F_0, VOICE QUALITY, AND DANISH STØD REVISITED

Nina Grønnum

Department of Nordic Studies and Linguistics, University of Copenhagen

ninag@hum.ku.dk

ABSTRACT

Danish stød is a syllable prosody, hitherto described as a kind of creaky voice with tonal side effects. One such is the well established, though not ubiquitous, abrupt lowering of F_0 towards the end of the syllable. Eli Fischer-Jørgensen found, in the 1970s and 1980s, a difference also in the beginning of stressed syllables: F_0 is higher in the beginning of syllables with stød. Confirming her findings 40 years later was straightforward in read word lists and as a clear trend in non-scripted monologues. In dialogues, however, the modest effect of stød on onset F_0 is overshadowed by the greater variation in overall frequency range.

Higher F_0 onset is present also in unstressed syllables with stød, though of lesser magnitude.

The results support a characterization of Danish stød as compressed, rather than creaky voice. They also demonstrate the viability of controlled speech materials for investigations of finer prosodic detail.

Keywords: Fundamental frequency, voice quality, stød, Danish.

1. INTRODUCTION

In her very comprehensive study of stød (EMG of various laryngeal muscles, glottography, oral airflow, subglottal pressure, intensity, F_0, inter alia) Eli Fischer-Jørgensen [4, 5] found that in about half of the 1700 items she examined, from a total of 15 speakers, stød is accompanied by a lowering of F_0 at the end of the stressed syllable. She attributed that to increased constriction of the glottis. Smith [26] and Riber Petersen [25] made similar acoustic observations, as did Hans Basbøll and I in more recent investigations [12-15]. But EFJ notes very explicitly that the most stable difference between words with and without stød (present in 96% of the word pairs analysed) consists in a *higher* F_0 level in the *beginning* of the stressed syllable with stød. This latter observation has been overlooked ever since, also – embarassingly – by myself.

There were four types of words, with and without stød, in EFJ's data: monosyllables, disyllables with long stressed vowel, disyllables with short stressed vowel followed by a sonorant consonant, plus a few trisyllabic nonce words, 178 word pairs in all with 8-10 tokens per word. The majority of these words were read aloud from word lists. Measurements were based on mingographic print-outs from an analog pitchmeter. To establish averages, EFJ transferred F_0 curves of words of a given type, say monosyllables, for a given speaker, from the mingograms to tracing paper and drew a visually determined average through the stød and non-stød tracings, respectively. The difference in Hz between two average tracings was measured at the point in time within the first 100 msecs where the separation between them was at maximum. No attempt was made to average over speakers.

There is considerable dispersion in EFJ's data. F_0 in the beginning of stressed syllables with stød is higher by between 3 and 31 Hz, and this wide range of values is not unambiguously correlated with any other variable, be it gender or word type.

Apart from calling attention to EFJ's significant investigation, 3-4 decenniums old by now, the analysis presented here widens the scope, introducing non-scripted speech and including unstressed syllables.

For a brief introduction to the phonetics of Danish, see [10]. Recent accounts of the phonetics of stød can be found in [12-15], and stød in all its amazing phonological and morphological complexity is described in Basbøll [1].

2. MATERIAL, SPEAKERS AND PROCEDURE

The material for this investigation comes from the monologues and dialogues in the DanPASS corpus, [11]. The monologues are instructions to an unseen receiver who could offer no feedback (as if spoken to the recipient's telephone answering service), calling for a fairly distinct style of speech. One of these instructions directed the listener from four different points of departure to various locations on a map of a fictitious town. – The participants in the dialogues communicated over headphones. Each had a hand-drawn map

of a landscape with a number of distinct landmarks. The giver had a route on the map; the follower did not. Their task was to collaborate to reproduce the giver's trail on the follower's map. The two maps were not exactly identical, which made for a lively exchange of questions and answers. There were four different pairs of maps.

