ARTICLE

Three quarters of a century of phonetic research on common Danish stød

Nina Grønnum

Department of Nordic Studies and Linguistics, University of Copenhagen, Emil Holms Kanal 2, 2300 Copenhagen S., Denmark
Email: ninag@hum.ku.dk

Abstract

Despite many attempts, we do not yet know how to answer the question exhaustively and adequately: What is common Danish stød phonetically? Specifically, is there an underlying physiological constancy behind the considerable acoustic variability? And is there a common denominator to the different acoustic manifestations that might explain their perceptual equivalence? Systematic acoustic, physiological, and perceptual investigations of common Danish stød now span (a little more than) three quarters of a century. This is a review of that research, to determine what may be considered established phonetic facts about common Danish stød and the open questions that remain for future research.

1. Introduction

This is a review of literature reporting experiments on the physiology, acoustics, and perception of common Danish stød, published during the past three quarters of a century.

The term COMMON DANISH STØD is a translation of FÆLLESDANSK STØD – as opposed to VESTJYSK STØD 'West Jutlandic stød'. When specification is not required, common Danish stød is rendered as STØD henceforth. It is found everywhere in Denmark north of a wavy line, the STØD BORDER, running from the west, between the islands Rømø and Mandø in the Wadden Sea, across Southern Jutland, South Funen, and the southernmost part of Zealand to the town of Præstø in southeast Zealand (Ringgaard 1973:22). Note, however, that stød is not distributed in the same manner in different Danish regions.

Some dialects south of the stød border have a tonal accent, as for instance on Rømø and Ærø. Others have no accent at all, whether tonal or stød, as for instance on Lolland and Bornholm (Ringgaard 1973).

Some authors merely characterize their speakers as Danish, others distinguish between speakers from Copenhagen and those who speak Standard Danish from

a regional background. The only dialect study is that of Ringgaard (1960), discussed in Section 4.1.4.

Section 2 is a brief description of the phonetic and phonological properties of stød as they are generally recognized today. Section 3 recounts a few older phonetic descriptions and experiments that serve to put into perspective the twenty-two modern investigations recounted in Section 4. Section 5 is a summary, and Section 6 contains a discussion.

2. Properties of common Danish stød

Stød is a challenging subject, partly due to its phonetic properties, as evidenced in Sections 3 and 4, but partly also due to the role it plays in the phonology of Danish.

2.1 Phonetic characteristics

The PROTOTYPICAL manifestation of stød is generally described as a particular kind of creaky voice or laryngealization. That is, vocal fold vibrations with irregular variation in periodicity as well as amplitude, commonly associated with a decrease of fundamental frequency (F_0) and intensity toward the end of the syllable. However, the manifestation may be weaker and resemble modal voice: merely a somewhat compressed voice quality, lacking vibratory irregularity and F_0 perturbation at the end. Stød is perceptible in whisper but disappears in song, at least in *bel canto* style singing. To set it apart from creaky voice or laryngealization as described in other languages, the manifestation of stød is termed CREAKINESS here. It is very reminiscent of Livonian stød, though structurally different (Kiparsky 2017). However, Livonian has become extinct since the last speaker passed away in 2013 (Charter 2013). Danish stød may therefore now be unique among the languages of the world.

Stød is conventionally marked with a superscript glottal stop symbol – after the vowel if it is long, as in [þeːˀn] *ben* 'leg', or else after the first postvocalic sonorant consonant, as in [benˀ] *bind* 'bandage' – even though stød is not, except in its strongest manifestation as under emphasis, produced as a glottal stop.

2.2 Phonological characteristics

Stød is a PROPERTY OF SYLLABLES. For a syllable to have stød it must satisfy two phonological conditions, one segmental, one prosodic. (i) It requires a long sonority rhyme, either a long vowel – as in [seːˀ] *se* 'see' – or else a short vowel followed by a sonorant consonant – as in [halˀ] *hal* 'hall'. In either case, one or more consonants may follow in the syllable – as in [seːˀn] *sen* 'late' and [halˀm] *halm* 'straw'. (ii) The syllable may not be unstressed. With long sonority rhymes and stress (whether primary or secondary), syllables have STØD-BASIS in Basbøll's account of stød (2005:277–278). Note that vowel length is distinctive in Danish, as in [ˈviːlə] *hvile* 'rest' versus [ˈvilə] *ville* 'would', and vowels with stød are long vowels phonologically. There is no corresponding length distinction in Danish consonants.

The fundamental PHONOLOGICAL GOVERNING PRINCIPLE is that a syllable with stød-basis has stød if it is not penultimate in the lexeme, hence for instance

[t͡sanˀ] *tand* 'tooth' with stød but [ˈt͡sanə̥] *tante* 'aunt' without stød. There are lexically marked exceptions to this principle in certain monosyllables with a short vowel ending in one sonorant consonant, such as [føl] *føl* 'foal' and [vɛn] *ven* 'friend' without stød, and in many disyllables ending in /əl ən ər/, such as [ˈaŋˀg̊əl] *ankel* 'ankle', [ˈlæːˀjən] *lagen* 'sheet', and [ˈfeːˀb̥ɐ] *feber* 'fever' with stød. In inflection and derivation complex morphological principles apply, giving rise to stød contrasts such as [ˈhuːˀsð̩] *huset* 'the house' with stød (from INDEF.SG. [ˈhuːˀs] *hus* 'house') versus [ˈhuːsð̩] *huset* 'housed' without stød (from INF [ˈhuːsə] *huse* 'house'). See further Basbøll (2005, 2008, 2014, 2021, 2022) and Grønnum, Pharao & Basbøll (2020).

Phonetic investigations of stød may be undertaken to illuminate questions about its physiology, acoustics, or perception. But the point of departure may also be a quest for a phonetic underpinning of a particular phonological analysis. Thus, for instance, building on prior investigations of vowel and consonant duration and the timing of the onset of creakiness (Petersen 1973; Fischer-Jørgensen 1987), Basbøll (1988) proposed that Danish syllables have moraic structure, and that stød is a property of the second mora in bimoraic syllables. He specified that non-sonorant consonants cannot be moraic in Danish, and that trimoraic syllables are excluded (Basbøll 1989, 2005:292, 2021). In other words, [seːˀ] *se* 'see' and [seːˀn] *sen* 'late' are equally bimoraic, as are [halˀ] *hal* 'hall' and [halˀm] *halm* 'straw'. Basbøll also suggested that consonants with stød are long phonologically. His analyses occasioned additional investigations of duration (Grønnum & Basbøll 2001) as well as two perceptual experiments (Grønnum & Basbøll 2002, 2003).

3. Older phonetic descriptions and investigations, 1743–1943

Jens Pedersen Høysgaard (1743, 1747, 1769) was the first to identify stød phonetically and to recognize its distinctive function. The term STØD is also due to him. It means 'push' or 'thrust', and he described it impressionistically as resembling a very little hiccup (*et meget lidet hik*) that arises when the pharynx closes and cuts off the breath (*svælget lukker sig for ånden* – Fischer-Jørgensen surmises that 'Høysgaard . . . hardly made any distinction between pharynx and larynx' 1987:72).

For a century and a half after Høysgaard, stød was commonly assumed to be a glottal stop. That was, for instance, the opinion of highly influential Otto Jespersen (1897–1899:297), one that he maintained in the German edition of his book (Jespersen 1913:78–79), although Rousselot's (1897–1901) exploratory kymograph experiments had shown that the vibrations in the larynx hardly ever died away completely in words with stød (Rousselot 1901:873–879). That same observation was subsequently made repeatedly – albeit in experiments of somewhat limited scope in terms of number of speakers and recorded words – thus also in the last investigations to rely exclusively on the kymograph in stød investigations, that is, Ekblom (1933) and Abrahams (1949). See Fischer-Jørgensen (1987:72–74) for a fuller account of older descriptions and instrumental investigations of stød.

The kymograph was not a very sensitive or accurate instrument, though it functioned reliably enough to show that stød is not normally a full glottal stop. Smith (1944) is the first phonetic investigation that also employed more advanced

methods and equipment, but it took more than a decade before the kymograph was irrevocably retired and became a museum piece. Thus Ringgaard (1960) was the last to employ it in some of his initial experiments around 1958.

4. Modern instrumental investigations

Between 1944 and 1978, three Danish doctoral dissertations and eight papers saw the light of day. Then there was a lull until Eli Fischer-Jørgensen's monumental report about experiments – spanning more than a decade since 1974 – appeared in 1987. Another lull ended in 2001 when Hans Basbøll and I published the first of six papers. The last period contains the four most recent contributions to empirical phonetic stød research. If I am not guilty of an oversight, there are no more modern empirical phonetic (acoustic, physiological, or perceptual) analyses of common Danish stød than the twenty-two books and papers reviewed below.

4.1 1944–1978

4.1.1 Smith 1944
Svend Smith's (1944) dissertation – in the first major, systematic, and detailed experimental phonetic investigation of Danish stød – challenged the generally accepted view of the larynx as the primary source of stød. In what for his time was a groundbreaking and difficult interdisciplinary undertaking, he recorded electromyographic (EMG) activity from surface electrodes placed on the lower abdomen of five speakers. Readers appreciate the precariousness of the enterprise when they learn that in normal upright bodily position, the abdominal activity involved in maintaining this posture drowns out whatever specific activity speech production might entail. Only lying down did Smith obtain valid recordings from the electrodes, registered by an oscillograph, and captured on film. The recorded material consisted of eleven pairs of stød/non-stød words (five monosyllabic, five disyllabic, and one trisyllabic pair), such as [vɛn] *ven* 'friend', [vɛnˀ] *Vend!* 'Turn!' and [ˈʁoːsən] *rosen* 'the rose', [ˈʁoːˀsən] *rosen* 'the praise'. Each pair was pronounced several times, in isolation, by five speakers. Numerous recordings were discarded due to speaker errors and technical glitches, yielding merely eighty-three pairs of words suitable for analysis. Six were whispered. In fifty-six instances, three of them whispered, the word with stød exhibited an abrupt increase in abdominal EMG activity, succeeded by an equally abrupt decrease. However, Smith was unable to synchronize EMG and sound recordings and therefore could not be sure at what point in time the higher EMG activity began in words with stød. Nevertheless, he proposed that the causative factor in stød production is a ballistic contraction of the expiratory muscles. This sudden contraction creates a steep increase in subglottal pressure that in its turn (though not infallibly) induces a secondary, proprioceptive reflexive innervation of the vocal folds, giving rise to creakiness. The subsequent equally abrupt reduction in subglottal pressure results in a reduction of intensity and F_0.

In a subsequent experiment, Smith investigated F_0 in stød versus non-stød. Six speakers produced some forty pairs of words with and without stød, as well as a few short sentences. They were recorded with a microphone whose vibrations were

transferred to an oscillograph and captured on film. The vibrations on the film strip were analysed with a Meyer-Schneider pitch meter, a time-consuming manual procedure. There was no significant F_0 difference at the beginning of words with stød and non-stød. The F_0 fall often observed at the end of syllables with stød is not a tonal phenomenon *per se*, notes Smith, but is to be explained in conjunction with the reduction in amplitude, and both are a consequence of the reduction in subglottal pressure that results from the abrupt cessation of innervation of the expiratory musculature. Finally, Smith observed a tendency for greater vibratory amplitude at the beginning of syllables with stød, indicative of greater tension, a tendency that became the subject of a third experiment.

Smith recorded ORAL AIRFLOW VELOCITY in unvoiced initial consonants in three speakers on a kymograph. He found some evidence of greater flow velocity that he interpreted as greater articulatory energy at the beginning of words with stød. In the same vein, the only voiced initial consonant in the recorded material, [l], exhibited vocal fold oscillations of greater amplitude in words with stød, also indicative of more articulatory energy.

Smith introduced the concept of PHASES in stød production. The first phase is the increase of EMG activity in the abdominal musculature. The second phase is the succeeding decrease of EMG activity – often, but not invariably, accompanied by irregular vocal fold vibrations. This concept of a biphasic stød resonated with most researchers following Smith but with a slightly different, that is, acoustic connotation. The FIRST PHASE became the interval with modal voice, the SECOND PHASE became the interval with creakiness.

