

# Individual variability in the use of tonal and non-tonal cues in intonation<sup>a)</sup>

Na Hu and Amalia Arvaniti<sup>b)</sup> 

Centre for Language Studies, Radboud University, Postbus 9103, 6500 HD Nijmegen, The Netherlands

[na.hu@ru.nl](mailto:na.hu@ru.nl), [amalia.arvaniti@ru.nl](mailto:amalia.arvaniti@ru.nl)

**Abstract:** Greek uses  $H^*$ ,  $L + H^*$ , and  $H^* + L$ , all followed by L-L% edge tones, as nuclear pitch accents in statements. A previous analysis demonstrated that these accents are distinguished by F0 scaling and contour shape. This study expands the earlier investigation by exploring additional cues, namely, voice quality, amplitude, and duration, in distinguishing the pitch accents, and investigating individual variability in the selection of both F0 and non-F0 cues. Bayesian multivariate analysis and hierarchical clustering demonstrate that the accents are distinguished not only by F0 but also by additional cues at the group level, with individual variability in cue selection. © 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

[Editor: Irina A. Shport]

<https://doi.org/10.1121/10.0028613>**Received:** 12 April 2024 **Accepted:** 22 August 2024 **Published Online:** 11 September 2024

## 1. Introduction

Research on intonation has largely focused on the study of F0, the prime intonation exponent, with influential models, such as the Autosegmental-Metrical theory of intonational phonology (Pierrehumbert, 1980; Ladd, 2008), representing intonation categories, e.g., pitch accents, by their F0 characteristics. Consequently, the phonetics of intonation is often studied with the expectation that category differences are reflected in F0, with little consideration of additional correlates. Furthermore, research on intonation has largely focused on aggregate results that often reinforce a sense of uniformity across study participants (Arvaniti, 2016).

Recent studies, however, indicate that intonation categories are realized using various correlates, including voice quality, and the duration and amplitude of segments [e.g., Breen *et al.* (2010) on American English, Arvaniti *et al.* (2016) on Polish, and Roessig *et al.* (2022) on German]. As there is also evidence that these correlates are perceptually relevant (Shang *et al.*, 2024), in the remainder of the paper we refer to them as *cues*. Arvaniti *et al.* (2024) show that such cues not only provide much-needed redundancy but may also enhance intonational contrasts or be in a cue-trading relationship with F0: for example, in their study of Greek pitch accents, the F0 of unaccented syllables was high before the high scaled  $H^* + L$  but low before  $H^*$ , while vowels accented with  $L + H^*$  showed increased duration when the accent was produced with a lower F0 peak or a shallower than average dip.

Additionally, it is becoming increasingly clear that there is substantial inter-speaker variability in phonetic realization, particularly in the value ranges used by different speakers. This has been demonstrated with segmental cues, such as VOT duration [e.g., Chodroff and Wilson (2017)], and prosodic cues, such as the degree of articulatory strengthening at phrasal onset (Fougeron and Keating, 1997) and the articulatory kinematics of phrase finality (Byrd *et al.*, 2006). Relatively fewer studies have examined the number of cues used by individual speakers for intonation, and many are small in scope [e.g., Dahan and Bernard (1996) on French emphatic accents and Cangemi *et al.* (2015) on German pitch accents] or focus only on F0 differences [e.g., Niebuhr *et al.* (2011)].

Here, we extend this line of research by exploring individual variability in the selection of both tonal and non-tonal cues in intonation, using three pitch accents— $H^*$ ,  $L + H^*$ , and  $H^* + L$ —found in nuclear position in Greek declaratives where they are followed by L-L% edge tones; they are illustrated in Fig. 1. Arvaniti *et al.* (2024) show that these accents differ in F0 scaling ( $H^* < L + H^* < H^* + L$ ) and overall contour shape:  $L + H^*$  is a rise that starts with a marked F0 dip, while  $H^*$  and  $H^* + L$  are a low gradual fall and a high steep fall, respectively. Here, we expand that investigation by examining group-level patterns in the use of three non-F0 cues, duration, voice quality, and amplitude, and individual variability in tonal and non-tonal cue selection.

<sup>a)</sup>A preliminary analysis was presented at *Acoustics 2023 Sydney*, 4–8 December 2023, Sydney, Australia.

