Beyond F0: Voice-Quality Bundles as Prominence Cues in German

Christiane Ulbrich, Jörn Krantz, Sophie Repp, Heiko Seeliger

University of Cologne

`culbrich@uni-koeln.de`, `jkrantz@uni-koeln.de`, `sophie.repp@uni-koeln.de`,
`h.seelige@uni-koeln.de`

Abstract

This study investigates whether high and low pitch accents vs. post-nuclear targets exploit different Voice Quality (VQ) parameters to enhance prominence, addressing a gap in the understanding of the multidimensional nature of prosodic prominence. We measured harmonic-noise-ratio (HNR), cepstral peak prominence smoothed (CPPs), jitter, shimmer and spectral tilt in nuclear accents vs. post-nuclear targets in German rising and falling *wh*-questions to determine if pitch height relates to specific VQ bundles or if these VQ parameters independently mark linguistically relevant prosodic categories. We found that spectral parameters (CPPs, spectral tilt) and, to a lesser extent, HNR differ significantly between high and low nuclear pitch accents, whereas only HNR showed significant differences for high and low post nuclear targets. These results suggest that specific VQ bundles associate with different intonational elements in German and contribute differentially to the signaling of prosodic prominence.

Index Terms: voice quality, pitch height, pitch accent, post-nuclear tonal target, prominence

1. Introduction

Previous research on prominence marking in intonation languages has focused on fundamental frequency, duration and overall energy as prominence-lending cues, with pitch accents, phrase accents and boundary tones being linked to pitch modulation [1], [2]. Regarding VQ, some research has found differences between pitch-accented and unaccented syllables, suggesting that VQ cues prominence [3], but other research reports the absence of effects [4]. This study addresses the interaction of a range of VQ measures and pitch height in nuclear pitch accents – which are typically associated with prosodic prominence – vs. post-nuclear pitch targets in falling and rising *wh*-questions in German, thus investigating the multidimensionality of prosodic prominence as reflected in the bundling of different phonetic characteristics of an utterance.

1.1. Phonetic realization of pitch accents in German

German intonation exhibits contrasts between several high and low pitch accents, differing in tonal scaling and temporal alignment. Féry [5] distinguishes medial/late peaks vs. early-peak accents (also [6]). The key difference, relevant for the present study, lies in the alignment of the F0 peak relative to the stressed syllable. Early peaks, realized with relatively high F0 before the falling or low F0 in the stressed syllable, are associated with low prosodic prominence, signaling (degrees of) discourse givenness. Medial peaks, with high F0 in the vowel of the stressed syllable resulting in a high tonal are associated with high prosodic prominence, signaling semantic focus. Many subsequent studies show that focus elicits higher F0 peaks and greater pitch excursions, whereas given post-focal elements are downstepped or compressed, indicating that German encodes information structure both categorically and gradiently [1].

The turning-point-based description just reviewed does not fully capture how the different accent types differ acoustically. The Tonal-Center-of-Gravity (TCoG) approach [7, 8] also quantifies this: early-peak accents show an earlier temporal TCoG compared to medial/late-peak accents, reflecting their lower prominence. Contrastive focus further increases prominence through longer durations and larger pitch excursions [9, 10], shifting the TCoG toward later and higher regions, indexing increased prosodic prominence.

Recent work highlights the influence of speech act type on prosodic prominence: prosodic cues to focus and givenness behave differently in non-assertive acts like exclamations or (biased) questions, compared to statements [10], [11], [12]. For instance, while German statements typically allow deaccentuation of given constituents, this process is reduced in non-assertive speech acts. Still, contrast and givenness can modulate pitch-accent height and alignment, revealing a dynamic interaction between prosodic prominence and illocutionary force. Furthermore, the “classic” association of high and late peak accents with increased prominence and low or early peaks with reduced prominence or givenness in statements (see above), has been shown to not straightforwardly apply to rising question contours, where low nuclear accents may function as prominence-bearing categories marking focus [10], [12]. The occurrence of low pitch targets marking focus has been interpredted as a prominent diversion from a rising baseline [11], [13].

