LINGUISTICA URALICA LX 2024 4 : 241—270 https://dx.doi.org/10.3176/lu.2024.4.01

*TUULI TUISK* (Copenhagen—Tartu—Riga),
*NICOLAI PHARAO* (Copenhagen)

# UNVEILING TONAL CONTRASTS IN THE BALTIC REGION: EXPLORING STØD IN LIVONIAN SPONTANEOUS SPEECH

**Abstract.** This paper presents findings for the tonal contrast and phonation differences between words with and without stød in Livonian spontaneous speech. Livonian differentiates between two contrastive phonological tones: the broken tone or stød and the plain tone. Stød is similar to the Danish stød in some respects and is said to be part of the tone systems of the phonologies of languages in the Baltic region. The findings show that the tonal contrast between words with and without stød tends to be neutralised in Livonian spontaneous speech, but there are individual differences between speakers and also differences between men and women. The most common non-modal phonation period categories in words with stød are creaky and tense. The results also indicate that stød disappears when the word has no prominence.

**Keywords:** Livonian, fundamental frequency, Generalized Additive Models, stød, non-modal phonation, prominence, spontaneous speech.

## 1. Introduction

The languages around the Baltic Sea constitute an interesting linguistic area where the Indo-European and Finno-Ugric languages meet. This region is also called the Circum-Baltic area which has been an arena for intensive linguistic contact, as well as migrations, colonizations and expansions (see, e.g., Koptjevskaja-Tamm & Wälchli 2001; Wälchli 2011). However, it has never been an economically, politically, culturally or linguistically united region. There is a multitude of interactions within specific ethnic communities and languages or language groups across this area, displaying extensive diversity. An example of such a contact-rich area is the region where Latvian, Livonian, and Estonian languages are spoken (Stolz 1991). It has been stated that the convergence in the Circum-Baltic area works primarily on a micro-level and reflects language contact between groups of people and maximally of two or three languages (Koptjevskaja-Tamm & Wälchli 2001). The region exhibits moderate genetic diversity, but it is the continuity of contacts over a long period of time that is

Received 13 May 2024, accepted 11 November 2024, available online 10 December 2024.

regarded as the predominant characteristic. Current understanding suggests the Finnic speakers arrived at the Baltic Sea approximately 3200 to 2800 years ago, encountering Baltic tribes en route (see Lang 2018; Grünthal & Heyd & Holopainen & Janhunen & Khanina & Miestamo & Nichols & Saarikivi & Sinnemäki 2022).

A number of linguistic features serve to mark the Circum-Baltic area as a buffer zone between the languages of the Standard Average European type and Central Eurasia type (Wälchli 2011). In comparison with other well-known linguistic areas, the number of languages involved in the Circum-Baltic area is relatively small. Depending on how they are counted, there are around 20—30 languages (Wälchli 2011 : 325—326): Baltic: Lithuanian (including Žemaitian), Latvian (including Latgalian), and extinct Old Prussian; Finnic: Livonian, Estonian (North & South), Votic, (Ingrian), Finnish, Karelian (three major varieties), Veps; Germanic: Swedish (including Gutnish), Danish, German, Low German, Yiddish; Slavic: Russian, Belorussian, Polish, Kashubian; Indo-Aryan: varieties of Romani; and Turkic: Karaim. In other words, this area has a large number of well-documented dialects and contact varieties, many of which are endangered or extinct, and therefore we can assume that there is a high degree of language contact in the region.

It has been pointed out that there are no isoglosses covering all the Circum-Baltic languages (Koptjevskaja-Tamm & Wälchli 2001) and only two isoglosses have been proposed that could unite this area: polytonicity and word order in possessive noun phrases (Koptjevskaja-Tamm 2006). Polytonicity, i.e. the existence of tonal suprasegmental oppositions, in the Circum-Baltic area has been suggested to have at least three different origins and involve languages from the Scandinavian, Baltic and Finnic areas. Three different layers of polytonicity in the Circum-Baltic languages have been highlighted: (1) tones on long syllable nuclei in the Baltic languages, (2) overlength in Southern Finnic and some Latvian dialects, and (3) phonologization of secondary stress in Scandinavian (Koptjevskaja-Tamm & Wälchli 2001 : 640—646). Among these three layers, the initial one is a distinct archaism, representing a remnant of Late Indo-European polytonicity, whereas the other two are relatively recent developments. There is no evidence indicating that the development of polytonicity in any of these three areas was d i r e c t l y triggered by the influence of either of the other areas. However, some degree of mutual influence among the three polytonal areas is acknowledged.

The Danish phenomenon s t ø d is central in Danish phonology (for the prosodic theory of stød see Goldshtein 2023) and the term is often used to describe similar features found in Livonian. In Danish, stød has traditionally been described as a kind of creaky voice with tonal side effects, i.e., non-modal voice with aperiodicity and irregular amplitude, often accompanied by a fundamental frequency perturbation, particularly an abrupt and brief dip in fundamental frequency (e.g., Fischer-Jørgensen 1989; Grønnum & Basbøll 2001; Grønnum & Vazquez-Larruscaín & Basbøll 2013). Contemporary studies on stød demonstrate that stød is a multifaceted prosodic phenomenon with numerous phonetic associations, defying reduction to a single measure (for the phonetic research on common Danish stød see the comprehensive review by Grønnum 2022). The articulation of stød exhibits significant variability, spanning a spectrum from subtly compressed vocal quality to markedly creaky phonation (Grønnum & Basbøll 2007 : 200).

Livonian stød is reminiscent of Danish stød. In Livonian word prosody, stød is generally referred to as ”the broken tone” (in Livonian *katkāndõks tūoņ* or *murdtõd tūoņ*) and is described as contrasting with ”the plain tone” (*pīldzit tūoņ* or *vientõd tūoņ*). In both languages, stød can distinguish between meanings of words (here are some examples in orthography, where stød has been indicated by an apostrophe: Danish *læ'ser* 'reads' with stød vs. *læser* 'reader' no stød, Livonian *kuo'nnõ* 'at home' with stød vs. *kuonnõ* 'frog, PSg' no stød). Just like in Danish, Livonian stød requires a certain amount of voiced material in order to be realised. The stød-basis is a long stressed syllable with a long monophthong, diphthong, or triphthong, or alternatively a short vowel followed by a voiced geminate consonant or consonant cluster (here are some examples in IPA, where stød is marked with the superscripted symbol of glottal stop: [roːˀ] 'money', [kuoˀiːgə] 'ship, PSg', [iˀzːzə] 'father, PSg', [kɑˀnʲːdʲi] 'hen, PPl'). While not consistently emphasised in the overviews and studies of Livonian phonetics and phonology, words with stød must be stressed to manifest the characteristic features most clearly. The interpretation of the situation of Danish and Livonian, as well as the origin of stød in these languages, differs considerably among researchers (for Livonian, see section below). It has been suggested that the tonal word accent systems in the Circum-Baltic languages exhibit certain similar characteristics and that stød in both Danish and Livonian are a part of this picture (e.g., Kiparsky 2017; Riad 2000). We shall see that the data obtained for spontaneously spoken Livonian does not unequivocally support this model.

The Baltic languages are described as having tone oppositions, but the varieties differ, however, in how many tones they have, how these tones are phonetically realised, and in which context they apply. The traditional Latvian system, retained in some Central dialects, is described as having three syllable tones: (1) level or drawling, (2) falling, (3) broken (e.g., Kariņš 1996; Koptjevskaja-Tamm 2006; LvdaF 2013; LVG 2013). Acoustically, the level tone is characterised by level or rising fundamental frequency and in the case of the falling tone the fundamental frequency rises at the beginning of a syllable and then drops sharply. The broken tone is similar to the falling tone except that occasionally irregular phonation or glottal stop is present (e.g., Kariņš 1996; Laua 1984). In communication, ternary oppositions like these are typically not realised, and the majority of Latvian dialects have simplified the original three-tone system to a two-tone contrast, although the dialects differ in which two of the three tones have been merged together.

Lithuanian is described as having two contrastive tones (e.g., Dogil 1999; Balode & Holvoet 2001; Koptjevskaja-Tamm 2006), both in the standard language and in the two main dialect varieties — Žemaitian and Aukštaitian — and their realisation is closely related to the system of word stress. When stress falls on a syllable containing a long vocalic segment, it must be realised either as an acute (s h a r p, f a l l i n g) or a circumflex (e v e n, d r a w n) tone or syllable accent. The distinction between tones is clearest in western Lithuania, particularly in the Northern Žemaitian dialects, where the acute tone is realised as a broken tone. This broken tone has an initial rise in fundamental frequency and intensity, followed by a glottal stop, after which the remaining part of the vowel is much lower in intensity and fundamental frequency (Balode & Holvoet 2001 : 73).

Together with Estonian and Low Latvian (Central and Livonianized, also called Tamian, dialects), Livonian belongs to the group where the reduction of a non-initial syllable was compensated by a secondary lengthening of the initial syllable (Koptjevskaja-Tamm & Wälchli 2001). However, secondary lengthening is a more extensive phenomenon in Livonian and can only be partly explained diachronically by the reduction of non-initial syllables. Livonian is the only Finnic language to differentiate between two contrastive phonological tones. Various studies of both controlled and spontaneous speech have addressed the descriptions of the Livonian stød and presented the most characteristic acoustic reflexes of it. For example, words with stød have a rising-falling or falling shape of the pitch contour (Posti 1936; 1937a; 1937b; Penttilä & Posti 1941; Pajupuu & Viitso 1986) and a higher fundamental frequency at the beginning of syllables (Vihman 1971) and intensity peaks occur early in the primary-stressed syllable of words with stød (Teras & Tuisk 2009). The characteristic shape of the pitch contour sometimes occurs together with creaky phonation, also known as laryngealization or glottalisation, i.e., vocal fold vibrations with irregular variation in periodicity as well as amplitude (Vihman 1971; Teras & Tuisk 2009). Analyses of intensity have revealed that words with stød have a decrease in intensity (Vihman 1971) or an irregular intensity contour (Tuisk 2015). Words with stød and without stød can also differ in duration: words with stød have a decrease in primary stressed syllable duration (Suhonen 1982; Tuisk & Teras 2009). However, there can also be a decrease of the duration of the second syllable in words with stød, while the primary stressed syllable duration remains the same for words with and without stød (Tuisk 2015).

Out of these characteristics, a distinctive pitch contour and laryngealization are most common. In terms of acoustics, Livonian stød is always associated with a tonal difference, and characterized by a tonal contour with a rapid and extensive rise, often followed by a rapid fall. The f0 peak or turning point generally occurs early in the primary-stressed syllable.

In the case of Livonian, under the term ”laryngealization” there are various descriptions of glottal modification that could be categorised under the collective term ”non-modal phonation”. Non-modal phonation is characterised by certain acoustic, physiological and perceptual differences from the modal phonation state (e.g., Laver 1980; Gordon & Ladefoged 2001). Livonian non-modal phonation has been described as ranging from no audible or weakly audible glottal effect up to complete glottal closure (e.g., Tuisk 2014).

Spectral tilt is one of the acoustic parameters that differentiates phonation types in a number of languages. The study on Livonian stød carried out by Balodis (2018) tested the method of Esposito (2004). The data used in this pilot study were taken from monologues of one male speaker and analysed two different rounded back vowels in CVV syllables. The study used six measurements of spectral tilt to determine whether Livonian stød was associated with any non-modal phonation, specifically creaky voice. According to the results, it was concluded that one measurement of spectral tilt (H1-H2, i.e., first harmonics — second harmonics; H1-H2 reflects glottal constriction or open quotient, where lower values indicate greater constriction, e.g., Keating & Garellek 2015) out of six can be used to test for creaky voice in Livonian. It was shown that creaky voice never occurred with vowels in non-stød syllables.