<sup>b)</sup>Author to whom correspondence should be addressed.

## 2. Methods

### 2.1 Participants and materials

The dataset was that of [Arvaniti et al. \(2024\)](#). Here, we present the gist of the methodology; for details, see the OSF repository.

The dataset comprises recordings from 13 native speakers of Standard Greek (10 females, mean age 34) without self-reported speech or hearing disorders. Recordings were conducted in quiet environments using a DAT recorder at a sampling rate of 44.1 kHz.

The participants read question-answer pairs. Following the analysis of [Arvaniti and Baltazani \(2005\)](#), each question was formulated to evoke one of three pragmatic situations, thereby eliciting one of the target accents in the response [cf. [Roessig et al. \(2022\)](#)]. All responses were declarative statements ending in a fall; see [Table 1](#) for examples. H\* was elicited using questions prompting broad focus statements, indicating the accented item is new in discourse. H\* + L was elicited using questions seeking information the addressee considers obvious or predictable. L + H\* was elicited using questions prompting narrow focus, as L + H\* indicates corrective or contrastive information.

The responses consisted of one or two content words ([Table 1](#)). One-word responses carried one of the target accents; two-word responses had a prenuclear L\* + H on the first content word and the target accent on the last. The words carrying the target accents were stressed on the antepenult, penult, or ultima. The dialogues were interspersed with fillers, and read four times across four blocks (one per repetition). This yielded 936 tokens (13 participants \* 3 accents \* 3 stress locations \* 2 contexts [presence/absence of prenuclear L\* + H] \* 4 repetitions); 92 tokens were discarded due to excessive background noise, disfluencies, and extensive stretches of creak. The naturalness and pragmatic appropriateness of the remaining tokens were assessed auditorily by the second author, a native speaker of Greek. All were deemed adequate exemplars of the intended tunes and thus suitable for further analysis. The analyzed dataset comprised 844 tokens (272 H\*s, 274 H\*+Ls, and 298 L+H\*s).

### 2.2 Measurements

The F0-related measurements were derived using Functional Principal Component Analysis (FPCA) ([Ramsay and Silverman, 2005](#)). FPCA was conducted by [Arvaniti et al. \(2024\)](#) and the output was used here.

FPCA mathematically represents each input curve using [Eq. \(1\)](#), where  $f(t)$  represents the modelled F0 curve, approximated by adding principal component curves (PC curves:  $PC1(t)$ ,  $PC2(t)$ , etc.) with different scores ( $s1$ ,  $s2$ , etc.) to the mean curve of all input curves ( $\mu(t)$ ),

$$f(t) \approx \mu(t) + s1 \times PC1(t) + s2 \times PC2(t) + \dots \tag{1}$$

The PC curves represent dominant modes of variation among the input curves such as variations in F0 shape and scaling. The score of a PC curve denotes the extent to which the curve contributes to approximating the input curve; it is unique for each input curve and characterizes that curve's shape. As the PC scores are numerical values, using them as the dependent variable with the effect of interest (e.g., accent type) as the predictor in a statistical model provides an understanding of that effect's impact on curve shape ([Gubian et al., 2015](#)).

In [Arvaniti et al. \(2024\)](#), FPCA was conducted following [Gubian et al. \(2015\)](#); this involved curve smoothing, landmark registration with the accented vowel onset as landmark, and functional PCA (for details, see the OSF repository). This analysis was done on F0 curves spanning a three-syllable window consisting of the accented syllable of each target word and the two unstressed syllables preceding it, as this window captured both F0 changes on unaccented syllables under the influence of the upcoming accent (H\*, L + H\*, or H\* + L) and varying degrees of tonal crowding as the accented syllable approached the utterance end.