To eliminate any influence from differences in onset consonant type or in vowel height on F_0, syllable initial consonants should be of similar type (similar glottal configuration and similar vocal tract constriction), and the vowels of similar tongue/jaw height in any pair of words to be directly compared, cf. [16, 19, 27]. Perfect segmental matches with stød and non-stød in both stressed and posttonic syllables do not exist in the corpus, but there are target words which satisfy conditions to an acceptable degree. There are four such words in the dialogues: *vandfald, bananpalme; lastbil, diamantmine* 'water fall, banana palm; truck, diamond mine;' and two in the monologues: *posthus, fængsel* 'post office, jail.'

[ˈvanˌfalˀ ˌbaˈnæːˀnˌpʰalmə]
[ˈlasd̥ˌbiːˀl ˌd̥iaˈmanˀd̥ˌmiːnə]
[ˈpʰʌsd̥ˌhuːˀs ˈfɛŋˀsl̥]

The latter two are less satisfactory, given the different posttonic onset consonants and the absence of a vowel sound proper in posttonic *-sel*, but no other two words come closer. Henceforth the words are quoted in orthography, with an apostrophe to designate stød. Except *fæn'gsel*, they are compounds with stress on the first part, and reduced (secondary) stress on the second part.

The landmark names in the dialogue maps occur twice each in a list of words which participants read aloud after completion of the navigation, but otherwise they are very unevenly distributed across speakers, absent with some, occurring many times with others: Some speakers are very voluble, others tend to be brief and matter-of-fact. Givers also tend to mention a landmark more often than do followers. In the monologues, different participants chose different routes through the town, and not everybody passed by the post office and/or the jail.

In standard Copenhagen Danish F_0 is prototypically higher in the posttonic than in the preceding stressed syllable, cf. [9], but pragmatic effects may occasionally induce a fall from stressed to posttonic syllable. Words in the corpus annotated in this manner would blow the variance on posttonic F_0 onset measurements out of all proportion and were excluded. Likewise, words marked as having extra prominence (which affects F_0 of the stressed syllable and particularly the magnitude of the F_0 rise to the posttonic) were left out. Even so, there are generally enough instances of each word to permit viable statistical analysis of the results.

All 16 speakers in the monologues (five women, 11 men), and 14 in the dialogues (seven of each sex) were selected for analysis here. Four women and six men appear in both recordings. Measurements were performed in Praat [3].

EFJ measured F_0 differences within the first 100 msecs after stressed vowel onset. Many syllable rhymes, particularly in posttonic syllables, in the DanPASS corpus are not that long. To ensure identical analysis conditions, the average over the first three complete vibratory cycles (bypassing initial perturbations in the transition from the preceding consonant, if any) were read off the Praat screen in semitones (ST) re 100 Hz. To calculate grand means over all speakers, the lowest ST value observed in each participant was transposed to zero and all other values adjusted accordingly.

3. RESULTS

3.1. Words read in isolation

EFJ's findings are clearly reproduced here, cf. Table 1.

Table 1: Vowel onset values in semitones in four words read in isolation.

	vand	fal'd	intvl
mean	**1,25**	**6,63**	**5,38**
N	27	27	27
var	1,179	4,598	3,392
	na'n	pal	
mean	**2,45**	**6,11**	**3,66**
N	27	27	27
var	1,990	2,305	1,660
	last	bi'l	
mean	**1,65**	**6,10**	**4,45**
N	28	28	28
var	2,125	6,244	5,074
	man't	mi	
mean	**2,62**	**5,87**	**3,25**
N	28	28	28
var	2,998	4,593	3,172

Student's *t*-tests (two-tailed, unpaired, equal or unequal variance as the case may be) prove the F_0 onset of *-na'n-* with stød to be significantly higher than that of stødless *vand-* (by 1,2 ST, $p<0,01$), and *-man't-* onset is significantly higher

than *last-* (by 0,97 ST, p<0,05). The posttonic syllables exhibit the same trend, but the differences between stød and non-stød are rather small (0,52 ST, NS and 0,23 ST, NS, respectively).

