Language and Cognition (2025), 17, e9, 1–25
doi:10.1017/langcog.2024.62

ARTICLE

Individual variability and the H* ~ L + H* contrast in English

Riccardo Orrico, Stella Gryllia, Jiseung Kim and Amalia Arvaniti

Centre for Language Studies, Radboud University, Nijmegen, The Netherlands
Corresponding author: Riccardo Orrico; Email: riccardo.orrico@ru.nl

(Received 04 August 2023; Revised 25 September 2024; Accepted 29 September 2024)

Abstract

The H* ~ L + H* pitch accent contrast in English has been a matter of lengthy debate, with some arguing that L + H* is an emphatic version of H* and others that the accents are phonetically and pragmatically distinct. Empirical evidence is inconclusive, possibly because studies do not consider dialectal variation and individual variability. We focused on Standard Southern British English (SSBE), which has not been extensively investigated with respect to this contrast, and used Rapid Prosody Transcription (RPT) to examine differences in prominence based on accent form and function. L + H*s were rated more prominent than H*s but only when the former were used for contrast and the latter were not, indicating that participants had expectations about the form–function connection. However, they also differed substantially in which they considered primary (form or function). We replicated both the general findings and the patterns of individual variability with a second RPT study which also showed that the relative prioritization of form or function related to participant differences in empathy, musicality and autistic-like traits. In conclusion, the two accents are used to encode different pragmatics, though the form–function mapping is not clear-cut, suggesting a marginal contrast that not every SSBE speaker shares and attends to.

Keywords: H* ~ L + H*; phonetics; pragmatics; individual variability; empathy; autistic-like traits; musicality

1. Introduction

This paper is concerned with the role of individual variability in the perception of the accentual contrast between H* and L + H* in British English and what this variability tells us about the contrast itself. The phonological distinction between H* and L + H* was posited, using these terms, by Pierrehumbert (1980). Although Pierrehumbert's intonation system focused on American English, its categories, including H* and L + H*, are often taken to apply to English in general, because of intonation's 'high degree of uniformity […] across most varieties of English' (Ladd, 2022: 249; for a similar stance, see also Ladd, 2008, ch. 3; Gussenhoven, 2016). Below we explain the nature of the contrast and review the empirical evidence for it.

Phonetically, both H* and L + H* involve a pitch peak aligned with the accented syllable. L + H* is realized as a sharp rise starting low in the speaker's range, while H* starts from a higher point and rises more gradually (Brugos et al., 2006, ch. 2.5). Production studies further suggest that the accentual peak is lower in H* than L + H* (Iskarous et al., 2023; 2024), but possible alignment differences remain uncertain: a comparison of Silverman and Pierrehumbert (1990) and Arvaniti and Garding (2007) suggests earlier peak alignment for L + H*, but Steffman et al. (2022) report the opposite. The pragmatics of the accents are addressed in Pierrehumbert and Hirschberg (1990): H* signals to the listener to add the accented item to the mutual belief space, while L + H* signals that the accented item, *and not some other relevant item*, should be in the mutual belief space. Thus, both H* and L + H* signal speaker commitment, though for L + H* this commitment is accompanied by *contrast*. In Pierrehumbert and Hirschberg (1990), contrast is not equated with correction but construed as any proposition conveying that an alternative proposition does not hold (cf. Krifka, 2008; Molnár, 2002). For instance, in (1), reproduced from Pierrehumbert and Hirschberg (1990: 297), the proposition in B's turn is contrastive because it selects one of the many possible reasons why the lamp under discussion stands up: thus, in B's turn both *weighs* and *ton* would typically bear a L + H* accent, though *It weighs a ton* does not correct a previous commitment about the lamp (cf. Bartels & Kingston, 1996). Following Pierrehumbert and Hirschberg (1990), in the remainder of the paper, we will use *contrast* to refer to both corrective and contrastive uses of accentuation.

(1) B is unpacking a new desk lamp.
A: But how does it [the desk lamp] stand up?
B: Feel that base. It *weighs* a *ton*.

Although the distinction between H* and L + H* is often treated as undisputed, it is not unequivocally supported by empirical evidence. Production studies indicate that both H* and L + H* can be used with contrastive and non-contrastive functions to mark both foci and topics. This applies both to controlled speech, investigated by Metusalem and Ito (2008) using a Discourse Completion Task, and the spontaneous yet formal style investigated by Hedberg and Sosa (2007) and Im et al. (2023), who examined political debates and TED talks, respectively. Perception studies support these findings, showing that L + H* is more likely to lead to a contrastive interpretation but the use of H* does not preclude it. For instance, Watson et al. (2008) used eye-tracking to investigate H* and L + H* and concluded that L + H* creates a strong bias for contrast, but H* is compatible with both new and contrastive referents. Stronger evidence for the contrastive use of L + H* comes from studies in which it is followed by deaccenting. Thus, in Ito and Speer's (2008) eye-tracking study, the presence of a L + H* on the adjective in adjective–noun pairs led to faster processing of contrastive referents, provided the L + H* was followed by a deaccented noun. Similarly, Kurumada et al. (2012) found that in utterances like *It looks like a zebra*, H*s on the verb and following noun are interpreted affirmatively (*it looks like a zebra, and it is*), while a nuclear L + H* on the verb followed by noun deaccenting triggers a contrastive interpretation by evoking a negative alternative (*it looks like a zebra, but it*

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

is not). Given such results which show variability and sensitivity to context, it is not surprising that H* and L + H* were the most frequent point of disagreement among early MAE_ToBI annotators (Syrdal & McGory, 2000). For similar reasons, Brugos et al. (2006, ch. 2.5) caution that 'both [H* and L + H*] can be used in a variety of contexts, and a specific context will not necessarily lead all speakers to select the same intonation contour'. The recognition of this variability, however, goes against the assumption that the two accents form distinct phonological categories, since their differences in form do not result in consistent differences in pragmatic interpretation.

The picture becomes more complicated if one considers recent studies on dialectal variation which cast doubt on the assumption that the intonational system of English is largely uniform. For American English, Burdin et al. (2018) report differences in both the frequency and realization of L + H* among speakers of Jewish English, African American English and Appalachian English (see also Holliday, 2021a, 2021b). Crucially, other studies suggest that some American English varieties may lack the H* ~ L + H* contrast altogether (see Arvaniti & Garding, 2007, on Minnesota English; Kim & Arnhold, 2024, on Canadian English).

The approach to the H* versus L + H* contrast in the literature on British English is even more varied. Studies that assume a uniform intonation system across English varieties adopt Pierrehumbert's analysis, largely concluding that H* and L + H* are not distinct (e.g., Dilley et al., 2005; Ladd, 2008; Ladd & Morton, 1997; Ladd & Schepman, 2003).[1] Ladd and Morton (1997) found that stepwise increases in peak height, intended to create a continuum from H* to L + H*, were perceived gradiently. Ladd and Schepman (2003) report that consecutive high accents show a consistent F0 dip aligned with the onset of the second accent and that listeners use the location of this dip to determine syllable boundaries. Thus, they conclude that the F0 dip should be part of the representation of all high accents, a proposal that implies British English does not make a distinction between H* and L + H*. Taken together, these findings support the contention of Ladd (2008, *inter alia*) that H* and L + H* are not distinct categories.

A final complication relates to individual variability. It is well established that individuals vary in how they process linguistic information in ways that relate to cognitive characteristics, such as memory and attention (for reviews, see Kidd et al., 2018; Yu & Zellou, 2019). Such effects are likely to be considerable for intonation, where both production and perception are probabilistic (Calhoun, 2010; Kurumada & Roettger, 2022) and unconstrained by lexical meaning. If so, then differences between individuals could have a sizeable impact on how intonation contrasts are produced and perceived.

In the present study, we investigated three traits that could be sources of individual variability: musicality, autistic-like traits and empathy. We chose these traits because they have been linked to the processing of phonetic and pragmatic information, both of which are of relevance in the perception of pitch accents. Studies on music and language have linked musicality to the ability to discriminate and reproduce phonetic differences: individuals with greater musical ability are better able to detect pitch differences in speech (Schön et al., 2004, on French; Cui &

[1]We note that this conclusion is not espoused by many British analyses of intonation which, instead, see the rise as epiphenomenal and thus analyze the accents as falls, with low falls largely corresponding to H* and high falls largely corresponding to L + H*; see O'Connor and Arnold (1973), Cruttenden (1997), Grabe et al. (2000), Gussenhoven (2004), and (Gussenhoven, 2016) for relevant discussions.

Kuang, 2019, on English) and imitate stress in L2 (Cason et al., 2020, on French learners of English). Autistic-like traits in neurotypical adults have also been linked to the processing of phonetic cues, in that more autistic-like traits correlate with weaker integration between phonetic cues and higher-order information (Stewart & Ota, 2008; Yu & Zellou, 2019). As Bishop et al. (2020) show, this sensitivity extends to prosody, in that individuals with fewer autistic-like traits are more attuned to the prosody-meaning mapping. Finally, more empathetic individuals show higher sensitivity toward pragmatic information, most likely as a result of their greater ability to understand what other people feel or think (Baron-Cohen & Wheelwright, 2004). The role of empathy in processing pragmatics extends to intonation, with more empathetic individuals attending more to intonation information in order to extract meaning in both L1 and L2 (on L1, see Esteve-Gibert et al., 2020; Orrico & D'Imperio, 2020; on L2, see Casillas et al., 2023).

