

On utterance-final intonation in tonal and non-tonal dialects of Kammu

David House¹, Anastasia Karlsson², Jan-Olof Svantesson², Damrong Tayanin²

¹ Dept of Speech, Music and Hearing, CSC, KTH, Stockholm, Sweden

² Dept of Linguistics and Phonetics, Centre for Languages and Literature, Lund University, Sweden

Abstract

In this study we investigate utterance-final intonation in two dialects of Kammu, one tonal and one non-tonal. While the general patterns of utterance-final intonation are similar between the dialects, we do find clear evidence that the lexical tones of the tonal dialect restrict the pitch range and the realization of focus. Speaker engagement can have a strong effect on the utterance-final accent in both dialects.

Introduction

Kammu, a Mon-Khmer language spoken primarily in northern Laos by approximately 600,000 speakers, but also in Thailand, Vietnam and China, is a language that has developed lexical tones rather recently, from the point of view of language history. Tones arose in connection with loss of the contrast between voiced and voiceless initial consonants in a number of dialects (Svantesson and House, 2006). One of the main dialects of this language is a tone language with high or low tone on each syllable, while the other main dialect lacks lexical tones. The dialects differ only marginally in other respects. This makes the different Kammu dialects well-suited for studying the influence of lexical tones on the intonation system.

In previous work using material gathered from spontaneous storytelling in Kammu, the utterance-final accent stands out as being especially rich in information showing two types of focal realizations depending on the expressive load of the accent and the speaker's own engagement (Karlsson et al., 2007). In another study of the non-tonal Kammu dialect, it was generally found that in scripted speech, the highest F0 values were located on the utterance-final word (Karlsson et al., 2008). In this paper, we examine the influence of tone, focus and to a certain extent speaker engagement on the utterance-final accent by using the same scripted speech material recorded by speakers of a non-tonal dialect and by speakers of a tonal dialect of Kammu.

Data collection and method

Recordings of both scripted and spontaneous speech spoken by tonal and non-tonal speakers of Kammu were carried out in November, 2007 in northern Laos and in February, 2008 in northern Thailand. 24 speakers were recorded ranging in age from 14 to 72 years.

The scripted speech material was comprised of 47 read sentences. The sentences were composed in order to control for lexical tone, to elicit focus in different positions and to elicit phrasing and phrase boundaries. Kammu speakers are bilingual with Lao or Thai being their second language. Since Kammu lacks a written script, informants were asked to translate the material from Lao or Thai to Kammu. This resulted in some instances of slightly different but still compatible versions of the target sentences. The resulting utterances were checked and transcribed by one of the authors, Damrong Tayanin, who is a native speaker of Kammu. The speakers were requested to read each target sentence three times.

For the present investigation six of the 47 read sentences were chosen for analysis. The sentences are transcribed below using the transcription convention for the tonal dialect.

- 1) nàa wèet hmràŋ
(she bought a horse)
- 2) nàa wèet hmràŋ yǐm
(she bought a red horse)
- 3) téek pháan tráak
(Teek killed a buffalo)
- 4) téek pháan tráak yían
(Teek killed a black buffalo)
- 5) Ò àh tráak, àh sǎŋ, àh hyár
(I have a buffalo, a pig and a chicken)
- 6) Ò àh hmràŋ, àh mǝew, àh prùul
(I have a horse, a cat and a badger)

Sentences 1 and 2 contain only words with a low lexical tone while sentences 3 and 4 contain only words with a high lexical tone. Sentences 2 and 4 differ from 1 and 3 only in that they end with an additional color adjective following the noun (red and black respectively). Sentences 2 and 4 were designed to elicit focal accent on the final syllable. Sentences 5 and 6 convey a listing of three nouns (animals). The nouns all have high lexical tone in sentence 5 and low lexical tone in sentence 6.

Of all the speakers recorded, one non-tonal speaker and four tonal speakers were excluded from this study as they had problems reading and translating from the Lao/Thai script. All the other speakers were able to fluently translate the Lao/Thai script into Kammu and were included in the analysis. Thus there were 9 non-tonal speakers (2 women and 7 men) and 10 tonal speakers (6 women and 4 men) included in this study. The speakers ranged in ages from 14 to 72.

The subjects were recorded with a portable Edirol R-09 digital recorder and a lapel microphone. The utterances were digitized at 48 KHz sampling rate and 16-bit amplitude resolution and stored in .wav file format. Most of the speakers were recorded in quiet hotel rooms. One speaker was recorded in his home and one in his native village.

