

Article

Stød Timing and Domain in Danish

Jailyn M. Peña

Department of Linguistics, New York University, New York, NY 10003, USA; jailynpena@nyu.edu

Abstract: This study investigates the timing of stød, a type of phonological nonmodal phonation related to creaky voice in Danish, relative to the syllable. Stød-bearing syllables are characterized by high fundamental frequency (F0) and modal phonation at the beginning of the syllable followed by nonmodal, often creaky phonation and low F0 towards the end of the syllable (the stød phase proper). However, the timing of these two phases relative to the syllable and to each other has been debated. To investigate this, F0 throughout the word and the timing of the stød phase proper relative to the syllable were analyzed in five types of monosyllabic words. The results show that across word types the first stød phase (high F0) coordinates with the syllable rhyme onset, whilst the second phase is timed to the center of the sonorant rhyme, in contrast to previous hypotheses of stød timing. This relationship is formalized using the framework of Articulatory Phonology. In doing so, two additions to the theory are proposed to account for the biphasic nature of stød and the timing of the stød phase proper relative to the syllable.

Keywords: Danish; stød; creaky phonation; nonmodal phonation; voice quality; articulatory phonology; suprasegmental timing



Citation: Peña, Jailyn M. 2022. Stød Timing and Domain in Danish. *Languages* 7: 50. <https://doi.org/10.3390/languages7010050>

Academic Editors: Christina Esposito and Sameer Khan

Received: 15 April 2021

Accepted: 10 February 2022

Published: 24 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

This paper investigates the timing of stød, a type of phonological, nonmodal phonation related to creaky voice in Danish (Fischer-Jørgensen 1989; Grønnum 1998; Basbøll 2005), relative to the syllable. Previous work has shown that stød is best characterized as biphasic in its realization, with high fundamental frequency (F0) and modal phonation at the beginning of stød-bearing syllables followed by a second phase often characterized by irregular vibratory pulses and nonmodal, creaky phonation (Fischer-Jørgensen 1989). This second phase has been termed the ‘stød phase proper’ or simply ‘stød’ in previous literature, and its timing relative to the syllable has been the subject of much debate, with some research claiming it is linked to a mora, and other literature claiming it is a property of the syllable (Fischer-Jørgensen 1989; Grønnum and Basbøll 2001; Basbøll 2005; Grønnum et al. 2013).

This study investigates the timing of the two stød phases relative to each other and to the syllable. Here, three timing hypotheses between the stød phase proper and the syllable are investigated, as well as independent evidence for morae in Danish from durational data, following Broselow et al. (1997) and Morén and Zsiga (2006). The results of this study suggest that the relationship between stød and the syllable is not modulated by morae. Furthermore, the timing relationship between stød and the syllable can be modeled using the framework of Articulatory Phonology (Browman and Goldstein 1986, 1989a, 1989b, 1990, 1992; Browman et al. 1990). To do so, two theoretical additions are proposed. First, I propose that the GLOTTAL tier in traditional frameworks of Articulatory Phonology can be subdivided into two tiers, the PHONATION tier and the F0 tier. This division allows for the phonological specification of multiple glottal configurations in the production of laryngeal/glottal phenomena, each of which may have independent effects on phonation and pitch. I motivate this division by showing its necessity to account for stød’s biphasic realization. Second, I propose that, analogous to the c-center effect wherein onset clusters act as a single unit relative to a following vowel gesture, the sonorant portion of the syllable

rhyme, henceforth referred to as the sonorant rhyme, acts as a single unit relative to stød's gestural constellation in Danish. This allows for the implementation of a new articulatory landmark, the sonorant rhyme center, which can enter into coordination relationships with other articulatory gestures.

In the remainder of the introduction, I first outline the acoustic and articulatory differences between modal and nonmodal phonation, particularly as it pertains to stød in Danish, before reviewing three possible timing relationships between stød and the syllable. I then outline the research questions of this paper before reviewing the key aspects of Articulatory Phonology (Browman and Goldstein 1986, 1989a, 1989b, 1990, 1992) relevant to this study. Section 2 details the methods of this study before the results of the analyses are presented in Section 3. The results are discussed in Section 4, before the conclusion in Section 5.

1.1. Acoustics and Articulation of Modal Phonation, Creaky Phonation, and Stød

Some languages use phonation phonologically to create lexical contrasts, e.g., Gujarati (Khan 2012), Mazatec (Garellek and Keating 2011), White Hmong (Esposito and Khan 2012; Garellek 2012), and Zapotec (Avelino 2010; Esposito 2010). This paper focuses on one such language, Danish, which contrasts modal phonation with a type of nonmodal phonation related to creaky voice, termed 'stød' in previous literature (Ege 1965; Basbøll 1985, 2005; Fischer-Jørgensen 1989; Grønnum 1998).

Acoustically, modal phonation, the most common vocal register used in speech, is best described in contrast to nonmodal phonation. Modally phonated segments, particularly sonorants, are typically characterized by higher F0, higher and more regular intensity, and more periodicity in the acoustic signal than segments produced with nonmodal phonation (Garellek 2015, 2019; Esling et al. 2019). Articulatorily, this is achieved with regular, periodic vocal fold vibration during pulmonic egressive airflow with minimal or no laryngeal constriction (Esling et al. 2019). In contrast, segments produced with creaky phonation, a type of nonmodal phonation, are characterized as having lower F0, more irregular intensity, and less periodicity in the acoustic signal than modally phonated ones (Keating et al. 2015; Davidson 2020).

Traditionally, the acoustic differences between modal and creaky phonation have been attributed to differing states of the glottis (Catford 1964, 1977). However, recent evidence suggests that the articulation of creaky phonation often involves constriction of multiple laryngeal articulators, not just the vocal folds at the glottis, and that differences in the precise articulation can lead to distinct acoustic outputs (Esling et al. 2019). For example, Esling et al. (2019) review three types of creaky phonation, which they refer to as creaky voice, also called 'glottal fry' or 'vocal fry' (Hollien et al. 1966; Hollien and Michel 1968), harsh voice (Laver 1980), and ventricular voice (Laver 1980), each of which has a distinct articulatory state of the larynx associated with it, though there is overlap across the three types. According to Esling et al., in the production of creaky voice, not only are the vocal folds shortened and thickened, generally due to thyroarytenoid muscle contraction, but the ventricular folds also often encroach upon the vocal folds. This increases the vibratory mass during voicing, lowering speakers' F0 and increasing irregularity in vocal fold vibration. In contrast, what Esling et al. refer to as harsh voice, with greater noise and aperiodicity in the acoustic signal, results from additional aryepiglottic constriction and ventricular adduction. Narrowing of these structures constricts the region immediately above the glottis, which they claim induces vibration in epilaryngeal structures during voicing, introducing more noise into the acoustic signal. For a full review of the different types of nonmodal phonation, see Esling et al. (2019).

Returning to the language at hand, in many Danish dialects, including the standard and dialects spoken around Copenhagen, the capital, the nonmodal phonation type stød phonologically contrasts with modal phonation, as show below in Table 1. Stød is represented by a superscript [ʔ]. In terms of its distribution in the lexicon, stød can only occur in stressed syllables with either a long vowel (Danish has contrastive vowel length) or

a short vowel followed by a sonorant coda consonant (V+S), a requirement referred to as ‘stød-basis’ (Basbøll 1985). Stød occurs in both monomorphemic and polymorphemic words, and in monosyllabic and polysyllabic words. It also displays complex interactions with morphology. As Basbøll (1985) notes, in most cases when an affix is attached to a stem in Danish, a stem that already had stød retains it, and a stem that did not have stød remains stød-less. However, some suffixes have been analyzed as causing stød-removal from a monosyllabic stem, some condition stød-addition, and some apply these processes inconsistently across the lexicon. Therefore, to limit the scope of this study, here, only monosyllabic words are examined; for a detailed account of stød’s phonological distribution, its interactions with morphology, and its distribution in polysyllabic words, see Basbøll (2005).

Table 1. Minimal pairs in Danish, contrasting in phonation. Stød-bearing words are on the left, and stød-less words are on the right.

Stød-Bearing Words	Stød-Less Words
[vɛnʔ] ‘turn (imperative)’	[vɛn] ‘friend’
[kʰʌmʔ] ‘come (past)’	[kʰʌm] ‘come (imperative)’
[ˈlɛ:ʔ.sɐ] ‘reads (present)’	[ˈlɛ:.sɐ] ‘reader’

Previous acoustic work has shown that syllables with stød are typically characterized by modal phonation and high or rising F0 in the first part of the rhyme, followed by a second phase, the stød phase proper, or simply stød (Fischer-Jørgensen 1989). This phase can manifest with a variety of acoustic correlates, including irregular vocal fold vibration/aperiodicity in the spectrogram (creaky phonation), a local dip in F0, irregular amplitude, decreased intensity compared to modal phonation, and (rarely) full glottal closure (Fischer-Jørgensen 1989; Grønnum et al. 2013). However, studies have found a high degree of variability in stød realization, particularly with respect to the stød phase proper, and not all of these acoustic correlates may be present in any given instantiation (Fischer-Jørgensen 1989; Grønnum and Basbøll 2001; Grønnum et al. 2013). Descriptions of a prototypical stød phase, the most common manifestation, resemble descriptions of prototypical creak as described by Keating et al. (2015), often occurring with some combination of irregular vocal fold vibration and aperiodicity in the spectrogram, low F0, irregular amplitude, and decreased intensity. A maximally ‘strong’ stød phase is reported to result in full glottal closure, though this is said to be relatively rare and possibly related to focus (Grønnum et al. 2013). Finally, ‘weaker’ instantiations without obvious acoustic correlates are also reported (Fischer-Jørgensen 1989; Grønnum and Basbøll 2001, 2003a, 2003b; Grønnum et al. 2013). For these instantiations, stød is still said to be perceptually salient to Danish listeners, despite the lack of obvious acoustic correlates. These descriptions have led researchers to conclude that stød is related to but not exactly creaky phonation.

Articulatorily, stød has been linked to multiple laryngeal articulators, including the cricothyroid, the vocal folds, and the ventricular folds (Fischer-Jørgensen 1989).¹ Fischer-Jørgensen (1989), in an electromyographic study of seven speakers, found that the cricothyroid was more contracted in the first half of stød-bearing rhymes than in the second half, and more contracted at the beginning of stød-bearing rhymes than in stød-less rhymes, where there was no contraction. Fischer-Jørgensen (1989) also reports on a fiberoptic study, which examined the vowels [i:] and [i:ʔ] for six speakers. She found that whilst the position of the vocal folds at the beginning of [i:] and [i:ʔ] was similar, five of the six participants showed transverse contraction of the vocal folds at the end of [i:ʔ], and all speakers showed evidence of ventricular fold constriction, evidenced by decreased distance between the ventricular folds, at the end of [i:ʔ]. However, the degree of constriction varied across speakers. For the one speaker who did not show transverse vocal fold constriction, the ventricular folds were nearly completely covering the vocal folds.

These various articulators have also been implicated in producing different acoustic correlates of stød. Fischer-Jørgensen notes that cricothyroid contraction lengthens the vocal folds, making them thinner and more tense, which she asserts is responsible for raising speakers' F0 in the first part of stød-bearing syllables' rhymes (Fischer-Jørgensen 1989; Woodson et al. 1989; Esling et al. 2019). In contrast, Fischer-Jørgensen notes that the beginning of vocal fold contraction aligns closely in time with the point at which F0 began to decrease abruptly for speakers, and that the point of maximum vocal and ventricular fold contraction corresponds to the F0 and intensity minimums for speakers.

In sum, prior research indicates that stød is articulatorily and acoustically biphasic, consisting of cricothyroid contraction which raises F0 in the first part of the syllable rhyme followed by a stød phase proper with vocal and ventricular fold contraction, leading to low F0, low intensity, and often creaky phonation in the second part of the rhyme. In the next section, I discuss how these findings have been used to inform the most comprehensive theory of stød distribution and timing in Danish, as well as how this theory parallels another proposal for suprasegmental-to-segmental timing in Thai.

1.2. Suprasegmental-to-Segmental Timing

In this section, I discuss two analyses of suprasegmental-to-segmental timing, both of which propose that a suprasegmental phonological feature is timed with respect to morae. The first analysis I discuss is the timing of stød in Danish, and the second analysis is the timing of the falling high-low (HL) tone in Thai. Crucially, these two analyses differ in how morae are formulated, and this difference in moraic formulation predicts different timing relationships between the suprasegmental feature and the syllable. The purpose of investigating two ways of formulating morae comes from the results of recent studies on stød onset timing which do not support the traditional moraic analysis put forth by Basbøll (1985, 2005) (see Section 1.2.1). Therefore, both methods of formulating morae are investigated in the analysis here with respect to stød timing in Danish to determine if an alternative way of conceptualizing morae, as in Thai, may better capture stød timing relative to the syllable than the analysis put forth by Basbøll (2005). This would allow for a moraic interpretation of the data, even if it does not support the timing relationship between the stød phase proper and the syllable that has been previously proposed.

1.2.1. Moraic Hypothesis of Stød Timing

The articulatory and acoustic findings reviewed in Section 1.1 have been used to support the most comprehensive phonological theory of stød distribution in Danish, henceforth referred to as the moraic hypothesis (Basbøll 1985, 2005). Under this hypothesis, stød is the default phonation for monosyllabic words. All stød-bearing, monosyllabic words are bimoraic, and stød is a suprasegmental feature licensed by the second mora in the syllable. In this analysis of Danish syllable structure, the first mora is projected by a vocalic element, and the second mora is projected by the phonological element that satisfies the conditions of stød-basis—either the second half of a long vowel or a sonorant consonant in a V+S sequence (Basbøll 1985, 2005). In contrast to these bimoraic, stød-bearing words, all stød-less monosyllabic words are considered monomoraic, including CVS words without stød in the native or native-like vocabulary. For these words, Basbøll proposes that the sonorant coda is extra-prosodic. Consequently, all monosyllabic words that meet the requirements of stød-basis have stød, the default phonation, unless the sonorant coda is extra-prosodic, and those that do not are categorically excluded from hosting stød. This is shown below in Table 2. Here, only one onset consonant (C), coda obstruent (O), or sonorant (S) is shown for brevity, though clusters are permissible.

With respect to stød timing, the moraic hypothesis explicitly integrates the acoustic findings of Fischer-Jørgensen, positing that the stød phase proper should begin in the middle of a long vowel, or, for syllables with a V+S rhyme, at the onset of the coda sonorant. Concurrently, the moraic hypothesis frames the domain of stød as the second half of a long vowel or a sonorant in a V+S sequence.

Table 2. Distribution of stød-bearing and stød-less syllables in monosyllabic Danish words. ‘—’ indicates illegal syllabic configurations in Standard Danish.

Syllable Type	Stød	Non-Stød
CV	—	[ja] ‘yes’
CVO	—	[hɛsd̥] ‘horse’
CVS	[vɛn [?]] ‘turn (imperative)’	[vɛn] ‘friend’
CV:	[b̥y: [?]] ‘town’	—
CV:O	[lɔ: [?] s] ‘lock’	—
CV:S	[b̥e: [?] n] ‘leg/bone’	—

However, recent experimental work has not corroborated this timing relationship between the stød phase proper and the syllable. In fact, for as long as stød has been studied acoustically, a great deal of variability in the stød phase’s timing has been robustly attested both within and across speakers (Fischer-Jørgensen 1989; Grønnum and Basbøll 2001; Grønnum et al. 2013). Fischer-Jørgensen (1989) explicitly highlights this variability, though she asserts that the onset of the stød phase proper generally falls around the middle of a long vowel or near the onset of a coda sonorant in a V+S rhyme. Grønnum and Basbøll (2001) also found that the onset timing of the stød phase proper did not differ between syllables with long vowels and those with V+S rhymes. However, this was likely due to a high degree of variability in stød onset timing, not because these two points were equidistant from the vowel onset. In fact, a close examination of the data in Grønnum and Basbøll (2001) reveals that the onset of the stød phase proper in syllables with long vowels usually occurs around the middle of the vowel, but that in V+S rhymes, the onset usually occurs in the last third of the vowel, not at the onset of the coda sonorant (see Grønnum and Basbøll 2001, Table 13, p. 247). Furthermore, a confounding factor in this study is that the long stød-bearing vowels occurred in the first syllable of disyllabic words, whereas stød-bearing V+S rhymes occurred in both monosyllabic and disyllabic words.

Also of note is that even early acoustic work has acknowledged the pervasive variability in stød offset timing, though it has never been subject to serious investigation. Fischer-Jørgensen (1989), Grønnum and Basbøll (2001), and Grønnum et al. (2013) note that various correlates of stød, including irregular vocal fold vibration, low F0, and low intensity, can spread from the second half of a long vowel or a sonorant in a V+S rhyme to following segments or even to a post-tonic syllable. These descriptions of spreading imply that the domain of stød is the syllabic configuration that satisfies stød-basis, that is, either the second half of a long vowel or a coda sonorant in V+S rhymes, and frame the presence of stød outside of these intervals as irrelevant to its phonological domain.

Another issue with the moraic hypothesis in Danish is the independent motivation for morae. As stated by Grønnum and Basbøll (2001), morae do not play any role in Danish other than to explain the distribution and timing of stød. For example, poetry in Danish counts syllables, not morae, and stress in Danish is said to be free, or quantity insensitive (Grønnum 1998; Kuznetsova 2018).² This lack of phonological evidence has led researchers to look for other evidence for morae in Danish. For example, Basbøll (1985) predicts that moraic coda sonorants in stød-bearing V+S rhymes will be longer, and thus quantitatively heavier, than non-moraic coda sonorants in stød-less rhymes. Therefore, the coda [n] in words such as [vɛn[?]] ‘turn (imperative)’ are predicted to be longer than the coda [n] in words such as [vɛn] ‘friend’. However, Grønnum and Basbøll (2001) did not find a significant difference in coda sonorant durations between stød-bearing and stød-less words. Follow-up perception experiments conducted by Grønnum and Basbøll (2003a, 2003b) testing the cognitive status of stød-bearing syllables as biphasic also did not yield evidence that the onset of the stød phase proper is perceptually timed to a second mora.

Based on these results, recent work (Grønnum and Basbøll 2007; Grønnum et al. 2013; Grønnum 2015) has described stød as a feature of the syllable without referring to lower-level linguistic structure such as morae. In fact, this work goes so far as to explicitly state that morae in Danish are merely a theoretical convenience with no psychological reality

for Danish speakers. This re-formalization of *stød* as a property of the syllable, however, remains agnostic as to the timing relationship between the syllable and the *stød* phase proper. Furthermore, the dearth of research specifically on the variation in *stød* timing has led Grønnum et al. (2013) to note that, at present, there is no evidence that *stød* realization is systematic.

