Creaky voice in South Swedish accent 1

Anna Hjortdal
Centre for Languages and Literature, Lund University, Sweden
anna.hjortdal@ling.lu.se

Abstract

Pitch and voice quality are increasingly understood as closely intertwined. While Swedish and Norwegian word accents have traditionally been understood in terms of pitch, Danish stød, which is systematically related when it comes to distribution and function, has been described as a type of creaky voice. According to the Laryngeal Articulator Model (LAM), both pitch lowering and creaky or harsh voice can be the acoustic outcomes of tightening the laryngeal constrictor mechanism. Laryngeal constriction has been proposed as the articulatory gesture behind word accents and stød. The present study investigated creaky voice in South Swedish word accents. Harmonics-to-noise ratio was significantly lower and jitter significantly higher in accent 1 compared to accent 2 stressed vowels. Further, jitter and shimmer was higher and spectral tilt was lower in sonorant consonants following stressed vowels. The results suggest that prototypical creaky voice is another cue to accent 1 in South Swedish and is in line with proposals that pitch falls in word accents correspond to laryngeal constriction.

Introduction

Most Mainland North Germanic language varieties have an accent distinction realised as different prosodic patterns on words, typically termed 'accent 1' and 'accent 2' (Gårding, 1977). The distribution is similar across language varieties and is to a large extent determined by phonology and morphology (Basbøll, 2005; Bruce, 1977; Gårding, 1977; Riad, 2012). There is high variability in realisation. In Norwegian and Swedish, word accents are typically understood as different pitch contours on words but with varying pitch patterns across dialects. In Danish, there is a distinction between a creaky voice feature, 'stød', and modal voice, 'non-stød' which some see as language-specific realisations of accent 1 and 2 respectively (Gårding, 1977; Öhman, 1967).

Within the framework of the Laryngeal Articulator Model (LAM) (Esling et al., 2019), both falling pitch and creaky voice could be understood as correlates of laryngeal constriction. Hypothesising that Swedish and Norwegian word accents and Danish stød are different acoustic outcomes of laryngeal constriction, creaky voice was investigated in South Swedish word accents.

South Swedish

In South Swedish disyllabic words, accent 1 is realised as a falling pitch in the first, stressed syllable whereas accent 2 is realised as a rise or rise-fall. This is in contrast to Central Swedish in which pitch falls from the prestressed to the stressed syllable in accent 1 and falls in the stressed syllable in accent 2 (Bruce, 1977, 2010). Svensson Lundmark et al. (2017) reported that creaky voice was present in a significantly higher proportion of accent 1 than accent 2 words. This was especially prominent in stressed vowels. Comparison with pitch patterns suggested that creaky voice was related to sharply falling or low pitch.

The laryngeal articulator model

According to the Laryngeal Articulator Model (LAM) (Esling et al., 2019), the lower vocal tract is a complex articulator, its importance corresponding to that of the tongue in the upper vocal tract. Within this framework, vowel, tonal and phonatory qualities are intertwined. The laryngeal articulator itself corresponds to the aryepiglottic muscle which is part of the aryepiglottic folds located above the epiglottis, i.e. supraglottally. Tightening the laryngeal constrictor mechanism, leads to thickening and 'bunching' of the vibrating structures which lowers the vibration frequency of vocal folds and thus pitch. It is associated with creaky and harsh voice quality (Esling et al., 2019).

Öhman's intonation model

Proposing a unifying model of the physiological basis of Scandinavian word accents, Öhman (1967) suggested that the diverse realisations of word accents, including the Danish stød/non-stød distinction, share a physiological basis. Pitch falls are 'negative pulses' imposed on 'phonatory stress pulses'. The phonatory stress pulses are centred at the beginning of stressed syllables and have rising pitch. The negative pulses, in turn, yield a pitch fall. The negative pulses are described as 'laryngeal consonants' and compared to glottal stops "only softer than the Danish ones" (Öhman, 1967, p. 29).

