Stop! Hey, what's that sound? the representation and realization of Danish stops
Puggaard-Rode, R.

Stop! Hey, what's that sound?

The representation and realization of Danish stops

Published by

LOT
Binnengasthuisstraat 9
1012 ZA Amsterdam
The Netherlands

phone: +31 20 525 2461
e-mail: lot@uva.nl
http://www.lotschool.nl

Cover illustration: Piedmont Photo
ISBN: 978-94-6093-417-9

DOI: https://dx.medra.org/10.48273/LOT0631
NUR: 616

Stop! Hey, what's that sound?

The representation and realization of Danish stops

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 11 januari
klokke 16.15 uur

door

Rasmus Puggaard-Rode

geboren te Varde, Denemarken
op 8 juli 1992

Promotor: Prof.dr. Janet Grijzenhout

Copromotor: Dr. Bert Botma

Promotiecommisie: Prof.dr. Paul Boersma
(University of Amsterdam)

Prof.dr. Yiya Chen

Dr. Willemijn Heeren

Prof.dr. Martin Krämer
(Arctic University of Norway – Tromsø)

Contents

Acknowledgments

This dissertation is the product of the labor and love of many people who helped me along the way. First off, I would like to thank my wonderful supervisors, Janet Grijzenhout and Bert Botma, for their support and guidance throughout the process. I would also like to thank my collaborators on research that went into the dissertation: Camilla Søballe Horslund, Henrik Jørgensen, and Yonatan Goldshtein. Thank you for always being available for questions, discussions, and useful comments, both in connection with papers that we have worked on and in general. I genuinely hope that our collaborations have only just begun. Yonatan has been a fantastic friend since I first discovered linguistics ten years ago, and has been more influential on my way of thinking than I could ever put into words.

Many people read early versions of one or more of the chapters, and gave helpful comments: Camilla, Yonatan, Stig Eliasson, Míša Hejná, Sam Kirkham, Jan Heegård Petersen, Nicolai Pharao, and Ruben Schachtenhaufen. Dirk Jan Vet contributed an enormously useful Praat script, and Paul Boersma, Laura Smorenburg, and Cesko Voeten all gave helpful tips about aspects of the statistical modeling and acoustic processing; Andreea Geambasu gave helpful advice on the presentation of the online appendices. Thanks to Søren Sandager Sørensen for always being available to discuss *any* aspect of the Danish language.

The contents of this dissertation have been greatly improved by the peer review system. I would like to thank all peer reviewers who read associated papers, including the reviewers for a paper submitted to *Journal of the International Phonetic Association* which never materialized. I would also like to thank the editors for these publications and submissions: Hartmut Haberland, Nanna Haug Hilton, James Kirby, Remco Knooihuizen, Jeff Mielke, Hans Van de Velde, and Elizabeth Zsiga. Thanks also to everyone who showed up to presentations and posters and gave feedback.

Thanks to the wonderful people I have surrounded myself with while working on this dissertation: Astrid van Alem, András Barany, Meike de Boer, Hang Cheng, Lis Kerr, Menghui Shi, Matthew Sung, Xander Vertegaal, Andrew Wigman, Sarah von Grebmer zu Wolfsthurn, Jiang Wu, Cesko, Laura, and the rest of my PhD cohort. Thanks to the Tromsø/Verona phonologists: Birgit Alber, Eirini Apostopoulou, Andrea Brugnoli, Joachim Kokkelmans, Martin Krämer, Anna Katharina Pilsbacher, Tor Håvard Solhaug, Olga Urek, Kate Zhakun. Thanks to the Aarhus dialectologists: Torben Arboe Andersen, Nina Grøftehauge, Inger Schoonderbeek Hansen, Kirsten Lyshøj, Ove Rasmussen, Viggo Sørensen. Thanks to my colleagues at UvA, Silke Hamann and Marijn van 't Veer. My understanding of language and linguistics was shaped at the fantastic Linguistics department in Aarhus. Thanks to teachers (Peter Bakker, Anna Jespersen, Bill McGregor, Jan Rijkhoff, Jakob Steensig, Ethan Weed), co-students (in particular to Yonatan, Søren, Katrine Garly, Kasper Fyhn, Mads Lundegaard, Tina Raakjær Mølgaard, Jonathan Puntervold, Nikolas Tranæs), my fellow DanTiN members, and everyone who hung out in Lingoland between 2012–2018. Thanks to Sidsel Rasmussen for pointing me in the direction of this PhD position. Finally but very importantly, thanks to my students from the past eight years. You have immeasurably improved my understanding of the subject matter.

For obvious reasons, writing a PhD in 2020–2022 has not exactly been what I expected when I started in 2018. My office has been my dining room table, the fun aspects of conferences and summer schools were essentially scrapped, and the lion's share of the dissertation has been written on a tablet. So a brief note of thanks to the services that made this bearable: the reading groups of the Tromsø/Verona phonology circle and the Aarhus-based Sounds of Language and Speech group who moved their activities online, the Linguistics Twittersphere, Gonge and Meilandt who brought Pink Cheese Quiz into our living rooms. A huge, huge thanks to my roomie, office mate, and perhaps my most essential coauthor during this whole debacle – our cat Kipawsky, who moved in during summer 2020. I eventually decided to edit out all his contributions. I am sure he meant well, but I found them completely incomprehensible, as if he simply wandered aimlessly over the keyboard. Here is a sample, which strikes me as pure gibberish: ‘88888888888888888,hlb.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l.,l’.

En stor tak til min mor og far, til Mikkel, Signe, Lærke og Oscar, og resten af min familie og venner i Danmark for at være en enorm støtte gennem hele projektet og for at være gode til at holde kontakten selv når jeg ikke har været det.

Finally, thanks to my wife, Rikke, who came with me on this journey. I feel weirdly safe moving around the world with you pursuing an elusive academic career, because I can be certain you will end up doing important work no matter where we might find ourselves. Thanks for being the coolest person I know, and thanks for agreeing to spend your life with me. For what it’s worth, this book is for you.

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 [ʊ̯ ɣ̯ 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 representation 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	[ʊ̯]			[ʏ̯]	[ɪ̯]		[ɐ̯]	

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.

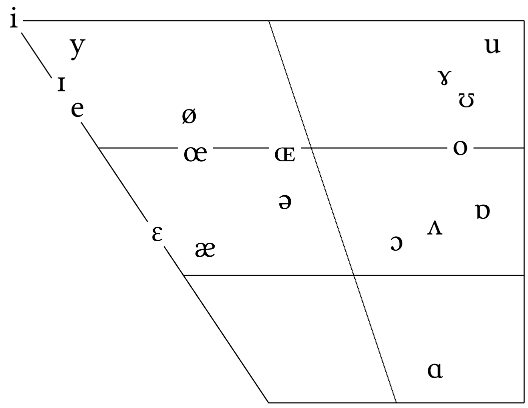


Figure 1.1: *Overview of vowel symbols used to transcribe Modern Standard Danish throughout the dissertation.*

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. [ɑ̰]) instead,

[5]Some of these contrasts are likely cued by other acoustic features than just $F1$–$F2$ 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]
[ʋ]	[v]	[v]
[ɣ̯]	[ɣ]	[ð]
[ʊ̯ ɪ̯]	[ʊ ɪ]	[w j]
[ɐ̯]	[ɐ]	[ʌ̯]
[æ]	[æ]	[æ ~ a]
[ɑ]	[a]	[ɑ]
[œ]	[œ]	[œ ~ œ̞]
[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 ~ œ̞], 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 [ʋ] and [ɣ̞] 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).

Part I

The big picture

CHAPTER 2

Previous research on Danish stops

2.1 Preliminaries

Standard Danish has six contrastive stops /b d g p t k/. In pre-vocalic onset position, the distinction between the two laryngeal series is uncontroversially based on aspiration.[1] /b d g/ are realized as voiceless unaspirated, /p t k/ are realized as voiceless aspirated. /b d g/ are sometimes claimed to have voiced medial allophones. /t/ has a salient affricated release, and is usually transcribed phonetically with superscript *s* [tˢ]. The most common phonological analysis of Danish stops suggests that the unaspirated stops show complex alternations with semivowel and zero allophones. All of these patterns are subject to variation in traditional dialects, although existing sources focus mostly on variation in categorical phonological patterns.

In this chapter, I aim to provide a comprehensive literature review of the history, phonetics, phonology, and variation of Danish stops. As discussed in the previous chapter, these are a very well-described set of sounds, although a number of open questions still remain. Relevant parts of the history of the language are covered in Section 2.2, since

[1]This is at least the case at the phonetic level (see Section 2.4.3).

diachronic developments will be important in interpreting the standard analysis of synchronic stop gradation, and will play a key role in the reanalysis proposed in Chapter 3; see Section 1.3 for a more general sketch of the history of Danish. In Section 2.3, I cover the results of previous studies of acoustic and articulatory phonetics. In Section 2.4, I turn to phonology, covering positional allophones and distributional properties of the stops, and the mechanisms that have been taken to be relevant in their underlying representation, before moving to the complex gradation processes whereby the voiceless unaspirated stops are usually taken to have semivowel or zero allophones in weak prosodic position. Finally, I give an overview of what we know about social and regional variation in the phonology and phonetics of stops, as well as phonetic reduction, on the basis of the existing literature in Section 2.5. Many of these topics will also be relevant in later chapters, but some are simply covered in this chapter in order to give as comprehensive an overview as possible of previous research on Danish stops.

2.2 History

In this section, I will briefly cover the history of the Danish stops and their main (proposed) allophones, namely semivowels. There is a long established and broad research tradition in Germanic historical linguistics, so this overview will not be comprehensive; I mostly stick to presenting facts which will become important in later chapters. Here, I will cover the historical sources of the current Modern Standard Danish stops and their allophones (Section 2.2.1), and recently attested changes in these (2.2.2).

2.2.1 Historical sources of stops and semivowels

This section covers the gross historical changes that led to the current inventory of stops and semivowels. There are of course counterexamples, and several well-described patterns of sounds behaving differently in e.g. consonant clusters, which I will not cover here. For more comprehensive overviews of these sound changes, see e.g. Brøndum-Nielsen (1928–1973), Skautrup (1944–1970), Hansen

(1962–1971), Brink and Lund (1975), Bandle et al. (2002–2005), and Hjorth (2017–2022).

Ancient Nordic largely retained the consonant inventory from Proto-Germanic (Braunmüller 2002), including the set of stop phonemes that is still retained in the modern Scandinavian languages: /b d g p t k/. These were also the stops found in Old Danish (Brøndum-Nielsen 1928–1973).

Old Danish /p t k/ developed from the Proto-Indo-European (PIE) voiced stops *b d g, which lost their voicing as a consequence of Grimm's law (Ringe 2006). These are now aspirated in onset position, but there is disagreement in the literature regarding when they developed aspiration, as discussed further in Sections 2.2.2 and 3.5.3; in German, they variably developed further into affricates as part of the Second Consonant Shift. An example is PIE *dn̥ǵʰwéh_2- > Proto-Germanic *tungōn- > Norse *tunga* > Modern Standard Danish [ˈtʰɔŋ] *tunge* 'tongue', cf. German [ˈtsʊŋə] *Zunge* (Ringe 2006: 81; DSL 2018).

Old Danish /b d g/ developed from different sources, but mostly from PIE breathy voiced stops *bʱ dʱ gʱ, which in some contexts were retained as stops in Proto-Germanic.[2] An example is PIE *gʱóstis 'stranger' > Proto-Germanic *gastiz 'guest' > Norse *gestr* > Modern Standard Danish [kɛst] *gæst* (Ringe 2006: 97; DSL 2018). Many of the current words with initial /b d g/ were later borrowings. They are no longer phonetically voiced in onset position, and it is unclear when voicing was lost (as discussed further in Section 3.5.3). The laryngeal contrast in especially /b p/ has historically been rather unstable.

The PIE voiceless stops *p t k developed into Proto-Germanic voiceless fricatives *f θ x with Grimm's law, and as a result of Verner's law, these were voiced in medial and final position *v ð ɣ; both developments can be seen in in the development of PIE *ph_2tḗr 'father', which further developed into *faþḗr, and is reflected in Norse as *faðir* and Danish as [fɑː] *far*, or [ˈfæːɣ̞ɐ] *fader*, which is still used in e.g. biblical contexts (Ringe 2006: 102).[3] The dental fricative [θ], which was retained in Old Danish, hardened to a voiceless stop in Early Modern Danish

[2] According to Hansen (1962–1971: II), they were variably assibilated during this stage.

[3] [ɣ̞] is the modern reflex of [ð], as discussed in Section 2.2.2 below.

and in the other East Scandinavian languages (Boeck 2018), significantly increasing the number of /t/-initial words; see for example Norse *þing* > Modern Standard Danish [tʰeŋˀ] *ting* 'thing'. A similar development took place in Frisian (Laker 2014).

Germanic developed a quantity contrast between singleton and geminate stops. The sources of gemination are somewhat controversial, and there are many open questions (Goblirsch 2018). The mechanisms differed between branches of Germanic, but a common suggested mechanism was the development of geminates from stop + nasal clusters; this is referred to as Kluge's law (Kluge 1884; Lühr 1988). In the East Nordic languages, geminates further developed from some stop + glide clusters, where the glide was originally retained but is lost in Modern Standard Danish; see Proto-Germanic *ligjaną > Norse *liggja* > Modern Standard Danish [ˈlekə] *ligge* 'lie' (Ringe 2006: 253; DSL 2018). Other geminates derived from /h/ + stop clusters (Nielsen and Stoklund 2018). Old Danish had a distinction between voiceless geminates [pː tː kː] and voiced geminates [bː dː gː], but this contrast was rather unstable, and neutralized in favor of the voiceless set sometime early in the Middle Danish period (Sørensen 2012). Occasionally, this happened in concert with compensatory lengthening of the preceding vowel, as in Norse *skegg* 'beard' > Modern Standard Danish [skɛːˀk] *skæg*, cp. Norse *bekkr* 'stream' > Modern Standard Danish [pɛk] *bæk* (DDO 2018).

In the beginning of the Middle Danish period (ca. 1100–1500), final devoicing leads to positional neutralization of the laryngeal contrast in stops (Frederiksen 2018). First attested around the year 1400 and continuing throughout Early and Late Modern Danish (Boeck 2018; Brink and Lund 2018), a number of important lenition processes take place, including what is known as *spirantsvækkelsen* 'spirant weakening' and *klusilsvækkelsen* 'plosive weakening'. Spirant weakening describes a series of related changes, whereby voiced fricatives (found only post-vocalically) changed into approximants or were lost. The processes are shown in (1), following (Frederiksen 2018).

(1)	[v]	→	[w]		
	[ð]	→	Ø or [j]		
	[ɣ]	→	Ø	/	_ [+high]
			[j]	/	_ [-back, -round]
			[w]	/	_ [+back]
			[j] or [w]	/	_ [+round]

The different outcomes for [ð] in (1) were largely regionally determined, and not consistent; [ð] was often retained. The precise place of articulation of [ɣ] likely already varied greatly by place of articulation of the preceding vowel; the tongue position in velars is generally highly dependent on surrounding vowels (e.g. Vilain et al. 1998), and the distribution in (1) is similar to the well-known distribution of *ich-Laut* and *ach-Laut* in German; cf. Hall's (1989) rule of German Fricative Assimilation. The consequences of this are discussed in further detail in Section 3.5.2.5 below.

A corresponding plosive weakening chain was consistent across the Danish-speaking area, although realized in various ways in different regional varieties. Here, I cover the outcome in the area around Copenhagen, but see Section 2.5.3 for an overview of different outcomes of plosive weakening. In Copenhagen Danish, plosive weakening was a chain of sound changes with much the same end result as spirant weakening: following a series of well-motivated sound changes (see Section 3.5.2), the stops [p t k] lenited to voiced fricatives in post-vocalic position, following the patterns in (2).

(2)	[p]	→	[v]
	[t]	→	[ð]
	[k]	→	[ɣ]

These voiced fricatives subsequently lenited to approximants in analogous ways to spirant weakening (1), although as with spirant weakening, [ð] was largely retained as a voiced fricative (until the 20th century; see Section 2.2.2). The change [p] → [ʊ] was presumably largely complete in earlier stages of the language, but this change was later rolled back in many varieties including Modern Standard Danish, such that most of these words are again realized with [p] (see Chapter 3, in particular Section 3.5.2.5).

The geminate stops underwent degemination sometime after the plosive weakening chain was completed, and were not affected by it; in West Jutlandic varieties, these developed into preglottalized stops, a phenomenon known as *vestjysk stød* 'West Jutlandic stød' (Ringgaard 1960). West Jutlandic stød parallels preaspiration in Icelandic and Faroese (e.g. Thráinsson 1978; Page 1997).[4] In Modern Standard Danish, these are plain voiceless stops [p t k].

2.2.2 Recent changes in stops and semivowels

According to Brink and Lund (2018), voicing in /b d g/ was lost sometime before 1700. This conclusion is rather controversial. Several scholars have argued that Proto-Germanic did not have distinctive voicing, but rather distinctive aspiration (e.g. Iverson and Salmons 1995, 2003a; Honeybone 2002). The bulk of the modern daughter languages do indeed have an aspiration-based contrast. However, there *are* traces of a voicing-based contrast in regional varieties of Danish (see Setion 6.6), and the results of spirant weakening and plosive weakening are arguably easier to explain if a voicing-based contrast is assumed in earlier stages of the language. This is tricky to resolve, but will be discussed further in Chapter 3, in particular Section 3.5.3. If we assume that Brink and Lund (2018) are right, and Danish had voicing until relatively recently, it is unclear if aspiration in /p t k/ developed simultaneously or was already present.

