THE INFLUENCE OF EMPATHY AND AUTISTIC-LIKE TRAITS IN PROMINENCE PERCEPTION

Riccardo Orrico, Stella Gryllia, Jiseung Kim, Amalia Arvaniti

Radboud University, Netherlands
riccardo.orrico@ru.nl, stella.gryllia@ru.nl, jiseung.kim@ru.nl, amalia.arvaniti@ru.nl

ABSTRACT

We report an investigation of how individual listeners' characteristics influence the assessment of prominence. Following the RPT paradigm, 82 Standard Southern British English speakers listened to utterances and marked the words that sounded prominent to them. We focused on words produced with either H* or L+H*, which appeared in either contrastive or non-contrastive pragmatic conditions. Both pragmatics (contrastive vs. non-contrastive accent) and phonetics (H* vs. L+H*) systematically guided participants' judgment. L+H*s and contrastive accents were more likely to be marked as prominent, but some listeners prioritized phonetics and others pragmatics, and these differences were related to their empathy and autistic-like traits: the higher a listener's empathy, the more they relied on pragmatics, while the higher the autistic-like traits the smaller the distinction between the two accents. Our results suggest that both acoustic and pragmatic salience are instrumental for prominence assessment, but listeners weigh these cues in different ways.

Keywords: individual characteristics, prominence, pragmatics, Rapid Prosody Transcription, English

1. INTRODUCTION

A growing body of research has been devoted to the study of individual variability, pointing to a heterogeneous nature of language processing. This variability has been linked to individual cognitive styles [1], such as empathy and autistic-like traits. Empathy has been linked to the ability to interpret the interlocutor's intentions and therefore to pragmatic skills [2]. With respect to prosody, empathy has been shown to correlate with the ability to derive meaning from intonation [3, 4]. On the other hand, individuals with more autistic-like traits show poorer pragmatic abilities [2], heightened sensitivity to acoustic cues [1, 5], and poorer high order cognitive ability [6] compared to those with fewer such traits.

Here we present a Rapid Prosody Transcription (RPT) study [7, 8] that investigated how individual differences in empathy and autistic-like traits affect prominence assessment. Previous RPT studies [9, 10, 11] have reported individual variability in RPT responses. The results of [9] and [10] suggest that listeners differ as to how they weigh the available cues when assessing prominence, while [11] reports that individuals with more autistic-like traits show poorer prominence perception. These differences may be due to the broad definition of prominence as the property of standing out in an utterance by means of phonetic, phonological, pragmatic, or lexical cues. Given the interplay of so many co-occurring cues, individuals with different cognitive styles may prioritize different cues when making judgments about prominence.

We examined prominence ratings during RPT as a function of pitch accent type (focusing on H* and L+H*), considering separately each accent's F0 shape and pragmatic function (for details, see 2.2.1). H* and L+H* were chosen because their prominence relates to both their shape and their pragmatics [12]. Further, since prominence perception does not rely exclusively on psychophysical properties [13], we focused on Standard Southern British English (SSBE), to avoid possible regional differences in the accents' function, as function plays a part in how the accents' prominence is assessed (cf. [14]). We hypothesized that neurotypical individuals with different processing styles would employ different strategies in assessing prominence. Processing styles were operationalized as degree of empathy and extent of autistic-like traits as measured by the Empathy Quotient (EQ) and the Autism Quotient (AQ) questionnaires respectively (for details, see 2.3). We predicted that, when assessing prominence, individuals with high EQ would rely more on pragmatic information, while individuals with high AQ would rely more on phonetics.

2. METHODS

2.1. Participants

We used Prolific (https://www.prolific.co/) to recruit 85 native monolingual speakers of SSBE without (self-reported) speech or hearing disorders: they were all brought up in monolingual households and had learned languages other than English through formal instruction only. Participants were paid £18 to complete the experiment. We report results from 82 participants (47 F, 19–50 years of age), after excluding one participant who faced technical problems and two participants with more than 10% of outlier responses (they selected all words in utterances with 1-5 words, or more than 85% of words in utterances with more than 5 words).

2.2. RPT Stimuli

The stimuli were 86 utterances produced by eight SSBE speakers (5 F). Twenty-two utterances were excerpts from read speech in a reading task, while the remaining 64 were from various spontaneous speech tasks. The total number of words in the 86 utterances was 879. The utterance length was 3–24 words (Mean: 10.2; SD: 4.7).