In a study using fibre optics, Lindqvist-Gauffin (1972b) reported laryngeal constriction for falling pitch in Stockholm Swedish accent 2. He suggested that pitch drops in Swedish word accents are produced by means of a laryngeal closing gesture; a sphincteric constriction at the aryepiglottic folds originating as a protective closure to avoid choking on food or water and recycled for articulation (Lindqvist-Gauffin, 1972a). According to Lindblom (2009), Lindqvist-Gauffin's ideas are in line with later work of Esling and colleagues. Esling et al. (2019) also suggested that the Danish creaky voice feature 'stød', which has been described as having different acoustic outcomes such as a glottal stop, creaky voice, an intensity drop and pitch effects (Fischer-Jørgensen, 1989; Hansen, 2015), is consistent with laryngeal constriction.

Voice quality and accentuation

According to the Voice Prominence Hypothesis (Ní Chasaide et al., 2013), accentuation can be expressed

through a variety of parameters in addition to pitch. Examining declaratives, WH questions and Yes/No questions in Connemara Irish, Ní Chasaide et al. (2013) found that where pitch prominence is not employed, other parameters such as voice quality are more prominent and that speakers may vary in the extent to which different parameters are exploited in expressing accentuation.

Present study

Within the framework of the Laryngeal Articulator model, both word accents and stød might be understood as varying degrees of laryngeal constriction, resulting in different acoustic outcomes and cue salience.

The present study investigated voice quality in South Swedish word accents. Since laryngeal constriction can result in pitch lowering as well as creaky and harsh voice (Esling et al., 2019), creaky or harsh voice is expected to cooccur with falling pitch. In South Swedish, pitch is falling in the stressed syllable in accent 1 words but rising in the stressed syllable in accent 2 words. More creaky voice is therefore expected in accent 1 compared to accent 2 stressed syllables.

Method

Participants

Thirteen speakers of South Swedish produced minimal pairs with accent 1 and accent 2. Mean age was 27.2, SD = 6.8, 5 female. All had lived the majority of their childhood in Scania. Participants gave informed consent and filled out a form. Data from one participant was excluded because the task was misunderstood.

Stimulus words

Stimuli were 6 Swedish minimal pairs differing only with respect to word accent to minimise interactions with oral resonances. The stimulus words were mono- and disyllabic nouns with accent 1 and accent 2, respectively, both in definite form. All stimulus words had short vowels, either /a/, /ʉ/ or /ɔ/. Half were followed by sonorants, the other half by obstruents. Stimulus words were incorporated into carrier sentences ending in prepositional phrases to avoid effects of terminal juncture intonation (Bruce, 1977). The order of stimulus words varied so that for half the stimuli, the accent 1 sentence was read before the accent 2 sentence and vice versa to avoid systematic influence from intonation patterns. This order varied between participants. All minimal pairs had the same spelling. To avoid ambiguity about which word participants were intended to read, each stimulus sentence was preceded by an English translation, taking advantage of knowledge the vast majority of Swedes (86 %) have of English as a second language[1]. Below is an example of a carrier sentence for the accent 1 stimulus word *tomten* 'plot of land':

1) *I saw the plot of land by the forest.*
   Jag såg tomten vid skogen

[1] *Special Eurobarometer 386: Europeans and their Languages.* 2012. Conducted by TNS Opinion & Social at the request of European Commission.

Table 1. Stimulus words were six Swedish minimal pairs with accent 1 and accent 2 respectively. Accent 1 is marked as ´and accent 2 as `.