The FPCA output is shown in [Fig. 2\(a\)](#). Following [Arvaniti et al. \(2024\)](#), we focused on PC1 and PC2 which capture most of the accent-related curve variance in the dataset. PC1 primarily reflects curve scaling, with higher scores (red lines) resulting in higher scaling and lower scores (blue lines) resulting in lower scaling. PC2 reflects contour shape,

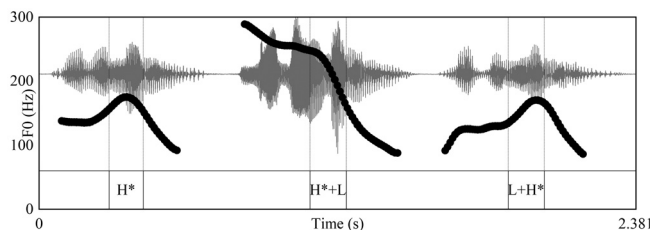


Fig. 1. Illustration of the three accents on the test word [laðo'lemono] “oil-lemon [sauce],” as produced by study speaker M13. Vertical lines mark the stressed syllable.

Table 1. Glosses of example dialogues (with explanatory notes in square brackets), phonetic transcriptions of the responses, and autosegmental representations of their tunes.

	Prenuclear L*+H Absent	Prenuclear L*+H Present
H*	<p>Q. <i>What's this?</i></p> <p>A. <i>Oil-lemon [sauce].</i></p> <p>[[laðo'lemono]ip]IP</p> <p style="text-align: center;">            </p> <p style="text-align: center;"><b>H*</b>    L-L%</p>	<p>Q. <i>What will you have?</i></p> <p>A. <i>Seabream in oil-lemon [sauce].</i></p> <p>[[sina'yriða laðo'lemono]ip]IP</p> <p style="text-align: center;">                       </p> <p style="text-align: center;">L*+H    <b>H*</b>    L-L%</p>
H*+L	<p>Q. <i>What should I make with so many lemons?</i></p> <p>A. <i>Lemonade.</i></p> <p>[[lemo'naða]ip]IP</p> <p style="text-align: center;">            </p> <p style="text-align: center;"><b>H*+L</b>   L-L%</p>	<p>Q. <i>What should I do with all these lemons?</i></p> <p>A. <i>Make lemonade.</i></p> <p>[[kane lemo'naða]ip]IP</p> <p style="text-align: center;">                       </p> <p style="text-align: center;">L*+H    <b>H*+L</b>   L-L%</p>
L+H*	<p>Q. <i>Did you say their son's eyes are brown?</i></p> <p>A. <i>Blue! [Aren't you listening?]</i></p> <p>[[yala'na]ip]IP</p> <p style="text-align: center;">            </p> <p style="text-align: center;"><b>L+H*</b> L-L%</p>	<p>Q. <i>Are their daughter's eyes brown?</i></p> <p>A. <i>They're blue! [Don't you remember her?]</i></p> <p>[[ine yala'na]ip]IP</p> <p style="text-align: center;">                       </p> <p style="text-align: center;">L*+H <b>L+H*</b> L-L%</p>

with higher scores leading to a rise-fall shape with a high, late peak, and lower scores resulting in a plateau with a low, early peak. The scores of these two PCs were included into the current analysis of individual variability.

As mentioned, the non-F0 cues included duration, amplitude, and voice quality. These were extracted from accented vowels, not the three-syllable window used for F0, as the effect of the accent's identity on preceding unaccented syllables is unknown for Greek. Duration and RMS amplitude were automatically extracted using PRAAT (Boersma and Weenink, 2023). For voice quality, we used the accented vowel's mean H1-H2 (Keating et al., 2011), corrected for formants and bandwidth using the Iseli et al. (2007) algorithm. This was calculated by averaging the measures across 12 time-intervals automatically obtained using PRAATSAUCE (Kirby, 2018).

