-and-burn agriculture.

The text can be then seen as a list of successive events. Some speakers give a lot of additional information about the events.

4.3 Prosodic structuring of rice-stories

For the present pilot study, seven tonal and four non-tonal speakers were chosen. They showed less disfluences than others.

We know from read speech that phrasing and focussing are achieved by right rising boundary tones. The prosodic units prosodic phrase and prosodic utterances are also found for spontaneous speech. In spontaneous speech there is a high tendency for usptepping of phrase boundary tones within utterances, with highest f0 values utterance-finally. This is similar to patterns found in enumeration in read speech (Figures 1 and 2). What is more interesting is that utterance final rises (sentence accents) are also upstepped in relation to each other. Thus, there is an upstepping of phrase boundaries within utterances and upstepping of sentence accents between utterances, see Figures 5 and 6. In the figures, only final sentence accents are shown. Prosodic events between the accents are not shown and it has to be kept in mind that several words and phrases occur between the sentence accents displayed in the figures.

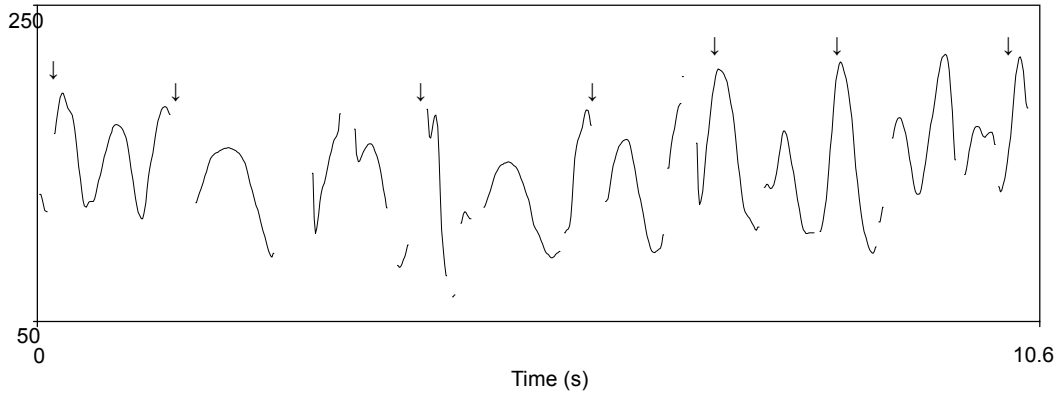

Figure 5. *Structuring into prosodic paragraphs, non-tonal speaker. Only sentence accents are shown. Arrows show sentence accents in final position in prosodic paragraphs. The words shown are [psɨm] [cməəl, hreʔ] [hɛɛl, həət, kiaw, cʔoʔ] [hntaar, hic] [gual, tiʔ, guum] [ɟam, ruŋ] [hmpɨal, muut, ʔɛɛp niʔ] [plant] [sow, field] [weed, harvest, harvest, barn] [dry, pound] [rice-mortar, hand, winnow] [soak, steam] [turn, knead, rice-basket that].*

