Intonation interpretations are talker-sensitive, but not talker-specific
Chigusa Kurumada (University of Rochester) and Andrés Buxó-Lugo (University of Maryland)
ckuruma2@ur.rochester.edu

One of the persistent puzzles for theories of intonational phonology and its processing pertains
to the phonetic variability in speech sounds [1-3]. Even a seemingly straightforward
distinction (e.g., question vs. statement) exhibits substantial variability conditioned on
physiological and socio-indexical features (e.g., talker’s height, gender, age, dialect) (Fig.1). The
current study investigates why listeners are, nonetheless, surprisingly good at deciphering
intonational meaning through examining roles of expectation adaptation and generalization.

Experiment 1 (Single talker): Two talkers (a female and a male) recorded pairs of a
statement (e.g., *It’s cooking.*) and a question (*It’s cooking?*) utterances. We manipulated the F0
and syllable duration so intonation contours gradually shifted from the typical falling to the rising
tone values (Fig. 2). A resulting 11-step continuum was then normed by 120 native speakers of
American English to determine the maximally ambiguous stimulus (i.e., for which listeners were
least certain about the intended meaning being a statement or a question).

S(subject)s (N=240) were randomly assigned to one of the four between-subject conditions (2
talkers * Statement-biasing vs. Question-biasing). In Pre-Test, 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 provided 2AFC (Question vs. Statement) judgments. In Exposure, Ss continued
to provide 2AFC judgments in the same format and received feedback: 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. Finally, in Post-Test, Ss repeated the procedure identical to the Pre-Test (22 trials).
There was no lexical overlap between the Exposure and Test items. Results in the Post-Test
revealed that Ss in the Question-biasing vs. Statement-biasing conditions provided opposing
interpretations for the previously ambiguous items (p<.001, Fig. 3). The “steeper”
categorization curves with near-categorical judgments at the two ends of the continuum in the
Post-Test (cf. the Pre-Test) suggest that intonation interpretations were re-tuned for a particular
exposure talker; Simply increasing or decreasing base response rates for the
question/statement meanings according to the Exposure items alone would not predict this.

Experiment 2 (Two talker): Results of Exp.1 are compatible with two possible mechanisms:
the retuning of intonation interpretations is strictly talker specific, and listeners begin the process
anew for each talker. Alternatively, the adjusted expectations can, at least partially, generalize
to a novel talker. To distinguish these, we modified Exp.1 by including both the female and male
talkers in the Pre/Post-Tests. Listeners heard only one talker throughout Exposure, and we
compared Pre and Post-Test responses for both talkers to test if they retuned their
interpretations for both talkers, or for just the talker they heard throughout Exposure. Results
suggest that there is a significant effect of cross-talker generalization (Fig.4). The effect was
further modulated by an asymmetry between the two talkers (i.e., the female exposure talker
triggered more generalization) and by the presentation order (i.e., responses to the talker who
appeared immediately after the Exposure exhibited more generalization).

The talker-sensitive processing of intonation is likely a key to mapping the variable phonetic
input to their abstract (phonological) representations [4]. The novel finding of cross-talker
generalization strongly suggests that intonation interpretations undergo constant adjustments
across conversational turns in a given context. This continuous re-tuning likely facilitates
effective intonation interpretations given that talkers generally converge to each other’s prosodic
production in the course of an interaction [5,6], abiding by common contextual constraints.
Figure 1. Acoustic variability and category overlap in the production of Question vs. Statement utterances (Buxo-Lugo & Kurumada, 2018)

Figure 2. Example 11-step continuum from the female talker (blue = Step 1, red = Step 11)

Figure 3. Proportions of "Question" responses in Experiment 1. The two between-subject conditions (female talker vs. male talker) are plotted in the left and the right panels, respectively. The dotted lines represent the Pre-Test categorization judgments, and the solid lines represent the Post-Test categorization judgments.

Figure 4. Proportions of "Question" responses in Experiment 2. The two between-subject conditions (female exposure talker vs. male exposure talker) are plotted in the left and the right panels, the two within-subject conditions (female test talker vs. male test talker) are plotted in the top and the bottom panels. The cases in which the exposure and the test talkers do not match (framed with black lines) demonstrate effects of cross-talker generalization.