However, as I discuss below, other languages besides Danish also evoke morae to explain suprasegmental-to-segmental timing relationships. In the next section I review how the coordination of the falling HL tone in Thai parallels *stød* coordination in Danish, and how adopting a different approach to moraic formulation, as proposed by Morén and Zsiga (2006), may allow for the preservation of a moraic analysis of *stød* timing in Danish.

1.2.2. Tonal Coordination in Thai and the Shared Mora Hypothesis

The use of morae as a prosodic mediator between a suprasegmental feature and its host syllable has been used to explain the timing of other suprasegmental phenomena relative to segments in languages besides Danish, namely tone. Here, I discuss one such analysis, that of tonal coordination in Standard Thai (Morén and Zsiga 2006; Karlin and Tilsen 2015), focusing on the distribution and timing of the HL tone, which strongly resembles that of *stød* in Danish.

Thai is a tonal language with five phonological tones, including a falling HL tone, which can only manifest on CVS, CVV, CVVO, and CVVS syllables,³ similar to the distribution of *stød*-bearing monosyllabic words in Danish (see Table 2). Morén and Zsiga (2006), in an acoustic study of tonal coordination in Thai, found that the fall in pitch in an HL tone is timed with respect to the onset of the coda sonorant in monosyllabic CVS words, which they claim is usually commensurate with the middle of the syllable rhyme, or to the middle of the syllable rhyme in CVV, CVVO, and CVVS words. Morén and Zsiga interpret these findings to mean that the L phase of the HL tone is timed to a second mora in a bimoraic syllable, and crucially that the second mora can be shared by multiple segments (Broselow et al. 1997). Thus, the syllable rhyme of CVV, CVVO, and CVVS words is divided in half, with each half corresponding to a mora. For CVV words, this effectively divides the long vowel in half, with each half corresponding to one mora. For CVVO and CVVS words, the syllable rhyme is divided in half, with the first half consisting of only a portion of the long vowel, corresponding to the first mora. In contrast to CVV words, however, in CVVO and CVVS words, the second half of the rhyme consists of the final portion of the long vowel (usually the last third or so) as well as the following coda consonant. Therefore, the second mora in CVVO and CVVS words is always shared by two segments, the latter part of the long vowel and the coda consonant. This accounts for why the L tone coordinates with the middle of the syllable rhyme in these syllables, and not the middle of the long vowel, in contrast to CVV words. This analysis of moraic formulation will henceforth be referred to as the shared mora hypothesis.

Morén and Zsiga motivate this analysis of morae in Thai by pointing out three aspects of their data. First, across the different word types that bear the HL tone, (CVS, CVV, CVVO, CVVS), the syllable rhyme durations are quite similar. Second, coda consonants in CVVO and CVVS words are shorter than in CVO and CVS words in their study, respectively. Third, vowels in CVVO and CVVS words are shorter than in CVV words. Together, Morén and Zsiga interpret these results such that morae in Thai have a preferred duration independent of the syllabic configuration of the syllable, and furthermore, that the phonetic duration of segments is modulated by this specification. This leads to long vowels in CVVO and CVVS words shortening relative to those in CVV words and coda consonants shortening relative to those in CVO and CVS words to fit this preference.

In sum, previous literature on the timing of multiple suprasegmental features relative to their host syllable has proposed that the relationship is mediated by morae, specifically ones with biphasic realizations such as *stød* in Danish and the HL tone in Thai (Basbøll 1985, 2005; Morén and Zsiga 2006). However, the proposals for these two languages differ in how they conceptualize morae. Therefore, given the recent findings that cast doubt on

the timing relationship between the stød phase proper and morae in Danish, in this study, I investigate whether reconceptualizing how morae are formulated provides a fruitful alternative to the traditional formulation of morae in Danish (Basbøll 1985, 2005), allowing a moraic interpretation of the data to be preserved under empirical investigation.

1.3. Research Questions and Predictions

In this section, I outline the research questions of this study and the predictions of three hypotheses of stød timing relative to the syllable. To restrict the domain of inquiry whilst still allowing for the results to be generalizable, here, I investigate stød timing in four types of monosyllabic words, CVS[?], CV:[?], CV:[?]O and CV:[?]S. In the following sections, the term stød will be used to refer to the stød phase proper, unless it is made clear from context that it is referring to both phases of stød (e.g., when discussing the distribution of stød-bearing syllables in the lexicon). The research questions are as follows:

1. Across word types (CVS[?], CV:[?], CV:[?]O and CV:[?]S), which hypothesis of stød timing best captures the relationship between the stød phase proper and the syllable?
2. How can stød timing relative to the syllable be modeled?
3. Is there independent evidence from segmental durations (vowel duration, coda sonorant duration, and overall sonorant rhyme duration) for morae in Danish?

Regarding the first question, three hypotheses of stød onset timing are investigated and compared in this study, as shown in Table 3. To determine which hypothesis best represents the timing relationship between stød and its host syllable, I adopt the approach in Shaw et al. (2011) that measures of relative stability rather than absolute stability are a reliable index of coordination between two gestures. Shaw et al. (2011) implemented this idea in a study on Moroccan Arabic onset cluster gestural coordination. In this study, they measured the temporal distance between various articulatory landmarks of the onset consonant (cluster) to a fixed anchor, the offset of the vowel, to assess which measure induced the least variability across words with different numbers of onset consonants (see Section 1.4 for labels of landmarks). In their results, they found that the temporal distance between the release of the onset (cluster) and the anchor remained relatively stable across words with one to three onset consonants, in contrast to the temporal distance between the target of the first onset consonant gesture or the center of the consonant gestures to the anchor, evidenced by lower relative standard deviations. Following from this, here, I assume that the timing relationship between stød and the syllable which induces the least variability in stød onset timing, both within and across word types, best represents the phonological organization of stød relative to the syllable. This view permits variance in the actual realization of stød, suitable for empirical data, as the stability of each measure is assessed by comparing heuristics of variance (here, standard deviation) and does not require that stød aligns perfectly with a given phonological point.

The first timing hypothesis examined here is the moraic analysis put forth by Basbøll (1985, 2005), which predicts that the onset of stød relative to the syllable will align with the onset of a sonorant coda in words with V+S rhymes (CVS[?] words) and with the middle of the long vowel in CV:[?], CV:[?]O, and CV:[?]S words. This hypothesis also predicts that stød offset will align with the offset of the sonorant coda in CVS[?] words and with the offset of the long vowel in CV:[?], CV:[?]O, and CV:[?]S words, providing evidence that the domain of stød is the sonorant coda in words with V+S rhymes or the second half of a long vowel.

Table 3. Predictions of the moraic hypothesis, shared mora hypothesis, and sonorant rhyme center hypothesis on stød timing across word types. Cells are combined across rows if the hypothesis makes the same predictions about various word types.

Word Type	Moraic Hypothesis	Shared Mora Hypothesis	Sonorant Rhyme Center Hypothesis
CVS [?]	Stød onset = Coda onset Stød offset = Coda offset	Stød onset = Coda onset Stød offset = Coda offset	
CV:?		Stød onset = Vowel midpoint Stød offset = Vowel offset	Stød onset = Sonorant rhyme midpoint
CV:? [?] O	Stød onset = Vowel midpoint Stød offset = Vowel offset		Stød offset = Sonorant rhyme offset
CV:? [?] S		Stød onset = Sonorant rhyme midpoint Stød offset = Sonorant rhyme offset	

A second possibility is that stød aligns with the second mora, but that morae are formulated in Danish via the shared mora hypothesis, as in Thai. The differences in moraic formulation between this hypothesis and the moraic hypothesis are shown below in Figure 1. According to this hypothesis, the first mora in a stød-bearing word is projected by either a short vowel in a V+S rhyme or by part of a long vowel, but the second mora may be shared between multiple segments. Applying this principle to Danish, I follow the analysis in Basbøll (1985, 2005) that only sonorant codas can contribute to a mora, in contrast to obstruent coda consonants. This formulation of morae differs from Thai, in which both coda sonorants and coda obstruents can contribute to a mora. Based on this hypothesis, stød is predicted to begin with the onset of a sonorant coda in CVS[?] words and in the middle of the vowel in CV:?[?] and CV:?[?]O words. For CV:?[?]S words, this hypothesis predicts that stød will align with the middle of the sonorant portion of the syllable rhyme, which includes the sonorant coda, mirroring the timing relationship between the HL tone and the syllable in Thai (Morén and Zsiga 2006). This hypothesis also predicts that across all word types, stød offset will align with the end of the sonorant rhyme. In this analysis, the domain of stød is the sonorant rhyme.

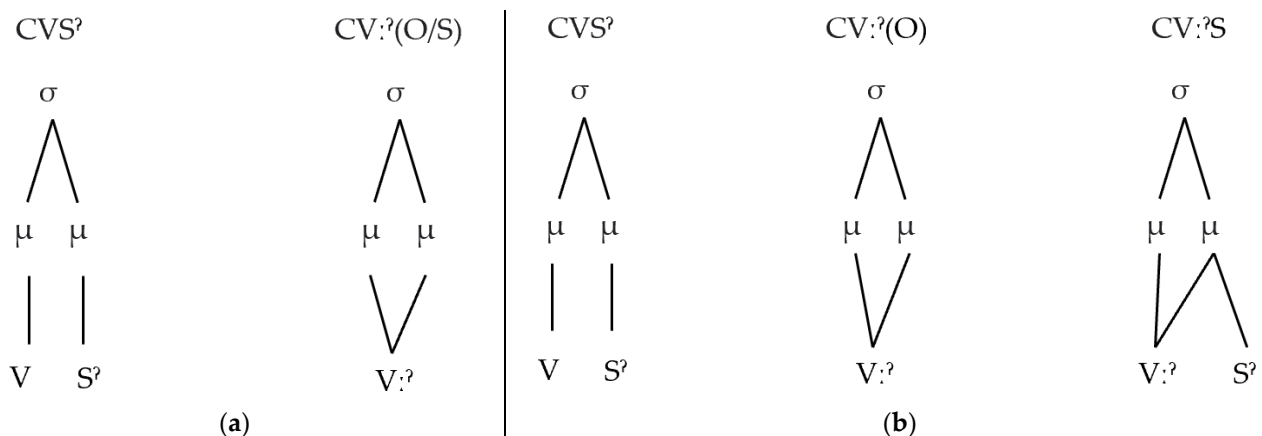


Figure 1. Phonological structure of bimoraic, monosyllabic words in Danish, based on two hypotheses of moraic formulation. (a) The moraic hypothesis, based on morae formulation in Basbøll (1985, 2005). In CVS[?] words, stød is proposed to time to the coda sonorant. In CV:?[?], CV:?[?]O, and CV:?[?]S words, stød is proposed to start in the middle of the long vowel and end at the end of the long vowel. (b) The shared mora hypothesis. In CVS[?] words, stød is proposed to time to the coda sonorant. In CV:?[?] and CV:?[?]O words, stød is proposed to start in the middle of the long vowel and end at the end of the long vowel. For CV:?[?]S words, the second mora is shared by the long vowel and coda sonorant, and stød is proposed to start in the middle of the sonorant rhyme and end at the end of the coda sonorant.

The third hypothesis of stød onset timing investigated here is that stød is a property of the rhyme and not mediated by morae, as stated in Grønnum and Basbøll (2007), Grønnum et al. (2013), and Grønnum (2015). This hypothesis, which I refer to as the sonorant rhyme center hypothesis, is a formalization of the previous claims that stød is a property of the syllable, not of morae. Here, I formalize this hypothesis by assuming that, like the shared mora hypothesis, if stød is a property of the rhyme then it is a property of the sonorant rhyme where it can be acoustically realized, and it is timed relative to the midpoint of this domain. Similar to the shared mora hypothesis, this hypothesis predicts that in CV:?, CV:O and CV:S words, stød onset will align with the midpoint of the sonorant rhyme. In contrast to the shared mora hypothesis, this analysis also predicts that stød onset will align with the midpoint of the sonorant rhyme in CVS? words, rather than the onset of the sonorant coda, and these points need not coincide. Finally, as with the shared mora hypothesis, this hypothesis predicts that stød offset will align with the end of the sonorant rhyme across all word types, and that the domain of stød is the sonorant rhyme. These timing predictions are spelled out in Table 3.

Regarding the final research question, given the similarity between the various hypotheses' predictions on stød timing, it is imperative to investigate independent evidence for morae in Danish. Therefore, durational measures to support the existence of morae will be examined. If morae are formulated via the moraic hypothesis, then we predict that coda sonorants in CVS? words will be longer than those in CVS words, following previous arguments that mora-projecting sonorants are quantifiably "heavier" than non-moraic ones in Danish (Basbøll 2005). If, however, morae are formulated via the shared mora hypothesis, we predict that sonorant rhyme durations across all word types will be similar, vowel durations in CV:S words will be shorter than in CV:? and CV:O words, and coda sonorant durations in CV:S words will be shorter than in CVS? words. These findings would provide evidence that morae have a preferred duration in Danish that phonetically conditions segmental durations, as in Thai.

1.4. Overview of Articulatory Phonology

In the final section of the introduction, I outline the framework of Articulatory Phonology to foreground the discussion of modelling stød timing in terms of gestural coordination (Browman and Goldstein 1986, 1989a, 1989b, 1990, 1992, 2000). An advantage of conceptualizing stød as a laryngeal gesture in an articulatory framework is that its timing relative to the syllable can be represented in ways that are compatible with both moraic and non-moraic analyses of stød timing, allowing for the direct comparison of the hypotheses put forth in Section 1.3. This is because Articulatory Phonology departs from autosegmental phonological theories in defining speech segments (including suprasegmentals) as discrete gestures made by articulators in the vocal tract, which can coordinate with other gestures (Browman and Goldstein 1986, 1989a, 1989b, 1990, 1992, 2000) or to morae (Gao 2008; Karlin and Tilsen 2015). Multiple gestures can also be coupled together into a constellation in order to produce gesturally complex segments (e.g., aspirated stops, nasals), which is ideal for capturing the biphasic nature of stød.

In Articulatory Phonology, gestures are defined by their location of constriction (CL) and the degree of constriction (CD). Five independent articulators are identified in the production of speech sounds, the lips (LIPS), the tongue tip (TT), the tongue body (TB), the velum (VEL), and the glottis (GLO). These articulators can be specified for CL values such as [protruded], [labial], [dental], [alveolar], [palatal], [velar], [uvular], and [pharyngeal]. There are five possible CD values: [closure], [critical], [narrow], [mid], and [wide]. For vowels, these can be combined (e.g., [wide mid]) when needed to indicate height distinctions between vowels. Independent gestures can be represented as timed relative to each other in the production of speech, as shown in the gestural score in Figure 2 for the word [søn] 'son' in Danish. Time is represented on the *x*-axis and the articulators on the *y*-axis. As shown in this figure, the voiceless alveolar fricative [s] is produced by a constellation of two gestures, the [critical] gesture of the tongue tip on the TT tier and the spreading of the

glottis on the GLO tier, as shown by the dotted-and-dashed lines. [œ] is also produced with a constellation of two gestures, the tongue body gesture on the TB tier and the lip protruding gesture on the LIPS tier in solid lines. Finally, the coda [n] is produced with another constellation of two gestures, an alveolar closure involving the tongue tip on the TT tier and a [wide] gesture on the VEL tier, shown in dashed lines.

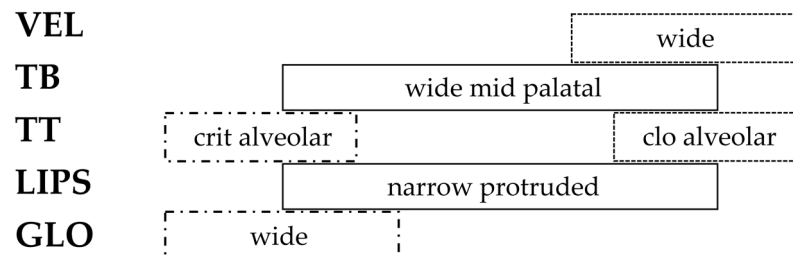


Figure 2. Gestural score for the word [sœn] ‘son’ in Danish.

The gestural score in Figure 2 also shows that, in addition to gestures coupling together into constellations corresponding to a single segment, gestures related to different segments are also coordinated relative to each other in speech (Browman and Goldstein 1990). Traditionally, gestures are described as following a 360° trajectory from beginning to end. Along this trajectory, certain gestural landmarks, which can enter into coordination relationships with other gestures, can be identified (Gafos 2002). The relevant landmarks, as defined in Gafos (2002), are given below in Figure 3.

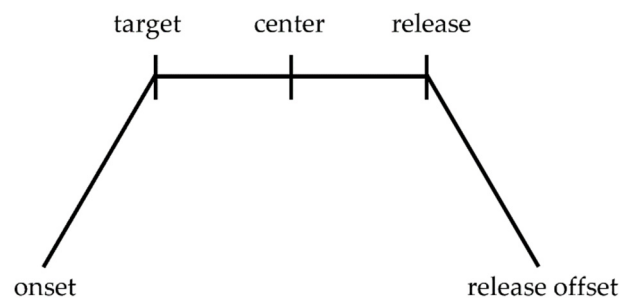


Figure 3. Gestural landmarks, as defined in Gafos (2002).

Using this notation, coordination relationships between adjacent segments can be defined by stating the coordination relationship between the landmark of one gesture and the landmark of another. For example, in a CV sequence in English, the onset of the vowel has been found to coordinate with the center of the consonant (Browman and Goldstein 1990). Furthermore, in some languages, including English (Browman and Goldstein 1988; Marin and Pouplier 2010), German (Pouplier 2012), and Romanian (Marin 2013), the most stable coordination relationship between an onset cluster and the following vowel is between the center landmark of the entire onset cluster and the onset of a vowel gesture. This effect has been termed the c-center effect in previous literature. In contrast, in VC sequences the release of the vowel gesture coordinates with the center of the coda consonant gesture (Browman and Goldstein 1990). Phenomena similar to the c-center effect have not been found between coda clusters and preceding vowels (Browman and Goldstein 1988, 2000; Honorof and Browman 1995). These coordination relationships are shown in Figure 4. I return to the modeling of stød using an articulatory framework in the discussion, in which I show how stød coordination relative to the syllable can be represented gesturally by proposing a new articulatory landmark analogous to the c-center, the sonorant rhyme center.

