Stop! Hey, what's that sound? the representation and realization of Danish stops

Puggaard-Rode, R.

CHAPTER 1

Introduction

1.1 Old and new insights

In the chorus of their 1966 hit single "For What It's Worth", the band Buffalo Springfield famously asked listeners a question which is repeated in the title of this dissertation. Although the band was (probably) not referring to a class of speech sounds, the same question (with a slightly shifted meaning) has been asked by linguists for centuries. Oral stops[1] appear deceptively simple, but closer inspection reveals that they allow for complex patterns of variability in articulatory implementation, acoustic signatures, and phonological behavior. These are patterns which phonetic and phonological theory need to account for.

The articulatory goal of a stop is a complete closure at some place of articulation, which is often acoustically cued only by silence. The place of the occlusion is acoustically cued only by a very brief release burst; or by a very rapid transition in the formants generated in the supraglottal cavities as the tongue and lips move from a complete

[1] Throughout this dissertation, I use term 'stops' to refer to oral stops, and 'nasals' to refer to nasal stops.

occlusion to a more open approximation; or by a combination of burst and formant transitions. Different stop types can be contrasted by a variety of glottal states during the closure, and these are mostly acoustically cued by minor changes in voice quality in adjacent sonorant sounds. Ladefoged (1975: 174) writes that "(i)n many cases, a consonant can be said to be a particular of way of beginning or ending a vowel"; this is especially true of stops.

Much is already known about the phonetics of Danish stops. We have a good understanding of the timing of closure (Fischer-Jørgensen 1954, 1972b; Andersen 1981a) and release (Fischer-Jørgensen 1954, 1979; Mortensen and Tøndering 2013), the spectral characteristics of releases (Fischer-Jørgensen 1954), perturbations in fundamental frequency F_0 after stops (Fischer-Jørgensen 1969; Jeel 1975; Petersen 1983), muscular and laryngeal activity in stop production (Frøkjær-Jensen et al. 1971; Fischer-Jørgensen and Hirose 1974; Hutters 1985), and phonetic reduction of stops (Pharao 2009, 2011).

Danish displays an intriguing and complex set of alternations between voiceless unaspirated stops [p t k] and semivowels [o̯ ɤ̯ ɪ̯], and according to a long-standing tradition of Danish phonological analysis, they are assumed to derive from the same phonemes /b d g/ (Uldall 1936; Hjelmslev 1951; Jakobson et al. 1951; Rischel 1970a; Basbøll 1975, 2005; Grønnum 2005). Danish has often been brought up in discussions of the phonetic underpinnings of phonological features. Different facts about Danish phonetics have been taken as evidence of (often conflicting) positions regarding the underlying representation of laryngeal contrast (Keating 1984a; Goldstein and Browman 1986; Kingston and Diehl 1994; Iverson and Salmons 1995; Beckman et al. 2013).

Both the recent and more distant history of the stops are well-described (Brøndum-Nielsen 1928–1973; Hansen 1962–1971; Brink and Lund 1975); the broad strokes of regional variation in stops are covered (Bennike and Kristensen 1898–1912); and recent studies have uncovered socially stratified variation in stop realization (Pharao and Maegaard 2017; Lillelund-Holst et al. 2019).

There are probably very few languages of the world whose stops are as well-described as Danish. In part due to this state of affairs, there are also many open questions about their phonetic and phonological

behavior, and how they vary across different speakers and varieties. In this dissertation, I take up some of these questions.

1.2 Structure of the dissertation

The dissertation is divided into two parts. In Part I, I focus on the 'big picture'. In Chapter 2, I review the existing literature in detail, and discuss abstract properties of the stops and their underlying representation on the basis of existing literature. The phonological account of stop–semivowel alternations mentioned above has remained largely unchanged in nearly a century, in spite of sound changes in that period of time (and prior) which have drastically altered the relevant allophones. In Chapter 3, I argue that these changes have made the traditional phonological account untenable; I suggest that from a synchronic point of view, the alternations must be considered suppletive, and that they are best understood in light of the diachronic pressures which produced them.