In Paul Kiparsky's view (2017), Livonian stød is a fundamentally tonal phenomenon. He analyses it as a falling (HL) tone, where a short high-toned syllable acquires a second, low-toned mora. In his interpretation, the glottal constriction and shortening that often accompany stød are phonetic enhancements of the distinctive falling pitch contour.

Regarding the historical connection between Livonian stød and Latvian tonal accents, scholarly opinions differ. Some researchers suggest that stød developed strictly through Latvian influence, attributing its development to language contact and the contrast of word tone, which is atypical in Finno-Ugric languages (e.g., Thomsen 1890; Kettunen 1925; Winkler 1999; 2000; Kiparsky 2017). Conversely, other perspectives posit that the phenomenon emerged independently in Livonian and Latvian due to the different functions of their suprasegmental systems (e.g., Posti 1942; Koptjevskaja-Tamm & Wälchli 2001). As for Livonian and Latvian, we share the notion that stød emerged independently in these languages, yet we acknowledge a significant Latvian influence on the characteristic acoustic features of Livonian stød.

It is important also to note the broader historical and linguistic context of the Livonian language in Latvia. Livonian was historically spoken in regions across Latvia, particularly in Northern Vidzeme and the northern coast of Courland around the Gulf of Riga, extending to south-western part of Estonia. The assimilation of Livonians and Latvians in Northern Vidzeme, leading to a language shift, occurred predominantly during the second third of the 19th century. In Courland, this transition took a slightly longer time. However, by at least the late 19th century, the Livonian community had become entirely bilingual (Ernštreits 2012). The focus of our study is on Courland Livonian. Bilingualism emerged primarily due to Livonian never attaining official status in institutions, schools, or churches; instead, it was mainly used within familial and social contexts and cultural activities. Following World War II, the bilingual Livonian community gradually shifted towards predominantly using Latvian. Over time, Latvian replaced Livonian in household communication, with proficiency in Livonian retained mostly by the older generation, born before or between the two World Wars. The recordings used for our analyses span from the 1970s to the 1990s, during which time the influence of Latvian is likely to have increased. Our data not only show pronunciation discrepancies among speakers but also variations within individuals. Seven of our speakers were born prior to World War I, while three were born during or after it. They represent a generation that was highly proficient in Livonian during their formative years, yet they were entirely bilingual. By the time of recording, they were likely even more influenced by Latvian. While we do not have independent evidence (from analysis of their grammar or lexical choices), it seems highly likely that their native fluency in Livonian was affected by Latvian influence, even if we do not know how inclined they were to maintain their native language in various contexts or switch into the dominant Latvian of the surrounding community. Following World War II, the Livonian population experienced a rapid decline. This decline can be attributed to various factors such as emigration, deportation, reluctance to openly identify with one's ethnic background due to fear, and the Soviet authorities' often unwillingness to acknowledge "Livonian" as an official ethnicity. While Livonian community and cultural activities resumed in the 1970s, the confidence to use Livonian in daily interactions was diminished.

Despite belonging to different language families, Livonian and Latvian contain traces of mutual contact and both languages have been givers and receivers. Livonian played a significant role in shaping the development of Latvian and its dialects as Latvian expanded into Livonian territories in Vidzeme, Kurzeme, and Zemgale. As a result, Livonian left its mark on the phonetics, grammar, and lexicon of Standard Latvian and its variations (Rudzīte 1996; Ernštreits & Kļava 2014). Evidence of Livonian influence on Latvian can be observed in the Livonianized dialects of Vidzeme and Kurzeme, as well as in numerous toponyms of Livonian (e.g., Rudzīte 1994 : 289). It is plausible to assume that the influence between languages was reciprocal, as evidenced by the replacement of the Livonian rising-falling tone with a falling tone as generally in Latvian tonal patterns. It has been observed that in the Livonianized dialect of Latvian, spoken in Courland and Northern Vidzeme, the falling and broken tones have sometimes merged into the falling tone. A similar pattern might exist in Livonian as well. However, sometimes, an intriguing difference remains: while in the Latvian Livonianized dialect, the falling and broken tones combine into a single falling tone, in Livonian, words with stød can exhibit a rising-falling tone (as shown, for example, by Penttilä & Posti 1941; Pajupuu & Viitso 1986 and in our results).

As a final note regarding the connections between the different kinds of polytonicity in the Circum-Baltic area, it should be noted that, inspired by Kiparsky's analysis of Livonian stød, there have been proposals (e.g., Riad 2000; Morén 2005a; 2005b) that Danish stød is the manifestation of a HL tonal pattern compressed within one syllable. Studies on Danish stød on this matter have found no phonetic reality to support this suggestion in Danish (e.g., Grønnum & Vazquez-Larruscaín & Basbøll 2013; Grønnum 2014). The overall f0 patterns are found to be identical in words with and without stød. It has been concluded that laryngealization as an autonomous syllable prosody in Danish, orthogonal to pitch and intonation, suffers none of the shortcomings attached to a tonal representation.

## 2. Implications for the current study and research aims

A number of conclusions from previous studies give rise to new research questions addressed in the current study. This research aims to analyse the Livonian stød involving a larger amount of data and discusses the roles of spontaneous speech, prominence and non-modal phonation in the phonetic realisation of stød.

The primary source for phonetic data in investigating the Livonian stød has mostly been carefully structured frame sentences, read phrases or single test words (e.g., Vihman 1971; Pajupuu & Viitso 1986; Lehiste & Teras & Pajusalu & Tuisk 2007; Lehiste & Teras & Ernštreits & Lippus & Pajusalu & Tuisk & Viitso 2008; Teras & Tuisk 2009). Several studies based on data from spontaneous speech have investigated disyllabic words with a specificsyllabic structure (e.g., Tuisk & Teras 2009; Tuisk 2015) or single primary stressed syllables of disyllabic words (Balodis 2018).

The motive for analysing data from spontaneous speech comes from the complex nature of this type of speech. Spontaneous speech is optimised for communication and can exhibit greater suprasegmental variability than read speech (e.g., Cucchiarini & Strik & Boves 2002; Labov 2001; Shriberg 2005). Prosodically, read speech is characterised by lower articulation rate, a lack of segmental reduction, fewer hesitations and a different pause structure (e.g., Levin & Schaffer & Snow 1982; Laan 1997). The most characteristic acoustic reflexes of the Livonian stød have been found in controlled speech. The aim of the current study is to investigate acoustically the f0 contours in Livonian words in spontaneous speech. The study compares words with and without stød elicited from the speech of five native female and five native male speakers. The realisation of the Livonian stød in a larger number of words with different syllabic structures and varying degrees of prominence will be investigated. The lack of analysis of monosyllabic words motivates the involvement of such words for the current analysis.

It is unclear how contrastive f0 is realised in Livonian spontaneous speech with respect to different degrees of prominence. The notion of linguistic or prosodic prominence has been continuously debated and received a wide number of contradicting or unspecific definitions (see Wagner & Origlia & Avezani & Christodoulides & Cutugno et al. 2015 and Cangemi & Baumann 2020 for discussions). The most widely accepted and very generic definition is provided by Terken and Hermes (2000) who define prominence as ”a property of a linguistic entity relative to an entity or a set of entities in its environment” and say that ”a linguistic entity is prosodically prominent when it stands out from its environment by virtue of its prosodic characteristics” (Terken & Hermes 2000 : 89). The comparison between words with different degrees of prominence allows us to investigate how a prosodically manifested contrast like stød interacts with prosodic prominence.

It has been pointed out in earlier studies that stød is realised also as a change in the phonation pattern — irregular vibrations of the vocal folds, i.e. glottalization or laryngealization. The classification of this non-modal phonation period would provide more information about the characteristics of stød.

The research questions of the present study are the following:

(1) How different are the f0 contours in words with and without stød in spontaneous speech?

(2) Do different degrees of prominence affect the realisation of words with stød?

(3) What kind of non-modal phonation occurs in words with stød?

## 3. Materials and methodology

The data used for this study were taken from archival sound recordings of ten Livonian speakers. The recordings are housed at the University of Tartu Archives of Estonian and Kindred Languages (Lindström & Lippus & Tuisk 2019; Tuisk & Pajusalu 2022 : 324—327). Table 1 displays details regarding the speakers and the timing of the recordings, with an asterisk denoting uncertain information. The interviews were conducted in the 1970s, 1980s, and 1990s and feature speakers talking about their daily lives, including work, family, and more. The age range of the five female speakers and five male speakers 64 to 94 years old. Eight speakers are from the eastern part of the Livonian Coast (villages Sīkrõg, Kuoštrõg, Pitrõg, Vaid, Kūolka). Two speakers are originally from the westernmost village Īra of the eastern part that is also referred to as a transition area between the western and eastern dialect area.

*Table 1*

**Speakers involved in the study**

| Gender | Speaker | Dialect area | Village | Time of recording | Age at the time of recording |
|---|---|---|---|---|---|
| female | F1 | western | Īra (> Kūolka) | 1986 | 82 |
| female | F2 | eastern | Sīkrõg | 1997 | 82 |
| female | F3 | eastern | Kuoštrõg | 1976 | 86 |
| female | F4 | eastern | Pitrõg | 1976 | 94* |
| female | F5 | eastern | Vaid | 1986 | 68 |
| male | M1 | eastern | Vaid** | 1983 | 86*** |
| male | M2 | eastern | Vaid | 1986 | 76**** |
| male | M3 | western | Īra (> Kūolka) | 1976 | 64 |
| male | M4 | eastern | Sīkrõg | 1986 | 77 |
| male | M5 | eastern | Kūolka | 1986 | 65 |

*or 92/93 years old; **or Īra; ***or 87; ****or 86

Altogether, 1342 mono-, di- and trisyllabic words were included in the study (see Table 2). Words with four syllables were left out as there were only two of them. The minimum number of words per speaker was 108 and the maximum number was 203. The selection of data was kept varied and consists of nouns, verbs, pronouns, adverbs, prepositions and conjunctions. The structure of the stressed syllable of the words was primarily long: a long open syllable (e.g., *nēḑi* [neːdʲi] ’them’, *pī’lõ* [piːˀlə] ’to stand’) or a long closed syllable (*võib* [vɯip] ’can’, *si’gžõ* [siˀgʒə] ’in autumn’). Our dataset also included words with short initial syllables and these words are without stød, such as *vanā* [vɑnɑˑ] ’old’. Studies have shown that in spontaneous speech, these words often exhibit an f0 peak at the start of the second syllable (e.g., Tuisk 2012), while in controlled speech, the f0 peak may occur early or late within the first syllable (e.g., Lehiste & Teras & Ernštreits & Lippus & Pajusalu & Tuisk & Viitso 2008). However, it has also been noted (Posti 1942) that a stød-like pitch contour appears in the long vowels of the second syllable in such words. We included these words as well to determine if their overall f0 contours have any effect on f0 of other types of words.