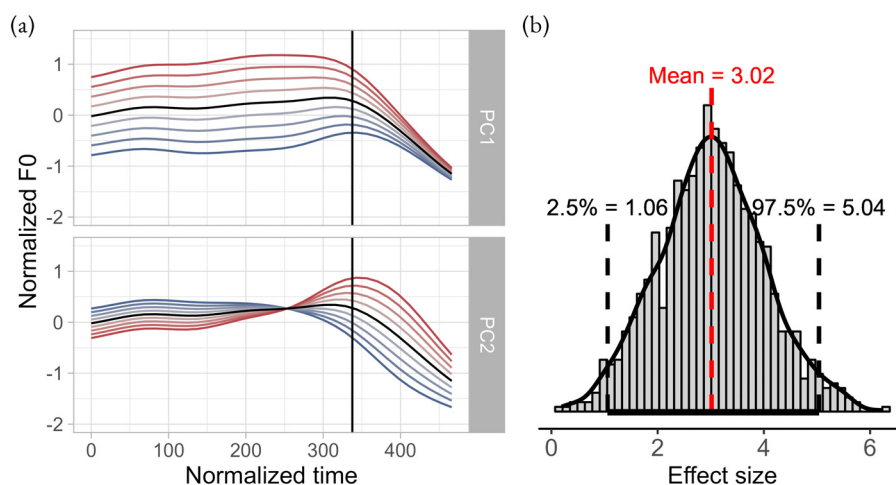


Fig. 2. (a) Color-coded curves illustrate the effect of PC1 and PC2 on the mean curve (solid black line); the vertical line indicates the onset of the accented vowel; for details see text (plot reproduced with permission from Arvaniti et al. (2024)). (b) Probability distribution based on simulated data; the red dashed line marks the distribution mean; the horizontal solid black bar indicates the 95% Credible Interval with the lower and upper bounds marked by the dashed black lines.

### 2.3 Statistical analysis

The aim of our statistical modelling was to estimate the effect of accent type ( $H^*$ ,  $L + H^*$ ,  $H^* + L$ ) on PC scores, duration, amplitude and H1-H2. To avoid confirmation bias, data categorization by accent type (henceforth *Accent*) was based on the *dialogue* in which each target word was elicited. Thus, for instance, all retained accents elicited in answers to *what's this?* were classed as  $H^*$ s, even when they were not prototypical of the  $H^*$  category; this is illustrated in Fig. 1, where the  $H^*$  is more dipped than typically expected for this accent.

The effect of *Accent* was estimated by fitting Bayesian multivariate mixed-effect models in R (R Core Team, 2020) using the *BRMS* package (Bürkner, 2018), the wrapper package for the probabilistic programming language *STAN* (Carpenter *et al.*, 2017). For the code, data, and a brief explanation of the difference between Bayesian and frequentist approaches to statistical inference visit the OSF repository.

The dependent variables were the accented vowel duration, mean H1-H2, and RMS amplitude, and the PC1 and PC2 scores of the F0 curves as estimated in Arvaniti *et al.* (2024), all normalized using z-scores across all speakers, a procedure that facilitates between-speaker comparisons [cf. Lorenzen *et al.* (2023)]. The constant effect was *Accent*, with three levels (HL, LH, and H) and H as the reference level. The random effects were *participant* and *item*, with *item* capturing differences in stress location (antepenult, penult, and ultima) and the absence or presence of an  $L^* + H$  prenuclear accent. Models were built top-down, starting with the model including *Accent* and by-participant and by-item random intercepts and by-participant and by-item random slopes for *Accent*. Each model component was evaluated by comparing models with and without it using Bayes factors ( $BF_{10}$ ) (Lee and Wagenmakers, 2014). A component was excluded if there was no evidence supporting it (when  $BF_{10} < 1$ ) (Jeffreys, 1998). The models were fitted using 4 chains, 10 000 iterations each, including 4000 warm-ups, and uninformative priors, with normal distributions with a mean of 0 and a standard deviation of 5.

To explore individual variability in cue selection, we conducted a hierarchical cluster analysis on the effect of *Accent* on individual speakers' acoustic measures, following Lorenzen *et al.* (2023). Unlike those authors, who used point estimates, namely the estimated means of posterior distributions [red dashed line in Fig. 2(b)], we based our analysis on the 95% CrIs of posterior distributions [horizontal solid black bar in Fig. 2(b)]. This method provides a more accurate description of the effect, as it encompasses 95% of the most likely values for the model coefficient given the data (Westfall and Henning, 2013). A 95% CrI excluding 0 suggests a 95% probability that the effect of interest is present [see Fig. 2(b)]. Conversely, a 95% CrI including 0 suggests that the effect of interest is likely absent.