Using the WaveSurfer speech analysis program (Sjölander and Beskow, 2000) the waveform, spectrogram and fundamental frequency contour of each utterance was displayed. Maximum and minimum F0 values were then annotated manually for successive syllables for each utterance. For sentences 1-4, there was generally very little F0 movement in the syllables leading up to the penult (pre-final syllable). Therefore, measurements were restricted to maximum F0 values on the pre-penultimate syllables, while both maximum and minimum values were measured on the penultimate and ultimate syllables. For sentences 5 and 6, the initial syllable and the three nouns often showed considerable F0 movement and were therefore provided with both maximum and minimum F0 values. For the other syllables which showed little F0 movement, measurements were again restricted to F0 maximum values.

To be able to compare the different speakers, the measured F0 values in Hz were converted to semitones (St) and then normalized in accordance with Fant and Kruckenberg (2004). A fixed semitone scale is used where the unit St is defined by

$$St = 12[\ln(Hz/100)/\ln 2] \quad (1)$$

which results in a reference value of St=0 semitones at 100 Hz, St=12 at 200 Hz and St=-12 at 50 Hz. Normalization is performed by subtracting each subject's average F0 in St (measured across the three utterances of each target sentence) from the individual St values.

Results

Plots for sentences 1-4 showing the F0 measurement points in normalized semitones are presented in Figure 1 for the non-tonal dialect and in Figure 2 for the tonal dialect. Alignment is from the end of the sentences. Both dialects exhibit a pronounced rise-fall excursion on the utterance-final syllable. There is, however, a clear difference in F0 range between the two dialects. The non-tonal dialect exhibits a much wider range of the final excursion (6-7 St) than does the tonal dialect (3-4 St).

If we are to find evidence of the influence of focus on the F0 excursions, we would expect differences between sentence pairs 1 and 2 on the one hand and 3 and 4 on the other hand where the final words in 2 and 4 receive focal accent. In the non-tonal dialect (Figure 1) we find clear evidence of this difference. In the tonal dialect there is some evidence of the effect of focal accent but the difference is considerably reduced and not statistically significant (see Table 1).

Table 1. Normalized maximum F0 means for the sentence pairs. 1st and 2nd refers to the F0 max for the 1st and 2nd sentence in the pair. Diff is the difference of the normalized F0 values for the pair, and the ANOVA column shows the results of a single factor ANOVA. Measurement position in the utterance is shown in parenthesis.

Sentence pair	1 st	2 nd	diff	ANOVA
Non-tonal				
1-2 (final)	1.73	3.13	-1.40	p<0.05
3-4 (final)	1.49	2.84	-1.35	p<0.05
5-6 (1 st noun)	1.60	2.63	-1.03	p<0.05
5-6 (2 nd noun)	2.47	1.86	0.61	n.s.
5-6 (final)	2.02	4.00	-1.98	p<0.05
Tonal				
1-2 (final)	0.56	0.76	-0.20	n.s.
3-4 (final)	0.76	1.54	-0.78	n.s.
5-6 (1 st noun)	2.01	1.22	0.79	p<0.05
5-6 (2 nd noun)	2.33	1.42	0.91	p<0.05
5-6 (final)	1.59	1.98	-0.39	n.s.

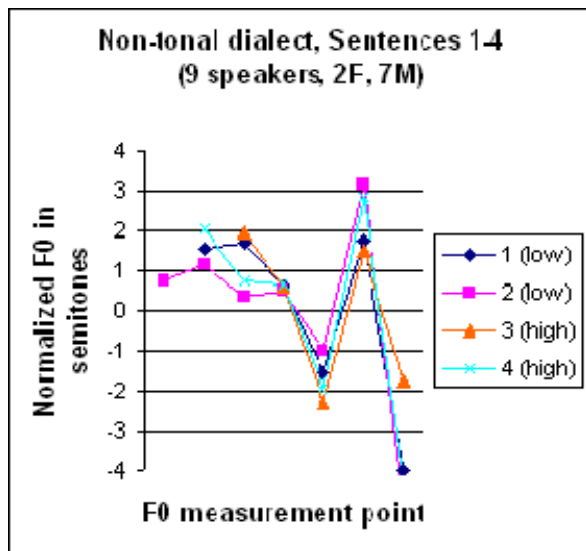


Figure 1: Normalized F0 measurement points for sentences 1-4 from nine speakers of the non-tonal dialect. Lexical tone in parenthesis refers to the tonal dialect.