*Table 2*

**Distribution of the data**

| Speaker | Words | Non-stød | | Stød | |
|---|---|---|---|---|---|
| | | Number of words | % | Number of words | % |
| **Female** | Monosyllabic | 172 | 13 | 149 | 11 |
| | Disyllabic | 206 | 15 | 61 | 5 |
| | Trisyllabic | 33 | 2 | 6 | 0.5 |
| | | **411** | **30** | **216** | **16.5** |
| **Male** | Monosyllabic | 168 | 13 | 181 | 13 |
| | Disyllabic | 225 | 17 | 97 | 7 |
| | Trisyllabic | 38 | 3 | 6 | 0.5 |
| | | **431** | **33** | **284** | **20.5** |
| | **Total:** | **842** | **63** | **500** | **37** |

The data was manually segmented and acoustically analysed in Praat (Boersma & Weenink 2022—2023). Spontaneous speech was annotated (by the first author) at the following levels: (1) utterances (defined by pauses), (2) words within the utterances, (3) the voiced part of the target word (used for analysing the fundamental frequency pattern), (4) non-modal phonation period in words with stød (each period was audiovisually classified for non-modal phonation, such as creaky, tense, breathy/whispery, glottal stop), and (5) the number and structure of syllables in each word (e.g., *1LC2SO* denotes a disyllabic word with a **l**ong **c**losed **1**st syllable followed by a **s**hort **o**pen **2**nd syllable).

The fundamental frequency of mono- and disyllabic words was examined by taking a time series of f0 for each word.[1] Ten equidistant f0 measures per voiced part of a word were taken with pitch floor set to 65 Hz and pitch ceiling to 500 Hz. In the case of monosyllables, the voiced part of a word consisted of the syllable rhyme. In disyllables, the second syllable rhyme was included as well. Word-initial consonants were excluded from the time-series measures, as they are not part of the syllable rhyme. The f0 measures were statistically analysed using generalized additive models (GAMs) (similar methods are used, for example, by Kaland & Swerts & Himmelmann 2023). GAMs provide a way of analysing f0 as a smooth curve using time series data and do not assume linear relationships (Lin & Chang 1999). This should enable detection of a rising-falling tonal contour for words with stød, as described in the introduction. For trisyllabic words, the f0 contour analysis yielded no clear results, so we excluded these words from the f0 analysis. However, they were still examined for non-modal phonation.

In this study prosodic prominence was operationalized at the level of the utterance, with utterances being delimited by pauses. Any word to be used for analysis was classified according to three levels: the most prominent word in the utterance, words with default level of prominence or words that have lost stress and therefore were labelled as not prominent (or unstressed) words. The decision to assign three levels of prominence, rather than just two, was based on previous experience working with Livonian spontaneous speech. Livonian speech has more than just a binary distinction, showing slightly greater nuance in prominence. Therefore, the ternary distinction was selected. Dividing prominence into three levels was also essential for identifying the occurrence of non-modal phonation. The goal was to determine at which level of prominence non-modal phonation is most likely to appear.

With respect to the non-modal phonation of stød, a subjective audiovisual labelling was carried out by the first author in consultation with Gert Foget Hansen at The University of Copenhagen, who has specialised in the phonetics of Danish stød, including auditory and acoustic cues to phonation types. The period of non-modal phonation in mono-, di- and trisyllabic words with stød was annotated in a separate tier as (1) not present (lar_n), (2) maybe present (lar_m), and (3) present (lar_y). If some kind of non-modal phonation occurred, it was categorised accordingly, e.g., lar_y_creaky, lar_y_tense, lar_m_creaky, lar_m_tense. The phonetic descriptions of voice quality by John Laver (1980), Gordon & Ladefoged (2001) and Keating & Garellek (2015) were taken as

[1] The script is adapted from Anja Arnhold's webpage https://sites.ualberta.ca/~arnhold/praatScripts.html and modified by Nicolai Pharao.

a supportive guide. Acoustic properties that reflect different types of non-modal phonation are (1) irregular waveform, low irregular f0 and intensity for c r e a k y voice, (2) mid/high and regular f0 with a regular waveform for t e n s e voice, (3) irregular waveform and aperiodic energy for b r e a t h y or w h i s p e r y voice, and (4) complete glottal closure with an absence of vocal fold vibration for g l o t t a l s t o p (Laver 1980; Gordon & Ladefoged 2001; Keating & Garellek 2015).

Examples of different types of non-modal phonation from out data are presented in the figures below.[2]

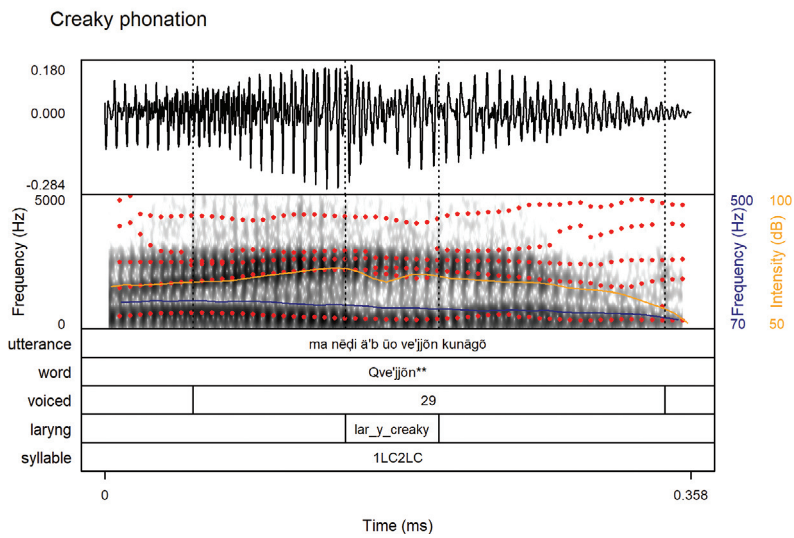


*Figure 1*. **Waveform and spectrogram of an example of the creaky phonation period pronounced by speaker M5. lar_y_creaky indicates the creaky period in the word** *ve'jjõn* **[veˀjjən] 'to fish, APP'.**

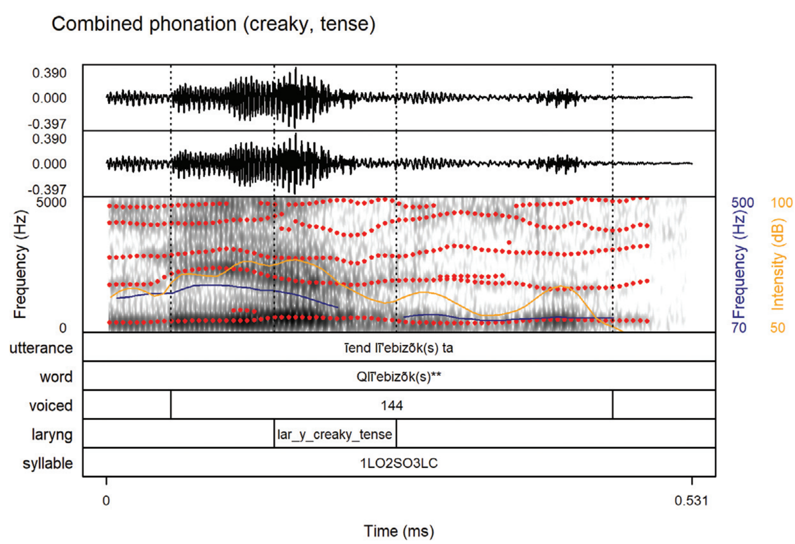


*Figure 2*. **Waveform and spectrogram of an example of the creaky and tense phonation period pronounced by speaker M3. lar_y_creaky_tense indicates the combined period in the word** *lī'ebizõks* **[liːˀebizəks] 'fat, InmSg'.**

[2] The figures were created with the R package praatpicture (Puggaard-Rode 2024).

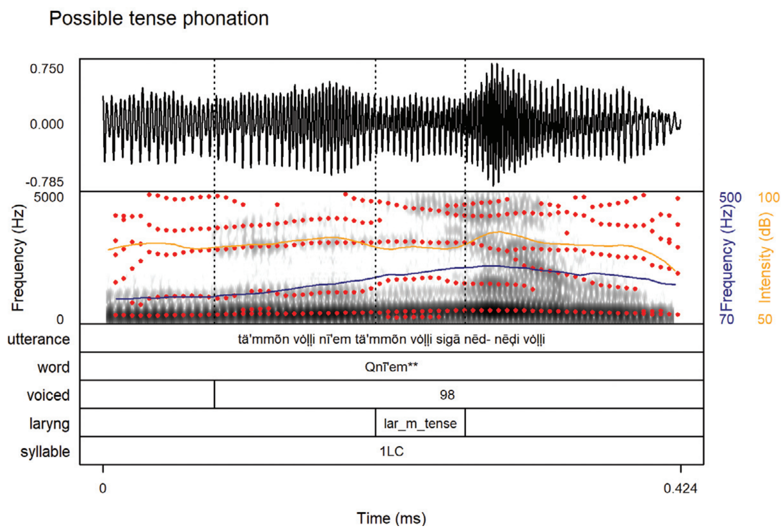


*Figure 3*. **Waveform and spectrogram of an example of the possible tense phonation period pronounced by speaker F1. lar_m_tense indicates the possible creaky period in the word *nī'em* [niːˀem] 'cow'.**

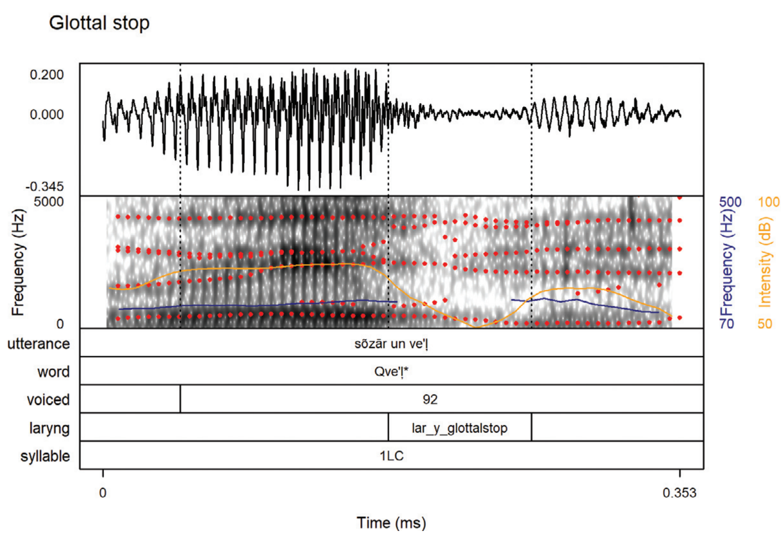


*Figure 4*. **Waveform and spectrogram of an example of the glottal stop pronounced by speaker F3. lar_y_glottalstop indicates the glottal stop in the word *ve'ļ* [veˀlʲ] 'brother'.**

## 4. Results

### 4.1. Fundamental frequency

To analyse the fundamental frequency contours associated with the stød contrast, GAMs were fit using the mgcv package (Wood 2011) in R (4.2.2). The ten equidistant time points were used as a smoothing predictor in an interaction with a factor Wordtype which indicated whether words had stød or not. Fundamental frequency measures were filtered to exclude outliers of more than two standard deviations from the mean for each point. The basic model was fit to monosyllabic and disyllabic words separately. The resulting smooths are shown in Figure 5 together with difference plots (both made using the itsadug package (van Rij & Wieling & Baayen & van Rijn 2022 in R). For each panel, the numbers on the x-axis indicate the point in the word, and the y-axis shows the fundamental frequency in hertz (with different ranges for men and women to highlight the tonal contours). The smoothed curves indicate the predicted fundamental frequency contours across speakers. Contours for men and women are shown separately. The smooth plots also show 95% confidence intervals indicating where there may be overlaps in the measures of the two contours. This is further illustrated in the difference plots accompanying the smooths. They show which intervals during the word can be deemed to be statistically significantly different, namely in areas where the confidence intervals do not include 0 difference. These intervals are also marked with red on the x-axis in the difference plots.