To prepare the data for hierarchical clustering we extracted posterior samples for the grand mean and the group-specific deviations. We then derived the posterior distribution of speaker-specific slopes for *Accent* on each acoustic measurement by adding the group-specific deviations to the grand mean, and calculated the 95% CrIs of the posterior distribution using the `median_qi()` function from the *tidybayes* package. Finally, the 95% CrIs was coded as follows:

- 1 if located to the right of zero, suggesting a positive *Accent* effect (e.g., longer duration);
- 0 if containing zero, indicating that the *Accent* affect was absent;
- -1 if located to the left of zero, indicating a negative *Accent* effect.

The resulting codes were analyzed using the `scipy.cluster.hierarchy` module in *PYTHON*, which produced a dendrogram, representing the hierarchical organization of clusters in the data. To present all possible combinations of the acoustic cues in expressing the three pitch accents, we used the number of the smallest branches in the dendrogram as the cluster count, where speakers under the same branch use the same cues.

## 3. Results

### 3.1 Group level effects

The final model (derived from Bayes Factor analysis) contained the constant effect of *Accent* on all five measurements, with varying slopes for *Accent* by participant but not by item. This indicates that the effect of *Accent* varied across participants, but remained consistent across items.

Figure 3(a) depicts the 95% CrIs of the posterior probability distributions of the slope parameters  $\beta$  in the final model. These parameters indicate the group-level effects of  $H^* + L$  and  $L + H^*$  on each measurement, representing for each measurement the average difference between each of those accents and  $H^*$ .  $L + H^*$  had higher PC1 and PC2 scores compared to  $H^*$  (PC1: Estimate = 0.64, 95% CrI [0.28, 0.99]; PC2: Estimate = 1.41, 95% CrI [0.90, 1.92]). Based on Fig. 2(a), this result suggests that  $L + H^*$  had higher F0 scaling and a more scooped shape than  $H^*$  [as in Arvaniti *et al.* (2024)].  $L + H^*$  also had higher amplitude (estimate = 0.77, 95% CrI [0.36, 1.19]) and longer duration (estimate = 0.80, 95% CrI [0.36, 1.23]) than  $H^*$ , but did not differ from it in H1-H2 (estimate = 0.38, 95% CrI [-0.01, 0.77]).  $H^* + L$  had a higher PC1 score than  $H^*$  (estimate = 1.45, 95% CrI [1.13, 1.77]), suggesting it was scaled higher [as in Arvaniti *et al.* (2024)], but did not differ from  $H^*$  in terms of PC2 (estimate = 0.41, 95% CrI [-0.01, 0.82]), i.e., in shape.  $H^* + L$  was also produced with higher H1-H2 (estimate = 0.36, 95% CrI [0.15, 0.58]) and higher amplitude (estimate = 0.46, 95% CrI [0.08, 0.83]) than  $H^*$ , but did not differ from it in duration (estimate = 0.34, 95% CrI [-0.05, 0.74]).