The 19th century saw the loss of the final vestiges of velar fricatives in Copenhagen Danish (Brink and Lund 1975). [ɣ] was either lost or replaced entirely by [j w] following the distribution in (1). [x], which occurred only before [t] in post-vocalic consonant clusters, hardened to [k].

A number of changes affected [ð], which is usually considered an allophone of /d/ (see Section 2.4): sometime before the 19th century, it weakened to a very open approximant or semivowel (the 'soft d'; see Brotherton and Block 2020), which is described as alveopalatal by Grønnum (1998). There is no fitting IPA symbol, so usually [ð] (or

[4]In fact, Kortlandt (1985) argues on the basis of a.o. West Jutlandic stød that preaspiration and glottalization in present-day Germanic languages are reflexes of glottalic stops in PIE.

[ð̞]) are used, mostly for historical reasons. The approximant has more recently developed a secondary dorsal articulation (Brink and Lund 2018); an ultrasound tongue imaging study by Siem (2019) suggests that this dorsal gesture is actually often of a greater magnitude than the coronal gesture. I follow Schachtenhaufen (2022) in using [ɤ̯] to transcribe the semivowel as realized in Modern Standard Danish.[5] The soft d participates in a number of frequent and obligatory schwa assimilation processes of the form /əɤ̯, ɤ̯ə/ ⟶ [ɤ], so it is often realized as fully syllabic (e.g. Brink and Lund 1975: 191ff., Basbøll 2005: 293ff.). In a recent development, the soft d has begun to assimilate not just with schwa, but also with preceding front vowels, such that it appears as syllabic *and* stressed; Brink (2013) mentions neutralization of the contrast between /i e æ/ in *bid bed bad* 'bite pray bathe', all of which can be realized as [pɤˀ].

2.2.3 Summary

The Modern Standard Danish initial unaspirated stops mostly developed from PIE breathy-voiced stops, and there is no agreement in the literature regarding when they lost their voicing. The current initial aspirated stops mostly developed from PIE voiced stops. The PIE voiceless stops developed into fricatives, but the coronals later hardened to aspirated stops. In medial and final position, both fricatives and stops weakened dramatically in earlier stages of Danish, eventually eliding or becoming approximants. In Germanic, medial and final geminates developed from a number of sources; these resisted post-vocalic lenition, and are now retained as medial and final voiceless unaspirated stops.

[5] This seems a better fit than [ð̞], but is still rather imprecise. The sound is centralized relative to the canonical position of [ɤ] (Juul et al. 2016; see also Section 1.4), and the tongue blade is somewhat raised. Anecdotally, non-native listeners often perceive the soft d as lateral, possibly due to the coronal component.

2.3 Phonetics

This section provides an overview of existing research on the phonetics of stops in Modern Standard Danish. This research can be divided into a number of subcategories. I will discuss, in turn, existing research on aspiration and voice onset time (Section 2.3.1); closure voicing (Section 2.3.2); actions of the glottis and larynx during stop production (Section 2.3.3); affrication during release (Section 2.3.4); articulatory force (Section 2.3.5); and finally, perturbations of fundamental frequency (F_0) resulting from the laryngeal setting of preceding stops (Section 2.3.6).

2.3.1 Aspiration and voice onset time

Grønnum (2005: 303ff.) has claimed that the only relevant difference between the two laryngeal stop series is aspiration. This applies only to onset position, as there is no (phonetic) laryngeal contrast in syllable-final stops;[6] the phonological situation may be more complicated, as discussed in Section 2.4.4 and further in Chapter 3. According to Grønnum, the only articulatory difference between /b d g/ and /p t k/ in onset position is the degree of glottal opening at the time of release. For /p t k/, the glottis is spread widely apart; for /b d g/, the vocal folds are adducted and tensed. As we will see in the coming paragraphs and chapters, there are a number of reasons why the claim that this is the only relevant difference must be an oversimplification.

Jespersen (1906: 60ff.) made a distinction between two sets of aspirated sounds, *bepustede* ('blown') and *beåndede* ('breathed'). The 'blown' aspirated sounds are found in strong, stressed position, and 'blown' probably refers to the salient burst that is found in this position.[7] The 'breathed' aspirated sounds are found in unstressed onset position, and have no salient burst, but they still have delayed voicing onset. This idea illustrates an important point about the phonetics of Danish stops: aspiration is always possible in onset

[6] Syllable-final stops may also be aspirated, especially if they are in phrase-final position, but this is not contrastive (Grønnum 2005: 49).

[7] According to Jespersen (1906: 61), at the time of release, the "air pressure behind the closure is so much stronger than that outside, that an explosion happens and a bang is heard" (my translation).

position, regardless of stress or position within the phrase. This is illustrated by words like [pʰæˈtʰeːˀtʰisk] *patetisk* ‘emotive, pathetic’.

A common method for classifying laryngeal stop contrasts based on pre-voicing and aspiration is voice onset time (VOT), popularized by Lisker and Abramson (1964). In this seminal paper, the authors investigate stop contrasts in 12 different languages, and conclude that stops generally fall into four categories, which can mostly be summarized with reference to the timing of voicing onset: voiceless unaspirated, voiced unaspirated, voiceless aspirated, and voiced aspirated. In voiced stops, vocal fold vibration begins at some point during the closure (negative VOT); in voiceless unaspirated stops, vocal fold vibration begins very shortly after the release (near-zero positive VOT); in voiceless aspirated stops, vocal fold vibration is delayed (high positive VOT). Only voiced aspirated stops cannot be easily defined on the basis of VOT. VOT measurements remain very prevalent in the phonetics literature, and as more languages have been studied, Lisker and Abramson’s neat three-way VOT distinction has been put increasingly into question. In Cho and Ladefoged’s (1999) study of 18 languages, they make a four-way categorization of positive VOT: unaspirated, slightly aspirated, aspirated, and highly aspirated. Ladd (2011) points out that as more and more languages are added to these typologies, any kind of categorization of the VOT continuum is likely to become increasingly moot. On a related note, in recent years, large-scale studies using VOT measurements have often relied exclusively on positive VOT (e.g. Stuart-Smith et al. 2013, 2015; Chodroff and Wilson 2017; Chodroff et al. 2019). There are a number of reasons for this: negative VOT is difficult to measure automatically (Sonderegger and Keshet 2012), and determining the starting point for measuring negative VOT is problematic in anything but absolute initial position.

Studies of the relative timing of voicing onset in Danish stops precede Lisker and Abramson (1964). The duration of aspiration in Danish stops (i.e. positive VOT) is measured by both Abrahams (1949) and Fischer-Jørgensen (1954), and VOT in Modern Standard Danish has been measured in several studies since. The results of some of these studies are compared in Table 2.1. Fischer-Jørgensen (1954) worked with a corpus of roughly 3,000 spectrograms from 10 speakers, although it is not clear how many of these were measured to arrive

Table 2.1: *Overview of mean positive voice onset time in ms in different studies. The numbers in parentheses indicate ranges of individual averages. Mortensen and Tøndering (2013) do not provide absolute means, but only means by vowel height and vowel aperture. Range of means by vowel height are reported here.*

	FJ (1954)	FJ (1980:a)	FJ (1980:b)	FJ (1980:c)	MT (2013)
/b/	14	17	22	12	13–19
/d/	17	23	23	24	18–26
/g/	23	30	34	30	27–36
/p/	66 (53–97)	60 (51–76)	69 (54–76)	93 (78–103)	65–70
/t/	79 (64–98)	75 (71–89)	84 (75–93)	115 (88–135)	84–86
/k/	74 (60–91)	73 (64–96)	81 (64–94)	106 (94–118)	63–80

at the VOT means in Table 2.1. Her results conform well to those of Abrahams (1949). Based on this, she classifies /p t k/ as strongly aspirated, and /b d g/ as unaspirated or lightly aspirated.

It is not clear from Fischer-Jørgensen (1954) how aspiration is delimited, but this is something she has discussed at length in later writings (Fischer-Jørgensen and Hutters 1981). In later papers, and likely also in the 1954 paper, her acoustic landmark for delimiting aspiration is the appearance of higher formants. This makes the measurements reported by Fischer-Jørgensen (1979, 1980) relatively difficult to compare to typological overviews of VOT, since recent studies commonly use the onset of periodicity in the waveform to delimit aspiration, and Fischer-Jørgensen and Hutters (1981) show that there are non-trivial and unpredictable differences between these two methods. Francis et al. (2003) find that the onset of periodicity in the waveform is the landmark that most closely corresponds to the physiological onset of vocal fold vibration; however, Fischer-Jørgensen and Hutters (1981) rightly point out that measurements made at the onset of higher formants are more stable across vowel heights. It is worth pointing out that relying on higher formants only works when measuring VOT in CV-sequences, and that waveforms display higher temporal accuracy than spectrograms.

Fischer-Jørgensen (1979, 1980) investigated the relationship between a number of temporal cues to the laryngeal contrast in stops. She measured both VOT, closure duration, and following vowel duration. She found consistently longer VOT and shorter vowel duration in /p t k/, which is in line with expectations; she also found consistently shorter closure duration in /p t k/, which is cross-linguistically uncommon (this is discussed in more detail in Section 2.3.5). Interestingly, she also found social stratification of the aspiration cue. She split her informants into three groups: a) non-Copenhageners, b) older Copenhageners (born before 1938), and c) younger Copenhageners (born after 1938). She only worked with a total of 18 participants, so it is a limited data basis, but her results quite stably show that Copenhageners had longer VOT values than non-Copenhageners, and that VOT was longest for younger Copenhageners. This suggests a change in progress at the time: very strong aspiration is a recent Copenhagen-based innovation. This is in line with an historical account where the transition from a voicing system to an aspiration system is relatively recent (Brink and Lund 2018; see Section 2.2).

Mortensen and Tøndering (2013) is the most recent account of VOT in Danish stops, based on roughly 3,000 tokens of stressed CV syllables in the DanPASS corpus of spontaneous speech (see Section 4.5.1). Mortensen and Tøndering follow Fischer-Jørgensen in using the appearance of higher formants as their acoustic landmark for VOT measurements. Interestingly, the results of Mortensen and Tøndering's VOT measurements are more in line with Fischer-Jørgensen's (1980) findings for older speakers outside the Copenhagen area. This is obviously inconsistent with an account of VOT increasing with time, but there may be a number of possible explanations. For example, Mortensen and Tøndering's measurements come from spontaneous recordings, and Fischer-Jørgensen's measurements are seemingly from read word lists. This highlights the difficulty with comparing VOT measurements across studies with different designs.

In Chapter 5, I also report VOT measurements of /p t k/ from the DanPASS corpus, using the onset of periodicity in the waveform to delimit aspiration; unlike Mortensen and Tøndering (2013), these measurements are not exclusively taken from stressed syllables. Unsurprisingly, I find average VOT values that are significantly shorter than those in Table 2.1. These are later compared to VOT measurements from traditional Jutlandic varieties of Danish in Section 6.5.

A number of linguistic and extralinguistic factors are known to influence VOT, and many of these are covered in some detail in Section 6.5. Mortensen and Tøndering's (2013) study explicitly investigates the effect of following vowel height on VOT, comparing the traditional IPA-based classification (high, mid, and low) with a physiological four-way classification of aperture in Danish vowels suggested by Grønnum (2005: 105). They find similar effects across the two different ways of categorizing vowels, but only find a significant influence of vowel height/aperture on VOT in /b d g/, where VOT is generally longer before high/close vowels, and shorter before low/open vowels. Fischer-Jørgensen (1980) found this in both aspirated and unaspirated stops. Additionally, Andersen (1981a) investigated the influence of speech rate on the articulation of /p/, and found that VOT was shorter in quick speech. He also found shorter closure duration in quick speech, and less glottal aperture.

Fischer-Jørgensen (1972d) reports a series of perception experiments investigating the relative weighting of cues in determining laryngeal and place contrasts in Danish stops. She shows that an aspiration phase is a necessary and sufficient cue for perceiving the laryngeal contrast. Long VOT is not found to be sufficient: the study compares perception of stops with regular aspiration noise to stops with an equally long silent phase preceding voicing onset, and finds that /p t k/ are only consistently correctly identified if aspiration noise is present during the release. She furthermore finds that a threshold of roughly 35 ms VOT is required for consistent correct identification of /p t k/.

The alveolar stops have drawn particular interest in the area of second language acquisition, in part due to problems caused by the salient affrication of /t/ in Modern Standard Danish (see Section 2.3.4). For this reason, further studies have measured VOT in /d t/: Garibaldi and Bohn (2015) report surprisingly high mean VOT values of 28 ms and 140 ms for /d t/, respectively; Yan (2016) reports mean VOT values of 15 ms and 96 for /d t/ respectively; Puggaard (2020c) reports a mean VOT of 93 m in /t/. Given the nature of the material, these values are expected to be higher than in the spontaneous speech measured by Mortensen and Tøndering (2013), but even with that caveat, the values reported by Garibaldi and Bohn (2015) are extremely high.

In syllable-initial clusters of aspirated stops and sonorant consonants, the sonorants are usually said to devoice categorically as an allophonic realization of aspiration. This observation goes at least as far back as Jespersen (1890, 1906), who notes voiceless realizations of /l r n/ after underlyingly aspirated stops. He also notes that /r/ devoices after /f/. Finally, he notes that /j/ alternates with a voiceless palatal fricative [ç] in words like [ˈkʰjoːl̩] *kjole* 'dress', which Jespersen transcribes as [ˈkçoːlə].[8] Note that the phonetic alveopalatal fricatives and affricates [ɕ tɕ] are usually analyzed as underlying clusters of /sj/ and /tj/, and this can also be considered a case of phonologized devoicing (Basbøll 2005; Grønnum 2005). More recent treatments of Danish phonetics maintain that non-nasal sonorants devoice after aspirated stops, but do not mention nasals (Heger 1981; Grønnum 2005). Grønnum's transcriptions in particular indicate that devoicing of non-nasal sonorants is categorical following /p t k f s/.

In a recent article, Juul et al. (2019) investigate the extent of sonorant devoicing acoustically by comparing the duration of voiced and voiceless portions in e.g. /gæ glæ kæ klæ/. They do not find evidence of categorical devoicing – there is little difference in the duration of the voiced portion in sets like /glæ klæ/. As such, they suggest to dispense with the tradition of transcribing this contrast as [klæ kl̥æ]. Unfortunately, they do not look into the spectral characteristics of the release; in a study of the same phenomenon in English, Tsuchida et al. (2000) found that liquids are partially devoiced following aspirated stops; the entire voiceless release in clusters like /kl/ is lateral, but voicing does set in during the final third of the lateral; they found only negligible devoicing after voiceless fricatives. This discrepancy suggests that the explanation for sonorant devoicing lies in articulatory overlap: liquids (partially) overlap with stop aspiration, but can hardly overlap with fricatives. The measurements given by Juul et

[8]Jespersen's transcription is 'translated' from Dania. He further notes that sonorant consonants can devoice in coda following stød, but intuitively, this does not happen in Modern Standard Danish.

al. (2019) suggest that the situation in Danish may be similar to the situation in English. As I discuss further in Section 2.4.3.1, the (lack of) categoricity in sonorant devoicing has implications for how the stops can be represented phonologically (see Puggaard-Rode et al. forthc.).

2.3.2 Closure voicing

Voicing is not generally considered an important cue to the laryngeal contrast in Danish stops. Because of this, little has been written about closure voicing. To my knowledge, no quantitative studies have been carried out investigating the topic. Abrahams (1949), Fischer-Jørgensen (1954) and Spore (1965) all briefly mention the occurrence of intervocalic voicing; post-pausally, stops are never voiced, but all stops show some degree of closure voicing intervocalically. This has a universal phonetic explanation: there is high subglottal pressure immediately after a vowel, which naturally results in voicing during (at least) the initial part of a post-vocalic closure, unlike closures that are not post-vocalic (Westbury and Keating 1986). Fischer-Jørgensen (1954, 1980) assumes that medial stops before inflectional suffixes are archiphonemic – in the 1954 paper, she uses the notation /B D G/ – and writes that these are "often completely voiced" (1954: 44) and "almost always pronounced as (very) weakly voiced" (1980: 208). Keating et al. (1983) write that /b d g/ are spirantized medially, and that /p t k/ are voiced medially; this is an oversimplification both phonetically and phonologically, and one I will return to in Section 2.4.

As such, several scholars have mentioned in passing that Danish stops are voiced intervocalically, either categorically or near-categorically. However, Jessen (1999, 2001) and Beckman et al. (2013) assume that Danish stops are categorically voiceless in all positions. Neither camp provide empirical evidence for their claims. As we will see in Chapter 4, this significantly influences how the laryngeal contrast in Danish stops is modeled phonologically (as also discussed below in Section 2.4.3.1). The overall lack of interest in closure voicing in Danish is surprising; the two major corpus studies of phonetic reduction in Danish (Pharao 2009; Schachtenhaufen 2013) consider many patterns of variation in the realization of consonants, but do not look into voicing.