Concluding the results of all three experiments, Svend Smith's final words (1944:120) are that '... Stødet herefter kan defineres fysiologisk som en *Trykaccent, en særlig markerende Bevægelse foretaget ved en stødagtig Fremhæven af Lydmasse*' [... the stød henceforth can be defined physiologically as a *stress-accent, a particular marking motion brought about by a push-like highlighting of sound mass*].

In the years following Smith's experiments, the explicit assumption became that the vocal folds (the vocalis muscles) are active, not passive participants in stød production. Common to every phonetic investigation of stød after Smith's is the assertion that the vibratory pattern of the vocal folds in syllables with stød deviates in SOME MANNER from normal, modal vocal fold vibrations, and subsequent researchers set out to characterize in more detail the nature of this deviation. Some have also tried to establish the physiological mechanism underlying it.

4.1.2 Fischer-Jørgensen 1955

Eli Fischer-Jørgensen (1955) measured the duration of long vowels without stød, long vowels with stød, and short vowels in two triads – [ˈpʰiːb̥ɐ] *piber* 'pipes', [ˈpʰiːˀb̥ɐ] *piber* 'whines', [ˈpʰib̥ɐ] *pipper* 'chirps', and [ˈlɛːsɐ] *læser* 'reader', [ˈlɛːˀsɐ] *læser* 'reads', [ˈlɛsɐ] *læsser* 'loads (PRS)' – recorded once in isolation by ten speakers. The same triads plus [ˈhyːlə] *hyle* 'howl', [ˈhyːˀlɐ] *hyler* 'howls', [ˈhylə] *hylde* 'praise (INF)' were embedded in short sentences and recorded three times by three speakers each. In isolation, long vowels with stød averaged 82% of the duration of long vowels without stød. In sentence context they averaged

89% of long stødless vowel duration. The short vowels averaged 50% in isolation and 60% in sentences, respectively, of long stødless vowel duration. Fischer-Jørgensen (1987) contains a more comprehensive investigation of vowel duration; see Section 4.2.1.

4.1.3 Faaborg-Andersen 1957

During a pioneer investigation of EMG activity in intrinsic laryngeal muscles in thirty-two healthy subjects and twenty-three patients with unilateral vocal fold paresis, Knud Faaborg-Andersen (1957) also recorded a pair of words with and without stød, [moːˀɐ̯] *mord* 'murder' and [moːɐ̯] *mor* 'mother', from six healthy speakers. (These words would be pronounced [moɐ̯ˀ] and [moɐ̯] today.) He obtained action potentials from two needle electrodes, one inserted via the mouth into the vocalis muscle and one inserted from the outside, through the skin and the cricothyroid cartilage, into the cricothyroid muscle. He obtained a total of nine and ten recordings from the vocalis and cricothyroid muscles, respectively. Table 11 (Faaborg-Andersen 1957:78) lists the maximum amplitude of the electric action potential in µV in the two muscles in the production of [moːˀɐ̯] and [moːɐ̯], respectively, in these nineteen instances. I have calculated the difference in maximum action potential between the two words. In the vocalis muscle, the average maximum action potential is 56.3 µV higher in the word with stød, whereas in the cricothyroid muscle, it is only 11.3 µV higher in the word with stød. However, the dispersion in the data is considerable and the difference is not statistically significant. I think the conclusion should be that Faaborg-Andersen demonstrated a TENDENCY for higher activity in the vocalis muscle in stød than in non-stød, but no corresponding increase in the cricothyroid muscle. Faaborg-Andersen did not measure F_0, therefore we cannot know whether the (in retrospect) surprising absence of substantial increase in cricothyroid activity indicates that there was no F_0 difference between the two words.

4.1.4 Ringgaard 1960

In another pioneer research effort, Kristian Ringgaard (1960) investigated West Jutlandic stød. It occurs in the dialects of West Jutland with an extension towards the east at Vejle and Horsens and reaches across the Little Belt to the northernmost part of Funen (Ringgaard 1973:24). Lying north of the stød border, these regions also have the common stød, though with a different distribution than in the Copenhagen standard language. Monosyllables with a short vowel succeeded by a sonorant consonant and an obstruent, such as *pels* 'fur' and *flink* 'nice', invariably have stød in Copenhagen, [pʰɛlˀs] and [fleŋˀɡ̊]. In Jutland and on Funen they are without stød, [pʰɛls] and [fleŋɡ̊].

West Jutlandic stød occurs in stressed syllables before /p t k/ in intervocalic, or originally intervocalic, position (Ringgaard 1960:10). That is, it occurs in disyllables and in monosyllables that derive from apocopated disyllables, but never in original monosyllables, as in [ˈsɔʔɡ̊ə] *sukker* 'sighs (PRS)' and apocopated [sɔʔɡ̊] *sukke* 'sigh (INF)', the latter contrasting with a stødless original monosyllable [ˈsɔɡ̊] *suk* 'sigh (N)'.

On a kymograph, seven speakers recorded isolated words with West Jutlandic stød, with common stød, and without stød. Ringgaard found the kymograms well suited to establish durational relations, but they could not provide unambiguous information about the ARTICULATION of West Jutlandic stød, except that it is different from the common Danish stød. He then proceeded with spectrography of his own speech at Svend Smith's Institut for Talelidende [Institute for the Speech Impaired] in Hellerup, the only phonetics laboratory in Denmark with a spectrograph in 1958. Ringgaard discovered that in the common stød of [d̪uˈkʰæːˀd̪] *dukat* (a gold coin), voicing continued until the onset of the final stop consonant, and the release of the stop consonant was affricated. That is, glottis must have been open at the time of release. West Jutlandic stød, on the other hand, exhibited a *'lydløst' tomrum* ['soundless' void] at the end of the vowel, obliterating the vowel formant transitions to the consonant, and no noise whatever after the release of the stop consonant (Ringgaard 1960:76).

Ringgaard was unable to determine what caused this soundless void, a vocal fold occlusion or EKSPIRATIONSSTANDSNING [cessation of expiration]. To settle the issue, he had an X-ray movie made at Århus Municipal Hospital of his vocal folds, illuminated from the back of the neck, moving at 48 frames per second. He pronounced [kʰad̪] *kat* 'cat' three times and [kʰaʔd̪] *katte* 'cats' six times. I regret that with my untrained eye I am unable to interpret the X-ray still pictures reproduced in Figures 27 and 28 (1960:78–81). But in a footnote (1960:77) Ringgaard quotes Chief Physician E. Mosekilde, who states that

> in the word [kʰad̪] *kat* 'cat' the false and true vocal cords only just come together while preserving their natural relief. In [kʰaʔd̪] *katte* 'cats' the occlusion is far more massive … The true and false vocal cords and the sinus of Morgagni are completely obliterated [my paraphrased translation].

Ringgaard concludes that West Jutlandic stød is a GLOTTAL STOP. He reasons that such a firm glottal occlusion must involve increased subglottal pressure.

Subglottal pressure was recorded via a 2 mm thick syringe, inserted into Ringgaard's larynx just below the cricoid cartilage, and printed on an electrocardiograph, while he pronounced isolated words without stød, with common stød, and with West Jutlandic stød. He found that in stødless words as well as words with common stød, subglottal pressure rose and fell in a smooth curve with peak pressure around the midpoint in the vowel. But in West Jutlandic stød, subglottal pressure was higher and the peak later in the vowel. He had expected subglottal pressure to increase even further, after the occlusion in the glottis. But it did not. During the glottal and succeeding oral occlusions, subglottal pressure fell back to its level of equilibrium. That could only happen if the respiratory organs performed a sudden INSPIRATORY movement.

Accordingly, in the last experiment in this impressive series, X-ray diaphragm kymography revealed that in four productions of [kʰaʔd̪] *katte* 'cats' Ringgaard's diaphragm performed a downwards, that is, an inspiratory movement during the glottal and oral occlusions, only reverting to expiration at the end of the word.

Ringgaard concluded the results of his challenging experiments, noting that West Jutlandic stød, that is, the glottal stop, appears to be both a very energetic and a very complex articulation.

I have no reason to doubt Ringgaard's methods and results. But I believe a cautious caveat is in order here, one that I could voice about many physiological investigations of speech. Because the invasive procedures are intensely uncomfortable, it is difficult to recruit enough subjects to confidently allow generalization of the results to all speakers of the language.

Finally, I want to point out that even more remarkable than the phonetic nature of West Jutlandic stød is the fact that West Jutland has two phonetically different – and phonologically distinctive – laryngeal phenomena, a glottal stop and a common Danish stød, and in almost perfectly mutually exclusive contexts. This is a rather unique situation. Nevertheless, West Jutlandic stød is alive in the dialects to this day (Torben Arboe, Yonatan Ungermann Goldshtein, and Inger Schoonderbeek Hansen, personal communication, March 2022).

4.1.5 Lauritsen 1968

Margaret Lauritsen's (1968) short paper is from a Berkeley Phonology Laboratory Report that I have not been able to access. I quote her results from Fischer-Jørgensen (1987:77–78). Lauritsen published 33 spectrograms of words with and without stød, produced by one speaker. At the end of long vowels with stød, the spectrograms revealed a clear reduction in intensity, sometimes creakiness, and falling F_0. Importantly, Lauritsen noted that the timing of creakiness is variable. In a long vowel it may spill over into a succeeding sonorant consonant, and in a sonorant consonant it may show up already in the preceding short vowel. She further noted that stød is audible in whispered speech through a reduction of the intensity of the whispery noise. I may add that the audibility of stød in whisper is something Danish speakers can easily verify for themselves. Contrasts like [vɛl] *vel* 'well' versus [vɛlˀ] *væld* 'abundance', [hans] *hans* 'his' versus [hanˀs] *Hans* (a boy's name), and [ˈmæːlɐ] *maler* 'painter' versus [ˈmæːˀlɐ] *maler* 'paints' are perfectly perceptible in whispered speech, in isolation as well as in running speech.

4.1.6 Vihman 1971

In an appendix to her dissertation on Livonian phonology, Marilyn Vihman (1971) compared Danish and Livonian stød. Livonian stød is always associated with a tonal difference, one that may even substitute for creakiness. That is not the case in Danish. One speaker recorded a series of minimal pairs, words with and without stød, in isolation and in a small frame sentence. Vihman found creakiness as well as decreasing F_0 at the end of most words with stød. But there were also instances of stød where no creakiness and no F_0 decrease were detectable in the spectrograms. Those words, however, exhibited higher F_0 at vowel onset than did non-stød words. Later in the vowel, F_0 converged with the non-stød pattern.

Vihman's observations were confirmed by nearly every acoustic analysis that followed.

4.1.7 Petersen 1973

Pia Riber Petersen (1973) analysed F_0, intensity, and a few spectra in disyllables from six speakers. As did Lauritsen (1968), she found considerable variability in the acoustic manifestation of stød in long vowels: explicit or less explicit creakiness, that started earlier or later in the vowel, and was of shorter or longer duration. One subject (labelled '22') stood out from the others by the complete absence of any visible signs of creakiness. However, his stød syllables exhibited higher F_0 than non-stød syllables.

Petersen subsequently ran a listening test, made up of recordings of a pair of words, [ˈlɛːˀsɐ] *læser* 'reads' and [ˈlɛːsɐ] *læser* 'reader', spliced out from their sentence context, by speaker 22 and by another male speaker with acoustically explicit stød. [ˈlɛːˀsɐ] was very often identified as stødless [ˈlɛːsɐ] *læser* 'reader', whereas only rarely was stødless [ˈlɛːsɐ] identified as [ˈlɛːˀsɐ] *læser* 'reads'. With that caveat, however, speaker 22's renditions of [ˈlɛːˀsɐ] *læser* 'reads' were identified correctly more often than those of the acoustically more explicit speaker.