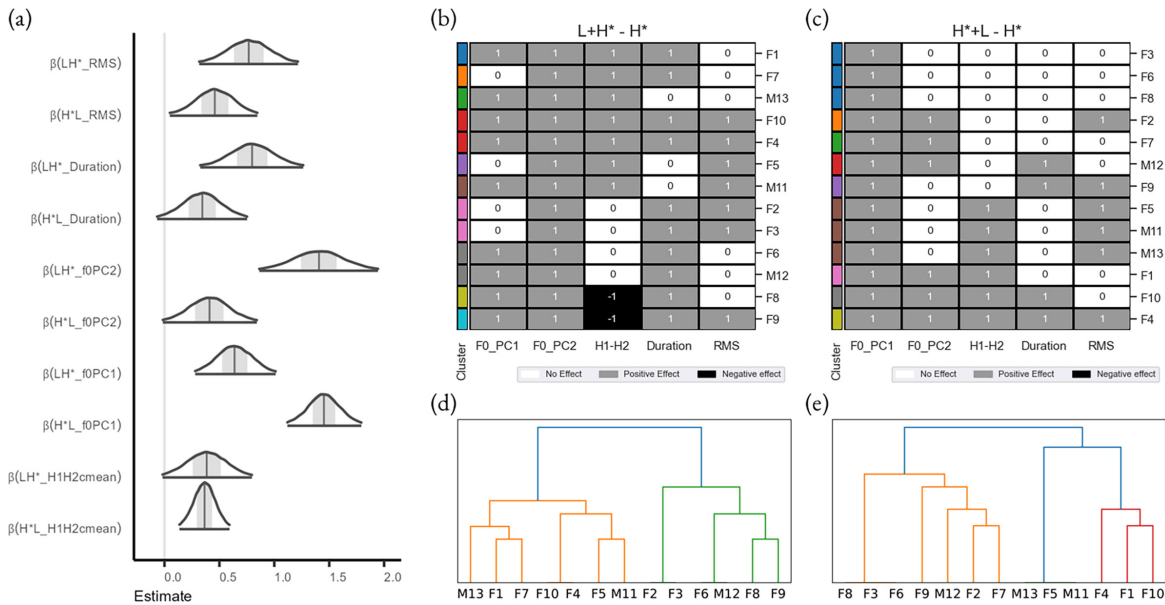


Fig. 3. (a) Posterior probability distributions of the slope parameters  $\beta$ ; (b) and (c) use of the five traits (columns) by individual speakers (rows), grouped by cluster (colored tabs), in realizing L + H\* (b) and H\* + L (c) in comparison to H\*; (d) and (e) dendrograms generated by hierarchical clustering (with speaker IDs on the x axis) for L + H\* (d) and H\* + L (e) in comparison to H\*.

### 3.2 Individual variability

Figures 3(b)–3(e) show the coding of the 95% CrI for the effect of L + H\* and H\* + L (relative to H\*) on each measurement for each speaker. For L + H\*, Fig. 3(b) shows that higher PC2 scores, indicating a more scooped curve, were present in the data of all 13 speakers; longer duration was the second most frequent feature, present in the data of 10 speakers, while higher F0 scaling (higher PC1 scores) and greater amplitude were less frequently used (9 and 7 speakers, respectively); H1-H2 was not consistently used. Figure 3(c) shows that higher PC1 scores, reflecting higher scaling, were used by all speakers for H\* + L relative to H\*, while PC2, H1-H2 and amplitude were used only by 6 speakers each, and duration was rarely used (by just 4 speakers). The variable use of cues gives rise to a large number of clusters [colored tabs in Fig. 3(b) and 3(c)]: 10 clusters for L + H\* vs H\* and 9 for H\* + L vs H\* [see also Fig. 3(d) and 3(e)]. However, as discussed, some cues dominate.

## 4. Discussion and conclusion

We investigated the role of tonal and non-tonal cues in phonetically encoding intonation categories, and the extent to which individuals vary in their use of such cues, by examining three pitch accents in Greek, H\*, L + H\*, and H\* + L.

At the group-level, Bayesian mixed-effect linear models confirmed the finding of Arvaniti et al. (2024) that H\*+L has a similar shape to H\* but with higher F0 scaling, while L + H\* is primarily distinguished from H\* by its scooped shape. The accents also differed in other dimensions: compared to H\*, L + H\* was accompanied by longer duration and greater amplitude of the accented vowel, while H\* + L exhibited greater amplitude and breathier voice. For L + H\*, these results confirm the importance of duration, which Arvaniti et al. (2024) show to be in a cue-trading relationship with F0. The finding that, at the group level, non-tonal cues consistently contribute to the realization of pitch accents provides further evidence that F0 is not the sole exponent of intonation [cf. Roessig et al. (2022)] or its only perceptual cue [cf. Shang et al. (2024)]. Given such results, the investigation of non-tonal cues should be given due attention. We note that the use of these additional cues was robust despite our accent classification approach which allowed for variability in production.

At the individual level, hierarchical clustering showed that speakers differed in the number of F0 and non-F0 cues they used for each accent, though only three speakers relied exclusively on F0, and only for H\* + L. Such individual variability has been previously demonstrated: Cangemi et al. (2015) showed that duration was used only by some speakers to differentiate focus types (hence, pitch accents) in German, while Cangemi and Grice (2016) and Niebuhr et al. (2011) found individual differences in the tonal cues differentiating intonation categories in Italian and German. In contrast, our participants were consistent in their use of F0, in that they all used the same dominant cue per accent: scaling for H\* + L, shape for L + H\*. Such consistency can be a heuristic for the essential differences between accents: for example, since all our speakers relied on F0 shape to differentiate L + H\* from H\*, but not all used scaling to do so, scaling is less likely to



be a defining characteristic of  $L + H^*$  and thus, we would argue, need not be included in this accent's phonological representation.