2.2.1 Categorization of the stimuli

Words with high and rising accents were categorized using phonetic and pragmatic criteria independently of each other. Phonetically, rising accents were annotated as L+H* if a deliberate dip was present at the accented syllable onset, or if the coda preceding a voiceless onset had low F0. Otherwise, accents were annotated as H*. Accents in absolute utterance initial position were classified as H* [15]. Accents followed by uptalk were not included as it would be impossible to separate the effect of the accent from that of the utterance-final rise on prominence assessment.

Pragmatically, words bearing H* or L+H* were classified as *contrastive* or *non-contrastive* within the context of their utterance. Only the orthographic transcripts were used for this categorization to avoid circularity between phonetic shape and pragmatic function in the annotation process. An accent was annotated as *contrastive* when it highlighted one item out of a small set of (mentioned or readily inferred) alternatives in discourse. Focus particles (e.g., *just*, *only*) and negative expressions (e.g., *don't* in *I don't know*) were also marked as contrastive. Finally, items were considered contrastive if they were part of an explicit parallelism in discourse (e.g., in *I want tea as well as coffee*, *tea* and *coffee* would be considered contrastive). All other high and rising accents were considered non-contrastive. This classification cuts across the categories *new*, *given*, *topic*, and *focus*, as it is possible to contrastively accent both new and given items as well as topics and foci, and (phonetic) accent type does not appear to vary systematically based on these pragmatic distinctions [16].

2.3. Additional tests

Participants completed the EQ [17] and AQ [18] questionnaires. The EQ score ranges from 0 to 80; it is based on 40 responses on a 4-point scale and represents the individual's level of emotional and cognitive empathy. The AQ score ranges from 0 to 50; it is based on 50 responses on a 4-point scale and is not a diagnostic test; rather, it positions adults of normal intelligence along a continuum measuring five traits associated with the autism spectrum disorder: social skills, communicative skills, attention to detail, attention switching, and imagination.

2.4. Experimental procedure

The study ran online using ROLEG developed at Radboud University. The self-paced RPT task had 2 practice trials followed by 86 trials in five blocks, giving participants 4 prompts for self-paced breaks. In each trial, participants would first hear an utterance while seeing it in writing on screen; then they heard a second repetition during which they could select which words sounded "highlighted, prominent, important, stressed, or emphasised" to them. The transcript did not include punctuation or capitalization other than apostrophes in contractions and possessives, and capitalization in proper nouns and the pronoun 'I'. Participants were instructed to select as many words as they saw fit but had to select at least one in each stimulus to proceed to the next trial. After the RPT task, participants were given access to the EQ and the AQ questionnaires on PsyToolkit (https://www.psytoolkit.org/, [19, 20]). Participants were asked to complete them within seven days after the RPT task to be remunerated.

2.5. Statistical analysis

The binary RPT responses (word marked as prominent or not) were analysed in three Generalized Linear Mixed-Effects Models (GLMM) using the lme4 package [21] in R [22]. The first model tested the relation between the RPT responses and the accents' phonetic and pragmatic classification. It included as fixed effects PHONETICS (H*, L+H*), PRAGMATICS (contrastive, non-contrastive) and their interaction. The random effects included random intercepts for TALKER, PARTICIPANT and ITEM (the accented word), and by-subject random slopes for ACCENT and PRAGMATICS. Two additional models tested the relations between the RPT responses, the accents' phonetic and pragmatic categorization, and the EQ and AQ scores. The EQ model included as fixed effects PHONETICS, PRAGMATICS, EQ, and the interaction between PRAGMATICS and EQ. The AQ model included PHONETICS, PRAGMATICS, AQ, and the interaction between PHONETICS and AQ. Both models included the random intercepts for SPEAKER, PARTICIPANT, and ITEM as random effects. The effects of EQ and AQ were tested in two separate models to avoid collinearity issues detected in preliminary analysis.

3. RESULTS

3.1. The effect of accent and pragmatics

Details about the data and analysis are available at osf.io/euwn8. The output of the first GLMM showed that both PHONETICS and PRAGMATICS were significant predictors of the binary responses (word selected as prominent or not): words annotated as L+H* or contrastive were more likely to be selected as prominent than those annotated as H* or non-contrastive ($\beta_{PHONETICS: L+H*}$ = 1.33, SE = 0.32, z = 4.11, *p* <.0001; $\beta_{PRAGMATICS: Contr}$ = 1.41, SE = 0.28, z = 5.02, *p* <.0001). The interaction between PHONETICS and PRAGMATICS was not significant (β = -0.34, SE = 0.46, z = -0.74, *p* > .4).

