

Prosody influences the listener's online representation of the speaker's thoughts

Ev Fedorenko¹, Lillia Cherkasskiy², Steven Piantadosi¹, Jonathan Scholz¹, and Rebecca Saxe¹

¹MIT; ²Stanford

A critical question in social cognition concerns whether people represent others' utterances as simply their literal meanings or as the speakers' beliefs about their meanings. Whereas representing utterances as others' beliefs may be socially-relevant, it is also computationally intensive. We hypothesized that the probability of representing the speaker's utterances as beliefs increases when some aspect of the mental state attribution (the owner of the mental state or the mental state itself) is salient (in focus). One way to increase the salience of some aspect of an utterance is via the use of prosodic cues. We therefore manipulated the prosodic contour of the phrase "I think" which preceded the speaker's utterances in naturalistic dialogs such that (a) both words were unstressed, (b) "I" was stressed (owner of the mental state in focus), or (c) "think" was stressed (mental state in focus). A control condition without "I think" was also included.

An acoustic analysis of the "I think" fragments of the materials (n=300) demonstrated that six acoustic features (duration, mean, max and min pitch, energy and power) distinguish among the three prosodic contours with the 81.3% accuracy (chance=33%). In a behavioral study we found that comprehenders distinguish the three contours of "I think": they infer the greatest confidence in the speaker's belief in the statement when "I" is stressed (but do not themselves form more confident opinions), followed by the unstressed condition, followed in turn by the "think"-stressed condition. In the fMRI study we used the functional localizer approach. We first identified brain regions that have been argued to be selectively engaged in representing other people's thoughts, using the false belief minus false photograph contrast (e.g., Saxe & Kanwisher, 2003; Perner 2007). We then compared the activation in these regions to "I think" sentences with different prosodic contours and to the control without "I think". Consistent with our hypothesis, these regions show a higher response to the conditions where "I" or "think" is stressed compared to exactly the same sentence if both words are unstressed, or to the condition without "I think" (the latter two are not different from each other, or from the fixation baseline). This is the first demonstration of the regions that have been implicated in representing others' thoughts interpreting utterances as others' beliefs in naturalistic dialogs.