PROSODIC SIGNALING OF INFORMATION AND DISCOURSE STRUCTURE FROM A TYPOLOGICAL PERSPECTIVE

Anastasia Karlsson[1], David House[2], Jan-Olof Svantesson[1]

[1]Lund University, Sweden, [2]KTH (Royal Institute of Technology), Stockholm, Sweden
anastasia.karlsson@ling.lu.se, davidh@speech.kth.se, jan-olof.svantesson@ling.lu.se

ABSTRACT

This study investigates the relationship between prosody and information/discourse structure in spontaneous spoken folk tales in the tonal Mon-Khmer language Northern Kammu, a language that behaves as a typical phrase language where available boundary tones are enhanced to mark information structuring. Topic is always placed before Comment by syntactic movement if necessary. There is a prosodic signaling of the boundary between Topic and Comment. Discourse structure is reflected in prosody, and we find higher boundary tones near the boundaries between Discourse Units. The results are discussed in terms of a typology of spoken discourse.

Keywords: prosodic typology, discourse structure, information structure, intonation, tone

1. INTRODUCTION

1.1 Prosody and information structure

The function of human communication is to organize and communicate thoughts [2]. In general terms, information structure reflects the status of content, as new or old. To convey information structure, languages can use different strategies, e.g. word order, morphology and prosody. The interaction between these strategies and information structure has been given considerable attention (e.g. [4, 13]), but little work has been done on information structuring in discourse.

The common element in nearly all theories about information structure is the twofold status of information, described as old/new, topic/comment or theme/rheme. Basically, the old is what has explicitly been mentioned previously or is supposed to be shared knowledge between the speakers. The new adds new information and "advances the discourse" [11].

When discourse moves forward, informational status of old and new changes as new information is introduced. Besides the structuring of informational flow on a local level, there is evidence that the utterances of a discourse are organized into more global units. Lehiste [9], Bruce [1] and others have shown that spoken paragraphs have a characteristic prosodic suprastructure as a result of demarcating topics, feedback seeking, and turntaking.

1.2 Tone and prosodic typology

Within prosodic typology, languages with lexical tones have traditionally been treated as belonging to a separate prosodic category, as it is usually assumed that the tones restrict the use of sentence intonation [3]. However, a series of studies on Kammu shows that the Northern Kammu dialect, which has a two-tone system (high and low tone) uses the same mechanisms for focusing and phrasing as non-tonal languages of the phrasal type [7], where the main function of sentence intonation is to mark prosodic boundaries. Phrase boundaries occur at the right edge of each prosodic phrase and are realized by a high (or high falling) pitch. The focused word is by default placed at the end of an utterance coinciding with the place of the boundary tone, and the pitch of the phrase boundary tone is raised. There is no additional tonal gesture for focal accent. Based on elicited speech it has been shown that prosodic boundaries in Kammu have three functions, marking prosodic phrases, focus and speaker engagement [5].

1.3 Prosodic boundaries in spontaneous monologues

A recent study of prosodic boundaries in Kammu investigated spontaneous accounts of rice growing, from the beginning of work in the field until the rice is cooked and eaten [8]. The accounts all take the form of a kind of listing of successive events. They have a very clear information structure where new information is placed at the end of each utterance; it is then repeated in the next utterance as old information and is followed in turn by new information.

The utterances comprising the accounts are also structured into larger Discourse Units (DUs) which are informatively coherent parts of discourse with a clear beginning and end [2]. The DUs fall naturally into thematic topics reflecting traditional Kammu agricultural activities. Results of the study showed that speakers upstepped the boundary tone on successive utterances up to the end of each DU.

The current study investigates the spontaneous telling of folk tales in order to find out if prosodic boundaries reflect information and discourse structuring in this speech genre as well. Folk tales have a more complex information structuring than do accounts of rice growing. There is no immediate succession of activities. Events, though chained, are elaborated with explanations, direct speech, and the addition of background information. Although having a common content, these narratives are freely recreated by each speaker to hold the listeners' attention.

2. SPEECH MATERIAL AND METHOD

2.1 Speech material

The speech material consists of recordings made by Kristina Lindell in the 1970s of four male speakers of Northern Kammu, and a recording of a fifth male speaker obtained in 2007 [7]. The five speakers tell a folk tale *Àay Cét Réey* 'Mr Seven Rice-cookers'. The material consists of monologues of a semi-spontaneous nature. The speakers show a high level of engagement and are very animated in some parts of the story.