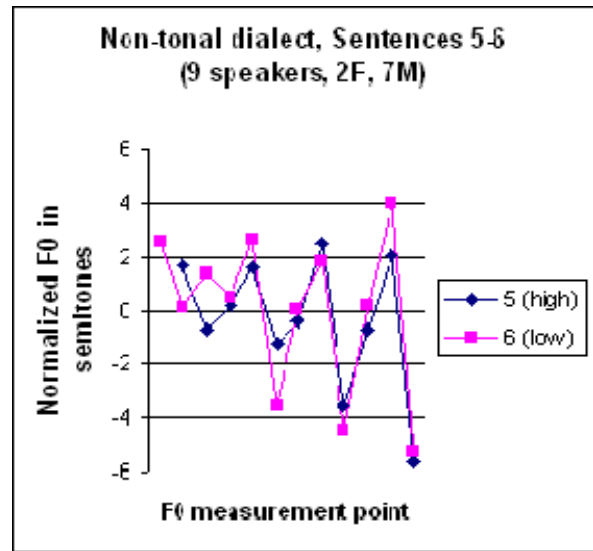


Figure 3: Normalized F0 measurement points for sentences 5-6 from nine speakers of the non-tonal dialect. Lexical tone in parenthesis refers to the tonal dialect.

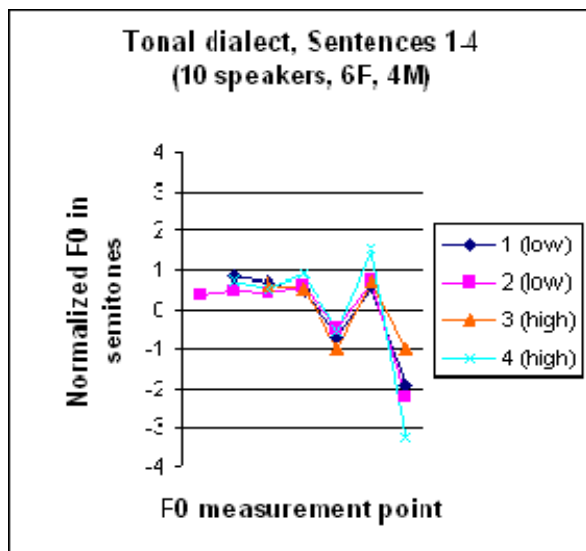


Figure 2: Normalized F0 measurement points for sentences 1-4 from ten speakers of the tonal dialect. Lexical tone is indicated in parenthesis.

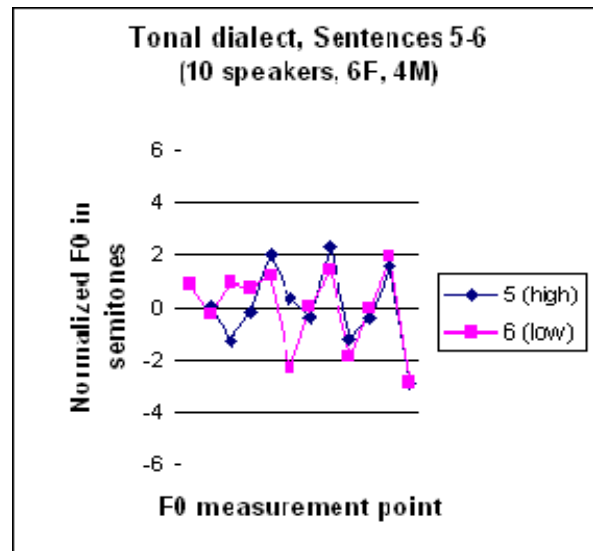


Figure 4: Normalized F0 measurement points for sentences 5-6 from ten speakers of the tonal dialect. Lexical tone is indicated in parenthesis.

Plots for sentences 5 and 6 are presented in Figure 3 for the non-tonal dialect and in Figure 4 for the tonal dialect. Alignment is from the end of the sentences. Both dialects show a similar intonation pattern exhibiting rise-fall excursions on each of the three nouns comprising the listing of the three animals in each sentence. As is the case for sentences 1-4, there is also here a clear difference in F0 range between the two dialects. The non-tonal dialect exhibits a much wider overall range (6-9 St) than does the tonal dialect (3-4 St).