In sum, while the phonetic difference between H* and L + H* is undisputed, scholars disagree on the validity of the H* ~ L + H* contrast, in that a clear distinction in the accents' functions is not supported by empirical evidence, thereby casting doubt on their forming a phonological contrast. The disagreements may relate to dialectal and individual differences, dimensions of variation that have sometimes been underestimated. To address both, here we focus on the processing of H* and L + H* in Standard Southern British English (henceforth SSBE), a variety on which there have been relatively few empirical studies, and additionally consider the role of individual variation. Thus, our results contribute to the debate on the status of H* and L + H* across English varieties and shed light on the role of individual variation in the processing of intonation.

Specifically, we examined H* and L + H* by adapting the Rapid Prosody Transcription (RPT) paradigm to our purposes. In typical RPT (Cole et al., 2010), linguistically untrained participants listen to utterances and mark on their orthographic transcripts the words they perceive as prominent; prominence is subsequently investigated through post-hoc analysis of parameters expected to affect prominence ratings (e.g., Baumann & Winter, 2018; Bishop, 2016; Cole et al., 2010; Im et al., 2023). Our own aim in using RPT was not to investigate possible cues to prominence but to explore the phonological status of H* and L + H*. We chose RPT because it is an indirect and not openly metalinguistic task that does not require participants to make difficult judgments about the meaning of the accents, but allows us to assess how different the accents sound to them. For this reason, our study is not concerned with all the possible parameters that could affect the relative prominence of L + H* and H* or with the prominence of these accents relative to other accents in the stimuli.

We used the prominence ratings of H* and L + H* as a way of understanding how salient the difference between the two accents is for SSBE listeners. We reasoned that if H* and L + H* are distinct categories, this would be reflected in bimodal prominence distributions, with L + H*-accented words being consistently rated more prominent than H*-accented items. In contrast, if H* and L + H* are variants of one category, the differences between them would be less salient, and this would be reflected in substantially overlapping, and potentially unimodal, prominence distributions (cf. Boomershine et al., 2008, for similar findings for segmental contrasts). As relative prominence can be related to function (cf. Im et al., 2023), we further hypothesized that if L + H*s are more prominent than H*s *regardless* of their function, this would support Ladd's (2008) contention that L + H*s are emphatic

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

versions of H*s and thus that L + H* and H* form a continuum (cf. Ladd & Morton, 1997). On the other hand, if prominence judgments take into account accent function (i.e., L + H*s are judged as more prominent only when they are also contrastive), this would suggest that phonetics and function are interpreted together, supporting the phonological separation of the two accents (Pierrehumbert & Hirschberg, 1990).

The above is the basis for the first study, reported in Section 2. The second study (Section 3) is a replication of the first, in which we additionally examined potential sources of individual variability in prominence assessment.

2. RPT study 1

2.1. Methods

2.1.1. Participants

Sixty participants, out of 85 recruited through Prolific (https://www.prolific.co/), completed the task. We report data from 47 of these after applying exclusion criteria based on language- or performance-related issues. Seven participants self-reported not being brought up in a monolingual SSBE household; the other six met one of the following exclusion criteria which indicated that they may not have given the task due attention: in more than 10% of the stimuli, they marked as prominent (i) all the words in utterances with up to five words, or (ii) more than 85% of the words in utterances with more than five words.

The 47 participants (19–47 years old, M = 33.77, SD = 7.7) were functional monolingual SSBE speakers; i.e., they had learned other languages through formal instruction only. Thirty were female, sixteen were male, and one was non-binary. None reported any history of speech or hearing disorders. Participants took approximately 30 minutes to complete the task and were remunerated for their participation.

2.1.2. Procedure

The study ran on Roleg, an online platform developed at Radboud University (https://www.roleg.nl/TaalExperiment/). It comprised 2 practice and 86 main trials, and included 4 self-paced breaks. In each trial, participants would first hear an utterance while seeing its transcript on screen; then they heard a second repetition during which they were asked to mark prominent words by clicking on a checkbox next to each word (Figure 1). We used typical RPT instructions (e.g., Cole et al., 2010), asking participants to mark words they heard as 'prominent, stressed, highlighted, important or emphasized'. They could select as many words as they saw fit but had to select at least one to proceed to the next trial (see https://osf.io/f7w9c/ for the full instructions). The transcripts shown to the participants did not include punctuation or capitalization, except that apostrophes were retained in contractions and possessives, and proper nouns and the pronoun *I* were capitalized.

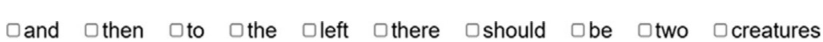

Figure 1. Transcript of a sample stimulus used in the RPT task.

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

2.1.3. Stimuli

The stimuli were 86 utterances selected from the data of 5 female and 3 male SSBE speakers, aged 18–54 years (M = 29.25, SD = 12.28), who had been recorded for a production study that included both read and unscripted speech (Kim et al., 2024); none was a professional talker. All stimuli were autonomous syntactic and prosodic entities; 22 were extracted from read speech and 64 from unscripted speech (see https://osf.io/f7w9c/ for additional information about the elicitation tasks). The total number of words in the 86 utterances was 879. The utterance length range was 3–24 words (M = 10.2; SD = 4.7), or 5–34 syllables (M = 13.3; SD = 6.5). The duration range was 0.52–6.8 s (M = 2.6, SD = 1.4).

Before the study, the stimuli were annotated both for the phonetic identity and pragmatics of the accents of interest here, namely 287 accents that had high or rising F0. The two annotations were done independently to avoid each influencing the other. This is explained in more detail below.

Phonetically, high and rising accents were categorized as H* or L + H*; the annotation was based on their F0 shape without taking into account their function in the utterance. Prosodic words were annotated as carrying a L + H* accent if they showed a deliberate F0 dip at the onset of the accented syllable. The dip had to be at a relatively low level in the utterance range (i.e., not the result of high and rising F0). Prosodic words with accented syllables that started with voiceless onsets were annotated as L + H*s if it could be ascertained that the preceding syllable deliberately ended in low F0; otherwise, the accent was annotated as H*. No distinction was made between downstepped !H* and non-downstepped H* (though see 2.1.4). Following MAE_ToBI conventions (Brugos et al., 2006), and in the absence of guidelines specific to SSBE, accents in the absolute utterance-initial position were classified as H*s. Accents immediately followed by uptalk were not considered in the analysis because it would not be possible to separate the effect of the accent from that of uptalk in the assessment of prominence. Finally, accents other than H* and L + H* were not considered in the present analysis, as the aim was not to conduct a full prominence study.

Pragmatically, all lexical items in the stimuli were categorized as *contrastive* or *non-contrastive* based solely on the fully punctuated orthographic transcript of the utterances.[2] Items were deemed contrastive if, in context, they generated a small set of explicitly mentioned or easily inferred alternatives (Pierrehumbert & Hirschberg, 1990; Krifka, 2008, see section 1). This classification cuts across the categories new, given, topic and focus, as contrast is orthogonal to other information structure dimensions (cf. Molnár, 2002). As an illustration, in (2), *under* is marked as contrastive since it is one of the possible ways to go around the lilies; in (3) *people* and *bags* were marked as contrastive, as they were two entities of a parallel construction and therefore contrasting with each other. In addition, focus particles, such as *just* and *only*, and negative expressions (e.g., *do not* in *I do not know*) were also marked as contrastive. All other items were marked as non-contrastive.

[2]We are aware that implicit prosody could influence pragmatic annotation (see Breen, 2014, for a review). To minimize its effect, annotators were instructed to avoid using their own implicit (or read-aloud) renditions of the utterances as a means of pragmatic classification. Additionally, they relied on a pragmatic annotation system devised with the help of two pragmatics experts, Chris Cummings and Hannah Rohde.

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

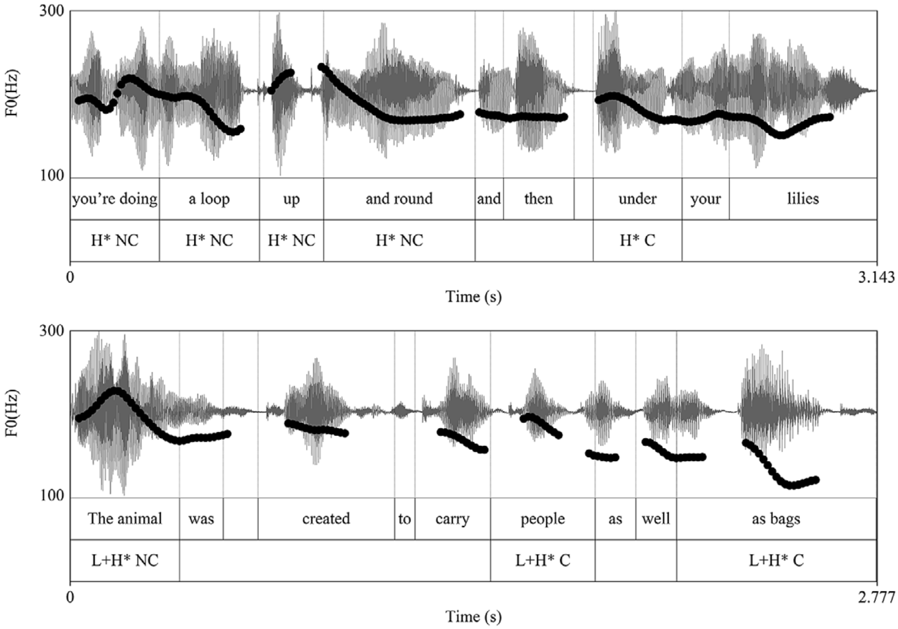

Figure 2. Waveforms and F0 tracks of two stimuli produced by two female speakers illustrating H* and L + H* accents with contrastive (C) or non-contrastive (NC) function.