Petersen furthermore found long vowels with stød in her material – where test items were embedded in semantically natural sentences – to have the same duration as long vowels without stød. She found the creakiness to take up about one third of the total long vowel duration and to begin at about 110 ms from vowel onset. Consonants with stød after short vowels were longer by about 35% than consonants without stød. Short vowels before consonants with stød were slightly shorter than before consonants without stød. F_0 maxima were higher in words with stød, and minima were lower. That is, the frequency range spanned was larger in stød than in non-stød words. Intensity displayed more of a fall in stød words, whether F_0 was also falling (as it was not with speaker 22).

4.1.8 Thorsen [Grønnum] 1974

I was intrigued by Petersen's results and ran a perceptual experiment. I found that when the end of a word with stød was cut off at the point in time where the vocal fold vibratory pattern changed visibly from modal voice to creakiness – or at any point earlier in time – stød was no longer perceptible. A similar test with stimuli from Petersen's speaker 22 revealed that even his shortest stimuli were infallibly perceived to have stød. I suggested that this was due to two features: his voice sounded somewhat strained or compressed in words with stød and they also began on a significantly higher F_0. In retrospect, I propose that these two features may be correlated; see the discussion in Section 6.1.

4.1.9 Fischer-Jørgensen & Hirose 1974a, 1974b

Eli Fischer-Jørgensen & Hajime Hirose (1974a) recorded EMG activity in labial and laryngeal muscles in stop consonants at the Haskins Laboratories in 1972. Among the words in the recorded material were also two pairs of stød/non-stød words, [man] *man* 'one/you' versus [manˀ] *mand* 'man' and [ˈpʰiːbɐ] *piber* 'pipes' versus [ˈpʰiːˀbɐ] *piber* 'whines'. In a brief follow-up paper (Fischer-Jørgensen & Hirose 1974b) they noted that the vocalis muscle exhibits a clear peak of activity in the stød words, absent in non-stød, with the one speaker whose recordings could

be reliably interpreted. The same speaker exhibited no difference in labial muscle activity, or in other words no evidence of a more forceful labial gesture in the onset of syllables with stød versus no stød.

4.1.10 Fischer-Jørgensen 1974

This is short note to say that the work that commenced with Hajime Hirose at the Haskins Laboratories in 1972 continued in May and June of 1974 in Copenhagen, where EMG recordings of the vocalis muscle were obtained from seven speakers in words with and without stød. Four speakers showed increased vocalis activity in vowels with stød, three speakers did not. The final words in this advance notice are: 'More details will be given in later reports' (Fischer-Jørgensen 1974:206), as indeed they were – in 1987.

4.1.11 Jørgensen 1978

John Jørgensen (1978) quotes Roman Jakobson's (1941) dictum that phonological phenomena that are rare in the languages of the world are acquired late by children. He recorded and transcribed the speech of four two-year-old toddlers. Each child produced around 125 short, non-imitated phrases – with two or three syllables per phrase. In a total of about one thousand syllables, only forty-one, equally distributed among the four children, contained a mistake in stød. There were twenty-five instances of no stød where one should have been, ten instances of stød where none should have been, and six instances of stød on the wrong segment in the syllable. Jørgensen suggests that this early mastery is due to stød being a prosodic phenomenon, and prosodic phenomena (for instance, intonation) are known to be acquired very early in children's linguistic development.

4.2 Fischer-Jørgensen 1987 (1989a, 1989b)

Fischer-Jørgensen (1989a) is a photocopy of a slightly revised version of the manuscript for 1987, and Fischer-Jørgensen (1989b) is an abbreviated version. I refer to the comprehensive 1987 annual report edition here, that has recently become available online. It reports the results of experiments conducted over more than a decade after 1974.

Various subgroups of fifteen speakers, some of whom spoke Standard Danish from a regional background, were analysed in the many different experiments reported in this mammoth paper. Eli Fischer-Jørgensen performed extensive acoustic analyses: duration, creakiness, F_0, intensity, and spectral composition. She filmed the vertical movement of the larynx, made fibre optic (laryngoscopic) videos of the larynx, recorded oral airflow, subglottal pressure, pharyngeal pressure, and she made palatograms. She obtained EMG signals from the vocalis, the cricothyroid, the interarytenoid, the posterior cricoarytenoid, and the lateral cricoarytenoid muscles. A microphone signal was also always recorded and synchronized with the physiological recordings.

4.2.1 Acoustic analyses

DURATION of vowels with stød in words read in isolation was less (by 15% on average) than of long vowels without stød. But in words recorded in sentences, vowels with and without stød were of equal duration. Consonants with stød were longer (by 33% on average) than consonants without stød, with a tendency for the preceding short vowel to be shorter before a consonant with stød. This is in line with Petersen's (1973) results and with Fischer-Jørgensen's own (1955) findings.

CREAKINESS was present in (only) 71% of 1700 words with stød, but there were considerable differences between speakers and even more considerable intraper-sonal differences, associated with differences in clarity of pronunciation across different recording sessions. Fischer-Jørgensen noted explicitly that there was not a single instance among the 1700 words where stød appeared as a glottal stop.

FUNDAMENTAL FREQUENCY was invariably, with every speaker, higher at the beginning of syllables with stød. At the end F_0 was lower in only about half of 1700 words analysed. The timing of F_0 lowering, when present, was variable.

INTENSITY was often higher at the beginning and nearly always, though not without exception, lower toward the end of a syllable with stød.

SPECTRAL ANALYSIS showed more energy in the upper part of the spectrum (that is, acoustic compression) in vowels with stød, mostly only toward the end of the syllable, because – as demonstrated by inverse filtering – the sound produced by the vocal folds has relatively more energy in the higher harmonics in creakiness than in modal voice.

In summary: Fischer-Jørgensen found the manifestation of stød to be extremely variable acoustically. However, whether creakiness was present or not, F_0 was invariably higher in the onset of syllables with stød, and intensity was nearly always lower at the end.

4.2.2 Physiological analyses

LARYNX VERTICAL MOVEMENT is reflected in the movement of the thyroid prominence (the Adam's apple). It was filmed in five male speakers – in profile – with a television camera and a video-recorder. A general tendency for the larynx to be lower in [u] than in [i] and [a] was confirmed, but no consistent difference was found between vowels with stød and vowels without stød.

FIBRE OPTIC VIDEO recordings of the vocal folds in seven speakers showed compression of the anterior part of the vocal folds toward the end of syllables with stød. Likewise, the ventricular folds were constricted, but with very considerable differences between speakers.

ORAL AIRFLOW was recorded in five speakers and showed a tendency for stronger flow in prevocalic unvoiced consonants in syllables with stød – reminiscent of Smith's (1944) observation – and a significantly diminished flow toward the end of the syllable.

SUBGLOTTAL PRESSURE was obtained in eight pairs of words from one speaker. It exhibited small but significant differences between stød and non-stød words: a steeper rise and a slightly, but consistently, higher peak pressure about 50 ms from vowel onset in syllables with stød. The reduction in airflow and the higher subglottal pressure together are indicative of increased resistance in the glottis – that is, the

opening between the vocal folds – during creakiness. The effect should be a lower PHARYNGEAL PRESSURE above the glottis, and that was what appeared in three speakers. But the difference in pharyngeal pressure was small between stød and non-stød and not always statistically significant.

PALATOGRAMS of five speakers revealed a larger area of contact between the tongue and the hard palate in high vowels with stød versus non-stød.

Together, the results of these experiments are indicative of the expenditure of more articulatory energy at the onset of syllables with stød – with a cautious caveat about the lips; see the discussion of Fischer-Jørgensen & Hirose (1974b) in Section 4.1.9.

EMG RECORDINGS were obtained by hooked wire electrodes inserted into the muscles, five in all. The INTERARYTENOID and POSTERIOR CRICOARYTENOID muscles are responsible for closing and opening the glottis, respectively. Recordings were obtained from three speakers and none of them revealed any activity specifically associated with stød. The LATERAL CRICOARYTENOID muscles bring the vocal processes together, increasing the medial compression of the vocal folds, but otherwise seem to have much the same function as the vocalis muscles. The VOCALIS muscles are active in all voiced sounds, relaxed in unvoiced sounds, and involved in the manifestation of glottal stops. The CRICOTHYROID muscles lengthen the vocal folds, making them thinner and tenser, thus increasing F_0. The vocalis muscles may also be involved when – within the chest register – F_0 rises, possibly to counteract the slight opening of the glottis that would otherwise result from tension of the cricothyroid muscles. Note that except for the interarytenoid, these muscles are all paired. That is, there are, for instance, two vocalis muscles, but it is customary to record from one side only in a pair and to refer to it in the singular.

VOCALIS AND CRICOTHYROID MUSCLES were recorded in seven speakers. Five speakers exhibited a markedly stronger vocalis activity in syllables with stød: a steep rise beginning on average 20–40 ms after vowel onset, a well-defined peak at about 100–130 ms, and a subsequent steep fall, yielding a total duration of 100–200 ms. The cricothyroid activity increased and decreased in harmony with F_0 movements. In words with stød, where F_0 was high at vowel onset, the cricothyroid peak was correspondingly early, slightly preceding the F_0 maximum. In words without stød, where F_0 performed a rising movement to the post-tonic syllable, the cricothyroid peak was correspondingly later. And, notably, the level of activity was NO HIGHER in the early (stød) cricothyroid peaks than in later (non-stød) peaks. In other words, the cricothyroid is associated with F_0 but not with creakiness *per se*, and vocalis was unambiguously a factor in the production of stød in these five speakers.

The situation was different for the remaining two speakers, who exhibited no apparent difference in vocalis activity between stød and non-stød words. In one speaker the electrode appears to have been dislocated, but the other speaker (labelled 'JR') likely did not employ his vocalis muscle in stød at all. Furthermore, what little creakiness he exhibited occurred on a rising F_0 and there was no F_0 drop at the end of the syllable. On the other hand, JR's ventricular folds were the most constricted among the seven speakers, almost completely blocking the view of his vocal folds. Fischer-Jørgensen does not say so, but it is implicit that whatever creakiness is present in JR's stød is a result of the shortening and thickening of the vocal folds that occur when the ventricular folds are constricted due to contraction of the ARYEPIGLOTTIC

SPHINCTER musculature. One might further speculate whether this ventricular constriction, present also to varying degrees in the other six speakers, may not also be a concomitant factor in THEIR stød production, in addition to the vocalis activity.

4.2.3 Stød versus other phonation types

Fischer-Jørgensen considered how stød is distinguished from similar non-modal voice qualities and phonation types. HARSH VOICE exhibits aperiodicity but not irregular variation in intensity. The first phase of a syllable with stød – that is, the beginning without creakiness – shares features with TENSE VOICE, for instance higher subglottal pressure and higher intensity. But it does not exhibit the constricted glottis that characterizes tense voice. Constricted glottis only appears during the second phase, with the creakiness. Stød shares many features with CREAK and CREAKY VOICE, but they never exhibit the reduction in overall intensity level very often found toward the end of stød syllables. Therefore, Fischer-Jørgensen (1987:183) concluded, 'Stød is not just creaky voice.'

4.2.4 Causal relations

The increase in FUNDAMENTAL FREQUENCY at the beginning of syllables with stød that Vihman (1971) and Petersen (1973) also found, cannot be due to an increase in SUBGLOTTAL PRESSURE, because the magnitude of that increase is insufficient to drive the magnitude of F_0 increase observed. Nor can a subsequent rapid decrease in subglottal pressure account for the frequently observed reduction of INTENSITY, because intensity often begins to lower before the peak in subglottal pressure is reached. But might a subglottal pressure increase induce at least a startle reflex, in terms of a vocalis contraction? Not as a general and consistent event, because in several individual instances, the vocalis activity peak preceded the subglottal pressure peak. Furthermore, contraction of the aryepiglottic sphincter musculature above the glottis, constricting the ventricular folds, can hardly be the result of changes in subglottal pressure. The conclusion is inescapable that vocalis muscle activity in stød is an independent activity and not a secondary effect of changes in subglottal pressure.