2.3.3 Actions of the larynx

Frøkjær-Jensen et al. (1971) studied the actions of the glottis during the production of Danish stops using a photo-electric glottograph. Comparing /b/ with /p/ in intervocalic onset position in stressed syllables, they find that there is a glottal opening gesture lasting throughout the closure phase of /p/, resulting in a fully spread glottis just after the stop release. This is unsurprising, and accounts for the aspirated release. More surprisingly, they also find a glottal opening gesture during the production of /b/, although this gesture has a shorter duration and smaller magnitude. The authors hypothesize that the smaller glottal opening gesture during /b/ follows naturally from the transition from vowel to consonant, and is not "effectuated by neural commands" (Frøkjær-Jensen et al. 1971: 134).

Hutters (1978, 1984, 1985) carried out similar studies using electromyography (EMG) and fiberoptic stills, and found support for Frøkjær-Jensen et al.'s findings: /b/ has a glottal opening gesture which peaks relatively early during the closure, and is of smaller magnitude than that of /p/. However, she questions Frøkjær-Jensen et al.'s proposed explanation that the small opening gesture in /b/ is a purely aerodynamic artifact of the vowel–consonant transition, since both her own EMG results and EMG results from Fischer-Jørgensen and Hirose (1974) show that the posterior cricoarytenoid muscles are active in achieving it.[9]

There are other reasons to doubt a purely aerodynamic explanation of the glottal opening gesture in /b/. In Westbury and Keating's (1986) model of closure voicing, they show that initial subglottal pressure in intervocalic position is high enough that vocal fold vibration should continue throughout a significant portion of the closure, assuming the vocal folds are adducted and tensed. By way of comparison, in the production of intervocalic /b/ in English, Hirose and Gay (1972) find no activity of the posterior cricoarytenoids, and Sawashima (1970) finds that there is usually no interruption of voicing and no arytenoid separation. Hutters (1985) proposes that the intervocalic

[9] Frøkjær-Jensen et al. (1971) was reprinted in 1973, and the reprint has an added note recognizing the forthcoming work of Fischer-Jørgensen and Hirose (1974), and that it throws doubt on their explanation of the findings.

glottal opening gesture in Danish is a measure taken to reinforce voicelessness in /b/, although she leaves the question relatively open. I return to this question in Chapter 4, when discussing why intervocalic voicing in Modern Standard Danish is so relatively rare.

The actions of the glottis result in aspiration, but also in other cues to the laryngeal contrast. Fischer-Jørgensen (1972d) reports that the first voicing pulses after /p t k/ are irregular in low vowels, and that it takes longer for higher vowel formants to appear after /p t k/ relative to /b d g/. Listeners demonstrably use these as cues to the laryngeal contrast before low vowels. Fischer-Jørgensen explains these phenomena with reference to the position of the vocal cords at the time of the burst.

Petersen (1983) has investigated whether there are vertical movements of the larynx during stop production in Danish. Petersen's motivation for looking into this is that larynx lowering has been found to correlate negatively with F_0 (e.g. Shipp 1975). Some studies suggest that F_0 is locally conditioned by the laryngeal setting of preceding stops in Danish (see Section 2.3.6), so Petersen tests whether larynx lowering is a likely explanation for this finding. He finds a (weak) correlation between local F_0 and larynx lowering, but does not find conclusive evidence of larynx lowering during /b d g/. Westbury (1983) identifies larynx lowering as one of several means to maintain voicing during closure; by lowering the larynx, the size of the supraglottal cavity is increased, which slows the increase in air pressure. This is perhaps not very relevant in Danish, where voicing plays little to no role in maintaining the laryngeal stop contrast (see Section 2.3.3 above).

I argue in later chapters that the seemingly marked laryngeal behavior found during stop production in Modern Standard Danish can help account for many of the interesting patterns we see in the acoustic signal, including the relative rarity of intervocalic voicing (Chapter 4) and the prevalence of affrication in aspirated stops (Chapter 5). Relatedly, I argue that different patterns of VOT, closure voicing, and affrication in regional varieties of Danish likely reflect differences in laryngeal behavior (Chapter 6).

2.3.4 Burst and affrication

Recall from Section 2.3.1 that Jespersen (1906) describes hearing an "explosion" immediately after the release of "blown" aspirated stops. This presumably refers to the release burst, which is well-known to be a crucial acoustic cue to place features of (particularly) aspirated stops (e.g. Blumstein and Stevens 1979). A salient release burst is the result of air escaping through a narrow gap becoming turbulent. The spectral characteristics of bursts are overall similar to fricatives at the same place of articulation, although significantly shorter in duration. Stop affrication serves to enforce cues for place of articulation, but also weakens cues for manner of articulation (see e.g. Repp et al. 1978).

Fischer-Jørgensen (1954) makes a number of observations about the acoustics of releases and release bursts, a.o. providing details about intensity and spectral characteristics of bursts. She finds that burst intensity decreases in the order /k/ > /p/ > /t/ although the difference between /k ~ p/ is blurred before rounded vowels. She also notes that the spectral characteristics of bursts are highly dependent on the quality of the following vowel. In a perception experiment, Fischer-Jørgensen (1972d) finds that there are sufficient place cues in the bursts of aspirated stops for listeners to determine the place of articulation if the aspiration phase is removed, and *vice versa*, that there are sufficient place cues in the aspiration phase if the burst is removed. The latter is especially true for /t/; if the aspiration phase of /t/ is superimposed on a bilabial or velar stop, it is still consistently perceived as /t/. Cues to the place of articulation of /b d g/ interact with the following vowel in complicated ways. Sometimes place is cued primarily from the characteristics of the burst; sometimes from the initial formant transitions in the following vowel. By comparison, formant transitions are a more stable cue to place of articulation in syllable-final stops (Fischer-Jørgensen 1972a), which are often unreleased except in phrase-final position (Grønnum 2005: 49).

Fischer-Jørgensen finds that the frequency range of the aspiration noise in /t/ is very similar (actually "exactly the same", 1954: 50) as the noise in /s/. Similar results had been found for /t/ in English, with the crucial difference that /s/-like noise in English /t/ continues only for roughly 25 ms (Jakobson et al. 1951), and /s/-like noise in Danish /t/

continues twice as long. Taking the VOT measurements reported in Table 2.1 into account, this implies that /t/-affrication is followed by aspiration proper, but that affrication takes up the bulk of /t/-releases. Fischer-Jørgensen finds no signs of affrication in /p/, and limited signs of affrication in /k/, where high-frequency noise is somewhat more prominent than in /h/ and /p/-releases.

Jespersen (1897–1899: 335) already noted that Danish /t/ was highly affricated, and that foreigners were likely to perceive it as an affricate before high front vowels (see also Hansen 1956: 56). He also noted that /t/ was affricated to some extent regardless of the following vowel, and suggested that this indicated a sound change in progress: /t/ ⟶ /ts/, as occurred earlier in High German with the Second Consonant Shift. He envisioned it as a sound change in progress for all aspirated stops, with /t/ ⟶ /ts/ being quite advanced, and /p k/ ⟶ /pf kx/ much less so.

Brink and Lund (1975) tracked the development of affrication in /t/ across more than a century's worth of recordings from Copenhagen. They found that it was already a widespread phenomenon in High Copenhagen Danish in the mid-19th century, and that it was an exceptionless feature of /t/ one century later. They even report cases of /t/ having lost closure altogether among younger speakers. Brink and Lund use [d^{s}h] when transcribing this sound, thus recognizing that the affricated portion of the release is followed by aspiration proper.[10] This transcription is in line with Hjelmslev's (1951) phonological analysis of the aspirated stops, where they are considered underlying clusters of /b d g/ + /h/ (see Section 2.4.2).

Basbøll and Wagner (1985) transcribe /t/ phonetically as [ts^{h}], which is in line with Fischer-Jørgensen's (1954) findings. It is unclear why, but in recent years, [t^{s}] has emerged as the standard transcription, indicating affrication but no aspiration (e.g. Grønnum 1998; Basbøll 2005). Schachtenhaufen (2022) has recently proposed transcribing the sound as a true affricate [ts]. This downplaying of aspiration is not in line with recent studies from Yan and Sloos (2019) and Puggaard (2020c), who both find that a sizable portion (20–25 ms on average) of /t/ releases is unaffricated.

[10] Brink and Lund's use of [d] does not indicate closure voicing; they transcribe using Dania, where sounds transcribed with [b d g] are not inherently voiced.

/t/ releases have been explored in some detail in previous studies, although the highly complex nature of the information in stop releases means that there are certainly still open questions. The properties of /p k/ releases remain largely unknown. In Chapter 5, I explore affrication patterns and spectral characteristics of aspirated stop releases further, and propose a method of data analysis that arguably allows us to statistically model this complex information without sacrificing the complexity of the data.

2.3.5 Articulatory force

Articulatory force is a problematic term in phonetics, because it has a large number of partially overlapping acoustic and articulatory correlates, some of which have been discussed already. Jaeger (1983) summarizes the phonetic properties that have been assumed to cue articulatory force in the literature. Fortis ('strong') consonants are variously claimed to have relatively greater pulmonic force; greater force or pressure of the articulators; rapid release of closures; longer duration; and no voicing. Because there is no general agreement in the literature about what articulatory force actually refers to, Henton et al. (1992) argue that the associated terms should be used very carefully. Henton et al. use the term only in the sense of increased respiratory effort, of which there are only few well-established examples, the most well-known one being the fortis stops of Korean.

Some approaches to phonological representation subsume the phonetic cues mentioned by Jaeger (1983) under a binary distinction between fortis and lenis.[11] In phonological descriptions of two-way laryngeal contrasts in stops, 'lenis' is usually straightforwardly taken to refer to /b d g/, and 'fortis' to /p t k/. For example, Kohler (1984) uses the feature [fortis] as a cover term for similar stop contrasts in Germanic languages, including Danish, where the contrasts involved have a wide variety of phonetic correlates. This is discussed further in Section 2.4.3.1.

[11] A number of largely equivalent terms are used in the literature; 'fortis' is sometimes used interchangeably with 'tense' or 'strong', and 'lenis' is sometimes used interchangeably with 'lax' or 'weak'.

Fischer-Jørgensen (1968, 1969) argues that articulatory force as a phonetic parameter should be kept separate from voicing and aspiration, since all three may show independent behavior. In acoustic studies, she finds that the closure duration in /b d g/ is consistently longer than /p t k/ (Fischer-Jørgensen 1972b), and EMG studies comparing /b ~ p/ show a tendency for greater organic pressure in /b/, although this is not significant for all speakers (Fischer-Jørgensen and Hirose 1974). In a questionnaire study, Fischer-Jørgensen (1972b) shows that speakers of Danish consistently judge /b d g/ to be produced with stronger organic pressure than /p t k/. She takes these findings as evidence that /b d g/ are actually 'more fortis' than /p t k/, at least as pertains to supraglottal articulatory force – although this is unlikely to affect perception much. Articulatory force, however, appears to be highly language-specific; Fischer-Jørgensen (1968) also finds that, on a number of parameters, the difference in articulatory force between French 'lenis' /b d g/ and 'fortis' /p t k/ is much greater than that between the Danish laryngeal series, and that all Danish stops are produced with less force than French /b d g/. Given that voicing and short duration are commonly associated with 'lenis' stops, it is remarkable how much both Danish laryngeal series resist voicing (Section 2.3.2). I will return to this problem in later chapters; in Chapter 5, I argue that the lack of sharp releases in /p t k/ may help to explain affrication patterns.

I am unaware of more recent research into articulatory force in Danish stops, but it is often mentioned in newer publications that Danish stops are all lenis (e.g. Basbøll and Wagner 1985; Basbøll 2005; Grønnum 1998, 2005). The most common way of transcribing the stressed syllable-initial allophones of /b d g/ is [b̥ d̥ g̊], with the devoicing diacritic used to indicate that they are phonetically lenis (Grønnum 1998). This strategy is also used in some traditions of narrow transcription of English (e.g. Lodge 2009). The syllable-initial allophones of /p t k/ are usually transcribed as [p^h t^s k^h], but it is sometimes claimed that more accurate transcriptions would be [b̥^{h} d̥^{s} g̊^{h}], since these are also phonetically lenis (e.g. Grønnum 1998; Basbøll 2005).

Schachtenhaufen (2022) proposes getting rid of the devoicing diacritic in Danish transcription. The *Handbook of the International*

Phonetic Association (IPA 1999: 24) briefly mentions the use of the devoicing diacritic with voiced obstruent symbols:

> "The voiceless diacritic can be used to show that a symbol that usually represents a voiced sound in a particular language on some occasions represents a voiceless sound."

This is clearly not the case in Danish (see Section 2.3.2 and Chapter 4), and IPA guidelines never indicate that the diacritic may be used to indicate articulatory force. As such, I am strongly in favor of the proposal to abandon [b̥ d̥ g̊] in favor of [p t k].

2.3.6 Fundamental frequency

It is well-known that laryngeal contrasts tend to influence F_0 in the initial part of a subsequent sonorant sound (e.g. House and Fairbanks 1953). This is well-established for languages with voicing-based contrasts (so-called 'true voicing' languages; Kirby and Ladd 2016), but has also been found for languages with aspiration-based contrasts (see Hanson 2009 and references therein), and in the case of Swiss German, even for a singleton–geminate contrast (Ladd and Schmid 2018). There is disagreement in the literature about the phonetic mechanism causing F_0-perturbations. Some argue that F_0 is lowered by voiced stops, due to e.g. the larynx lowering gesture discussed in Section 2.3.3 above (e.g. Kingston and Diehl 1994). Others argue that F_0 is raised locally by voiceless stops, since F_0 is similar after voiced stops and nasals (e.g. Hanson 2009). This suggests that voiceless stops are the ones showing exceptional behavior, since nasals should have no impact on F_0.

F_0-perturbations have sometimes been considered crucial in determining underlying representations, because they are a relatively stable feature of laryngeal two-way contrasts, regardless of how that contrast is otherwise realized. This prompted Keating (1984a) and Kingston and Diehl (1994) to incorporate it into their otherwise abstract [voice] features. On the other side of the debate, Goldstein and Browman (1986) assume that the F_0-perturbations are purely an artifact of glottal gestures.

Danish is interesting in this respect, since both laryngeal stop series are characterized by some degree of glottal opening (see Section 2.3.3). Several studies have investigated F_0-perturbations in Danish with inconclusive results. Fischer-Jørgensen (1969) finds no evidence of laryngeally induced differences in F_0 immediately after stops, while other studies have found evidence of higher F_0 after aspirated stops relative to unaspirated stops (Jeel 1975; Petersen 1978, 1983). The study by Petersen (1983) is of particular interest here, since he compares stops with a number of other consonants. He finds a tendency for negligibly higher F_0 after aspirated stops relative to unaspirated stops; however, he also finds a more stable tendency for increased F_0 after all obstruents relative to nasals. These findings suggest that the slight glottal opening gesture found for /b d g/ (see Section 2.3.3) causes a local spike in F_0, and the complete glottal opening gesture found for /p t k/ and voiceless fricatives causes a greater local spike in F_0. This is not well in line with a categorical featural explanation, such as Kingston and Diehl's (1994).[12] I return to the phonological implications of F_0-perturbations in Section 2.4.3.1.

2.3.7 Summary

Danish stops fall into two categories: unaspirated /b d g/ and aspirated /p t k/. Both categories have quite high VOT from a cross-linguistic perspective. Intervocalically, both series are described as showing some degree of initial closure voicing, and the medial allophones are generally (weakly) voiced. Both series are accompanied with a glottal opening gesture – likely to enforce voicelessness – but the magnitude of this gesture is quite small in /b d g/. There is salient affrication during the release in the realization of /t/, regardless of phonetic context. Surprisingly, the unaspirated series have longer overall longer closure duration and higher organic pressure, which are often taken as correlates of greater articulatory force. There is some evidence for differences in F_0-perturbations caused by the two laryngeal series, but also

[12]Kingston and Diehl (1994) cite Petersen (1983) and others as evidence for [voice] acting as an F_0-depressor in Danish, but do not engage with the intricacies of Petersen's findings.

evidence that F_0 is locally raised after both stop series, possibly due to the glottal opening gestures.

2.4 Phonology

Having covered the existing literature on the phonetics of Danish stops, I now turn to the existing literature on their phonology. I cover which processes the stops participate in, and previous proposals regarding their underlying representation. This requires a general overview of the Danish consonant inventory and phonotactics, which I give in Section 2.4.1.

In Section 2.4.2, I review previous descriptions of the stops' combinatorial possibilities, which in some cases include consonant clusters that only exist at an abstract underlying level. In Section 2.4.3, I review accounts of how the stops are underlyingly represented, focusing on the laryngeal contrast and place contrasts. Finally, alternations between stops and semivowels remain a tricky problem in Danish phonology, and in Section 2.4.4, I provide an overview of how this phenomenon (which is usually called *consonant gradation*) has previously been analyzed in the literature. Chapter 3 is further dedicated specifically to this issue.

2.4.1 Consonant inventory and positional allophones

Positional allophones in Danish are usually described in terms of 'strength' rather than syllabic position. Following Jakobson et al. (1951), 'strong' allophones are found in absolute initial position, and 'weak' allophones are found in either syllable-final position or syllable-initially in unstressed syllables before the central vowels [ə ɐ], as well as unstressed [i] in some specific morphemes.

The strong consonant allophones are shown in Table 2.2, along with the phonemes that they are commonly taken to represent. The place labels given in the table are simplified for reasons of space, and diacritics are kept to a minimum; for more information on the transcription conventions and more precise place labels, see Section 1.4.