The findings for fundamental frequency show several interesting tendencies. As for contrast between words with stød and without stød, the only clear difference is found in disyllabic words in the speech of men (see also Tuisk & Pharao 2024). Note also that the contours for monosyllabic and disyllabic words without stød for the male speakers are (qualitatively) identical: a shallow rise towards the middle of the word followed by steeper fall. The results for the female speakers are somewhat surprising. In monosyllabic words, we would expect higher f0 and distinctive tonal movement for words with stød (as shown, for instance, by Vihman 1971). Instead, a nearly level contour was found in monosyllabic words with stød, and with a lower f0 as well. The non-stød monosyllables show a level-falling contour for the female speakers. The patterns for the monosyllables produced by the male speakers are similar to those of the female speakers, although there is no difference in average f0 between the stød and non-stød words. What is very surprising is that in disyllables, words with and without stød have a similar tonal contour — a somewhat steep fall — in the recordings of the female speakers. Comparison of female and male speakers shows that women's pronunciation differs from men's pronunciation of the disyllabic words. Men have a clearly rising-falling f0 curve in disyllabic words with stød, while women have a falling curve. The results remained consistent even after excluding words with short initial syllables from the dataset. This shows that in the case of our dataset, including words (without stød) with short initial syllables did not have any effect on the f0 contours in general. There is quite some interspeaker as well as intraspeaker variation in this data, which is examined more closely below.

### 4.1.1. Interindividual variation in tonal contours

The patterns observed for the men and the women who participated in the interviews indicate differences at the group level, notably that whereas the five women may have neutralised the difference between words with and without stød at least as far as the tonal contour is concerned, the men appear to retain a difference, at least in the disyllabic words. This generalisation covers some differences between the individuals and also does not take into account differences in the tonal contours associated with levels of prominence. Figure 6 below shows the smoothed contours for the disyllabic words for each female participant in the study with separate panels for three levels of prominence: none, normal and extra prominence.

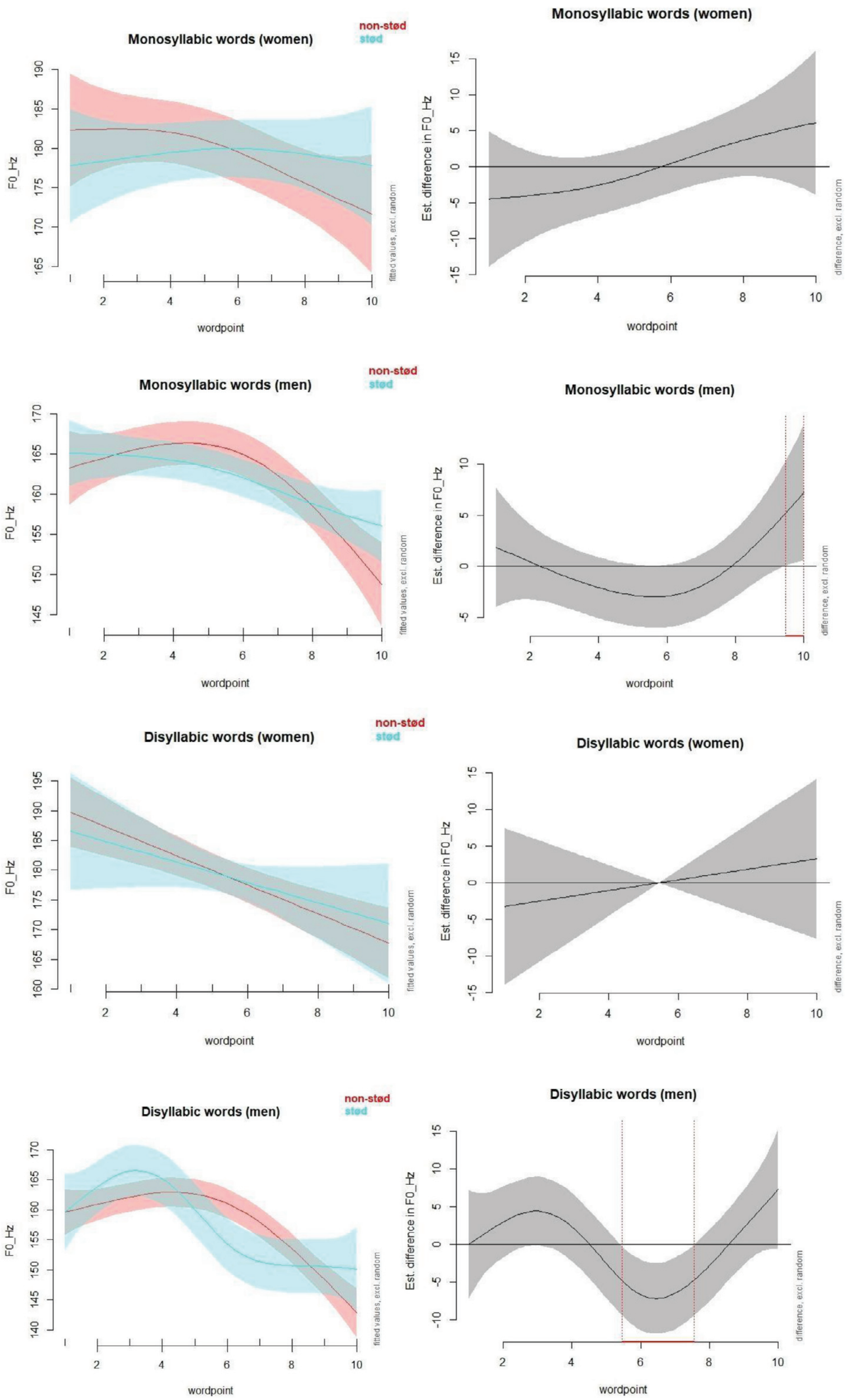


*Figure 5*. **Predicted fundamental frequency contours with 95% confidence intervals and difference plots for monosyllabic and disyllabic words with and without stød across speakers.**

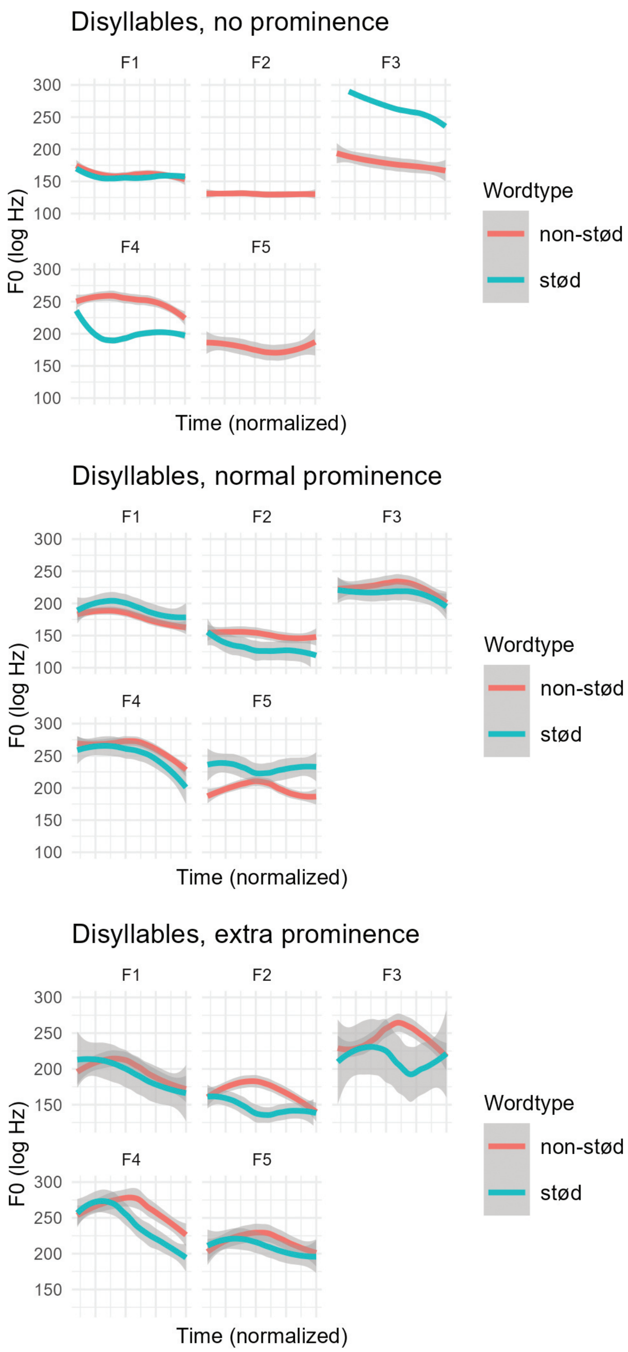


*Figure 6*. **Smoothed f0 contours with confidence intervals of the disyllabic words with three levels of prominence for female speakers.**

While there would appear to be no clear difference between words with and without stød for words with no or normal prominence, the two women F3 and F4 do exhibit the pattern with the tonal peak and a sharp fall in the first syllable in words with stød with extra prominence that was generally found for the men above in words with stød.

Below we look at the contours for men's productions of disyllabic words again with separate panels for separate levels of prominence (see Figure 7). As was the case with the five women, there is no (real) difference in contours for words with no prominence. The pattern with a peak in the first syllable and the lower tone in the second syllable in words with stød, compared to a slower rise to a peak in the second syllable in words without stød, which was also shown above, is clearest in the words with extra prominence with four of the five men (M1, M2, M3 and M5) exhibiting this pattern. Three of the speakers (M1, M2 and M4) also exhibit this pattern for words with stød under normal prominence.

Above, we focused on the disyllabic words, because they revealed the most interesting patterns at the group level. Below, we show the results for individuals for monosyllabic words, again as a function of level of prominence. In Figure 8, the tonal contours of the words produced by the five women are presented.

There does not appear to be any difference between monosyllabic words with and without stød when it comes to the tonal contour under no prominence. Speakers F1, F3 and F4 show similar contours for words with normal prominence. Speakers F3 and F5 are the only ones that show any apparent differences in tonal contour for monosyllabic words with extra prominence. However, note that the smooths for both have large confidence intervals indicating large variability in the measures and therefore not a certain indication of different contours. Note that the scale on the y axis of monosyllabic words with extra prominence is expanded compared to the figures of no prominence and normal prominence. This is because speaker F5 exhibited a wide range of fundamental frequency values.

Looking at the men we again find that there does not seem to be a discernible difference between words with and without stød when the monosyllabic words are unstressed, at least with respect to the tonal contour (see Figure 9 below). The evidence for the rise-fall pattern is not so clear, although speaker M2 does seem to produce it, particularly under extra prominence. Speaker M4 may also be said to have the rise-fall pattern in words with stød under normal prominence. Speaker M3 exhibits the most similar contours under all three prominence levels (with a contour rise at the end of words with stød with normal prominence).

Taken together, these analyses show that the stød contrast is absent when words have lost their stress. Furthermore, it appears that the expected tonal contour associated with stød (as based on previous studies) mainly occurs in extra prominent position in the utterance, and that even speakers who only rarely use the rise-fall tonal contour in words with stød may do so in the extra prominent condition.