The pragmatic categorization was then matched with the phonetic categorization to give four categories: contrastive L + H*s, non-contrastive L + H*s, contrastive H*s, and non-contrastive H*s. Figure 2 shows the waveforms and F0 tracks of examples (2) and (3).

(2) You're doing a loop up and round and then **under** your lilies

(3) The animal was created to carry **people** as well as **bags**

The phonetic and pragmatic annotations of the stimuli were initially done by the last author following the criteria mentioned above. Additionally, the second author annotated the stimuli for pragmatics, and the third author did the same for phonetics. All annotators worked independently but followed the same criteria. Unweighted *Cohen's Kappa* was calculated as a measure of reliability. For the phonetic annotation, the *Kappa* score was calculated considering whether a word was labeled as H*, as L + H*, or not labeled at all; the agreement was very high (0.85, *C.I.* = 0.81–0.89). For the pragmatic annotation, the *Kappa* score was calculated considering whether a word was labeled as contrastive or non-contrastive (only for words annotated as H* or L + H*); the agreement was substantial (0.7, *C.I.* = 0.61–0.79).

Table 1 shows the distribution of the accents across the four categories. Six accents, four H*s and two L + H*s, were removed from the analysis because Roleg reported aggregate responses for words that appeared more than once in a given utterance. Consequently, it was impossible to determine which instance of the word the participant had reacted to. This resulted in the analysis of 281 accents.

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

Table 1. Accent distribution by phonetic and pragmatic classification

		Phonetics		
		H*	L + H*	Total
Pragmatics	Contrastive	50	59	109
	Non-contrastive	143	35	178
Total		193	94	287

2.1.4. Analysis of the stimuli

The phonetic differences between the accent categories were examined by analyzing the F0, duration and amplitude of the accented syllables, to determine whether the phonetics and pragmatics of the accents varied independently, as per our annotation, thereby allowing for the independent investigation of each parameter's contribution to prominence. We note that, due to the relatively small sample of data, these results are presented with caution.

Accented syllable F0 was extracted in Praat (Boersma & Weenink, 2023), using the Python library *parselmouth* (Jadoul et al., 2018) with an octave cost of 0.1; F0 values were taken every 0.005 seconds for female and 0.01 seconds for male speakers, with customized F0 ranges, determined by the F0 minima and maxima across the stimuli of each speaker. We next ran a series of Generalized Additive Mixed Models (GAMMs; Wood, 2011; Wood, 2017) in R (R Core Team, 2020) to test for F0 differences between the phonetic and pragmatic classifications of the accents. We selected the model with the best fit using the function *compareML()* from the R package *itsadug* (van Rij et al., 2022). This model included the parametric factors for Phonetics (H*, L + H*) and Pragmatics (Contrastive, Non-contrastive), smooth terms for Time by both Phonetics and Pragmatics and random smooths for Speaker over Time by both Phonetics and Pragmatics. Syllables categorized as carrying a L + H* showed a larger rise-fall movement and a later peak than H*s; in contrast, the pragmatic categorization into contrastive and non-contrastive accents did not yield substantial differences in F0 shape (see Figure 3). These results match those of Kim et al. (2024) who followed a similar procedure with a much larger corpus.

Accented syllables were annotated in Praat using standard criteria of segmentation (Machač & Skarnitzl, 2009). Their duration was extracted using a Praat script and z-scored by utterance. Durations were then analyzed using a Linear Mixed-effect Model (LMM) with Phonetics, Pragmatics and their interaction as fixed effects, and Speaker as random intercept. The model summary showed that neither the phonetic nor the pragmatic categories were differentiated by means of duration ($\beta_{phonetics: L + H^*} = 0.21, t = 1.32, p = 0.19; \beta_{pragmatics: contrastive} = 0.10, t = 0.69, p = 0.49; \beta_{phonetics: L + H^* \times pragmatics: contrastive} = -0.11, t = -0.48, p = 0.63$).

The Root Mean Square (RMS) amplitude of the accented syllable was extracted in Praat and then normalized by dividing the obtained value by the RMS of the whole utterance. RMS was analyzed in a linear model having Phonetics, Pragmatics and their interaction as dependent variables (random intercepts yielded convergence issues). Both Phonetics and Pragmatics were significant, but not their interaction ($\beta_{phonetics: L + H^*} = 0.34, t = 4.89, p < 0.001; \beta_{pragmatics: contrastive} = 0.13, t = 2.11, p = 0.035; \beta_{phonetics: LH \times pragmatics: contrastive} = -0.17, t = -1.72, p = 0.086$).

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

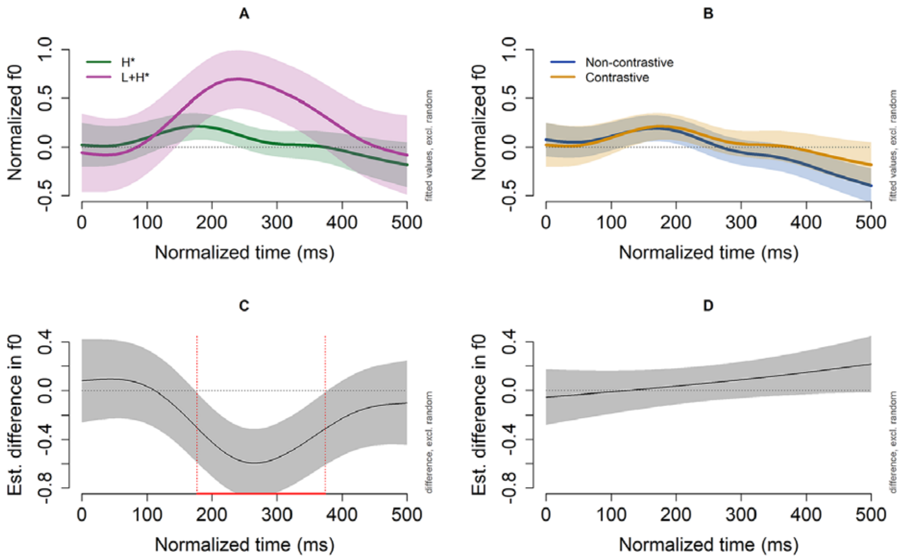

Figure 3. Visualization of GAMMs results: predicted F0 values (a) and estimated difference (c) as a function of phonetics, and pragmatics (b and d, respectively). The shaded areas refer to the 95% confidence interval. The difference is significant if zero is not included in the 95% confidence interval, as marked by the red lines.

Finally, we considered three more stimuli properties that might affect prominence ratings: downstepping, nuclearity and deaccenting of the material following the accent (cf. *inter alia* Turnbull et al., 2017; Im et al., 2023). As can be seen in Table 2, the distributions of these features across categories follow trends typical for English. A portion of the accents were downstepped and this applied mostly to H*s. In addition, there were more prenuclear than nuclear accents but there was no substantial difference between proportions of nuclear and prenuclear accents across categories. Finally, contrastive L + H*s were the accents most frequently followed by deaccenting.

In short, differences in F0 were consistent with H* and L + H* categories. In addition, L + H*s and contrastive accents had higher amplitude than H*s and non-contrastive accents, respectively, while other characteristics of the accents did not follow a consistent pattern that could have affected the outcome of our main study. Finally, as the analysis indicated that the accent form and function were independent of one another, we concluded that the phonetics and pragmatics of the accents were

Table 2. Counts and percentages of accents in the stimuli that were downstepped, nuclear or followed by deaccenting

Phonetics	Pragmatics	Count	Downstep	Nuclear accent	Deaccenting of following material
H*	C	50	16%	26%	10%
	NC	143	37%	41%	1%
L + H*	C	59	5%	31%	37%
	NC	35	6%	26%	9%

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

not conflated by annotation and thus that the role of each on prominence assessment could be independently investigated.

2.1.5. Processing of responses and statistical analysis

Following previous RPT studies (Baumann & Winter, 2018; Cole et al., 2019), we ran Generalized linear Mixed-effect Models (GLMMs) using the R package *lme4* (Bates et al., 2015), with the RPT response (word selected or not as prominent) as dependent variable. The fixed effects included Phonetics (H*, L + H*), Pragmatics (contrastive, non-contrastive) and their interaction. The random effects included random intercepts for Speaker, Listener and Item (accented word), and the by-Listener random slopes for Phonetics and Pragmatics.

Additionally, for each test item, we calculated *prominence scores (p-scores)*, i.e., the percentage of participants who marked that item as prominent (Cole et al., 2010; see also Cole & Shattuck-Hufnagel, 2016). We use *p-scores* primarily for result visualization.

2.2. Results

Both the distribution of *p-scores* (Figure 4) and the GLMM (Table 3) showed that L + H* accented words were rated prominent significantly more often than H* accented words, while contrastive words were rated prominent significantly more often than those with non-contrastive function. There was no interaction between Phonetics and Pragmatics. Contrastive items accented with L + H* were the most prominent, while H* accented non-contrastive items were the least prominent; both non-contrastive L + H*s and contrastive H*s had overlapping distributions of *p-scores* (see Figure 4c).