Table 2.2: *Strong allophones of Danish consonants and the phonemes they are commonly assumed to represent.*

	Labial	Alveolar	Palatal	Dorsal	Glottal
Unaspirated stop	[p] /b/	[t] /d/		[k] /g/	
Aspirated stop	[p^{h}] /p/	[t^{h}] /t/		[k^{h}] /k/	
Affricate			[tɕ] /tj/		
Fricative	[f] /f/	[s] /s/	[ɕ] /sj/		[h] /h/
Nasal	[m] /m/	[n] /n/			
Approximant	[ʋ] /v/	[l] /l/	[j] /j/	[ʁ] /r/	

The alveopalatals [tɕ ɕ] are usually (unproblematically) treated as surface mergers of underlying /tj sj/ clusters (e.g. Basbøll 2005). This is phonetically reasonable, and serves to explain why [j] is found after all obstruents syllable-initially *except* [t^{h} s], which would otherwise be a structural gap.

As discussed in Section 2.3.2, the unaspirated stops are often transcribed as [b̥ d̥ g̊], but I follow the proposal of Schachtenhaufen (2022) in abandoning this convention and representing them with [p t k] instead, for reasons outlined above and in Chapter 4. Further, as discussed in Section 2.3.4, the strong allophone of /t/ has been transcribed in a number of different ways, but in recent years most Danish phoneticians have converged on [t^{s}]. I go for the simple solution [t^{h}] here to emphasize the class behavior of the aspirated stops; this problem will be discussed further in Chapter 5.

[ʋ] and [ʁ] are sometimes described as voiced fricatives (e.g. Heger 1981; Haberland 1994; Grønnum 1998), and [ʋ] is often transcribed with the [v] symbol in other sources. However, both Grønnum (1998) and Basbøll (2005) explicitly note that both sounds lack friction,

Table 2.3: *Weak allophones of Danish consonants and the phonemes they are commonly assumed to represent.*

	Labial	Alveolar	Palatal	Dorsal
Stop	[p] (/b ~ p/)	[t] /t/		[k] (/g ~ k/)
Fricative	[f] /f/	[s] /s/	[ɕ]* /sj/	
Nasal	[m] /m/	[n] /n/		[ŋ] (/ng/)
Approximant	[ʋ]* /v/	[l] /l/	[j]* /j/	[ʁ]* /r/
Semivowel	[ʊ̯] (/v ~ b ~ g/)	[ɣ̯] (/d/)	[ɪ̯] (/j ~ g/)	[ɐ̯] /r/

logically making them approximants.[13] I stick to the more precise phonetic transcriptions here. There are no IPA symbols for uvular or pharyngeal approximants, but a narrow transcription of the rhotic could be [[ʁ̞ ~ ʁ̞̙ ~ ʕ̞]].

The weak allophones are shown in Table 2.3, along with the phonemes they are commonly taken to represent. Phonemic associations which will later be contested are given in parentheses. Allophones which are found only marginally in weak position are marked with an asterisk *. As above, more precise place labels and discussion of transcription conventions can be found in Section 1.4.

There is only one series of surface stops in weak position. This does not (necessarily) mean that the contrast is neutralized; as discussed in Section 2.4.4 and Chapter 3, the 'traditional' phonological analysis of Danish holds that the semivowels sometimes derive from /b d g/.

The semivowels frequently syllabify in unstressed syllables due to schwa assimilation, e.g. /rə/ ⟶ [ɐ] (see Brink and Lund 1975: ch. 32; Basbøll 2005: ch. 11; Schachtenhaufen 2010b). Grønnum (e.g. 1998) assumes that only [ɐ̯] is a phonetic semivowel, and otherwise uses the approximant symbols [w j] rather than [ʊ̯ ɪ̯]. As discussed in Section 2.2, the so-called 'soft d' [ɣ̯] is a very open semivowel with

[13] Basbøll (2005) uses [v], which has the unfortunate consequence that he must postulate a feature [voice] which is only distinctive for labiodental fricatives.

both coronal and dorsal articulatory components (Siem 2019; Brotherton and Block 2020). Grønnum (1998) describes it as alveopalatal and does not mention a dorsal gesture; Basbøll describes it as alveolar with a secondary velar component (2005; see also Ejstrup 2010); Brink and Lund (2018) describe the secondary articulation as pharyngeal. In fact, the precise articulatory targets of the coronal and dorsal gestures are unknown, and difficult to pinpoint since the tongue is relatively far from the roof of the mouth at all points. The more commonly used transcription is [ð], and Basbøll (2005) proposes the narrow transcription [[ð̠ ˠ̞]], under the (arbitrary) assumption that the dorsum most closely approximates the velum. I follow Schachtenhaufen (2022) in using a vocalic transcription, as this is more in line with the phonetic substance, and also highlights the class behavior of the four semivowels. Schachtenhaufen (forthc.) proposes the narrow transcription [[ɤ̯̈]].

The central approximants [ʋ j ʁ] and the fricative [ɕ] are found only marginally in weak position. [ʋ j] are only found after [l] in a small number of words, such as [ulˀʋ] *ulv* 'wolf' and [ʋælˀj] *valg* 'choose'. These clusters were historically more wide-spread, but have been lost in many cases where they are still reflected in writing, e.g. [hælˀ] *halv* 'half'. Particularly the remaining [lj] clusters are highly unstable, with realized [j] only found in highly distinct speech (see relevant entries in the pronunciation dictionary of Brink et al. 1991). [ʁ] is only found in weak position in a small set of imperatives (Basbøll 2018). Danish infinitive verbs are usually minimally disyllabic and have underlying word-final schwa; the regular imperative is formed by removing the final schwa, which in some cases leads to clusters with highly marked sonority slopes. In very conservative Danish, the imperative of a verb like [ˈkʰlætʁɐ] *klatre* 'to climb' would thus be monosyllabic [kʰlætʁ̥] *klatr!* 'climb!' with devoiced [ʁ̥]. The syllable structure of these imperatives runs counter to that predicted by any sonority hierarchy (see Parker 2002). Most speakers nowadays use syllabic weak allophones to avoid sonority reversals, as in [ˈk^{h}lætɐ].[14] Finally, [ɕ] is only found in weak position in a small number of recent loans, such as [hæɕ] *hash* 'hashish'.[15]

[ŋ] is sometimes assumed (by e.g. Grønnum 2005) to be derived from underlying /ng/. This is discussed in further detail in Section 2.4.2.2. The phonemic associations of particularly final [p k ʊ̯ ɪ̯] are problematic and highly complex, with multiple potential phonemes associated with single allophones, and often no sure way to determine the underlying representation of allophones. This is discussed further in Section 2.4.4 and Chapter 3.

2.4.2 Combinatorial possibilities

In this section, I will discuss the phonotactics of stops. I review possible initial and final consonant clusters and analyses of these, as well as medial behavior, in Section 2.4.2.1. Subsequently, I discuss phonological analyses that have assumed abstract consonant clusters including stops in Section 2.4.2.2.

2.4.2.1 Surface clusters

Danish phonotactics are relatively permissive. Most possible combinations of stop + sonorant are found syllable-initially. Vestergaard (1967) gives the following attested stop-initial consonant clusters:

(3)

/b/	+	/	j	r	l			/
/d/	+	/	j	r		v		/
/g/	+	/	j	r	l		n	/
/p/	+	/	j	r	l			/
/t/	+	/	j	r		v		/
/k/	+	/	j	r	l	v	n	/

[14] The pattern is similar in other sonority-reversing imperatives. The imperative of [ˈʋeklə ~ ˈʋekl̩] *vikle* 'to wrap' is [ʋekl̥] *vikl!* 'wrap!' in highly conservative Standard Danish, but disyllabic [ʋekl̩] is much more common in Modern Standard Danish. Due to schwa assimilation in the infinitive, this may lead to neutralization between some infinitive–imperative pairs.

[15] Some speakers nativize this word as [hæs] to avoid [ɕ] in weak position (Brink et al. 1991).

The possible combinations in (3) are in agreement with versions of the sonority hierarchy that treat (voiceless) obstruents as a class (e.g. Basbøll 1977; Clements 1990); stops are never found before other voiceless obstruents, but frequently before sonorants. There are a number of further restrictions, many of which can be explained with reference to the Obligatory Contour Hierarchy (OCP; Goldsmith 1976): /b p + v/, /d t + l/, and /d t + n/ are not found due to shared place of articulation. /n/ is more restrictive than non-nasal sonorants, and /m/ is not found in clusters at all, which is not too surprising: nasals are often taken to be the 'least sonorous sonorants' (Krämer and Zec 2020), but coronals are generally very phonotactically permissive (e.g. Yip 1991). The lack of /gv/-clusters is more difficult to explain. There are at least two examples of syllable-initial (but not word-initial) /gv/-clusters, derived from the same root: [leŋ.ˈkʋist] *lingvist* 'linguist' and [leŋ.kʋi.ˈstik] *lingvistik* 'linguistics'. This suggests that the lack of word-initial /gv/-clusters is an accidental gap.

/s/ occurs quite freely in initial clusters, and most of the combinations in (3) are also found after /s/, exceptions being [stʋ-] and [skn-].[16] The laryngeal contrast in stops is neutralized after /s/, and only (phonetically) unaspirated stops are found. Uldall (1936) argues that only /p t k/ are found after /s/ underlyingly, since [skʋ-] *is* a possible cluster, as in e.g. [skʋæt] *skvat* 'wimp' – like /kv-/, but unlike /gv-/ (although note the exception above).

It is cross-linguistically common for /s/ to display this kind of phonotactic behavior (e.g. Goad 2012). This is often modeled phonologically by assuming that /s/ is not attached to the syllable, but rather to a higher level of prosodic structure (e.g. Goldsmith 1990). Basbøll's (e.g. 1994, 1999, 2005) model of Danish (and general) phonotactics, the Sonority Syllable Model, gets around this problem by assuming an unconventional set of 'order classes'. They are defined as in (4), which shows the order classes from most to least sonorous, where each lower order class is a proper subset of higher classes. [ ] covers all segments, regardless of underlying representation.

(4) [+vocoid] > [+sonorant] > [+voice] > [-spread glottis] > [ ]

[16] [skl-] is marginal and found only in [skleˈʁoːsə] *sklerose* 'sclerosis'.

More sonorous segments occur closer to the center of the syllable, and less sonorous segments occur closer to the edges. This is a maximal general sonority hierarchy; in Danish, the three innermost classes can be collapsed with minimal loss in explanatory value, as in (5).

(5) [+voice] > [-spread glottis] > [ ]

Basbøll assumes that the features in (4) and (5) need to apply at the level of positional allophone. As mentioned above, only unaspirated stops, which are [-spread glottis], are found after /s/. As such, his Sonority Syllable Model predicts clusters with peripheral /s/ without additional formal machinery. On the other hand, the model also predicts a large number of non-occurring clusters, and has trouble explaining why /s/ is the only possible [+spread glottis] segment in three-member clusters; at least /f/ and /h/ should be equally likely.

Vestergaard (1967) also gives a list of possible final clusters. He lists phoneme combinations, but due to the problems with determining phonemes from weak allophones (see Sections 2.4.1 and 2.4.4), I list positional allophones instead. Recall that there is no laryngeal contrast in weak position. Possible two-member clusters with final stops are listed in (6):

(6)	[	ʊ̯		ɐ̯	l	m				s			] + [p]
	[	ʊ̯	ɪ̯	ɐ̯	l	m	n	ŋ	f	s	p	k	] + [t]
	[		ɪ̯	ɐ̯	l			ŋ		s			] + [k]

[t] occurs very freely as the final member of final clusters. The only weak allophone not found before [t] in monomorphemes is [ɤ̯].[17] The lack of [ɤ̯t] clusters can be historically explained with reference to the OCP, since [ɤ̯] developed from an alveolar stop (see Section 2.2). Synchronically, such an OCP constraint cannot carry too much weight, as [ɤ̯t] should be no problem to produce; such clusters are indeed found in certain polymorphemic monosyllables, such as [leːˀɤ̯+t] *ledt* 'cruel (indefinite neuter)'.

/p k/ only cluster with homorganic nasals, whereas all nasals are found before /t/. There are similar restrictions for semivowel + stop

[17] The marginal final allophones [ɕ ʋ ʁ] are also not found before [t].

clusters. This kind of phenomenon has often been explained with reference to 'coronal underspecification' (e.g. Avery and Rice 1989; Paradis and Prunet 1989); if [t] is assumed to have no underlying place features, there can be no requirements for place sharing. As is the case syllable-finally, [s] occurs freely before all stops, but otherwise, obstruents are only found before [t]. Stop + stop clusters occur freely, as long as the final segment is [t]. Apart from coronal underspecification, another likely explanation for this freedom is analogy with polymorphemic syllables: [-t] is a neuter suffix in adjectives, so it is found freely in word-final position at the surface level.

Final three-member clusters are much more free than initial ones, and Basbøll's Sonority Syllable Model cannot explain the order of these. Most clusters ending in /-st -sk/ are possible, including clusters of three obstruents, such as [tʰɑkst] *takst* 'fare'. The situation becomes all the more complex if polymorphemic syllables are taken into account. These complex final clusters will not be discussed further in this dissertation, but they are discussed in great detail by Basbøll (2005: ch. 7).

When clusters of two homorganic stops are found at morpheme boundaries, they are realized as geminates, as in e.g. [pʰukːɒːˀˈʁoːɤ] *Puggaard-Rode*. If the second stop is aspirated, the result is an aspirated geminate, as in [ˈpʰlæstekːʰiʁuːˀ] *plastikkirurg* 'plastic surgeon'. Compounding and derivational morphology do not otherwise affect phonotactics, although they do affect allophone selection in the traditional analysis of gradation patterns, as discussed in Section 2.4.4.

2.4.2.2 Underlying clusters

Early attempts at categorizing the distinctive phonemes of Danish were much more interested in distributional properties than distinctive features (Fischer-Jørgensen 1952). Accounts by Uldall (1936), Hjelmslev (1951), and to a lesser extent Basbøll (1968) seem primarily interested in arriving at the lowest possible number of phonemes, and are relatively unconcerned about the level of abstraction needed to arrive at that number.

Hjelmslev (1951) and Basbøll (1968) both assume that there are only three phonemic stops /b d g/, and that surface aspirated stops are underlyingly clusters of /b d g/ + /h/. This solution is abandoned by Basbøll in later writings (e.g. Basbøll and Wagner 1985), because it causes serious problems in explaining the distribution of /h/. Hjelmslev (1951) acknowledges this problem, and solves it by assuming that initial aspirated stops are underlyingly /h/-initial, such that a word like [kʰænˀ] *kan* 'can' is underlyingly /hgand/.[18] Nowadays, presumably everyone assumes that there are two series of phonemic stops in Danish, and aims to distinguish them by means of e.g. distinctive features instead, as discussed in Section 2.4.3.1.

Hjelmslev (1951) also assumed that stød is not an underlying property but rather structurally derived from e.g. final consonant clusters. This led him to propose that words with stød but no clear 'stød basis' according to his criteria must have abstract underlying final clusters, specifically of the type /ld rd nd/. An example is the word [lønˀ] *løn* 'salary', which Hjelmslev assumes is underlyingly /lønd/; cp. words with similar surface structure but no stød, like [kul] *guld* 'gold', which he assumes is underlyingly /gul/. Hjelmslev's analysis of stød was not particularly influential. The current mainstream account (as laid out in Basbøll 1985, 2003, 2005) affords no special role to final consonant clusters, but assumes that stød follows from a bimoraic rhyme, i.e. a long vowel, diphthong, or short vowel + sonorant consonant; Basbøll assumes that obstruents are never moraic in Danish.[19] In this account, words like [kul] *guld* 'gold' lack stød because the final sonorant is underlyingly specified as extraprosodic, and thus not parsed as moraic.

Another abstract consonant cluster rooted in structuralist phonology which has had more staying power is the analysis of [ŋ] as underlying /ng/ (e.g. Uldall 1936; Vestergaard 1967; Grønnum 2005). One (abandoned) reason for this analysis was the glossematic principle that all phonemes must appear in both onset and coda position (Hjelmslev 1936). A more compelling reason is that high

[18]This kind of analysis would probably be deemed overly abstract by most phonologists nowadays, but if the goal is to arrive at the lowest possible number of phonemes, Hjelmslev is undeniably successful: he ends up with a system of only 18 phonemes in Danish.

[19]See Vázquez-Larruscaín (2021) for an argument that all consonants are moraic, and a 'stød filter' rules out obstruents with stød.

vowels are lowered before any nasal + consonant cluster *and* before [ŋ], causing neutralization between high and mid-high vowels, as in (7).

(7) /i y u/ →[e ø o] / _ C[+nasal] C)σ, *or*
_ [ŋ])σ

Similarly, the vowel length contrast is neutralized before final consonant clusters *and* [ŋ], where only short vowels are found. Both rules can be formalized more economically if we assume that [ŋ] is underlyingly a nasal + consonant cluster.

These patterns are undoubtedly found because [ŋ] historically developed from nasal + stop clusters. Since Chomsky and Halle (1968), it has been a topic of much debate how much diachronic information is available to current speakers of a language. Hale and Reiss (2000) and Scheer (2015) argue that there is no principled limit on the abstractness of underlying representations, while Blevins (e.g. 2004) argues that explanations for phonological patterns should generally be found in diachrony, and that diachronic information is not available to current speakers. I tend strongly towards the second position here: even if [ŋ] behaves like a nasal + stop cluster, that is not sufficient evidence that speakers still store it as such.