In addition to pitch accents, theoretical models often posit *phrase accents* (e.g., H- or L-), which occur between the last pitch accent in an intonation phrase and the boundary tone. The phonological status of phrase accents is contested: some authors defend them as distinct edge tones associated with phrase structure [14], while others argue that they are superfluous or even epiphenomenal, with their effects explained by boundary-tone timing or implementation [15]. The association of phrase accents to prominence is largely unexplored: studies investigating boundary phenomena do not examine phrase accent and boundary tone independently of each other. Recent evidence suggests that intonational rises—including boundary rises—can enhance listeners’ attention and memory for entire prosodic domains (e.g., serial-recall, attention-orienting tasks; [16]). This raises the possibility that such rises may support the mental discourse prominence of the content of whole phrases. In our study, we investigate nuclear accents and phrase accents, which we pre-theoretically call *post-nuclear targets* (remaining agnostic regarding their phonological status).

1.2. Voice Quality in high vs. low pitch accents

Spectral characteristics are similar to F0 or duration in that they change under accentuation. For several languages, it has been observed that accented and stressed syllables show higher energy in upper frequency bands and a less steep spectral tilt than their unstressed or unaccented counterparts, reflecting increased articulatory effort and more peripheral vowel articulation (e.g., Dutch, Italian, Spanish, German: [17] [18] [19] [20], [21], [22]). This suggests that VQ is tightly linked to prosodic prominence. According to the *Voice Prominence Hypothesis* [23], the prominence of a syllable indeed depends on several voice-source parameters (also [24]).

However, findings for specific measures of VQ are inconsistent. For German, [4] found no difference in H1*–H2* (the corrected difference in amplitude between the first and second harmonics) between accented and unaccented syllables, while [3] report H1*–H2* as well as HNR (harmonic-to-noise-ratio) to be higher in accented syllables in German.

Perception studies indicate that a less steep spectral tilt is associated with higher perceived prominence in both German [25] and Dutch [18]. [26] confirmed spectral tilt effects in accentuation in English, but [27] and [28] could not empirically validate this. Furthermore, recent studies for Dutch show that the reliability of spectral tilt varies with vowel quality [29]. Vowel-dependent effects have also been observed in other languages. For instance in North American English, low vowels are more likely to be perceived as creaky [30], suggesting that VQ cues interact with the segmental makeup of the utterance. These patterns can be traced to underlying articulatory mechanisms: creaky voice and glottalization tend to co-occur with larynx raising—shortening the vocal tract and raising formant frequencies—hereas breathy phonation is often associated with larynx lowering [31]. Taken together, this cross-linguistic evidence indicates that vowel quality can substantially modulate the acoustic and perceptual expression of VQ, complicating comparisons across VQ measures.

In sum, the role of VQ measures is less well understood than that of F0, duration, and overall energy. Acoustic measurements for VQ are less standardized and less robustly defined [28], [32], although certain measures like HNR or H1*–H2* are being more routinely investigated now.

Regarding the interaction of VQ and F0, recent research in speech production and perception suggests that there is a systematic but non-linear relationship, whose shape has been described as being that of a "wedge-shaped" function [32], [33], [34]: In mid-F0 range, the voice is most periodic; CPPs and HNR are at their highest, indicating low noise and high regularity; H1*–H2* is typically mid-to-high. At the F0 extremes, the voice becomes non-modal, leading to a drop in periodicity and changes in glottal tension. H1*–H2* tends to be low at both high-F0 and low-F0, HNR and CPPs are low at low-F0, due to high irregularity and aperiodicity. They are highest in the modal range but can remain relatively high in high F0 if the voice is highly periodic but tense. Thus, H1*–H2* reflects glottal tension, while CPPS/HNR primarily reflect the degree of aperiodicity/noise, which is most prominent at the low F0 extreme.