There are further potentially related parameters. They are partly acoustic (high F_0 and high intensity at syllable onset; creakiness, lowering of intensity, lowering of F_0, and increased energy in the upper part of the voice source spectrum toward the end of the syllable), partly physiological (articulatory energy; increased vocalis and crico-thyroid muscle activity; constriction of the ventricular folds). Fischer-Jørgensen ascertained that high F_0 and cricothyroid activity are obviously associated. (I would add that higher intensity may be merely a function of higher F_0, not an independent factor, because when frequency increases, so does intensity, exponentially to the power of two, *ceteris paribus* (Divell 2010).) But the F_0 lowering and the creakiness, often observed in the second phase of stød syllables, cannot be due to a cricothyroid relaxation that already begins early in the first phase. Falling F_0 and creakiness are more likely the result of a strong contraction of the vocalis muscle (and the lateral cricoarytenoid), often accompanied by constriction of the ventricular folds. In other words: cricothyroid activity at syllable onset increases F_0, and vocalis activity causes creakiness and lowers F_0 toward the end of syllables with stød.

4.3 Grønnum & Basbøll 2001–2012, Grønnum 2015

In a series of investigations from 2001 to 2012, Hans Basbøll and I addressed various acoustic and perceptual issues, before my provisionally last phonetic investigation of stød in 2015.

4.3.1 Grønnum & Basbøll 2001

In Grønnum & Basbøll (2001) we tested the phonetic reality of Basbøll's (1988) proposal that consonants with stød are phonologically long, and we also recorded a small corpus to measure vowel duration. Consonants with and without stød were measured in five different conditions: (i) word finally in utterance final position, (ii) word finally in utterance medial position, (iii) word medial position between a stressed and a post-tonic vowel, (iv) word medial position between a vowel and a consonant, and (v) syllabic when schwa is assimilated. Three vowel qualities ([i y ɛ]) were tested short, long with stød, and long without stød. Seventy-one different words in all, each embedded in six different semantically normal sentences yielded 426 different utterances, recorded by five speakers.

VOWEL DURATION measurements in five speakers across three vowel qualities averaged 128 ms for long vowels without stød, 121 ms for vowels with stød, and 90 ms for short vowels. The 7 ms difference between long vowels with and without stød is not statistically significant and they are of essentially equal duration. Short vowel duration is 70% of long vowel duration, in agreement with both Petersen (1973) and Fischer-Jørgensen (1987).

SONORANT CONSONANT DURATION measurements did not confirm Petersen's (1973) and Fischer-Jørgensen's (1987) findings that consonants with stød are longer than consonants without stød, not across positions. And particularly, word final position turned out to be intriguing. In utterance medial position, word final [nˀ] and [lˀ] were between 4 and 27 ms longer than [n] and [l], but utterance finally, [nˀ] and [lˀ] were between 19 and 48 ms SHORTER, not just relative to [n] and [l] in the same context but compared with [n] and [l] in all other positions, as shown in Table 1.

We wondered why utterance final consonants with stød should be so short and suggested that opposing forces in stød production and pause anticipation might be at work. We recalled Fischer-Jørgensen's (1987) laryngoscopic investigation that showed a transverse narrowing of the glottis (that is, compression of the anterior part of the vocal folds) during the stød phase, and we suggested two different scenarios (Grønnum & Basbøll 2001:242).

(i) If a narrower glottis is a prerequisite for stød, then abduction of the vocal cords, i.e., opening of the glottis, preparatory to the pause, will inhibit creaky voice more than modal voice, and the vibrations will accordingly cease sooner.

(ii) Subglottal pressure decreases before a pause. Therefore, volume velocity through the glottis diminishes, and the aerodynamic driving force in the vibrations – the so-called Bernoulli Effect (Bernoulli 1738; Grønnum 2005:69) – will not be sufficient to maintain vibrations against the adductive forces involved in stød. In other words, vibrations are interrupted by a glottal closure.

Table 1. Duration of [n] and [l], with stød and without stød, in five different contexts. The difference is statistically significant if $p \leq 0.01$. Data extracted from Tables 9 and 10 of Grønnum & Basbøll (2001:243)

Position	[nˀ] and [lˀ] duration in ms	[n] and [l] duration in ms	Difference in ms	*p*-value
Word final utterance final	36	63	−27	≤ 0.01
Word final utterance medial	68	55	13	≤ 0.01
Word medial before a vowel	58	49	9	
Word medial before [s]	65	60	5	≤ 0.01
Syllabic consonants	96	94	2	

We noted that certain items in our recordings appeared to be instances of the latter scenario, others fit the first option better. SYLLABIC CONSONANTS with and without stød (in pairs like [ˈsg̊yll̩] *skylde* 'owe' and [b̥eˈsg̊ylˀl̩] *beskylde* 'accuse') did not exhibit consistent differences.

Overall, our data on consonant duration did not provide a general phonetic underpinning of a phonological analysis of consonants with stød as being long. The only condition in Table 1 that exhibited noticeably longer stød consonants was utterance medial, word final position (68–55 = 13 ms). In retrospect, it is perhaps relevant that in the sentences read by the speakers, these words were almost exclusively succeeded by a word beginning with an unvoiced obstruent. Therefore, spill-over of creakiness to a succeeding syllable could not occur. Perhaps for the stød to run its full course, the consonant carrying it may be slightly lengthened. There is some support for such an assumption in the values in Table 1 from word medial position before [s], where [nˀ lˀ] were very slightly but nevertheless significantly longer than [n l] (65–60 = 5 ms).

STØD ONSET TIMING AND DURATION measurements confirmed previous authors' observations of considerable variability in onset and duration of creakiness. It may start as early as 10 ms after vowel onset and as late as 130 ms, averaging at 67 ms across short and long vowels alike. Quite apart from the fact that creakiness is not universally present, we noted that this variability challenges the traditional view, developed over the years since Smith (1944), of a biphasic stød (in long vowels) that begins with modal voice and approximately halfway through the vowel turns into creakiness.

4.3.2 Grønnum & Basbøll 2002

In Grønnum & Basbøll (2002) we conducted a perceptual experiment to find out whether listeners perceive vowels with stød to be long, in accordance with the physical facts and their phonological status. We created forty disyllabic nonce words with four different stressed syllable types, Vːˀ, Vː, VCˀ, and VC, such as [ˈtsiːˀlən], [ˈtsiːlən], [ˈtsilˀən], and [ˈtsilən]. Each syllable type occurred in ten different instantiations, exemplified with VC, thus: [ˈtsilən ˈb̥eməð ˈsd̥ɛnəð ˈmyləð ˈg̊øləð ˈsølən ˈsd̥ulən ˈtsɔməð ˈkʰʁalən ˈtsʁaməð]. We recorded these 40 words embedded in semantically reasonably natural utterances. The words were then excised from their

Table 2. Judgements of similarity between different syllable types, by two groups of listeners, eight in each group, as a percentage of the potential total of 640 in each cell. Based on Table 3 of Grønnum & Basbøll (2001:357)

similarities:	V:ˀ = V:	VCˀ = VC	V:ˀ = VCˀ	V: = VC	V:ˀ = VC	V: = VCˀ
Group 1	**69%**	**82%**	46%	48%	**30%**	**26%**
Group 2	48%	54%	**72%**	**71%**	**29%**	**25%**

context and arranged in an ABX-test where X was to be judged as most resembling either A or B. A, B, and X were always different. For instance, the subjects would hear [ˈmylˀəð ˈmyːləð ˈmyːˀləð] and had to decide whether [ˈmyːˀləð] sounded more like [ˈmylˀəð] or more like [ˈmyːləð]. Twenty-four different ABX combinations, times ten instantiations, times two voices yielded 480 different ABX stimuli, arranged in a random sequence, that were presented to twenty-two subjects. Six listeners responded in a completely random fashion, but sixteen were consistent in their responses, albeit in two different ways, as evidenced in Table 2.

We interpreted values above 67% as similarity, below 33% as dissimilarity (bold in the table) and percentages between the two as indeterminacy. There were, as accident would have it, eight listeners in each group. Listeners in group 1 appeared to base their similarity judgements on vowel length: (words with) phonologically long vowels resemble each other, irrespective of the presence of stød (69%), and (words with) phonologically short vowels resemble each other, irrespective of the presence or absence of stød in the succeeding consonant (82%). Group 2 founded their judgements on the presence or absence of stød. That is, (words with) stød syllables resemble each other, irrespective of phonological vowel length (72%), and (words with) stødless syllables likewise resemble each other (71%). Both groups rejected any resemblance between (words with) syllables that are different with respect to both stød and phonological vowel length (30%/29%; 26%/25%). We concluded: 'In brief: to half of our subjects stød vowels resemble phonologically long stødless vowels, to the other half syllables with stød resemble each other, irrespective of the segmental composition.'

4.3.3 Grønnum & Basbøll 2003

New in this paper is a perception experiment to test whether listeners could discriminate different timings of onset of creakiness. From the recorded material in our 2001 experiment, we extracted twenty disyllabic words, four words from each of five speakers, instances of [ˈd̢iːˀsən] *disen* 'the fog', [ˈsb̥iːˀsɐ] *spiser* 'eats', [ˈviːˀsɐ] *viser* 'shows (PRS)', [ˈpʰiːˀb̥ɐ] *piber* 'whines', [ˈg̊yːˀsɐ] *gyser* 'shivers', [ˈlɛːˀsɐ] *læser* 'reads', and [ˈlɛːˀnɐ] *læner* 'leans'. Two items by each speaker had visible creakiness in waveform and spectrum and two did not, even though stød was unambiguously audible in every word.

Listeners' task was a visual analogue scaling where they had to mark, on 10 cm long lines on sheets of paper, where in the vowel they thought the stød began (early to the left, late to the right). Eighty-one university students participated.

No evidence of bipartition of long stød vowels was found. That is, responses did not cluster around the midpoint of the line. There was a clear overall tendency for perceived stød onset at the very beginning of the vowel. Nevertheless, some stimuli had a slightly more pronounced tendency for early judgements than others. We extracted seven acoustic parameters from the comprehensive selection available in Praat (Boersma & Weenink 2006): HNR (harmonicity-to-noise ratio) average across the vowel; HNR maximum; distance of HNR maximum from vowel onset; average F_0 through the vowel; F_0 maximum; F_0 minimum; F_0 range. We also measured vowel duration. The only factor that showed a modest correlation ($r = 0.75$) with listeners' responses was, in fact, vowel duration. Stød was perceived to begin earlier in the high vowels [iːʔ] and [yːʔ] that are intrinsically shorter than lower [ɛːʔ]. The listeners, apparently at a loss to do what was asked of them, searched for something else in the stimuli to differentiate them, and found the acoustic property that distinguishes [iːʔ yːʔ] from [ɛːʔ], namely duration. That was our conclusion in 2003. I would like to offer another tentative explanation now. If stød perception is essentially cued by higher onset F_0, perhaps as a concomitant of greater acoustic compression, the intrinsically higher F_0 of [iːʔ] and [yːʔ] versus [ɛːʔ] may have contributed to a perception of earlier stød onset.

In retrospect, I want to stress what I consider to be the most interesting aspect of this experiment, namely that listeners apparently did not perceive long vowels with stød as having two parts, but – by and large – perceived stød throughout the vowel.

4.3.4 Grønnum & Basbøll 2007

In Grønnum & Basbøll (2007) we reviewed the phonetic results from our previous investigations and concluded with a proposal for a characterization of stød in articulatory terms, thus (Grønnum & Basbøll 2007:200):

> The laryngeal activity is a ballistic gesture that – minimally – makes for a slightly compressed voice quality, at one end of a continuum, and – maximally – creates a distinctly creaky voice at the other. Under emphasis it may become a complete glottal closure. It is a property of the sonorant syllable rhyme. It is aligned with the onset of the rhyme. It is variable with respect to strength and to temporal extension.