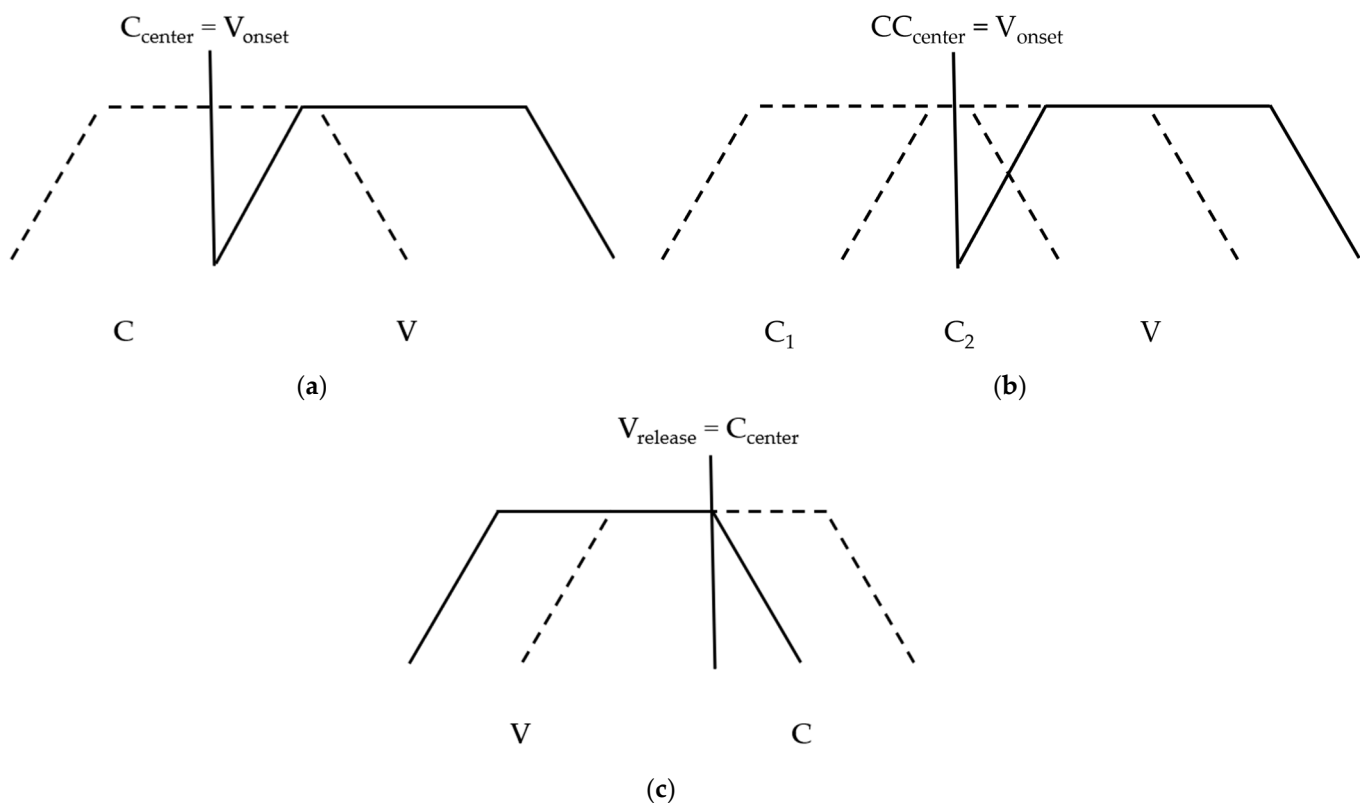


Figure 4. (a) CV coordination in English; (b) CCV (c-center) coordination in English; (c) VC coordination in English.

2. Materials and Methods

2.1. Participants

Nine native Danish speakers were recruited through word of mouth (4M, 5F, ages 25–28, mean age = 26.1). Eight were from Copenhagen or the greater-Copenhagen area and one (a male) was from Greve, a municipality southwest of Copenhagen, whose accent was judged to be similar to the Copenhagen accent by two Danish consultants.⁴ All speakers were living in Copenhagen at the time of recording and had lived in Copenhagen for at least three years. All the speakers were bilingual Danish–English speakers whose primary language was Danish, and only one speaker had lived outside of Denmark for longer than six months.⁵ The same speaker also had issues with their hearing at a young age but has since undergone a corrective procedure and has passed the national military hearing exam. The other participants did not report any hearing or speech disorders.

2.2. Stimuli

The stimuli consisted of one type of stød-less word, CVS (29 words), and of four types of stød-bearing monosyllabic words, CVS[?] (63 words), CV:[?] (21 words), CV:[?]O (40 words), and CV:[?]S (55 words). For CVS, CVS[?], and CV:[?]S words, the manner of articulation of the coda sonorant was varied between a nasal ([n] or [m]) or an approximant ([l]).⁶ Due to the relative infrequency of some types of monosyllabic words (e.g., there are fewer CV:[?] words in the lexicon than CVS[?] words) and segmentability concerns, it was not possible to elicit the same number of target words per word type, and some words were repeated in the task. For the full list of target words, see the Supplementary Materials.⁷

In sum, 179 target words with stød and 29 CVS words without stød were elicited per speaker, for a total of 208 words per speaker and 1872 tokens across speakers. Of these, all tokens for which the vowel could not be segmented from the following tautosyllabic sonorant were excluded (96 tokens). Tokens were also excluded if stød was masked by phrasal creak (55 tokens), if the word began with a period creak on a word-initial vowel (e.g., [ʔ:ʔl], ‘eel’, 44 tokens), or if the voice quality could not be categorized (3 tokens). Five additional words were excluded for all speakers, on the criteria that they were unknown to at least four of the speakers (45 words total). Individual tokens were excluded if a hesitation or mispronunciation occurred on the target word, the preceding word, or the following word, if the speaker was unfamiliar with the word, or if they placed excessive emphasis on the word, evidenced by noticeably louder intensity and a following pause (69 tokens) or if they were realized with a full glottal stop (100 tokens). After these exclusions, 1460 tokens remained. Of these, the remaining CVS[?], CV:[?], CV:[?]O, and CV:[?]S words were analyzed for stød presence. Words with only an F0 dip (76) were excluded because, in the absence of creaky phonation, it was unclear whether the F0 dip was due to regular sentence intonation or due to stød. Finally, words for which stød could not be reliably segmented (172) were also excluded, leaving 1212 words in the final dataset, 1014 of which had stød. The distribution of these words across word types is given below in Table 4.

Table 4. Stimuli by word type. The total number of tokens per word type included in the final study is given in parentheses below the example (total = 1212).

CVS	CVS [?]	CV: [?]	CV: [?] O	CV: [?] S
[gʊl] ‘gold’ (198)	[sɔ̌al [?]] ‘stable’ (331)	[hi: [?]] ‘hibernation’ (101)	[gʊ: [?] s] ‘goose’ (284)	[gɛ: [?] n] ‘gene’ (298)

2.3. Procedure

Speakers were recorded in a quiet room in Copenhagen, Denmark, using the Tascam-DR 40 recorder and a head-mounted Shure WH20 microphone at a sampling rate of 44.1 kHz. Speakers first filled out a questionnaire about their language background before reading blocks of ten sentences with embedded target words on a Microsoft PowerPoint slide. Each target word was embedded into a semantically meaningful sentence. If a target word was elicited twice, it was embedded in two unique sentences. To examine how often stød spread beyond the end of the target word, all CVS[?], CV:[?], and CV:[?]S words were followed by a /j/-initial word with a stød-less first syllable. /j/ was chosen as the initial consonant to allow the coda sonorant in CVS[?] and CV:[?]S words to vary between a nasal and an approximant whilst still being segmentable from the following word. These words varied in length from one syllable (*jeg* [jaj] ‘I’) to six syllables (*jernbaneselskabet* [ˈjæŋnɐːnəsɛlsqæ:[?]ɸɔ̌], ‘the railroad company’). The sentences were randomized into three unique orders, and every five blocks a screen appeared instructing the participant to take a break, though they were informed that they could take a break at any time. Speakers were instructed to read the sentences as if talking to a native Danish speaker, but no instructions about speech style or speech rate were given explicitly. If there was a disfluency on the target word or a surrounding word, speakers were asked to re-record the sentence. After the sentence reading task, speakers also read a word list containing the same target words, but these data are not reported here. In total, the study took approximately two hours.

2.4. Segmentation and Acoustic Analysis

2.4.1. Target Word Segmentation

Target words were segmented from the surrounding speech for analysis using the acoustic analysis software Praat (Boersma and Weenink 2021). For CVS[?], CV:[?], and CV:[?]S words, the following word was also segmented to examine the extent of stød spreading. For each target word, the sonorant rhyme, consisting of either a vowel (CV:[?] and CV:[?]O words)

or a vowel and a sonorant consonant (CVS, CVS', and CV:ʔS words) was then segmented from the syllable onset and any coda obstruents, if present. The onset of the vowel was identified as the onset of clear first and second formants (F1 and F2), characteristic of a vowel, and periodicity in the waveform, following Garellek (2012) and Styler (2017). This procedure was uniform across all words.

For CV:ʔO words ending in obstruents, the offset of the vowel was defined as the offset of F1 and F2, corresponding to the closure of a coda stop or the onset of aperiodic noise of a coda fricative. Stød-bearing words with coda sonorants (CVS' and CV:ʔS) were always followed by a /j/-initial word, characterized by a low F1 and a high F2 and F3. For CVS' and CV:ʔS words ending in nasals, the offset of the nasal was defined using both the spectrogram and the waveform to identify the point at which the waveform shape changed from relatively simple, characteristic of a nasal (Styler 2017), to more complex, characteristic of /j/, and F2 became visible in the spectrogram, indicating the beginning of [j] in the following word. For CVS' and CV:ʔS words ending in [l], also characterized by a low F1 and high F2, though not as high as the F2 of [j], the offset of liquid was defined as point at which F2 began to rise usually from a relatively high plateau to the onset of [j]. CVS words were followed by a variety of words, usually stop-initial. For these words, the offset of the nasal or [l] was taken to mark the end of the word. Finally, for CV:ʔ target words, the formant patterns of the long vowel impacted the formant transitions from the end of the target word to the onset of the following /j/. For CV:ʔ words with the front unrounded vowel [i:ʔ], F2 remained high throughout the vowel and the following [j]. However, F3 was observed to rise towards F4 before lowering at the onset of [j]. Concurrently, the vowel offset was defined as the point at which F3 and F4 diverged from each other, as shown in Figure 5. CV:ʔ words with back vowels and front rounded vowels exhibited a rise in F2 towards F3 at the end of the target word, indicating the onset of the following /j/. For these words, the offset of the vowel was defined as the point at which F2 and F3 began to diverge, signifying the onset of [j].

For words with complex rhymes (CVS, CVS' and CV:ʔS), the vowel and coda sonorant were also segmented from each other. When the coda sonorant was a nasal, the offset of the vowel/onset of the nasal was defined as the point at which F2 disappeared from the spectrogram, anti-formants characteristic of nasals appeared in the spectrogram, and the amplitude of the waveform decreased (Styler 2017). When the sonorant coda was [l], the formant transitions from vowel offset to the onset of the following [l] differed as a function of vowel backness and roundedness. First, the front unrounded vowels [i:], e:, ε:] and [a:]⁸ were produced with a very high F2, even greater than F2 for [l]. Concurrently, the boundary between the vowel and [l] was placed where F2 dipped from the offset of the vowel towards the onset of [l]. For words with the front rounded vowels [y:] and [ø:], which had lower F2 values than [l], the point at which F2 began to rise was selected as the onset of [l]. For the back rounded vowels [u:] and [ɔ:], the boundary was placed where F2 rose and came together with F3 in a pinch, indicating the onset of [l]. Finally, for the back and central unrounded vowels [ʌ] and [a], F1 was higher than the F1 of [l], and the boundary was placed where F1 dipped, indicating the onset of [l]. Examples are given in Figure 5.

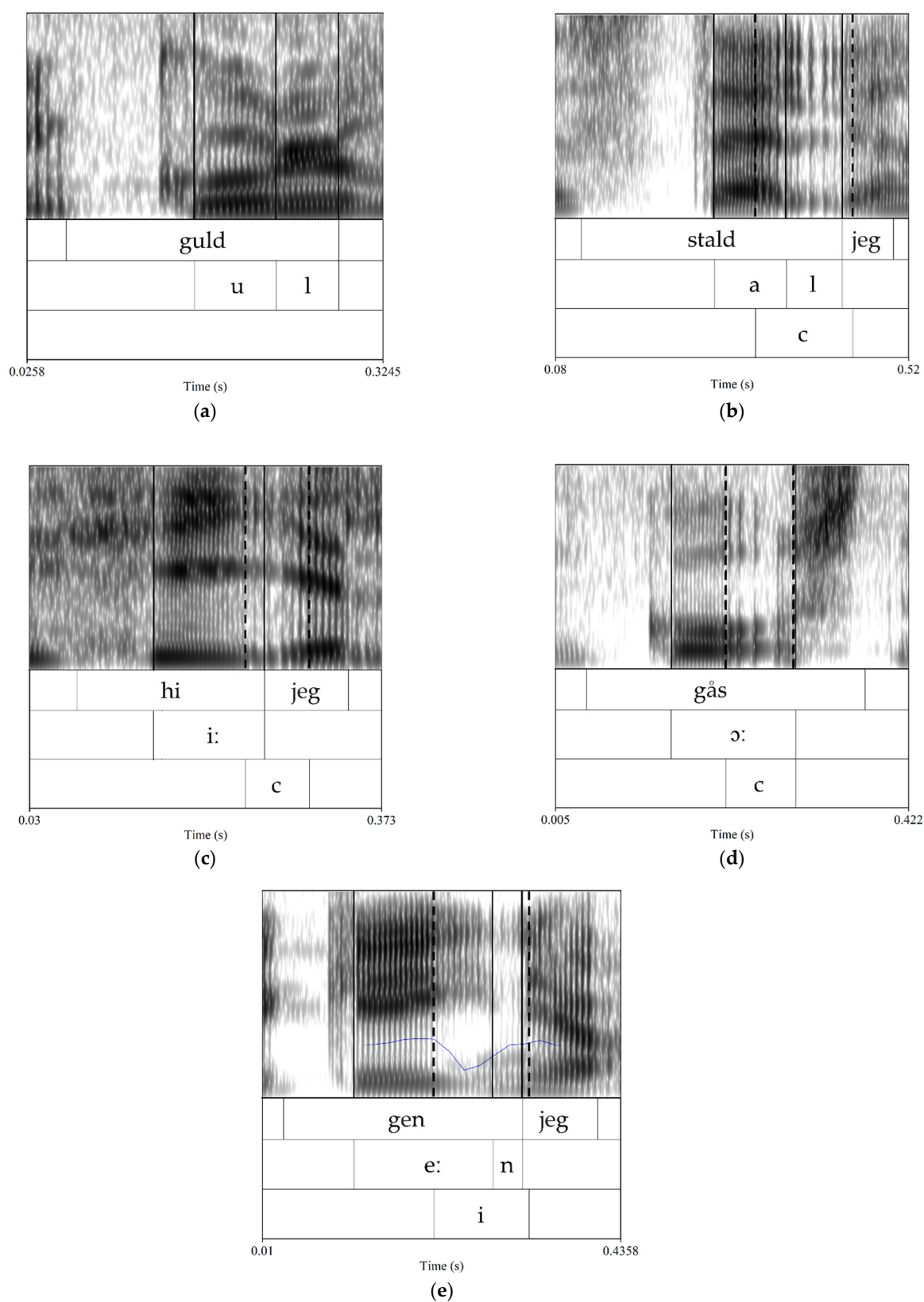


Figure 5. Segmental (solid lines) and voice quality (dashed lines) segmentations for CVS, CVSʔ, CV:ʔ, CV:ʔO, and CV:ʔS words. For CVSʔ, CV:ʔ, and CV:ʔS words the following word is also segmented. Phonation is marked on the third tier, where ‘c’ indicates intervals of creaky phonation and ‘i’ indicates intervals of low intensity. (a) CVS word [gʊl] ‘gold (n)’; (b) CVSʔ word [sɔ̌alʲ] ‘stable (n)’; (c) CV:ʔ word [hi:ʔ] ‘hibernation’; (d) CV:ʔO word [gʊ:ʔs] ‘goose’; (e) CV:ʔS word [gʊ:ʔn] ‘gene.’

2.4.2. Phonation Segmentation and Acoustic Analysis

First, to verify the biphasic realization of *stød* that has been previously reported, F0 measures were taken every millisecond over the course of the vowel for all word types as well as over the course of the sonorant coda for CVS, CVS[?], and CV:ʔS words using the straight algorithm in VoiceSauce (Shue et al. 2011). The data were then averaged over fifths of the vowel or sonorant, respectively, using a custom Python script to obtain the average F0 per fifth of the vowel/sonorant across word types.

Regarding the *stød* phase proper, *stød* can manifest with a variety of acoustic correlates, including low F0, decreased intensity, creaky phonation, and full glottal closure, whilst in other cases, it can be difficult to detect visually. Therefore, each target word was marked for multiple acoustic cues of *stød*, though in this analysis, only tokens with a visible period of creaky phonation in the spectrogram or a period of decreased intensity and F0 were included in the final analysis. Creak onset and offset were visually identified as the onset/offset of irregular amplitude in the waveform accompanied by irregular glottal pulses in the spectrogram and often perturbations in the pitch tracker (Keating et al. 2015). Tokens with low intensity and F0 but no visible creak (22 tokens) were included if there was an observable period of low intensity that could not be attributed to a coda consonant, evidenced by lightening of the spectrogram, a localized dip in pitch of at least 20 Hz rounded to the nearest hertz, and often significant perturbations in the formant tracker. Onset boundaries were placed where the formants began to lighten, and offset boundaries were placed where intensity began to increase, evidenced by darkening of the formants in the spectrogram. Examples are given below in Figure 5, with the target word written orthographically on the first tier, vowels and any sonorant consonants transcribed in IPA on the second tier, and creak (c) and intervals of low intensity (i) marked on the third tier. When two words are transcribed on tier 1, the first word is the target word.

3. Results

The format of this section is as follows. In Section 3.1, the results of the F0 analyses are presented. Section 3.2 consists of the *stød* onset analysis, which compares three measures of *stød* onset timing based on the moraic hypothesis, the shared mora hypothesis, and the sonorant rhyme center hypothesis. *Stød* offset timing is analyzed in Section 3.3, and segmental durations (vowel, coda sonorant consonant, and overall rhyme duration, respectively) are analyzed in Section 3.4. For each acoustic measure, linear mixed-effects models were run in *R* using the *lmer()* function in the *lme4* package (Bates et al. 2015)⁹, and post hoc Tukey tests (Bretz et al. 2010) were used to confirm statistical significance at the 0.05 level between levels of the categorical fixed effects, except for the F0 analyses, for which the 0.01 level was used due to the large number of tests that were performed, following Garellek (2012). For all the durational models and the *stød* onset/offset models, a random slope for word type by speaker and random intercepts for speaker and token were included. For the F0 models and the SD model, random intercepts for speaker and token were included.