From a biomechanical perspective, lower F0 combined with creaky voice reflects increased glottal constriction and reduced subglottal pressure, whereas breathy voice at higher F0 points to more abducted vocal folds and increased airflow; these physiological settings inherently bias the vocal system toward particular pitch–VQ combinations. As a result, what may appear acoustically as a pitch effect may in fact arise from underlying laryngeal biomechanics that co-determine both pitch range and VQ. Thus, the interaction of pitch and VQ is not merely additive but mutually constraining: the VQ used by a speaker can push the perceived pitch upward or downward, while simultaneously limiting which pitch targets are biomechanically achievable.

Importantly, [32] shows that the interaction of pitch and VQ is not merely biomechanical: speakers appear to coordinate pitch and laryngeal configuration to enhance the perceptual salience of prosodic contrasts. This suggests that listeners may integrate F0 and H1*–H2* jointly when evaluating prominence or accentuation. Furthermore, complex interactions of VQ and F0 have been reported in tone languages [33]. Also, perceived prominence depends, on the one hand, on the interaction of F0, duration, and variables reflecting articulatory effort, and on the other hand, on the interaction between prosodic categories (e.g., accentuation) and continuous acoustic variables (e.g., F0 maximum, duration, spectral energy distribution) [25], [35], which supports multidimensional models of prosodic prominence in which pitch and VQ cues are tightly coupled. Understanding the interactions of multiple prominence-lending sources is critical for developing comprehensive models of prosodic systems and for accounting for cross-linguistic variation in phonation contrasts. Crucially, this requires making explicit that VQ and pitch are neither fully independent nor fully dependent dimensions: they share a biomechanical substrate that creates systematic co-variation, yet each can also vary on its own and contribute distinct perceptual cues.

2. Study

To investigate whether particular VQ bundles are directly associated with pitch or rather correlate with a particular prosodic category, we investigated a range of VQ measures in high- and low-tone nuclear pitch accents in comparison to high- and low-tone post-nuclear targets in production data from three previous experiments. The experiments were originally designed to study prosodic reflexes of speech act type (exclamation, question) and information structure (focus, givenness) in German [11], [36]. From the productions in the experiments, we selected verb-second and embedded verb-final *wh*-questions, which had a falling or a rising final contour. In the rising contours, the nuclear pitch accents had a low target (L*), that is focus was marked by a low accent. In the falling contours, the nuclear pitch accents had a high target (H* (sometimes preceded by low leading tone)).

2.1. Participants and materials

We selected productions by 49 native speakers of German (24 female), who lived in Berlin at the time of recording and spoke Standard German. All *wh*-questions were transitive sentences with a direct object. Sentences were either verb-final (Exp. 2, 3), with the object followed by a participle and auxiliary, or verb-second (Exp. 1). Figure 1 illustrates pitch tracks for two verb-final *wh*-questions.

From the original dataset of 587 questions 97 data points were excluded due to diphthongization of otherwise stable vowels (e.g. in [iˈtaːli̯ən]) or due to a lack of a reliable F0 track.

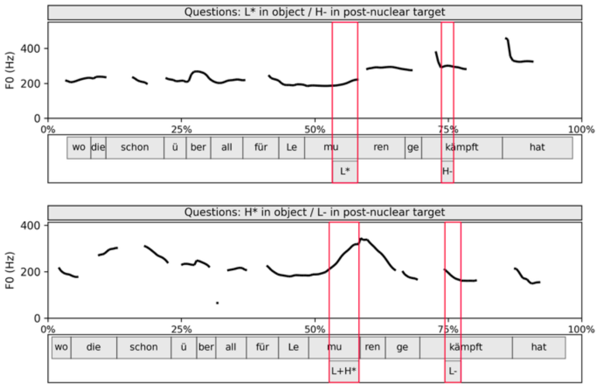

Figure 1: *F0 contour showing low (top) and high (bottom) nuclear pitch accent and high and low post-nuclear target.*