We interpreted the ballistic gesture as a lowpass filtered muscular response to a transient neural command, timed to coincide with the onset of the rhyme. The neural command may be stronger or weaker, making for more irregular or less irregular vocal fold vibration of shorter or longer duration, but when the command is executed, the speaker can no longer control the way the vocal folds respond to it, just as one can no longer control the trajectory of a tennis ball once it has bounced off the racket.

Figure 1 is an illustration of two very different stød manifestations and for comparison a word without stød, in words excised from larger utterance contexts. Note that stød is equally perceptible in the two leftmost words.
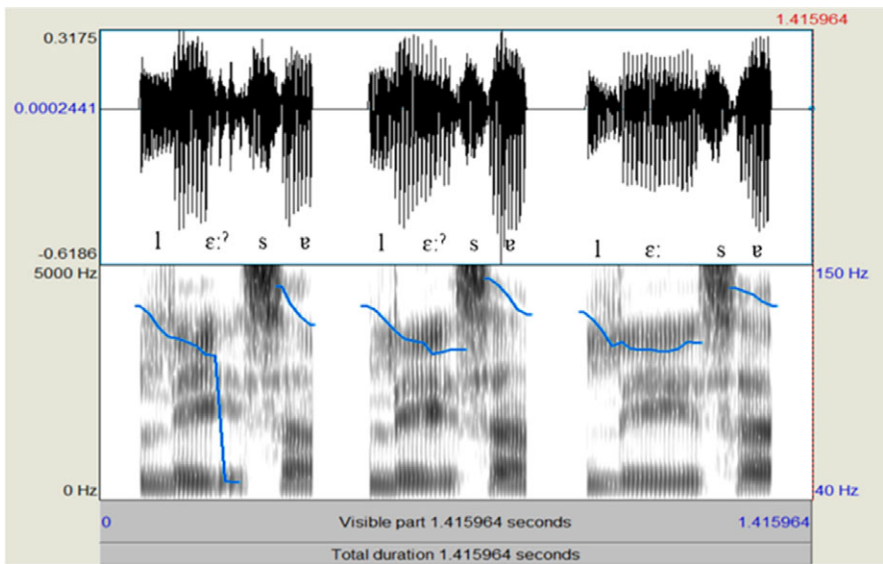
Figure 1. Praat screen dump (Boersma & Weenink 2006): waveform (above), spectrum (below), and F_0 tracing superposed on the spectrum of [ˈlɛːˀsɐ] *læser* 'reads' with explicit creakiness and F_0 perturbation (left), weak creakiness and no F_0 perturbation (middle), and [ˈlɛːsɐ] *læser* 'reader' without stød (right). Male speaker. Adapted from Grønnum, Vazquez-Larruscaín & Basbøll (2013:69).

Our proposal is consistent with the fact that speakers cannot choose to increase the duration of stød *ad infinitum*, the way one may choose to lengthen creaky voice at the end of an utterance. In the same vein, parents addressing their infants may produce extremely long stødless vowels, but they do not similarly lengthen vowels with stød (Bleses et al. 2008). Our account of the laryngeal activity is also consistent with the way stød behaves acoustically: more explicit or less explicit creakiness; variable timing of the onset of creakiness in waveform and spectrum; variable total duration that often makes the creakiness continue well into the following syllable. Furthermore, our proposal is consistent with Fischer-Jørgensen's (1987) EMG-data: the higher vocalis muscle activity in creakiness relative to modal voice increases and decreases (roughly) gradually. It would be quite noteworthy if the actual mechanical change in vocal fold vibration mode were not also (roughly) gradual. We had no concrete evidence to suggest some underlying systematic factor beneath the considerable variability in stød strength and timing, but we conjectured that analysis of stød in non-scripted speech would reveal variation of stød manifestation as a function, *inter alia*, of the degree of prominence on the syllable.

The paper also contains a report on a small and informal but revealing experiment. We excised the word [ˈlɛːˀsɐ] *læser* 'reads' (Figure 2, right) from an utterance and spliced out the [l] and the first half of the vowel, to the left of the dotted red line. On three separate occasions I presented the resulting truncated item (Figure 2, left) to three groups of Danish university students (about 15, 20, and 70 persons, respectively) in a lecture room. I told the audience that they would hear a word
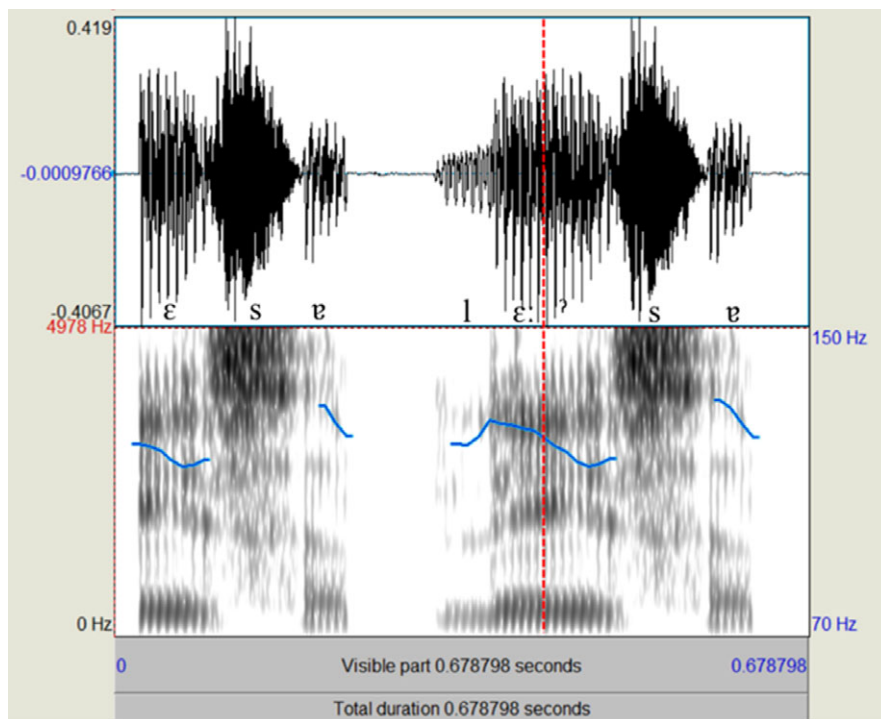
Figure 2. Praat screen dump (Boersma & Weenink 2006): waveform (above), spectrum (below), of [ˈlɛːˀsɐ] *læser* 'reads' on the right and a truncated version of the same word on the left. Male speaker. The vertical dotted line in the right-hand waveform indicates the point of incision. Adapted from Grønnum & Basbøll (2007:197).

spliced out from a larger context and that it might sound somewhat abrupt. I played the truncated word back five times and asked the audience: 'What do you hear? Could this be a Danish word?' Despite the creakiness (admittedly not very explicit though the stød was perceptibly unmistakable) in what was left of the vowel, it was unambiguously of short duration, and apparently duration trumps creakiness when the two are in conflict, because the audience unhesitatingly responded with a word in the Danish lexicon: [ˈɛsɐ] *esser* 'aces (N.PL)'. This is my redacted formulation of our result. Our reasoning in the paper is more complicated and somewhat convoluted.

Similarly, the [l] and the first half of the vowel were removed from the word [ˈlɛːˀnɐ] *læner* 'leans' (Figure 3, right). The creakiness has spilled over into the following nasal consonant. As such, the truncated item (Figure 3, left) has a perfect match in the lexicon, [ˈɛnˀɐ] *ænder* 'ducks', and that is how listeners responded. In other words, the accidental spill-over was perceived at face value as ordinary creakiness in the [n]. If that had not been the case, listeners would have had recourse to another short vowel lexical item, [ˈɛnɐ] *ender* 'ends (PRES/N.PL)'.
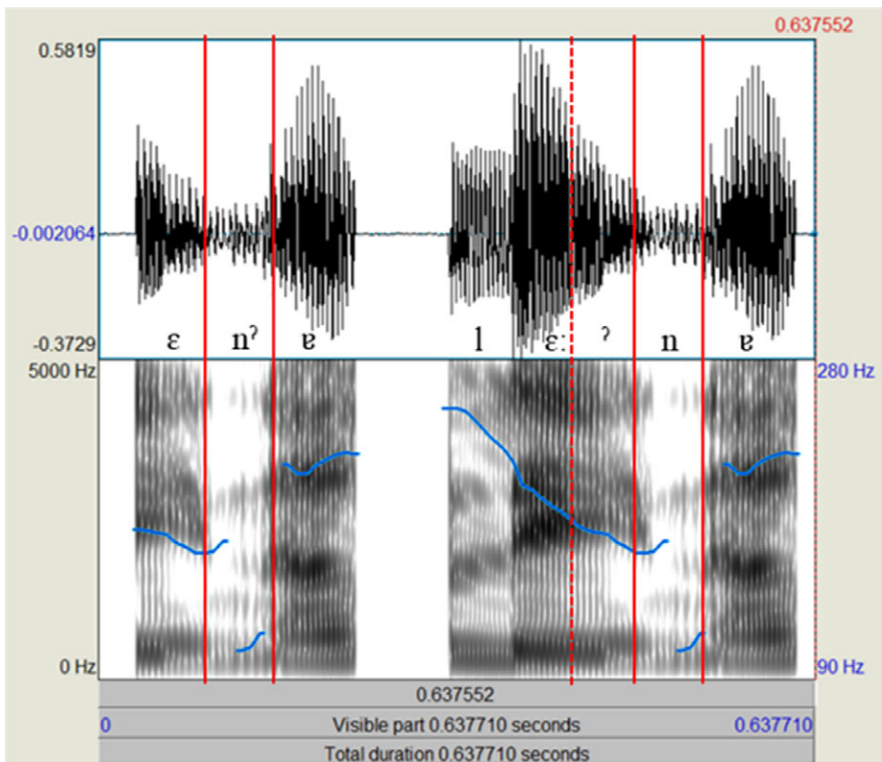
Figure 3. Praat screen dump (Boersma & Weenink 2006): spectrum (above), waveform (below), of [ˈlɛːʔnɐ] *læner* 'leans' on the right and a truncated version of the same word on the left. Female speaker. The solid lines enclose the nasal consonant, and the dotted line indicates the point of incision. Adapted from Grønnum & Basbøll (2007:198).

4.3.5 Grønnum & Basbøll 2012

This paper is not primarily a study in the phonetics of stød, but it contains a section about psycholinguistic issues (Grønnum & Basbøll 2012:38–40) with two observations that add to the phonetic picture of stød.

(i) We noted that words with stød enter a child's lexicon no later than corresponding words without stød. Nor do stød alternations in inflection delay children's acquisition of morphology (Kjærbæk & Basbøll 2010:15, 25).

That is in line with Jørgensen's (1978) observation that two-year-old toddlers make very few stød mistakes (see Section 4.1.11), and I take it to mean that the complex laryngeal articulation involved in stød is not an obstacle to children's acquisition of their mother tongue.

(ii) During the phonetic annotation of the DanPASS corpus (almost ten hours of non-scripted speech, monologues as well as dialogues, by 22 speakers; see Grønnum 2009), the transcribers (myself among them) found that despite highly variable acoustic manifestations, stød was very nearly always clearly identifiable perceptually and – even more noticeably – did not seem to suffer any weakening in less distinct and/or more rapid passages. This is in obvious contrast to the manifestation of most segments. Thus, for instance, obstruents vary from the most clearly pronounced prototypical stops and fricatives to the weakest possible approximants, and they may be deleted altogether. Stød is never likewise deleted.