### 4.2. Non-modal phonation

Words with stød were analysed with respect to phonation, specifically in order to investigate possible occurrence of non-modal phonation in these words. Non-modal phonation was established auditorily as well as on the basis of the waveform and spectrograms. The non-modal period appeared only on the (primary-)stressed syllables. The distribution between the presence, possible presence and lack of non-modal phonation in words with stød was 99 —

”yes” (lar_y), 175 — ”maybe” (lar_m), and 226 — ”no” (lar_n). Hence, in 45% of words with stød no period of non-modal phonation was detected. In 55% of words with stød, the presence or possible presence of a non-modal period was established. ”Possible presence” refers to instances where it was difficult to distinguish between non-modal phonation of stød and the natural creakiness or other non-modal sound of the voice of the speaker. The classification

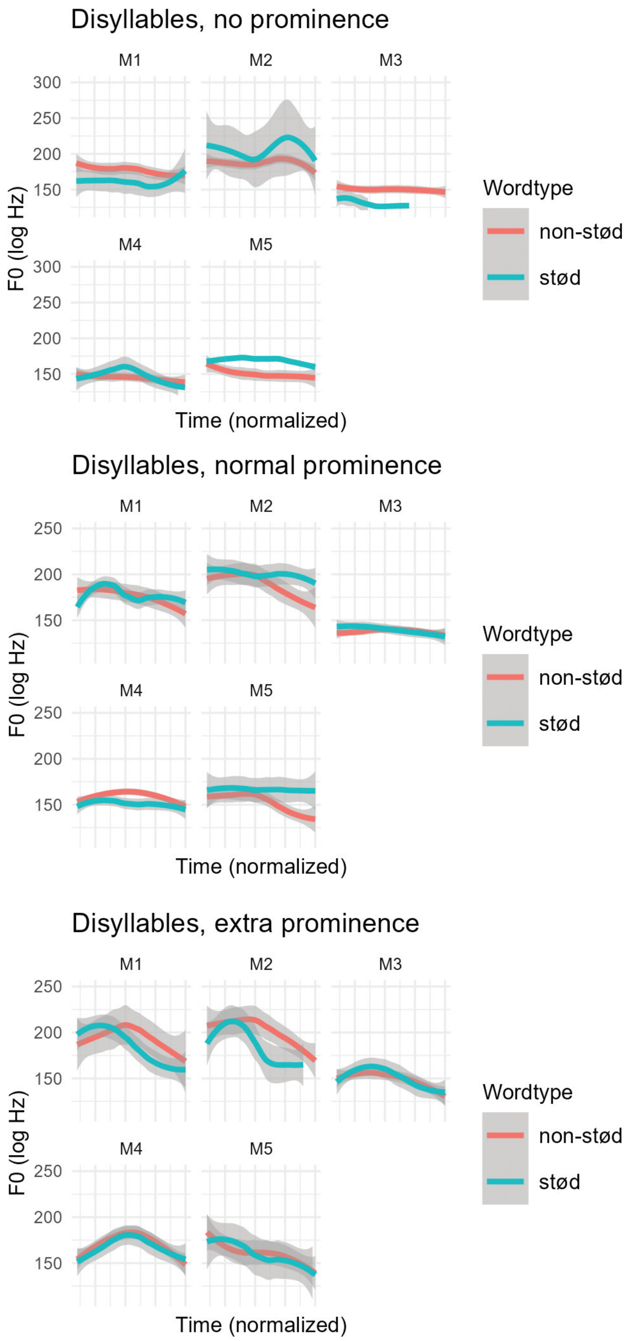


*Figure 7*. **Smoothed f0 contours with confidence intervals of the disyllabic words with three levels of prominence for male speakers.**

includes four types of non-modal phonation: (1) creaky, (2) tense, (3) breathy or whispery, and (4) glottal stop (see section 3). The distribution of non-modal phonation was analysed using Chi-Square tests. The independent variables for the analysis were the occurrence of non-modal phonation, individual speaker, syllable type, level of prominence and speaker's gender.

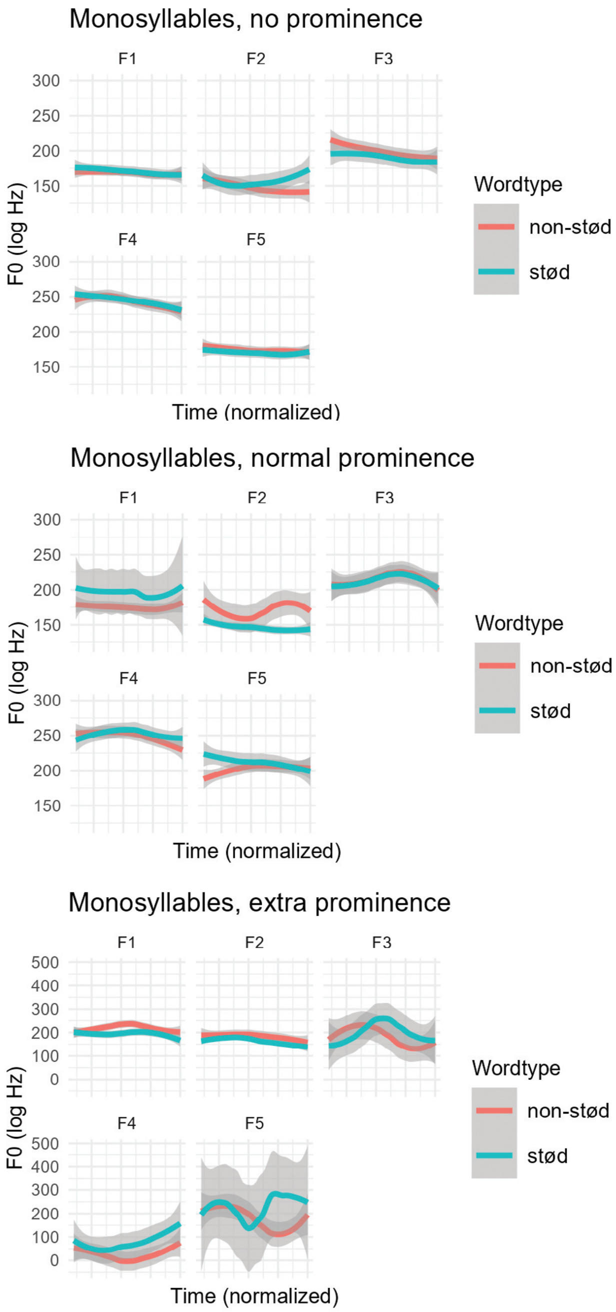


*Figure 8*. **Smoothed f0 contours with confidence intervals of the monosyllabic words with three levels of prominence for female speakers. The scale on the y axis is expanded compared to the other figures. This is due to the large range of fundamental frequency values obtained for speaker F5.**

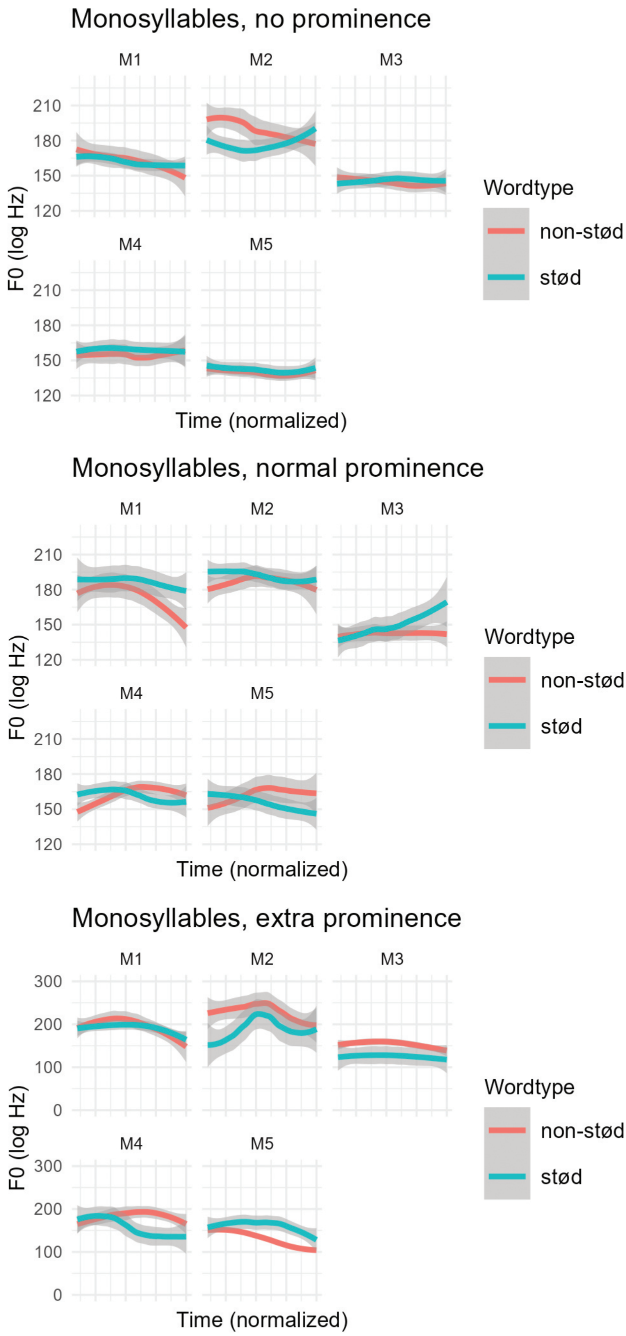


*Figure 9*. **Smoothed f0 contours with confidence intervals of the monosyllabic words with three levels of prominence for male speakers. Note that the scale on the y axis is expanded compared to the other figures. This is due to the higher average fundamental frequency values for speaker M2.**

This subset of the data consisted of 500 words with stød. Some kind of non-modal period was found in 274 of these words, always in syllables with primary stress. The distribution between the groups involving these tokens is presented in Table 3.

*Table 3*

**Classification of non-modal phonation**

| | creaky | tense | breathy | glottal stop | combined |
|---|---|---|---|---|---|
| lar_y | 45 | 37 | – | 2 | 15 |
| lar_m | 62 | 48 | 2 | – | 63 |
| **Total** | **107** | **85** | **2** | **2** | **78** |

Out of 274 actual and possible non-modal occurrences approximately 39% were classified as ”creaky” and 31% as ”tense”. In 28% of cases it was difficult to make a decision about the classification and they were labelled as ”combined”. Glottal stop occurred twice in the pronunciation of one female speaker in monosyllabic words *veʼļ* [veˀlʲ] ʼbrotherʼ and (*minā äʼb*) *lǟʼ* [læːˀ] ʼ(I donʼt) goʼ.

As for co-occurrence of non-modal phonation and tonal contour, there seems to be a tendency for them to interact. That is, if there was non-modal phonation phase then the tonal contour was falling or rising-falling. But in some cases, the tonal stød was audible without any non-modal phonation period. In rare cases, when the established non-modal period and rising-falling or falling contour was absent, the entire word was pronounced with a non-modal phonation.

The distribution of non-modal phonation with respect to gender is presented in Table 4.

*Table 4*

**Distribution of non-modal phonation**

| | **female** | | **male** | |
|---|---|---|---|---|
| | **Number of tokens** | **%** | **Number of tokens** | **%** |
| **lar_y** | 38 | 18 | 61 | 22 |
| **lar_m** | 80 | 37 | 95 | 33 |
| **lar_n** | 98 | 45 | 128 | 45 |
| **Total** | **216** | **100** | **284** | **100** |

Pearson’s Chi-Squared test was performed and Pearson’s residuals were calculated for each cell in the contingency table. Residuals between –1.96 and 1.96 are not statistically significant and suggest observed counts are close to expected counts. In our analysis, the Chi-Squared test and Pearson’s residuals show slightly different indications of statistical differences. This is likely because the Chi-Squared test evaluates the overall association between two categorical variables across the entire contingency table. The Pearson’s residuals, on the other hand, measure the contribution of each individual cell to the overall Chi-Squared statistic. We are cautious in interpreting the results related to residuals, as it is not always clear what is carrying the effect.

