Perception and adaptation of speech prosody in autistic adolescents

Chigusa Kurumada, Rachel Rivera, Christina Layton, Paul Allen, Loisa Bennetto
University of Rochester

Autistic children and adolescents often experience difficulties in recognizing the linguistic and affective meaning of speech prosody [1-3]. One possible, yet underexplored, reason for these challenges may be **the phonetic variability** present in everyday speech. Even seemingly straightforward distinctions (e.g., question vs. statement) exhibit substantial variability due to physiological and socio-indexical features of speakers, such as the talker's gender, age, and accent/dialect (**Fig.1**) [4,5]. In this study, we hypothesize that the difficulties autistic adolescents face in receptive prosody, compared to neurotypical controls, may stem from differences in their ability for rapidly **adapt to the cross-talker variability in prosody** [6].

Experiment: A male native speaker of American English recorded pairs of statement (e.g., *It's cooking.*) and question (e.g., *It's cooking?*) utterances. We manipulated the F0 and syllable duration so that 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 point at which the stimuli were maximally ambiguous (i.e., listeners were least certain about whether the intended meaning was a statement or question).

[Adaptation] 3 groups of subjects participated: 94 autistic adolescents (mean age = 14.9), 76 non-autistic adolescents (mean age = 15.2) and 100 non-autistic young adults (mean age = 20.3). In each, subjects were randomly assigned to the Question-biasing vs. Statement-biasing conditions. In Pre-Test, they heard 44 instances of one item type (e.g., *it's cooking*) sampled uniformly from the Statement-Question continuum and provided 2AFC (Question vs. Statement) judgments. In Training, they were assigned to one of the two between-subject conditions and received feedback: Those in the **Question-biasing** condition heard prototypical Statements (Step 1) and the prosodically ambiguous item (Step 6) disambiguated as Questions. Those in the **Statement-biasing** condition heard the prototypical Questions (Step 11) and the ambiguous items (Step 6) disambiguated as Statements. Finally, in Post-Test, all subjects repeated the task identical to the Pre-Test (44 trials). There was no lexical overlap between Training and Tests.

After a mere 6 min of training, the control subjects in the Question-biasing vs. Statement-biasing conditions provided opposing interpretations for the previously ambiguous items ($p<.001$, non-autistic young adult controls' responses shown in **Fig.3A**). Note that adaptation was most prominent for the mid-continuum, "ambiguous", stimuli (Steps 5-7); Judgments for the end-point stimuli remained near-categorical (Step 1 = statement, Step 11 = questions). This supports the control subjects' perceptual retuning based on the input, as simply adjusting base response rates for questions/statements to the training items alone would not predict this. Critically, **autistic adolescents did not adapt as significantly** (**Fig.3B**). A mixed effect model revealed a significant by-group difference in the adaptive shift of responses ($p<.02$).

[Discrimination] All participants also took part in an oddball discrimination task. 84 triplets of tone stimuli were created by extracting the final syllable from the 11-step adaptation test stimuli (i.e., "-ing" from "it's cooking"). In each triplet, either the first or the last tone was an oddball that was two steps apart from the other two (e.g., Step1-Step1-Step3). All three groups of participants identified the location of the oddball well above the chance accuracy of 50% (**Fig.3C**).

In sum, while autistic adolescents were as accurate as non-autistic controls in perceiving subtle prosodic variations, **their prosodic categorization is less adaptive to the input statistics**. This is critical because adolescence, to many, is a pivotal stage of expanding social network and the cross-talker variability in the input [7]. The accurate but less flexible perception in autistic adolescents could severely impede robust sound-meaning mapping in language communication. Possible sources of the reduced adaptivity in autism will be discussed.

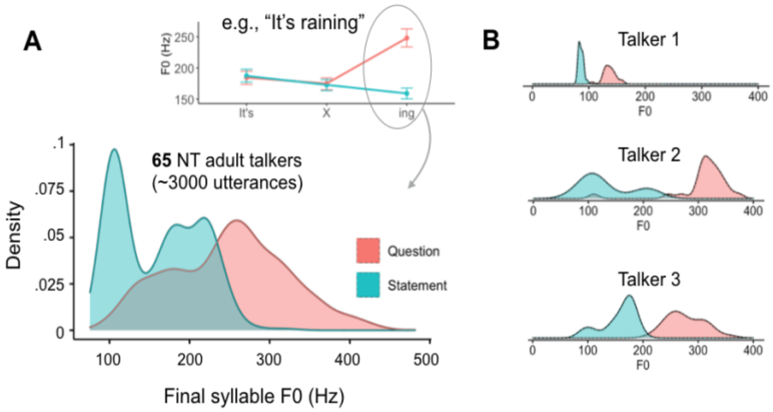

Fig. 1. **A**. Substantial category overlap in phonetic cue distributions from 65 native talkers (~3000 tokens), based on data from Xie et al., (2021) [5]. **B**. While better separation of question vs. statement observed within each talker, no simple baseline corrections can fully resolve the across-talker variability. This motivates the importance of adaptation in comprehension.