3.1. F0 Analyses

The purpose of the F0 analyses is twofold. The first analysis is to verify that F0 at the beginning of *stød*-bearing rhymes is higher than for *stød*-less rhymes, confirming previous findings that the first *stød* phase is characterized by modal phonation and high F0. The second analysis is to determine when in the rhyme F0 decreases, in order to shed light on the timing relationship between the first and second *stød* phases.

In the first F0 analysis, the average F0 for each fifth of the vowel was calculated across all speakers.¹⁰ Five linear mixed-effects regression models were then run, one for each fifth of the vowel, with F0 as the dependent variable and word type (CVS[?], CV:ʔ, CV:ʔO and CV:ʔS) as a categorical fixed effect.¹¹ In each model, the average F0 of CVS words, the only type of word without *stød*, was the baseline. Due to the large number of tests, *p* values under 0.01 were considered significant, following Garellek (2012). The results of the models

are shown below in Table 5. ‘↑’ indicates that the word type has a significantly higher average F0 than CVS words for the given interval, and ‘↓’ indicates that the word type has a significantly lower average F0 than CVS words. Dashes indicate no significant difference.

Table 5. F0 differences between stød-bearing words and CVS words per fifth of the vowel. ↑ indicates that the word type with stød has significantly higher F0 than CVS words during that time interval. ↓ indicates that the word type with stød has significantly lower F0 than CVS words during that time interval. Dashes indicate no significant difference.

F0 at Each Vowel Fifth (Compared to CVS)					
	1	2	3	4	5
CVS [?]	↑	↑	↑	—	—
CV:?	↑	↑	↑	—	—
CV:O	—	—	—	↓	—
CV:S	↑	↑	↑	—	↓

As shown in Table 5, CVS[?], CV:?, and CV:S words have a significantly higher average F0 than CVS words in the first 3/5 of the vowel, indicated by the upwards arrows (first fifth: CVS–CVS[?]: $\beta = 12.065$, $p < 0.001$; CVS–CV:?: $\beta = 24.685$, $p < 0.001$; CVS–CV:S: $\beta = 18.091$, $p < 0.001$; second fifth: CVS–CVS[?]: $\beta = 13.527$, $p < 0.001$; CVS–CV:?: $\beta = 26.364$, $p < 0.001$; CVS–CV:S: $\beta = 18.886$, $p < 0.001$; third fifth: CVS–CVS[?]: $\beta = 14.404$, $p < 0.001$; CVS–CV:?: $\beta = 21.578$, $p < 0.001$; CVS–CV:S: $\beta = 14.711$, $p < 0.001$). In contrast, the average F0 for CV:O words does not differ from CVS words in the first 3/5 of the vowel (first fifth: $\beta = 0.038$, $p = 0.99$; second fifth: $\beta = 2.160$, $p = 0.56$; third fifth: $\beta = -4.971$, $p = 0.20$). During the 4/5 interval, CVS[?] and CV:?, and CV:S words do not differ from CVS words, and CV:O words have a significantly lower average F0 than CVS words (CVS–CVS[?]: $\beta = 9.651$, $p = 0.03$; CVS–CV:?: $\beta = 4.46$, $p = 0.45$; CVS–CV:O: $\beta = -14.864$, $p < 0.01$; CVS–CV:S: $\beta = 1.384$, $p = 0.75$). Finally, in the last fifth of the vowel, only CV:S words differ from CVS words by having a significantly lower F0 (CVS–CVS[?]: $\beta = -1.438$, $p = 0.77$; CVS–CV:?: $\beta = -2.446$, $p = 0.72$; CVS–CV:O: $\beta = -8.885$, $p = 0.08$; CVS–CV:S: $\beta = -20.687$, $p < 0.001$). These results indicate that the average F0 for stød-bearing words is significantly higher than for stød-less words in the first part of the syllable rhyme, except for CV:O words. I will return to why this may be in the discussion.

In the second F0 analysis, to investigate the timing relationship between the first and second stød phases, the change in F0 over the course of the rhyme per word type was examined. To do so, the average F0 for each fifth of the vowel was calculated for each word type, as well as the average F0 per fifth of the coda sonorant for CVS, CVS[?], and CV:S words. Linear mixed-effects regression models were run for each word type with the average F0 per fifth of the interval, vowel, or sonorant, as the dependent variable and the time interval (first fifth, second fifth, etc.) as a categorical fixed effect. In these models the fixed effect of time interval was sum-coded. Significant effects thus represent intervals for which F0 differed significantly from the overall average F0 across all intervals for that word type. Post hoc Tukey tests were used to establish significant differences between adjacent intervals. The results of the models are shown below in Table 6, with p values under 0.01 considered significant, following Garellek (2012). For each comparison, ‘↑’ indicates that the average F0 of that interval was significantly higher than the overall average F0 of that word type, and ‘↓’ indicates that the average F0 of that interval was significantly lower. Parentheses indicate marginally significant effects (significant at the 0.05 level only), and dashes indicate no significant difference between the interval’s F0 average and the word type’s overall F0 average. For CV:O and CV:O words, the sonorant coda cells are greyed out, indicating the absence of a coda sonorant.

Table 6. F0 differences from the overall F0 average per word type. ↑ indicates that the average F0 for that interval is significantly higher than the overall average F0 of that word type. ↓ indicates that the average F0 for that interval is significantly lower than the overall average F0 of that word type. Arrows in parentheses indicate marginally significant effects (significant at the 0.05 level), and dashes indicate no significant difference. For CV:ʔ and CV:ʔO words, the sonorant coda cells are greyed out, indicating the absence of a coda sonorant.

F0 Differences from the Overall Average F0 per Word Type										
	Vowel					Sonorant				
	1	2	3	4	5	1	2	3	4	5
CVS	—	—	(↓)	↓	↓	—	—	↑	↑	—
CVSʔ	↑	↑	↑	↑	—	(↓)	↓	↓	↓	↓
CV:ʔ	↑	↑	(↑)	↓	↓					
CV:ʔO	↑	↑	—	↓	↓					
CV:ʔS	↑	↑	↑	↑	↓					

As Table 6 shows, over the course of the rhyme, the average F0 for stød-less CVS words follows a different trajectory than the average F0 for stød-bearing words. For CVS words, F0 during the first 2/5 of the vowel is not significantly different from the overall average F0 (first fifth: $\beta = 2.079$, $p = 0.10$; second fifth: $\beta = -1.251$, $p = 0.32$). F0 during the middle of the vowel is marginally lower than the overall average, but this effect did not reach the 0.01 threshold of significance (third fifth: $\beta = -2.882$, $p = 0.02$). Finally, F0 during the 4/5 and 5/5 vowel intervals is significantly lower than the overall average F0 (fourth fifth: $\beta = -3.418$, $p < 0.01$; fifth fifth: $\beta = -5.141$, $p < 0.001$) but significantly higher than the overall average during the 3/5 and 4/5 sonorant intervals (third fifth: $\beta = 4.086$, $p < 0.01$; fourth fifth: $\beta = -4.59$, $p < 0.001$). Post hoc Tukey tests comparing each interval's average F0 to the preceding interval's average F0 revealed no significant differences between adjacent intervals. This indicates that the change in F0 over time for CVS words was more gradual and less extreme compared to stød-bearing words, as shown in Figure 6. No other effects for CVS words reached significance.

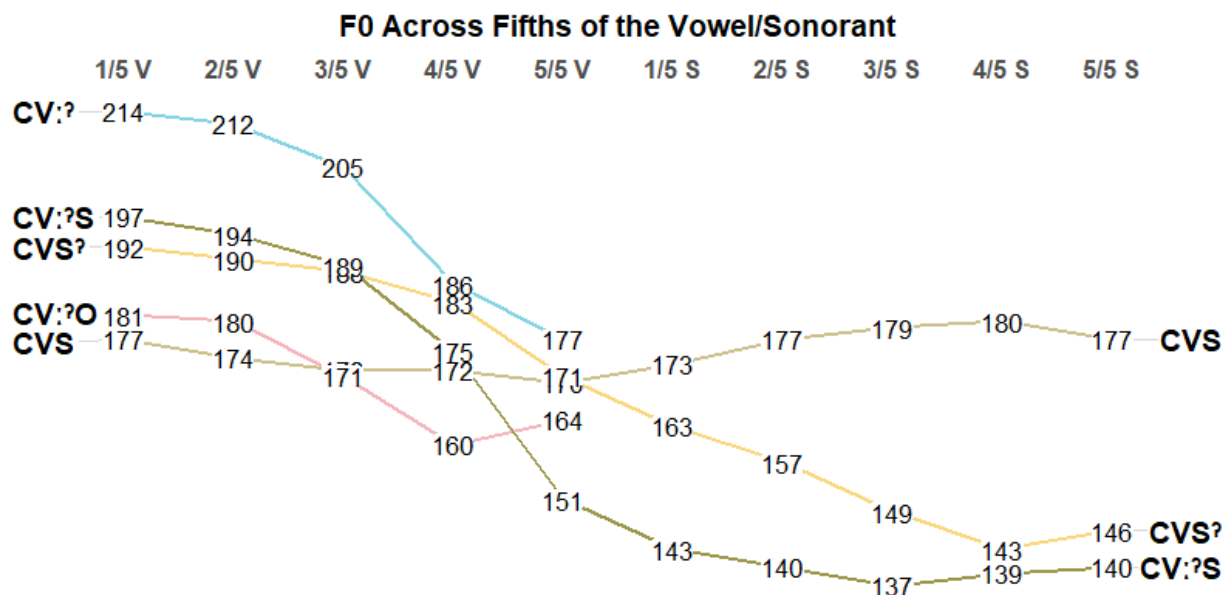


Figure 6. Average F0 in Hz across fifths of the vowel/sonorant by word type. 'V' indicates intervals of the vowel and 'S' indicates intervals of the sonorant. Numbers represent the average F0 in Hz per interval. For CV:ʔ and CV:ʔO words without sonorant coda consonants, the F0 contours end at the 5/5 V mark.

In contrast to stød-less CVS words, the average F0 during the first 2/5 of the vowel in CV:ʔ and CV:ʔO words is significantly higher than each word types' average F0 (first fifth: CV:ʔ: $\beta = 15.457, p < 0.001$; CV:ʔO: $\beta = 9.933, p < 0.001$; second fifth: CV:ʔ: $\beta = 13.306, p < 0.001$; CV:ʔO: $\beta = 8.555, p < 0.001$). F0 is also marginally higher than the average F0 during the middle of the vowel in CV:ʔ words, but this effect did not reach the 0.01 threshold of significance (third fifth: $\beta = 6.011, p = 0.03$). For CV:ʔ and CV:ʔO words, F0 is significantly lower than the average during the last 2/5 of the vowel (fourth fifth: CV:ʔ: $\beta = -12.891, p < 0.001$; CV:ʔO: $\beta = -11.057, p < 0.001$; fifth fifth: CV:ʔ: $\beta = -21.884, p < 0.001$; CV:ʔO: $\beta = -6.796, p < 0.001$). Post hoc Tukey tests comparing each interval's average F0 to the preceding interval's average F0 revealed a significant F0 decrease from the 3/5 to the 4/5 of the vowel in CV:ʔ words ($\beta = -18.902, p < 0.001$). For CV:ʔO words, post hoc Tukey tests revealed a marginally significant decreased in F0 from the 2/5 interval to the 3/5 of the vowel, and a significant decrease in F0 from the 3/5 to the 4/5 of the vowel (second fifth–third fifth: $\beta = -9.189, p = 0.03$; third fifth–fourth fifth: $\beta = -10.423, p < 0.01$).

Similarly, the average F0 during the first 4/5 of the vowel in CVSʔ and CV:ʔS words is significantly higher than each word types' average F0 (first fifth: CVSʔ: $\beta = 23.776, p < 0.001$; CV:ʔS: $\beta = 36.486, p < 0.001$; second fifth: CVSʔ: $\beta = 21.625, p < 0.001$; CV:ʔS: $\beta = 33.912, p < 0.001$; third fifth: CVSʔ: $\beta = 20.335, p < 0.001$; CV:ʔS: $\beta = 28.071, p < 0.001$; fourth fifth: CVSʔ: $\beta = 14.642, p < 0.001$; CV:ʔS: $\beta = 14.091, p < 0.001$). For CVSʔ words, F0 is marginally lower than the average during the first fifth of the sonorant and significantly lower than the average during every following interval (first fifth: $\beta = -5.089, p = 0.04$; second fifth: $\beta = -10.761, p < 0.001$; third fifth: $\beta = -19.471, p < 0.001$; fourth fifth: $\beta = -25.027, p < 0.001$; fourth fifth: $\beta = -22.463, p < 0.001$). Post hoc Tukey tests revealed a marginally significant decrease in F0 from the 4/5 vowel interval to the 5/5 interval, but this effect did not reach the 0.01 threshold of significance (fourth fifth–fifth fifth: $\beta = -12.209, p = 0.03$). For CV:ʔS words, F0 during the final fifth of the vowel is significantly lower than the overall average F0 ($\beta = -9.321, p < 0.001$) as well as during every interval of the sonorant coda (first fifth: $\beta = -17.929, p < 0.001$; second fifth: $\beta = -20.259, p < 0.001$; third fifth: $\beta = -23.311, p < 0.001$; fourth fifth: $\beta = -21.509, p < 0.001$; fifth fifth: $\beta = -20.231, p < 0.001$). Post hoc testing further revealed a significant decrease in F0 from the 3/5 vowel interval to the 4/5 interval and from the 4/5 vowel interval to the 5/5 interval (third fifth–fourth fifth: $\beta = -13.979, p < 0.01$; fourth fifth–fifth fifth: $\beta = -23.413, p < 0.001$). No other main effects or post hoc comparisons between adjacent time intervals were significant.

Overall, the second F0 analysis reveals F0 patterns across stød-bearing word types consistent with previous descriptions of stød, which differ from the F0 trajectory in stød-less CVS words. This is shown in Figure 6. For CVS words, F0 at the beginning of the word does not differ significantly from the average. F0 then decreases towards the end of the vowel, before increasing towards the end of the coda sonorant. In contrast, for stød-bearing words, F0 begins high before decreasing significantly near the center of the words' sonorant rhyme. For CV:ʔ and CV:ʔO words in this study, F0 is significantly higher than the average during the first 2/5 of the vowel and significantly lower during the final 2/5. F0 decreased significantly after the middle of the vowel in CV:ʔ words, whereas for CV:ʔO words, F0 began decreasing slightly earlier, after the first 2/5 of the vowel, though this effect was marginal, and continued decreasing during the middle of the vowel. For CVSʔ words, F0 is high during the first 4/5 of the vowel and low during the last 4/5 of the sonorant. In this study, vowels in CVSʔ words composed approximately 53% of the sonorant rhyme. The sonorant rhyme center therefore usually falls within the last fifth of the vowel. Concurrently, post hoc testing revealed a marginally significant decrease in F0 from the 4/5 to the 5/5 of the vowel. Finally, for CV:ʔS words, F0 is high for the first 4/5 of the vowel and low during the final fifth as well as during the entire coda sonorant. In this study, the vowel in CV:ʔS words composed approximately 69% of the sonorant rhyme with an average duration of 124 ms (mean sonorant rhyme duration = 180 ms). Therefore, the sonorant rhyme center usually falls within 4/5 interval of the vowel. Concurrently, post hoc tests revealed a significant F0 decrease from the 3/5 to the 4/5 and from the 4/5 to the

5/5 of the vowel. These results are shown in Figure 6. Average sonorant rhyme durations, vowel durations, and coda sonorant durations are given in Table 7.

Table 7. Average rhyme, vowel, and sonorant durations in milliseconds per word type. Greyed out cells indicate the lack of a sonorant coda for CV:ʔ and CV:ʔO words.

Word Type	Sonorant Rhyme Duration	Vowel Duration	Coda Sonorant Duration
CVS	139 ms	70 ms	70 ms
CVSʔ	167 ms	90 ms	77 ms
CV:ʔ	131 ms	131 ms	
CV:ʔO	145 ms	145 ms	
CVS	180 ms	124 ms	56 ms

Together, the two F0 analyses point towards the conclusion that stød is acoustically biphasic, with high F0 relative to stød-less CVS words in the first part of the rhyme and a significant decrease in F0 beginning around the middle of the sonorant rhyme. These results provide evidence for analyses of stød timing such that across word types, the second stød phase is related to the second half of the sonorant rhyme, regardless of word type. These results will be discussed further in the Section 4.

3.2. Stød Onset Timing

3.2.1. Comparing Different Hypotheses of Stød Onset Timing

In this section, only words with stød were analyzed ($n = 1014$). To analyze stød onset timing, relativized measures were adopted. By expressing stød onset timing as a proportion of the relevant phonological domain (the vowel, coda sonorant, or sonorant rhyme), the measures normalize stød onset timing across CV:ʔ, CV:ʔO, and CV:ʔS words, for which differences in phonetic vowel duration were found (see Section 3.4.1). To compare the timing predictions of the moraic hypothesis, the shared mora hypothesis, and the sonorant rhyme center hypothesis, three measures of stød onset timing were calculated, and the correspondence between measure and hypothesis is given in Table 8.

Table 8. Predictions made by the moraic hypothesis, shared mora hypothesis, and sonorant rhyme center hypothesis per word type. In each cell, the measure that is proposed to induce the least variability in stød onset timing, according to the stated hypothesis, is given for each word type.