Grønnum (2005: 308) points out that that the /g/ in /ng/ actually surfaces in some [ŋ]-final roots in few morphological contexts, namely before the verbalizing suffix [-ˈeːˀɐ] *-ere* and the demonym suffix [-ˈɛnˀsɐ] *-enser*, both of which cause stress shift, as in (8).

(8)	[tifˈtʰʌŋ]	*diftong*	'diphthong'
	[tiftʰʌŋˈkeːˀɐ]	*diftongere*	'diphthongize'
	[ˈkʰʌleŋ]	*Kolding*	(city name)
	[kʰʌleŋˈkɛnˀsɐ]	*koldingenser*	'person from Kolding'

These are taken as evidence that /g/ in /ng/-clusters deletes in final position after the application of a place assimilation rule, and /g/ surfaces only if resyllabified to a subsequent stressed syllable. The number of such alternations is extremely limited, and /g/ does not surface before unstressed syllables, as evidenced by the alternation in

(9) and several imperative–infinitive alternations in verbs with final [ŋ].

(9)	[ˈsɑŋˀ]	*sang*	‘song’
	[ˈsɑ.ŋɐ]	*sanger*	‘singer’

It can be discussed whether the alternations in (8) provide evidence that *diftong* and *Kolding* have /g/ specified at some level of representation, but it does not seem sufficient to posit that all instances of [ŋ] have underlying /g/. For comparison, consider the city names *Esbjerg* and *Hamborg* and their demonyms in (10):

(10)	[ˈɛspjæɐ̯ˀ]	*Esbjerg*	(city name)
	[ɛspjæɐ̯ˈkɛnˀsɐ]	*esbjergenser*	‘person from Esbjerg’
	[ˈhɑmpɒːˀ]	*Hamborg*	‘Hamburg’
	[hɑmpɒˈkɛnˀsɐ]	*hamborgenser*	‘person from Hamburg’

The names of *Esbjerg* and *Hamborg* are both in principle decomposable: *bjerg* translates as ‘mountain’, *borg* as ‘castle’. It is unlikely that speakers store them as compounds, though; the first syllables are cranberry morphemes, and the second parts are phonetically reduced and semantically hardly associated with these place names.[20] In their unreduced forms, these two words are realized as [pjæɐ̯ˀʊ̯] *bjerg* and [pɒːˀʊ̯] *borg*, respectively. Recall from Table 2.3 that [ʊ̯] is sometimes analyzed as a weak allophone of /g/; [k] surfacing in the demonyms *esbjergenser* and *hamborgenser* may be considered evidence that /g/ is still present in the underlying forms of these particular city names, even if it is never otherwise realized[21] – but it should certainly not be considered evidence that every instance of [ɐ̯] is derived from underlying /rg/.[22] Since the analysis does not work in (10), I propose that it also does not work in (8).

[20] In fact, the area around Esbjerg is notoriously flat.

[21] I would not actually argue in favor of this. The *-enser* suffix is unproductive and quite rare, even in demonyms; more likely, the demonyms are stored separately rather than actively derived.

[22] Note that there are multiple reasons to assume that [ɒː], as seen finally in *Hamborg*, is underlying /ɔr/. This is evidenced by fully productive verb alternations like infinitive [stɔːˀ] *stå* ‘to stand’ and present tense [stɒːˀ] *står*.

2.4.3 Underlying representation

Determining how phonemes are underlyingly represented is a key priority in phonology; most commonly, this is done with distinctive features, but I discuss other mechanisms as well. Two sets of features are relevant in distinguishing the Danish stops: place features, in order to group the stops into classes of /b p/, /d t/, and /g k/; and laryngeal features, in order to group the stops into classes of /b d g/ and /p t k/. There has been little discussion of how place of articulation is underlyingly represented, even though this is actually crucial in order to account for the gradation patterns discussed in Section 2.4.4 and Chapter 3. The specific nature of the laryngeal contrast has been subject to much discussion. This is due to a.o. differences in opinion about what phonological patterns should be accounted for, and how phonetically substantial the representational mechanism should be.

2.4.3.1 Representation of laryngeal contrast

Danish stops have frequently come up in discussions about how the distinction between /b d g/ and /p t k/ should be underlyingly represented. Kohler (1984) assumes the feature [fortis]; Keating (1984a) and Kingston and Diehl (1994) assume an abstract feature [voice] with variable phonetic correlates; Iverson and Salmons (1995), Basbøll (2005), Grønnum (2005), and Beckman et al. (2013) use [spread glottis] with clear phonetic correlates;[23] Goldstein and Browman (1986) use gestural scores; Puggaard-Rode et al. (forthc.) model the contrast as a quantity difference in subsegmental representational units. There are several other proposals concerning laryngeal representation in the literature that do not touch on Danish in particular, which I will not discuss here, including Chomsky and Halle (1968), Halle and Stevens (1971), Lombardi (1995), and Avery and Idsardi (2001).

Kohler's (1984) use of the feature [fortis] (a 'power' feature as opposed to a laryngeal feature) was already mentioned in Section 2.3.5 above. The feature is intended to broadly cover laryngeal contrasts in obstruents, but also to have clear phonetic correlates. Kohler clarifies some of these, and discusses some phonological processes associated with the feature. [+fortis] has two primary articulatory correlates:

[23]This approach is termed 'laryngeal realism' by Honeybone (2005).

1) Greater articulatory power in oral stricture formation, which is a.o. cued by quicker movement towards closure, and leads to the common sound changes of vowel shortening and vowel lowering before [+fortis] obstruents ('pre-fortis clipping'); and 2) greater activity of glottal muscles, implemented either as tensing of the vocal cords or wide opening of the glottis, and cued by e.g. closure voicing, aspiration, and F_0-perturbations. The feature can also be implemented in various other ways. Kohler notes that closure voicing is an "extreme manifestation" (1984: 163) of [-fortis], since air stream power and articulatory tension are necessarily very low during closure voicing.

Kohler's treatment of Danish is somewhat contradictory. On the one hand, he assumes that a lower degree of oral stricture in lenis stops (/b d g/) is the cause of the historical plosive weakening chain discussed in Section 2.2, as well as the stop–semivowel alternations still found synchronically (see Section 2.4.4 and Chapter 3): since /b d g/ are lenis, they are produced with less tight stricture than /p t k/, and are thus more likely to lose that stricture during quick, spontaneous speech (Kohler 1984: 156–158). On the other hand, Kohler is aware of and cites the research by Fischer-Jørgensen and Hirose (1974) showing that Danish /b d g/ are *not*, in fact, produced with a less tight stricture than /p t k/ (see Section 2.3.5). According to Kohler (1984: 164–165), this is due to a trade-off between the oral and glottal correlates of [fortis]: a long aspiration phase is correlated with short closure duration and *vice versa*, and since aspiration in Danish fortis stops is very long and prominent, closure duration is correspondingly short, and the force of closure is correspondingly weak. These two positions are seemingly incompatible; the first requires /b d g/ to have weaker relative organic pressure, while the second acknowledges that /p t k/ have weaker relative organic pressure and seeks to explain why. Finally, recall from Section 2.3.6 that studies of F_0-perturbations in Danish are inconclusive, and F_0 is locally raised after all stops. Kohler asserts to explain Danish stop gradation with [fortis], but the relevant phonetic correlates of [fortis] seemingly do not apply to the Danish contrast.

Keating (1984a) criticizes the idea that phonological features should necessarily be phonetically substantial. She argues that such an approach unavoidably results in too many features, which have no generalizability beyond one or a few languages. She argues for a three levels of representation: 1) a universal phonological level, where only possible phonologically active contrasts are specified; 2) a universal phonetic category level, where only possible phonetically distinct contrasts are specified; and 3) a continuous phonetic level, where the phonetic categories are implemented according to language-specific rules. At the phonological level, the abstract feature [voice] is sufficient to account for all two-way laryngeal contrasts in obstruents. This can be implemented with the phonetic categories {voiced} and {voiceless unaspirated}, or {voiceless unaspirated} and {voiceless aspirated}[24] – these contrasts should play no role at the phonological level. [voice] is argued to be active in at least three common phonological processes, regardless of which phonetic categories it maps onto: vowel duration effects, F_0-perturbations, and assimilation. Danish is specifically mentioned here as a language with progressive [voice] assimilation, as opposed to Polish, which has regressive [voice] assimilation. Keating does not go into any detail here, but she presumably refers to progressive devoicing of sonorants after aspirated stops (as discussed in Section 2.3.1) – i.e. progressive assimilation of [-voice]. Her account of assimilation is short, and misses important assymetries in laryngeal assimilation: [+voice] spreads in languages where [+voice] is cued with closure voicing, and [-voice] spreads in languages where [-voice] is cued with aspiration, as pointed out by Iverson and Salmons (1995). The direction of spreading is also asymmetrical, such that voicing usually spreads regressively (Lombardi 1999), while aspiration usually spreads progressively (Iverson and Salmons 1995).

Kingston and Diehl (1994) also argue that all Germanic languages have a [voice] distinction in obstruents, with two main correlates: [+voice] has earlier voicing onset relative to release than [-voice], and [+voice] always acts as an F_0-depressor.[25] There is certainly a difference in relative voicing onset between the two laryngeal series in Danish, but as mentioned several times above, there is little (and conflicting) evidence for laryngeally induced F_0-perturbations. Kingston and Diehl *do* assume a feature [spread glottis] for e.g. languages with three-way

[24] See Vaux and Wolfe (2005) for a discussion of the other logical combination, {voiced} and {voiceless unaspirated}, which is not mentioned by Keating (1984a).

[25] Recall from Section 2.3.6 that the directionality of F_0-perturbations assumed by Kingston and Diehl (1994) has been contested by e.g. Hanson (2009).

laryngeal contrasts and a number of Chinese languages. When they analyze Danish as having contrastive [voice], it is ostensibly due to categorical intervocalic voicing in Danish (ibid.: fn. 16). Recall from Section 2.3.2 that this is actually understudied; I return to this issue in Chapter 4, where it is shown that Modern Standard Danish does not have categorical intervocalic voicing in /b d g/. The specific issues with Kingston and Diehl's treatment of the Danish data would likely be solved if Danish was analyzed as having a [spread glottis]-based contrast rather than a [voice]-based contrast.

Iverson and Salmons (1995) opt for a less abstract representation of the laryngeal contrast. They use [spread glottis] as the feature responsible for the laryngeal contrast in most Germanic languages, with the notable exception of Dutch. This approach accounts for a number of phonological processes and restrictions. Key among them is devoicing of sonorants following [+spread glottis] segments, which is analyzed as progressive assimilation of [+spread glottis], and which does not occur in so-called 'true voicing' languages, where the laryngeal contrast is cued with closure voicing. However, recall from Section 2.3.1 that acoustic studies of both Danish (Juul et al. 2019) and English (Tsuchida et al. 2000) find that this devoicing process is much less categorical than often assumed. In particular, these studies found very little evidence for sonorant devoicing after voiceless fricatives. Iverson and Salmons' approach also accounts for the loss of aspiration in /s/ + stop clusters, by assuming that a single [+spread glottis] feature is linked to both consonants. Other work has also proposed that onsets can have only one glottal gesture or laryngeal feature (Browman and Goldstein 1986; Anderson and Ewen 1987; Kehrein and Golston 2004).

Beckman et al. (2013) follow Iverson and Salmons in assuming that only [spread glottis] is active in languages with aspiration-based stop contrasts. They go on to suggest, following a proposal by Chomsky and Halle (1968), that + and - values for features are converted to numerical values on a language-specific basis during phonetic implementation. In an aspiration language like German, /b d g/ are [-spread glottis] and /p t k/ are [+spread glottis]. During phonetic implementation, [+spread glottis] is converted to a high numerical value like [9sg] and [-spread glottis] to a low numerical value like [1sg]. The low value assigned to /b d g/ allows for passive voicing in intervocalic position.

In Danish, [+spread glottis] is also converted to a high value of [9sg], while [-spread glottis] is converted into a medium-high value like [5sg]. This ostensibly accounts for why passive voicing in Danish /b d g/ is blocked (following Jessen 2001) – recall that Kingston and Diehl (1994) sought to explain essentially the opposite pattern, namely that intervocalic voicing in Danish /b d g/ is categorical; recall also from Section 2.3.2 that there is actually no consensus about the extent of passive voicing in Danish. Kirby and Ladd (2018) provide a critique of Beckman et al.'s (2013), particularly as pertains to F_0-perturbations.

Grønnum (2005) also assumes that the relevant laryngeal feature is [spread glottis], and that this accounts for sonorant devoicing. She assumes a separate feature [aspirated], which is added to [+spread glottis] stops syllable-initially, as in (11) (ibid.: 403):

(11) [+spread glottis, -continuant] ⟶ [+aspirated] / σ(_

F_0-perturbations should be a problem to a [spread glottis] account, since they are often taken as evidence of a more abstract [voice] feature. However, as Goldstein and Browman (1986) point out, F_0-perturbations can be interpreted as a direct result of laryngeal activity during stop articulation. They argue that an articulatory account of F_0-perturbations provide a better explanation of the Danish data. Their proposal for underlying representation of the Danish laryngeal contrast is in line with their Articulatory Phonology framework (Browman and Goldstein 1986, 1992), which framework abandons discrete segments altogether, and instead assumes continuous underlying representations consisting of gestural scores of varying magnitudes.

Puggaard-Rode et al. (forthc.) account for the Danish laryngeal contrast using a novel representational framework called Q-CV. Q-CV is an extension of Q-theory (Inkelas and Shih 2017; Shih and Inkelas 2019a), which proposes a quantized approach to phonological representation. Traditional segments are divided into distinct units (subsegments) which may carry separate sets of distinctive features, and traditional segments are emergent from these (Shih and Inkelas 2019b). The earliest papers suggested that all traditional segments consist of three subsegments, but it has more recently been argued that quantity contrasts can be modeled by varying the number of subsegments (Garvin et al. 2018, 2020; Schwarz et al. 2019). The innovation of Q-CV is that subsegments are anchored by root nodes which are defined in terms of simple head–dependency relationships between C and V components, inspired by Dependency Phonology (e.g. Anderson and Ewen 1987). A root node C is defined as a complete closure with no outgoing air, and C_V is defined as a prominent constriction resulting in turbulent airflow. Puggaard-Rode et al. propose that most attested laryngeal contrasts in stops (see Ladefoged 1973; Henton et al. 1992) can be represented by varying the number and order of C and C_V nodes.

The plain unaspirated stops /b d g/ in Danish are represented as in (12), where 'pl' represents any place feature, and co-indexing indicates that the place feature is the same throughout.

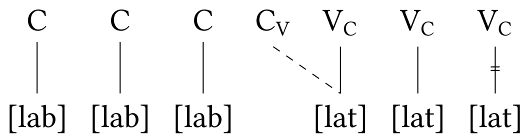
(12)
$$\begin{array}{ccc} \mathrm{C} & \mathrm{C} & \mathrm{C} \\ | & | & | \\ \mathrm{pl}_i & \mathrm{pl}_i & \mathrm{pl}_i \end{array}$$

The aspirated stops /p t k/ are represented as in (13). The prominent aspirated release is modeled as a bare C_V node with no associated place features.

(13)
$$\begin{array}{cccc} \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{C_V} \\ | & | & | & \\ \mathrm{pl}_i & \mathrm{pl}_i & \mathrm{pl}_i & \end{array}$$

An advantage of this approach is that it elegantly captures the gradient nature of sonorant devoicing in Danish. A [p^hl] sequence is shown in (14). Rather than a feature like [spread glottis] spreading from the stop to the liquid, partial liquid 'devoicing' is modeled as the feature [lateral] spreading into the bare C_V node of the stop release. (The other root nodes of the lateral are marked as V_C, which is defined as a constriction which does not significantly obstruct air flow.)

(14)



In (14), the final [lateral] feature delinks from its root node due to a process of 'compensatory shortening'. This happens because the feature [lateral] retains a fairly distinct spectral profile when voiceless (e.g. Maddieson and Emmorey 1984), unlike nasals, which do not have a significantly shortened voiced phase after aspirated stops (Juul et al. 2019).

The approach of Puggaard-Rode et al. can also account for why sonorant devoicing is negligible after /f s/. Voiceless fricatives, like aspirated stop releases, have C_V root nodes, but they have associated place features, which inhibit features like [lateral] from spreading to them.

Summing up, Kohler (1984) assumes that the laryngeal distnction is modulated with the feature [fortis], but his approach cannot account for findings relating to articulatory force in Danish stop production (see Section 2.3.5). Keating (1984a) and Kingston and Diehl (1994) assume a [voice]-based distinction, where feature values are cued by language-internal relative differences in voicing onset rather, laryngeally induced F_0-perturbations, and the presence of intervocalic voicing, rather than actual closure voicing. The first criterion certainly holds for Danish, but the second is not fully supported by the literature (see Section 2.3.6), and the extent of the third is rather overestimated (see Section 2.3.2 and Chapter 4). Iverson and Salmons (1995) assume the feature [spread glottis] modulates the contrast, and this ostensibly explains patterns of sonorant devoicing, although this effect has also been rather overestimated (see Section 2.3.1). Beckman et al. (2013) assume that [-spread glottis] in Danish is assigned a medium-high value during phonetic implementation, accounting for why passive intervocalic voicing is blocked – i.e. assuming a different set of phonetic facts than Kingston and Diehl (1994). Goldstein and Browman (1986) use gestural scores to represent the contrast, and argue that this can account for the gradient patterns of F_0-perturbations reported in the literature (see Section 2.3.6). Puggaard-Rode et al. (forthc.)