Accent 1	Accent 2
ánden 'the duck'	*ànden* 'the spirit
bácken 'the back'	*bàcken* 'the hill'
kúllen 'the litter'	*kùllen* 'the hill'
stúbben 'the stubble'	*stùbben* 'the stump'
tómten 'the plot of land'	*tòmten* 'Santa Claus'
skótten 'the shots'	*skòtten* 'the Scotsman'

Recording procedure

Participants were given lists with stimulus sentences and instructed to read the English sentences to themselves and subsequently read the Swedish ones out loud. They were asked to focus on the prepositional phrases to avoid sentence accent (Bruce, 1977) on the stimulus words. Recordings were made with a Sony PCM-D50 field recorder.

Acoustic measurements and analysis

Voice quality was measured in the vowel in the stressed syllable. Measurements were also made for sonorant consonants following short vowels. Danish stød is realised in either the latter half of long vowels or in sonorant consonants following short vowels (Fischer-Jørgensen, 1989), indicating that this might also be a relevant place to look for creaky voice in South Swedish.

Creaky voice was investigated by measuring harmonics-to-noise ratio (HNR), jitter, shimmer and spectral tilt. HNR is the difference in amplitude between harmonic and noisy parts of the voice spectrum and is typically lower for creaky voice (Garellek, 2019). Jitter and shimmer are indexes of the regularity in frequency and amplitude, respectively (Esling et al., 2019). Both are higher for harsh voice (Esling et al., 2019) and jitter is usually higher for creaky voice since voicing is less regular (Keating et al., 2015). For each stimulus word, HNR, jitter and shimmer were measured in the vowels and sonorant consonants, respectively, in the stressed syllable using a Praat script. According to Garellek (2019), the most reliable correlate of voice constriction is lower spectral tilt. For pressed or constricted voice, vocal fold closure is typically quicker than in normal voice and this closing pattern leads to a flatter intensity decrease for harmonics, i.e. lower spectral tilt (Kreiman & Sidtis, 2011). The most common measure of spectral tilt is H1-H2 which is the difference between the amplitudes of the first and second harmonics (Keating & Esposito, 2006). Measures over different frequencies, such as H1-A1, H1-A2 and H1-A3, are concerned with the difference between H1 and the harmonics closest to the first three formants and has been found to be relevant measures in some languages (Garellek, 2019; Keating & Esposito, 2006). Spectral tilt was measured using a Praat (Boersma & Weenink, 2020) script developed for the purpose by Chad Vicenik[2].

The term 'creaky voice' encompasses several different sub-categories of voice quality. 'Prototypical

[2]http://phonetics.linguistics.ucla.edu/facilities/acoustic/praat.html

creaky voice' is typically characterised by low and irregular F0 (seen in e.g. increased jitter) and a constricted glottis (seen in lower spectral tilt, e.g. lower H1-H2). Other types of creaky voice share some of these characteristics, but not all. In 'vocal fry', pitch is typically regular and less noisy than modular voice, 'unconstricted creaky voice' does not have lower spectral tilt and in 'tense/constricted voice', F0 is not low (Keating et al., 2015).

In total, voice quality for 144 words was measured. The data was analysed in linear-mixed effects models using the lmerTest package (Kuznetsova et al., 2017) in the R software (R Core R Core Team, 2017). Linear mixed-effects models have been argued to be superior to by-subject or by-item analyses because they simultaneously model variance associated with participants and stimulus words (Barr et al., 2013). By-item and by-subject random intercepts and random slopes for the fixed effects investigated were added.

Results

All means and standard deviations are reported in Table 2. In the stressed vowel, HNR was significantly lower and jitter significantly higher for words with accent 1 compared to words with accent 2. See Figures 1 and 2. In the sonorant consonants, jitter and shimmer were significantly higher in accent 1 words. There were no significant differences in the spectral tilt measures during the stressed vowel, but in the following sonorant consonant, spectral tilt was lower for accent 1 words. See Figures 1 and 3.

