The perception of prosodic prominence: continuous or categorical—and for whom?

Jason Bishop[1,2], Chen Zhou[1], Mei-Ying Ki[1]

[1]*CUNY Graduate Center (USA),* [2]*CUNY College of Staten Island (USA)*

Perceiving prosodic prominence, like perceiving other aspects of speech, emerges from some combination of (a) bottom-up cues related to the signal and (b) top-down cues related to: linguistic knowledge (e.g., phonology); experience using language (e.g., lexical stats); interpretation and various aspects of the speaking/listening context (Cole et al., 2019; Cole et al., 2014; Bishop et al. 2020). However, we know much less about the role played by properties of listeners themselves—i.e., individual differences. This paper is about such differences.

The variety of individual differences we focus on here is related to a particular set of cognitive mechanisms underlying perception and processing of language, what we will refer to (following work in sentence processing; e.g., Xiang et al., 2013) as *pragmatic skill*. Individuals differ in their level of this construct, i.e., in their propensity to integrate context, to engage in perspective taking (relevant to both Theory of Mind and empathy), and to interpret speakers' intensions. Work in recent years has shown that measures of pragmatic skill, such as those targeting neurotypical autistic traits (e.g., Bishop et al., 2020; Jun & Bishop, 2015) and empathy (e.g., Arvaniti et al., 2022), predict sensitivity to prosodic cues in various perception and processing tasks. The present study explored how such variation influences listeners' performance in Rapid Prosody Transcription (Cole et al., 2010), a speeded identification task widely used to test the perceptual salience of prosodic events, both phonetic and phonological. In a reanalysis of the data set reported in Bishop et al. (2020), we asked whether individual differences in pragmatic skill related to autistic traits (based on subscales of the AQ and BAPQ; Baron-Cohen et al., 2001a; Hurley et al., 2007) and empathy (the Reading the Mind in the Eyes task, henceforth "EYES"; Baron-Cohen et al., 2001b) mediate listeners' sensitivity to phonological structure during prominence perception. Our hypothesis, based on recent work (Arvaniti et al., 2022), was that measures of empathy in particular may predict a more gradient strategy, making listeners' identification of prominence more sensitive to the following distinctions: *accent status* (a word's status as prenuclear accented or nuclear accented in the ToBI framework; Beckman & Ayers, 1997); *accent type* (whether a word bears a !H*, H*, or L+H*); and phonetic variation (f0 peak height) for words bearing H* and L+H*. Data were based on the judgments of 160 American English-speaking listeners who heard approximately 5 minutes of (ToBI annotated) connected speech (see Bishop et al., 2020).

Mixed-effects regression analyses showed the following. First, higher scores on all measures of pragmatic skill were associated with the identification of an overall higher proportion of accented words (with small effect sizes), but no measure of pragmatic skill predicted differences in sensitivity to *accent status* contrasts (see Fig 1). Second, higher scores on EYES (our measure of empathy), but no other measure, were associated with significantly greater sensitivity to *accent type* (Fig 2). We also found EYES to be predictive of listeners' sensitivity to within-category phonetic cues; for words marked by L+H*, higher levels of empathy were associated with greater sensitivity to the height of f0 peaks (Fig 3). Finally, we also found EYES scores to interact significantly with lexical frequency, such that judgements by low-empathy listeners were more strongly influenced by this non-signal-based lexical variable (and we note this pattern has also been reported for clinical groups; Grice et al., 2016).