Word Type	Moraic Hypothesis	Shared Mora Hypothesis	Sonorant Rhyme Center Hypothesis
CVSʔ	Coda sonorant measure	Coda sonorant measure	
CV:ʔ		Vowel midpoint measure	
CV:ʔO	Vowel midpoint measure		Sonorant rhyme center measure
CV:ʔS		Sonorant rhyme center measure	

The first measure, henceforth the vowel midpoint measure, is a direct reflection of the timing predictions made by the moraic hypothesis for all words with long vowels and by the shared mora hypothesis for CV:ʔ and CV:ʔO words. For this measure, the stød onset timing is expressed as the difference between stød onset and the vowel midpoint as a percentage of the vowel's duration. Using this measure, the moraic hypothesis predicts that all words with long vowels will have stød onset times that cluster around 0%, indicating that stød onset is timed to the middle of the long vowel. The shared mora hypothesis

predicts that CV:ʔ and CV:ʔO words will have stød onset times of approximately 0%, whilst CV:ʔS words will have a later onset time.

$$\text{Vowel midpoint measure} = ((\text{stød onset} - \text{vowel midpoint}) / \text{vowel duration}) \times 100$$

The second measure, henceforth the coda sonorant measure, represents the predicted timing relationship between stød onset and the syllable in CVSʔ words by both the moraic hypothesis and the shared mora hypothesis. For this measure, stød onset is calculated as the difference between stød onset and the sonorant coda onset, expressed as a percentage of the coda sonorant duration. Both the moraic hypothesis and the shared mora hypothesis predict that CVSʔ words will have stød onset times that cluster around 0%, indicating that stød onset is timed to the onset of the sonorant coda.

$$\text{Coda sonorant measure} = ((\text{stød onset} - \text{sonorant coda onset}) / \text{sonorant coda duration}) \times 100$$

The last measure, the sonorant rhyme center measure, represents the predicted timing relationship between stød onset and the sonorant rhyme center as a percentage of the sonorant rhyme. This measure quantifies the predictions made by the shared mora hypothesis for CV:ʔS words, as well as those for all word types made by the sonorant rhyme center hypothesis. The shared mora hypothesis predicts that only CV:ʔS words will have stød onset times that cluster around 0%, indicating that stød onset is timed to the center of the sonorant rhyme. In contrast, the sonorant rhyme center hypothesis predicts that all word types will have stød onset times that cluster around 0%.

$$\text{Sonorant rhyme center measure} = ((\text{stød onset} - \text{sonorant rhyme midpoint}) / \text{sonorant rhyme duration}) \times 100$$

As stated in the introduction, variation in production is expected. It is therefore not the absolute timing of stød per utterance that matters, but rather the measure that minimizes variability across the word types. That is, like the c-center effect that has been found for vowel-onset cluster coordination, the measure of stød onset timing that induces the least amount of variability best represents how stød is timed relative to the syllable (Shaw et al. 2011). Here, I adopt standard deviations (SD) as my index of stability. For each word type, the mean and SD of each measure by hypothesis is given below in Table 9. These means and SDs are based on the formula specified in Table 8. For CVSʔ and CV:ʔS words, the measure which induces the least variability is bolded. For CV:ʔ and CV:ʔO words, the means and SDs predicted by all hypotheses are the same.

Table 9. Means and SDs for measures of stød onset timing by hypothesis and word type, as a percentage of the relevant phonological domain. For CVSʔ and CV:ʔS words, the measure which induces less variability is bolded, and columns are combined when multiple hypotheses predict the same timing relationship.

Word Type	Moraic Hypothesis	Shared Mora Hypothesis	Sonorant Rhyme Center Hypothesis
CVSʔ	$\mu = -12.81, SD = 36.85$		$\mu = -0.50, SD = 15.33$
CV:ʔ	$\mu = 9.08, SD = 18.16$		
CV:ʔO	$\mu = -4.36, SD = 16.82$		
CV:ʔS	$\mu = 19.50, SD = 22.03$	$\mu = -2.51, SD = 14.45$	

As shown in Table 9, for CVSʔ and CV:ʔS words, the sonorant rhyme center measure (bolded) reduces variability compared to the alternative moraic measure. To determine if this difference is significant, a linear mixed-effects model was run for CVSʔ and CV:ʔS words with speakers' SDs per measure type as the dependent variable and word type (CVSʔ versus CV:ʔS words) and stød measure type (coda sonorant measure for CVSʔ words and the vowel midpoint measure for CV:ʔS words, collectively referred to as the moraic

measure, versus the sonorant rhyme center measure for both word types) as categorical fixed effects. The sonorant rhyme center measure for CVS[?] words was the baseline, and speaker was included as a random intercept.

The results show that for both CVS[?] and CV:[?]S words, the sonorant rhyme center measure has significantly smaller SDs than the alternative moraic measure (Moraic measure: $\beta = 18.867$, $p < 0.001$) but that the sonorant rhyme center measure does not differ significantly between CVS[?] and CV:[?]S words ($\beta = -0.505$, $p = 0.825$), as shown below in Figure 7. An interaction between word type and stød onset measure type was found ($\beta = -11.732$, $p < 0.01$), indicating that the difference between stød onset measures for CV:[?]S words was significantly smaller than for CVS[?] words.

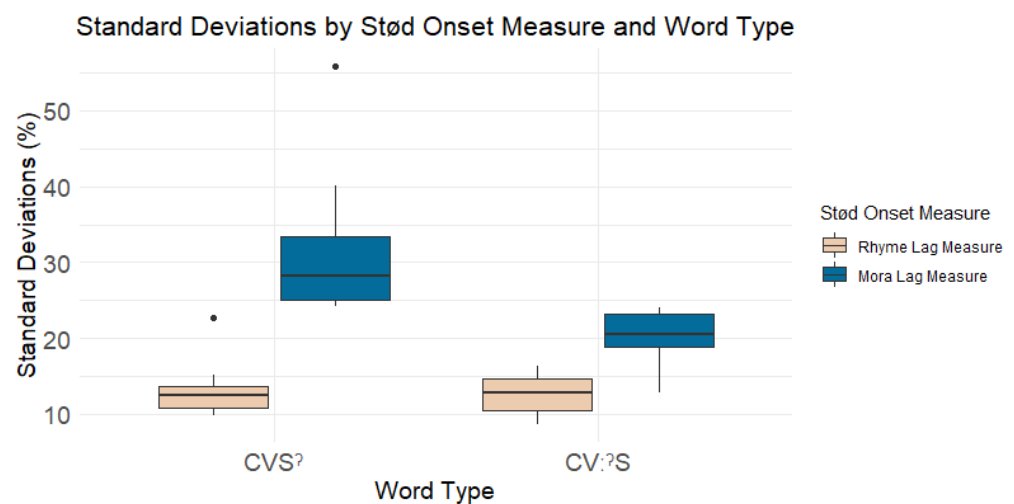


Figure 7. Standard deviations as a percentage of the relevant phonological domain by stød onset measure and word type.

3.2.2. Sonorant Rhyme Center Measure Analysis

The results in Section 3.2.1 indicate that the sonorant rhyme center measure represents the most stable timing relationship between stød and the syllable investigated here, evidenced by significantly lower SDs across word types. To determine whether the differences in stød onset timing between word types for the sonorant rhyme center measure are significant, a linear mixed-effects regression model was carried out with stød onset time as a percent difference from the rhyme center (the sonorant rhyme center measure) as the dependent variable and word type as a categorical fixed effect. To determine whether the average stød onset time across the different word types differed significantly from the sonorant rhyme midpoint, the fixed-effect word type was sum-coded and the dependent variable was scaled from -50 – 50% , centering the measure's means around 0.

Model results revealed that the model intercept was not significantly different from 0 ($\beta = -0.451$, $p = 0.834$), indicating that the average stød onset time across all word types does not differ significantly from the sonorant rhyme center. CV:[?] words have significantly later stød onset times than the overall average ($\beta = 10.947$, $p < 0.001$), and CV:[?]O have significantly earlier stød onset times ($\beta = -3.486$, $p = 0.03$). A post hoc Tukey test found that CV:[?] words have a later stød onset time than all other word types (CV:[?]–CVS[?]: $\beta = 10.320$, $p < 0.001$; CV:[?]O–CV:[?]: $\beta = -14.433$, $p < 0.001$; CV:[?]S–CV:[?]: $\beta = -12.032$, $p < 0.001$). No other significant effects were found. This is shown in Figure 8. Negative rhyme duration percentages indicate stød onset times in the first 50% of the rhyme, and positive percentages indicate stød onset times in the last 50%.

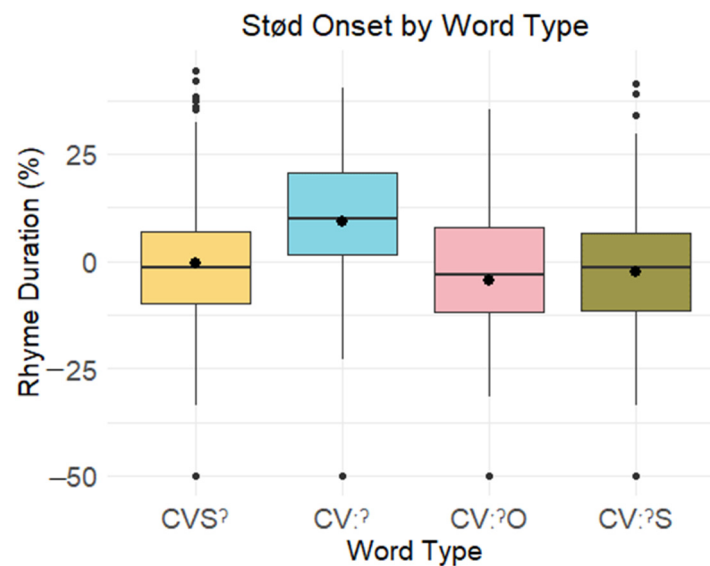


Figure 8. Stød onset (rhyme center measure) by word type, scaled from -50% to 50% . Negative rhyme durations indicate stød onset times in the first 50% of the rhyme. CV:O words have early stød onset times and CV:? words have late stød onset times compared to the overall average across word types, which does not differ significantly from 0. Black dots indicate the mean stød onset time per word type.

3.3. Stød Offset Timing

Turning to stød offset timing, preliminary visualization revealed that for all word types with fully sonorant rhymes (CVS?, CV:?, and CV:S), it is usually the case that stød offset occurred after the end of the word (88.36% of tokens). In contrast, CV:O words do in fact prevent stød from spreading due to the presence of a coda obstruent, which cannot host stød. Only two CV:O words had a stød offset time after the word's end due to lenition of the word-final stop in [sɖa:ʔb] 'staff/employees' and [se:ʔd] 'seen' to an approximant, allowing stød to spread beyond the target word. Therefore, unlike stød onset, stød offset cannot be analyzed using a relativized measure, since it consistently takes place at or after the word's end. Stød offset lag was thus calculated as the raw difference in milliseconds between the offset of stød and the end of the sonorant portion of the syllable rhyme for all word types.

A linear mixed-effects regression model was run with stød offset lag in milliseconds as the continuous dependent variable, word type as a categorical fixed effect, and CV:O words as the baseline, as they were expected to have stød offset times near 0 ms. The results show that CV:O tokens' stød offset times did not differ significantly from 0 ($\beta = -0.171$, $p = 0.943$). All other word types have significantly later stød offset times than CV:O words (CVS?: $\beta = 42.890$, $p < 0.001$; CV:?: $\beta = 50.572$, $p < 0.001$; CV:S: $\beta = 33.616$, $p < 0.001$). A post hoc Tukey test confirmed these results, additionally finding that CVS? and CV:? words have later offset times than CV:S words (CV:S–CVS?: $\beta = -9.274$, $p = 0.01$; CV:S–CV:?: $\beta = -16.956$, $p < 0.01$) but do not differ from each other (CV:–CVS?: $\beta = 7.682$, $p = 0.279$). These results are shown in Figure 9.

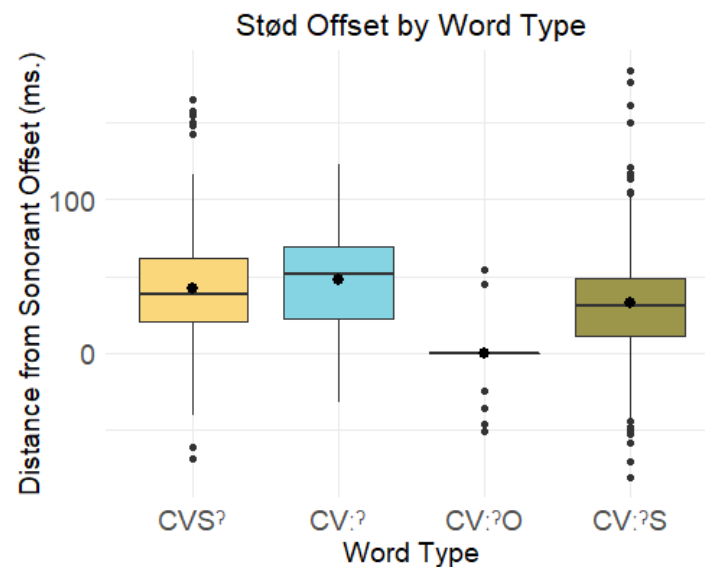


Figure 9. Stød offset in milliseconds by word type. CVS? and CV:? words have the latest stød offset times, which do not differ from each other. CV:S have earlier offset times, and CV:O words have the earliest offset times, which do not differ significantly from 0. Black dots indicate the mean stød offset time per word type.

3.4. Segmental Durations

3.4.1. Vowel Duration

The final set of segmental duration analyses was conducted to address the third research question: is there evidence from segmental durations for morae in Danish? The moraic hypothesis put forth by [Basbøll \(2005\)](#) does not make specific predictions about vowel duration across word types, but if Danish is a moraic language that formulates morae based on the shared mora hypothesis, then, following [Morén and Zsiga \(2006\)](#), vowels in CV:S words should be shorter than vowels in CV:? words, since the sonorant coda contributes weight to the second mora, which has a preferred length.

To investigate vowel duration, a linear mixed-effects model with vowel duration in milliseconds as a continuous dependent variable and the word type as categorical fixed effects was run with word type sum-coded. The results of the model show that all word types' vowel durations differ significantly from the overall average. Short vowels in CVS and CVS? words are significantly shorter than the overall average (CVS: $\beta = -41.778$, $p < 0.001$; CVS?: $\beta = -20.624$, $p < 0.001$), whereas long vowels in CV:?, CV:O, and CV:S words are significantly longer (CV:?: $\beta = 18.376$, $p < 0.001$; CV:O: $\beta = 33.684$, $p < 0.001$; CV:S: $\beta = 12.337$, $p < 0.001$). Post hoc Tukey tests further revealed significant differences between all word types except CV:? and CV:O words (CV:O–CV:?: $\beta = 15.038$, $p = 0.16$) and CV:? and CV:S words (CV:S–CV:?: $\beta = -6.038$, $p = 0.85$). Vowels in all types of stød-bearing words are also significantly longer than vowels in CVS words (CVS?–CVS: $\beta = 21.154$, $p < 0.001$; CV:–CVS: $\beta = 60.154$, $p < 0.001$; CV:O–CVS: $\beta = 75.462$, $p < 0.001$; CV:S–CVS: $\beta = 54.115$, $p < 0.001$), and long vowels in CV:?, CV:O, and CV:S words are longer than short vowels in CVS? words (CV:–CVS?: $\beta = 39.000$, $p < 0.001$; CV:O–CVS?: $\beta = 54.308$, $p < 0.001$; CV:S–CVS?: $\beta = 32.962$, $p < 0.001$). Finally, vowels in CV:S words are significantly shorter than vowels in CV:O words (CV:S–CV:O: $\beta = -21.346$, $p < 0.001$). This is shown below in Figure 10.

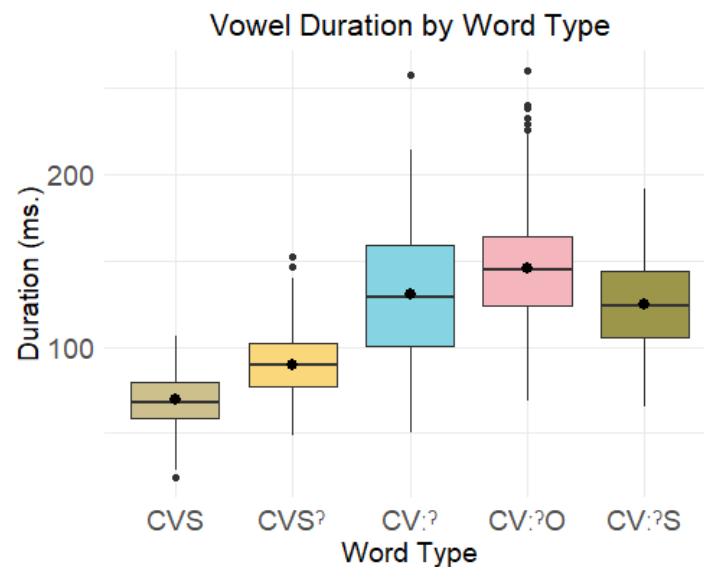


Figure 10. Vowel duration in milliseconds by word type. All differences are significant except the comparisons between CV:' and CV:'O words and between CV:' and CV:'S words. Black dots indicate the mean vowel duration per word type.

3.4.2. Coda Sonorant Duration

Sonorant consonant durations were analyzed to test the predictions of the moraic and the shared mora hypotheses. The moraic hypothesis predicts that non-moraic coda sonorants in CVS words are shorter than moraic codas in CVS' words. The shared mora hypothesis is agnostic to the relationship between coda sonorant durations in CVS and CVS' words. Rather, the shared mora hypothesis predicts that coda sonorants in CVS' words will be significantly longer than coda sonorants in CV:'S words. To investigate these predictions, a linear mixed-effects regression model with only CVS, CVS', and CV:'S words ($n = 827$) was run with sonorant coda duration in milliseconds as a continuous dependent variable and word type as a sum-coded categorical fixed effect. The results of the model show that coda sonorant durations in CVS words do not differ from the overall average ($\beta = 2.677$, $p = 0.20$). In contrast, coda sonorants in CVS' words were slightly longer than the overall average ($\beta = 6.852$, $p = 0.04$) and coda sonorants in CV:'S words were shorter than the overall average ($\beta = -12.363$, $p < 0.001$). Post hoc Tukey tests found a significant difference between sonorant durations in CVS and CVS' words compared to CV:'S words, but not compared to each other (CV:'S–CVS: $\beta = -15.041$, $p < 0.001$; CV:'S–CVS': $\beta = -21.893$, $p < 0.001$; CVS'–CVS: $\beta = 6.852$, $p = 0.08$). This is shown below in Figure 11.