The final dataset of 490 *wh*-questions contained 189 questions with a high nuclear pitch accent on the stressed syllable of the object noun and 301 questions with a low nuclear pitch accent. For the analysis of the post-nuclear target—the stressed but unaccented ultimate syllable of the lexical verb participle—we only considered verb-final questions (Exp. 2 and 3; $n = 341$) to keep the measurement area constant in relation to the end of the clause. Table 1 shows the vowel quality distribution.

Table 1: *Number of nuclear pitch accents and post-nuclear targets per vowel quality.*

Tone on	**Vowel Type**				
stressed syllable of nuclear pitch accent	[aː]	[eː]	[oː]	[uː]	[yː]
H*	61	16	54	25	33
L*	108	48	70	31	44
stressed syllable in post-nuclear target	[ɛ]	[ɔ]	[yː]		
H-	73	10	19		
L-	159	51	29		

2.2. Data analysis

All questions were annotated for syllable boundaries, nuclear pitch accent (type) (GToBI), as well as the vowel boundaries in the pitch accented word and the post-nuclear target. We analyzed HNR, jitter, shimmer, CPPs, H1*–H2* and also H1*–A3*. Extraction via Parselmouth [37] used Praat [38] defaults for harmonicity (cc), jitter, shimmer, and CPPs. H1*–H2* and H1*–A3* were calculated from LTAS and corrected for formant influence following the Iseli-method as described in [3]. For F0-dependent measures, we set a pitch floor/ceiling of 100/500 Hz (female) and 60/300 Hz (male). To ensure robustness, we verified that diverging HNR/CPPs trends remained consistent across different windowing settings..

For the statistical analysis, we employed mixed effects regression models and ran separate models for NUCLEAR PITCH ACCENT (H*, L*; data set Exp.1-3) and POST-NUCLEAR TARGET (H-, L-; data set Exp. 2-3). Because of the relevance of vowel quality for VQ, we included VOWEL as an experimental factor. SPEAKER, ITEM, and SEX were included as random factors. The models were fitted via the `lme4` package [39] in R 4.5.1 (R Core Team, 2025). To ensure best model fit, the models were fitted top-down using ANOVAs. In all cases, a simpler model including only random intercepts had the best model fit.

2.3. Results

Harmonics-to-Noise Ratio (HNR): Statistical analysis of HNR revealed significant differences for NUCLEAR PITCH ACCENT ($b = –1.88$, $t = –2.38$, $p < .05$), with H* showing higher harmonicity than L*. In POST-NUCLEAR position, HNR was the only parameter showing a significant effect ($b = –6.30$, $t = –3.21$, $p < .01$): low targets showed greater harmonicity than high targets.
Jitter and Shimmer: There were no significant effects.
Cepstral Peak Prominence smoothed (CPPs): CPPs differed significantly for NUCLEAR PITCH ACCENT ($b = 1.48$, $t = 2.84$, $p < .01$): L* showed higher values than H*. No effect for POST-NUCLEAR TARGETS was observed.
H1*–H2*: The first spectral tilt measure showed an effect for NUCLEAR PITCH ACCENT ($b = –2.49$, $t = –3.47$, $p < .001$): H* accents had higher H1*–H2* values than L* accents. No significant differences emerged for POST-NUCLEAR TARGETS.
H1*–A3*: The second spectral tilt measure revealed an inverse pattern to H1*–H2*: L* showed higher H1*–A3* values than H* ($b = 3.88$, $t = 2.62$, $p < .01$). Again, POST-NUCLEAR TARGETS showed no significant effect.