4.3.6 Grønnum 2015

Fischer-Jørgensen (1987) observed that F_0 is consistently higher at the beginning of syllables with stød. It is in fact the most stable acoustic difference between syllables with and without stød. In Grønnum (2015) I widened the scope of analysis to running, non-scripted speech and included measurements in post-tonic syllables. I culled the data from the DanPASS corpus (Grønnum 2009). In order that any observed differences in vowel onset F_0 be ascribed solely to the stød/non-stød difference, initial consonants in the syllable had to be of similar glottal and oral configurations (voiced or not, aspirated or not, fricative or stop) and the vowels should have the same tongue height. That left a meagre six acceptable words for analysis: [ˈvanˌfalˀ] *vandfald* 'waterfall', [b̥aˈnæːˀnˌpʰalmə] *bananpalme* 'banana palm', [ˈlasd̥ˌb̥iːˀl] *lastbil* 'lorry', [d̥iaˈmanˀd̥ˌmiːnə] *diamantmine* 'diamond mine', [ˈpʰʌsd̥ˌhuːˀs] *posthus* 'post office', and [ˈfɛŋˀsəl] *fængsel* 'prison'. By way of compensation, there were sixteen speakers and often more than one instantiation by the same speaker. In ISOLATED WORDS, onset F_0 in stressed syllables with stød ([ˈnæːˀn], [ˈmanˀd̥]) was significantly higher (by about one semitone) than without stød ([ˈvan], [ˈlasd̥]). There was a similar difference in the post-tonic syllable ([ˌfalˀ], [ˌpʰal]; [ˌb̥iːˀl], [ˌmiː]), but of lesser magnitude (about half a semitone) and not statistically significant. In words extracted from NON-SCRIPTED MONOLOGUES the same trends were observed, but the less constrained production in non-scripted speech, with its expanded frequency range, introduced a considerable dispersion in the measurements and the differences were not statistically significant. The lively exchanges between speakers in the NON-SCRIPTED DIALOGUES caused an even more considerable increase in overall frequency range, and semantic and pragmatic factors would consistently place some words higher in the speakers' range than others, thus obscuring any modest increase in F_0 caused by stød. In a parenthetical note: This is a neat demonstration that read speech and controlled speech materials are justified and necessary when looking at some of the finer details of prosody that are easily obscured in less constrained speech.

I concluded that despite its perceptual robustness, and apart from the stable F_0 difference at syllable onset, there is abundant evidence that the acoustic properties of stød are extremely variable: Vocal fold vibrations may or may not be explicitly irregular. The timing of the irregularity – when present – may be earlier or later in the sonorant syllable rhyme and it may continue into a succeeding post-tonic syllable. Likewise, syllables with stød often exhibit lowering of F_0 toward the end

of the rhyme. In retrospect, I suggest that the results support a characterization of stød as basically physiologically compressed voice that in its weakest form merely increases F_0 in the beginning, whereas stronger compression activates the vocalis muscles and turns the modal vibratory pattern into creakiness and/or lowers F_0. What governs this borderline between compressed modal-like vibratory patterns throughout the syllable versus explicit – occasionally weak – creakiness from some time after vowel onset is, however, an open and intriguing question.

4.4 The most recent studies, 2015–2022

Three studies are concerned with the nature and timing of the creakiness that characterizes stød, and Kirkedal (2016) studied the importance of stød for automatic speech recognition.

4.4.1 Hansen 2015

The point of departure for Gert Foget Hansen's (2015) dissertation is the perceptual equivalence of different acoustic manifestations of stød. He challenges Grønnum & Basbøll's (2007) suggestion that the action of the vocal folds is a ballistic response to a transient neural command. Instead, he proposes a controlled dynamic voice quality gesture, a progressively increasing and decreasing compression of the vocal folds (Hansen 2015:58–62). Whether or not creakiness occurs is a question of the degree of compression. (Note that Hansen and I were mutually unaware of each other's writing in 2015.) The difference between the two proposals, he suggests, is in the control mechanism, and in Hansen's view his hypothesis is better equipped to explain the perceptual equivalence of different manifestations: it is the DYNAMIC CHANGE – the increase and decrease – in compression that Danes perceive as stød, irrespective of its magnitude, that is, whether creakiness occurs or not.

Compressed vocal fold vibrations entail a relatively shorter open period during each vibratory cycle. The acoustic result is a voice source spectrum with harmonics of relatively greater intensity in the upper part of the spectrum, that is, acoustic compression. The intensity relation between the upper and lower parts of the source spectrum can be expressed formally in several different ways. Hansen selected three such formal expressions to apply in the acoustic analysis.

The material consisted of thirty-four words with stød and their non-stød counterparts, inserted in sentences, read four to six times each by one male speaker of Standard Copenhagen Danish.

After a series of painstakingly detailed – token by token – acoustic observations, it turned out that in a considerable number of instances, maximum acoustic compression occurred early in the syllable, before the onset of irregular vibrations. If acoustic compression is a consequence of physiological compression of the vocal folds, and if the observed maximum acoustic compression precedes the onset of creakiness, it follows that increased vocal fold compression cannot unambiguously be made responsible for creakiness. Some other factor must be at play. Hansen felt

compelled to abandon his hypothesis and refrained from speculating any further. He concluded (Hansen 2015:189):

> Når kompression ikke nødvendigvis hænger sammen med uregelmæssige svingninger og med intensitetsdykket sådan som hypotetiseret, så bidrager kompressionsforløbet ikke til at forklare den variabilitet der ses især med hensyn til forekomsten af uregelmæssige svingninger i forbindelse med stød. [When compression is not necessarily correlated with irregular vibrations and with the reduction of intensity as hypothesized, the course of compression does not contribute to explaining the variability observed, particularly with respect to the occurrence of irregular vibrations in connection with stød.]

4.4.2 Kirkedal 2016

Anders Søeborg Kirkedal (2016) found that supplied with information about stød, automatic speech recognition improves – in read as well as non-scripted speech. He tested seventeen acoustic parameters in a search for the one that contributes most to correct identification of stød in the speech signal. The highest ranking among the seventeen – even above measures of compression – are two expressions of phase distortion.

I think this is an interesting finding. Phase distortion is a derivative of phase shift, the time shift between the lowest harmonic in the spectrum and harmonics at higher frequencies. It is a direct reflection of changes in the intricate, three-dimensional vibratory pattern of the vocal folds: front-to-back, side-to-side, bottom-to-top. Any change from the modal vibratory pattern, through minimally compressed vocal fold vibrations to explicit creakiness, will involve phase shifts. But the vibratory pattern is not an independently controlled feature. It is a consequence of changes in other parameters, for instance vocal fold length and tension, the degree of transverse glottis opening, and compression of the vibrating mass in the vertical dimension when the ventricular folds are constricted. I suspect that phase shift and phase distortion are very sensitive measures of change in the vibratory pattern and might be detectable earlier than acoustic evidence of compression and irregular vibrations. If it were easy to extract from the speech signal, phase distortion might be a good indicator of the timing of the stød command.

4.4.3 Esling, Moisik, Benner & Crevier-Buchman 2019

Note that VOICE QUALITY in (the title of) this book refers not only to the vocal folds but encompasses also what is customarily called BASE OF ARTICULATION or ARTICULATORY SETTING. And LARYNGEAL is not to be understood literally either but encompasses the LARYNX as well as the EPILARYNX and the PHARYNX.

The book's central message is that the generally accepted description of speech as a source~filter function – a sound emitted by the glottis and filtered by the resonance chambers above the glottis – is an inadequate dichotomy. Modifications in the epilarynx and the pharynx – the LARYNGEAL CAVITY in the authors' terms – are not merely responsible for modifying resonance frequencies but are active participants in differentiating the many ways the vocal folds may be configured and

vibrate. In other words, the laryngeal cavity is an ARTICULATOR. The LARYNGEAL ARTICULATOR MODEL (LAM) has three components: (i) tongue retraction, (ii) narrowing of the aryepiglottic folds through contraction of the complex sphincter/constrictor musculature, and (iii) raising of the larynx. These three components, in various combinations and to varying degrees, are involved in all the different phonation types, with their different vocal and ventricular fold configurations. Exceptions are aspiration, modal voice, breathy voice, falsetto, and breathy falsetto: they are exclusively properties of the vocal folds, of their different degrees of tension and different degrees of opening between the arytenoids.

Among the accompanying media files are a wealth of laryngoscopic video recordings, including recordings of five stød and contrasting non-stød words spoken by two Danes. They 'illustrate how degrees of laryngeal constriction produce the phonetic effects that accompany instances of stød ... all consistent with narrowing/tightening of the laryngeal constrictor mechanism' (Esling et al. 2019:149, video legend). In the body of the text the authors state that the account of stød in Grønnum (1998), Grønnum, Vazquez-Larruscaín & Basbøll (2013), and Hansen (2015) 'is consistent with our description of laryngeal constrictor action' and that the stød/non-stød distinction 'is accounted for globally by laryngeal constriction and its co-related effects' (Esling et al. 2019:148). We are not told what these co-related effects are, but 'Fischer-Jørgensen (1989) offers a comprehensive instrumental account of stød effects, as does Hansen (2015), corresponding elegantly with our LAM interpretation' (Esling et al. 2019:148). They continue: 'Both also observe that the stops /t, k/ affricate when they occur initially in syllables containing stød – perhaps an enhancement in force across the syllable.' Neither Fischer-Jørgensen nor Hansen make any such statement. Nor could they have because the facts are different: /t/ is always affricated in syllable initial position – stød or no stød – and /k/ is never affricated.

There is no mention in the book of Fischer-Jørgensen's EMG recordings and hence no hint as to how laryngeal constrictor and vocalis muscle activity might interact in stød.

The authors state that, with the LAM model as the point of departure, much remains to be investigated. They note expressly (Esling et al. 2019:15) that the relation between F_0 and the laryngeal articulator is not well understood. Evidently, research in that area would be very relevant to the description of stød.

4.4.4 Peña 2022

Jailyn M. Peña's (2022) paper was published online on 24 February 2022. Accordingly, the 'three quarters of a century' in the title of this review is a liberal interpretation of what – at the latest possible stage in the preparation of the manuscript – became seventy-nine years.

Peña subscribes to the generally accepted view that syllables with stød have two phases, a first phase with modal voicing and high F_0, and a second phase typically with lower F_0, irregular vibrations, and lower intensity. She finds that the sonorant portion of the whole syllable rhyme 'acts as a single unit relative to stød's gestural constellation in Danish' (Peña 2022:2) and that the onset of the stød phase proper is timed relative to the centre of that unit. This is in explicit contradistinction to

Basbøll's mora analysis (see Section 2.2), where stød is a property of the second mora in bimoraic syllables. That is, the creakiness turns up in the second half of a long vowel or in the sonorant consonant after a short vowel. In Basbøll's analysis, any succeeding sonorant consonants, as in [b̥eːˀn] *ben* 'leg' and [jɛlˀm] *hjelm* 'helmet', are irrelevant in the mora account.

Peña analyses four different types of monosyllables with stød (C = any consonant, V = any vowel, S = sonorant consonant, O = obstruent consonant): 21 CVːˀ words (as in [søːˀ] *sø* 'lake'), 40 CVːˀO words (as in [huːˀs] *hus* 'house'), 55 CVːˀS words (as in [d̥æːˀl] *dal* 'valley'), 63 CVSˀ words (as in [nɛmˀ] *nem* 'easy'), and 29 CVS words without stød for reference (as in [sœn] *søn* 'son'). Nine Danish speakers recorded the words, embedded in natural sentences.

F_0 runs the expected course, that is, it begins higher in words with stød than in words without stød. Another expected result is the stød spill-over to the succeeding word, present in 88% of Peña's data.

Peña argues that if mora is the relevant temporal domain, stød should begin in the middle of a long vowel, or at the boundary between a short vowel and the succeeding sonorant consonant. If, on the other hand, the whole sonority rhyme is the relevant domain, stød should begin at the sonority rhyme centre. The predictions of the two approaches will be indistinguishable in CVːˀ and CVːˀO words, where the long vowel constitutes the whole sonority rhyme. But they make different predictions for CVːˀS and CVSˀ words: stød onset will be later than the vowel midpoint in CVːˀS words under the sonority rhyme centre hypothesis. And if the vowel is longer than the consonant, stød onset will come earlier than the V-S boundary in CVSˀ words.