2.2.1. Individual variability

Individual patterns were inspected using the by-Listener random slopes from the GLMM reported in Table 3 (see Drager & Hay, 2012). The slope value for a specific variable indicates the extent to which a participant differentiated prominence as a function of that variable: the higher the value, the more the participant relied on it. Thus, the slopes can be used as a proxy of each participant's relative reliance on phonetics and pragmatics. Figure 5 illustrates this point by showing the relationship between the slope values for Phonetics and Pragmatics. The cut-off for the groupings was determined using as reference the participant whose difference between the Phonetics and Pragmatics slopes was closest to 0: the participants within the 20^{th} percentile around this reference point were taken to have a relatively balanced approach toward the two cues; participants below and above the 20^{th} percentile were those deemed to have relied mainly on Pragmatics and Phonetics, respectively.

2.3. Interim discussion

Our main goal in this first study was to consider whether in SSBE, H* and L + H* are sufficiently distinct to receive different prominence ratings. Such a result would point to H* and L + H* being distinct categories, rather than forming a continuum,

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

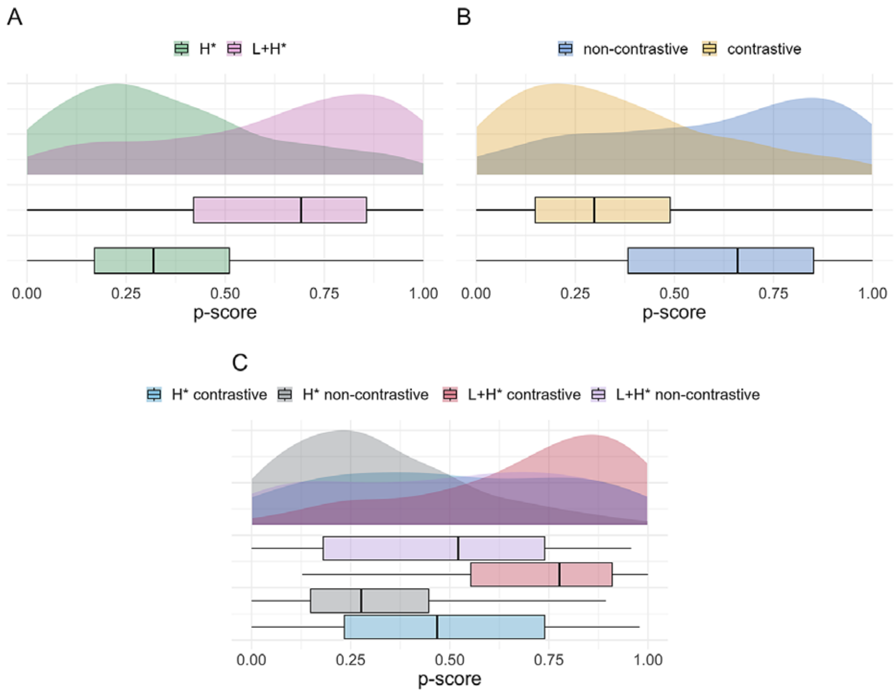

Figure 4. Density and box-whisker plots of *p-scores* as a function of Phonetics (a), Pragmatics (b), and their interaction (c).

Table 3. Summary of the GLMM output for Study 1

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	−1.41	0.30	−4.70	<.001
Phonetics: *L + H**	1.18	0.33	3.58	<.001
Pragmatics: *Contrastive*	1.15	0.29	3.98	<.001
Phonetics: *L + H** × Pragmatics: *Contrastive*	0.17	0.46	0.38	0.707

especially if their prominence separation requires a convergence of phonetic and pragmatic factors. Our results provide evidence that this was so.

The *p-score* distributions for the four subcategories created by crossing phonetics and pragmatics (Figure 4c) suggest that prominence assessment did not depend exclusively on either factor but on a combination of the two. Thus, our results agree with those of Turnbull et al. (2017), Cole et al. (2019), Bishop et al. (2020) and Im et al. (2023) on American English, and Baumann and Winter (2018) on German, and add comparable information about SSBE.

In brief, contrastive L + H*s and non-contrastive H*s were the most and least likely subcategories to be selected as prominent, respectively, with their scores creating a bimodal distribution. Contrastive H*s and non-contrastive L + H*s, on the other hand, spanned the entire distribution of *p-scores*, indicating that their prominence assessment was at chance level and likely influenced by numerous factors beyond accent identity and information structure. As noted in 2.1.4, factors such as RMS

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

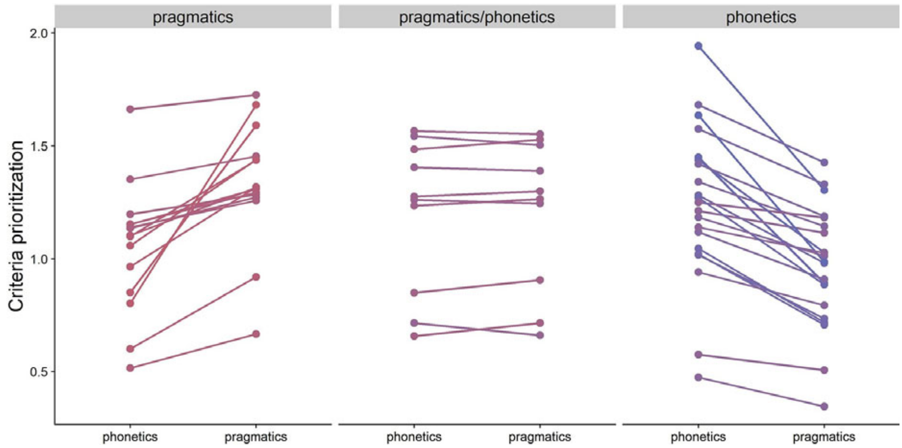

Figure 5. Random slope values for Pragmatics and Phonetics within the responses of individual participants (extracted from the model in Table 3). The panels show individuals grouped according to whether slopes for Pragmatics are higher than (left), about the same as (middle), or lower than those for Phonetics (right).

amplitude, the presence of downstep, and tonal context, including the deaccenting of following words, and the status of an accent as nuclear or prenuclear, do affect prominence ratings and were present in our stimuli (see Table 2). However, it is also reasonable to assume, based on the *p-score* distributions, that their effect was largely limited to the two subcategories in which phonetics and pragmatics did not match.

In conclusion, the bimodal distribution of *p-scores* for contrastive L + H*s and non-contrastive H*s is not compatible with a view that H* and L + H* form a continuum (cf. Ladd & Morton, 1997). If that were the case, we would expect a unimodal and skewed distribution of *p-scores.*

However, the picture is more complex than the aggregate results would suggest. The individual responses showed that listeners varied when weighing pragmatic and phonetic cues to prominence: some prioritized the former, others the latter, while a third group relied on both approximately equally. These results echo Baumann and Winter (2018) who found that some participants relied more on prosody and others on morphosyntactic properties when assessing prominence. Evidence for individual variability during RPT has also been reported by Bishop et al. (2020), who connected these differences to cognitive styles. This is explored in the second study.

3. RPT study 2

Our aim in conducting the second RPT study was twofold. First, we wished to examine whether the individual variation patterns would be replicated and if so, whether the aggregate findings were replicable. The second aim was to test whether musicality, empathy and autistic-like traits could be behind individual variability, as differences in these traits could affect the listeners' sensitivity to pragmatic and phonetic cues. Differences in sensitivity can be related to long-standing disagreements regarding the phonological status and function of H* and L + H*. To this end, we investigated the link between RPT responses and participants' empathy, autistic-like traits and musicality (see 3.1.2): we hypothesized that more empathetic listeners

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

would be more sensitive to and therefore more likely to rely on pragmatics, while listeners with more autistic-like traits or higher musical abilities would be more sensitive to and therefore rely more on phonetics. The study was preregistered on the OSF (Open Science Framework) platform: https://osf.io/enrtj.

3.1. Methods

3.1.1. Participants

Eighty-five participants, recruited through Prolific, took part in the study. We report results for 82 of them (47 female; 19–50 years old, *M*: 33.8, *SD*: 8.6). We excluded two participants who met the exclusion criteria mentioned in section 2.1.1, and one participant whose answers were not registered due to technical issues with the platform. Participants were remunerated for their participation.

3.1.2. Measures of individual characteristics and participant responses

The Empathy Quotient test (EQ, Baron-Cohen & Wheelwright, 2004) was used to assess whether empathy modulates sensitivity to pragmatic information during prominence assessment. EQ is a self-report questionnaire that measures cognitive and emotional empathy. It consists of 40 statements to which participants respond using a forced-choice 4-point Likert scale (*strongly disagree*, *slightly disagree*, *slightly agree*, *strongly agree*). Each statement is scored as 0 (non-empathic responses), 1 (somewhat emphatic responses), and 2 (most empathic responses); their sum gives a score ranging from 0 to 80. As noted in section 1, EQ had been previously used by Esteve-Gibert et al. (2020) and Orrico and D'Imperio (2020) to investigate differences in the processing of intonational meaning.

