

That was a question?: Accommodating variability in intonation interpretations

Andrés Buxó-Lugo & Chigusa Kurumada (University of Rochester)

abuxolug@ur.rochester.edu

To be able to comprehend spoken language, listeners must deal with the fact that no two words or utterances are ever spoken in the same way, especially when produced by different speakers [1]. A long-standing problem in communicatively enriched sentence comprehension is that there is **substantial variability** in how speakers use speech sounds such as intonational contours to encode pragmatic intentions [2]–[4]. While a recent proposal suggests that listeners adapt mappings between intonational cues (e.g., F0) and inferred pragmatic intentions in a speaker- and/or context-specific manner [5], underlying mechanisms for such adaptation remain unclear. The current study demonstrates that 1) listeners' intonation interpretation of questions vs. statements can adapt to the distributional input more rapidly than has been shown before; and 2) this adaptation process makes use of syntactic information to infer category membership.

[Materials] 6 tokens of the English construction “It’s X-ing” was recorded either with a rising or a falling utterance final boundary tone, corresponding to **Question** (e.g., It’s cooking?) and **Statement** (e.g., It’s cooking) interpretations (12 tokens each). We resynthesized 11-step continua of “it’s X-ing” sentences, gradually shifting F0 and constituent duration from the typical rising to the falling tone values (Fig. 1). Continua were then normed by 120 native speakers of American English to determine the maximally ambiguous stimulus (i.e., for which hearers were least certain about the intended interpretation).

Experiment 1: S(ubject)s (N=180) were randomly assigned to a Control condition or one of two treatment conditions: Statement-biasing and Question-biasing (Fig. 2). In a Pre-Exposure test phase all Ss heard 22 instances of one item type (e.g., *it’s cooking*), sampled uniformly from the Statement-Question continuum (twice per step), and completed a 2AFC task. During an Exposure phase, Ss continued to provide 2AFC judgments in the same format and received feedback. Ss in the **Control** condition only heard the most prototypical Statement and Question intonation contours (Steps 1 & 11, 15 sentences each). Ss in the **Question-biasing** condition heard prototypical Statements (i.e., Step 1) and the acoustically ambiguous item (Step 6), disambiguated as Questions. Ss in the **Statement-biasing** condition heard the prototypical Questions (Step 11) and the acoustically ambiguous items (Step 6), identified as Statements (Fig. 2). Finally in a Post-Exposure test, Ss completed another 22 trials that were identical to the Pre-Exposure test, to assess the possible adaptation resulting from the Exposure input. There was no lexical overlap between the Pre-/Post-Exposure and the Exposure phases.

Results show that Ss in the two treatment conditions provided opposing interpretations for the previously ambiguous items ($p < .001$, Fig. 2). No such shift was observed in the Control condition. The data suggest that listeners readily accommodated their overall categorization function based on the very restricted input given in the Exposure phase.

Experiment 2: Ss (N=120) repeated the Control and the Statement-biasing conditions in Exp.1 except that the Exposure items disambiguated as Question (Step 11 in both conditions) were syntactic interrogatives (e.g. “Is it X-ing?”). The same resynthesis method was used to ensure that intonational features remained identical for all the items across Exp.1 and Exp.2. If adaptation concerns intonational features alone, the results should replicate those in Exp.1 because all the relevant phonological information was unchanged. If adaptation interfaces with form-based inference, there should be enhanced adaptation in the Statement-biasing condition: The Exposure input supports the inference that the speaker would use the interrogative syntax to convey the question meaning, and hence declarative utterances (It’s an X) would be overall more likely to be Statements. The results supported the latter prediction (Fig. 3; $p < .001$).

Overall, the current results provide novel insights into the mechanism allowing for robust, pragmatically-rich, sentence comprehension despite the substantial variability in the input. Listeners do not only adapt their intonational interpretations, they also tune into how a given speaker combines multiple linguistic devices – intonation and syntax – to encode their intention.

References

- [1] D. R. Ladd, *Intonational phonology*, 2nd ed. Cambridge University Press, 2008.
- [2] A. Arvaniti and G. Garding, "Dialectal variation in the rising accents of American English," *Papers in Laboratory Phonology* 9, pp. 547–576, 2007.
- [3] C. G. Clopper and R. Smiljanic, "Effects of gender and regional dialect on prosodic patterns in American English," *Journal of Phonetics*, vol. 39, no. 2, pp. 237–245, 2011.
- [4] R. J. Podesva, "Salience and the social meaning of declarative contours: Three case studies of gay professionals," *Journal of English Linguistics*, vol. 39, no. 3, pp. 233–264, 2011.
- [5] C. Kurumada, M. Brown, and M. K. Tanenhaus, "Effects of distributional information on categorization of prosodic contours," *Psychonomic Bulletin and Review*, 2017.