In Part II, I report a series of quantitative case studies. Most studies on the phonetics of Danish stops are either introspective, based on auditory impression, or based on carefully read speech in a laboratory setting. In Part II of this dissertation, I rely on corpora of spontaneous speech, aided by recent advances in statistical modeling that allow the user to take into account many factors simultaneously. The interface between phonetics and phonology is foregrounded throughout the dissertation, but the chapters in Part II focus increasingly less on phonology and more on phonetics (and method development for corpus phonetics).

Chapter 4 presents a study on intervocalic closure voicing. This study uses the DanPASS corpus (Grønnum 2009). Intervocalic voicing has not previously been investigated in Danish; unlike the impression one gets from (most of) the literature, intervocalic voicing is found to be relatively rare in all Danish stops. The study both draws on and illuminates theoretical consequences of existing studies of e.g. glottal activity and F_0-perturbations in Danish stops. This, in turn, makes it possible to evaluate a number of proposals in the literature regarding the repre-

sentation of laryngeal contrast, since several of these proposals make explicit predictions about intervocalic voicing.

The same corpus is used in Chapter 5 to test an old prediction of Otto Jespersen's (1897–1899), namely that Danish was (or is still) undergoing a sound change similar to the Second Consonant Shift in German, whereby the aspirated stops [p^h t^h k^h] are becoming affricates. Particularly [t^h] is strongly affricated in Modern Standard Danish, but it has not previously been examined whether this phenomenon is related to phonetic environment, nor whether it is actually isolated to [t^h]. The spectral characteristics of stop releases are analyzed using a novel method for statistically modeling dynamic changes in speech spectra, viz. function-on-scalar regression (Reiss et al. 2010; Greven and Scheipl 2017a; Bauer et al. 2018). The chapter serves as a case study for this method, which I argue may be a solution to the problem of how to analyze the highly complex and multidimensional information in the spectrum.

Denmark has a rich tradition in dialectology; there are both thorough country-wide overviews of regional variation (Bennike and Kristensen 1898–1912; Brøndum-Nielsen 1927; Skautrup 1944–1970; Ringgaard 1971) and a large number of monographs or smaller studies describing individual dialects. These descriptions were mostly written within the structuralist framework of glossematics pioneered by Louis Hjelmslev and the Linguistics Circle of Copenhagen, which was one of the centers of structuralism in Europe in the 20th century. Phonology was a core topic in these descriptions, but they rarely included descriptions of phonetics, since glossematics was explicitly uninterested in phonetic substance (Hjelmslev 1943: 46). In Chapter 6, I describe a very large legacy corpus of tape recordings from the 1970s (DS 1971–1976), and use it to investigate variation in voice onset time and stop affrication in the traditional regional varieties of Danish spoken on the Jutland (Jylland) peninsula. This study sheds new light on long-standing gaps in our knowledge about phonetic variation in Danish, and also exposes some previously ignored correspondences between variation in synchronic phonetic implementation and variation in the outcome of attested sound changes.

Finally, in Chapter 7, I summarize and synthesize the proposals and results presented throughout the dissertation.

All chapters of the dissertation share a common topic, but several of them have also been published or submitted as individual publications. This applies to Chapters 3–6. Some of these publications are coauthored; some chapters combine multiple publications; some chapters combine already published research with novel research. Each chapter has a note on the first page describing where the research has been published and presented before, who else contributed to it, and where to find the underlying code and data. The chapters have all been reformatted and to some extent rewritten to fit the dissertation; this should be a coherent book in itself, so the chapters are updated with cross-references, etc.