F_0 at the onset of the posttonic syllable is the peak of a rising F_0 movement. But the onset of the preceding stressed syllable is not normally the lowest value in this F_0 pattern. It initiates a slight fall which only turns into a rise about halfway through the stressed syllable rhyme. Therefore, the F_0 intervals listed in the three tables here do not adequately reflect the magnitude of the perceived rise in pitch between the two syllables, which is anchored to the low turning point rather than the preceding, slightly higher, onset of the stressed syllable.

3.2 Words in context: non-scripted monologues

Table 2: Vowel onset values in semitones in two words from a non-scripted monologue.

	post	hu's	intvl
mean	**4,20**	**8,47**	**4,27**
N	15	15	15
var	8,97	12,27	5,837
	fæn'g	sel	
mean	**5,93**	**6,47**	**0,54**
N	17	17	17
var	9,57	12,52	2,057

The stressed syllable with stød, *fæn'g-*, has a higher F_0 onset than the one without stød, *post-*, and even more so than in the isolated words (1,73 ST, NS). The trend is the same in the posttonic syllables, and the difference between *-hu's* and *–sel* is substantially larger (2 ST, NS) than the corresponding values in Table 1 (0,52 ST and 0,23 ST, respectively). The less constrained production in non-scripted speech introduces a considerable variability in the data, as reflected by the large variances.

3.3 Words in context: non-scripted dialogues

The less distinct pronunciation in dialogues vis-a-vis isolated words read from a list entails smaller F_0 rises to the posttonic syllables. Thus, the intervals are all larger in Table 1 than in Table 3 by 2,30 ST, p<0,001; 1,37 ST, p<0,001; 0,81 ST, NS, and 1,49 ST, p<0,01, respectively.

The stressed syllables without stød, *vand-* and *last-*, have higher F_0 onsets than do *-na'n-* and *-man't-* by 1,09 ST, p<0,05 and 0,89 ST, NS, respectively. There can be no reason in the words themselves for this reversal – vis-a-vis the data in Table 1. It must be due to the wider context of the recordings.

Table 3: Vowel onset values in semitones in four words from non-scripted dialogues.

	vand	fal'd	intvl
mean	**3,51**	**6,59**	**3,08**
N	47	47	47
var	3,948	8,009	5,263
	na'n	pal	
mean	**2,42**	**4,71**	**2,29**
N	22	22	22
var	2,790	3,485	0,715
	last	bi'l	
mean	**3,35**	**6,99**	**3,64**
N	53	53	53
var	5,138	12,194	5,333
	man't	mi	
mean	**2,46**	**4,22**	**1,76**
N	32	32	32
var	7,676	8,707	2,560

Compare *vand-* and *last-* in Tables 3 and 1. They are located higher in the frequency range in the dialogues than in isolation (by 2,26 ST, p<0,001 and 1,70 ST, p<0,001, respectively). The same is not true of the stressed syllables of *bana'npalme* and *diaman'tmine* which appear at nearly the same frequencies in the two tables (2,45/2,42 ST and 2,62/2,46 ST, respectively). In other words, the reason the stressed syllables without stød have higher F_0 onset values than those with stød in the dialogues is that everything else is not equal: they are located higher in the overall frequency range.

Why are *vandfal'd* and *lastbi'l* located higher in the range than *bana'npalme* and *diaman't-mine*? (1) The context and the dialogue exercise were novel to the participants, something which may have caused them to speak in a slightly higher pitch range early in the recordings, until they found their bearings, and *vandfal'd* belongs in the first set of maps. (2) The truck (*lastbi'l*) is completely absent from the follower's map. That may have caused a livelier exchange, in higher pitched voices, than in the negotiation of other landmarks. – This line of reasoning introduces a pragmatic issue, however, which is beyond the scope of this paper and will not be pursued any further.

Since the posttonic syllables with stød, *-fal'd* and *-bi'l*, occur in those same two words which are raised in the frequency range, it is only to be expected that they should have higher onsets than *-pal-* and *-mi*, as indeed they do, and the differences are also considerably larger than in

the isolated words in Table 1 (1,88 ST, p<0,01 and 2,77 ST, p<0,01, respectively).

