Acoustic correlates of the Danish voice quality contrast for male and female talkers

Jailyn Peña
Institute for Phonetics and Speech Processing, Ludwig Maximilian University of Munich (Germany)

This study examines the production of the Danish phonological phonation contrast between modal voice and 'stød,' a non-modal voice quality related to creaky voice [1]. Previous research finds that stød syllables are realized with high F0 and modal voice at the beginning of the rhyme followed by a dip in fundamental frequency (F0) and intensity, increased aperiodicity, and often the percept of creaky phonation by the end of the rhyme [2, 3]. In contrast, modal syllables are produced with regular, periodic vocal fold vibration, a short falling or fall-rise F0 contour, and higher, relatively stable intensity [4]. However, prior research has not focused on how stød is differentiated from modal voice by spectral tilt and aperiodicity (noise) in the acoustic signal, measures of which successfully differentiate phonological voice quality types in other languages, e.g., White Hmong [5]. Additionally, though previous research notes that stød is highly variable, little investigation of potential sources of variation has been carried out. One such factor is gender. For example, Esposito [6] found that male speakers distinguish creaky and modal voice using H1-A3 whereas female speakers do so with H1-H2 in Santa Ana Del Valle Zapotec. The present study addresses these gaps in the literature by examining the acoustic correlates to the Danish phonation contrast, including measures of spectral tilt and noise, as well as whether speaker gender may be a contributing factor to the variation.

In this study, stød and modal syllables with short vowels and coda sonorants were elicited from 37 native Danish speakers (25F, 12M). Using VoiceSauce [7], F0, strength of excitation, eight measures of spectral tilt (uncorrected measures were used for the coda), four harmonics-to-noise ratio (HNR) measures, cepstral peak prominence, and segment durations were extracted. Two linear discriminant analyses (LDAs) were then run including both overall mean measures and delta measures capturing the change over time per correlate, cf. [8]. Linear mixed effects regression models were conducted on the top three acoustic correlates per linear discriminant (LD) to investigate differences based on phonation type and gender.

In both LDAs, LD1 was primarily related to gender (Figure 1). For the vowel LDA, LD1 was most strongly correlated with mean F0, mean HNR<1500 Hz, and mean H2*kHz-H5*kHz. For the coda LDA, LD1 was most strongly correlated with mean HNR<2500 Hz, mean H4-H2kHz, and mean H2-H4. In contrast, LD2 was primarily related to phonation, correlating strongly with measures of HNR only (HNR<1500/2500/3500 Hz). Linear regression analyses revealed that female talkers had higher F0 during the vowel, higher H2-H4 during the coda, and less aperiodic noise (lower HNR) and lower spectral tilt in high frequencies (H2*kHz-H5*kHz, H4-H2kHz) throughout the rhyme compared to male speakers. In particular, these findings are consistent with previous descriptions of tense voice during the vowel [9]. Stød vowels also have higher F0 than modal vowels, consistent with previous findings [1], and lower H2-H4 during the coda. Interactions also revealed that only female talkers had significantly higher H4-H2kHz in stød codas compared to modal ones, indicating an increase in breathiness [5, 7, 9], and that the difference in HNR between male and female talkers was greater in modal codas than stød codas (Figure 2).

Overall, the results indicate that lower HNR throughout the rhyme is the most consistent cue to the Danish phonation contrast across talker groups. Furthermore, female talkers produce vowels as acoustically more tense than male speakers (higher pitched, less noisy, lower spectral tilt in higher frequencies), with stød vowels being more tense than modal ones. Female talkers also produce stød codas with greater breathiness, whereas male talkers do not. Finally, speakers showed a smaller difference in HNR for stød codas than modal ones, indicating greater convergence towards an acoustic target when producing stød, similar to findings on creaky voice in other languages, e.g., [10]. Together, these results suggest that although there may be some overlap, male and female talkers may use somewhat different articulatory strategies to produce stød.