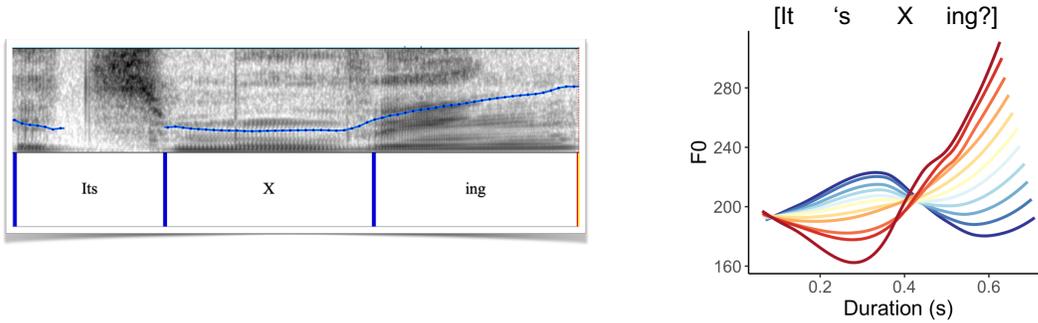


Figure 1. [left] Example contour of "It's X-ing?"; [right] resynthesized contour continuum, ranging from a prototypical Statement (the bottom) contour to a prototypical Question (the top) contour.

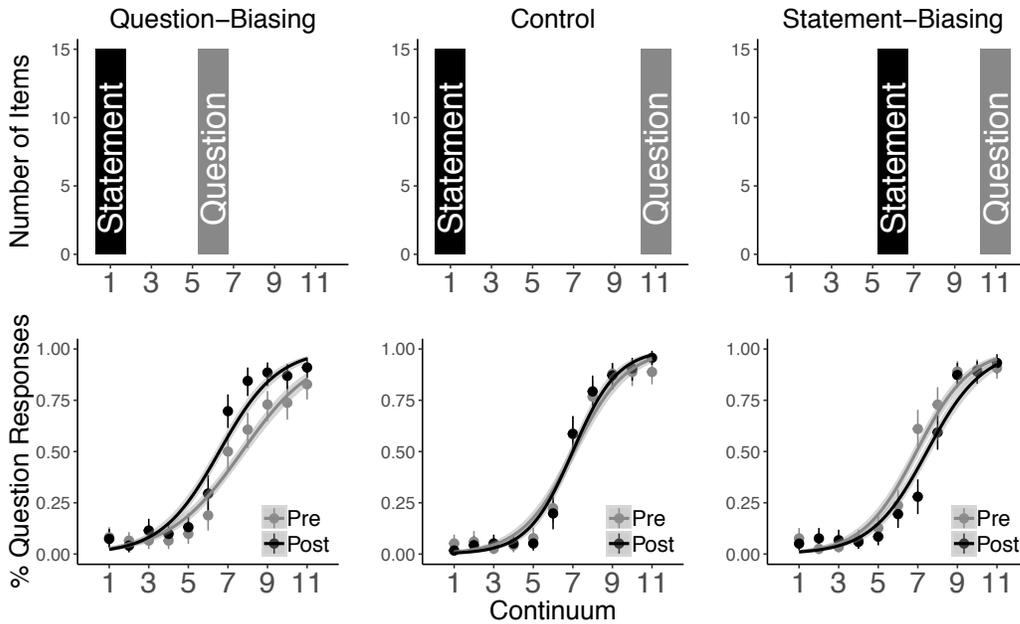


Figure 2. Top: Number of items per continuum step for each condition in Experiment 1. Five lexical items repeated thrice per sentence type. Bottom: Pre- and Post-Exposure responses for each condition.

Figure 3. Results for Experiment 2, Control and Statement-Biasing conditions. The pre-/post-test items were identical to Exp.1. The only change made was that the exposure items with Question feedback (Step 11) were syntactic interrogatives ("Is it X-ing?") in Experiment 2.