3.4 How far into the syllable?

How far into the syllable does the difference at vowel onset extend? Here is an indication: The two stressed syllables *vand-* and *-man't-* compare favourably: the same short vowel [a] with a well defined transition to the same postvocalic consonant [n]. Table 4 presents average F_0 values at the stressed vowel offset in *vandfald* and *diamantmine*, read in isolation.

Table 4: Vowel offset values in semitones in two stressed syllables from words read in isolation.

	vand	man't
mean	**1,59**	**3,96**
N	28	28
var	2,083	4,598

The difference (2,37 ST, p<0,001) is larger than the difference at vowel onset (1,37 ST, p<0,01), cf. Table 1. – Given the higher location of *vandfal'd* in the overall frequency range in the dialogues, cf. above, there is no point in repeating vowel offset measurements in these words in dialogue context.

If these observations generalize to a larger set of data, it would appear that the F_0 difference between stød and non-stød increases through the syllable, unless counteracted by increasing vocal cord compression, resulting in creak and/or F_0 lowering.

3.4 Summary of the results

There is no doubt – as shown already by Eli Fischer-Jørgensen – that syllables with stød have higher F_0 in the beginning than do syllables without stød, ceteris paribus. The effect is present in both stressed and unstressed syllables, but somewhat less explicit without stress. However, it is a rather modest difference, of an order of magnitude of one semitone, and it is small enough to be overshadowed by the greater variability in other prosodic parameters in non-scripted speech, notably overall F_0 speaking range. – This is a neat demonstration that read speech and controlled speech materials are justified and necessary in the analysis of some of the finer details of prosody. Automated datamining of the complete DanPASS corpus would probably not have revealed any significant increase of F_0 at the onset of syllables with stød.

4. DISCUSSION

In spite of its perceptual robustness, and apart from the small but stable F_0 difference at syllable onset, the acoustic properties of Danish stød are extremely variable, cf. [12-15]: Vocal fold vibrations may or may not be explicitly irregular; the timing of the irregularity – when present – may be earlier or later in the syllable rhyme, and it may continue into a succeeding posttonic syllable. Likewise, stød syllables may or may not exhibit the often quoted lowering of F_0 towards the end of the rhyme.

The acoustic properties of stød are very similar to what Blankenship [2] found in Mazatec and Mpi. She suggests that *laryngealization* is a more appropriate concept than creaky voice, as do Garellek and Keating [6], because laryngealized vowels do not consistently have an audible creak nor display irregular glottal pulses on a spectrogram. Note also the observation in Gerfen and Baker [7] that laryngealization in Coatzospan Mixtec is highly variable within and across speakers and often realized with very subtle F_0 and amplitude cues.

In [15] we suggested, along the same lines, that the distinctive and fundamental property of stød is not creak, but a somewhat compressed voice quality which in its stronger manifestation becomes creaky towards the end of the syllable and/or may produce a late and abrupt F_0 lowering. The results presented here support the characterization of stød as fundamentally compressed voice, assuming that in its weakest form the compression merely increases F_0, whereas a stronger manifestation turns the modal vibratory pattern into creak and/or lowers F_0, cf. [15; 8, 20]. – We ought to have added Eli Fischer-Jørgensen's observations to our list of arguments against a phonological representation of stød as essentially a L tone squeezed back into a H stressed syllable to form the second part of a bitonal HL pitch accent, as suggested in [17, 18, 21, 22-24].

5. ACKNOWLEDGEMENT

Thanks to Hans Basbøll for very helpful comments to the first draft.

5. REFERENCES

[1] Basbøll, H. 2005. *The phonology of Danish* Oxford: Oxford University Press.

[2] Blankenship, B. 2002. The timing of nonmodal phonation in vowels. *J. Phonet.* 30, 163-191.

[3] Boersma, P., Weenink, D. 2006. Praat: doing phonetics by computer [Computer program praat5404_win64] http://www.praat.org/ (accessed 29 December 2014).