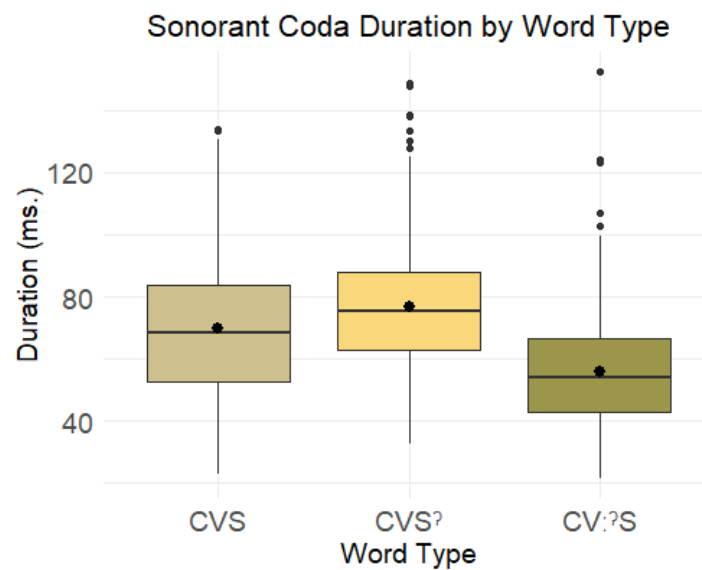


Figure 11. Sonorant coda duration in milliseconds by word type. CVS and CVS? words have significantly longer coda consonants than CV:~S words, but do not differ from each other. Black dots indicate the mean sonorant coda duration per word type.

3.4.3. Sonorant Rhyme Duration

The last durational analysis on overall sonorant rhyme duration was conducted to test the predictions of the shared mora hypothesis, which predicts that all stød-bearing word types will have approximately the same sonorant rhyme duration. To investigate differences between sonorant rhyme durations, a linear mixed-effects model was run with sonorant rhyme duration as a continuous dependent variable and word type as a sum-coded categorical fixed effect. The results show that sonorant rhyme durations for all word types differ significantly from the overall average. CVS, CV:~, and CV:~O words have significantly shorter sonorant rhymes (CVS: $\beta = -16.946$, $p < 0.001$; CV:~: $\beta = -28.103$, $p < 0.001$; CV:~O: $\beta = -13.059$, $p < 0.001$), whereas CVS? and CV:~S words have significantly longer sonorant rhymes compared to the overall average (CVS?: $\beta = 10.559$, $p < 0.001$; CV:~S: $\beta = 21.501$, $p < 0.001$). Post hoc Tukey tests revealed that sonorant rhymes in CVS, CV:~, and CV:~O words do not differ significantly from each other (CV:~–CVS: $\beta = -11.157$, $p = 0.56$; CV:~O–CVS: $\beta = 3.887$, $p = 0.97$; CV:~O–CV:~: $\beta = 15.044$, $p = 0.25$), but are all shorter than sonorant rhymes in CVS? and CV:~S words, which do not differ significantly (CVS?–CVS: $\beta = 127.505$, $p < 0.001$; CV:~S–CVS: $\beta = 38.447$, $p < 0.001$; CV:~–CVS?: $\beta = -38.662$, $p < 0.001$; CV:~O–CVS?: $\beta = -23.618$, $p < 0.001$; CV:~S–CV:~: $\beta = 49.604$, $p < 0.001$; CV:~S–CV:~O: $\beta = 34.560$, $p < 0.001$; CV:~S–CVS?: $\beta = 10.942$, $p = 0.18$). This is shown below in Figure 12.

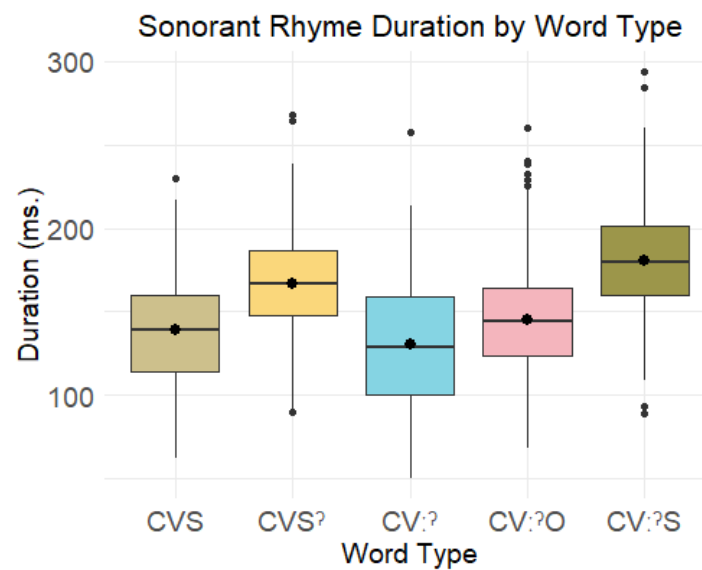


Figure 12. Sonorant rhyme duration in milliseconds by word type. CVS? and CV:°S words have the longest sonorant rhyme durations, which do not differ significantly. CVS, CV:°, and CV:°O words have shorter sonorant rhymes which do not differ significantly in duration. Black dots indicate the mean sonorant rhyme duration per word type.

4. Discussion

4.1. General Discussion

The goals of this study are threefold. The first goal is to determine which hypothesis of stød timing best captures the relationship between stød and the syllable. The second goal is to establish how to model the relationship between stød, a suprasegmental phonation type, and the syllable, using the framework of Articulatory Phonology. The final goal of this study is to assess if there is sufficient independent evidence to substantiate a moraic interpretation of stød timing. I address the first question here, before developing a sketch of glottal phonology in Section 4.2. I show how stød coordination can be modeled in an articulatory framework in Section 4.3. In Section 4.4, I focus on stød realization in CV:°O and CV:° words, the two word types for which stød onset differed significantly from the overall average, before concluding the discussion in Section 4.5, where I discuss the segmental duration analyses with respect to moraic analyses of stød timing.

Overall, the analyses confirmed previous findings that stød realization is biphasic with modal phonation and high F0 in the first part of the syllable rhyme followed by a stød phase proper in the second part of the syllable rhyme. Furthermore, the analyses here support a refinement of this broad statement, such that the first phase is timed relative to the first half of the sonorant rhyme, and the second phase is timed relative to the second half. The first F0 analysis confirmed that F0 in CVS?, CV:°, and CV:°S words is higher during the first 3/5 of the vowel than in CVS words. The exception to this pattern was that F0 at the beginning of CV:°O words did not differ significantly from stød-less CVS words. This is discussed further in Section 4.4. Furthermore, the second F0 analysis revealed that F0 decreases significantly near the middle of the sonorant rhyme across all stød-bearing word types. This result corroborates the findings of the stød onset timing analysis, which confirmed that for the two crucial word types, CVS? and CV:°S, the sonorant rhyme center measure minimized variability in stød onset timing across word types compared to the alternative measures (the sonorant coda measure for CVS? words, and the vowel midpoint measure for CV:°S words), evidenced by significantly lower SDs. The sonorant rhyme center measure also reduces differences in stød onset timing means across word types, all of which cluster around the center of the rhyme. Thus, the results here support the conclusion that the sonorant rhyme center hypothesis best represents the coordination relationship

between stød and the syllable, allowing for an interpretation of stød timing that is unified across word types.

The stød offset analysis also has significant implications. In the data presented here, stød spread to the following word in 88.36% of the data, indicating that spillover to a post-tonic syllable is in fact the norm rather than the exception. This is particularly notable for CV:ʔS words, as previous work has claimed that the relevant phonological domain of stød is the second half of a long vowel or a coda sonorant in V+S rhymes. This claim has framed stød in coda sonorants in CV:ʔS words as both irrelevant to the phonological specification of stød's domain and, implicitly, as deviant from an ideal stød realization. However, the data here suggest that, rather than being exceptional, stød is nearly always present in the entire latter portion of the sonorant rhyme, regardless of its segmental contents.

Together, the stød onset and stød offset timing analyses show that the sonorant rhyme center hypothesis best captures how stød is realized relative to the syllable. Furthermore, the analyses point towards a conceptualization of stød domain as the sonorant portion of the syllable rhyme, in line with Grønnum and Basbøll (2007), Grønnum et al. (2013), and Grønnum (2015), which unifies the predictions of stød timing across all stød-bearing syllables, regardless of their rhyme's segmental content. Furthermore, the timing relationship between stød and the syllable can be expressed quite simply using Browman and Goldstein's theory of Articulatory Phonology (Browman and Goldstein 1986, 1989a, 1989b, 1990, 1992, 2000). In the next subsection, I develop an account of stød timing in which I propose two additions to the theory. First, building on articulatory work from Esling et al. (2019), I develop a theory of glottal articulatory phonology that allows for the theoretical modeling of the biphasic nature of stød. Second, I posit the existence of a new articulatory landmark for coordination, the sonorant rhyme center, and show how integrating this new landmark allows for a straightforward account of stød timing relative to the syllable.

4.2. Developing a Theory of Glottal Articulation

In this section, I develop a theory of glottal articulatory phonology that allows for the theoretical modeling of the biphasic nature of stød. Note that whilst I focus on modeling stød in Danish, this proposal is meant to be a universal addition to Articulatory Phonology which is motivated by and can be used to model the acoustic realization of nonmodal phonation types in multiple languages.

There are several advantages to modeling stød timing using the framework of Articulatory Phonology. First, it dispenses with the need to propose that the relationship between stød and the syllable is mediated by a prosodic intermediary, such as morae. Second, by specifying that the gestures which produce stød have a specific coordination relationship with each other and with a new articulatory landmark, the sonorant rhyme center, the biphasic nature of stød can be straightforwardly accounted for. This is beneficial as it neatly explains why stød affects the first half of the sonorant rhyme differently than the second half, something that previous theories have had to merely stipulate. Third, this approach dispenses with the notion that stød realization is random and not under speakers' control (Grønnum et al. 2013), which, as this study shows, does not appear to be the case.

In order to model stød realization in Danish, I depart from previous theories of articulatory phonology which treat all glottal phenomena (aspiration, tone, voicing, etc.) as emerging from various states of the glottis (Catford 1964, 1977). Instead, I integrate the view that the larynx is a complex structure composed of multiple articulators (Esling et al. 2019). This allows for the larynx to have numerous possible configurations resulting from the coupling of various laryngeal/glottal gestures which contribute to the acoustic quality of phonation. Crucially, adopting this view allows for the phonological representation of multiple glottal articulators whose gestures can be represented on independent gestural tiers. Here, I operationalize this by dividing the GLO tier into two sub-tiers. The first tier represents articulatory gestures that primarily affect phonation, and the second tier represents articulatory gestures that primarily affect F0. This is shown below in Figure 13 with the addition of the PHON(ATION) sub-tier and the F0 sub-tier.

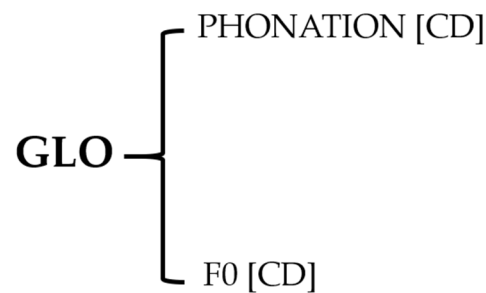


Figure 13. The GLO tier, subdivided into the PHON(ATION) tier and the F0 tier.

One motivating factor behind this division is the robust variability in non-modal phonation production, both within and across languages. For example, as [Fischer-Jørgensen \(1989\)](#) notes, Danish speakers consistently showed activation of the cricothyroid in the first part of stød-bearing syllables and increased vocal and ventricular fold constriction in the second part of the syllable. However, speakers varied in their use, and in the extent of their use, of other glottal articulators, such as the lateral cricoarytenoid. Cross-linguistically, [Keating et al. \(2011\)](#) found that the same phonation types (breathy/creaky) differed along multiple acoustic measures between Gujarati, Hmong, Mazatec, and Yi. This may indicate that speakers have phonologized a voice quality contrast, but the precise articulatory configuration that creates the contrast may vary, both within and across languages.

Another motivating factor for this division of the GLO tier is the complex interaction between nonmodal phonation and F0 cross-linguistically. [Esposito and Khan \(2020\)](#) systematically review languages that utilize phonation phonologically, showing that for some non-tonal languages, F0 is used phonetically as a cue for phonation, e.g., Javanese ([Brunelle 2010](#)) and Kedang ([Samely 1991](#)), whilst for others it is not, e.g., Gujarati¹² ([Khan 2012](#)), and for others, phonation is a cue for phonological tone, e.g., Mandarin ([Yang 2011](#); [Kuang 2017](#)) and Cantonese ([Yu and Lam 2014](#)). Furthermore, in some languages, phonation and tone are redundant cues, e.g., Green Hmong ([Andruski and Ratliff 2000](#)) and Vietnamese ([Alves 1995](#)), whereas in others they seem to act independently of each other, e.g., Mazatec languages ([Garellek and Keating 2011](#)). For example, in Yucatec Maya, creaky phonation is accompanied by a preceding phonological high tone ([Frazier 2013](#)). However, acoustically, creaky phonation is usually produced with low F0 ([Keating et al. 2015](#)), making an argument in which the preceding tone is an acoustic consequence of the same articulatory gesture that produces the following period of creaky phonation difficult. Rather, it seems more likely that separate laryngeal/glottal gestures are involved, one which results in high F0, and another which results in creaky phonation. More evidence that different articulators can be manipulated for phonation and F0 changes comes from [Garellek and Keating \(2011\)](#), who found that in Jalapa Mazatec, phonation did not affect speakers' F0. Together, these languages show that the interplay of phonation and F0 is complex and multidimensional, and languages differ drastically in the utilization of these acoustic dimensions and how they implement them articulatorily as well as in how they represent them phonologically.

The final argument for the division of the GLO tier comes from articulatory work by [Esling et al. \(2019\)](#). Traditionally in Articulatory Phonology, glottal phenomena are produced by varying degrees of constriction at the glottis, usually formalized of in terms of vocal fold constriction. However, Esling et al. have found robust evidence that glottal constriction is often the result of other glottal articulators' motion. That is, at the glottal place of articulation, the vocal folds are not always the primary articulators, independent of other laryngeal structures; rather, other laryngeal articulators often affect the state of the vocal folds, and therefore the glottis. For example, [Esling et al. \(2019\)](#) found that glottal stops, usually represented on the GLO tier with the constriction degree [clo(sure)], are often produced with ventricular fold–vocal fold coupling, in which the ventricular folds press down on the vocal folds, increasing the vibratory mass, thereby slowing and eventually ceasing vocal fold vibration, resulting in a full glottal stop. Likewise, a speakers' F0 can be raised by contraction of the cricothyroid, which lengthens and thins the vocal folds.

Accordingly, I have chosen to divide the GLO tier based on phonologized phonation and F0 targets rather than by particular articulators.

The possibility of independence between phonation type (modal, creaky, and breathy) and F0 (phonological pitch or not) cross-linguistically and articulatorily thus empirically motivates the division of the GLO tier as proposed here, allowing for a full and faithful account of the relationship between these two acoustic events. With this structure of the GLO tier in mind, in the subsequent section, I show how this addition to the theory allows for the biphasic realization of *stød* to be straightforwardly modeled in an articulatory framework. Note that even though here I am only concerned with modeling Danish, this structure may be used to model other languages with a variety of phonation and F0 contrasts, phonological or not. Cross-linguistic work should establish if further subdivisions of the GLO tier are needed, but for the current purposes, the subdivision into the PHON and F0 tiers are sufficient.

4.3. Developing a Model of Gestural Coordination between *Stød* and the Syllable

In this section, I present the formal analysis of how *stød* is gesturally coordinated relative to the syllable, as shown below in Figure 14. In these diagrams, gestures are assumed to be audible during their plateaus (from target to release). For readability, in these scores, I use Gafos (2002)-style notation to depict the coordination relationships between different gestures' articulatory landmarks but separate the gestures onto different tiers. Here, CVS words without *stød* are shown as a reference to orient the reader towards the notation scheme in Figure 14a, before examples of *stød*-bearing word are given in Figure 14b–e, and only the coordination relationships between segmental gestures and *stød* gestures are represented for succinctness. Between the tiers, solid vertical lines represent coordination between a *stød* gesture's and a segmental gesture's landmark, and dotted lines represent coordination between the two *stød* gestures themselves. Vowel gestures are given in solid lines, consonant gestures in dotted-and-dashed lines, and *stød* gestures in dashed lines. An audibility schematic is given below each diagram with the articulatory sonorant rhyme center labeled. When the sonorant rhyme consists of only one segment, as in CV:ʔ and CV:ʔO words, the sonorant rhyme center is the center of the plateau of the TB gesture. When complex, as in CVSʔ and CV:ʔS words, the sonorant rhyme center is calculated as the center of the plateaus of the vowel's TB gesture and the gesture that produces the constriction for the sonorant coda consonant in the oral tract. Note that the two *stød* gestures in Figure 14b–e are part of the same phonological phenomenon, as indicated by the fact that both are represented by dashed lines, and the F0 gesture is not posited to be a phonological tone. This contrasts with languages such as Yucatec Maya (Frazier 2013), in which creaky phonation is preceded by a phonological high tone. This difference could be represented with differing line types across the PHON and F0 tiers to show their independent phonological status in the language.