Fig. 1 shows the distribution of p-scores (the percentage of participants selecting a word as prominent) for the four combinations of PHONETICS and PRAGMATICS. The categories *H* non-contrastive* and *L+H* contrastive* peaked towards the low and high end of the p-score distribution respectively, while *H* contrastive* and *L+H* non-contrastive* largely overlapped and showed more variable responses than the first two categories.

3.2. Individual response patterns

Following [9] and [23], we observed individual patterns by inspecting the by-subject random slopes for PHONETICS and PRAGMATICS. This revealed that individual participants prioritized different criteria. To illustrate the three prototypical patterns that emerged, we show in Fig. 2 the random slope values by PHONETICS (x-axis) and PRAGMATICS (coloured dots) for participants P1, P2, and P3; the y-axis shows the predicted probabilities of the GLMM output variable, i.e., whether a word was marked as prominent (1) or not (0). P1 reliably differentiated the pragmatic conditions (represented by the vertical distance between the orange and purple dots); at the same time, this participant showed a less reliable distinction between the two phonetic conditions (represented by the relatively horizontal dashed lines linking the dots across the phonetic conditions, H* and L+H*), suggesting that they relied mainly on pragmatic information to mark prominence. On the other hand, P3 showed a robust distinction between H* and L+H* (as represented by the steeper dashed lines) but exhibited a less robust distinction based on pragmatics, especially in the case of L+H* accents (as represented by the closer vertical distance between the orange and purple dots for L+H*); this suggests they relied mainly on phonetic information to mark prominence. Finally, P2 differentiated the slope values for pragmatics *and* phonetics to a similar extent, indicating that they relied on both criteria relatively uniformly.

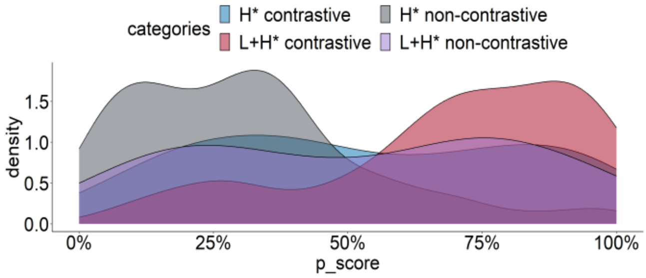

Figure 1: Density plots of p-scores for the four combinations of PHONETICS and PRAGMATICS.

In order to quantify the extent to which a participant relied on a specific criterion (phonetics or pragmatics), we extracted the *slope values* independently for the two variables (PHONETICS and PRAGMATICS) and added the individual slope values to the general slope of the effects predicted by the GLMM. The obtained *values* provide a quantitative measure of the relevance of the two criteria for each participant's responses: the higher the value of a criterion for a given participant, the more they relied on that criterion. Fig. 3 illustrates these *values* and shows that, while the three prototypical patterns can be identified (as shown in the three panels), the individual responses form a continuum extending from those who relied more on pragmatics (left panel; shown in shades of red) to those who relied more on accent (right panel; shown in shades of blue).

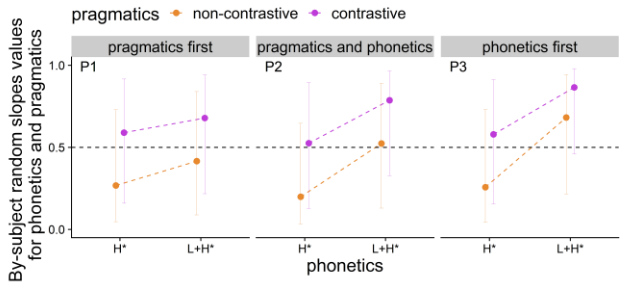

Figure 2: Random slope values for three participants showing prototypical response patterns: those who prioritized pragmatics (P1), those who prioritized phonetics (P3), and those who relied equally on both (P2). Points indicate slope values extracted from the GLMM; vertical lines indicate confidence intervals.