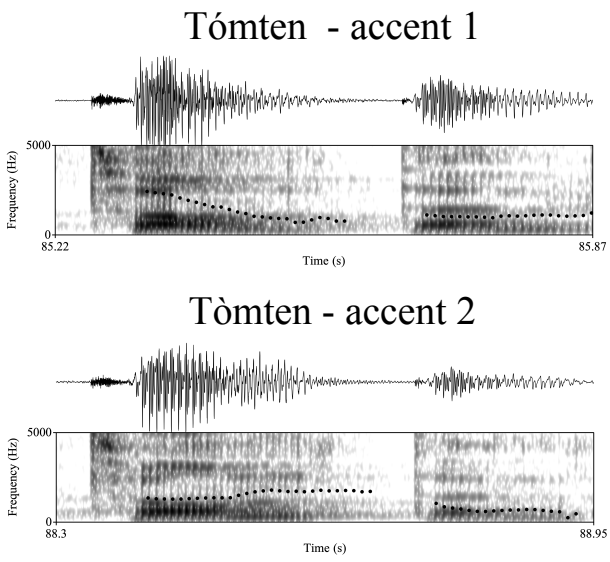

Figure 1. As can be seen above from a soundwave and spectrogram example from the data, more irregular pitch (higher jitter) and noise (lower HNR) accompanied the F0 fall during the stressed syllable in accent 1 compared to accent 2 words. Spectral tilt was lower during sonorant consonants following stressed vowels in accent 1 words.

Table 2. Results for HNR, jitter, shimmer and spectral tilt. Differences that are significant at the 0.05 level are in **bold**. Results for male and female speakers are collapsed.

Measure	Accent 1	Accent 2	p-value
Vowels			
HNR, dB	**7.8 ± 3.6**	**10.3 ± 4.2**	**0.012**
Jitter, %	**2.71 ± 1.5**	**1.7 ± 1.0**	**0.011**
Shimmer, %	12.2 ± 5.0	10.1 ± 4.9	N.S.
H1-H2, dB	1.6 ± 4.8	1.4 ± 4.5	N.S
H1-A1, dB	0.1 ± 8.7	-0.1 ± 8.7	N.S
H1-A2, dB	11.7 ± 8.9	12.8 ± 8.8	N.S.
H1-A3, dB	25.6 ± 8.5	26.6 ± 8.9	N.S.
Sonorant consonants			
HNR, dB	10.6 ± 4.1	13.7 ± 4.3	N.S.
Jitter, %	**2.3 ± 1.3**	**1.6 ± 0.9**	**0.042**
Shimmer, %	**12.9 ± 5.5**	**9.7 ± 5.1**	**0.009**
H1-H2, dB	**1.6 ± 7.3**	**4.9 ± 8.0**	**0.044**
H1-A1, dB	8.0 ± 11.1	12.1 ± 15.1	N.S.
H1-A2, dB	**23.8 ± 7.8**	**27.6 ± 8.4**	**0.019**
H1-A3, dB	32.1 ± 7.8	35.7 ± 10.0	N.S.

Discussion

The present study investigated voice quality in South Swedish, hypothesising that accent 1 stressed syllables would have more creaky voice than accent 2 stressed syllables. In accent 1 words, HNR was lower and jitter higher in stressed vowels than in accent 2, indicating more noise and irregularity. Sonorant consonants following stressed vowels in accent 1 words had higher jitter and shimmer and lower spectral tilt than sonorant consonants following stressed vowels in accent 2 words. Overall, the results support the previous finding that South Swedish accent 1 is associated with more creaky voice in the stressed syllable (Svensson Lundmark et al., 2017).

In the present study, different acoustic measures relating to creaky voice were investigated, enabling more detailed analysis of the type of creaky voice. The creaky voice in accent 1 is associated with the syllable with low or falling pitch (Svensson Lundmark et al., 2017) and there is more irregularity, noise and constriction (at least in sonorant consonants) than in corresponding accent 2 syllables. Together, these acoustic measures indicate that voice quality in South Swedish accent 1 stressed syllables has the characteristics of prototypical creaky voice.