2.2 Method

We analyze information structure, syntax and intonation separately and then compare them to determine the connections between them. This method allows us to point out the main features active in conveying information structure and avoids circularity. As we need a tool for discourse segmentation which is genre independent and to avoid subjective judgements as far as possible, we have segmented the discourse on the basis of syntactic surface events.

As spoken discourse is rich in self-corrections, incomplete constructions and other phenomena, traditional syntactic segmentation is not always applicable. The central unit for our syntactic segmentation is the sentence, consisting of a verb and its complements (if any).

Information flow is divided into Information Units (IUs). Each IU consists of given + new information or only new information. The principle for recognizing new and given is based on explicit (or anaphoric, or paraphrased) mentioning in discourse. An IU consists of Topic (what the proposition is about) and Comment (information about the Topic) or only Comment.

The five narratives were also divided into DUs based on discrete actions or scenes within the narratives. The first nine DUs were used in this study and comprised the following actions and scenes: 1) pregnancy 2) birth 3) eating 4) asking about the father 5) going to meet the father 6) the big tree 7) bathing 8) the merchants 9) going back home. All storytellers except Speakers 4 and 5 adhered to this scheme. Speaker 4 told scene 4 before scene 3, and Speaker 5 told a scene (5a) about the father trying to kill the son after scene 5, and then proceeded directly to scene 9, omitting scenes 6 to 8 (cf. Figure 1).

Intonation was labeled by measuring the F0 maxima in the last word of each utterance and within the sentence. Only F0 maxima were chosen as previous work has indicated that it is the high part of the prosodic boundary that changes as an expression of different functions [5, 7].

3. ANALYSIS

3.1 Informational structuring of narratives

Four main types of IUs are found for the present narratives: 1) $Topic_{old}$–Comment, 2) $Topic_{new}$–Comment, 3) $[Topic_{old}$–Comment$]_{repetition}$ and 4) Comment. Topic is always placed before Comment and can consist of new or given (or explicitly repeated) information. Many IUs lack Topic. Comment is always new information, except for the IUs which are repetitions of a previous IU. In the present investigation we merge the four types into two: Topic–Comment and Comment, and leave the more fine-grained analysis for the near future.

3.2 Discourse Units

All speakers follow the main content of the tale, and except for Speakers 4 and 5 they include the same nine DUs (see section 2.2 above).

3.3 Syntax

The default word order in Kammu is SVO. We find many cases of null subject sentences in our material, when the subject is recoverable from the context. There is not very much word-order variation in the stories; the cases we found are left dislocation of the verb (VSO) or of the object (OSV), and placing the object before the verb using a construction with the verb *màʌt* 'take' ((S) *màʌt* O V). By matching the levels of syntax and information structure we find that the left movement places the dislocated syntactic unit in the Topic position.

3.4 Intonation

We do not find a systematic raising of phrase boundaries at the end of IUs, as was found for elicited

speech and spontaneous accounts of rice growing. Instead, the F0 maximum tends to be within the Comment, near the Topic. We assume that it functions as marking the boundary between Topic and Comment, and we chose therefore to measure the F0 maximum in the last word of the Topic and in the first word of the Comment in each Topic–Comment unit.

4. RESULTS

4.1 Tonal marking of Information Structure

The F0 maxima of the first and last words of all Topics and Comments were measured for each speaker and converted to semitones. Expecting that the boundary between Topic and Comment was marked by an F0 rise, we computed the difference between the F0 maximum of the first word of the Comment and the F0 maximum of the last word of the Topic for each Topic–Comment unit. To compensate for the F0 difference due to the two lexical tones in Kammu, we computed the F0 mean value for High tone and Low tone words for each speaker and added the difference between them (given in Table 1) to the F0 values (in semitones) of Low tone words before calculating the Topic-to-Comment rise. The number of rises and falls are shown in Table 1. As seen there, the number of rises is about twice the number of falls. An exact binomial test shows this result to be highly significant ($p < 0.001$). At the same time, the table shows that there are quite large differences between the speakers.

4.2 Tonal marking of Discourse Units

There is a tendency for the rises between Topic and Comment to be larger than the falls. As seen in Table 1, the largest rise is larger (for most speakers much larger) than the largest fall. A possible interpretation of this finding is that the large rises have a communicative function related to the DUs, similar to the upstepping of prosodic boundaries found in spontaneous accounts of rice growing [8].

