

## Tonal cues to Topic and Comment in spontaneous Japanese and Mongolian narratives

Anastasia Karlsson<sup>1</sup>, David House<sup>2</sup>, Jan-Olof Svantesson<sup>1</sup>

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

Our study treats intonation on the level of spontaneous discourse in Japanese and Mongolian. Besides being verb-final and having formal markers of Topic, these languages share some important prosodic similarities. Both lack lexical stress. Phrasing is achieved by initial boundaries. The functional load of tone in marking words is high: Japanese has word accents, and Mongolian has an obligatory Accentual Phrase coinciding with the grammatical word and having an initial LH boundary (Karlsson 2014).

This study investigates tonal cues to the Topic-Comment dichotomy in spontaneous narratives. Topic is defined here as what the proposition is about and Comment is information about the Topic. The Japanese material consisted of twelve spontaneous narratives spoken by six speakers in pairs (four females, two males) having an average duration of 1 minute 28 seconds. The Mongolian material was spoken by four female speakers (two pairs). The narratives were divided into phrases and annotated for Topic and Comment. Mean F0 and final F0 peak values were measured independently for each phrase. For the Japanese narratives, each speaker produced a higher mean F0 for the phrases annotated as Topic than for those annotated as Comment with a mean difference of 1.84 semitones. The phrase-final F0 peak was also higher for Topic than Comment in all narratives except three (two speakers) with a mean difference of 1.36 semitones. Two Mongolian narratives have been analyzed. For one of them, the speaker produced a higher mean F0 for Topic than for Comment with a statistically significant mean difference of 1.08 semitones. There was no significant difference between Topic and Comment, for the second speaker's mean F0, nor for the phrase-final F0 peak for any of the speakers. Thus we see a difference between the two languages with Japanese using both overall phrasal F0 range and the level of local final highs to reflect the information structure. In Mongolian *read* speech, the verb and sometimes longer utterance-final parts are typically produced with low pitch. As the verb is usually described as a typical Comment, we expected to find a connection between pitch lowering and Comment also in *spontaneous* Mongolian narratives, but this was not the case.

The most striking difference between our elicited and spontaneous recordings is the frequently occurring high terminals in spontaneous speech. High terminals in Japanese have been described by e.g. Maekawa (2015). In the second part of our study we concentrate on high terminals in our Mongolian recordings. Narratives were segmented into "syntactic groups" (verb and its arguments). Mongolian has the Accentual Phrase, the Intermediate Phrase and Intonation Phrase (Karlsson 2014). In our spontaneous material, the Intermediate Phrase shows a high overlapping with syntactic groups. While Intonation Phrase is marked by L% in read speech, in the present material it is instead marked by H% coinciding with the end of Comment. These H% terminals may result from sequences of events in our narratives, marking continuation of informational flow. However, some speakers have H% even at the ends of thematic episodes. We know that terminal tones in discourse can be multifunctional (House et al. 2009). In our Mongolian material the different functions of high terminals can depend on more remote relationships between the phrases. Comments that belong to the same overarching thematic episode are often marked by upstepping their H%. Marking of the end of a thematic episode (when not L%) is done by a downstepped H%. New thematic episodes start with a considerable reset and/or are marked lexically.

In summary, the Topic-Comment dichotomy on the utterance level is articulated by changes in pitch range in Japanese and the level of local final highs. Mongolian spontaneous discourse does not display this difference between Topic and Comment. Comment is instead often marked by a high terminal and its pitch scaling can be shaped by long-term discourse relations.

House, D., Karlsson, A., Svantesson, J-O., and Tayanin, D. 2009. "The Phrase-Final Accent in Kammu: Effects of Tone, Focus and Engagement". In *Proceedings of Interspeech 2009, Brighton*, pp. 2439-2442.

Karlsson, A. 2014. "The intonational phonology of Mongolian". *Prosodic Typology 2: The Phonology of Intonation and Phrasing*, pp.187-215. Jun, Sun-Ah (ed). Oxford University Press.

Maekawa, K. 2015. "Functional difference between the two variants of rising-falling intonation in spontaneous Japanese monologue". In *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*.