The Autism Quotient test (AQ, Baron-Cohen et al., 2001) was used to assess whether autistic-like traits modulate sensitivity to phonetic detail during prominence assessment. AQ is not a diagnostic test for autism; rather, it positions neurotypical adults along a continuum measuring five traits associated with the Autism Spectrum Disorder: social skills, communicative skills, attention to detail, attention switching, and imagination. These five traits are measured by different AQ subscales: the questionnaire comprises 10 statements for each subscale (50 statements in total) to which participants respond using a forced-choice 4-point Likert scale (*strongly disagree*, *slightly disagree*, *slightly agree*, *strongly agree*); the score is calculated by assigning 0 (non-autistic-like response) or 1 (autistic-like response); the total score is the sum of all subscale scores and ranges from 0 to 50. The AQ has been used to examine individual variation connected to both phonetics (e.g., Stewart & Ota, 2008; Yu, 2010; Yu et al., 2013) and pragmatics (e.g., Bishop, 2016; Yang et al., 2018). Some studies have relied on AQ subscales; e.g., Yu et al. (2013) analyzed each subscale separately, while Bishop (2016) and Bishop et al. (2020) used only AQ-Communication. Others have used the aggregate score for their main hypotheses, reporting information about the subscales, to variable extent, as a post-hoc analysis (e.g., Stewart & Ota, 2008; Yang et al., 2018; Yu, 2010). Since no one subscale has been consistently linked to phonetics, we followed the latter approach and formulated our hypothesis considering the aggregate score. Finally, we note that AQ and EQ are negatively correlated with one another (Baron-Cohen & Wheelwright, 2004). However, since here we were interested in the effect of AQ on

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

sensitivity to phonetics and of EQ on sensitivity to pragmatics, the two tests do not tap into the same aspect of our study.

The Mini Profile of Music Perception Skills (Mini-PROMS, Law & Zentner, 2012; Zentner & Strauss, 2017) was used to assess whether musicality modulates sensitivity to phonetic detail during prominence assessment. The Mini-PROMS was chosen because it tests musical ability rather than musical training or love for music. It consists of four components: Melody (10 items), Metric Accent (10 items), Tempo (8 items), and Tuning (8 items). Participants listen to pairs of musical fragments and indicate whether they are the same or different using a 5-point Likert scale (*definitely same*, *probably same*, *I do not know*, *probably different*, *definitely different*). Correct answers receive 2 points (*definitely same/different*) or 1 point (*probably same/different*); incorrect and *I do not know* answers are awarded zero points. The score for each component is calculated as the sum of the points divided by 2; the total score ranges from 0 to 36 and is the sum of the component scores. Studies using Mini-PROMS as a predictor of the processing of prosody-related information include Foncubierta et al. (2020).

Participant responses to each of the above tests were normally distributed (see Figure 6) and covered most of the range of each test, from 16 to 70 for EQ ($M = 43.15$, $SD = 12.33$), from 1 to 42 for AQ ($M = 19.18$, $SD = 9.91$), and from 10.5 to 28 for MiniPROMS ($M = 19.66$, $SD = 3.97$). Cronbach's alpha showed high reliability for all three tests: EQ = 0.95 (C.I. = 0.93–0.97); AQ = 0.93 (C.I. = 0.90–0.95); Mini-PROMS = 0.97 (C.I. = 0.95–0.98). Pearson correlations between pairs of tests showed a negative correlation between EQ and AQ ($r(80) = -.58$, $CI = -0.71, -0.41$, $p < .001$), as expected, and no correlation between MiniPROMS and either EQ or AQ.

3.1.3. Stimuli and procedure

We employed the design and stimuli used in the first study; the platform issue concerning stimuli involving the same word had been resolved and thus all 287 accents were analyzed. Within a week of completing the RPT task, the participants also completed the AQ, EQ and MiniPROMS in random order chosen by themselves.

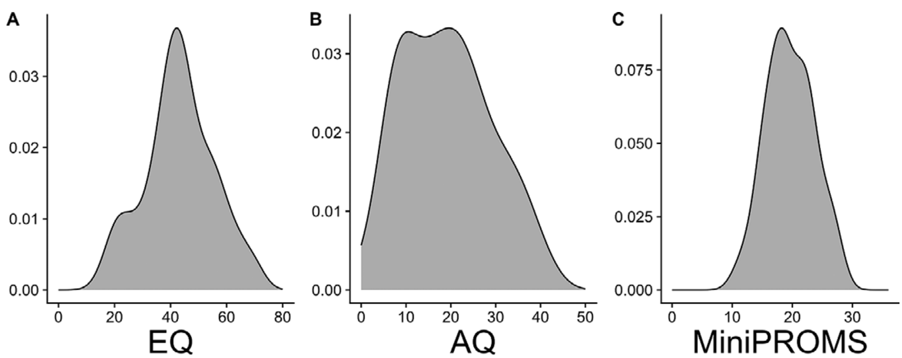

Figure 6. Score distributions for EQ (a), AQ (b) and MiniPROMS (c).

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

3.1.4. Recruitment and power analysis

We used a stop-go method to determine the sample needed to reach statistical power of 80% or higher (see https://osf.io/enrtj). Briefly, we paused the study after the first 25 participants and used their data to run simulations (for details see https://osf.io/f7w9c/). Following Vasishth and Gelman (2021), we calculated power by simulating datasets with increasing number of participants, going from 20 to 120.

Figure 7 shows the results of the simulations for the interactions between EQ and Pragmatics, AQ and Phonetics and MiniPROMS and Phonetics. They indicate that with the present sample size of 82 participants, we reach more than 90% power for the interaction between EQ and Pragmatics, and MiniPROMS and Phonetics. Somewhat lower power was predicted for the interaction between AQ and Phonetics (approximately 70%). We decided to stop the study without reaching the 80% threshold for AQ, as the power analysis indicated a smaller effect size than MiniPROMS and EQ and thus a lesser role of AQ overall.

3.1.5. Statistical analysis

The RPT responses were analyzed in the same way as in the first study (see 2.1.5). In addition, we fitted three more GLMMs to test the effects of EQ, AQ, and MiniPROMS. For EQ, the model included Phonetics, Pragmatics, EQ score, and the interaction between EQ score and Pragmatics as fixed factors. For AQ, the model included Phonetics, Pragmatics, AQ score, and the interaction between AQ score and Phonetics as fixed factors. For MiniPROMS, the model included Phonetics, Pragmatics, MiniPROMS, and the interaction between MiniPROMS and Phonetics as fixed factors. All three models had Item, Subject and Speaker as random intercepts.

3.2. Results

3.2.1. Prominence ratings

As shown in Table 4, the output of the first GLMM replicated the results of the first study (see also Figure 8). The same applied to individual differences (see Figure 9).

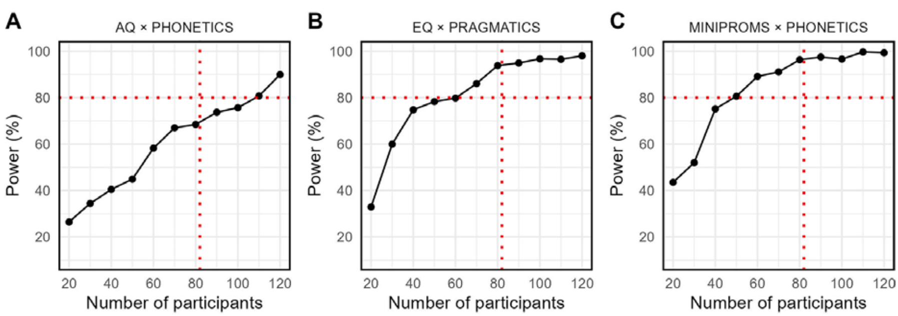

Figure 7. Power analysis output for the interaction AQ × Phonetics (a), EQ × Pragmatics (b) and MiniPROMS × Phonetics (c).

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

3.2.2. The role of individual characteristics on prominence ratings

The interaction between EQ and Pragmatics was significant (see Table 5 for the model summary). As illustrated in Figure 10, contrastive accents were more likely to be selected as prominent by individuals with higher EQ than those with lower scores, while the lower prominence of non-contrastive accents did not change as a function of EQ.

The interaction between AQ and Phonetics was also significant (see Table 6 for the model summary). As illustrated in Figure 11, H*s were more likely to be selected as prominent by participants with higher AQ than those with lower AQ, while the higher prominence of L + H* accents did not change as a function of AQ.

Finally, there was a significant interaction between MiniPROMS and Phonetics (see Table 7 for the model summary). As also shown in Figure 12, L + H*s were more

Table 4. Summary of the GLMM output for Study 2

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	−1.14	0.19	−6.08	<.001
Phonetics: *L + H**	1.33	0.32	4.11	<.001
Pragmatics: *Contrastive*	1.41	0.28	5.02	<.001
Phonetics: *L + H** × Pragmatics*: Contrastive*	−0.34	0.46	−0.74	0.458

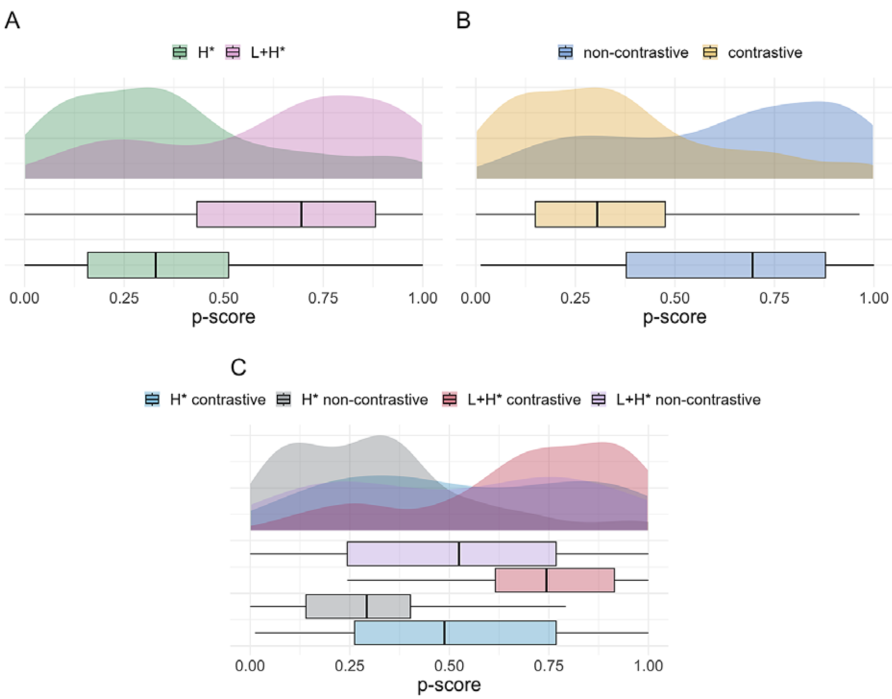

Figure 8. Density and box-whisker plots of *p-scores* as a function of Phonetics (a), Pragmatics (b), and their interaction (c).