1.3 The Danish language

Danish is a North Germanic language spoken by approximately 5.6 million people, mostly in the nation of Denmark (Eberhard et al. 2020). The synchronic North Germanic languages (Danish, Norwegian, Swedish, Icelandic, and Faroese) share a common parent language that first started diverging in the transitional period between Ancient Nordic and Old Nordic, around the year 1,000 AD. The divergence was mainly caused by innovations spreading from the south, i.e. the border area between Denmark and Germany (Birkmann 2002). The differences between Norwegian and the insular varieties (Icelandic and Faroese) were initially very minor, but as innovations spread from the south and contact across the Atlantic Ocean diminished, the peninsular varieties (Danish, Norwegian, and Swedish) grew increasingly similar (Gooskens 2020). The modern insular languages retain the complex inflectional morphology of Old Nordic, which has mostly been lost in the peninsular languages. Nowadays, the insular and peninsular languages are not mutually intelligible. Danish has been subject to a number of prominent sound changes, particularly consonant gradation and vowel reduction in unstressed syllables, which has made it an outlier in the peninsular group (Torp 1998). Consonant gradation is the commonly used cover term for a number of related lenition processes; these are discussed in much more detail in Section 2.4.4 and Chapter 3. The peninsular languages are still "mutually intelligible with some

effort" (Gooskens 2020: 763; see also Gooskens 2006; Gooskens et al. 2010).

The Danish language was characterized by extensive regional variation until fairly recently. Danish dialects are traditionally divided into three primary subgroups: *Eastern Danish*, as spoken on the island of Bornholm and in Scania (Skåne) in the southern part of present-day Sweden; *Insular Danish*, as spoken on the islands of Zealand (Sjælland), Funen (Fyn), and a number of smaller islands in the Baltic Sea; and *Jutlandic*, as spoken on the Jutland peninsula and a number of smaller surrounding islands (Skautrup 1944–1970; Kristiansen 2003a). The capital and largest city, Copenhagen (København), is located on the island of Zealand. Insular Danish, with Copenhagen as the primary center of innovation, first started emerging as a standard variety in the 16th century. This likely did little to reduce regional variation in the spoken language before the 19th century, when two societal developments led to the accelerated spread of Insular Danish (particularly High Copenhagen Danish) as a spoken standard variety: social and geographical mobility increased due to agrarian reforms, and an obligatory educational system was introduced. Around the 1960s, two further developments rapidly accelerated the leveling of traditional dialects: Standard Danish was adopted as a central tenet of government policy on mother-tongue education (Kristiansen 1990, 2003b), and private television sets (broadcasting in the standard language) became a common household appliance (Kristiansen 2003a). As a result, the traditional varieties quickly became moribund (Kristiansen 1998), and today, Denmark is a radically standardized speech community (Pedersen 2003). Features from the traditional dialects are still heard in some areas, but they have a very different social function compared to a century ago (Scheuer et al. 2019).

Modern Standard Danish is doing well in almost every measure of language vitality (e.g. Kirchmeier-Andersen 2007). It is an extremely well-described language, with a tradition for language description going back to at least the 17th century, and a tradition for language philosophy going back significantly further (Hovdhaugen et al. 2000). Danish phonetics are particularly well-described: Rischel (2000: 161) suggests that "among Western languages, Danish is second only to

English in its post-medieval tradition of painstaking phonetic observation".

1.4 Transcription conventions

Not all readers of this dissertation will be familiar with Danish. For this reason, I provide transcriptions of all words and morphemes mentioned throughout. I mostly use idealized phonetic transcriptions in square brackets [ ], which represent my own pronunciation in reasonably clear speech, using a pre-defined set of symbols to represent sounds that are positionally contrastive in my accent.[2] Minor, predictable variations in pronunciation due to e.g. phonetic context are not transcribed. I generally do not give phonemic transcriptions of words and phrases, although I frequently refer to phonemes in slant brackets / /; I clarify throughout how these should be understood. Especially /b d g p t k/ are often used to denote the stops when I focus on phonological contrast rather than phonetic realization; keep in mind /b d g/ are not systematically voiced, and are typically transcribed phonetically as [p t k]. When I do provide phonemic transcriptions of full words or phrases, they are always copied from other sources.

The consonant symbols I use in transcriptions are shown in Table 1.1. The table should mostly be self-explanatory, except perhaps for the four-way aperture distinction in central oral consonants. The approximant–semivowel distinction in Danish is somewhat unusual; semivowels have a more open approximation, and approximants and semivowels are generally found in different prosodic contexts. I return to this distinction in Sections 2.4.1 and 2.4.4, and in Chapter 3.