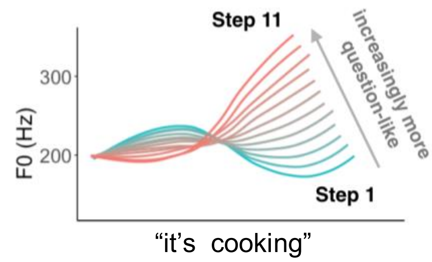

Fig. 2. Fundamental frequency (F0) and duration of an 11-step continuum of the stimuli. Step 1 and Step 11 were produced by a male native speaker of American English.

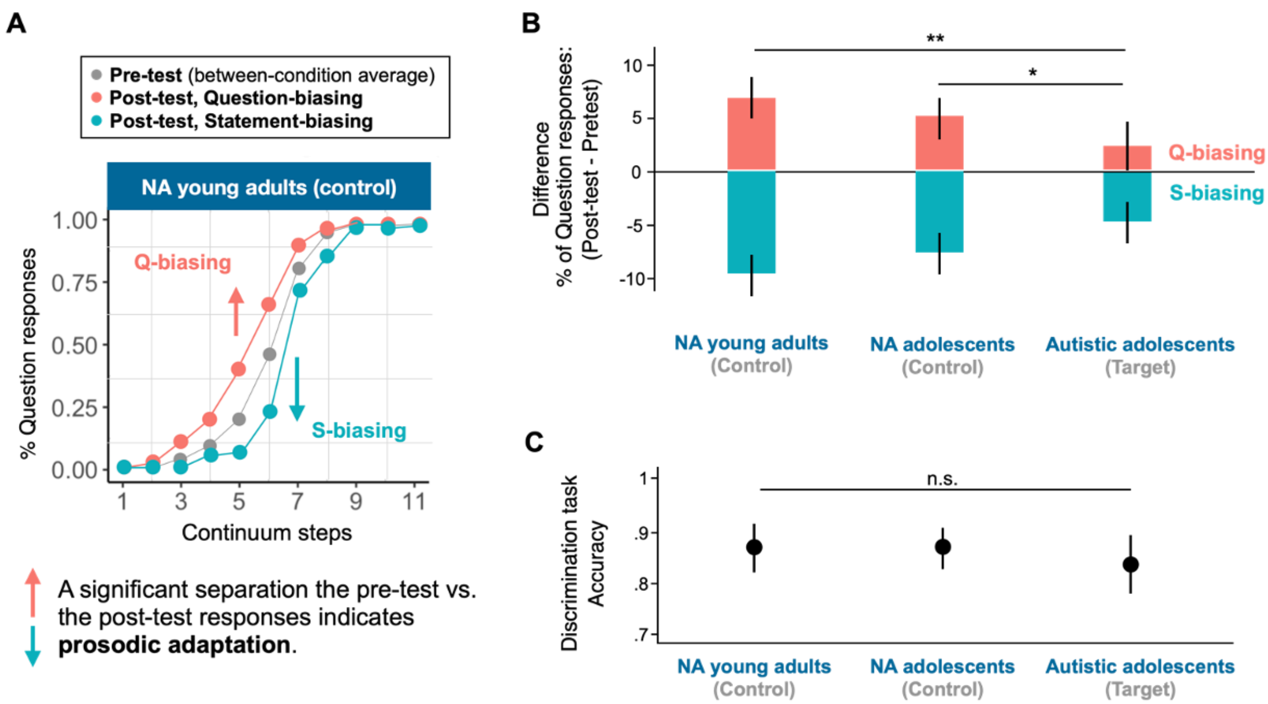

Fig. 3. Participants' responses. **A.** 2 groups of control subjects adapted their categorization responses in the opposing directions after the 6 min training between the Question-biasing (red) and in the Statement-biasing (green) conditions. Here we present data from the non-autistic young adults as an example. The gray line indicates the average pre-test responses. **B.** Autistic adolescents were less adaptive as compared to the two control groups. Error bars indicate bootstrapped 95% CIs. **C.** Discrimination task response accuracies did not differ across the three groups tested.

[1] Tager-Flusberg, H. (2000). Understanding the language and communicative impairments in autism. [2] McCann, H., & Peppe, S. (2003). Prosody in autism spectrum disorders: a criticl review. [3] Järvinen-Pasley et al., (2008). The relationship between form and function level receptive prosodic abilities in autism. [4] Clopper, C. & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. [5] Xie et al., (2021). Encoding and decoding of meaning of structured variability in intonational speech prosody. [6] Alispahic et al., (2023). Auditory perceptual learning in autistic adults. [7] Abrams et al., (2022). A neurodevelopmental shift in reward circuitry from mother's to nonfamilial voices in adolescence. *J. Neurosci*.