To eliminate the confounding influence from differences in vowel duration in the recorded material, within and across word types, stød onset times are represented as distances (i) to the centre of the long vowel (in CVːˀ, CVːˀO, and CVːˀS words), (ii) to the V-S boundary (in CVSˀ words), and (iii) to the sonority rhyme centre (in CVSˀ and CVːˀS words), and they are expressed as proportions of the total duration of the respective domains. An example is the formula in (1) for the relative distance of stød onset to the long vowel midpoint (Peña 2022:20):

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

The distance measure that yields the least amount of variability as expressed by its standard deviation is considered the best representation of how stød is timed in the syllable.

Table 3 presents the results, adapted from Peña's Table 9 (2022:20). Note that the rightmost column, distance from vowel onset, is my add-on.

Standard deviations are extremely large and the dispersion of the underlying data therefore considerable. That is only to be expected, given what we know about the variability in stød onset timing, from early to late after vowel onset, and anything in between. But standard deviations, even if very large, are nonetheless somewhat smaller in the 'distance-to-sonority-rhyme-centre' means. Peña concludes that the sonority rhyme centre represents the most stable timing relationship between stød and the syllable. That is, morae do not capture the reality of stød onset timing as well as do sonority rhyme centres.

Table 3. Means and standard deviations of normalized stød onset time relative to different points in time in the four syllable types listed on the left. From Peña's Table 9 (2022:20), with the add-on of the distance from vowel onset in the rightmost column. For more details see the text

Word type	Distance to Vː midpoint or to V-S boundary	Distance to sonority rhyme centre	Distance from vowel onset
CVːˀ	9.08 ms; SD = 18.16	identical	77.4 ms
CVːˀO	−4.36 ms; SD = 16.82	identical	66.2 ms
CVSˀ	−12.81 ms; SD = 36.85	−0.50 ms; SD = 15.33	80.1 ms
CVːˀS	19.50 ms; SD = 22.03	−2.51 ms; SD = 14.45	86.2 ms

Peña does not consider the inescapable implication for speech production: To produce different stød onset times in words with sonority rhymes of different duration, for instance [sĝɛlˀ] *skæl* 'dandruff' and [sĝɛlˀm] *skælm* 'rogue (N)', or [oɐˀ] *ord* 'word' and [oɐˀm] *orm* 'worm', the speaker must look ahead and anticipate the end of the sonority rhyme. Then the rhyme must be halved before the speaker can initiate the stød gesture. One wonders if that is really a likely production strategy.

Peña does not quote the non-normalized stød onset times – that is, the distance of stød from the onset of the vowel. But the AVERAGES of this measure can be retrieved from the values in her Table 7 (vowel and consonant durations, 2022:19) and the results in Table 9 (2022:20), reproduced in Table 3 here. For example, to calculate stød onset distance from vowel onset in CVːˀ words, I looked up the average duration of the long vowel in Peña's Table 7 (2022:19): 131 ms. Its midpoint is then at 65.5 ms. Onset distance from the midpoint is given as 9.08 ms. If we feed these values into formula (1) above, we get:

$$(2) \qquad 9.08 = \big((\text{stød onset} - 65.5)/131\big) \times 100$$

And then we rewrite the expression and find the distance of stød onset from vowel onset:

$$(3) \qquad \text{stød onset} = \big(9.08/100\big) \times 131 + 65.5 = 77.4 \text{ ms}$$

Likewise for the remaining three word types. This procedure does not yield standard deviations, so we cannot know how stød onset times relative to vowel onset would rank statistically compared with sonority rhyme centre measures.

I am inclined to believe that a fixed reference for the control and timing of creakiness in syllables with stød, like the onset of the stressed vowel, is a more likely candidate than an anchor that floats in time as a function of the specific segmental composition of the syllable sonority rhyme. This is an empirical question deserving to be settled.

5. Summary

As far as I know, there are no further published empirical phonetic stød investigations since 1944 other than those reviewed in Section 4 and whose results are summarized below.

5.1 Acoustic properties that distinguish syllables with stød

The list of differentiating qualities is long.

(i) F_0 is higher at the onset of the vowel (Vihman 1971; Petersen 1973; Thorsen 1974; Fischer-Jørgensen 1987; Grønnum 2015; Peña 2022).

(ii) The voice source spectrum contains relatively more energy in the upper part of the spectrum (Fischer-Jørgensen 1987; Hansen 2015).

(iii) Maximum acoustic compression often occurs early in the syllable before the onset of irregular vocal fold vibrations (Hansen 2015).

(iv) Spectral phase shift is more pronounced (Kirkedal 2016).

(v) Modal voice typically – but not universally – turns into creakiness at some point during the sonorant syllable rhyme (Lauritsen 1968; Vihman 1971; Petersen 1973; Fischer-Jørgensen 1987; Grønnum & Basbøll 2001, 2007; Peña 2022).

(vi) F_0 typically – but not universally – falls toward the end of the sonorant syllable rhyme (Smith 1944; Lauritsen 1968; Vihman 1971; Petersen 1973; Fischer-Jørgensen 1987).

(vii) Intensity nearly always falls toward the end of the sonorant syllable rhyme (Smith 1944; Lauritsen 1968; Petersen 1973; Fischer-Jørgensen 1987; Peña 2022).

(viii) Creakiness typically spills over into a succeeding syllable if its onset is a vowel or a sonorant consonant (Lauritsen 1968; Grønnum & Basbøll 2007; Peña 2022).

5.2 Physiological properties that distinguish syllables with stød

This list is a little longer.

(i) The expiratory musculature (the diaphragm) contracts and relaxes abruptly (Smith 1944). Nobody has since verified that, unless the tendency for greater oral airflow in syllable initial unvoiced consonants (Smith 1944; Fischer-Jørgensen 1987) is to be explained by higher expiratory activity.

(ii) Oral airflow diminishes toward the end of the syllable (Fischer-Jørgensen 1987).

(iii) Subglottal pressure is higher, indicating increased resistance in the glottis (not necessarily greater expiratory pressure) that in its turn derives from compression of the vocal folds (Smith 1944; Ringgaard 1960; Fischer-Jørgensen 1987).

(iv) The anterior parts of the vocal folds are visually compressed at the end of the syllable (Fischer-Jørgensen 1987).

(v) The cricothyroid muscle is active, its maximum activity preceding the onset of the vowel, resulting in higher F_0 at vowel onset (Fischer-Jørgensen 1987).

(vi) The aryepiglottic sphincter musculature contracts, specifically constricting the ventricular folds (Esling et al. 2019) – perhaps more so in speakers with weaker vocalis activity (Fischer-Jørgensen 1987).

(vii) The vocalis muscle is active in creakiness (Faaborg-Andersen 1957; Fischer-Jørgensen & Hirose 1974a, 1974b) – perhaps less so in speakers with more ventricular fold constriction (Fischer-Jørgensen 1987).

(viii) The creakiness in syllables with stød cannot be sustained *ad infinitum* (Grønnum & Basbøll 2007).

(ix) The contact area between the tongue dorsum and the hard palate is larger, indicative of the expenditure of more articulatory energy during creakiness (Fischer-Jørgensen 1987).

Considering the complexity of the acoustics and physiology of stød, as summarized above, it is thought-provoking that it should be acquired so early in children's linguistic development. Two-year-old children make hardly any stød mistakes (Jørgensen 1978).

5.3 Perceptual properties

The perception of stød is a much less explored field.

(i) Lauritsen (1968) observed that stød is audible in whisper due to reduction of noise intensity. This is something any Dane from north of the stød border can verify.

(ii) Grønnum & Basbøll (2002) found that vowels with stød were NOT perceived to have two parts or phases.

(iii) Grønnum & Basbøll (2003) found that listeners by and large perceive the onset of stød to coincide with stressed vowel onset.

(iv) Grønnum & Basbøll (2012) repeated the observation from the annotation of the DanPASS corpus that stød is perceptually very robust and withstands considerable variation in tempo and clarity of pronunciation.

(v) Stød is not heard in *bel canto* song. Fischer-Jørgensen (1987) mentions this fact – not that she performed any experiment – but this is something any Dane from north of the stød border is able to verify if only he can carry a tune.

6. Discussion

Is it possible to make an integrated whole of all the facts presented above? Three major topics invite speculation.

6.1 F_0 and acoustic compression

The first question concerns the RELATION BETWEEN F_0 AND THE ACOUSTIC COMPRESSION that results from a longer glottal closure time in each vibratory cycle. Fischer-Jørgensen (1987) found that the peak in cricothyroid muscle activity preceded the onset of the vowel in syllables with stød, accounting for the high F_0 at vowel onset. She also found acoustic compression to mostly coincide with creakiness. That is different from Hansen (2015) who found that peak acoustic compression often preceded the onset of creakiness. I do not know how to reconcile

the two accounts. How different are the physical facts behind 'mostly coincide' and 'often precede'? For the sake of the argument, I will adopt Hansen's (2015) account.

Fischer-Jørgensen leaves no doubt that creakiness is a result of increased vocalis muscle activity. If that is so, and if acoustic compression peaks earlier than the onset of non-modal vibrations, acoustic compression and non-modal vocal fold vibrations cannot have the same cause. But perhaps high F_0 and acoustic compression can both be attributed to the lengthening and tensing induced by the cricothyroid muscles. Even more tentatively: Could the acoustic compression be the independent variable, making higher F_0 the dependent variable? If so, any argument that high F_0 at syllable onset is a separate tonal gesture in stød production becomes void.

If indeed acoustic compression early in syllables with stød is the salient feature – common to all manifestations of stød and in conjunction with high F_0 – it could be THE property listeners extract from the stream of speech. Not the higher F_0 *per se*, and not the creakiness that is not always present anyway. That would explain the perceptual equivalence of the many different acoustic manifestations of stød. It may not be straightforward, but I believe the proposal is testable in experiments with synthetic speech where one can manipulate F_0, acoustic compression, and creakiness independently.

6.2 Ventricular folds and vocal folds

The second major question is the COORDINATION OF THE CONSTRICTION OF THE VENTRICULAR FOLDS AND VOCALIS MUSCLE ACTIVITY. Are they two independent gestures? Is there a trade-off relation between them? Fischer-Jørgensen's speaker JR apparently did not have any stød-specific vocalis muscle activity at all. The weak creakiness in his stød appeared due solely to the shortening and thickening of the vocal folds, resulting from constriction of the ventricular folds, presumably due to contraction of the aryepiglottic sphincter musculature.

6.3 Cricothyroid, ventricular folds and vocal folds

The third conundrum is HOW CRICOTHYROID ACTIVITY RELATES TO THE VENTRICULAR FOLD CONSTRICTION AND THE VOCALIS MUSCLES. Is one neural command imaginable that first induces the cricothyroid to create acoustic compression (hence also higher F_0), and then after some variable time lag triggers the sphincter and vocalis muscles to produce creakiness? If there are two separate neural commands, how are they coordinated? It must be a foolproof control, because we have never seen the non-creaky and the creaky intervals reversed in time. It will take a combination of Fischer-Jørgensen's (1987) EMG study (including the sphincter musculature) and Esling et al.'s (2019) laryngoscopic experiments to shed light on this question.

6.4 Two simpler questions

Why would there be more articulatory energy at the onset of syllables with stød? That is something that perhaps needs to be explored in more comprehensive experiments. But increased articulatory effort may account for an anecdotal and very

commonplace observation among Danish phoneticians: Ask linguistically naïve speakers of Danish what the difference is between pairs of words with and without stød, for instance [ˈmuːˀsən] *musen* 'the mouse' and [ˈmuːsən] *musen* 'the muse'. They invariably reply along the lines of *Der er ligesom mere tryk på* [ˈmuːˀsən] 'There is kind of like more stress on the mouse.'

Petersen ([1973](#)) found that out of context [ˈlɛːˀsɐ] *læser* 'reads' was often identified as stødless [ˈlɛːsɐ] *læser* 'reader', whereas only rarely was stødless [ˈlɛːsɐ] identified as [ˈlɛːˀsɐ] *læser* 'reads'. This irreversibility should be verified in a more comprehensive experiment.

Clearly, there is work to be done!

References

Abrahams, Henrik. 1949. *Etudes phonétiques sur les tendances évolutives des occlusives germaniques*. Aarhus: Universitetsforlaget.