The Danish rhotic is usually transcribed as [ʁ]. The sound is variously described as uvular (Basbøll 2005), suprapharyngeal (Grønnum 1998), or pharyngeal (Grønnum 2005), with no consensus in

[2] My accent is probably best described as young regionalized Modern Standard Danish with features from south–west Jutlandic and Aarhus Danish. The main differences from Copenhagen Danish are in intonation, which is never transcribed here. Vowel quality and length also occasionally differs, as do the exact patterns of stød. Sometimes transcriptions divert from my accent when discussing prominent patterns of variation or when transcribing speech from corpus recordings; this will be indicated clearly in the text.

Table 1.1: *Overview of consonant symbols used to transcribe Modern Standard Danish throughout the dissertation.*

	Bilabial	Labiodental	Alveolar	Alveopalatal	Palatal	Velar	Pharyngeal	Glottal
Stop, asp.	[pʰ]		[tʰ]			[kʰ]		
Stop, unasp.	[p]		[t]			[k]		
Affricate				[tɕ]				
Fricative		[f]	[s]	[ɕ]				[h]
Nasal	[m]		[n]			[ŋ]		
Approximant		[ʋ]			[j]		[ʁ]	
Lateral appr.			[l]					
Semivowel	[o̯]			[ɤ̯]	[i̯]		[ʁ̞]	

the literature about the exact place of articulation (Sobkowiak 2018). While [ʁ] is an imprecise transcription, it probably comes closest to the actual articulation without using diacritics. The semivowels fit awkwardly into the rigid system of consonant place distinctions in the International Phonetic Alphabet (IPA); some have prominent secondary articulations, as discussed in Chapter 3. All nasals, laterals, and semivowels may appear as syllabic, ostensibly due to assimilation with an adjacent underlying schwa (Brink and Lund 1975; Basbøll 2005; Schachtenhaufen 2010b).[3]

The vowel symbols I use in transcriptions are shown in Figure 1.1. These are plotted from measurements of the first two formants ($F1$–$F2$) from recordings of my own speech. Figure 1.1 includes syllabic variants of the semivowels, with the exception of the pharyngeal semivowel [ɐ], since the $F1$–$F2$ frequencies of this sound are indistinct from [ʌ] in my speech.[4] The vowel placements in Figure 1.1 are similar to those found

[3] In Schachtenhaufen's (2022) conventions for transcribing Modern Standard Danish, he makes no distinction between syllabic and non-syllabic sonorants, since it is unclear whether the phonological notion of 'syllabicity' has any stable acoustic cues in the language (Schachtenhaufen 2010a).

[4] [ʌ ~ ɐ] are structurally quite different, so I do distinguish between them in transcriptions. [ʌ] is a full lexical vowel while [ɐ] always derives from [ɐ̯] which has syllabified due to schwa assimilation.

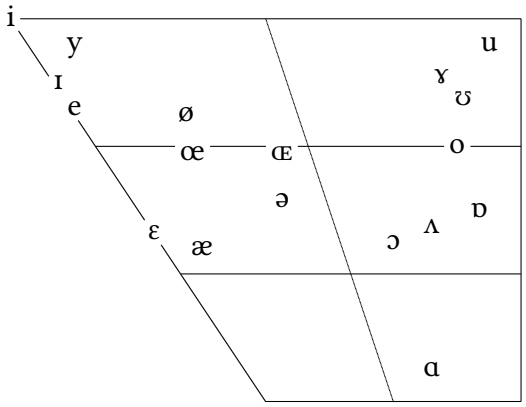

Figure 1.1: *Overview of vowel symbols used to transcribe Modern Standard Danish throughout the dissertation.*

in a corpus study by Juul et al. (2016), although my [ɤ ʌ] are relatively high and my [ɔ] is relatively low compared to their results. The upper quadrants are extremely crowded in Modern Standard Danish,[5] and some vowel symbols are consequently quite far removed from their conventional positions in the IPA quadrilateral.