Table 1: Mean difference between high and low tones, number of Topic to Comment rises and falls, and largest rise and fall for each speaker

Speaker:	1	2	3	4	5
tone difference (st)	4.25	2.06	3.30	2.40	3.44
number of rises	23	17	24	27	29
number of falls	13	8	18	7	18
largest rise (st)	8.59	12.54	6.02	11.11	9.96
largest fall (st)	3.96	4.21	4.28	2.40	2.69

In order to test this hypothesis, we investigated the distribution of the large rises at the Topic–Comment border. A "large rise" was defined as one whose size (after compensation for the influence of lexical tones) was at least 3 semitones. The reason for this is that previous research [12] has shown that Northern Kammu speakers perceive words differing by about 3 semitones from each other as belonging to different tone categories.

For each large rise, we measured the distance to the border of a DU (as defined in 2.2 above), assigning the distance 0 to those Topic–Comment units that are closest to the DU border, distance 1 to those with one Topic–Comment unit intervening between them and a DU border, etc. The distance distribution of those Topic–Comment units that are marked by large rises was then compared with the distribution of all Topic–Comment units using the one-sample Kolmogorov–Smirnov test. The result showed that, except for Speaker 3, the Topic–Comment units that are marked by large rises are closer to the DU borders than other Topic–Comment units ($p < 0.01$ for Speakers 4 and 5; $p < 0.05$ for Speakers 1 and 2). The fact that these Topic–Comment units cluster around the DU borders suggests that the high rises have a communicative function demarcating DUs even in these spontaneous and otherwise rather unstructured narratives.

As an example, the distribution of the large rises for Speaker 5 is shown in Figure 1, illustrating the tendency for large rises to cluster around the DU borders. In this figure, small rises and falls are represented as dots, and large rises as bars proportional to the size of the rise. The figure also shows the borders and numbers of the DUs as given in section 2.2.

Figure 1: Distribution of large Topic-to-Comment rises for Speaker 5. See the text for details.
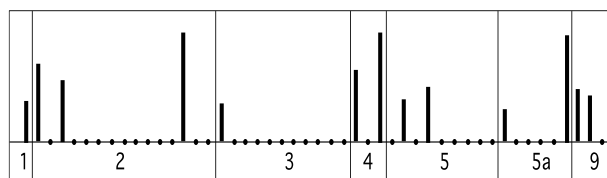
5. DISCUSSION

5.1 Typological implications

Prosodic typology divides languages into intonation, phrase and tone languages [3, 6]. In intonation languages with both pitch accents and boundary tones, signaling of information structure is spread between the two and new information is often marked by the

addition of a pitch accent on the focused word. In phrase languages, signaling of information structure by intonation is suppressed. Despite the occurrence of lexical tones in Northern Kammu, it behaves as a typical phrase language, lacking pitch accents and marking information structuring by enhancing available boundary tones. Topic is always placed before Comment, and syntactic movement is used to move a unit from its default position in order to place it in Topic position. Prosody thus gets a role equivalent to syntax and morphology, constraining the way of expressing information structure.

The boundary between Topic and Comment is generally signalled by upstepping of F0. The following example illustrates shifts in informational status of sentence units in two adjacent sentences. The first sentence only consists of Comment. In the second sentence, *tráak* 'buffalo' becomes a Topic$_{old}$ and is marked by an upstepping of F0 between the final high of the Topic and the beginning of the Comment *pláh* 'put'. There is also syntactic left dislocation of the object *tráak* 'buffalo' to the position before the verb *pláh* 'put' to get it in the Topic position.

[ʔəə, ɲíp tráak mòoj tóo.]$_{comment}$
Oh, grab buffalo one Classifier
'Then he grabbed a buffalo.'

[mʌʌt tráak nìʔ]$_{topic}$ [pláh kə́əj màh knìʔ.]$_{comment}$
take buffalo that put on rice that
'[He] spread the buffalo on the rice.'