In these gestural scores, *stød* gestures are represented as having the constriction degree value [crit]. For the F0 tier, this is a somewhat arbitrary label meant to evoke the fact that the cricothyroid is contracting, which lengthens and tenses the vocal folds and thus raises F0; currently, there is a dearth of research showing whether different degrees of cricothyroid constriction are distinguished phonologically by speakers. On the PHON tier, the CD value [crit] is nonarbitrary and is based on articulatory descriptions of glottal constriction in Esling et al. (2019) for various laryngeal phenomena, including the production of [h], aspiration, voicelessness, breathy phonation, whispery phonation (similar to breathy phonation but slightly more constricted), creaky phonation, and glottal stops. Based on their descriptions, creaky phonation is produced with greater glottal constriction than [h], aspiration, voicelessness, breathy phonation, and whispery phonation but less constriction than full glottal stops. Concurrently, the second *stød* gesture, which is often acoustically linked to creaky phonation (Fischer-Jørgensen 1989), is denoted by the CD value [crit], indicating that the glottal aperture is small enough to impede airflow and cause perturbations in vocal fold vibration, leading to aperiodicity.¹³

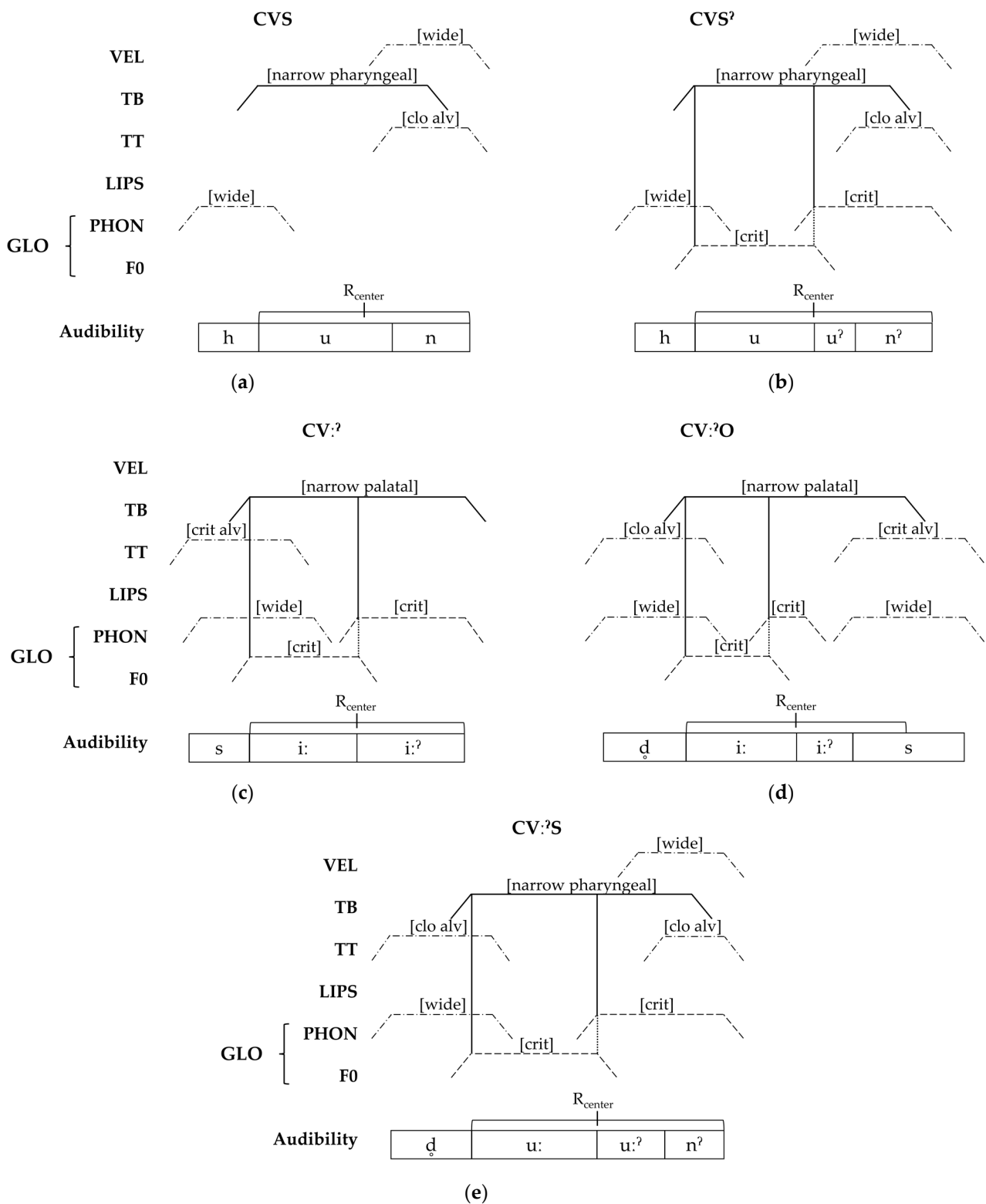


Figure 14. Gestural scores of stød coordination for CVS[?], CV:[?], CV:[?]O, and CV:[?]S words. An example of a CVS word is provided for reference. (a) Gestural coordination for CVS word [hun] ‘she’; (b) Gestural coordination in the CVS[?] word [hun[?]] ‘dog’; (c) Gestural coordination in the CV:[?] word [si:[?]] ‘sieve’; (d) Gestural coordination in the CV:[?]O word [d̥i:[?]s] ‘fog’; (e) Gestural coordination in the CV:[?]S word [d̥u:[?]n] ‘down (feathers)’.

The results of this study provide support for a biphasic analysis of stød, such that across word types, the first and second stød gestures are sequentially coordinated. This is shown in Figure 14b–e by the gestural constellation on the PHON and F0 tiers. Across the word types, the target of the first stød gesture (cricothyroid contraction), represented on the F0 tier, is coordinated with respect to the target of the vowel gesture on the TB tier, as shown by the solid vertical line connecting the two. This coordination relationship between the vowel and stød accounts for the presence of modal phonation with high F0 in the first half stød-bearing rhymes.

With regards to the second stød phase, this study finds that across word types, the PHON gesture of stød (ventricular fold–vocal fold coupling, following from Fischer-Jørgensen (1989)’s articulatory study of stød) is coordinated with respect to the center of the sonorant rhyme, as shown in Figure 14b–e. This finding necessitates the addition of a new articulatory landmark, the sonorant rhyme center, which can enter into coordination relationships with other gestures. This relationship is shown by the solid vertical line connecting the second stød gesture on the PHON tier and sonorant rhyme center on the TB tier. This relationship between stød and the sonorant rhyme is reminiscent of the c-center effect found for complex onsets. That is, just as the timing of vowel gestures is most stable relative to the center of an onset consonant (cluster) gesture(s), the timing of the target of the stød phase proper (the point at which it becomes audible) is most stable relative to the center of the sonorant rhyme, which, like onsets, may be complex.

Finally, turning to the relationship between the two stød gestures themselves, the results of the second F0 analysis show that across word types the decrease in F0 aligns with the onset of the stød phase proper at the sonorant rhyme center. This could be for two reasons. The first is that the F0 stød gesture releases at the sonorant rhyme onset, leading to a decrease in F0. The second is that the F0 stød gesture does not release at the sonorant rhyme center, but the second stød gesture on the PHON tier is what leads to the F0 decrease, as creaky phonation is associated with low F0. Differentiating between these two possibilities is difficult when lacking articulatory data. Therefore, here, I choose to represent the relationship between the two stød gestures such that the onset of the second stød gesture, represented on the PHON tier, coordinates with respect to the release of the first stød gesture on the F0 tier, as shown by the dashed vertical line connecting the two stød gestures in Figure 14b–e. This posits a maximally strong interpretation for the F0 decrease in the middle of the rhyme—that is, at the sonorant rhyme center, the F0 gesture is at the release point, and the PHON gesture reaches its target, leading to a decrease in F0. Further articulatory work on the exact coordination between the two stød gestures is needed to confirm this coordination relationship.

Thus, by proposing that the sonorant rhyme is a formal phonological unit that can enter into timing relationships with other gestures, the notion of morae can be dispensed with for this particular phenomenon, and the stød gestural constellation can be coordinated relative to the syllable without a prosodic intermediary. Furthermore, the biphasic nature of stød with its modal phonation and high F0 in the first half of the rhyme and nonmodal, often creaky phonation in the second half can be straightforwardly accounted for as a natural byproduct of the specific timing relationships between the gestures that produce stød without having to specify that Danish has a phonological high tone. Finally, this theory dispenses with the notion that stød realization is random and not under speakers’ control, providing an alternative theory of coordination that is predictive of stød coordination in other types of words.

4.4. Stød in CV:ʔO and CV:ʔ Words

As shown in Figure 14b–e, the target of the second stød phase is posited to be phonologically linked to the sonorant rhyme midpoint. This timing relationship holds well for CVSʔ and CV:ʔS words. However, the results of the stød onset analysis revealed that the second stød phase began significantly earlier in CV:ʔO words and significantly later CV:ʔ

words, compared to the overall average. What, then, may be responsible for these start times?

First, with respect to CV:ʔO words, it is notable that the coordination relationship between stød and the sonorant rhyme, as shown in Figure 14d, predicts that the second stød phase should only be audible for less than half of the vowel. This is due to coarticulatory overlap between the end of the vowel gesture and the beginning of the voiceless coda gesture, which is assumed to acoustically mask the end of the vowel. The center of the articulatory sonorant rhyme therefore occurs after the auditory center of the vowel (Browman and Goldstein 2000). Thus, if stød does gesturally coordinate with the sonorant rhyme center, this predicts later stød onset times for CV:ʔO words than for other words with long vowels. Despite this prediction, the acoustic data in this study suggest that the onset of the second stød phase in CV:ʔO words occurs earlier than the overall average across word types. What, then, leads to this early onset?

One explanation for this could be that the stød phase proper begins early relative to the articulatory sonorant rhyme center to maximize its perceptibility in CV:ʔO words. This is in line with other studies which have shown that gestures may realign in order to increase the recoverability of a segment (Löfqvist and Yoshioka 1980, 1981; Löfqvist and McGarr 1987; Munhall and Löfqvist 1992; Byrd 1996; Chitoran et al. 2002). For example, stop–stop clusters, in which the release of the first stop is its only percept, exhibit less gestural overlap within the cluster than in other phrasal positions to preserve the release of the first stop (Byrd 1996; Chitoran et al. 2002).

Similarly, I propose that here the [wide] gesture of the coda obstruent (coda obstruents in Danish are voiceless) on the PHON tier is what causes the gestural realignment between the second stød gesture and the rhyme. If the second stød gesture were to align with the sonorant rhyme center in CV:ʔO words, its duration would be relatively short due to the presence of the following [wide] gesture of the coda obstruent. This could make recoverability more difficult for listeners. By beginning the second stød gesture earlier, the duration of the second stød phase increases, thereby increasing its perceptibility, as shown in Figure 15. This early coordination of the second stød gesture also explains why F0 at the beginning of CV:ʔO and CVS words is not significantly different. If the PHON stød gesture is coordinated relatively early, then low F0, an acoustic correlate of creaky phonation, may depress the F0 effects of the first stød gesture to the point that the average F0 across the first 3/5 of the vowel do not differ significantly from stød-less words.

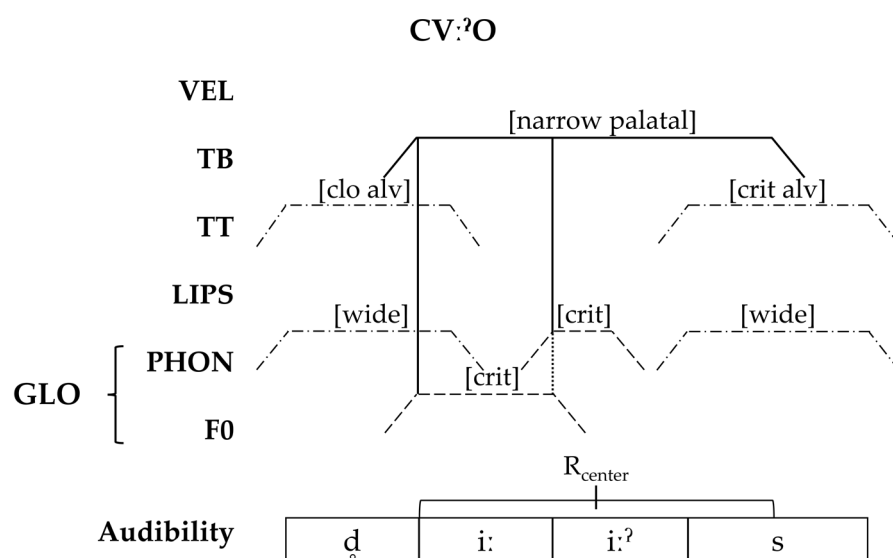


Figure 15. The phonetic realization of stød onset in CV:ʔO words, such as [d̥i:ʔs] ‘fog.’ Note that stød begins early relative to the sonorant rhyme center, possibly for perceptual reasons.

Other evidence that supports the hypothesis that speakers make articulatory adjustments to increase stød's perceptibility comes from the durational data. As shown in the vowel duration analysis, long vowels in CV:[?]O and CV:[?] words are phonetically the longest vowels analyzed here. This is somewhat unexpected, given that in many languages vowels in closed syllables tend to be phonetically shorter than in open syllables (e.g., Thai). It may therefore be the case that speakers phonetically lengthen vowels in CV:[?]O words in order to increase the domain during which stød can be realized. This argument is supported by typological data, such that there is a cross-linguistic tendency for phonologically complex tones to only appear in syllables with phonetically long sonorant rhymes (Gordon 2004; Zhang 2004). Furthermore, this vowel lengthening is similar to what has been found for other languages, such as Hausa (Zhang 2002). Zhang (2002) found that in Hausa, vowels in CVO syllables with the complex falling HL tone are phonetically longer than vowels in CVO syllables with simple H or L tones.¹⁴ Zhang interprets these findings to mean that the lengthening occurs to support the HL tone, which requires a longer duration to be realized than an H or L tone. Similarly, the lengthening of long vowels in CV:[?]O words in Danish may be interpreted as compensatory, given that stød cannot spread to a post-tonic syllable. Therefore, to increase stød duration, speakers may lengthen vowels in CV:[?]O words.

Turning to stød in CV:[?] words, a gestural score of the phonetic realization of stød in CV:[?] words is shown below in Figure 16. Note that the second stød phase begins after the sonorant rhyme center. Here, the discussion is more speculative, though of note is that during the first stød phase, F0 begins higher in CV:[?] words than any other word type. In fact, the F0 in CV:[?] words is significantly higher than the F0 in CVS and CV:[?]O words during the first 3/5 of the vowel, before decreasing sharply after the 3/5 interval of the vowel. The late onset of the second stød phase may therefore be a phonetic consequence of the relatively high starting point of F0 in CV:[?] words. Why, then, does F0 begin so high in CV:[?] words?

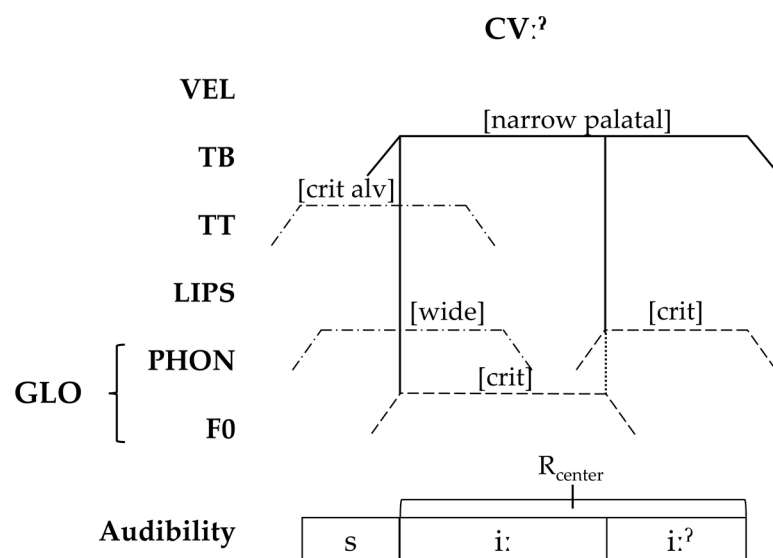


Figure 16. The phonetic realization of stød onset in CV:[?] words, such as [si:] 'sieve.' Note that stød begins late relative to the sonorant rhyme center, possibly for perceptual reasons.

Again, one potential explanation is a perceptual one. Of the stød-bearing words examined here, CV:[?] words have the shortest sonorant rhyme, which did not differ significantly from CVS or CV:[?]O words. Furthermore, the stød offset analysis revealed that stød does not spread more outside of CV:[?] words than other types of words, e.g., CVS[?] words. This indicates that in CV:[?] words, stød is being realized over a relatively short domain, which speakers may try to compensate for by enhancing the phonetic realizations of the stød phases to aid listeners' perception of stød. This phonetic enhancement could result in relatively high F0 at the beginning of stød-bearing syllables, which in turn may lead to

late onset times for the second stød phase. However, this explanation is still speculative, and further investigation of the relationship between the sonorant rhyme duration, F0, and creak onset timing in CV:ʔ words is needed.

4.5. *Durational Data with Respect to Morae in Danish*

In the final section of the discussion, I discuss the results of the durational analyses as they pertain to the moraic hypothesis and the shared mora hypothesis of stød timing.

First, coda sonorant durations in CVS and CVSʔ words were compared, following the prediction from Basbøll (1985, 2005) that moraic codas are longer than nonmoraic codas. The results here show that coda sonorants CVS and CVSʔ words do not differ significantly in duration, replicating previous findings (Grønnum and Basbøll 2001). This study thus finds no independent evidence to support a moraic interpretation of the data, given the traditional view of moraic formulation.

Second, duration measures were examined under the working hypothesis that similar sonorant rhyme durations across all word types indicate that morae have a predefined length in Danish, as in Thai (Morén and Zsiga 2006). Following from this, vowels in CV:ʔS words are expected to be shorter than vowels in CV:ʔ and CV:ʔO words, and coda sonorants in CV:ʔS are expected to be shorter than those in CVSʔ words, in order to preserve the relatively fixed length of the mora. Given that obstruents in Danish do not contribute to morae, unlike in Thai, vowels in CV:ʔO words and CV:ʔ words are predicted to be similar in duration.

The results of the rhyme duration analysis revealed significant differences such that CVSʔ and CV:ʔS words have significantly longer sonorant rhyme durations than CV:ʔ and CV:ʔO words. This is likely due to the fact that CVSʔ and CV:ʔS words have two sonorant segments in the rhyme, whereas CV:ʔ and CV:ʔO words only contain one. Furthermore, the vowel duration analysis revealed no significant differences in vowel duration between CV:ʔ and CV:ʔS words. With respect to the coda sonorant duration analysis, the results revealed that coda sonorants in CV:ʔS words are significantly shorter than coda sonorants in CVSʔ words, as predicted by the shared mora hypothesis. However, the coda sonorants do not shorten to the point that the sonorant rhyme in CV:ʔS words is commensurate in length to that of CV:ʔ words, casting doubt as to whether the shortening of coda sonorants in CV:ʔS words in Danish is due to an abstract phonological preference to preserve the duration of mora, as in Thai.

Overall, this study does not find sufficient evidence to interpret the acoustic findings as supporting either moraic analysis of stød timing. Of the four acoustic measures investigated, only one result was in line with the predictions of the shared mora analysis. Furthermore, as previously stated, morae do not play any role in Danish other than to explain the distribution of stød-bearing syllables. For example, stress in Danish is free and is not sensitive to syllable weight (Grønnum 1998). I therefore reject a moraic interpretation of the data here and conclude that a gestural theory of stød coordination relative to the syllables' sonorant rhyme better captures the timing relationship between stød and the syllable.

5. Conclusions

In conclusion, this study contributes to the body of work finding that stød in Danish is realized as biphasic, with modal phonation and high F0 at the beginning of stød-bearing rhymes and non-modal, often creaky phonation at the end of stød-bearing rhymes. The study also finds new evidence that the stød phase proper is timed with respect to the center of the sonorant rhyme across word types examined here and not to morae. This finding is formalized in the discussion, where I expand upon previous theories of articulatory phonology in order to capture the biphasic nature of stød realization. This expansion of the theory has implications for theories of Articulatory Phonology, allowing for the modeling of many complex glottal phenomena which have yet to receive adequate attention under this framework. A new articulatory landmark, the sonorant rhyme center, is also posited, which can enter into coordination relationships with other gestures. Finally, this study rejects

a moraic interpretation of the data, showing that there is no unambiguous, independent evidence for morae in Danish.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/languages7010050/s1>. The complete word list is attached as an excel document and expanded examples of segmentation are included as PNG files.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of New York University (protocol code 19-3248, approved 3 June 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Due to restrictions on data distribution by the E.U., the data are not available online. The sound files and text grids are available to researchers upon request.