3.2.1 The effect of EQ and AQ

EQ participants' scores ranged from 16 to 70 (Mean = 43.2, SD = 12.3). AQ scores ranged from 1 to 42 (Mean = 19.2, SD = 9.9). Cronbach's Alpha [24] showed high reliability for both (0.95 for EQ; 0.93 for AQ). A Pearson correlation showed that AQ and EQ were negatively correlated (r(80) = -.58, p <.001).

The GLMM output revealed a significant interaction between EQ and PRAGMATICS (β = .006, SE = .003, z = 2.22, *p* < .03), indicating that the higher a participant's EQ, the more likely they were to mark contrastively accented words as prominent. AQ and PHONETICS also interacted (β = -.008, SE = .004, z = -2.20, $p < .03$), so that the higher a participant's AQ, the more likely they were to mark H*s as prominent, thereby reducing the distinction between the two phonetic conditions. The effects of EQ and AQ are illustrated in Fig. 4.

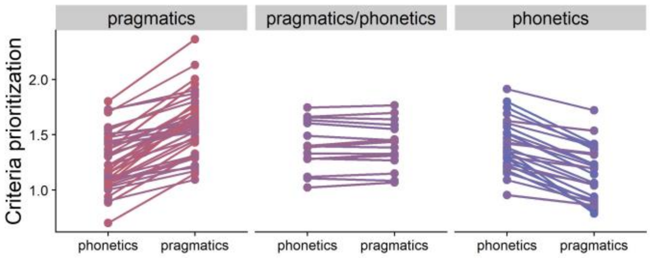

Figure 3: Slope values for PHONETICS and PRAGMATICS for individual responses. Gradience from red (left) to blue (right) reflects a continuum from prioritizing pragmatics to prioritizing phonetics.

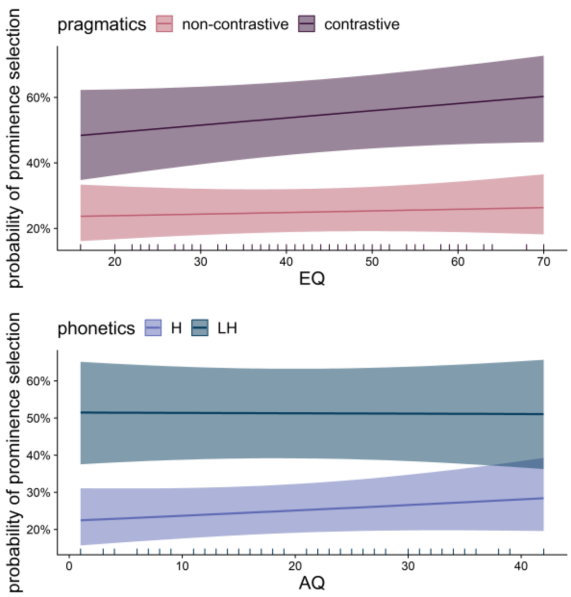

Figure 4: Predicted probabilities of RPT responses for the interaction between PRAGMATICS and EQ (top) and PHONETICS and AQ (bottom).

4. DISCUSSION AND CONCLUSIONS

We tested the relationship between prominence assessment and cognitive styles (operationalized as empathy and autistic-like traits) by considering as predictors the phonetics differences between H* and L+H* accents and the pragmatic information they convey (contrastive or non-contrastive).

The aggregate results (Fig. 1) showed that both phonetic and pragmatic information was exploited when making judgments about prominence, and the two factors worked synergistically in driving these judgments. This suggests that phonetic information was not a sufficient cue to prominence; rather, the role of accent shape was weighed together with other cues within the utterance.

Our results also suggest that these cues were not exploited by the participants in the same way. While both phonetic and pragmatic cues were used by all participants to some extent, many (though not all) individuals prioritized either phonetics or pragmatics to some extent. These different priorities may be the lead cause behind the individual differences in RPT responses reported in [9] and [10]: variation can be explained by individual characteristics like those examined in the present study. Participants with higher EQ were more likely to rely on pragmatics than those with lower EQ, supporting previous findings that empathetic individuals are highly sensitive to pragmatic information [3, 4]. In addition, participants with higher AQ made less reliable distinctions between the phonetic conditions. This finding is in line with previous studies: [11] reported poorer prominence perception for high AQ listeners (though the effect was mainly observed in unaccented words) and linked this result to lower sensitivity to the prosody-meaning mapping. In our study, high AQ listeners showed an increased tendency to mark accented words as prominent, regardless of accent type. This can be interpreted as a tendency to perceive as prominent any prosodically strong word and a poorer integration of F0 shape with context.