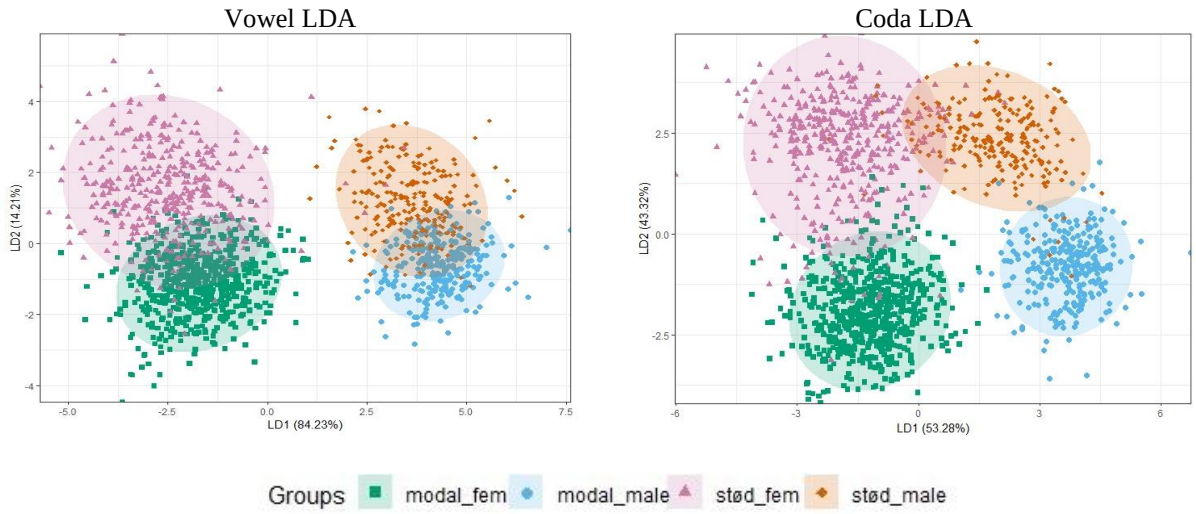

Fig.1: LD1-LD2 spaces for the vowel (left) and coda (right) LDAs.

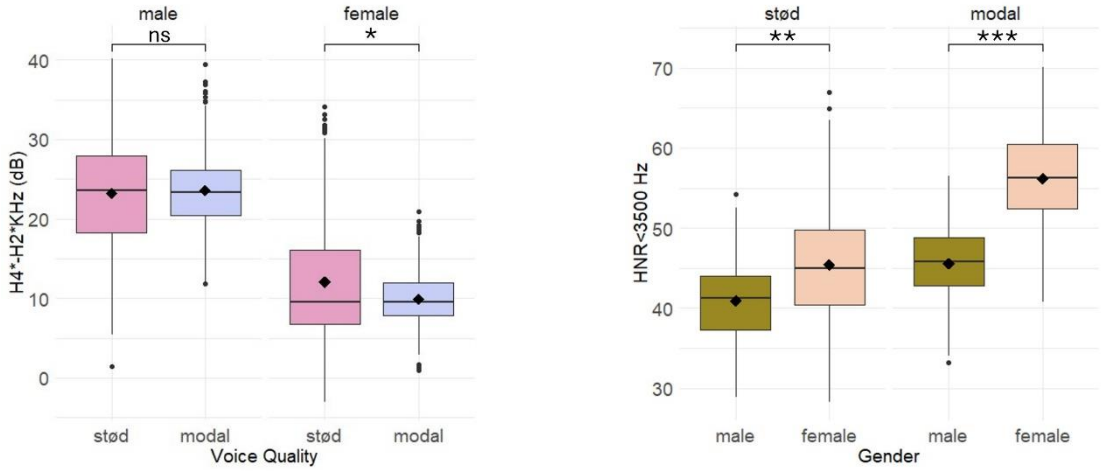

Fig.2: Gender x Phonation Interactions for H4-H2kHz (left) and HNR<3500 Hz (right) during the coda.

References

[1] N. Grønnum, "Three quarters of a century of phonetic research on common Danish stød," *Nordic Journal of Linguistics*, pp. 1–32, 2022, doi: 10.1017/S0332586522000178.

[2] N. Grønnum, M. Vazquez-Larruscaín, and H. Basbøll, "Danish Stød: Laryngealization or Tone," *Phonetica*, vol. 70, no. 1–2, pp. 66–92, 2023, doi: 10.1159/000354640.

[3] E. Fischer-Jørgensen, "A phonetic study of the stød in Standard Danish," *Annual Report of the Institute of Phonetics University of Copenhagen*, vol. 21, pp. 55–265, 1897, doi: 10.7146/aripuc.v21i.131884.

[4] N. Thorsen, "On the Variabilty in F₀ Patterning and the Function of F₀ Timing in Languages where Pitch Cues Stress," *Phonetica*, vol. 39, no. 4–5, pp. 302–316, 1982, doi: 10.1159/000261669.

[5] M. Garellek and C. M. Esposito, "Phonetics of White Hmong vowel and tonal contrasts," *Journal of the International Phonetic Association*, vol. 53, no. 1, pp. 213–232, 2023, doi: 10.1017/S0025100321000104.

[6] C. M. Esposito, "Variation in contrastive phonation in Santa Ana Del Valle Zapotec," *Journal of the International Phonetic Association*, vol. 40, no. 2, pp. 181–198, 2010, doi: 10.1017/S0025100310000046.

[7] Y. L. Shue, P. Keating, C. Vicenik, and K. Yu, "VoiceSauce: a program for voice analysis," in *17th International Congress of Phonetic Sciences*, 2011, pp. 1846–1849. doi: 10.1121/1.3248865.

[8] M. Garellek, "Acoustic Discriminability of the Complex Phonation System in !Xóõ," *Phonetica*, vol. 77, no. 2, pp. 131–160, 2020, doi: 10.1159/000494301.

[9] J. Kuang, "The Tonal Space of Contrastive Five Level Tones," *Phonetica*, vol. 70, pp. 1–23, 2013, doi: 10.1159/000353853.

[10] M. Blomgren, Y. Chen, M. L. Ng, and H. R. Gilbert, "Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers," *The Journal of the Acoustical Society of America*, vol. 103, pp. 2649–2658, 1998, doi: 10.1121/1.422785.