3. Discussion

The present study investigated VQ parameters in German high and low nuclear pitch accents and post-nuclear targets to assess how VQ contributes to prominence marking depending on pitch height in rising and falling questions in German. In this study, prosodic prominence is understood as the functional salience of a linguistic unit within its discourse context, which is formally manifested through specific tonal configurations and, further enhanced by coordinated VQ modulations. When prominence is absent, as in post-nuclear targets, these VQ patterns diminish or disappear. Thus, VQ parameters seem to support prominence marking, but their bundling varies systematically with the tonal shape of the nuclear pitch accent. There are systematic differences across multiple VQ parameters: High nuclear pitch accents show higher HNR, higher H1*–H2*, and lower CPPs and H1*–A3* compared to low nuclear pitch accents, while jitter and shimmer remain unaffected. This configuration indicates that high nuclear accents are characterized by greater harmonicity, steeper spectral tilt at lower frequencies (higher H1*–H2*), and reduced overall cepstral balance (lower CPPs). Low pitch accents display the inverse pattern: lower harmonicity, flatter low-frequency spectral tilt, and greater cepstral balance. This differentiation was largely absent in post-nuclear targets, where only HNR shows an effect – albeit with low targets exhibiting higher values than high targets, reversing the pattern for nuclear pitch accents.

The differentiation of VQ in nuclear pitch accents but not in post-nuclear targets suggests that speakers deploy coordinated adjustments of multiple (supra-)laryngeal and cues to enhance prominence distinctions. Post-nuclear material, which is inherently less prominent and often compressed or deaccented, apparently does not require or exhibit comparable VQ modulation. This pattern aligns with the *Voice Prominence Hypothesis* [23], which posits that prominence manifests through coordinated voice-source parameters. Indeed, our findings indicate that VQ varies independently of pitch height *per se*, and that the VQ differences cannot be reduced physiological consequences of pitch manipulation. Recall that previous research has documented systematic covariation between pitch and VQ, where higher F0 correlates with breathier phonation and lower F0 with creakier voice [34], [40]. Our results show a more complex picture: High pitch accents come with higher harmonicity and steeper low-frequency spectral tilt (higher H1*–H2*), while also showing lower overall cepstral balance (lower CPPs). This configuration does not straightforwardly align with a breathy-high/creaky-low dichotomy. Instead, German speakers seem to actively manipulate laryngeal tension, glottal configuration, and possibly supralaryngeal resonance independently of pitch scaling to create distinct VQ profiles for low vs. high pitch accents.

The observed VQ differences between pitch accents reflect systematic variations in laryngeal dynamics, consistent with the F0-VQ covariation described by [32] and the "wedge-shaped function" shape [33]. Crucially, the diverging patterns observed in our data—specifically higher HNR for H* and higher CPPs for L*—are interpreted as a physiological result of laryngeal dynamics rather than an artifact of measurement.While HNR reflects overall periodicity and is enhanced by the increased laryngeal tension and vocal effort typical for high pitch targets in prominence marking, CPPs is highly sensitive to vocal stability and regularity.The higher CPPs values for L* accents align with a more stable, modal phonation often found in the mid-to-low f0-range, where laryngeal tension is lower compared to f0-extremes. This configuration indicates that high nuclear accents are characterized by greater harmonicity and steeper spectral tilt at lower frequencies (higher H1*-H2*), but reduced overall cepstral balance (lower CPPs). Conversely, low pitch accents display lower harmonicity but greater stability and a flatter low-frequency spectral tilt [41].

The interpretation of the spectral tilt measure H1*−H2* further highlights distinct glottal configurations that adhere to F0-dependent strategies. The lower H1*−H2* in low pitch accents suggests greater glottal constriction and a steeper spectral slope. This finding is consistent with the observation that low F0 targets often invoke phonatory adjustments toward the creaky/tense quality (low H1*−H2*) to acoustically mark the pitch boundary [33]. In contrast, the higher H1*−H2*, values observed in high pitch accents signify reduced glottal constriction. While high F0 can lead to tense voice (low H1*−H2*), our result suggests that for these specific accents, the mechanism favors a less constricted setting (higher H1*−H2*), despite the high F0.