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

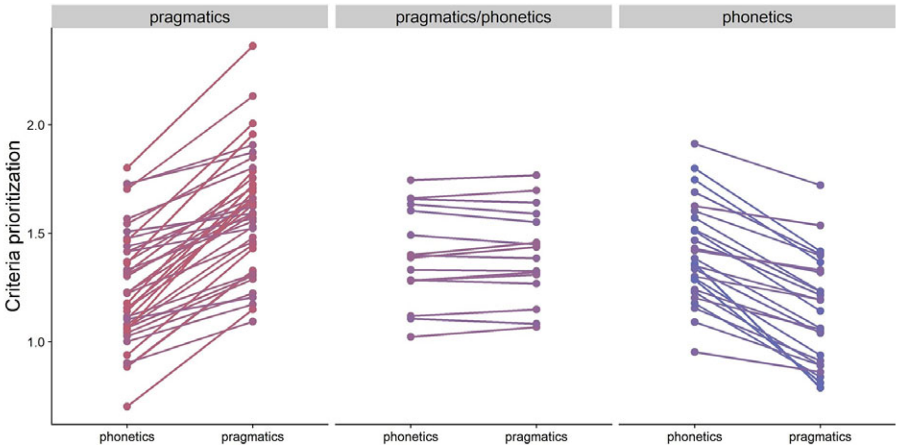

Figure 9. Random slope values for Pragmatics and Phonetics within the responses of individual participants (extracted from the model in Table 4). The panels show individuals grouped according to whether slopes for Pragmatics were higher than (left), similar to (middle), or lower than those for Phonetics (right).

Table 5. Summary of the GLMM testing the effect of EQ on RPT responses

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	−1.21	0.31	−3.79	<.001
Phonetics: *L + H**	1.15	0.22	5.06	<.001
Pragmatics: *Contrastive*	1.04	0.25	3.99	<.001
EQ	.002	.006	0.42	0.67
EQ × Pragmatics: *Contrastive*	.006	.003	2.22	.027

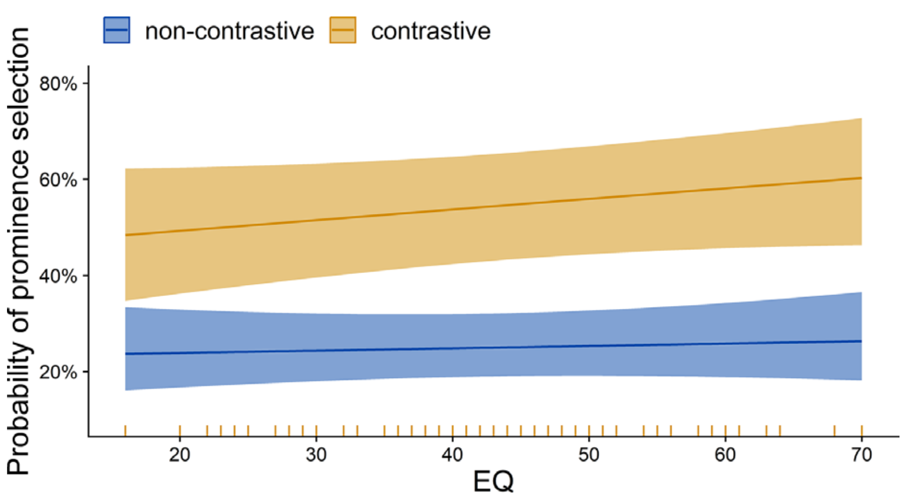

Figure 10. Probability of prominence selection as a function of EQ and Pragmatics. Shaded areas around the regression lines refer to 95% confidence intervals.

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

likely to be selected by participants with higher MiniPROMS scores, while H*s were less likely to be selected as prominent, regardless of MiniPROMS score.

3.2.3. Interim discussion

The second study replicated both the aggregate results and the individual variation patterns of the first. In addition, it shed light on the sources of these individual differences.

The interaction between EQ and Pragmatics showed that more empathetic listeners were more likely to mark contrastive accents as prominent. This trend

Table 6. Summary of the GLMM testing the effect of AQ on RPT responses

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	−1.25	0.23	−5.40	<.001
Phonetics: *L + H**	1.31	0.24	5.48	<.001
Pragmatics: *Contrastive*	1.28	0.22	5.80	<.001
AQ	.008	.008	1.01	.31
AQ × Phonetics: *L + H**	−.008	.004	−2.20	.028

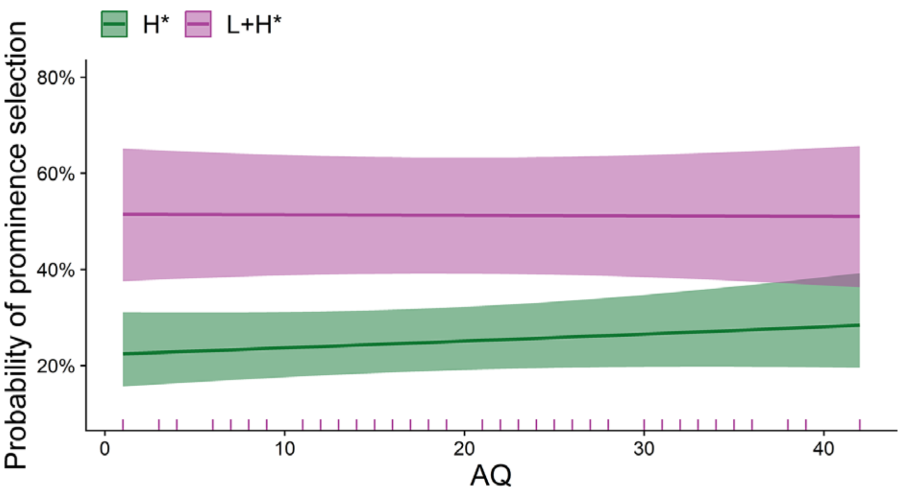

Figure 11. Probability of prominence selection as a function of AQ and Phonetics. Shaded areas around the regression lines refer to 95% confidence intervals.

Table 7. Summary of the GLMM testing the effect of MiniPROMS scores on RPT responses

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	−1.23	0.42	−2.96	.003
Phonetics: *L + H**	1.59	0.29	2.02	.043
Pragmatics: *Contrastive*	1.28	0.22	5.79	<.001
MiniPROMS	.007	.019	0.35	0.72
MiniPROMS × Phonetics: *L + H**	.029	.009	3.13	.002

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

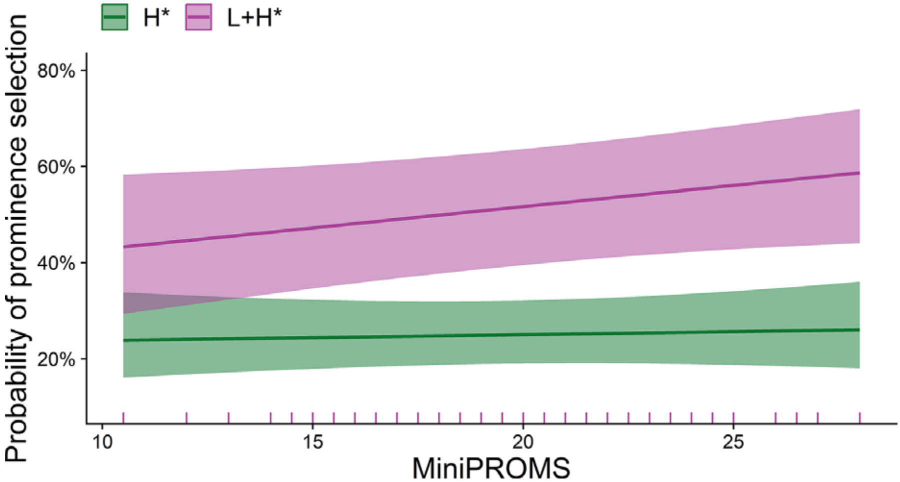

Figure 12. Probability of prominence selection as a function of MiniPROMS scores and Phonetics. Shaded areas around the regression lines refer to 95% confidence intervals.

indicates that these individuals were more sensitive to pragmatic differences, as we had hypothesized. This result agrees with earlier studies on the role of empathy in the processing of pragmatics, whether empathy was directly measured using the EQ (Esteve-Gibert et al., 2020; Orrico & D'Imperio, 2020), or inferred from either the AQ total score, as in Yang et al. (2018), or the AQ Communication subscale, as in Bishop (2016).

