

The role of head and eyebrow movements for the production and perception of prosodic prominence: Evidence (mostly) from Swedish news readings

Gilbert Ambrazaitis
Linnaeus University, Växjö, Sweden

in collaboration with David House, KTH Stockholm,
and Johan Frid, Lund University Humanities Lab

This talk focuses on the role of head and eyebrow gestures for the multimodal construction of prominence. I will summarize results from two recent studies (in progress or submitted) on the co-occurrence of gestures and pitch accents in Swedish news readings. The first study – on prominence production – asks whether the realization of the accentual fall (HL) and the following so-called big-accent rise (H) in the Swedish word accents (Accent 1, Accent 2) is varied as a function of accompanying head and eyebrow gestures. The purpose of this study is to evaluate the hypothesis that prominence production displays a cumulative relation between the acoustic and the kinematic dimensions of spoken language, especially focusing on the clustering of different types of gestures (head, eyebrows), at the same time asking if lexical-prosodic features would interfere with this cumulative relation. The material tested were 60 brief news readings from Swedish television, comprising about 12 minutes of speech from five news presenters (two female). The results reveal a significant trend for larger f_0 rises (in semitones) when a head movement accompanies the accented word, and even larger when an additional eyebrow movement is present. This trend is observed for accentual rises that encode phrase-level prominence, but not for accentual falls that are primarily related to lexical prosody. Moreover, the trend is manifested differently in different lexical-prosodic categories (words with Accent 1 vs. Accent 2, and with one vs. two lexical stresses). The overall result is generally in line with a number of related previous studies, but it also provides a novel type of support for a cumulative-cue hypothesis, as it considers the clustering of gestures (head plus eyebrows).

The second study focuses on the audio-visual integration of pitch accents and head movements in prominence perception. Using a web-based rating task (involving 85 participants: 44 in an audio-visual and 41 in an audio-only condition), we collected prominence ratings for a selection of clips taken from our news reading materials (218 words). A series of analyses were performed on this data set, overall suggesting a small but significant effect of the visual modality, as a stronger perceptual distinction was made between words with and without head movements in the audio-video condition compared to the audio-only condition. Again, these results are in line with previous studies, but they add a couple of novel insights, as they, for instance, provide evidence from a rather uncontrolled rating setting outside the lab. An important contribution that is made by both studies is that they extend previous research efforts, that were mostly based on artificial gestures (implemented in virtual agents) or deliberately performed gestures, by testing ecologically valid speech, involving spontaneously produced gesture – even if the speech is well-prepared and read (news readings). If there is time, I will present a brief overview of our beginning attempts to include data from spontaneous dialogue.