Acknowledgments: I would like to thank Lisa Davidson, Juliet Stanton, Gillian Gallagher, Marc Garellek, the members of the NYU PEP Lab, participants at the 2020 Annual Meeting on Phonology and the 2020 Acoustical Society Meeting (“Acoustics Virtually Everywhere”), and two anonymous reviewers for their feedback on this work.

Conflicts of Interest: The author declares no conflict of interest.

Notes

- ¹ In contrast to the clear and consistent effects of the cricothyroid, the vocal folds, and the ventricular folds on stød production across speakers, the other articulators discussed by Fischer-Jørgensen (1989), including the lateral crico-arytenoid, were used less consistently across speakers and their effects were less clear, leading me to put aside the issue of how they contribute to stød production for future research.
- ² Ito and Mester (2015) provide an account for Danish stress in which Danish is quantity sensitive and trochaic. However, as Kuznetsova (2018) points out, there are many exceptions to the generalizations they make in their analysis, leading me to set aside this work for now.
- ³ Here I use Morén and Zsiga’s notation of VV rather than V: to maintain theoretically neutral as to whether or not two adjacent vowels with the same quality comprise a single phonological unit in Thai.
- ⁴ Another speaker (male) from Frederikssund, a municipality northwest of Copenhagen, was also recruited but was excluded from the data for having a distinctly different accent from the other speakers.
- ⁵ The speaker lived in Idaho, U.S. for one year from ages 15–16.
- ⁶ The study was also designed to balance for vowel quality and coda sonorant manner of articulation, but these factors are not discussed here.
- ⁷ CVO words were also elicited but are not included in the analysis here.
- ⁸ As noted in Grønnum (1998), the phonetic pronunciation of [a:] is much higher than the transcription implies and is closer to [e:].
- ⁹ Models in which the fixed effect was sum-coded were run twice to obtain the coefficients and *p* values for every level of the fixed effects compared to the overall average, since one level of each fixed effect is always omitted from the model output. This did not change the coefficients or *p* values of the models.
- ¹⁰ Given that both males and females produced all word types, F0 was not normalized to semitones.
- ¹¹ For the model on the first fifth of the vowel, the voicing of the onset consonant was also included in a separate model as a fixed effect. The results of this model showed that vowels following voiceless onset consonants had significantly higher average F0s, but this effect did not differ across word types, and did not change the results presented in this paper. I there put aside the effect of onset voicing for now to focus on effect of word type on F0.
- ¹² In Esposito and Khan (2020), Danish is listed as a language with contrastive voice quality that does not utilize F0 as a cue for phonation, in contrast to other non-tonal languages which do have F0 specifications as part of their phonological representation (e.g., Javanese, Kedang). I believe that, contrary to Esposito and Khan (2020), Danish is more like the latter than the former set of languages.
- ¹³ To fill out the complete scale of possible CD values on the PHON tier, following descriptions in (Esling et al. 2019), here I assume that [h], aspiration, and voicelessness have CD values of [wide], indicating the glottis is spread and airflow is unimpeded. Breathy voice, which Esling et al. consider to be slightly less constricted than whispery voice, has a CD value of [mid], indicating slightly

more constriction, and whispery voice has a constriction value of [narrow], indicating even more constriction. Finally, a CD value of [clo] is used to indicate a glottal stop.

14 I would like to thank an anonymous reviewer for bringing this paper to my attention.

References

- Alves, Mark. 1995. Tonal Features and the Development of Vietnamese Tones. *Hawaii'i Working Papers in Linguistics* 27: 1–13.
- Andruski, Jean E., and Martha Ratliff. 2000. Phonation Types in Production of Phonological Tone: The Case of Green Mong. *Journal of the International Phonetic Association* 30: 37–61. [\[CrossRef\]](#)
- Avelino, Heriberto. 2010. Acoustic and Electroglottographic Analyses of Nonpathological, Nonmodal Phonation. *Journal of Voice* 24: 270–80. [\[CrossRef\]](#) [\[PubMed\]](#)
- Basbøll, Hans. 1985. Stød in Modern Danish. *Folia Linguistica* 19: 1–50. [\[CrossRef\]](#)
- Basbøll, Hans. 2005. *The Phonology of Danish*. Oxford: Oxford University Press.
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software* 67: 1–48. [\[CrossRef\]](#)
- Boersma, Paul, and David Weenink. 2021. Praat: Doing Phonetics by Computer [Computer program] (version 6.1.54). Available online: <http://www.praat.org/> (accessed on 10 April 2021).
- Bretz, Frank, Torsten Hothorn, and Peter Westfall. 2010. *Multiple Comparisons Using R*, 1st ed. Boca Raton: Chapman and Hall/CRC. [\[CrossRef\]](#)
- Broselow, Ellen, Su-I. Chen, and Marie Huffman. 1997. Syllable Weight: Convergence of Phonology and Phonetics. *Phonology* 14: 47–82. [\[CrossRef\]](#)
- Browman, Catherine P., and Louis Goldstein. 1986. Towards an Articulatory Phonology. *Phonology* 3: 219–52.
- Browman, Catherine P., and Louis Goldstein. 1988. Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica* 45: 140–55. [\[CrossRef\]](#) [\[PubMed\]](#)
- Browman, Catherine P., and Louis Goldstein. 1989a. Articulatory Gestures as Phonological Units. *Phonology* 6: 201–51. [\[CrossRef\]](#)
- Browman, Catherine P., and Louis Goldstein. 1989b. *Gestural Structures and Phonological Patterns*. Haskins Laboratories Status Report on Speech Research. New Haven: Haskins Laboratories, pp. 1–23.
- Browman, Catherine P., and Louis Goldstein. 1990. Tiers in Articulatory Phonology, with Some Implications for Casual Speech. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, pp. 341–76.
- Browman, Catherine P., and Louis Goldstein. 1992. Articulatory Phonology: An Overview. *Phonetica* 49: 155–80. [\[CrossRef\]](#)
- Browman, Catherine P., and Louis Goldstein. 2000. Competing Constraints on Intersegmental Coordination and Self-Organization of Phonological Structures. *Les Cahiers de l'ICP. Bulletin de La Communication Parlée* 5: 25–34.
- Browman, Catherine P., Louis Goldstein, John Kingston, and Mary E. Beckman. 1990. Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. *Tiers in Articulatory Phonology, with Some Implications for Casual Speech* 341: 376.
- Brunelle, Marc. 2010. The Role of Larynx Height in the Javanese. In *Austronesian and Theoretical Linguistics*. Edited by Raphael Mercado, Eric Potsdam and Lisa de Mena Travis. Amsterdam: John Benjamins, pp. 7–24.
- Byrd, Dani. 1996. Influences on Articulatory Timing in Consonant Sequences. *Journal of Phonetics* 24: 209–44. [\[CrossRef\]](#)
- Catford, John Cunnison. 1964. Phonation Types: The Classification of Some Laryngeal Components of Speech Production. In *In Honor of Daniel Jones*. Edited by David Abercrombie, Dennis Butler Fry, Peter MacCarthy, Norman Carson Scott and John Leslie Melville Trim. London: Longmans, pp. 26–37.
- Catford, John Cunnison. 1977. Mountain of Tongues: The Languages of the Caucasus. *Annual Review of Anthropology* 6: 283–314. [\[CrossRef\]](#)
- Chitoran, Ioana, Louis Goldstein, and Dani Byrd. 2002. Gestural Overlap and Recoverability: Articulatory Evidence from Georgian. *Gussenhoven, Warner, Papers in Laboratory Phonology* 7: 419–47.
- Davidson, Lisa. 2020. The Versatility of Creaky Phonation: Segmental, Prosodic, and Sociolinguistic Uses in the World's Languages. *Wiley Interdisciplinary Reviews: Cognitive Science* 12: e1547. [\[CrossRef\]](#)
- Ege, Niels. 1965. The Danish Vowel System. *GENGO KENKYU (Journal of the Linguistic Society of Japan)* 47: 21–35.
- Esling, John, Scott Moisk, Allison Benner, and Lise Crevier-Buchman. 2019. *Voice Quality. The Laryngeal Articulator Model*. Cambridge: Cambridge University Press.
- Esposito, Christina M. 2010. Variation in Contrastive Phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association* 40: 181–98. [\[CrossRef\]](#)
- Esposito, Christina M., and Sameer ud Dowla Khan. 2012. Contrastive Breathiness across Consonants and Vowels: A Comparative Study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42: 123–43. [\[CrossRef\]](#)
- Esposito, Christina M., and Sameer ud Dowla Khan. 2020. The Cross-Linguistic Patterns of Phonation Types. *Language and Linguistics Compass* 14: e12392. [\[CrossRef\]](#)
- Fischer-Jørgensen, Eli. 1989. Phonetic Analysis of the Stød in Standard Danish. *Phonetica* 46: 1–59. [\[CrossRef\]](#) [\[PubMed\]](#)
- Frazier, Melissa. 2013. The Phonetics of Yucatec Maya and the Typology of Laryngeal Complexity. *Language Typology and Universals* 66: 7–21. [\[CrossRef\]](#)
- Gafos, Adamantios I. 2002. A Grammar of Gestural Coordination. *Natural Language & Linguistic Theory* 20: 269–337.

- Gao, Man. 2008. Mandarin Tones: An Articulatory Phonology Account. Unpublished Doctoral dissertation, Yale University, New Haven.
- Garellek, Marc. 2012. The Timing and Sequencing of Coarticulated Non-Modal Phonation in English and White Hmong. *Journal of Phonetics* 40: 152–61. [\[CrossRef\]](#)
- Garellek, Marc. 2015. Perception of Glottalization and Phrase-Final Creak. *The Journal of the Acoustical Society of America* 137: 822–31. [\[CrossRef\]](#) [\[PubMed\]](#)
- Garellek, Marc. 2019. The Phonetics of Voice. In *The Routledge Handbook of Phonetics*, 1st ed. Edited by William F. Katz and Peter F. Assmann. Abingdon-on-Thames: Routledge, pp. 75–106. [\[CrossRef\]](#)
- Garellek, Marc, and Patricia Keating. 2011. The Acoustic Consequences of Phonation and Tone Interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41: 185–205. [\[CrossRef\]](#)
- Gordon, Matthew. 2004. Syllable Weight. In *Phonetically Based Phonology*. Edited by Bruce Hayes, Robert Kirchner and Donca Steriade. Cambridge: Cambridge University Press, pp. 277–312.
- Grønnum, Nina. 1998. Danish. *Journal of the International Phonetic Association* 28: 99–105. [\[CrossRef\]](#)
- Grønnum, Nina. 2015. F0, Voice Quality, and Danish Stød Revisited. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Paper Number 0423:1-5. Glasgow: The University of Glasgow.
- Grønnum, Nina, and Hans Basbøll. 2001. Consonant Length, Stød and Morae in Standard Danish. *Phonetica* 58: 230–53. [\[CrossRef\]](#)
- Grønnum, Nina, and Hans Basbøll. 2003a. Stød and Length in Standard Danish: Experiments in Laboratory Phonology. In *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Universitat Autònoma de Barcelona, pp. 455–58.
- Grønnum, Nina, and Hans Basbøll. 2003b. Two-Phased Stød Vowels—A Cognitive Reality. In *Proceedings from Fonetik 2003, Lövånger, Sweden, 2–4 June 2003*. Edited by Mattias Helder. Umeå: Umeå University, vol. 9, pp. 33–36.
- Grønnum, Nina, and Hans Basbøll. 2007. Danish Stød: Phonological and Cognitive Issues. In *Experimental Approaches to Phonology*. Oxford: Oxford University Press, pp. 192–206.
- Grønnum, Nina, Miquel Vazquez-Larruscaín, and Hans Basbøll. 2013. Danish Stød: Laryngealization or Tone. *Phonetica* 70: 66–92. [\[CrossRef\]](#)
- Hollien, Harry, and John F. Michel. 1968. Vocal Fry as a Phonational Register. *Journal of Speech and Hearing Research* 11: 600–4. [\[CrossRef\]](#)
- Hollien, Harry, Paul Moore, Ronald W. Wendahl, and John F. Michel. 1966. On the Nature of Vocal Fry. *Journal of Speech and Hearing Research* 9: 245–47. [\[CrossRef\]](#) [\[PubMed\]](#)
- Honorof, Douglas N., and Catherine P. Browman. 1995. The Center or Edge: How Are Consonant Clusters Organized with Respect to the Vowel? In *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, Sweden, 13–19 August 1995*. Edited by Kjell Elenius and Peter Branderud. Stockholm: Stockholm University, vol. 3, pp. 552–55.
- Ito, Junko, and Armin Mester. 2015. The Perfect Prosodic Word in Danish. *Nordic Journal of Linguistics* 38: 5–36. [\[CrossRef\]](#)
- Karlin, Robin, and Sam Tilsen. 2015. The Articulatory Tone-Bearing Unit: Gestural Coordination of Lexical Tone in Thai. *Proceedings of Meetings on Acoustics* 22: 1–9.
- Keating, Patricia A., Marc Garellek, and Jody Kreiman. 2015. Acoustic Properties of Different Kinds of Creaky Voice. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, UK, August 10–14.
- Keating, Patricia, Christina M. Esposito, Marc Garellek, and Jianjing Kuang. 2011. Phonation Contrasts Across Languages. In *Proceedings of the International Congress of Phonetic Sciences XVII*. Edited by Wai-Sum Lee and Eric Zee. Hong Kong: City University of Hong Kong, vol. 108, pp. 903–16.
- Khan, Sameer ud Dowla. 2012. The Phonetics of Contrastive Phonation in Gujarati. *Journal of Phonetics* 40: 780–95. [\[CrossRef\]](#)
- Kuang, Jianjing. 2017. Covariation between Voice Quality and Pitch: Revisiting the Case of Mandarin Creaky Voice. *The Journal of the Acoustical Society of America* 142: 1693–706. [\[CrossRef\]](#)
- Kuznetsova, Natalia. 2018. What Danish and Estonian Can Show to a Modern Word-Prosodic Typology. In *The Study of Word Stress and Accent*. Edited by Rob Goedemans, Jeffrey Heinz and Harry van der Hulst. Cambridge: Cambridge University Press, pp. 102–43.
- Laver, John. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.
- Löfqvist, Anders, and Nancy McGarr. 1987. Laryngeal Dynamics in Voiceless Consonant Production. In *Laryngeal Function in Phonation and Respiration*. Edited by Thomas Baer, Clarence Sasaki and Katherine Harris. Boston: College Hill, pp. 391–402.
- Löfqvist, Anders, and Hirohide Yoshioka. 1980. Laryngeal Activity in Swedish Voiceless Obstruent Clusters. *Journal of the Acoustical Society of America* 68: 792–801. [\[CrossRef\]](#)
- Löfqvist, Andres, and Hirohide Yoshioka. 1981. Laryngeal Activity in Icelandic Obstruent Production. *Nordic Journal of Linguistics* 4: 1–18. [\[CrossRef\]](#)
- Marin, Stefania. 2013. The Temporal Organization of Complex Onsets and Cudas in Romanian: A Gestural Approach. *Journal of Phonetics* 41: 211–27. [\[CrossRef\]](#)
- Marin, Stefania, and Marianne Pouplier. 2010. Temporal Organization of Complex Onsets and Cudas in American English: Testing the Predictions of a Gestural Coupling Model. *Motor Control* 14: 380–407. [\[CrossRef\]](#)
- Morén, Bruce, and Elizabeth Zsiga. 2006. The Lexical and Post-Lexical Phonology of Thai Tones. *Natural Language & Linguistic Theory* 24: 113–78.
- Munhall, Kevin, and Anders Löfqvist. 1992. Gestural Aggregation in Speech: Laryngeal Gestures. *Journal of Phonetics* 20: 111–26. [\[CrossRef\]](#)

- Poupplier, Marianne. 2012. The Gestural Approach to Syllable Structure: Universal, Language- and Cluster-Specific Aspects. In *Speech Planning and Dynamics*. Edited by Susanne Fuchs, Melanie Weirich, Daniel Pape and Pascal Perrier. Frankfurt am Main: Peter Lang, pp. 63–96.
- Samely, Ursula. 1991. Kedang (Eastern Indonesia) Some Aspects of Its Grammar. In *Forum Phonetikum*. Hamburg: Buske Verlag.
- Shaw, Jason A., Adamantios I. Gafos, Philip Hoole, and Chakir Zeroual. 2011. Dynamic Invariance in the Phonetic Expression of Syllable Structure: A Case Study of Moroccan Arabic Consonant Clusters. *Phonology* 28: 455–90. [[CrossRef](#)]
- Shue, Yen-Liang, Patricia Keating, Chad Vicenik, and Kristine Yu. 2011. VoiceSauce: A Program for Voice Analysis. Paper presented at 17th International Congress of Phonetic Sciences, Hong Kong, China, August 17–21; pp. 1846–49.
- Styler, Will. 2017. On the Acoustical Features of Vowel Nasality in English and French. *The Journal of the Acoustical Society of America* 142: 2469–82. [[CrossRef](#)] [[PubMed](#)]
- Woodson, Gayle E., Oommen Mathew, Franca Sant’Ambrogio, and Giuseppe Sant’Ambrogio. 1989. Effects of Cricothyroid Muscle Contraction on Laryngeal Resistance and Glottic Area. *Annals of Otology, Rhinology & Laryngology* 98: 119–24. [[CrossRef](#)]
- Yang, Ruo-Xiao. 2011. The Phonation Factor in the Categorical Perception of Mandarin Tones. In *ICPhS XVII*. Hong Kong: City University of Hong Kong, pp. 2204–7.
- Yu, Kristine M., and Hiu Wai Lam. 2014. The Role of Creaky Voice in Cantonese Tonal Perception. *Journal of the Acoustical Society of America* 136: 1320–33. [[CrossRef](#)]
- Zhang, Jie. 2002. *The Effects of Duration and Sonority on Countour Tone Distribution: A Typological Survey and Formal Analysis*. New York: Routledge.
- Zhang, Jie. 2004. The Role of Contrast-Specific and Language-Specific Phonetics in Contour Tone Distribution. In *Phonetically-Based Phonology*. Edited by Bruce Hayes, Robert Kirchner and Donca Steriade. Cambridge: Cambridge University Press, pp. 157–90.