Finally, the inverse pattern found in H1*−A3*, where low pitch accents display higher values, implies that these accents distribute spectral energy differently, maintaining relatively more energy in the mid-frequency formant regions. This energy distribution may be the outcome of adjustments in laryngeal tension that alter the filtering properties of the vocal tract or reflect specific articulatory strategies (e.g., tongue positioning) related to the realization of the low F0 and its associated acoustic filtering properties.

4. Theoretical evaluation and conclusion

Integrating our results with existing work on German intonation has important implications for models of prosodic prominence – especially when going beyond the "classic" analysis of prosodic contours in falling statements. Recall that in statements, (contrastive and non-contrastive) focus is marked by high pitch accents, whereas low accents are more frequently associated with givenness or reduced prominence. In rising questions, however, low pitch accents are themselves prominence-bearing and may mark focus. Our research shows that VQ bundles accompany this differential prominence realization in rising vs. falling contours.

Our finding converges with research on multidimensional prominence, which shows that prominence arises from coordinated adjustments across several acoustic dimensions, [27]. The VQ bundles observed here – harmonicity, spectral tilt, and cepstral balance – form structured configurations rather than independent correlates. Such configurations may enhance the perceptual distinctiveness of accent categories, consistent with cue-integration accounts in prominence perception [42]. Whether listeners actively use these cues for accent identification or prominence judgments is an issue that must be investigated in future research.

Another empirical desideratum is the role of VQ in prominence marking across different speech acts. We tested information-seeking *wh*-questions here. Other types of questions, for instance rhetorical questions, have been shown to differ in VQ in comparison to information-seeking questions [43]. Future research will have to clarify whether VQ bundles that are associated with prominence are stable across pragmatic contexts or whether they show an interaction with the illocutionarily induced VQ modulation by speakers.

A central theoretical question concerns whether VQ variation reflects phonological control or automatic phonetic implementation. The distinction aligns with models separating abstract prosodic categories from their gradient phonetic realization [44]. While the systematicity of our patterns suggests a degree of phonological specification, VQ modulation may still result from biomechanical interactions between laryngeal configuration and F0 control *in the presence of prominence*. This latter view begs of course the question of which phonetic characteristics are a "true" ingredient of prominence, which characteristics are not, and why. Future research looking at VQ in high and low boundary tones, and on prenuclear pitch accents will be informative here and provide a more comprehensive understanding of position-dependent phonetic variation, thereby complementing existing work on prosodic strengthening across phrase positions [45]. These reflections align with studies showing increased reliance on spectral cues when F0 information is reduced or ambiguous [40].

In conclusion, our findings establish that VQ systematically differentiates high and low nuclear pitch accents in German through coordinated bundles of laryngeal parameters. This variation appears position-specific and independent of pitch height alone, suggesting functional linguistic control. Whether this constitutes a consciously manipulated prominence cue or an automatic implementation detail, and whether it meaningfully impacts listener perception, remain open empirical questions warranting further investigation.

5. Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 281511265 – SFB 1252 Prominence in Language (project A06), University of Cologne.

6. References

[1] S. Baumann and M. Grice, “The intonation of accessibility,” *Journal of Pragmatics,,* vol. 38, no. 10, p. 1636–1657, 2006.

[2] C. Féry and F. Kügler, “Pitch accent scaling on given, new and focused constituents in German,” *Journal of Phonetics,* vol. 36, no. 4, p. 680–703, 2008.

[3] S. Roessig, B. Winter and D. Mücke, “Tracing the phonetic space of prosodic focus marking,” *Frontiers in Artificial Intelligence,* vol. 5, no. 842546, 2022.

[4] C. Mooshammer, “Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German,” *JASA,* vol. 127, no. 2, pp. 1047-1058, 2010.