The nouns in sentence 5 have high tone in the tonal dialect, while those in sentence 6 have

low tone. A comparison of the F0 maximum of the nouns in the three positions for the non-tonal dialect (Figure 3) shows that the noun in the first and third position has a higher F0 in sentence 6 than in sentence 5 but in the second position the noun has a higher F0 in sentence 5 than in sentence 6. A comparison of the F0 maximum of the nouns in the three positions for the tonal dialect (Figure 4) shows that the F0 maximum for the high tone (sentence 5) is higher than the low tone (sentence 6) in the first and second position but not in the third position.

Discussion

In general terms, the results presented here show considerable similarities in the basic intonation patterns of the utterance-final accent in the two dialects. There is a pronounced final rise-fall excursion which marks the end of the utterance. The presence of lexical tone in the tonal dialect does, however, restrict the intonation in certain specific ways. The most apparent overall difference is found in the restricted range of the tonal dialect. This is an interesting, but perhaps not such an unexpected difference. As the tonal differences have lexical meaning for the speakers of the tonal dialect, it may be important for speakers to maintain control over the absolute pitch of the syllable which can result in the general reduction of pitch range.

The lexical tones and the reduction of pitch range seem to have implications for the realization of focal accent. In the non-tonal dialect, the final color adjectives of sentences 2 and 4 showed a much higher F0 maximum than did the final nouns of sentences 1 and 3. Here the speakers are free to use rather dramatic F0 excursions to mark focus. The tonal speakers, on the other hand, seem to be restricted from doing this. It is only the final color adjective of sentence 4 which is given a markedly higher F0 maximum than the counterpart noun in sentence 3. Since the adjective of sentence 4 has high lexical tone, this fact seems to allow the speakers to additionally raise the maximum F0. As the adjective of sentence 2 has low lexical tone, the speakers are not free to raise this F0 maximum. Here we see evidence of interplay between lexical tone and intonation.

In the listing of animals in sentences 5 and 6, there is a large difference in the F0 maximum of the final word in the non-tonal dialect. The word “badger” in sentence 6 is spoken with a much higher F0 maximum than the word “chicken” in sentence 5. This can be explained by the fact that the word “badger” is semantically marked compared to the other common farm animals in the list. It is quite natural in Kammu farming culture to have a buffalo, a pig, a chicken, a horse and a cat, but not a badger! Some of the speakers even asked to confirm what the word was, and therefore it is not surprising if the word often elicited additional speaker engagement. This extra engagement also shows up in the tonal speakers’ versions of “badger” raising the low lexical tone to a higher F0 maximum than the word “chicken” in sentence 5 which has high lexical tone. Here, speaker engagement is seen to override the to-

nal restriction, although the overall pitch range is still restricted compared to the non-tonal dialect. Thus we see an interaction between speaker engagement, tone and intonation.

Conclusions

In this study we see that the general patterns of intonation for these sentences are similar in the two dialects. However, there is clear evidence of the lexical tones of the tonal dialect restricting the pitch range and the realization of focus, especially when the lexical tone is low. Speaker engagement can have a strong effect on the utterance-final accent, and can even neutralize pitch differences of high and low lexical tone in certain cases.

Acknowledgements

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.

References

- Fant, G. and Kruckenberg, A. (2004) Analysis and synthesis of Swedish prosody with outlooks on production and perception. In Fant G., Fujisaki H., Cao J. and Xu Y (eds.) *From traditional phonology to modern speech processing*, 73-95. Foreign Language Teaching and Research Press, Beijing.
- Karlsson A., House D., Svantesson J-O. and Tayanin D. (2007) Prosodic phrasing in tonal and ton-tonal dialects of Kammu. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, 1309-1312.
- Karlsson A., House D. and Tayanin, D. (2008) Recognizing phrase and utterance as prosodic units in non-tonal dialects of Kammu. In *Proceedings, FONETIK 2008*, 89-92, Department of Linguistics, University of Gothenburg.
- Sjölander K. and Beskow J. (2000) WaveSurfer - an open source speech tool. In *Proceedings of ICSLP 2000, 6th Intl Conf on Spoken Language Processing*, 464-467, Beijing.
- Svantesson J-O. and House D. (2006) Tone production, tone perception and Kammu tonogenesis. *Phonology* 23, 309-333.