Pearson’s Chi-Squared test with Yates’ continuity correction showed a significant relationship between non-modal phonation and syllable type ($\chi 2$ (df = 1, N = 500), = 6.9667, $p = 0.008$). Pearson’s residuals are presented in Table 5.

While most Pearson’s residuals were close to zero, suggesting no substantial deviation from expected counts, the residual for ”open syllable”/”modal phonation” was −1.87, indicating a lower-than-expected count. The residual for ”open syllable”/”non-modal phonation” was 1.69, suggesting a higher-than-expected count. The residuals for closed syllables are close to zero, indicating that the

Table 5

**Pearson's residuals (phonation and syllable type)**

| Group | Category | Pearson residual |
|---|---|---|
| Closed syllable | Modal phonation | 0.83 |
| Closed syllable | Non-modal phonation | –0.76 |
| Open syllable | Modal phonation | –1.87 |
| Open syllable | Non-modal phonation | 1.69 |

observed counts align well with the expected counts. According to that, "non-modal phonation" and "possible non-modal phonation" are more likely to occur in long open syllables.

There was no effect of speaker gender on the presence/possible presence and lack of the non-modal phonation period ($\chi2$ (df = 1, N = 500), = 1, $p = 0.98$), i.e., the presence/possible presence and lack of the non-modal phonation was not related to the gender of the speaker. Pearson's residuals show that there is no evidence of a significant association between gender (female/male) and the categories modal and non-modal phonation based on these residuals (Table 6).

Table 6

**Pearson's residuals (phonation and speaker's gender)**

| Group | Category | Pearson residual |
|---|---|---|
| Female | Modal phonation | 0.04 |
| Female | Non-modal phonation | –0.03 |
| Male | Modal phonation | –0.03 |
| Male | Non-modal phonation | 0.03 |

There was an effect of individual speakers on the occurrence or absence of non-modal phonation ($\chi2$ (df = 9, N = 500), = 25.808, $p < 0.05$), indicating that some speakers were more likely to have a non-modal phonation. Pearson's residuals are presented in Table 7. They showed that most individuals exhibited residuals close to zero, suggesting little to no significant deviation. However, M3 showed a residual of –1.87 for category "modal phonation", which is near the significance threshold.

Table 7

**Pearson's residuals (phonation and individual speakers)**

| Individual | Residual for modal phonation | Residual for non-modal phonation |
|---|---|---|
| F1 | –0.49 | 0.44 |
| F2 | 1.33 | –1.21 |
| F3 | –0.28 | 0.26 |
| F4 | –1.18 | 1.08 |
| F5 | 1.04 | –0.95 |
| M1 | 1.33 | –1.21 |
| M2 | 1.63 | –1.48 |
| M3 | –1.87 | 1.70 |
| M4 | –1.26 | 1.14 |
| M5 | 0.25 | –0.23 |

As for prominence (no prominence vs. normal and extra prominence), there was a significant effect ($\chi2$ (df = 1, N = 500), = 16.26, $p < 0.001$), i.e., non-modal phonation was more likely to occur when the word had either normal or extra prominence. Pearson's residuals analysis revealed significant deviations in the "no prominence" row (see Table 8). According to that, if the word has no prominence, then it is more likely not to have non-modal phonation.

*Table 8*

**Pearson's residuals (phonation and prominence)**

| Group | Category | Pearson residual |
|---|---|---|
| Prominence | Modal phonation | –1.91 |
| Prominence | Non-modal phonation | 1.74 |
| No prominence | Modal phonation | 2.38 |
| No prominence | Non-modal phonation | –2.16 |

The findings indicate a strong association between the "no prominence" category and the observed frequencies. In contrast, the residuals for the "prominence" row did not exceed the significance threshold, indicating no substantial deviations from expected counts.

The average duration of the non-modal phonation (i.e. lar_y) in the syllable is 74 ms. The location of the period varied within the syllable.

In Table 9, findings for individual speakers are presented. The table shows whether a speaker had non-modal phonation in more than 50% of the stød words they produced or not. In addition, we looked at the level of prominence for the syllables with non-modal phonation. Female speakers F1, F3 and F4, and male speakers M3 and M4 had non-modal phonation in more than 50% of words with stød. Conversely, female speaker F2, along with male speakers M1 and M2, had non-modal phonation in less than 50% of the stød words they produced. Female speaker F5 and male speaker M5 had an approximately equal distribution of stød words with and without non-modal phonation. Concerning level of prominence, we found the following distribution: F1, F3, F4, F5, M2, M3, M4, and M5 predominantly exhibited non-modal phonation in words with normal or extra prominence. Speaker F2 tended to produce non-modal phonation periods equally across words regardless of level of prominence. Male speaker M1 appeared to be the least affected by any prominence.

*Table 9*

**Occurrences of non-modal phonation period in words with stød**

| Speaker | | Over 50% occurrences of non-modal phonation period | Over 50% occurrences of non-modal phonation period in words with normal or extra prominence |
|---|---|---|---|
| female | F1 | + | + |
| | F2 | – | +/– |
| | F3 | + | + |
| | F4 | + | + |
| | F5 | +/– | + |
| male | M1 | – | – |
| | M2 | – | + |
| | M3 | + | + |
| | M4 | + | + |
| | M5 | +/– | + |

## 5. Discussion

The findings concerning tonal patterns and non-modal phonation observed in spontaneous Livonian speech offer a basis for deeper exploration concerning the linguistic context of Livonian, particularly the interaction with Latvian. The results of this study highlight the usefulness of studying spontaneous speech to fully understand the phonetics of Livonian stød, including the importance of prominence and speaker diversity. The following discussion involves the sociolinguistic aspects and influence of Latvian, background of the speakers, non-modal phonation and the role of prominence in spontaneous vs. read speech.

The overall observations regarding tonal contours in disyllabic words from our dataset (Figure 5) indicate that female speakers exhibit the characteristic pattern of the Livonianized dialect of Latvian, i.e. a merged pattern with only the falling contour. This finding implies that these female speakers may have been influenced by speakers of the Latvian dialect. Courland Livonian has never been and is not completely homogeneous and contacts with Latvians were customary, so an influence can be expected. On the other hand, the five male speakers in our dataset demonstrated rising-falling f0 contours in disyllabic words. This phenomenon may be attributed to their backgrounds and social context. Specifically, all these men were actively involved in fishing during their youth. Fishing, along with agriculture and animal husbandry, constituted a major economic sector in Courland. Typically, the crews were composed of Livonian men, and it is highly probable that the language spoken aboard the boats was Livonian. Therefore, it is likely that these men retained the pronunciation they used while conversing during their fishing expeditions at sea.

Our data show variations in pronunciation among speakers, as well as within individuals. The factors influencing how individual speakers phonetically produce stød and why some maintain tonal distinctions remain somewhat ambiguous. Several speakers in our study have previously served as participants in previous Livonian stød research. Yet, comparing their controlled speech with their spontaneous speech yields conflicting findings. Specifically, female speakers exhibit rising-falling tonal patterns in controlled speech but not in spontaneous speech, whereas men typically maintain the rising-falling pattern in both types of speech. Tuisk & Teras (2009) conducted an analysis of disyllabic words with and without stød in spontaneous speech, focusing on stressed words positioned at the beginning and middle of sentences. Within this dataset, female speaker F3 and male speaker M4 (who also contributed to the present study) did not exhibit any laryngealization. However, our current dataset reveals that both speakers displayed some form of non-modal phonation in their pronunciation of words with stød. Discrepancies between the findings of the two studies may indicate the necessity for a larger dataset to draw comprehensive conclusions about individual speakers. The dataset in our current analysis involves more words in different positions in the utterance, encompassing monosyllabic and trisyllabic words as well, which allows for a broader analysis of the pronunciation of individual speakers. It should be noted that we are not the first people to find less use of locally marked forms in spontaneous speech than in word list materials. Stuart-Smith & Timmins & Tweedie (2007) found that working-class boys in Glasgow would produce more non-standard forms in the word list readings. They analyse it as performing working class identity in opposition to a field worker from a university. The difference between elicited and spontaneous speech described here for Livonian may be interpreted in a similar fashion: the reading of word lists for a researcher brings out stereotypical speech.

As for individuals, there is reason to believe that the background of each speaker can help to explain some of the patterns observed. For example, male speaker M4 in our study was a Livonian scholar and served as a long-term valuable informant for linguists. It is likely that he was conscious of the significance of stød and committed to preserving its traditional pronunciation. Also, female speakers F3 and F5 were long-term informants for many researchers who worked on Livonian and they tried to keep apart their Livonian and Latvian pronunciation.

There is limited information available concerning the manifestation of stød in the pronunciation of speakers with different dialectal backgrounds. Researchers have primarily been hindered by the scarcity of sound recordings from the western region of the Livonian coast, limiting their ability to compare them with recordings from the eastern Livonian villages. In phonetically transcribed texts, stød is generally indicated in all texts, but it remains unclear whether it reflects the actual pronunciation of the speaker or is marked to maintain consistency between texts from the western and eastern parts. The current data suggests that dialect background may play a role in the pronunciation of one speaker, who did not show a clear tonal contrast. The female speaker F1 is originally from the westernmost village Īra of the eastern Livonian Coast. This area has sometimes been considered to be a transition area or Central Livonian dialect area, because it shares common features with both eastern and western Livonian (Viitso 1999). There was no significant difference between the tonal contours in words with and without stød in her speech, regardless of level of prominence. This aligns with Kettunen's (1925 : 4) conclusion, which posited that there were dialectal variations in the expression of the Livonian stød, with its most pronounced occurrence observed in East Livonian. However, this explanation is not entirely supported by the pronunciation of the male speaker M3, who is from the same dialect area. He exhibited rising-falling tonal contour in disyllabic words with an extra prominence and falling contour in words with no or normal prominence. Also, both speaker F1 and speaker M3 demonstrated a higher occurrence of non-modal phonation than average, indicating that non-modal phonation serves as the primary distinguishing feature in their pronunciation of words with stød in spontaneous speech.

Next, the tonal pattern and non-modal phonation will be discussed. Both a rising-falling or falling tonal pattern and non-modal phonation have been identified as significant components of the Livonian stød. Furthermore, research has emphasised that the distinctive pitch contour serves as the most consistent indicator and may act as a sufficient cue for stød even in the absence of non-modal phonation. Studies have demonstrated that words with stød typically exhibit an early f0 peak within the initial third of the first syllable (Tuisk & Teras 2009). It has been noted that even if this peak occurs later, it still occurs earlier (at 61—66% into the word) than the delayed peak found in non-stød words. However, the current findings show that the rising-falling pitch pattern is not consistently present in spontaneous speech. Instead, a predominantly falling or level pitch contour can occur without any discernible f0 peak, especially in monosyllabic words. The results suggest that the rising-falling contour occurs more in controlled speech. Our results also do not support Lauri Posti's claim that there are three tones in Livonian: rising, broken (stød) and falling (e.g., Posti 1936, 1942). According to him, falling tone could not be treated as a broken tone and he remained convinced that there is a special falling tone in Livonian (Posti 1937a; 1937b). In his study with Aarni Penttilä (Penttilä & Posti 1941) only rising and broken tones are discussed, but this was probably due to unprocessed materials.

The non-modal phonation observed in long syllable nuclei in words with stød is also recognized as one of the distinctive characteristics of Livonian stød. Earlier research has commonly labelled this phenomenon as laryngealization, glottalization, or creaky voice. It has been indicated that it occurs more frequently in controlled speech than in spontaneous speech (e.g., Teras & Tuisk 2009; Tuisk & Teras 2009), but datasets analysed from individual speakers have indicated that it can also occur less frequently in controlled speech (e.g., Vihman 1971). Our current data, obtained from the spontaneous speech of ten speakers, show that for some speakers a distinctly delineated non-modal phonation phase was identified in fewer than 50% of words with stød. The significant occurrence of potential laryngealization suggests that non-modal phonation occurs in natural speech more often than expected. We also note that even in cases where it was impossible to determine a fixed non-modal phonation phase in the speech of some speakers, the entire word with stød was pronounced with non-modal phonation.