Table 2.4: *Overview of proposals regarding the representation of laryngeal contrast in Danish.*

Proposal	**Feature or mechanism**	**Phonetic correlates**	**Explained phonological patterns**
Kohler (1984)	[fortis]	Articulatory force in oral stricture and glottal activity	Stop gradations
Keating (1984a)	[voice]	No correlates required	Progressive devoicing of sonorants
Kingston and Diehl (1994)	[voice]	Relative voicing onset and F_0-perturbations	Intervocalic voicing
Iverson and Salmons (1995)	[spread glottis]	Aspiration	Progressive devoicing of sonorants
Beckman et al. (2013)	[spread glottis]	Aspiration	No passive voicing
Goldstein and Browman (1986)	Gestural scores	Voicing, aspiration, F_0-perturbations	–
Puggaard-Rode et al. (forthc.)	Additional C_V root node	Aspiration	Gradient sonorant devoicing

represent the contrast with the presence or absence of a bare C_V root node in the novel framework of Q-CV. This approach can account for the gradient nature of sonorant devoicing found by Juul et al. (2019). In Table 2.4, I give an overview of the different proposals.

The proposals regarding laryngeal representation aim to account for different, sometimes conflicting, phonetic phenomena and phonological processes. Some issues could be resolved if we had a clearer idea about the extent of passive voicing in /b d g/; in Chapter 4, I summarize the predictions of different approaches regarding intervocalic voicing, and discuss them again in the light of new data from a corpus study.

2.4.3.2 Representation of place contrasts

Grønnum (2005) follows Chomsky and Halle (1968) and most later work in distinctive feature theory (see Broe 1992) in using place features that refer to the active articulator: /b p/ are [labial], /d t/ are [coronal], and /g k/ are [dorsal]. Basbøll (2005), somewhat untraditionally, refers to the passive articulator in phonological feature labels, such that /d t/ are [alveolar] and /g k/ are [velar]. In spite of this difference at face value, the choice of active vs. passive articulator has no major consequences for their respective phonological analyses, at least not as pertains to stops. As we will see in the next section, this is partially because neither of them attempt to account for the concomitant changes in place of articulation that result from stop–semivowel alternations. I argue in Section 3.4.2 that the labels LABIAL, CORONAL, and DORSAL (as used in Feature Geometry; e.g. Sagey 1986) can more readily account for the stop–semivowel alternations than the features used by Basbøll and Grønnum.

2.4.4 Gradation patterns

The Danish stop gradation patterns have already been briefly covered in Section 2.4.1 above, and a critique and reanalysis will follow in Chapter 3. Here, I will cover the history of the traditional analysis of Danish stops and their allophony, and the main arguments usually presented in favor of the analysis. I will also exemplify the relevant patterns.

Uldall (1936) hinted at part of the analysis when he proposed that the phonemes /d g/ were realized as [t k] in strong position and [ð ɣ] in weak position.[26] As Jakobson et al. (1951) point out, this results in an analysis where the phonemic affiliation of an unaspirated [t] depends on its prosodic position – [t] in strong position derives from underlying /d/, while [t] in weak position derives from underlying /t/. Interestingly, around the same time, Martinet (1937: 41ff.) noted the possi-

[26] I use the symbols [ð ɣ] here as the sounds were likely less vocalic at the time.

Table 2.5: *Realizations of stop phonemes, following Grønnum's (2005) analysis.*

Phoneme	*Strong*	*Weak*
/b/	[p]	[p ~ ʊ̯]
/d/	[t]	[ɤ̯ ~ Ø]
/g/	[k]	[k ~ ɪ̯ ~ ʊ̯ ~ Ø]
/p/	[pʰ]	[p]
/t/	[tʰ]	[t]
/k/	[kʰ]	[k]

bility of an analysis similar to Uldall's, but argued explicitly against it on the grounds that it would be excessively abstract.

Hjelmslev (1951) also analyzes [ð ɣ] as allophones of /d g/, although recall from Section 2.4.2.2 that he assumed only one series of stop phonemes. As a result, for Hjelmslev, [ð ɣ] in weak position are analyzed as /d g/, while [t k] in weak position in weak position are analyzed as abstract clusters of /hd hg/.

Rischel (1970a) was the first to propose a generative analysis of the full system, drawing on morphophonological alternations. Rischel's analysis still relied heavily on the presence of the 'soft g', a velar voiced fricative or approximant [ɣ ~ ɰ]. Recall from Section 2.2 that the soft g is no longer found in Modern Standard Danish, although it would have been found in conservative varieties at the time. Rischel also incorporated alternations between /g/ and [ɪ̯ ʊ̯]. This indicates that his analysis covers a previous diachronic stage, where stylistic alternations between [ɣ ~ ɰ] and the semivowels [ɪ̯ ʊ̯] were still observed. This latter point is made explicit in the similar analysis by Basbøll (1975: 65–67), where he mentions stylistic alternations of the form [lɔːˀv ~ lɔːˀʊ̯] *lov!* 'promise!' and [lɔːˀɣ ~ lɔːˀʊ̯] *låg* 'lid'. Such stylistic alternations were a strong argument in favor of the traditional analysis, but they are no longer found in Modern Standard Danish.

The analysis given by Grønnum (2005) is similar to Richel's and the early analysis by Basbøll, except it no longer relies on the soft g, which had been completely lost at this point. Grønnum's analysis yields the strong and weak realizations of stop phonemes seen in Table 2.5. Since stylistic alternations that support the traditional analysis are no longer observed, support for the analysis can only be found in morphophonological alternations. Evidence for the traditional analysis of /p t k/ comes from alternations with e.g. the verbalizing suffix [-ˈeːˀɐ] *-ere*, which causes stress shift, as in (15).

(15)	[kæˈlʌp]	*galop*	'gallop (n.)'
	[kælʌˈpʰeːˀɐ]	*galopere*	'to gallop'
	[ʋæt]	*vat*	'cotton wool'
	[ʋæˈtʰeːˀɐ]	*vattere*	'to apply cotton wool'
	[lɑk]	*lak*	'lacquer (n.)'
	[lɑˈkʰeːˀɐ]	*lakere*	'to lacquer'

[p ~ ʊ̯] alternations are found to varying extents throughout the speech community, but are generally considered less 'standard' than the other stop–glide alternations (as evident from relevant entries in the pronunciation dictionary of Brink et al. 1991). When /b/ can be a realized as [ʊ̯], there is stylistic variation between [ʊ̯ ~ p]; this is unlike /d ɡ/, where replacing semivowels with stops is generally ungrammatical.[27] Evidence for [p ~ ʊ̯] alternations comes from strong declension in verbs, in particular the present participle [-t] *-t* suffix and the past tense [-tə] *-te* suffix; usually, strong allophones are found in coda position before these suffixes, as in (16).[28] The strong verbal declension is both irregular and unproductive.

(16)	[ˈkʰøːøp ~ ˈkʰøːʊ]	*købe*	'to buy'
	[ˈkʰøptə, *ˈkʰøʊ̯tə]	*købte*	'bought'

Note that the vowel copy in [ˈkʰøːøp] and syllabicity of [ʊ̯] in [ˈkʰøːʊ] are due to schwa assimilation; see Brink and Lund (1975: ch. 32), Basbøll (2005: ch. 11), and Schachtenhaufen (2010b, 2012, 2013) for overviews of these complex phenomena. Note also that [ˈkʰøʊ̯tə] is certainly acceptable in some varieties of Danish, but in these varieties, there is

[27] A notable counterexample is found in the suffix [-ɤ] *-et*, which may be either a past participle suffix in verbs or a neuter definite suffix in nouns. This suffix displays stylistic (and largely regional) variation between [-ɤ ~ -ət] (Petersen et al. 2021).

[28] Note that weak allophones are found before the otherwise homophonous neuter gender [-t] *-t* suffix in adjectives.

no evidence to suggest that the consonant in question is underlyingly /b/.

Evidence for stop–semivowel alternations in /d/ generally comes from unproductive morphological derivations in Latinate loanwords, as in (17).[29] As with [-ˈeːˀɐ] *-ere*, the suffixes [-iˈtʰeːˀt] *-itet* and [-ˈik] *-ik* both cause stress shift.

(17)	[soˈliɣ̞ˀ]	*solid*	'solid'
	[solitiˈtʰeːˀt]	*soliditet*	'solidity'
	[meˈtʰoːɤ]	*metode*	'method'
	[metʰoˈtik]	*metodik*	'methodology'

/d/ is often left unrealized, because there are several phonological environments where [ɣ̞] is not allowed. For example, /d/ is usually deleted before [t], as in (18), where the neuter suffix [-t] *-t* is added to a [ɣ̞]-final root. This is ostensibly due to an OCP constraint, but recall from Section 2.4.2.1, though, that there is little reason why a synchronic OCP constraint should rule out a surface cluster of [ɣ̞t]. More likely, this is the result of an OCP constraint operating historically.

(18)	[soˈliɣ̞ˀ]	*solid*	'solid'
	[soˈlit]	*solidt*	'solid (neu.)'

[ɣ̞] is consistently not found in coda position after other consonants, but in Rischel's (1970a) analysis, /d/ is Ø-realized but phonologically present in a number of words where it emerges in the derivational morphology. His examples come from the adjectivizing suffixes [-i] *-ig* and [-isk] *-isk*, as in (19).

(19)	[sønˀ]	*synd*	'sin'
	[ˈsønti]	*syndig*	'sinful'
	[joːˀɐ̯]	*jord*	'earth'
	[ˈjoɐ̯tisk]	*jordisk*	'earthly'

[29] I will use the established term Latinate (e.g. Plag 1999: 54ff.) throughout the dissertation to refer to loanwords of either Greek or Latin origin.

A problem with this, as Rischel also notes, is that these suffixes are usually preceded by weak consonant allophones (including [ɣ̞]) in other words, as demonstrated in (20).

(20)	[tyɣ̞ˀ]	*dyd*	'virtue'
	[ˈtyːɣ̞i]	*dydig*	'virtuous'
	[ˈjøːɤ]	*jøde*	'Jew'
	[ˈjøːɣ̞isk]	*jødisk*	'Jewish'

This could be an argument in favor of analyzing the [t ~ Ø] alternations in (19) as evidence for suppletive roots rather than the result of a synchronically active phonological process.

Evidence for stop–semivowel alternations in /g/ comes from both strong declension of verbs and derivations of Latinate loanwords. In accordance with historical spirant weakening as shown in (1), a very rough system for the realization of weak /g/ can be expressed as in (21).

(21)	/g/	⟶	[ɪ̯]	/	after front vowels
			[ʊ̯]	/	after back vowels

Note, however, that /g/ is not realized if the process in (21) results in a diphthong with minimal movement, i.e. *[iɪ̯ yɪ̯ uʊ̯ oʊ̯].[30] In some cases, the patterns in (21) only reflect older stages of Danish; historical changes in vowel quality have not necessarily led to corresponding changes in /g/-allophone. An example is the word [ˈstɑːɪ] *stege* 'fry' which has a back vowel in Modern Standard Danish, but surfaces with a front vowel and [k] when it undergoes (strong) declension, as in [stekt] *stegt* 'fried'. This suggests that the rule in (21) may not be synchronically active.

Both [ɪ̯] and [ʊ̯] alternate with [k] in the morphology of words like *bage* 'to bake'. Here, morphophonologically induced vowel shortening causes vowel backing [æ ⟶ ɑ] in the highly lexicalized compound

[30] [eɪ̯ øɪ̯] are found in conservative Copenhagen Standard Danish, but not in younger Copenhagen Standard Danish or in my variety. In conservative Copenhagen Standard Danish, the words *hø* 'hay' and *høg* 'hawk' are pronounced as [høːˀ] and [høːˀɪ̯], respectively; in younger Copenhagen Standard Danish, the distinction is neutralized, and both are pronounced [høːˀ] (Schachtenhaufen 2020–). In (my variety of) Jutlandic Standard Danish, however, there seems to be a distinction between length and overlength in such pairs: [høːˀ] *hø*, [høːːˀ] *høg*.

bagværk 'baked goods' and in the strong declinations, as seen in (22); unlike in the *stege* example above, this [ɑ] clearly patterns as a back vowel.

(22)	[ˈpæːɪ]	*bage*	'to bake'
	[ˈpɑʊ̯ʋæɐ̯k]	*bagværk*	'baked goods'
	[ˈpɑktə]	*bagte*	'baked'

Evidence for [k ~ Ø] alternations are found in the derivational morphology of some Latinate loanwords, as in (23).

(23)	[filoˈloːˀ]	*filolog*	'philologist'
	[filoloˈkiːˀ]	*filologi*	'philology'

This particular alternation between [oˈloːˀ] and [oloˈkiːˀ] is found in several pairs of words denoting scientific fields and practitioners of those fields, such as *antropolog* ~ *antropologi* 'anthropologist ~ anthropology', *psykolog* ~ *psykologi* 'psychogist ~ psychology', *fonolog* ~ *fonologi* 'phonologist ~ phonology'. It is possible that speakers have reanalyzed the original stress-shifting [-iːˀ] *-i* suffix as a stress-shifting [-kiːˀ] *-gi* suffix in this group of words.

Rischel (1970a: 472) derives the gradation patterns with the ordered set of rules in (24).

(24) a. [+obstruent, -continuant] ⟶ [-voiced] / strong position
b. [+voiced] ⟶ [+continuant]

He assumes that the feature modulating the laryngeal contrast is [voiced]. In (24-a), non-continuant obstruents are devoiced in strong position. In (24-b), any remaining [+voiced] sound, i.e. any [+voiced] obstruent in weak position, is weakened to a continuant.

Grønnum (2005: 402) derives the gradation patterns with the simple rule in (25).

(25) [-spread glottis, -continuant] ⟶ [+voiced, +continuant] / _)σ

In prose, unaspirated stops are realized as voiced continuants syllable-finally.

These rules have little explanatory value and appear functionally unintuitive, but otherwise capture the concomitant changes in manner features quite well. They do not, however, capture the changes in place features. Grønnum even specifically notes that this rule captures the process /g/ → [ɣ], but requires another step (which she does not formalize) to derive the semivowels that are actually realized; in other words, the rule requires speakers to reconstruct a lost allophone in the process of derivation. This problem is not limited to /g/; in fact, the rules cannot straightforwardly account for the places of articulation of any of the derived semivowels. I discuss this problem further in Section 3.3.2.

Pharao (2004) reports a psycholinguistic production study where speakers were asked to form morphologically complex nonce words in order to test the productivity of the gradation patterns. For example, a participant would be asked to form a verb from the nonce noun [slyɣ̯ˀ] with the verbalizing [-ˈeːˀɐ] suffix; [slyˈteːˀɐ] would be taken as evidence for the productivity of stop gradation, [slyˈɣ̯eːˀɐ] as evidence against it. His results show that productivity of the gradation pattern is speaker-specific: some reproduce it, others do not. In any case, this may not be evidence of a productive phonological rule, but could just as well be considered evidence of highly specific productive morphological schemas, following Bybee (e.g. 1985, 2001; Bybee and Slobin 1982)

Basbøll (1968, 2005, 2015) assumes a strict division between phonology and morphophonology, and does not require core phonology to account for stop gradation. He requires that all allophones of the same phoneme share distinctive feature(s), and that a phoneme can be determined for each allophone on the basis of purely phonological principles. As such, he analyzes /ɣ̯/ as a separate phoneme, and [ʊ̯ ɪ̯] as allophones of /v j/ – always. He also has a further level of representation, where morphophonemes like |b d g| can be realized as phonemes like /v ɣ̯ j/. Morphophonemes are established from phonemes on the basis of morphological alternations. In this model, the units of lexical storage are morphophonemes, but Basbøll remains skeptical of the psychological validity of this construct. Basbøll's analysis is otherwise essentially identical to the analysis of Grønnum (2005), but the division of labor between different grammatical modules is quite different.

Ács et al. (2008; see also Ács and Jørgensen 2016) reanalyze the stop–semivowel alternations in a pair of papers that weigh the Natural Phonology notion of biuniqueness very highly; namely, the notion that phone–phoneme correspondences are as transparent and uniform as possible (e.g. Stampe 1969). In their analysis, each distinctive sound translates into a separate phoneme. Accordingly, they propose a much higher number of distinctive phonemes, but do not require an elaborate set of rules to arrive at positional allophones.

In this section, I have provided an overview of the traditional analysis of Danish stop gradation. I will return to this analysis in Chapter 3, where I argue that the traditional analysis reflects sound change rather than synchronic phonology, and that the patterns of stop gradation do not play a role in synchronically active phonology.

2.4.5 Summary

Danish positional allophones are best analyzed as belonging to a 'strong' series and a 'weak' series. Two laryngeal series of stops are found in strong position, and only one is found in weak position. Danish is a phonotactically quite permissive language, and the stops are found in a large number of surface clusters. The stops have been assumed to play a role in a number of abstract clusters, and it is often assumed that [ŋ] is derived from an underlying /ng/-cluster; I have argued against such an analysis. A number of different representational mechanisms have been claimed to modulate the laryngeal contrast in stops – [fortis], [voice], [spread glottis], gestural scores, and subsegmental quantity. Most of these accounts run afoul of some of the phonetic facts presented in Section 2.3 above. Finally, there is a long tradition for subsuming aspirated stops in strong position with unaspirated stops in weak position, and unaspirated stops in strong position with semivowels in weak position. I have given an overview of the (morpho)phonological evidence in favor of this analysis, and will return to arguments against this analysis in Chapter 3.