Additionally, some speakers, e.g., F4 and F10, used more cues than others, such as F3 and F6. At present, it is unclear why some speakers use multiple cues, while others do not. Previous research suggests systematicity in such cue variability (Chodroff and Wilson, 2017), but our results do not provide supporting evidence: e.g., speakers F3, F6, and F8 used only F0 scaling to differentiate  $H^* + L$  from  $H^*$ , but multiple cues to differentiate  $L + H^*$  from  $H^*$ . It is possible that clearer patterns will emerge with larger speaker groups, and the investigation of individual traits (such as musicality and autistic-like traits) which may explain the observed differences in cue selection in production, as has already been shown for perception; for instance, Orrico *et al.* (2023) found that autistic-like traits affect the extent to which participants attend to phonetic detail in intonation processing. If production and perception are linked [Beddor (2009) and Harrington *et al.* (2008), among others] we could expect that individual variability in production correlates with speakers' cognitive characteristics and their sensitivity to the same cues in perception. In turn, this could suggest that the distinction between tonal categories is more pronounced for speakers who use multiple cues. However, evidence for this hypothesis is limited [e.g., Shultz *et al.* (2012)], indicating that further research is needed to better understand the link between production and perception, as well as the influence of cognitive characteristics on individual variability. Additionally, insights could be gained by incorporating into the clustering process *the extent* to which individual speakers use specific cues (something we have not addressed here).

In conclusion, the present study showed that F0 cues, despite being the most consistently used, are not the only means of encoding intonation categories. Non-F0 parameters, here duration, amplitude, and voice quality, also contribute. This finding adds to a growing understanding that F0 is not the sole exponent of intonation. Finally, our results reveal individual variability in cue selection. Using multiple cues can highlight the essential differences between categories, but also prompts questions about the potential sources of individual variability in production and its influence on perception.

#### Acknowledgments

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 A.A.).

#### Author Declarations

##### *Conflict of Interest*

The authors have no conflicts of interest to declare.

##### *Ethics Approval*

This research has been approved by the Research Ethics Advisory Group, School of Humanities, University of Kent, UK (Approval No. 0701819). This approval was accepted by the Ethics Assessment Committee Humanities (EACH), Faculty of Arts, Radboud University, The Netherlands.

#### Data Availability

The data and scripts are available at the OSF repository: <https://osf.io/rm46h/>.

#### References

- Arvaniti, A. (2016). "Analytical decisions in intonation research and the role of representations: Lessons from Romani," *Lab. Phonol.* 7(1), 6.
- Arvaniti, A., and Baltazani, M. (2005). "Intonational analysis and prosodic annotation of Greek spoken corpora," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, edited by Sun-Ah Jun (Oxford University Press, Oxford, UK), pp. 84–117.
- Arvaniti, A., Katsika, A., and Hu, N. (2024). "Variability, overlap, and cue trading in intonation," *Language* 100(2), 265–307.
- Arvaniti, A., Zygis, M., and Jaskuła, M. (2016). "The phonetics and phonology of the Polish calling melodies," *Phonetica* 73(3–4), 338–361.
- Beddor, P. S. (2009). "A coarticulatory path to sound change," *Language* 85(4), 785–821.
- Boersma, P., and Weenink, D. (2023). "Praat: Doing phonetics by computer (version 6.3.10) [computer program]," <https://www.fon.hum.uva.nl/praat/> (Last viewed 20 May 2023).
- Breen, M., Fedorenko, E., Wagner, M., and Gibson, E. (2010). "Acoustic correlates of information structure," *Lang. Cognit. Processes* 25(7), 1044–1098.
- Bürkner, P. C. (2018). "Advanced Bayesian multilevel modeling with the R package brms," *R J.* 10(1), 395–411.
- Byrd, D., Krivokapić, J., and Lee, S. (2006). "How far, how long: On the temporal scope of prosodic boundary effects," *J. Acoust. Soc. Am.* 120(3), 1589–1599.
- Cangemi, F., and Grice, M. (2016). "The importance of a distributional approach to categoricity in autosegmental-metrical accounts of intonation," *Lab. Phonol.* 7(1), 9.
- Cangemi, F., Krüger, M., and Grice, M. (2015). "Listener-specific perception of speaker-specific productions in intonation," in *Individual Differences in Speech Production and Perception*, edited by S. Fuchs, D. Pape, C. Petrone, and P. Perrier (Peter Lang, Lausanne, Switzerland), pp. 123–145.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. (2017). "Stan: A probabilistic programming language," *J. Stat. Soft.* 76(1), 1–32.