The findings indicate that creaky voice is an additional cue to accent 1 in South Swedish. This is in line with the Voice Prominence Hypothesis (Ní Chasaide et al., 2013), according to which accentuation can be expressed through a variety of different acoustic cues. In addition to pitch and creaky voice, duration has also been found to be a relevant cue to accent distinction in Swedish. Ambrazaitis and Tronnier (2021) reported shorter vowel and consonant durations for the stressed syllable in accent 1 as compared to accent 2. Future studies could examine the perceptual relevance of creaky voice and duration for speakers of South Swedish.

More creaky voice in accent 1 is also consistent with laryngeal constriction as has been proposed for Swedish word accents (Lindqvist-Gauffin, 1972a) and Danish stød (Esling et al., 2019). Constriction per se was not significantly different during stressed vowels but rather only during sonorant consonants following stressed vowels, suggesting that constriction is imposed on words as they unfold. The present study was concerned with acoustics, but future studies could employ articulatory techniques to further investigate the research question.

Riad (2009) proposed that stød developed from an accent type similar to the Eskilstuna variety where creaky voice/stød often cooccurs with a steep pitch fall. Contrary to Danish, however, the stød has not been phonologized in the Eskilstuna dialect (Riad, 2000; Riad, 2009). If both stød and word accents are the acoustic outcomes of different degrees of laryngeal constriction (Esling et al., 2019; Lindqvist-Gauffin, 1972a; Öhman,

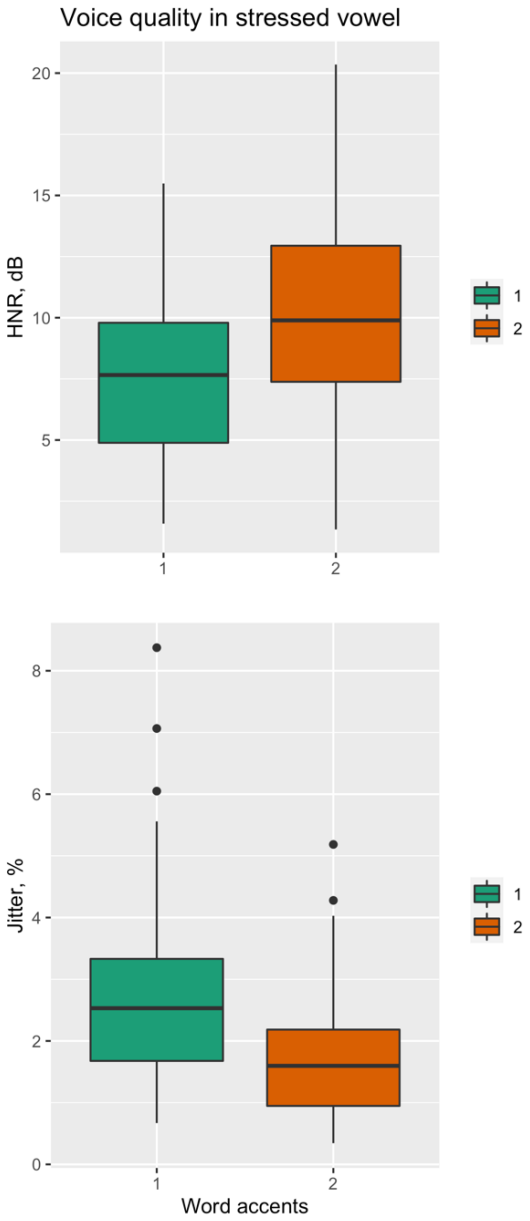

Figure 2. HNR was significantly lower and jitter significantly higher during stressed accent 1 vowels, indicating more noise and irregularity.