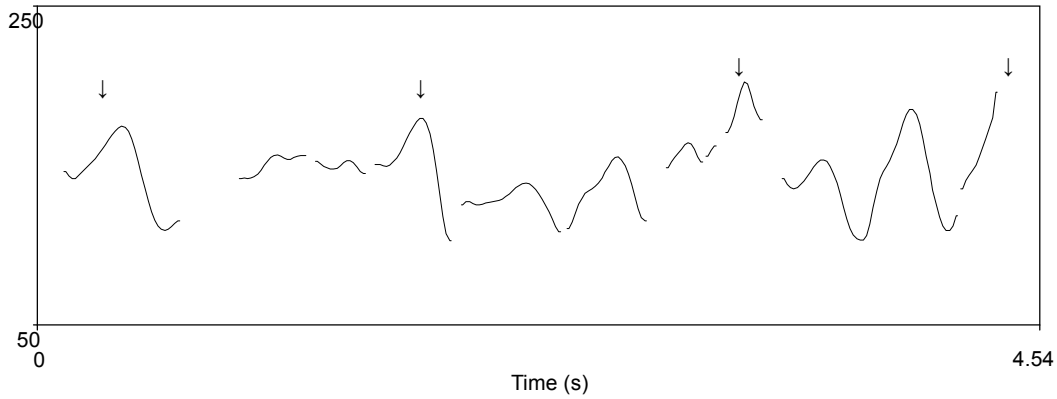

Figure 6. *Structuring into prosodic paragraphs, tonal speaker. Only sentence accents are shown. Arrows show sentence accents finally in prosodic paragraphs. The words shown are [rɨam$_H$] [pɔɔk$_H$, reʔ$_H$, puur$_H$] [cmɔɔl$_L$, hɛɛl$_H$, hic$_H$] [kuum$_L$, ruŋ$_L$, pəʔ$_L$] [clear$_H$] [burn$_H$, field$_H$, burn$_H$] [sow$_L$, weed$_H$, pound$_H$] [winnow$_L$, steam$_L$, eat$_L$]*

The upstepping of sentence accents takes place until a certain utterance and then the same pattern begins again. This pattern is somewhat more difficult to observe for tonal speakers when the final sentence accent is on a word with another lexical tone than on the preceding sentence-accented word. When we have sequences of words with only high or low lexical tones in these positions, the pattern comes up more clearly. The repeated upstepping of sentence accents indicates that utterances are organised to form larger speech units, speech paragraphs.

4.4 Interplay between prosodic and discourse structures

To find out if and how these prosodic paragraphs are connected with discourse structure they are matched with the lexical level, as shown in (1). In the brackets only utterance-final words within each prosodic paragraph are shown. Each bracket represents a prosodic paragraph. They are prosodically marked by upstepping of sentence accents and the highest F0 value at the last word. Lexical tones are given for tonal speakers. The question mark is written in unclear cases, because of influence from lexical tones on the intonation. For most of the speakers only parts of their monologues were chosen for the investigation. The prosodic paragraphs comprise different numbers of words, not shown here.

(1) Parts of discourse realised as separate prosodic paragraphs.

Speaker 1 (non-tonal)

[clear] [burn, sow] [sow, come out, rice, time arrives, ripe] [harvest, cut, dry, pound]

Speaker 2 (non-tonal)

[field-house] [burn] [sow, weed] [pound] [winnow, grain, soak] [steam, rice]

Speaker 3 (non-tonal)

[clearing trees, burn, burn, burn][2]

Speaker 4 (non-tonal)

[plant, sow, field] [weed, harvest, harvest, barn] [dry, pound] [rice-mortar, hand, winnow] [soak, steam] [turn kneed, rice basket]

Speaker 5 (tonal)

[clear^{H}] [burn^{H}, burn^{H}] (sow^{L})? [weed^{H}, harvest^{H}, beat^{H}] (barn^{L})? [pound^{H}] [rice grain^{H}, winnow^{L}] [soak^{L}, steam^{L} rice^{L}]

Speaker 6 (tonal)

[clear^{H}] [burn^{L}, burn^{H}] (sow^{L})? [weed^{H}, harvest^{H}] [beat^{H}] (barn^{L})? [pound^{H}] [winnow^{L}] [soak^{L} steam^{L}] [rice^{L}]

Speaker 7 (tonal)

[clear^{H}] [burn^{H}, burn^{H}] [house^{L}, sow^{L}, grow^{L}, harvest^{H}, return^{H}] [dry^{H}, pound^{H}] [winnow^{L}] [soak^{L}] [steam^{L},rice^{L}, eat^{L}]

Speaker 8 (tonal)

[field^{H}, clear^{H}, burn^{L}, burn^{H}] [sow^{L}, weed^{H}, weed^{L}, harvest^{H}] [pound^{H}] [winnow^{L}, soak^{L}, steam^{L} rice^{L}, rice^{L}, eat^{L}]

Speaker 9 (tonal)

[clear^{H}] [burn^{H}, sow^{L}, weed^{H}, harvest^{H}, pound^{H}] [steam^{L}, eat^{L}]

Speaker 10 (tonal)

[clear^{H}] [burn^{H}, field^{H}, burn^{H}] [sow^{L}, weed^{H}, harvest^{H}, pound^{H}] [winnow^{L}] [steam^{L}, eat^{L}]

Speaker 11 (tonal)

[clear^{H}] [burn^{L}, burn^{H}] [sow^{L}, weed^{H}, bud^{L}, put out ear^{H}, ripe^{L} ()]? [pound^{H}] [harvest^{H}, pound^{H}] [winnow^{L}, rice^{L}]