Taken together, our results show that individual characteristics related to cognitive styles shape the way individuals process linguistic information and predict the extent to which listeners evaluate cues of different nature (as opposed to relying only on acoustic cues) when making judgments about prominence. Avenues for future research include testing additional individual characteristics that have been argued to play a role in linguistic processing (e.g., musicality [25] and language background [4]). It is also important to investigate whether these individual differences reflect distinct processing styles or simply the way participants interpret the RPT task. Revisiting the definition of prominence by taking into account crucial grammatical and language-specific aspects (cf. [13]) are necessary steps to tackle this question: a rigorous understanding of what prominence is would allow for a thorough investigation of its processing mechanisms.

5. ACKNOWLEDGEMENTS

The financial support of the European Research Council through grant ERC-ADG-835263 (SPRINT) to Amalia Arvaniti is hereby gratefully acknowledged. We thank Na Hu, Katherine Marcoux, and Sofia Sialiaki for their help with the preparation of the experiment and participant recruitment.

6. REFERENCES

[1] Yu, A. C., Zellou, G. 2019. Individual differences in language processing: Phonology. *Annual Review of Linguistics* 5, 131–150.

[2] Baron-Cohen, S. 2004. *The Essential Difference*. Penguin UK.

[3] Esteve-Gibert, N. J., Schafer, A., Hemforth, B., Portes, C., Pozniak, C., D'Imperio, M. 2020. Empathy influences how listeners interpret intonation and meaning when words are ambiguous. *Memory & Cognition* 48(4), 566–580. doi: http://doi.org/10.3758/s13421-019-00990-w

[4] Orrico, R., D'Imperio, M. 2020. Individual empathy levels affect gradual intonation-meaning mapping: The case of biased questions in Salerno Italian. *Laboratory Phonology* 11(1), 1–39.

[5] Yu, A. C. 2010. Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PloS One* 5(8), [Online]. doi: https://doi.org/10.1371/journal.pone.0011950

[6] Stewart, M. E., Ota, M. 2008. Lexical effects on speech perception in individuals with "autistic" traits. *Cognition* 109(1), 157–162.

[7] Cole, J., Mo, Y., Hasegawa-Johnson, M. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1(2), 425–52.

[8] Cole, J., Shattuck-Hufnagel, S., 2016. New Methods for Prosodic Transcription: Capturing Variability as a Source of Information. *Laboratory Phonology* 7(1): 8. doi: 10.5334/labphon.29

[9] Baumann, S., Winter, B. 2018. What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *Journal of Phonetics* 70, 20-38.

[10] Arvaniti, A., Gryllia, S., Zhang, C., Marcoux, K. P. 2022. Disentangling emphasis from pragmatic contrastivity in the English H*~L+ H* contrast. *Proceedings of Speech Prosody 2022*. doi: 10.21437/SpeechProsody.2022-170

[11] 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.

[12] Ladd, D. R., Morton, R. 1997. The perception of intonational emphasis: continuous or categorical?. *Journal of phonetics* 25(3), 313–342.

[13] Ladd, D. R., Arvaniti, A. 2023. Prosodic prominence across languages. *Annual Review of Linguistics* 9:171–193.

[14] Arvaniti, A., Garding, G. 2007. Dialectal variation in the rising accents of American English. *Papers in Laboratory Phonology*, 9: 547–576.

[15] Brugos, A., Shattuck-Hufnagel, S., Veilleux, N. 2006. *Transcribing Prosodic Structure of Spoken Utterances with ToBI* (MIT Open Courseware). http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/index.htm.

[16] Molnár, V. 2002. Contrast from a contrastive perspective. *Language and Computers* 39, 147-162.

[17] Baron-Cohen, S., Wheelwright, S. 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders* 34(2), 163–175.

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

[19] Stoet, G. 2010. PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods* 42(4), 1096–1104.

[20] Stoet, G. 2017. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology* 44(1), 24–31.

[21] Bates, D., Mächler, M., Bolker, B. Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 1(1), 1–48. Available: https://www.jstatsoft.org/v067/i01

[22] R Core Team, *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020, [Online]. Available: http://www.r-project.org/.

[23] Drager, K., Hay, J. 2012. Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change* 24(1), 59–78.

[24] Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: https://doi.org/10.1007/BF02310555

[25] Schön, D., Magne, C., Besson, M. 2004. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* 41(3), 341–349.