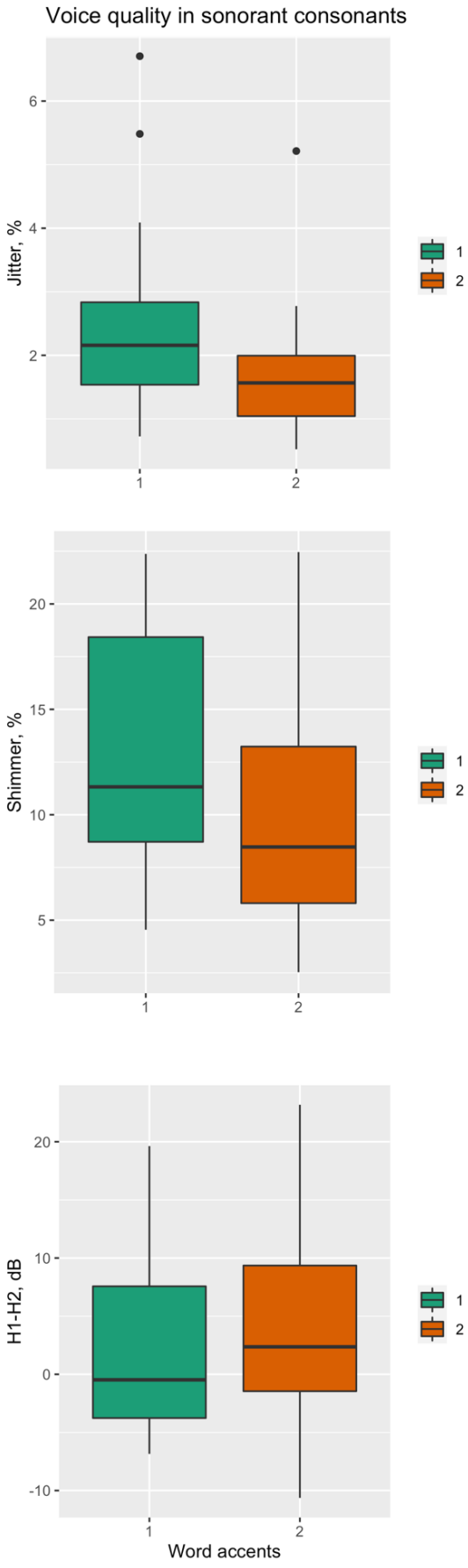

Figure 3. Jitter and shimmer was significantly higher and H1-H2 significantly lower in sonorant consonants following stressed vowels in accent 1 words.

1967), perhaps different acoustic cues of the same articulatory gesture have been phonologized in the different language varieties - although with remnants of associated acoustic correlates such as creaky voice for sharply falling pitch.

Öhman's (1967) laryngeal consonants could perhaps be understood within the framework of Articulatory Phonology. In Articulatory Phonology, phonological units are articulatory gestures rather than acoustic features (Hall, 2010). Physiologically, gestures consist of the formation and subsequent release of 'constrictions'. The constrictions are specified for location in the vocal tract and degree of narrowness (Hall, 2010). Svensson Lundmark (2020) compared Öhman's pulses to 'tone gestures' (Gao, 2010) and suggested that Swedish word accents could be understood as falling or rising tone gestures. Öhman's (1967) laryngeal consonants could be understood as 'laryngeal constriction' articulatory gestures.

In conclusion, the study showed that voice quality is more creaky in South Swedish accent 1 stressed syllables compared to corresponding accent 2 syllables. The findings indicate that creaky voice is another acoustic cue to accent 1 in South Swedish. The results are also in line with proposals that pitch falls in word accents correspond to tightening of the laryngeal constriction mechanism (Esling et al., 2019; Lindqvist-Gauffin, 1972a; Öhman, 1967).

Acknowledgements

I would like to thank Merle Horne, Mechtild Tronnier and Mikael Roll for valuable comments on previous versions of this paper. This work was supported by the Swedish Research Council (grant number 2018.00632).

References

Ambrazaitis, G., & Tronnier, M. (2021). Segmental durations as a correlate of Swedish word accents: Evidence from Stockholm and Scania Swedish. Working Papers: Proceedings of Fonetik 2021, Lund.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Basbøll, H. (2005). *The Phonology of Danish*. Oxford University Press.