[5] C. Féry, German Intonational Patterns. Niemeyer, 1993.

[6] T. Rathcke and J. Harrington, “The variability of early accent peaks in Standard German,” *Phonetica,* vol. 67, no. 1-2, pp. 85-107, 2010.

[7] J. Barnes, N. Veilleux, A. Brugos and S. Shattuck-Hufnagel, “Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology,” *Laboratory Phonology,* vol. 3, no. 2, pp. 337-383, 2012.

[8] F. Cangemi, A. Albert and M. Grice, “Modelling intonation: Beyond segments and tonal targets.,” in *Proc. of the 19th ICPhS*, Melbourne, 2019.

[9] K. Görs and O. Niebuhr, “Hocus Focus – How prosodic profiles of contrastive focus emerge and change in different elicitation contexts,” in *Speech Prosody*, 2012.

[10] H. Seeliger and S. Repp, “Contrast, givenness, and (the lack of) accent shift and deaccentuation in non-assertive speech acts,” *Laboratory Phonology,* vol. 14, no. 1, pp. 1-46, 2023.

[11] S. Repp, “The prosody of *wh*-exclamatives and *wh*-questions in German: Speech act differences, information structure, and sex of speaker,” *Language and Speech,* vol. 63, no. 2, pp. 306–361, 2020.

[12] S. Repp and H. Seeliger, “Contrast and givenness in biased declarative questions,” in *Proc. of the 20th ICPhS*, 2023.

[13] F. Kügler and S. Genzel, “On the prosodic expression of pragmatic prominence: The case of pitch register lowering in Akan,” *Language and Speech,* vol. 55, no. 3, pp. 331–359, 2012.

[14] M. Grice, D. R. Ladd and A. Arvaniti, “On the place of phrase accents in intonational phonology,” *Phonology,* vol. 17, no. 2, pp. 143–185, 2000.

[15] T. Cho and J. McQueen, “Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress,” *Journal of Phonetics,* vol. 33, no. 2, pp. 121-157, 2005.

[16] M. Lialiou, M. Grice, C. Röhr and P. B. Schumacher, “Auditory processing of intonational rises and falls in German: Rises are special in attention orienting,” *Journal of Cognitive Neuroscience,* vol. 36, no. 6, pp. 1099–1122, 2024.

[17] A. M. C. Sluijter and V. J. van Heuven, “Spectral balance as an acoustic correlate of linguistic stress,” *JASA,* vol. 100, no. 4, pp. 2471–2485, 1996.

[18] A. M. C. Sluijter and V. J. van Heuven, “Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch,” *Phonetica,* vol. 54, no. 1, pp. 1–28, 1997.

[19] G. Dogil und B. Williams, “The Phonetic Manifestation of Word Stress, ” in *Word Prosodic Systems in the Languages of Europe,* Mouton de Gruyter, 1999, pp. 273-311.

[20] S. Wiener, A. A. Bramlett, B. Brown und J. Dueck, “Acoustic analysis and perception ratings of first and second language speakers’ Italian lexical stress, ” *Applied Linguistics*, 2025.

[21] K. Broś, “Acoustic cues to stress perception in Spanish – a mismatch negativity study,” in *Proc. of INTERSPEECH*, Dublin, 2023.

[22] G. Aronov and A. Schweitzer, “Acoustic correlates of word stress in German spontaneous speech,” in *Proceedings of the P&P*, 2016.

[23] A. Ní Chasaide, I. Yanushevskaya and C. Gobl, “Voice source dynamics in intonation,” in *Proc. of the 17th* ICPhS, 2011.

[24] Î. Yanushevskaya, C. Gobl and A. Ní Chasaide, “Voice quality and F0 cues for affect expression: Implications for synthesis,” in *Proc. of INTERSPEECH*, 2005.