Further, our results indicate that both musicality and autistic-like traits reflect sensitivity to phonetic information though, in different ways and to a different extent. The interaction of MiniPROMS scores and Phonetics strongly confirmed our prediction: participants scoring high on musicality were more likely to mark as prominent words accented with L+H* relative to those with low scores, indicating greater sensitivity to the differences between H* and L + H*. Our results agree with those of previous studies that musical abilities play an important role in the processing of phonetic information (Cui & Kuang, 2019; Schön et al., 2004).

Finally, the interaction of AQ scores and Phonetics supported our hypothesis, in that it showed a link between AQ and sensitivity to phonetic detail, though not in the direction we had anticipated. Participants with higher AQ scores were more likely to mark H*-accented words as prominent relative to those with lower scores, suggesting that high AQ individuals were sensitive to small phonetic changes that may not be particularly salient to those with lower AQ. This finding supports previous research showing that the differences in the perception of phonetic information as a function of autistic-like traits may not depend on higher auditory sensitivity, but on a different way of processing higher-order information (Yu & Zellou, 2019).

In brief, our findings confirmed the existence of individual variability in RPT responses detected in the first study and showed that it is related to differences in cognitive styles. These differences mean that participants are more sensitive to either pragmatic or phonetic information. We contend that this sensitivity leads listeners to prioritize different aspects of information in the signal when assessing prominence.

4. General discussion

This paper addressed the long-standing debate concerning the phonological status of H* and L + H* accents in English; we investigated their relative salience, as reflected in *p-scores*, as a means of understanding their relationship in SSBE. We reasoned that if L + H* is an emphatic variant of H* (cf. Ladd & Morton, 1997) it should be rated more prominent than H* regardless of pragmatic function, while the *p-scores* of H*- and L + H*-accented items should form a unimodal distribution.

Our results, however, showed that both phonetics and pragmatics affected the responses. By and large, this result is compatible with other recent findings that showed phonetic properties, phonological status and information structure affect how prominence is assessed (c.f. Im et al., 2023; Turnbull et al., 2017). As a result of these factors, the *p-scores* of our four categories spanned the entire *p-score* range, indicating that some accented words were not selected by anyone and others were selected by all participants – an outcome that highlights the multiple influences on RPT responses.

Critically, the responses did not form a continuum interpretable as the result of additive contributions of phonetics and pragmatics toward increased prominence. Rather, the distribution of the *p-scores* was bimodal, indicating that H* and L + H* were processed as separate entities. In turn, this suggests that H* and L + H* were phonologically interpreted: listeners' expectations were that form and function would be matching, leading to clear differences in *p-scores*, while mismatches created uncertainty. In this respect, the results support the view that the accents are not only phonetically distinct (cf. Kim et al., 2024), but also that each phonetic realization is preferentially connected to a distinct information-structure related function.

However, this conclusion is modulated by variation in individual responses, which showed that some participants prioritized pragmatics over phonetics, others did the opposite, while a few participants weighed both cues (almost) equally. The variable response patterns indicate that the differences between H* and L + H* accents were not equally salient to all participants. Our second study suggested that this was due to individual differences associated with empathy, which modulated how salient pragmatic differences were, and musicality and autistic-like traits, both affecting salience based on accent phonetics.

We contend that these differences among individuals result in their being more or less sensitive to the overall distinction between H* and L + H*. For some individuals, one dimension is less salient than the other, while for others both dimensions carry similarly low or similarly high weight. Consequently, individuals may reach different generalizations regarding this accentual contrast: some acquire it because both the phonetic and pragmatic differences are salient to them; others do not because they fail to integrate the phonetic and pragmatic information into distinct accentual categories. Further, the replication of both the aggregate and individual results indicates a stable variation in the population, a status quo that may also lie behind the long-standing disagreements among linguists about the status of H* and L + H*.

In sum, while the aggregate results favor a contrast between H* and L + H*, the variable interpretation that speakers have of these accents suggests that some SSBE speakers may not acquire the distinction between H* and L + H* as a contrast (cf. Arvaniti & Garding, 2007, on the lack of H* ~ L + H* in Minnesotan English and Kim & Arnhold, 2024, for similar conclusions regarding Canadian English). In segmental phonology, similar cases are not unusual and are referred to as *marginal*

https://doi.org/10.1017/langcog.2024.62 Published online by Cambridge University Press

contrasts (cf. Ladd, 2006; Scobbie, 2007). For example, while the Italian mid-vowels [e] ~ [ɛ] and [o] ~ [ɔ] are involved in minimal pairs in most Italian dialects, native speaker awareness of the two contrasts – i.e., of the mapping between each phonetic category and the lexicon – is low, suggesting they are likely to be ignored during lexical identification (Renwick & Ladd, 2016). Hualde et al. (2017) further argue that individual variability is an indicator of contrast marginality: those authors examined Canadian Raising and found that Chicago speakers distinguish minimal pairs such as *writer* [rʌɪɾɚ] ~ *rider* [raɪɾɚ] in their productions, but show individual variation in perceiving them. The characteristics of marginal segmental contrasts amply relate to the findings reported here about H* ~ L + H*. The link between phonetics and function in differentiating the accents is what characterizes them as contrastive. However, not all speakers are equally sensitive to the differences, therefore the contrast shows signs of marginality, at least in SSBE.

5. Conclusion

We tested the hypothesis that H* and L + H* form a continuum in SSBE and assessed this hypothesis by considering the extent to which the accents' prominence, as reflected in RPT responses, differs based on their phonetic differences and pragmatic functions. Our results clearly showed that L + H* contrastive and H* non-contrastive accents have distinct *p-score* distributions, suggesting that the differences between the two accents are salient to SSBE listeners, so long as phonetics and pragmatics match. In turn, the presence of this distinction, coupled with the indistinguishable *p-scores* of H* contrastive and L + H* non-contrastive accents, supports the existence of prototypical uses of H* and L + H*. This is indicative of contrast. However, the presence of individual differences among our listeners suggests that the contrast is marginal and its presence in a given speaker's grammar depends on individual cognitive traits, here empathy, musicality and autistic-like traits: combinations of these traits allow some individuals to integrate phonetic and pragmatic information to form categories, while others do not. We argue that the presence of these individual differences in how salient the distinction between H* and L + H* appears to individuals may be the key to understanding this long-debated issue: seeing the H* ~ L + H* contrast as marginal can explain the disagreements in the literature regarding these accents as stemming from the speakers' variable use and understanding of the accents.

Data availability statement. The data and scripts are publicly available and can be found at https://osf.io/f7w9c/.

Acknowledgments. We thank Na Hu, Katherine Marcoux, Sofia Sialiaki and Cong Zhang for help with various aspects of the studies reported here, and Chris Cummins and Hannah Rodhe for their input and help with devising the pragmatic annotation scheme. Finally, we thank the editors and two anonymous reviewers for their helpful comments and suggestions.

Funding statement. This research is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement No. ERC-ADG-835263 to Amalia Arvaniti).

Competing interest. The author(s) declare none.

References

Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. In J. Cole & J. Hualde (Eds.), *Laboratory Phonology* 9 (pp. 547–576). De Gruyter Mouton.

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, 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

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, 5–17. https://doi.org/10.1023/A:1005653411471

Bartels, C., & Kingston, J. (1996). Salient pitch cues in the perception of contrastive focus. *University of Massachusetts Occasional Papers in Linguistics*, 22(1), 1–25.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Baumann, S., & Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *Journal of Phonetics*, 70, 20–38. https://doi.org/10.1016/j.wocn.2018.05.004

Bishop, J. (2016). Individual differences in top-down and bottom-up prominence perception. *Proceedings of Speech Prosody* 2016, 668–672. https://doi.org/10.21437/SpeechProsody.2016–137

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, Article 100977. https://doi.org/10.1016/j.wocn.2020.100977

Boersma, P., & Weenink, D. (2023). Praat: Doing phonetics by computer [Computer program]. Version 6.3.13, retrieved 31 July 2023 from http://www.praat.org/

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. In A. Lahiri, P. Avery, B. E. Dresher & K. Rice (Eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 145–172). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110208603.2.145

Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass*, 8(2), 37–50. https://doi.org/10.1111/lnc3.12061

Brugos, A., Shattuck-Hufnagel, S., & Veilleux, N. (2006). Transcribing Prosodic Structure of Spoken Utterances with ToBI. MIT Open Courseware. Retrieved 2 August 2023, from 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.

Burdin, R. S., Holliday, N., & Reed, P. E. (2018). Rising above the standard: Variation in L+H* contour use across 5 varieties of American English. *Proceedings of Speech Prosody* 2018, 354–358. https://doi.org/10.21437/SpeechProsody.2018-72

Calhoun S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language*, 86, 1–42.

Casillas, J. V., Garrido-Pozú, J. J., Parrish, K., Arroyo, L. F., Rodríguez, N., Esposito, R., Chang, I., Gómez, K., Constantin-Dureci, G., Shao, J., Rascón, I., & Taveras, K. (2023). Using intonation to disambiguate meaning: The role of empathy and proficiency in L2 perceptual development. *Applied PsychoLinguistics*, 44(5), 913–940. https://doi.org/10.1017/S0142716423000310

Cason, N., Marmursztejn, M., D'Imperio, M., & Schön, D. (2020). Rhythmic abilities correlate with L2 prosody imitation abilities in typologically different languages. *Language and Speech*, 63(1), 149–165. https://doi.org/10.1177/002383091982633

Cole, J., Hualde, J. I., Smith, C. L., Eager, C., Mahrt, T., & de Souza, R. N. (2019). Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of Phonetics*, 75, 113–147. https://doi.org/10.1016/j.wocn.2019.05.002

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. https://doi.org/10.1515/labphon.2010.022

Cole, J., & Shattuck-Hufnagel, S. (2016). New methods for prosodic transcription: Capturing variability as a source of information, *Laboratory Phonology*, 7(1), 8. https://doi.org/10.5334/labphon.29

Cruttenden, A. (1997). *Intonation*. Cambridge University Press.