Basbøll, Hans. 1988. The Modern Danish stød and phonological weight. In Pier Marco Bertinetto & Michele Loporcaro (eds.), *Certamen Phonologicum: Proceedings from the Cortona Phonology Meeting*, 119–152. Turin: Rosenberg & Sellier.

Basbøll, Hans. 1989. Phonological weight and Italian raddoppiamento fonosintattico. *Rivista di Linguistica = Italian Journal of Linguistics* **1**, 5–31.

Basbøll, Hans. 2005. *The Phonology of Danish*. Oxford: Oxford University Press.

Basbøll, Hans. 2008. Stød, diachrony and the non-stød model. *North-Western European Language Evolution (NOWELE)* **54/55**, 147–189.

Basbøll, Hans. 2014. Danish stød as evidence for grammaticalization of suffixal positions in word structure. *Acta Linguistica Hafniensia* **46**(2), 137–158.

Basbøll, Hans. 2021. Danish stød in the light of morae, the Weight Law and sonority (strength): A personal view. In Patrizia Noël Aziz Hanna & Laura Catherine Smith (eds.), *Linguistic Preferences* (Trends in Linguistics: Studies and Monographs 358), 143–170. Berlin: De Gruyter.

Basbøll, Hans. 2022. Phonology and phonetics, a recurrent theme in European structuralisms: The case of Otto Jespersen and André Martinet. In Lorenzo Cigana & Frans Gregersen (eds.), *Studies in Structuralism*, 247–282. Copenhagen: Det Kongelige Danske Videnskabernes Selskab.

Bernoulli, Daniel. 1738. *Hydrodynamica, Sive de Viribus et Motibus Fluidorum Commentarii* [Hydrodynamics, or commentaries on the forces and motions of fluids]. Strasbourg: Johannis Reinholdi Dulseckeri.

Bertelsen, Henrik. 1920 (1979). *Danske Grammatikere IV* [Danish grammarians IV]. Copenhagen: Gyldendal. Reprinted 1979 by Det Danske Sprog og Litteraturselskab, Copenhagen: Reitzels Boghandel.

Bleses, Dorthe, Werner Vach, Malene Slott, Sonja Wehberg, Pia Thomsen, Thomas O. Madsen & Hans Basbøll. 2008. Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language* **35**, 619–650.

Boersma, Paul & David Weenink. 2006. Praat: Doing phonetics by computer (Computer program praat5404_win64). Available at https://www.praat.org/.

Charter, David. 2013. Death of a language: Last ever speaker of Livonian passes away aged 103. *The Times*, 5 June 2013. Available at https://www.thetimes.co.uk/article/death-of-a-language-last-ever-speaker-of-livonian-passes-away-aged-103-8k0rlplv8xj.

Divell, Mike. 2010. Introduction to the physics of waves and sound. Available at https://people.finearts.uvic.ca/~aschloss/course_mat/MU207/divell.physics_of_waves.pdf.

Ekblom, Richard. 1933. Om de danska accentarterna [On the Danish accents]. *Språkvetenskapliga Sällskapets i Uppsala Förhandlingar, Uppsala Universitets årsskrift* **5**(1), 1–12.

Esling, John Henry, Scott Reid Moisik, Allison Benner & Lise Crevier-Buchman. 2019. *Voice Quality: The Laryngeal Articulator Model* (Cambridge Studies in Linguistics 162). Cambridge, UK: Cambridge University Press. Media files available at https://drive.google.com/drive/folders/1-Te5DhQ6NApJoDXJOePR5c_HUOBLJsec.

Faaborg-Andersen, Knud. 1957. *Electromyographic Investigation of Intrinsic Laryngeal Muscles in Humans: An Investigation of Subjects with Normally Movable Vocal Cords and Patients with Vocal Cord Paresis*. Copenhagen: Møllers Bogtrykkeri.

Fischer-Jørgensen, Eli. 1955. Om vokallængde i dansk rigsmål [On vowel length in Danish standard language]. *Nordisk Tidsskrift for Tale og Stemme* **15**, 33–56.

Fischer-Jørgensen, Eli. 1974. Electromyographic investigation of Danish consonants, stress, and stød. *Annual Report of the Institute of Phonetics, University of Copenhagen* **8**, 203–206. Available at https://tidsskrift.dk/ARIPUC.

Fischer-Jørgensen, Eli. 1987. A phonetic study of the stød in Standard Danish. *Annual Report of the Institute of Phonetics, University of Copenhagen* **21**, 55–265. Available at https://tidsskrift.dk/ARIPUC.

Fischer-Jørgensen, Eli. 1989a. *Phonetic Analysis of the Stød in Standard Danish*. Turku: University of Turku.

Fischer-Jørgensen, Eli. 1989b. Phonetic analysis of the stød in Standard Danish. *Phonetica* **46**, 1–59.

Fischer-Jørgensen, Eli & Hajime Hirose. 1974a. A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production. Haskins Laboratories: Status Report on Speech Research SR 39/40, 230–254.

Fischer-Jørgensen, Eli & Hajime Hirose. 1974b. A note on laryngeal activity in the Danish 'stød'. Haskins Laboratories: Status Report on Speech Research SR 39/40, 255–259.

Grønnum, Nina. 1998. Danish. *Journal of the International Phonetic Association* **28**, 99–105.

Grønnum, Nina. 2005. *Fonetik og Fonologi*. Copenhagen: Akademisk Forlag.

Grønnum, Nina. 2009. A Danish phonetically annotated spontaneous speech corpus (DanPASS). *Speech Communication* **51**, 594–603. Corpus available at https://danpass.hum.ku.dk.

Grønnum, Nina. 2015. F0, voice quality, and Danish stød revisited. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th Congress of Phonetic Sciences* (ICPhSXVIII), paper number 0423. Glasgow: The University of Glasgow. Available at https://www.internationalphoneticassociation.org/icphs-proceedings-ICPhS2015/Papers/ICPHS0423.pdf.

Grønnum, Nina & Hans Basbøll. 2001. Consonant length, stød and morae in Standard Danish. *Phonetica* **58**, 230–253.

Grønnum, Nina & Hans Basbøll. 2002. Stød and length: Acoustic and cognitive reality? In Bernard Bel & Isabelle Martin (eds.), *Proceedings of the Speech Prosody 2002 Conference*, 355–358. Aix-en-Provence: Laboratoire Parole et Langage.

Grønnum, Nina & Hans Basbøll. 2003. Stød and length in Standard Danish: Experiments in laboratory phonology. In Maria-Josep Solé, Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 455–458. Universitat Autónoma de Barcelona.

Grønnum, Nina & Hans Basbøll. 2007. Danish stød: Phonological and cognitive issues. In Maria-Josep Solé, Patricia Speeter Beddor & Manjaru Ohala (eds.), *Experimental Approaches to Phonology*, 192–206. Oxford: Oxford University Press.

Grønnum, Nina & Hans Basbøll. 2012. Danish stød: Towards simpler structural principles? In Oliver Niebuhr (ed.), *Understanding Prosody: The Role of Context, Function, and Communication*, 27–46. Berlin: De Gruyter.

Grønnum, Nina, Nicolai Pharao & Hans Basbøll. 2020. Stød in unexpected morphological contexts in Standard Danish: An experimental approach to sound change in progress. *Nordic Journal of Linguistics* **43**, 147–180.

Grønnum, Nina, Miguel Vazquez-Larruscaín & Hans Basbøll. 2013. Danish stød: Laryngealization or tone. *Phonetica* **70**(1–2), 66–92.

Hansen, Gert Foget. 2015. Stød og stemmekvalitet: En akustisk–fonetisk undersøgelse af ændringer i stemmekvaliteten i forbindelse med stød [Stød and voice quality: an acoustic phonetic investigation of changes

in voice quality in connection with stød]. Ph.D. dissertation, Det Humanistiske Fakultet, Københavns Universitet, Copenhagen.

Høysgaard, Jens Pedersen. 1743. *Concordia res parvæ crescunt, eller Anden Prøve af Dansk Orthographie* [In concord small things grow, or second specimen of Danish orthography]. Copenhagen. Reprinted in Bertelsen (1920:219–248).

Høysgaard, Jens Pedersen. 1747. *Accentuered og Raisonnered Grammatica* [Accentuated and reasoned grammar]. Copenhagen. Reprinted in Bertelsen (1920:249–488).

Høysgaard, Jens Pedersen. 1769. *Første Anhang til den Accentuerede Grammatika* [First supplement to the accentuated grammar]. Copenhagen. Reprinted in Bertelsen (1920:507–550).

Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautegesetze*. Uppsala: Språkvetenskapliga Sällskapets i Uppsala Förhandlingar 1940–1942.

Jespersen, Otto. 1897–1899. *Fonetik*. Copenhagen: Det Schubotheske Forlag.

Jespersen, Otto. 1913. *Lehrbuch der Phonetik: Zweite Auflage*. Leipzig and Berlin: B.G. Teubner.

Jørgensen, John. 1978. Stød and unexpected substitutions in the speech of Danish 2-year-old children. *Annual Report of the Institute of Phonetics, University of Copenhagen* 12, S25–S29. Available at https://tidsskrift.dk/ARIPUC.

Kiparsky, Paul. 2017. Livonian stød. In Wolfgang Kehrein, Björn Köhnlein, Paul Boersma & Marc van Oostendorp (eds.), *Segmental Structure and Tone*, 195–210. Berlin: De Gruyter.

Kirkedal, Andreas Søeborg. 2016. Danish stød and automatic speech recognition. Ph.D. dissertation, Copenhagen Business School.

Kjærbæk, Laila & Hans Basbøll. 2010. Morfologisk udvikling hos danske børn i alderen 0–10 år: Substantivernes pluralisdannelse. [Morphological development in children aged 0–10 years: Noun plural formation]. In Hans Vejleskov (ed.), *Børnesprog: Fra det 12. Nordiske Symposium om Børnesprogsforskning* [Children's language: From the 12th Nordic Symposium on Children's language research], 15–27. Copenhagen: Forlaget UCC.

Lauritsen, Margaret. 1968. A phonetic study of the Danish stød. Project on Linguistic Analysis, Phonology Laboratory Report, Department of Linguistics, University of California Berkeley, Reports Second Series 7, D1–D12.

Peña, Jailyn M. 2022. Stød timing and domain in Danish. *Languages* 7, 50. Available at https://doi.org/10.3390/languages7010050.

Petersen, Pia Riber. 1973. An instrumental investigation of the Danish 'stød'. *Annual Report of the Institute of Phonetics, University of Copenhagen* 7, 195–234. Available at https://tidsskrift.dk/ARIPUC.

Ringgaard, Kristian. 1960. *Vestjysk stød* [West Jutlandic stød]. Aarhus: Universitetsforlaget i Aarhus.

Ringgaard, Kristian. 1973. *Danske dialekter: En kortfattet oversigt* [Danish dialects: A brief overview]. Copenhagen: Akademisk Forlag.

Rousselot, l'abbé Pierre-Jean. 1897–1901. *Principes de phonétique expérimentale* [Principles of experimental phonetics]. Tome 1 (1897), Paris & Leipzig: H. Welter. Tome 2 (1901, new edition 1925), Paris: Didier & Toulouse: Privat.

Smith, Svend. 1944. Bidrag til Løsning af Problemer vedrørende Stødet i dansk Rigssprog: En eksperimentalfonetisk Studie [Contributions to the solution of problems concerning the stød in Standard Danish: An experimental phonetic investigation]. *Nordisk Tidsskrift for Tale og Stemme* 8(1–2), Copenhagen: Kaifer.

Thorsen, Nina Grønnum. 1974. Acoustical and perceptual properties of the Danish stød. *Annual Report of the Institute of Phonetics, University of Copenhagen* 8, 207–213. Available at https://tidsskrift.dk/ARIPUC.

Vihman, Marilyn. 1971. Livonian phonology, with an appendix on stød in Danish and Livonian. Ph.D. dissertation, University of California, Berkeley.