I use the following symbols for distinctive prosodic characteristics: primary stress [ˈ], vowel length [ː], and stød [ˀ]. There are of course many 'phonetic levels' (e.g. Jespersen 1906; Hansen and Lund 1983), but it is not clear that more than two levels of stress, i.e. 'stressed' and 'unstressed', have a distinctive function (see e.g. Rischel 1970b, 1972, 1983). Stød is a constrastive prosodic feature which is usually realized as laryngealization during the latter part of long vowels or during post-vocalic sonorants (Grønnum and Basbøll 2003; Grønnum et al. 2013); the morphophonological patterns of stød are inordinately complex and riddled with exceptions and idiosyncrasies (Basbøll 1985, 2003, 2005; Goldshtein forthc.). The superscripted glottal stop symbol [ˀ] is routinely used to transcribe stød, but Schachtenhaufen (2022) has recently suggested using the diacritic for creaky voice (e.g. [o̰]) instead,

[5]Some of these contrasts are likely cued by other acoustic features than just *F*1–*F*2 frequencies.

Table 1.2: *Differences between transcription conventions in this dissertation and those of Schachtenhaufen (2022) and Grønnum (1998).*

This dissertation	**Schachtenhaufen**	**Grønnum**
[p t k]	[p t k]	[b̥ d̥ g̊]
[tʰ]	[ts]	[tˢ]
[ʊ]	[v]	[v]
[ɣ̞]	[ɣ]	[ð]
[ʊ̯ ɪ̯]	[ʊ ɪ]	[w j]
[ɐ̯]	[ɐ]	[ʌ̯]
[æ]	[æ]	[æ ~ a]
[ɑ]	[a]	[ɑ]
[œ]	[œ]	[œ ~ œ̝]
[ɒ ~ ʌ]	[ɒ]	[ɒ ~ ʌ]

since [ˀ] is not an official IPA symbol. I stick with the more traditional transcription here, as stød has a variety of possible phonetic cues in addition to creak or laryngealization (Hansen 2015), and I think it is better understood as a property of higher-order prosodic constituents rather than the sonorant sound it happens to primarily attach to.

There have been many competing standards of phonetic transcription in Danish. Jespersen (1890) developed a phonetic alphabet, *Dania*, which remained very popular in the discipline throughout the 20th century. Dania is still often used today, although it has increasingly been replaced by the IPA. I will not be using Dania in this dissertation, but I will occasionally be 'translating' Dania transcriptions from earlier research into IPA. In order to avoid confusion for readers familiar with a different tradition of Danish IPA transcription, it is worth briefly commenting on some of these. Table 1.2 compares the transcription conventions in this dissertation with the ones recently proposed by Schachtenhaufen (2022) and with Grønnum's (1998) *Illustrations of the IPA* entry for Danish.

I follow Schachtenhaufen (2022) in collapsing the distinctions between [æ ~ a] and [œ ~ œ̝], since neither are contrastive in

my speech.[6] Schachtenhaufen (2022) further proposes collapsing the [ɒ ~ ʌ] distinction, as this is the only pair of full vowels in Modern Standard Danish that does not show a contrastive length distinction; [ɒ] is always long, and [ʌ] is always short. I retain this distinction; impressionistically, I find that there is a significant difference in quality between the two.[7] Finally, I do not follow Schachtenhaufen in transcribing [ɑ] as a front vowel, since my realization of this vowel is generally quite back (i.e. *F*2 is quite low), and it usually patterns phonologically as a back vowel. I return to this point in Chapter 3.

Differences in transcription of consonants is covered in more detail when I discuss the inventory from a phonological perspective in Section 2.4.1, and conventions will be discussed throughout the dissertation. The phonological discussion in Chapter 3 supports representing [ʊ] and [ɤ̞] as sonorants rather than obstruents; the corpus study of intervocalic voicing in Chapter 4 supports representing [p t k] as baseline voiceless; and the corpus study of spectral properties in stop releases in Chapter 5 arguably does not support treating [tʰ] as an affricate.

[6]The merger between [æ ~ a] is well-known and has been discussed by e.g. Brink and Lund (1975), Grønnum (1995), and Juul et al. (2016).

[7]Schachtenhaufen (2022) argues that the difference in quality between [ɒ ~ ʌ] is simply due to length, and that all short–long pairs differ in quality. While this is true, the difference between [ɒ ~ ʌ] in *F*1–*F*2 frequency is greater than other short–long pairs (Juul et al. 2016).