**Contextual inferences in comprehension and acquisition of contrastive prosody**

Listeners use prosody to determine information structure of a sentence, and thereby infer speaker meanings. (e.g., Pierrehumbert & Hirschberg, 1990; Büring, 2003). Contrastive accent (B-accent in Jackendoff's (1972) term; L+H* in the TOBI transcription system) is known to evoke a contrast set and a presupposition that this contrast is relevant in the context at hand. However, the actual acoustic realizations of contrastive accent are notoriously variable (e.g., Ladd, 2008). We present experimental evidence that both children and adults overcome the massive variability and noise in the prosodic signal through rational inference.

Experiment 1 tested whether form-based inferences fine-tune prosodic interpretation in adult listeners. We used the construction "It looks like an X" pronounced either with 1) nuclear H* on the final noun (**Noun-focus prosody**), indicating that the referent is likely to be X, or 2) contrastive L+H* on "looks" (**Verb-focus prosody**), implying otherwise (e.g., *It LOOKS like a butterfly but it is not one*). The task was a 2AFC decision between a picture of the mentioned referent X (e.g., a butterfly) and a picture that visually resembled X (e.g., a moth). Participants were randomly assigned to one of two between-subjects conditions. In the **Prosody-only condition**, participants heard "It looks like an X" with either Noun-focus or Verb-focus prosody (12 each). In the **Prosody+Form condition**, participants heard these two prosodic patterns (8 each) as well as 8 instances of "It's an X" (e.g., "It's a BUTTERFLY!") with nuclear accent on the final noun. **[Results]** Although Verb-focus prosody elicited more "not X" responses in both conditions ($p<.001$), participants were also more likely overall to respond "not X" in the Prosody+Form condition than in the Prosody-only condition ($p<.05$), presumably on the pragmatic expectation that the speaker would have said "It's an X" if she had meant so.
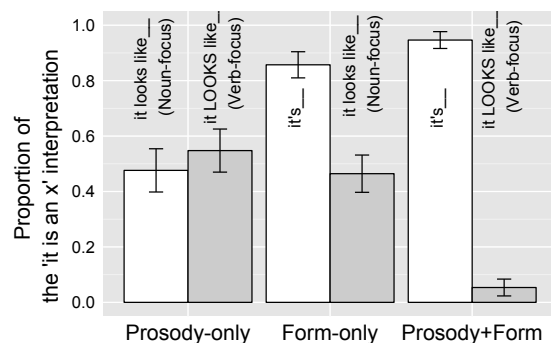
In **Experiment 2**, 4-year-olds were tested in a similar 2AFC picture-selection paradigm. They selected a likely referent based on a puppet's utterances. In the **Prosody-only condition**, as in Exp1, the puppet always said "It looks like an X", encoding the two meanings (i.e., the referent is X or not) using Noun-focus vs. Verb-focus prosody. In the **Form-only condition**, the puppet instead used two constructions: "It's an X" vs. "It looks like an X" (with Noun-focus prosody). In the **Prosody+Form condition**, the puppet used both constructions and marked "It LOOKS like an x…" with Verb-focus prosody. **[Results]** 4-year-olds did not distinguish between the two prosodic patterns in the Prosody-only condition (Fig.1). However, when the input also included a more reliable formal contrast (Prosody + Form condition), children were almost at ceiling in interpreting "It LOOKS like an x…" as indicating a contrastive meaning (i.e., …**but it is not one**). This supports a view that the presence of an unambiguous expression ("It's an X") encourages children to infer that contrastive prosody signals a distinct speaker meaning.

Prosodic input is thus interpreted conditionally on what linguistic cues – structural as well as prosodic – are used by a specific speaker. 4-year-olds, who are otherwise insensitive to the contrast evoking function of L+H*, can bootstrap their knowledge through an inference based on the formal contrast. We conclude that the prosody-pragmatics interface in natural language is not a one-to-one mapping between phonological representations and information structure; rather, it is a part of a bigger inference problem in which listeners integrate information from multiple sources to infer what the speaker means to say in a particular context.

References

Büring, D. (2003). On d-trees, beans, and b-accents. *Linguistics & Philosophy, 26 (5)*, 511-545.

Jackendoff, R. (1972). *Semantics in generative grammar*. Cambridge, MA: MIT Press.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, and M. Pollack (eds), *Intentions and plans in communication and discourse* (pp.271–311). MIT Press.