[4] Fischer-Jørgensen, E. 1987. A phonetic study of the stød in standard Danish. *Ann. Rep. Inst. Phonet., Univ. Copenh.* 21, 55-265.

[5] Fischer-Jørgensen, E. 1989. Phonetic analysis of the stød in standard Danish. *Phonetica* 46, 1-59.

[6] Garellek, M., Keating, P. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *J. Int. Phonet. Ass.* 41/2, 185-205.

[7] Gerfen, C., Baker, K. 2005. The production and perception of laryngealized vowels in Coatzospan Mixtec. *J. Phonet.* 33, 311-334.

[8] Gordon, M., Ladefoged, P. 2001. Phonation types: a cross-linguistic review. *J. Phonet.* 29, 383-406.

[9] Grønnum, N. 1990. Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German. *Phonetica* 47, 182-214.

[10] Grønnum, N. 1998. Danish. *J. Int. Phonet. Ass.* 28, 99-105.

[11] Grønnum, N. 2009. A Danish phonetically annotated spontaneous speech corpus (DanPASS). *Speech Comm.* 51, 594-603. www.danpass.hum.ku.dk/danpass_webpage/danpass.htm

[12] Grønnum, N., Basbøll, H. 2001. Consonant length, stød and morae in standard Danish. *Phonetica* 58, 230–253.

[13] Grønnum, N., Basbøll, H. 2002. Stød and length: acoustic and cognitive reality? *Proc. Speech Prosody* Aix-en-Provence, 355–358.

[14] Grønnum, N., Basbøll, H. 2003. Stød and length in standard Danish: experiments in laboratory phonology. *Proc. 15th ICPhS* Barcelona, 455–458.

[15] Grønnum, N., Vazquez-Larruscaín, M., Basbøll, H. 2013. Danish Stød: Laryngealization or Tone. *Phonetica* 70, 66–92.

[16] Hombert, J.-M. 1978. Consonant Types, Vowel Quality and Tone. In: Fromkin, V.A. (ed), *Tone, A Linguistic Survey* New York: Academic Press, 77-112.

[17] Itô, J., Mester, A. 1997 Stø^{3}det i dansk. Handout. *Scandinavian Summer School in Generative Phonology*, Hvalfjarðarströnd.

[18] Itô, J., Mester, A. (Un)accentedness and the perfect prosodic word. manuscript. *Ling. Inquiry* (submitted).

[19] Jeel, V. 1975. An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants. *Ann. Rep. Inst. Phonetics, Univ. Copenhagen* 9, 191-211.

[20] Ladefoged, P., Maddieson, I. 1996. *The sounds of the world's languages* Oxford: Blackwell.

[21] Morén, B. 2005. Danish stød and Eastern Norwegian pitch accent: the myth of lexical tones. *13th Manchester Phonology Meeting* (MFM13), Manchester. Power Point Presentation available at http://www.hum.uit.no/a/moren/cv_files/morenMFM13.pdf (accessed June 2013).

[22] Riad, T. 2000. The Origin of Danish Stød. In: Lahiri, A. (ed), *Analogy, levelling, markedness: principles of change in phonology and morphology*. The Hague: Mouton de Gruyter, 261-300.

[23] Riad, T. 2005. Historien om tonaccenten. *Studier i svensk språkhistorie* 8, 1-27.

[24] Riad, T. 2009. Eskilstuna as the tonal key to Danish. *Proc. Fonetik 2009*, 22nd Swedish Phonet. Conf., Stockholm, 12–17. Available at http://www2.ling.su.se/fon/fonetik_2009/proceedings_fonetik2009.pdf (accessed June 2013).

[25] Riber Petersen, P. 1973. An instrumental investigation of the Danish stød. *Ann. Rep. Inst. Phonetics, Univ. Copenhagen* 7, 195-234.

[26] Smith, S. 1944. *Bidrag til Løsning af Problemer vedrørende Stødet i Dansk Rigssprog* København: Kaifer.

[27] Thorsen, N. (same as Grønnum, N.) 1979. Interpreting raw fundamental frequency tracings of Danish. *Phonetica* 36, 57-78.