- Chodroff, E., and Wilson, C. (2017). "Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English," *J. Phon.* **61**, 30–47.
- Dahan, D., and Bernard, J. M. (1996). "Interspeaker variability in emphatic accent production in French," *Lang. Speech* **39**(4), 341–374.
- Fougeron, C., and Keating, P. A. (1997). "Articulatory strengthening at edges of prosodic domains," *J. Acoust. Soc. Am.* **101**(6), 3728–3740.
- Gubian, M., Torreira, F., and Boves, L. (2015). "Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts," *J. Phonetics* **49**, 16–40.
- Harrington, J., Kleber, F., and Reubold, U. (2008). "Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study," *J. Acoust. Soc. Am.* **123**(5), 2825–2835.
- Iseli, M., Shue, Y.-L., and Alwan, A. (2007). "Age, sex, and vowel dependencies of acoustic measures related to the voice source," *J. Acoust. Soc. Am.* **121**(4), 2283–2295.
- Jeffreys, H. (1998). *The Theory of Probability*, 3rd ed. (Oxford University Press, Oxford, UK).
- Keating, P., Esposito, C., Garellek, M., Khan, D., and Kuang, J. (2011). "Phonation contrasts across languages," in *Proceedings of the 17th International Congress of Phonetic Sciences*, pp. 1046–1049.
- Kirby, J. (2018). Praatsauce: Praat-based tools for spectral analysis, <https://github.com/kirbyj/praatsauce> (Last viewed 4 September 2024).
- Ladd, D. R. (2008). *Intonational Phonology* (Cambridge University Press, Cambridge, UK).
- Lee, M. D., and Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course* (Cambridge University Press, Cambridge, UK).
- Lorenzen, J., Roessig, S., and Baumann, S. (2023). "Redundancy and individual variability in the prosodic marking of information status in German," in *Proceedings of the 20th International Congress of Phonetic Sciences*, edited by R. Skarnitzl and J. Volín, pp. 1320–1324.
- Niebuhr, O., D'Imperio, M., Gili Fivela, B., and Cangemi, F. (2011). "Are there 'shapers' and 'aligners'? Individual differences in signalling pitch accent category," in *Proceedings of the 17th International Congress of Phonetic Sciences*, pp. 120–123.
- Orrico, R., Gryllia, S., Kim, J., and Arvaniti, A. (2023). "The influence of empathy and autistic-like traits in prominence perception," in *Proceedings of the 20th International Congress of Phonetic Sciences*, edited by R. Skarnitzl and J. Volín, pp. 1280–1284.
- Pierrehumbert, J. (1980). "The phonology and phonetics of English intonation," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Ramsay, J. O., and Silverman, B. W. (2005). *Functional Data Analysis*, 2nd ed. (Springer, New York).
- R Core Team (2020). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria), <https://www.r-project.org/> (Last viewed 4 September 2024).
- Roessig, S., Winter, B., and Mücke, D. (2022). "Tracing the phonetic space of prosodic focus marking," *Front. Artif. Intell.* **5**, 842546.
- Shang, P., Roseano, P., and Elvira-García, W. (2024). "Dynamic multi-cue weighting in the perception of Spanish intonation: Differences between tonal and non-tonal language listeners," *J. Phonetics* **102**, 101294.
- Shultz, A. A., Francis, A. L., and Llanos, F. (2012). "Differential cue weighting in perception and production of consonant voicing," *J. Acoust. Soc. Am.* **132**(2), EL95–EL101.
- Westfall, P. H., and Henning, K. S. S. (2013). *Understanding Advanced Statistical Methods* (CRC Press, Boca Raton, FL).