Cui, A., & Kuang, J. (2019). The effects of musicality and language background on cue integration in pitch perception. *The Journal of the Acoustical Society of America*, 146(6), 4086–4096. https://doi.org/10.1121/1.5134442

Dilley, L. C., Ladd, D. R., & Schepman, A. (2005). Alignment of L and H in bitonal pitch accents: Testing two hypotheses. *Journal of Phonetics*, 33(1), 115–119. https://doi.org/10.1016/j.wocn.2004.02.003

Drager, K., & Hay, J. (2012). Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, 24(1), 59–78. https://doi.org/10.1017/s0954394512000014

Esteve-Gibert, N., Schafer, A. J., 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, 566–580. https://doi.org/10.3758/s13421-019-00990-w

Foncubierta, J. M., Machancoses, F. H., Buyse, K., & Fonseca-Mora, M. C. (2020). The acoustic dimension of reading: Does musical aptitude affect silent reading fluency?. *Frontiers in Neuroscience*, 14, 513019. https://doi.org/10.3389/fnins.2020.00399

Grabe, E., Post, B., Nolan, F., & Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics*, 28(2), 161–185. https://doi.org/10.1006/jpho.2000.0111

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge, UK: Cambridge University Press.

Gussenhoven, C. (2016). Analysis of intonation: The case of MAE_ToBI, *Laboratory Phonology*, 7(1), 10, 1–35. https://doi.org/10.5334/labphon.30

Hedberg, N., & Sosa, J. M. (2007). The prosody of topic and focus in spontaneous English dialogue. In C. Lee, M. Gordon, & D. Büring (Eds.), *Studies in Linguistics and Philosophy*, Vol. 82 (pp. 101–120). Springer. https://doi.org/10.1007/978-1-4020-4796-1_6

Holliday, N. (2021a). Perception in black and white: Effects of intonational variables and filtering conditions on sociolinguistic judgments with implications for ASR. *Frontiers in Artificial Intelligence*, 4, 642783. https://doi.org/10.3389/frai.2021.642783

Holliday, N. (2021b). Intonation and referee design phenomena in the narrative speech of black/biracial men. *Journal of English Linguistics*, 49(3), 283–304. https://doi.org/10.1177/00754242211024722

Hualde, J. I., Luchkina, T., & Eager, C. D. (2017). Canadian Raising in Chicagoland: The production and perception of a marginal contrast. *Journal of Phonetics*, 65, 15–44. https://doi.org/10.1016/j.wocn.2017.06.001

Im, S., Cole, J., & Baumann, S. (2023). Standing out in context: Prominence in the production and perception of public speech. *Laboratory Phonology*, 14(1), 1–62. https://doi.org/10.16995/labphon.6417

Iskarous, K., Cole, J., & Steffman, J. (2024). A minimal dynamical model of Intonation: Tone contrast, alignment, and scaling of American English pitch accents as emergent properties. *Journal of Phonetics*, 104, 101309. https://doi.org/10.1016/j.wocn.2024.101309

Iskarous, K., Steffman, J., & Cole, J. (2023). American English pitch accent dynamics: A minimal model configurations. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 1469–1473).

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58(2), 541–573. https://doi.org/10.1016/j.jml.2007.06.013

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, 71, 1–15. https://doi.org/10.1016/j.wocn.2018.07.001

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, 22(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kim, J. & Arnhold, A., (2024). Phonetic and phonological aspects of prosodic focus marking in Canadian English, *Laboratory Phonology* 15(1), 1–38. https://doi.org/10.16995/labphon.9316

Kim, J., Hu, N., Gryllia, S., Orrico, R., & Arvaniti, A. (2024). Delineating H* and L+H* in Southern British English. *Proceedings of Speech Prosody* 2024, 1185–1189. https://doi.org/10.21437/SpeechProsody.2024-239

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3–4), 243–276.

Kurumada, C., Brown, M., & Tanenhaus, M. (2012). Pragmatic interpretation of contrastive prosody: It looks like speech adaptation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 34, 647–652. https://escholarship.org/uc/item/6jw49594

Kurumada, C., & Roettger, T. B. (2022). Thinking probabilistically in the study of intonational speech prosody. *Wiley Interdisciplinary Reviews: Cognitive Science*, 13(1), e1579. https://doi.org/10.1002/wcs.1579

Ladd, D. R. (2006). "Distinctive phones" in surface representation. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory Phonology*, 8 (pp. 3–26), De Gruyter Mouton. https://doi.org/10.1515/9783110197211.1.3

Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511808814

Ladd, D. R. (2022). The trouble with ToBI. In J. Barnes & S. Shattuck-Hufnagel (Eds.), *Prosodic Theory and Practice*. MIT Press. https://doi.org/10.7551/mitpress/10413.003.0009

Ladd, D. R., & Morton, R. (1997). The perception of intonational emphasis: Continuous or categorical?. *Journal of Phonetics*, 25(3), 313–342. https://doi.org/10.1006/jpho.1997.0046

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, 31(1), 81–112. https://doi.org/10.1016/S0095-4470(02)00073-6

Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *PLoS ONE*, 7, Article e52508. https://doi.org/10.1371/journal.pone.0052508

Machač, P., & Skarnitzl, R. (2009). *Principles of Phonetic Segmentation*. Epocha Publishing House. https://doi.org/10.1159/000331902

Metusalem, R., Ito, K. (2008). The role of L+H* pitch accent in discourse construction. *Proceedings of Speech Prosody* 2008, 493–496. https://doi.org/10.21437/SpeechProsody.2008-110

Molnár, V. (2002). Contrast – From a contrastive perspective. In H. Hasselgård, S. Johansson, B. Behrens, & C. Fabricius-Hansen, (Eds.), *Information Structure in a Crosslinguistic Perspective* (Language and Computers: Studies in Practical Linguistics 39), (pp. 147–161). Amsterdam/New York: Rodopi.

O'Connor, J. D., & Arnold, G. F. (1973). *Intonation of Colloquial English* (2nd ed.). Longman.

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. https://doi.org/10.5334/labphon.238

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* [Unpublished doctoral dissertation]. Massachusetts Institute of Technology.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in Communication* (pp. 271–311). https://doi.org/10.7916/D8KD24FP

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

Renwick, M. E., & Ladd, D. R. (2016). Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts. *Laboratory Phonology*, 7(1):19. https://doi.org/10.5334/labphon.17

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. https://doi.org/10.1111/1469-8986.00172.x

Scobbie, J. M. (2007). Interface and overlap in phonetics and phonology. In G. Ramchand & C. Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces* (pp. 17–52). https://doi.org/10.1093/oxfordhb/9780199247455.013.0002

Silverman, K. E. A., & Pierrehumbert, J. B. (1990). The timing of prenuclear high accents in English. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech* (pp. 72–106). Cambridge: Cambridge University Press.

Steffman, J., Shattuck-Hufnagel, S., & Cole, J. (2022). The rise and fall of American English pitch accents: Evidence from an imitation study of rising nuclear tunes. *Proceedings of Speech Prosody 2022*, 857–861. https://doi.org/10.21437/SpeechProsody.2022-174

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition*, 109(1), 157–162. https://doi.org/10.1016/j.cognition.2008.07.010

Syrdal, A. K., & McGory, J. (2000). Inter-transcriber reliability of ToBI prosodic labeling. *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP 2000)*, (3), 235–238. https://doi.org/10.21437/ICSLP.2000-521

Turnbull, R., Royer, A. J., Ito, K., & Speer, S. R. (2017). Prominence perception is dependent on phonology, semantics, and awareness of discourse. *Language, Cognition and Neuroscience*, 32(8), 1017–1033. https://doi.org/10.1080/23273798.2017.1279341 van Rij, J., Wieling, M., Baayen, R., & van Rijn, H. (2022). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.4.1.

Vasishth, S., & Gelman, A. (2021). How to embrace variation and accept uncertainty in linguistic and psycholinguistic data analysis. *Linguistics*, 59(5), 1311–1342. https://doi.org/10.1515/ling-2019-0051

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+ H*. *Cognitive Science*, 32(7), 1232–1244. https://doi.org/10.1080/03640210802138755

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

Yang, X., Minai, U., & Fiorentino, R. (2018). Context-sensitivity and individual differences in the derivation of scalar implicature. *Frontiers in Psychology*, 9, 1720. https://doi.org/10.3389/fpsyg.2018.01720

Yu, A. C. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PLoS ONE*, 5(8), e11950. https://doi.org/10.1371/journal.pone.0011950

Yu, A. C., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and "autistic" traits. *PLoS ONE*, 8(9), e74746. https://doi.org/10.1371/journal.pone.0074746

Yu, A. C., & Zellou, G. (2019). Individual differences in language processing: Phonology. *Annual Review of Linguistics*, 5, 131–150. https://doi.org/10.1146/annurev-linguistics-011516-033815

Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*, 1400, 33–45. https://doi.org/10.1111/nyas.13410

Cite this article: Orrico, R., Gryllia, S., Kim, J., & Arvaniti, A. (2025). Individual variability and the H* ~ L + H* contrast in English, *Language and Cognition*, 17, e9, 1–25. https://doi.org/10.1017/langcog.2024.62