Additionally, it is apparent that non-modal phonation does not fall into a singular category. While the subjective nature of this evaluation may be a point of contention, it still offers a solid foundation for a more thorough examination of phonation in similar contexts in the future. Livonian non-modal phonation encompasses not only creaky voice but very likely also other qualities. Regarding future research in this area, one could propose the hypothesis that there exist a greater variety of categories for non-modal phonation in spontaneous speech compared to controlled speech. In the current study, we decided not to use spectral tilt measures related to voice quality and leave these for future studies. Spectral tilt measures are sensitive to many aspects of recording and can be influenced by differences in vowel quality which would be difficult to control for.

Previous research on Livonian stød has explored how stød is manifested as a complete closure, sometimes debating whether it represents its most defining realisation of stød (e.g., Kettunen 1925; de Sivers 1965). Our general conclusion is that the phonetic expression of stød as a glottal stop is infrequent in speech and tends to occur more often in controlled speech. For instance, Vihman (1971 : 293) noted that in her Livonian data of isolated words from one female speaker, the structure where stød was consistently observed as a complete glottal closure predominantly comprised mono- and disyllabic words with short vowel + obstruent forms. In our dataset, we observed the glottal closure in just two monosyllabic words featuring a short vowel + sonorant, as well as a long monophthong (*veʼļ* [veˀlʲ] ʼbrotherʼ and (*minā äʼb*) *lǟʼ* [læːˀ] ʼ(I donʼt) goʼ). In the speech of the female speaker F3, both words appeared in the final position of the utterance and were pronounced with normal and extra prominence. Although based on only two occasions, this suggests that the occurrence of the glottal stop in this context is not primarily determined by syllable structure, but rather by the level of prominence on the word and also the position of the word in the utterance.

The findings of the current study confirm that the duration of the non-modal phonation phase in words with stød varies and it is not possible to predict its exact location in the word. This is in line, for example, with Vihman's and Lehiste's findings from Livonian controlled speech (Vihman 1971; Lehiste & Teras & Ernštreits & Lippus & Pajusalu & Tuisk & Viitso 2008) as well as the findings of Grønnum & Vazquez-Larruscaín & Basbøll on Danish (Grønnum & Vazquez-Larruscaín & Basbøll 2013). Regarding the location of the non-modal phonation, previous studies on controlled speech have generally found a fixed alignment of this period with the peak of the pitch curve. This is not what we have found in our data from spontaneous speech. The location can vary within the syllable and appears to be unrelated to the timing of the f0 peak.

Finally, it is essential to highlight that the data analysed in this study are derived from spontaneous speech. Controlled speech data may offer clearer stød patterns, but this is likely because previous research based on Livonian spontaneous speech has examined a limited set of target words spoken by an even smaller number of individuals. These studies mainly provide insights into the general level of prominence, which may include accentuation or stress. For instance, Tuisk (2015) investigated the spontaneous speech of six Livonian speakers, and found an early fall in tonal contour for disyllabic words with stød in stressed position and a delayed fall for those without stød across all speakers. However, our research indicates that tonal patterns are more nuanced, as tone contours can vary across different levels of prominence. Lehiste & Teras & Ernštreits & Lippus & Pajusalu & Tuisk & Viitso (2008) examined disyllabic unstressed words from the spontaneous speech of the male speaker M5, whose speech is also analysed in our current study. In their investigation, M5 demonstrated an early peak in fundamental frequency and a falling tone in both words with and without stød. Conversely, our analysis of the same speaker reveals a consistent level contour with no significant peak for words with stød, and a falling or level-falling contour for words without stød in instances of normal or no prominence at all. The results derived from our present data show that prominence plays a role in the realisation of stød. When a word in the utterance was articulated with prominence, there was a greater tendency for non-modal phonation or a rising-falling tonal contour to manifest.

## 6. Conclusions

This paper presents research investigating the phonetic manifestation of the Livonian stød, based on the most extensive amount of acoustic data collected from spontaneous speech of five female and five male speakers. The fundamental frequency contours associated with the stød contrast were analysed using time series data and Generalized Additive Models. A non-modal phonation associated with words with stød were observed and the role of prominence was discussed.

The general results indicate that tonal distinction between words with and without stød seems to be mostly eliminated in spontaneous conversation. The only clear and consistent tonal contrast was found in disyllabic words in the speech of men, while women had neutralised the distinction between words with and without stød. The dominant language, Latvian, is likely to have influenced the pronunciation of the female speakers, at least in terms of tone.

In the Livonianized dialect of Latvian, spoken in Courland and Northern Vidzeme, the falling and broken tone have sometimes merged into the single falling tone.

As for prominence, the contrast between words with and without stød is most apparent when the words have extra prominence and seem to be categorically absent in words with no stress. Concerning the non-modal phonation observed in words with stød, it is more likely to occur when the word carries a certain level of prominence. The most common non-modal phonation period categories in words with stød were ”creaky” and ”tense”. The glottal stop variant is very rare in Livonian spontaneous speech.

The acoustic and auditory analyses provided herein underscore a following observation: Livonian stød cannot be simply defined by two or three stable characteristics. Beyond its phonological intricacy, various conditions must be considered for stød to manifest. Its realisation is influenced by factors such as speech style (spontaneous vs. read speech), prominence, the pronunciation of individual speakers and the influence of Latvian, as well as the interplay among these elements. While spontaneous speech often displays greater variability compared to read speech, it can be deduced that certain characteristics may neutralise in spontaneous speech. As for non-modal phonation, this study charts the course for some future lines of inquiry.

**Acknowledgements.** The publication costs of this article were covered by the Estonian Academy of Sciences.

This study was supported by the Estonian Research Council grant (project PUTJD1107 ”Stød in the languages around the Baltic Sea”) and partly (discussion part) by the project “Improving access to a critically under-resourced language: AI-based approaches for producing and obtaining Livonian content” financed by the Recovery and Resilience Facility/NextGeneration EU (LU-BA-PA-2024/1-0056). We would like to thank Gert Foget Hansen for his valuable guidance in the analysis of phonation. We would also like to acknowledge two reviewers for their insightful comments and suggestions. All remaining mistakes are entirely our own.

**Addresses**

Tuuli Tuisk
University of Copenhagen, University of Latvia, University of Tartu
tuuli.tuisk@ut.ee

Nicolai Pharao
University of Copenhagen
nicolaip@hum.ku.dk

## Abbreviations

**APP** — Active Past Participle; **f0** — Fundamental Frequency; **HL** — High-Low; **InmSg** — Instrumental Singular; **PPl** — Partitive Plural; **PSg** — Partitive Singular.

**LvdaF 2013** — Latviešu valodas dialektu atlants. Fonētika. Apraksts, kartes un to komentāri. Rīga: Latvijas Universitātes Latviešu valodas institūts, 2013; **LVG 2013** — Latviešu valodas gramatika, Rīga: Latvijas Universitātes Latviešu valodas institūts, 2013.

## REFERENCES

B a l o d e, Laimute & H o l v o e t, Axel 2001, The Lithuanian language and its dialects. — Circum-Baltic Languages. Volume 1: Past and Present, Amsterdam—

Philadelphia: John Benjamins (Studies in Language Companion Series 54), 41—79.

Balodis, Uldis 2018, Non-modal phonation associated with stød vowels in Livonian. — ESUKA 9 (2), 165—187.

Boersma, Paul & Weenink, David 2022—2023, Praat: doing phonetics by computer [Computer program]. http://www.praat.org/.

Cangemi, Francesco & Baumann, Stefan 2020, Integrating phonetics and phonology in the study of linguistic prominence. — Journal of Phonetics 81, 1—6.

Cucchiarini, Catia & Strik, Helmer & Boves, Lou 2002, Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. — Journal of the Acoustical Society of America 111 (6), 2862—2873.

Dogil, Grzegorz 1999, Baltic languages. — Word Prosodic Systems in the Languages of Europe, Berlin—New York: Mouton de Gruyter (Empirical Approaches to Language Typology [EALT]), 877—896.

Ernštreits, Valts 2012, Livonian in the 21st century. — Les langues finno-ougriennes aujourd'hui I, Paris (Études finno-ougriennes 44). https://doi.org/10.4000/efo.675.

Ernštreits, Valts & Kļava, Gunta 2014, Grammatical changes caused by contact between Livonian and Latvian. — ESUKA 5 (1), 77—90. https://doi.org/10.12697/jeful.2014.5.1.05.

Esposito, Christina M. 2004, Santa Ana del Valle Zapotec phonation. — UCLA Working Papers in Phonetics 103, 71—105.

Fischer-Jørgensen, Eli 1989, A Phonetic Study of the Stød in Standard Danish, Turku: University of Turku.

Goldshtein, Yonatan 2023, Outline of a Prosodic Theory of Stød. PhD Thesis, Aarhus University.

Gordon, Matthew & Ladefoged, Peter 2001, Phonation types: a cross-linguistic overview. — Journal of Phonetics 29 (4), 383—406.

Grønnum, Nina 2022, Three quarters of a century of phonetic research on common Danish stød. — Nordic Journal of Linguistics, 1—32. https://doi.org/10.1017/S0332586522000178.

Grønnum, Nina & Basbøll, Hans 2001, Consonant Length, Stød and Morae in Standard Danish. — Phonetica 58 (4), 230—253. https://doi.org/10.1159/000046177.

—— 2007, Danish stød: phonological and cognitive issues. — Experimental Approaches to Phonology, Oxford: Oxford University Press, 192—206.

Grønnum, Nina & Vazquez-Larruscaín, Miguel & Basbøll, Hans 2013, Danish Stød: Laryngealization or Tone. — Phonetica 70 (1—2), 66—92. https://doi.org/10.1159/000354640.

Grünthal, Riho & Heyd, Volker & Holopainen, Sampsa & Janhunen, Juha & Khanina, Olesya & Miestamo, Matti & Nichols, Johanna & Saarikivi, Janne & Sinnemäki, Kaius 2022, Drastic demographic events triggered the Uralic spread. — Diachronica. Volume 39, Issue 4, Aug 2022, 490—524. https://doi.org/10.1075/dia.20038.gru.

Kaland, Constantijn & Swerts, Marc & Himmelmann, Nicolaus P. 2023, Red and blue bananas: Time-series f0 analysis of contrastively focused noun phrases in Papuan Malay and Dutch. — Journal of Phonetics. Volume 96, January 2023. https://doi.org/10.1016/j.wocn.2022.101200.

Kariņš, Arturs Krišjānis 1996, The Prosodic Structure of Latvian. PhD dissertation, University of Pennsylvania.

Keating, Patricia & Garellek, Marc 2015, Acoustic analysis of creaky voice. https://www.researchgate.net/publication/270760538_Acoustic_analysis_of_creaky_voice.

Kettunen, Lauri 1925, Untersuchung über die livische Sprache I. Phonetische Einführung. Sprachproben, Tartu (ACUT B VIII).

Kiparsky, Paul 2017, Livonian stød. — Segmental Structure and Tone, Berlin—Boston: Walter De Gruyter, 195—210.