[25] S. Baumann and B. Winter, “What makes a word prominent? Predicting untrained German listeners’ perceptual judgments,” *Journal of Phonetics,* vol. 70, pp. 2038, 2018.

[26] N. Campbell, and M. E. Beckmann, “Stress, prominence, and spectral tilt,” in *Proc. INT*, 1997.

[27] G. Kochanski, E. Grabe, J. Coleman and B. Rosner, “Loudness predicts prominence: Fundamental frequency lends little,” *JASA,* vol. 118, no. 2, pp. 1038–1054, 2005.

[28] M. Heldner, “Spectral emphasis as a perceptual cue to prominence,” *TMH-QPSR,* vol. 42, no. 1, p. 51–57, 2001.

[29] G. G. Severijnen, H. R. Bosker and J. M. McQueen, “Acoustic correlates of Dutch lexical stress re-examined: Spectral tilt is not always more reliable than intensity,” in *Speech Prosody*, 2022.

[30] Gerratt, B. R. and J. Kreiman, “Toward a taxonomy of nonmodal phonation,” *Journal of Phonetics,* vol. 29, no. 4, pp. 365–381, 2001.

[31] J. H. Esling, “Pharyngeal consonants and the aryepiglottic sphincter,” *Journal of the IPA,* vol. 26, no. 2, pp. 65–88, 1996.

[32] P. Keating, J. Kuang, M. Garellek and C. Esposito, “A cross-language acoustic space for vocalic phonation distinctions,” *Language,* vol. 99, no. 2, pp. 351–389, 2023.

[33] J. Kuang, “Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice,” *JASA,* vol. 142, no. 3, p. 1693–1706, 2017.

[34] P. Keating, M. Garellek and J. Kreiman, “Acoustic properties of different kinds of creaky voice,” in *Proceedings of the 18th International Congress of Phonetic Sciences*, 2015.

[35] A. Eriksson and M. Heldner, “The acoustics of word stress in English as a function of stress level and speaking style,” in *Proc. of the 18th ICPhS*, 2015.

[36] S. Repp and H. Seeliger, *Calibrating prosodic prominence in complex pragmatic settings: Prenuclear, nuclear and postnuclear prosodic effects of speech act andcomposite information structural categories (±given, ±contrast),* Submitted, 2025.

[37] Y. Jadoul, B. Thompson and B. De Boer, “ Introducing parselmouth: A python interface to praat,” *Journal of Phonetics,* vol. 71, pp. 1–15, 2018.

[38] P. Boersma and D. Weenink, “Praat: doing phonetics by computer, ver. 6.4.47, Computer program,” praat.org.

[39] D. Bates, M. Mächler, B. Bolker and S. Walker, “Fitting linear mixed-effects models using lme4,” *Journal of Statistical Software,* vol. 67, no. 1, pp. 1–48, 2015.

[40] J. Kuang and M. Liberman, “Integrating voice quality cues in the pitch perception of speech and non-speech utterances,” *Frontiers in Psychology,* vol. 9, 2147, 2018.

[41] J. Hillenbrand and R. A. Houde, “Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech,” *Journal of Speech, Language, and Hearing Research,* vol. 39, no. 2, pp. 311–321, 1996.

[42] J. Bishop and P. Keating, “Perception of pitch location within a speaker’s range: Fundamental frequency, voice quality and speaker sex,” *JASA,* vol. 132, no. 2, pp. 1100–1112, 2012.

[43] B. Braun, N. Dehé, J. W. D. Neitsch and K. Zahner, “The Prosody of Rhetorical and Information-Seeking Questions in German,” *Language and Speech,* vol. 62, no. 4, pp. 779–807, 2019.

[44] D. R. Ladd, Intonational Phonology, Cambridge: Cambridge University Press, 2008.

[45] T. Cho, “Prosodic boundary strengthening in the phonetics–prosody interface,” *Language and Linguistics Compass,* vol. 10, no. 3, pp. 120–141, 2016.