Taken together, these findings suggest that pragmatic skill—and a measure of empathy in particular—not only influences the perception of prominence, but does so in ways more closely associated with intonational structure (pitch accent types, f0) than with metrical structure (accent status). Moreover, these individual differences related to f0 level contrasts (!H*, H*, L+H*) and phonetic accent peaks (for L+H*) suggests that a long-standing question about how categorically emphasis is perceived (cf Ladd & Morton, 1997) may have a listener-specific (but highly systematic) answer. We discuss implications of these patterns for the categorical versus gradient perception of prosody, and possible constraints on the interpretational effects of some intonational contrasts in (American) English.
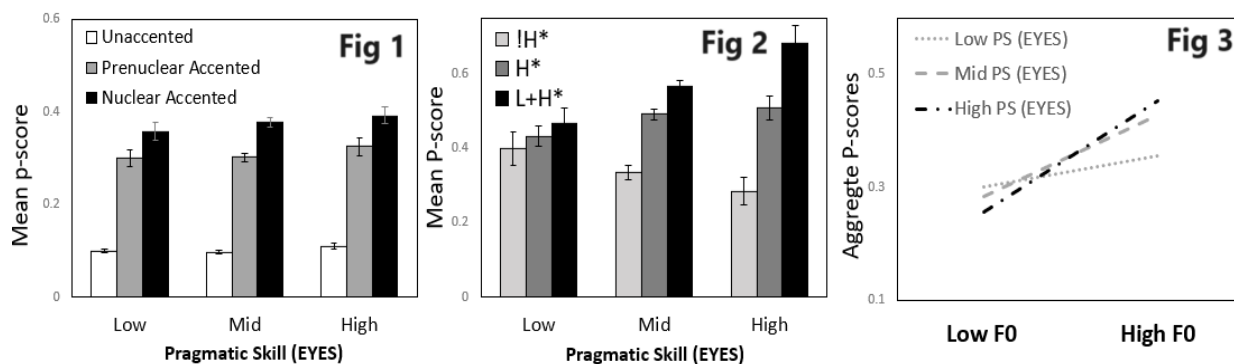
Fig 1: P-scores (proportion words marked prominent) as a function of accent status and EYES scores (higher EYES scores indicate greater pragmatic skill related to empathy). **Figure 2**: P-scores as a function of accent type and EYES scores. **Figure 3**: Aggregate P-scores for words marked by L+H* with f0 peaks that are 1 standard deviation below/above the mean peak height. [Note: binning of EYES scores is for visualization only; this was a continuous variable in all mixed-effects regression models]

References

Arvaniti, A., Gryllia, S., Zhang, C., Marcoux, K. (2022). Disentangling emphasis from pragmatic contrastivity in the English H* ~ L+H* contrast. *Proc. Speech Prosody 2022*, 837-841.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001a). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *J. of Autism and Developmental Disorders*, 31(1), 5-17.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I. (2001b). The "Reading the Mind in the Eyes" Test revised version. *J. of Child Psychology & Psychiatry*, 42, 241-251.

Beckman, M., & Ayers Elam, G. (1997). Guidelines for ToBI labeling (Version 3). Unpublished manuscript, The Ohio State University.

Bishop, J., Kuo, G., & Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from Rapid Prosody Transcription. *Journal of Phonetics,* 82, 100977.

Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2010). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1(2), 425–452.

Cole, J., Hualde, I., Smith, C., Eager, C., Mahrt, T., de Souza, R. (2019). Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of Phonetics* 75, 113-147.

Cole, J., Mahrt, T., Hualde, J.I. (2014) Listening for sound, listening for meaning: Task effects on prosodic transcription. Proc. Speech Prosody 2014, 859-863.

Grice, M., Krüger, M. & Vogeley, K. (2016). Adults with Asperger syndrome are less sensitive to intonation than control persons when listening to speech. Cult. Brain 4, 38–50.

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *J. of Developmental Disorders*, 37(9), 1679–1690.

Jun, S-A. & Bishop, J. (2015). Prominence in relative clause attachment: Evidence from prosodic priming. In L. Frazier & E. Gibson (Eds.): ***Explicit and implicit prosody in sentence processing: Studies in honor of Janet Dean Fodor***. Springer.

Ladd, D.R. & Morton, R. (1997). The perception of intonational emphasis: continuous or categorical? Journal of Phonetics, 25, 313-342.

Xiang, M., Grove, J., & Giannakidou, A. (2013). Dependency-dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology*, 4, 708.