Boersma, P., & Weenink, D. (2020). *Praat: doing phonetics by computer (Version 6.1.16)*. https://praat.org

Bruce, G. (1977). *Swedish word accents in sentence perspective* (Vol. XII). CWK Gleerup.

Bruce, G. (2010). *Vår fonetiska geografi*. Studentlitteratur AB.

Esling, J. H., Moisik, S. R., Benner, A., & Crevier-Buchman, L. (2019). *Voice Quality: The Laryngeal Articulator Model*. Cambridge University Press. https://doi.org/https://doi.org/10.1017/9781108696555

Fischer-Jørgensen, E. (1989). *A phonetic study of the stød in standard Danish*. University of Turku Phonetics.

Gao, M. (2010). *Articulatory Phonology (AP) and tonal alignment: further testing of a proposed AP model of tone-to-segment alignment in Mandarin Chinese* Tone and Intonation in Europe (TIE4), Stockholm, Sweden.

Gårding, E. (1977). *The Scandinavian Word Accents* (Vol. XI). CWK Gleerup.

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. F. Assmann (Eds.), *The Routledge handbook of phonetics*. Routledge Taylor and Francis.

Hall, N. (2010). Articulatory Phonology. *Language and Linguistics Compass*, *4*, 818-830. https://doi.org/10.114 1/.1749-818x 2010,00236.x

Hansen, G. F. (2015). *Stød og stemmekvalitet: En akustisk-fonetisk undersøgelse af ændringer i stemmekvaliteten i forbindelse med stød* [Doctoral thesis, Copenhagen University]. Copenhagen.

Keating, P., & Esposito, C. (2006). Linguistic Voice Quality. *UCLA Working Papers in Phonetics*, *105*, 85-91.

Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. ICPhS, Glasgow.

Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies*. Wiley-Blackwell.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss

Lindblom, B. (2009). F0 lowering, creaky voice, and glottal stop: Jan Gauffin's account of how the larynx works in speech. *Proceedings, FONETIK 2009, Dept. of Linguistics, Stockholm University*.

Lindqvist-Gauffin, J. (1972a). A descriptive model of laryngeal articulation in speech. *STL-QPSR*, *13*(2-3), 1-9.

Lindqvist-Gauffin, J. (1972b). Laryngeal articulation studied on Swedish subjects. *STL-QPSR*, *13*(2-3), 10-27.

Ní Chasaide, A., Yanushevskaya, I., Kane, J., & Gobl, C. (2013). The Voice Prominence Hypothesis: the Interplay of F0 and Voice Source Features in Accentuation. *Proceedings of Interspeech 2013*, 3527-3531. https://doi.org/10.21437/Interspeech.2013-759

Öhman, S. (1967). Word and sentence intonation: A quantitative model. *STL-QPSR*, *4*, 1-6.

R Core Team. (2017). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. http://www.R-project.org/

Riad, T. (2000). The Origin of Danish stød. In A. Lahiri (Ed.), *Analogy, Levelling, Markedness*. Mouton de Gruyter.

Riad, T. (2009). Eskilstuna as the tonal key to Danish. *Proceedings, FONETIK 2009, Dept. of Linguistics, Stockholm University*.

Riad, T. (2012). Culminativity, stress and tone accent in Central Swedish. *Lingua*, *122*(13), 1352-1379. https://doi.org/10.1016/j.lingua.2012.07.001

Svensson Lundmark, M. (2020). *Articulation in time: Some word-initial segments in Swedish* Lund University]. Lund.

Svensson Lundmark, M., Ambrazaitis, G., & Ewald, O. (2017). Exploring multidimensionality: Acoustic and articulatory correlates of Swedish word accents. *Interspeech*, 3236-3240. https://doi.org/10.21437/Interspeech.2017-1502