2.5 Variation

The dicussion of phonetics and phonology of Danish stops has so far been based on a rather narrow conception of 'Danish'. The studies summarized in Section 2.3 mostly characterize the phonetics of distinct Standard Copenhagen Danish as spoken in the laboratory. Descriptions of the phonology of Danish (Section 2.4) are also frequently concerned with a conservative, distinct form of Standard Copenhagen Danish. This is not an accident: modern-day Denmark has sometimes been described as one of the most radically standardized speech communities in Europe (Pedersen 2003, but see also Maegaard and Monka 2019). This does not, however, mean that Danish is without internal variation – nor that the traditional dialects, many of which are now moribund, have gone undescribed or undocumented. Peculiarities in the realization of stops play a salient role in certain sociolinguistic varieties of Danish, and differences in the phonetic realization and the phonological trajectories of stops abound in regional varieties and traditional dialects.

In Section 2.5.1 below, I summarize research intophonetic reduction of stops in Modern Standard Danish. Section 2.5.2 summarizes research into the social significance of different realizations of stops, while Section 2.5.3 provides a brief summary of existing descriptions of regional variation in stop phonetics and phonology.

2.5.1 Phonetic reduction

Pharao (2009, 2011, 2012) and Schachtenhaufen (2013) both studied reduction patterns in spontaneous Modern Standard Danish on the basis of the DanPASS corpus (Grønnum 2009; see Section 4.5.1). These in-depth studies are based on the relatively narrow transcriptions accompanying the recordings, rather than on the actual acoustic signal. Pharao focuses only on consonants, while Schachtenhaufen gives a more holistic overview of how all distinctive sounds behave in spontaneous speech. In what follows, I use parentheses around transcriptions to indicate that they are neither phonetic nor phonemic, but rather refer to study variables: (p t k) refer to stops that would be realized as aspirated in careful distinct speech, and (b d g) as unaspirated. This does not presuppose a phonological analysis.

The results of Pharao and Schachtenhaufen, insofar as they pertain to stops, are similar. Both find quite low rates of reduction in (p t k) and (b). (g) is reduced at very high rates – roughly half of all word-internal tokens of (g) are reduced in some way, although outright deletion is rare. The mechanisms responsible for (b g) reduction seem to be similar. (b g) both show highest reduction rates intervocalically, and speakers who reduce at high rates tend to do so for both sounds. The mechanism responsible for reduction in (d) is seemingly different. (d) deletes at quite high rates (22% of all word-internal tokens), while other forms of (d)-reduction are rare (Pharao 2011). The most common context for (d)-deletion is interconsonantal position, presumably as a result of cluster simplification. This may be because coronal stops are found frequently in affixes. Probabilistic deletion of particularly coronal stops in clusters is a well-known phenomenon in English (e.g. Guy 1991; Tanner et al. 2017; Holtz 2021), where coronal stops are also common in affixes.

Except in (d), reduction typically entails an increase in aperture: stops are realized as fricatives, or less commonly sonorants. Relatively speaking, (p t k) are much more likely than (b d g) to reduce to voiceless fricatives. Apart from deleting very commonly, (d) is commonly realized as a tap or semivowel [ɾ ɹ], and less commonly as a fricative [ð] (Schachtenhaufen 2013: 196).

Pharao et al. (2017) report a study testing listeners' ability to perceive words with reduced (g). They show that regardless of whether a categorical or continuous measure of reduction is used, listeners show significantly better perception of non-reduced (g). However, reduced and non-reduced (g) are both correctly perceived at high rates.

2.5.2 Social variation

Characteristics in the realization of /t/ which differ saliently from Modern Standard Danish have often been attributed to varieties spoken in ethnically diverse neighborhoods in Copenhagen ('multiethnolect'; e.g. Quist 2000). Features from these varieties are overall negatively stereotyped, and other speakers of Danish make quite specific characterizations of people on the basis of what is sometimes referred to as the 'street' register (Møller 2009). A key feature that distinguishes these varieties from others is prosody: Copenhagen's multiethnolect is perceived as staccato, and does not maintain a contrast between long and short vowels (Pharao and Hansen 2005; Møller 2009; Hansen and Pharao 2010).

Another key feature is palatalized affricated release of /t/, often transcribed [t^j] (e.g. Maegaard 2007; Madsen 2008, 2013; Lillelund-Holst and Pharao 2014; Hyttel-Sørensen 2017); <tj> is also frequently used in place of <t> in informal written speech parodying this variety (Stæhr 2015). Hyttel-Sørensen (2017) reports that the use of [t^j] is much more common among male speakers of multiethnolect, while Ag (2010) finds that [t^j] is frequently used among adolescent female speakers of multiethnolect. Lillelund-Holst et al. (2019) show that the use of [t^j] is associated with traits like non-Danish ethnicity and "playing tough" in the speech of adolescent girls, regardless of whether they also use the prosodic frame associated with multiethnolect; however, listeners judge the combination of [t^j] and a non-multiethnolect prosodic frame to be unnatural.

Pharao et al. (2014) showed that a fronted variant of [s] with high spectral center of gravity, sometimes transcribed [s^+], signals a feminine and homosexual register in male speakers when combined with a fairly neutral 'modern' Copenhagen prosodic frame. However, [s^+] is also a feature of multiethnolect, and any association with femininity and homosexuality disappears when combined with a multiethnolect prosodic frame. In a follow-up study, Pharao and Maegaard (2017) showed that male speakers using [s^+] are less likely to be evaluated as feminine and homosexual if listeners have already heard that speaker use [t^j]; if the order is swapped, and the speakers are first heard using [s^+], the use of [t^j] makes no difference. Speakers are also more likely to be evaluated as immigrants or 'gangsters' if they are heard using [t^j] before [s^+], but not if the order is swapped.

2.5.3 Regional variation

In the traditional regional varieties of Danish, many of which are now extinct or moribund (see Section 1.3), the stops often show different phonological behavior from Standard Danish, and the patterns of

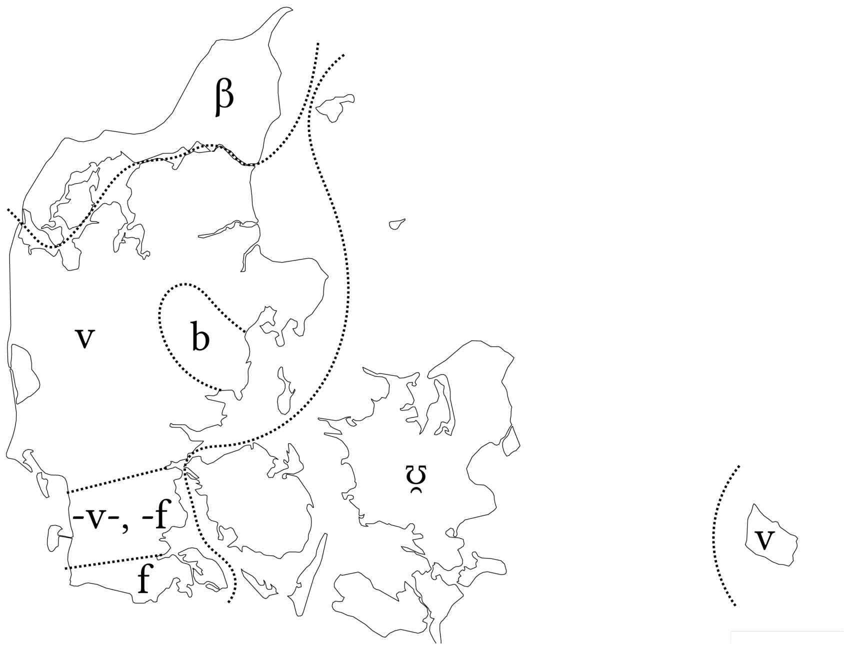


Figure 2.1: *Map showing varying outcomes of /p/ weakening in traditional varieties of Danish. Adapted from Bennike and Kristensen (1898–1912: map 49) in slightly simplified form.*

plosive and spirant weakening discussed in Section 2.4.4 often had different outcomes than in Modern Standard Danish.

The current stop–semivowel alternations in Modern Standard Danish are largely a result of plosive and spirant weakening, which was found to various extents throughout the country. An overview of the precise outcomes throughout the country is given visually in maps 42–60 in Bennike and Kristensen (1898–1912), some of which are reproduced below; see also K4.1–K4.3 in Skautrup et al. (1970–) for more recent maps covering the Jutland peninsula, specifically. Sørensen (2012) also gives a detailed account of sound changes in different varieties.

In earlier stages of Standard Danish, Old Danish weak /p/ lenited into [ʊ̯], although this was later largely rolled back, such that Modern Standard Danish has relatively free variation between [p ~ ʊ̯] (see

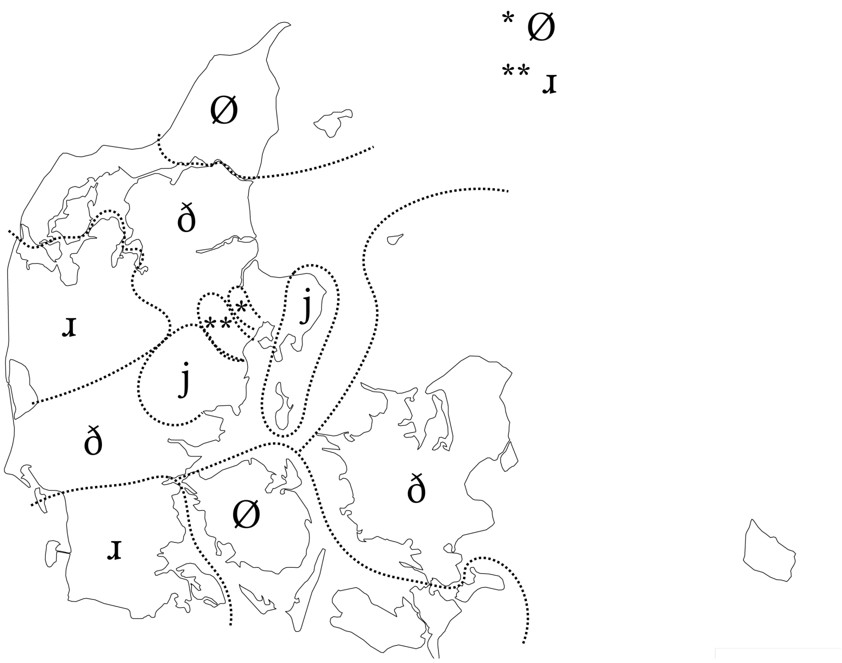


Figure 2.2: *Map showing varying outcomes of /t/ weakening in traditional varieties of Danish. Adapted from Bennike and Kristensen (1898–1912: map 50) in slightly simplified form.*

Section 2.4.4). In different parts of the country, the sound developed consistently into bilabial and labiodental fricatives [β v f], as shown in Figure 2.1.[31] This map is limited to the present-day national borders, but the Danish-speaking area continues somewhat further south in the Jutland peninsula. The original map includes a small area further south, near present-day Viöl in Germany, where a stop [b ~ p] was retained.

Old Danish /t/ weakened into a semivowel [ɣ̞] in Modern Standard Danish.[32] In other parts of the country, it weakened in a variety of different ways, as mapped in Figure 2.2. Old Danish /t/ was generally

[31] The transcriptions in Figure 2.1 are ‘translated’ from Dania transcriptions; note that there is no way to determine whether the sound transcribed as [b] was actually voiced, nor whether the sound transcribed as [v] was actually a fricative.

[32] The development from [ð] to [ɣ̞] is quite recent (Brink and Lund 2018), so [ð] is used in Figure 2.2.

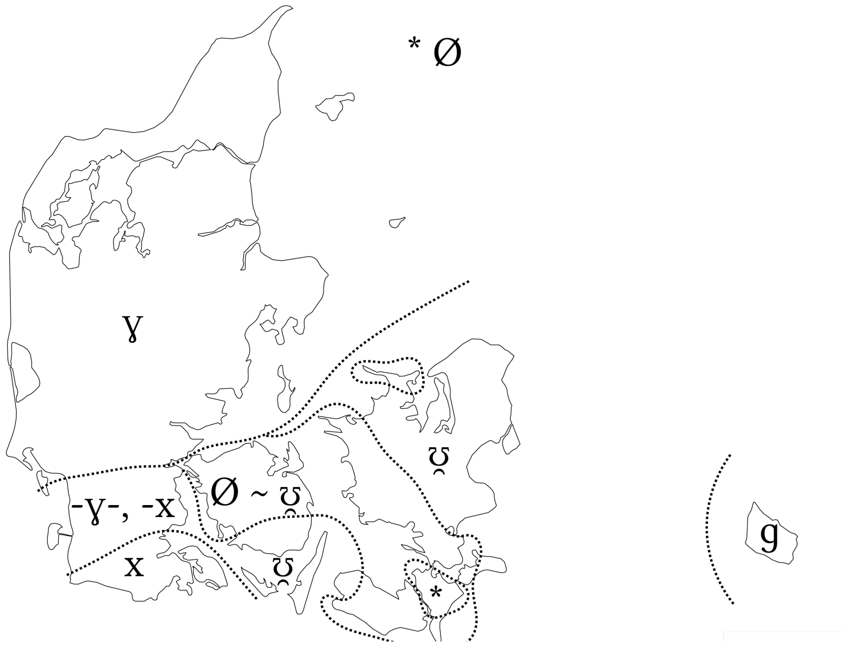


Figure 2.3: *Map showing varying outcomes of /k/ weakening after back vowels in traditional varieties of Danish. Adapted from Bennike and Kristensen (1898–1912: map 51) in slightly simplified form.*

quite unstable. In addition to the soft d (the precise pronunciation of which may vary across regions), it developed into a number of different glides or semivowels [ɪ̯ ʊ̯ ɣ ɹ],[33] and was lost entirely in large parts of the country. Similar to the development of Old Danish /p/, a small area in present-day Germany retained a stop [t] in some contexts.

Just as the development of Old Danish /k/ in Modern Standard Danish varied significantly by phonetic context, there are different regional outcomes by phonetic context. The map in Figure 2.3 shows

[33]It is difficult to say exactly what the transcription [r] in Bennike and Kristensen (1898–1912) refers to, but Veirup (1955) specifically describes it as a "fronted soft d", and classifies it as Dania [ṛ], which Jespersen (1890: 49) describes as very similar to American *r*, i.e. [ɹ].

the different outcomes after back vowels.[34] In some varieties the sound was lost entirely, in others it developed into fricatives [ɣ x] or the semivowel [ʊ̯]. In the area marked [Ø ~ ʊ̯], Old Danish [k] was lost completely after [u ɑ]. The map in Figure 2.4 shows the different outcomes after front vowels. Vocalization after front vowels resulted in [ɪ̯]. In areas marked [Ø ~ ɪ̯], the sound was lost entirely after high vowels [i y]. Old Danish /p k/ showed some class behavior, in that reduction to fricatives had similar outcomes – in areas where /p/ developed into a voiceless fricative [f], /k/ also developed into a voiceless fricative [x], and *vice versa* for voiced fricatives. This may be due to historical differences in closure voicing in these varieties, as I discuss in Section 6.6.

Old Danish medial geminates were generally retained as stops in insular dialects, including Standard Danish (as discussed in 2.2.1), but underwent a number of lenition processes in peninsular Danish. In some varieties, Old Danish [kː] developed into [c] or [ɣ] on a lexical basis, while [tː] developed in different ways, most often into [ɪ̯]. Recall from Section 2.2.1 that the laryngeal contrast in geminates was lost early on in the variety that developed into Standard Danish, due to devoicing of [bː dː gː]. In Jutlandic varieties, however, [bː dː] instead degeminated early on, and underwent the same plosive weakening processes as singletons; this caused some systematic differences between these varieties and Standard Danish. [gː] also degeminated in most varieties, but in some western and southern varieties, it devoiced instead, and merged with [kː], as in Standard Danish (Sørensen 2012).

Stops in simple onset have generally been quite stable, with a few systematic exceptions. Onset /d/ lenited to a voiced fricative in northern parts of the Jutland peninsula, and in a few areas it developed further into a rhotic or was lost completely. /g k/ underwent a number of palatalization processes before front vowels, such that they were realized as [gj kj] in most of the country (although not the variety from which Modern Standard Danish developed); in some areas, this resulted in alveopalatal affricates, fricatives, or glides [tɕ ɕ dʑ j].

[34] As in Figure 2.1, I cannot determine whether the sound transcribed as [g] was actually voiced. The sound transcribed as [x] may have been uvular [χ].