There is no one-to-one overlapping between speakers but similarities in the interplay between prosodic and discourse structures are striking. Thus, the prosodic paragraphs correspond to the following four main discourse topics: [clear the field]$_1$ [burn the field]$_2$ [[[weed] harvest] [pound]]$_3$ [[winnow] [eat]]$_4$. Some speakers mark more topics, but the main four topics given

[2] Three different verbs are used to show successive stages in burning. For this speaker only a short part of his monologue was chosen.

here are always marked prosodically. The topic [pound] gets the highest F0 values for all speakers except one. The high level can be connected to the kind of activity described: the pounding demands hard beating and it may be reflected in the extra rising of F0 on this word. The topics include several activities mainly as follows: [clear the field] includes [find the field and clear it, dry the felled trees], [burn the field] includes [burn, (burn) and burn again], [pound] includes [sow, weed, ripe, harvest, put into the barn, dry], [eat] includes [winnow, steam, put on the table, eat]. Some speakers give many more details. The last topic about rice is often produced with downstepped tones and a reduced F0 range.

5 Conclusion and discussion

The present investigation of spontaneous speech in dialects of Kammu is based on the assumption about the same underlying intonational structure in tonal and non-tonal dialects. The assumption is proved by acoustic investigations of Kammu read speech (Karlsson et al. 2007, Svantesson et al. 2009, House et al. 2009, Karlsson et al. 2010). It is found that the underlying intonation is the same for the two types of dialects, occurrence of lexical tones being the reason for phonetic differences between the dialects. The prosodic structure and its functions identified for read speech is the base for the proposed investigation of spontaneous material. Spontaneous monologues allow observing relations between utterances. It is found that utterances are prosodically organised into speech paragraphs that coincide fairly well with shifting of related topics in monologues. Thus we find discourse topics with internal coherence on both prosodic and thematic levels. Kammu speakers mark the coherent topic by upstepping of sentence accents and demarcate the topic shift by the highest tonal rise at the end of the topic. This is in some way similar to what was found for Swedish: the shift of topic is often marked by a final high F0 (Bruce 1998). However, this gesture is optional in Swedish. In Kammu the high rise at topic shift is structurally dependent: the available prosodic structure is used for discourse purposes by strengthening of utterance boundaries at topic shift. Previously (House at al. 2009) we described the sentence accent as multifunctional, signalling utterance boundary, focus and speaker engagement. Based on spontaneous speech we can now add a fourth function of sentence accent, signalling of topic shift.

The present description is the first step into the analysis of spontaneous Kammu speech. We need to quantify the description if possible and to investigate other types of discourse.

Acknowledgments

The work reported in this paper has been carried out within the research project, "Separating intonation from tone" (SIFT), supported by the Bank of Sweden Tercentenary Foundation with additional funding from Crafoordska stiftelsen.

Bibliography

Bruce, G. (1982). Textual aspects of prosody in Swedish. *Phonetica*, 39. 274-287.

Bruce, G. (1998). *Allmän och svensk prosodi*. Praktisk lingvistik 16. Lund: Department of Linguistics and Phonetics, Lund University.

House, D., Karlsson, A., Svantesson, J. & Tayanin, D. (2009). The phrase-final accent in Kammu: effects of tone, focus and engagement. *Proceedings of Interspeech 2009*, Brighton, 2439-2442.

Karlsson, A., House, D., Svantesson, J. & Tayanin, D. (2010). Influence of lexical tones on intonation in Kammu. *Proceedings of Interspeech 2010*, Makuhari, Japan, 1740-1743.

Karlsson, A., House, D., Svantesson, J. & Tayanin, D. (2007). Prosodic phrasing in tonal and non-tonal dialects of Kammu. *Proceedings of 16th International congress of phonetic sciences*, Saarbrücken, 1309-1312.

Lehiste, I. (1975). The phonetic structure of paragraphs. In A. Cohen and S. Nooteboom (eds.), *Structure and Process in Speech Perception*, 195-206. Berlin: Springer-Verlag.

Svantesson, J-O. (1983). *Kammu phonology and morphology*. Lund: Gleerup.

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

Svantesson, J-O, House, D., Karlsson, A. & Tayanin, D. (2009). Reduplication with fixed tone pattern in Kammu. *Proceedings of Fonetik 2009*, 82-84. Stockholm University.

Swerts, M. & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, 37 (1), 21-43.