Koptjevskaja-Tamm, Maria 2006, The Circle That Won't Come Full: Two Potential Isoglosses in the Circum-Baltic Area. — Linguistic Areas. Conver-

gence in Historical and Typological Perspective, London: Palgrave Macmillan UK, 182—226. https://doi.org/10.1057/9780230287617_8.
Koptjevskaja-Tamm, Maria & Wälchli, Bernhard 2001, The Circum-Baltic languages: An areal-typological approach. — Circum-Baltic Languages. Typology and Contact. Volume 2: Grammar and Typology (Studies in Language Companion Series 55), Amsterdam—Philadelphia: John Benjamins, 615—750. https://doi.org/10.1075/slcs.55.
Laan, Gitta P. M. 1997, The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. — Speech Communication 22, 43—65.
Labov, William 2001, Principles of Linguistic Change. Social Factors, Malden, MA: Blackwell Publishers, Inc. (Language in Society 29).
Lang, Valter 2018, Läänemeresoome tulemised, Tartu: Tartu Ülikooli Kirjastus (Muinasaja teadus 28).
Laua, Alīse 1984, Latviešu literārās valodas fonētika, Rīga: Zvaigzne.
Laver, John 1980, The Phonetic Description of Voice Quality, Cambridge: Cambridge University Press (Cambridge Studies in Linguistics 31).
Lehiste, Ilse & Teras, Pire & Ernštreits, Valts & Lippus, Pärtel & Pajusalu, Karl & Tuisk, Tuuli & Viitso, Tiit-Rein 2008, Livonian Prosody, Helsinki (MSFOu 255).
Lehiste, Ilse & Teras, Pire & Pajusalu, Karl & Tuisk, Tuuli 2007, Quantity in Livonian: Preliminary Results. — LU XLIII, 29—44.
Levin, Harry & Schaffer, Carole A. & Snow, Catherine 1982, The Prosodic and Paralinguistic Features of Reading and Telling Stories. — Language and Speech 25(1), 43—54. https://doi.org/10.1177/002383098202500104.
Lin, Xihong & Zhang, Daowen 1999, Inference in generalized additive mixed models by using smoothing splines. — Journal of the Royal Statistical Society. Series B: Statistical Methodology 61 (2), 381—400. https://doi.org/10.1111/1467-9868.00183.
Lindström, Liina & Lippus, Pärtel & Tuisk, Tuuli 2019, The online database of the University of Tartu Archives of Estonian Dialects and Kindred Languages and the Corpus of Estonian Dialects. — Plurilingual Finnic. Change of Finnic languages in a multilinguistic environment, Helsinki (Uralica Helsingiensia 14), 327—350.
Morén, Bruce 2005a, Danish Stød and Eastern Norwegian Pitch Accent: The Myth of Lexical Tones. Presented at 13th Manchester Phonology Meeting (MFM13), Manchester, England.
—— 2005b, Danish Stød and Eastern Norwegian Pitch Accent: Prosody, Morphology and Non-tonal Lexical Specification. Presented at 11. møte om norsk språk — 11MONS, Bergen, Norway.
Pajupuu, Hille & Viitso, Tiit-Rein, 1986, Livonian polyphthongs. — Estonian Papers in Phonetics 1984—1985, Tallinn, 96—130.
Penttilä, Aarni & Posti, Lauri 1941, Über die steigende und sog. Stoss-intonation im Livischen. — FUF XXVII, 235—272.
Posti, Lauri 1936, Liivin kielen intonaatioista. — Vir., 314—326, 384.
—— 1937a, Lisää liivin intonaatioista. — Vir., 189—200, 263.
—— 1937b, Vieläkin liivin intonaatioista. — Vir., 448—451.
—— 1942, Grundzüge der livischen Lautgeschichte, Helsinki (MSFOu 85).
Puggaard-Rode, Rasmus 2024, praatpicture. A library for making flexible Praat Picture-style figures in R. — Proceedings of the 13th International Seminar on Speech Production (ISSP 2024), Autrans, 145—148.
Riad, Tomas 2000, The origin of Danish stød. — Analogy, Levelling, Markedness. Principles of Change in Phonology and Morphology, Berlin—New York: Mouton de Gruyter (Trends in Linguistics. Studies and Monographs [TiLSM] 127), 261—300.
van Rij, Jacolien & Wieling, Martijn & Baayen, Harald R. & van Rijn, Hedderik 2022, itsadug. Version 2.4.1. Interpreting Time Series and Autocorrelated Data Using GAMMs. https://cran.r-project.org/web/packages/itsadug/itsadug.pdf.
Rudzīte, Marta 1994, Latviešu un lībiešu valodas savstarpējā ietekme. — Lībieši. Rakstu krājums, Rīga: Zinātne, 288—319.

—— 1996, Latviešu un lībiešu valodas kontaktu atspulgi. — Baltu filoloģija VI. Rakstu krājums, Rīga: Latvijas Universitāte, 3—7.

Shriberg, Elizabeth 2005, Spontaneous Speech: How People Really Talk, and Why Engineers Should Care. — Proceedings of the 9th European Conference on Speech Communication and Technology 2005, Lisboa, 1781—1784.

de Sivers, Fanny 1965, Le coup de glotte en letton et en live. — Ceļi XII, 51—56.

Stolz, Thomas 1991, Sprachbund im Baltikum? Estnisch und Lettisch im Zentrum einer sprachlichen Konvergenzlandschaft, Bochum: Brockmeyer (Bochum-Essener Beiträge zur Sprachwandelforschung XIII).

Stuart-Smith, Jane, & Timmins, Claire & Tweedie, Fiona 2007, 'Talkin' Jockney'? Variation and change in Glaswegian accent. — Journal of Sociolinguistics 11 (2), 221—260. https://doi.org/10.1111/j.1467-9841.2007.00319.x.

Suhonen, Seppo 1982, Über die Quantität der livischen Vokale. — Voces amicorum Sovijärvi. In Honorem Antti Sovijärvi septuagesimum annum agentis die XXII mensis aprilis anno MCMLXXXII, Helsinki (MSFOu 181), 295—306.

Teras, Pire & Tuisk, Tuuli 2009, The characteristics of stød in Livonian. — Proceedings of the Xth Conference, Helsinki 2008, Frankfurt: Peter Lang Verlag, 217—226.

Terken, Jacques & Hermes, Dik 2000, The Perception of Prosodic Prominence. — Prosody: Theory and Experiment. Studies Presented to Gösta Bruce, Dordrecht: Springer (Text, Speech and Language Technology 14), 89—127. https://doi.org/10.1007/978-94-015-9413-4_5.

Thomsen, Vilhelm 1890, Beröringer mellem de finske og de baltiske (litauisk-lettiske) Sprog, København.

Tuisk, Tuuli 2012, Tonal and Temporal Characteristics of Disyllabic Words in Spontaneous Livonian. — LU XLVIII, 1—11.

—— 2014, Survey of research on Livonian prosody. — ESUKA 5 (1), 261—292. https://doi.org/10.12697/jeful.2014.5.1.15.

—— 2015, Acoustics of stød in Livonian. — Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow: University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0356.pdf.

Tuisk, Tuuli & Pajusalu, Karl 2022, A century of Livonian studies at the University of Tartu. — ESUKA 13 (1), 307—333. https://doi.org/10.12697/jeful.2022.13.1.12.

Tuisk, Tuuli & Pharao, Nicolai 2024, Time-series f0 analysis of contrastive tones: the case of Livonian natural speech. Conference paper accepted for the Speech Prosody 2024 conference in Leiden.

Tuisk, Tuuli & Teras, Pire 2009, The Role of Duration Ratios and Fundamental Frequency in Spontaneous Livonian. — LU XLV, 241—252.

Vihman, Marilyn May 1971, Livonian Phonology, with an Appendix on Stød in Danish and Livonian. Dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Linguistics in the Graduate Division of the University of California, Berkeley.

Viitso, Tiit-Rein 1999, Keskliivi. — Õdagumeresoomõ veeremaaq. Läänemeresoome perifeeriad, Võro (Võro Instituudi Toimõtiseq 6), 158—169.

Wagner, Petra & Origlia, Antonio & Avesani, Cinzia & Christodoulides, George & Cutugno, Francesco et al. 2015, Different part of the same elephant: A roadmap to disentangle and connect different perspectives on prosodic prominence. — Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow: University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0202.pdf.

Wälchli, Bernhard 2011, The Circum-Baltic languages. — The Languages and Linguistics of Europe. A Comprehensive Guide, Berlin—Boston: De Gruyter Mouton, 325—340. https://doi.org/10.1515/9783110220261.325.

Wiik, Kalevi 1989, Liivin katko, Turku: Turun Yliopisto.

Winkler, Eberhard 1999, Katkeintonatsiooni tekkimisest läänemeresoome keeltes. — Õdagumeresoomõ veeremaaq. Läänemeresoome perifeeriad, Võro (Võro Instituudi Toimõtiseq 6), 201—206.

—— 2000, Zum Stoßton im Ostseefinnischen. — Aspekte baltischer Forschung, Essen (Ernst-Moritz-Arndt-Universität Greifswald. Schriften des Instituts für Baltistik 1), 344—363.

W o o d, Simon N. 2011, Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. — Journal of the Royal Statistical Society. Series B: Statistical Methodology 73 (1), 3—36. https://doi.org/10.1111/j.1467-9868.2010.00749.x.

*ТУУЛИ ТУЙСК* (Копенгаген—Рига—Тарту), *НИКОЛАЙ ФАРАО* (Копенгаген)

## ТОНАЛЬНЫЕ КОНТРАСТЫ В БАЛТИЙСКОМ РЕГИОНЕ: ИЗУЧЕНИЕ СТЁДА В ЛИВСКОЙ СПОНТАННОЙ РЕЧИ

В статье представлены данные о тональном контрасте и фонетических различиях между словами со стёдом и без него в спонтанной речи ливов. В ливском языке различают два контрастных фонологических тона: ломаный тон, или стёд, и простой тон. Стёд в некоторых отношениях похож на датский стёд и считается частью тоновой системы фонологий языков Балтийского региона. Результаты исселдования показывают, что тональный контраст между словами со стёдом и без него имеет тенденцию нейтрализоваться в спонтанной речи ливов, но существуют индивидуальные различия между говорящими, а также различия между мужчинами и женщинами. Наиболее распространенные немодальные категории периода фонации в словах со стёдом — скрипучие и напряженные. Результаты также свидетельствуют, что стёд исчезает, если слово не имеет заметного значения.

*TUULI TUISK* (Kopenhaagen—Riia—Tartu), *NICOLAI PHARAO* (Kopenhaagen)

## TOONIVASTANDUS LÄÄNEMERE REGIOONIS: KATKETOON LIIVI KEELE SPONTAANSES KÕNES

Artiklis analüüsitakse toonivastandust ja fonatsiooni erinevusi liivi spontaanse kõne püsi- ja katketooniga sõnades. Liivi keeles eristatakse kaht kontrastset fonoloogilist tooni: püsitooni ja katketooni ehk *stød*'i. Katketoon sarnaneb teatud mõttes taani keele *stød*'iga ja see on osa Läänemere piirkonna keelte toonisüsteemidest. Meie uurimistulemused näitavad, et liivi spontaanses kõnes kipub katke- ja püsitooniga sõnade tonaalne kontrast neutraliseeruma. Samal ajal esineb hääldususes nii kõnelejate individuaalseid erinevusi kui ka meeste ja naiste vahelisi erinevusi. Katketooniga sõnade mittemodaalse fonatsiooni perioodi iseloomulikumad kategooriad on kärisev ja pingne või pingul fonatsioon. Tulemused näitavad ka seda, et kui sõna pole lausungis rõhuline või prominentne, siis katketooni iseloomulikud tunnused kaovad.