5.2 Prosodic evidence for hierarchic structure of folk tales

Not only information structure, but also discourse structure is reflected in prosody. Thus, we find higher boundary tones near the boundaries between DUs. In addition, the large rises occur mostly at the boundaries between DU4 and DU5 and between DU6 and DU7 (see section 2.2). The prosody thus indicates that some of the DUs are superimposed on the others, suggesting a hierarchy in the structuring of DUs. This can be analyzed in reference to Propp's model of folk tales [10]. Stories are composed of sequences of events and actions ("functions" in Propp's terms) and these constitute episodes of the story. In our tales we find three episodes, Birth, Meeting with father and Meeting with the merchants. These episodes are composed of sequences of events. Prosodically, the episodes are marked by higher F0 boundaries at their borders. This suggests the relevance of a finer hierarchy than the division into DUs used here, with "events" grouped into

"episodes". Propp's model is based on the written form of narratives. We have found prosodic evidence for this hierarchical structuring of tales in the oral form of narratives.

5.3 Speech genre

These results indicate that there are differences between the prosodic signaling of information structure in accounts of rice growing and in folk tale narration. This can be explained by differences in information structuring between speech genres. In accounts of familiar activities, such as rice growing, preplanning as an effect of familiarity leads to systematic long-term mutual relations of boundary tones marking information units and thematic shifts. Structural means to signal information structure are uniform through this kind of narrative.

In tale telling we find more short-term prosodic relations reflecting less planned sequences of events, with the expression of information structure varying more. As a result we find more frequent changes of word order and fewer long-term prosodic relations. The main function of prosody becomes the marking of boundaries between Topic and Comment.

"Episodes" are, however, predefined for this folk tale and we find more systematic prosodic signaling of their borders. This is in fact similar to the pre-planned episodes in the accounts of rice growing. The difference is that while in the rice growing accounts, the episodes are composed of sequences of actions with clear endings, the episodes in the folk tale are less homogeneous as they are composed of events, actions, direct speech and parenthetic information.

6. CONCLUSIONS

Our results have several typological implications. We show how a phrase language with lexical tones marks information structure. Kammu uses available tonal boundaries to mark two kinds of events, boundary between Topic and Comment and borders between larger Discourse Units.

Topic always precedes Comment in Kammu, and the word order reflects this structure by using syntactic movement of units to the Topic position when necessary. We also find prosodic evidence for Propp's [10] semantic schema for tales, structuring them into events that constitute episodes. We find that prosodic boundaries in Kammu tales indicate a hierarchic structuring of events grouped into episodes.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] Bruce, G. 1982. Textual aspects of prosody in Swedish. *Phonetica* 39, 274–287.

[2] Chafe, W. 2001. The analysis of discourse flow. In: Schiffrin, D., Tannen, D., Hamilton, H.E. (eds), *The Handbook of Discourse Analysis.* Oxford: Blackwell, 673–687.

[3] Féry, C. 2010. The intonation of Indian languages: an areal phenomenon. In: Hasnain, I., Chaudhury, S. (eds), *Problematizing Language Studies. Festschrift for Ramakant Agnihotri,* Delhi: Aakar Books, 288–312.

[4] Fiedler, I., Schwarz, A. (eds). 2010. *The Expression of Information Structure. A documentation of its Diversity across Africa.* Berlin. Humboldt University.

[5] House, D., Karlsson, A., Svantesson, J.-O., Tayanin, D. 2009. The phrase-final accent in Kammu: effects of tone, focus and engagement. *Proceedings of Interspeech 2009*, Brighton, UK, 2439-2442.

[6] Jun, S.-A. 2005. Prosodic typology. In: Jun, S.-A. (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing.* Oxford: Oxford University Press, 430–458.

[7] Karlsson, A., House, D., Svantesson, J.-O. 2012. Intonation adapts to lexical tone: the case of Kammu. *Phonetica* 69, 28–47.

[8] Karlsson, A., Svantesson, J.-O., House, D. 2013. Multifunctionality of prosodic boundaries in spontaneous narratives in Kammu. *Proceedings Prosody-Discourse Interface 2013,* Leuven.

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

[10] Propp, V. 1958. *Morphology of the Folk-Tale,* Bloomington, Ind.: Indiana University Press. (Originally published as *Morfologiya skazki*, Leningrad 1928.)

[11] Steedman, M., Kruijff-Korbayová, I. 2001. Two dimensions of Information Structure in relation to Discourse Semantics and Discourse Structure. In: Kruijff-Korbayová, I., Steedman, M. (eds), *Proceedings of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics,* 1-9.

[12] Svantesson, J.-O., House, D. 2006. Tone production, tone perception and Kammu tonogenesis. *Phonology* 23, 309–333.

[13] Zimmermann, M., Féry, C. (eds). 2010. *Information Structure,* New York: OUP Inc.