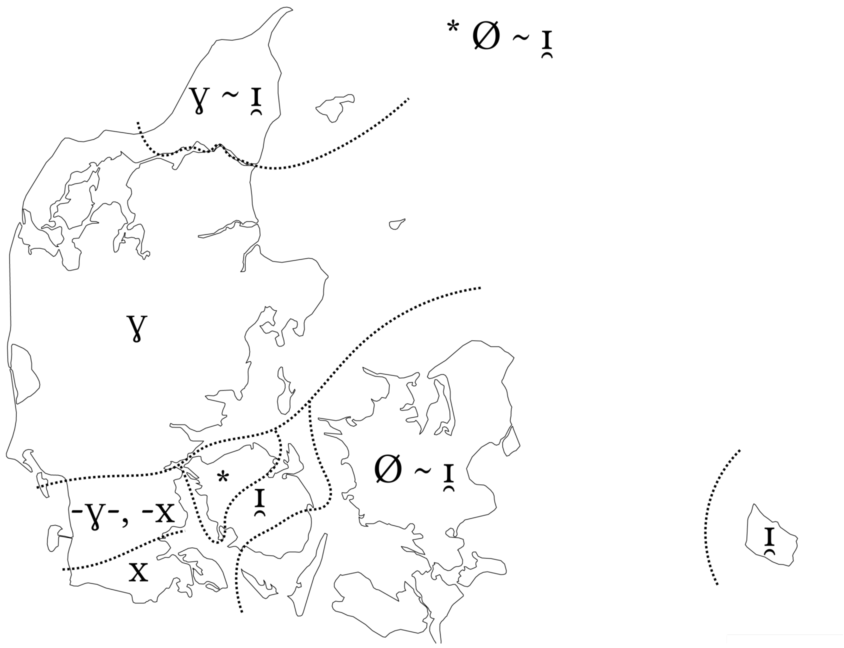


Figure 2.4: *Map showing varying outcomes of /k/ weakening after front vowels in traditional varieties of Danish. Adapted from Bennike and Kristensen (1898–1912: map 52) in slightly simplified form.*

Early descriptions of Danish phonetics sometimes described stød as a full glottal closure (e.g. Sweet 1874: 97, who describes it as "the sound produced in coughing"). In Modern Standard Danish, stød is better described as laryngealization rather than full glottal closure (e.g. Fischer-Jørgensen 1987, 1989), although full glottal closure may well be found in other varieties (e.g. Sørensen 2012). In several regional varieties, particularly in northern Jutland, stød is realized as a so-called 'parasitic plosive' (*klusilspring*), i.e. a dorsal stop, the precise place of articulation of which is determined by the preceding vowel. Parasitic plosives were first described by Nielsen (1947), and have been discussed several times since (Andersen 1955; Søndergaard 1970; Nielsen 1978; Ejskjær 1990, 2006; see Liberman 2006 for an argument that parasitic plosives are not related to stød). Andersen

(1972) argued that parasitic plosives are 'reduction diphthongs', i.e. vowels which undergo a change in the feature [±vocalic]. These vowels originally underwent syllable-finally devoicing (due to stød), and later developed secondary articulatory characteristics like friction or complete closure.[35] Mortensen (2012) analyzes this as a case of high vowel devoicing.

On the topic of stød, recall from Section 2.2.1 that some Western Jutlandic varieties have 'vestjysk stød', i.e. preglottalization of obstruents, as the modern reflex of Old Danish geminates (Skautrup 1928; Ringgaard 1960, 1974; Ejskjær 1967, 1990). In some varieties, preglottalization even co-occurs with parasitic plosives.

Overall, as should be clear from this brief overview, regional differences in the phonological development of stops are quite well-described. The situation is different for phonetic variation. The Danish dialectological tradition has largely been couched within the structuralist tradition of glossematics, which is explicitly not interested in phonetic substance (Hjelmslev 1943).

There has, however, been some interest in the variable realization of /t/. /t/-affrication (as discussed in 2.3.4) is known to be regionally delimited, although there is no consensus in the literature about which varieties lack affrication. Brink and Lund (1975: 353) describe it as missing from "all the country's dialects" (translation mine).[36] A distinctly non-affricated variant of /t/ is known colloquially as 'dry t'. Dry t is mentioned in an encyclopedia article on aspirated stops (Petersen 2009b), which associates the variant with northern Jutland; cp. the corresponding article on affricates, which makes reference to /t/ in Modern Standard Danish (Petersen 2009a). Petersen et al. (2021: 156ff.) describes non-affricated /t/ as a feature of western rather than northern Jutland. Grønnum (2005: 51) describes it as a feature of a "high and formal style" (translation mine); affrication *is* a relatively recent development, so non-affricated /t/ is conservative in a sense, but recall from Section 2.3.4 that affrication was already exceptionless in Copenhagen Danish by the early 1950s (Brink and Lund 1975). When descrip-

[35]Indeed, in recordings of these varieties, seemingly free variation is often heard between parasitic palatal and velar fricatives and parasitic plosives.

[36]This is a dubious claim, especially since Standard Copenhagen Danish is of course also a dialect.

tions of relevant regional varieties mention articulation of /t/, they simply say that it is fronted relative to Standard Danish (Nielsen 1984), or retracted relative to Standard Danish (Espegaard 1995), respectively. I published a small-scale study on the phonetics of /t/ throughout the Jutland peninsula, showing that lack of affrication is found to varying extents throughout the area (Puggaard 2018a; see also Puggaard 2018b). This study also found an interesting interaction between VOT and affrication; more noisy stop releases were found to be more common in varieties that also had relatively long VOT.

I return to regional variation in stops in Chapter 6, where I focus mostly on phonetic differences, which are certainly the most understudied. Using a large legacy corpus called *Dialektsamlingen* 'the dialect collection' (DS 1971–1976; see Section 6.2.2), I shed light on regional differences in VOT and affrication patterns in stops, and discuss the variable patterns of closure voicing in the data. I focus particularly on the varieties of Jutland, which show relatively limited influence from Copenhagen Danish in the corpus recordings.

2.5.4 Summary

In spite of the relatively unified account of stops given before this section, variation actually abounds. The unaspirated stops /b g/ commonly reduce in spontaneous speech, and /d/ is often lost in clusters. A palatalized variety of /t/ is widespread in Copenhagen's multiethnolect, signaling complex social meaning. Other regional varieties than Modern Standard Danish have been subject to sometimes very different patterns of diachronic change, leading in turn to very different stop gradation patterns. Stød sometimes interacts with stops in complex ways in regional varieties. /t/-affrication, which is salient in Standard Copenhagen Danish, is not found throughout the entire speech community, although it remains unclear precisely where it is found and where it is not.

CHAPTER 3

The synchrony and diachrony of stop gradation

3.1 Introduction

In Section 2.4.4, I gave an overview of the traditional account of the positionally determined stop lenition patterns found synchronically in Standard Danish. This process is usually referred to as 'stop gradation' in the literature on the topic (see e.g. Rischel 1970b). The traditional analysis of the process links aspirated stops [pʰ tʰ kʰ] in 'strong' position to unaspirated stops [p t k] in 'weak' position as realizations of the phonemes /p t k/, and voiceless unaspirated stops [p t k] in strong position to semivowels [ʊ̯ ɣ̯ ɪ̯] in weak position as realizations of the phonemes /b d g/ (Uldall 1936; Rischel 1970a; Basbøll 1975, 2005; Grønnum 2005). As mentioned in Section 2.4.1, in the context of Danish, 'weak' position refers to coda or onset before schwa and (in some specific morphemes) [i], and 'strong' position refers to the onset otherwise (e.g. Jakobson et al. 1951). In Section 2.4.4, I reviewed

The research reported in this chapter is collaborative work with Camilla Søballe Horslund and Henrik Jørgensen. Sections 3.2–3.4 are based on published work (Horslund et al. 2021a, 2022) in rewritten form. The account presented here has been developed throughout a number of presentations (Horslund et al. 2020, 2021b; Puggaard-Rode et al. 2021, 2022c).

evidence in favor of the traditional analysis, which is found mainly in the irregular morphology of a small subset of the Danish lexicon. This chapter has two aims: I will argue that the traditional analysis is not suitable for Modern Standard Danish, and that the synchronic state of affairs follows from a series of consecutive sound changes. These changes are individually phonetically well-motivated, but cannot be motivated as a collective synchronic process.

There is an important problematic aspect of the traditional analysis which I have not yet been covered. The semivowels [ʊ̯ ɪ̯], which alternate with voiceless unaspirated stops [p k], also alternate with the approximants [ʋ j] in the same environments. This leads to neutralizations where the underlying form cannot be determined by phonological means. Some underlying forms can be determined with reference to morphology, but many cannot; many of the lexical items which surface with [ʊ̯ ɪ̯] do not participate in any relevant alternations, since these alternations are only found in irregular morphology. Examples include [kʁɑːˀʊ̯] *grav* 'grave' and [pʁɑːˀʊ̯] *brag* 'bang' with surface [ʊ̯], and [hɑɪ̯ˀ] *haj* 'shark' and [kʰʋɛːˀɪ̯] *kvæg* 'cattle' with surface [ɪ̯]. The spelling may help us determine whether the surface sound historically developed from /v j/ or /b ɡ/, but there are no hints in synchronic morphophonology.

Another potential problem is that the traditional analysis assumes phonemes with allophones which do not share any discernible phonological features, or indeed any common phonetic properties. It is a matter of theoretical debate whether or not this is actually a problem. In some 'substance-free' approaches to phonology (e.g. Mielke 2008; Iosad 2017), it is sufficient evidence of phonological categoryhood that sounds show stable alternations; other frameworks require phonological processes to be 'natural' (see e.g. Postal's 1968 Natural Condition). I discuss this problem further in Section 3.3.2 below.

Building on a previous reanalysis of synchronic stop gradation by Ács et al. (2008) couched in Natural Phonology, I suggest below that neither the phonological module nor the morphological module needs to account for the alternations in Modern Standard Danish. Instead, I place the burden in the lexicon; in other words, I propose that the inflections and derivations resulting in the relevant alternations are not morphologically derived – the evidence presented to the language learner in favor of such an analysis is simply insufficient for this – but rather stored separately in the lexicon. The following quote from Linell (1975: 261) about the goals of phonological theory is insightful in this regard:

> "We are not primarily interested in making all possible structural ('significant') generalizations about phonology (…) Instead, we are interested in those generalizations that a speaker–listener can reasonably make."

Consider the structural generalization that [k ~ s] sometimes alternate in English in a limited subset of words like *electric* ~ *electricity* and *opaque* ~ *opacity*; this process is often referred to as 'velar softening'. Chomsky and Halle (1968) attempted to account for velar softening with synchronic phonological rules, but it is not at all clear that modern speakers actually make such a generalization (Postal 1968). If phonology is to be considered a module of grammar, then synchronic phonological processes are limited to those that have a cognitive basis; structural generalizations may be irrelevant if they are not evident to speakers, for example, if they are due to sound change.[1] The traditional analysis of Danish stop gradation captures a structural generalization, but this generalization is arguably not phonological in the cognitive sense. As an alternative to the traditional analysis, I propose a new analysis of phoneme–allophone correspondences in Danish, where [ʊ̯ ɪ̯] in particular are never associated with the stops /b ɡ/, but always with the approximants /v j/, respectively.

I further argue that the structural generalization captured by the traditional analysis is much better accounted for with reference to the historical trajectory of Danish, and to well-understood constraints on articulation and perception. A central tenet of the Evolutionary Phonology framework (e.g. Blevins 2004, 2015) is that synchronic patterns resulting from well-understood sound changes do not need to be accounted for in the synchronic grammar (see also e.g. Ohala 1990a).

[1] Delimiting the influence of history and cognition on phonological processes is a tricky matter, since many processes are compatible with both explanations. Beguš (2022) calls this the 'duplication problem'.

In the second half of this chapter, inspired by Blevins' (2004) typology of sound changes, I will outline the historical trajectory that led to the synchronic stop gradation patterns, and the well-known phonetic pressures that may have led to them; in particular, the pressure against obstruent voicing.

In Section 3.2 below, I briefly recap the overview of stop gradation given in Section 2.4.4) and introduce some further complicating factors. In Section 3.3, I discuss in detail the problems with the traditional analysis of stop gradation, which I propose ultimately makes the analysis unlearnable. I present an alternative analysis in Section 3.4, which largely relies on suppletion, and discuss how this solves the problems discussed in the preceding section. In Section 3.5, inspired by the framework of Evolutionary Phonology, I outline the historical changes that led to the current system, and describe in detail the articulatory and perceptual pressures that likely caused these changes. Finally, in Section 3.6, I summarize the main claims of the chapter.

3.2 The alternations

Many of the alternations relevant to stop gradation were introduced in Section 2.4.4. I will briefly recap the most important points here.

The aspirated stops in strong position [pʰ tʰ kʰ] alternate with unaspirated stops in weak position [p t k], and these are assumed in the traditional analysis to be realizations of the phonemes /p t k/. An example of [k ~ kʰ] alternation is seen in (1), where we see alternation with the unproductive stress-shifting derivational suffix [-ˈænˀt] *-ant*.

(1)	[pʰʁɑkˈtʰik]	*praktik*	'internship'
	[pʰʁɑktʰiˈkʰænˀt]	*praktikant*	'intern'

The voiceless unaspirated stops in weak position [p t k] alternate with either stops, semivowels, or zero in weak position. Strong [p] is usually also realized as [p] in weak position, but shows stylistic alternation with [ʊ̯] in a number of lexical items; these are assumed to derive from the phoneme /b/. An example is shown in (2) with strong past tense declension, as described in Section 2.4.4.

(2)	[ˈskæːʊ sɑ ~ ˈskæːæp sɑ]	*skabe sig*	‘act out’
	[ˈskɑptə sɑ]	*skabte sig*	‘acted out’

Strong [t] alternates with [ɣ̯] in weak position, and with zero in clusters before coronal consonants; these are assumed to derive from the phoneme /d/. An example of both allophones is shown in (3), which shows alternation with the unproductive stress-shifting derivational suffix [-iˈtʰeːˀt] *-itet*, as described in Section 2.4.4.

(3)	[kʁɑˈʋiɣ̯ˀ]	*gravid*	‘pregnant’
	[kʁɑʋitiˈtʰeːˀt]	*graviditet*	‘pregnancy’

In weak position, strong [k] alternates with [ʊ̯] after (historically) back vowels, with [ɪ̯] after (historically) front vowels, with zero after high vowels, and with [k] in consonant clusters; these are assumed to derive from the phoneme /g/. These alternations are found in strong verb declensions and irregular derivational morphology, none of which are productive. I gave an example in Section 2.4.4, example (4) showing all three overt allophones in one lexical item; it is repeated here for ease of reference.

(4)	[ˈpæːɪ]	*bage*	‘to bake’
	[ˈpɑʊ̯ʋæɐ̯k]	*bagværk*	‘baked goods’
	[ˈpɑktə]	*bagte*	‘baked’

Two other (assumed) phonemes participate in similar alternations, namely /v r/. Strong [ʋ] alternates with weak [ʊ̯], and strong [ʁ] alternates with weak [ɐ̯], as in (5), where we see alternations with the stress-shifting verbalizing suffix [-ˈeːˀɐ] *-ere*.[2]

(5)	[kʰuɐ̯ˈsiʊ̯ˀ]	*kursiv*	‘italics’
	[kʰuɐ̯siˈʋeːˀɐ]	*kursivere*	‘italicize’
	[kʰuɐ̯ˀ]	*kur*	‘cure (n.)’
	[kʰuˈʁæːˀɐ]	*kurere*	‘to cure’

[2] The different vowel quality of the suffix in [kʰuˈʁæːˀɐ] *kurere* ‘to cure’ is the result of r-coloring, whereby adjacent /r/ changes the quality of surrounding vowels in largely predictable ways; see Basbøll (1972, 2005: ch. 5) for more details.

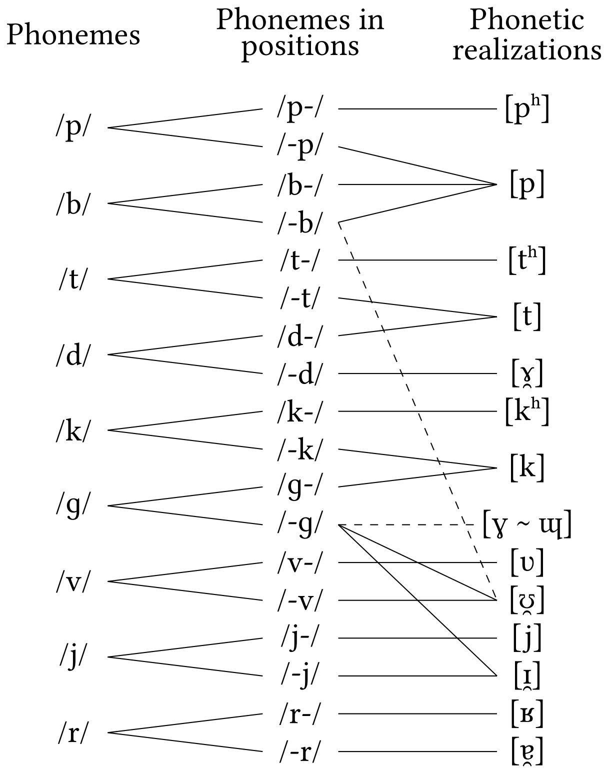


Figure 3.1: *Overview of phonemes and positional allophones according to the traditional analysis.*

This analysis of /r/ works well, and poses no problems for the analysis of stops, so I will not discuss /r/ further below. The analysis of /v/ poses a problem for the traditional account, as [ʋ̯] is now a potential allophone of either /g v/, or possibly /b/.

The proposed phoneme /j/ poses a different problem. /j/ is often assumed to have the strong allophone [j] and the weak allophone [ɪ̯]. To my knowledge, the two never alternate, as no words ending in [ɪ̯] participate in stress-shifting alternations. However, the analysis is straightforward, since the phonetic difference between [j ɪ̯], if there is a consistent difference at all, is minuscule. This is problematic for the traditional analysis, as [ɪ̯] is potentially an allophone of either /g j/.

Figure 3.1 gives an overview of the relevant phonemes proposed in the traditional analysis of Danish consonants and their allophones in strong and weak position, including the stylistic alternations between [p] and [ʊ̯], and the obsolete alternations between [k] and the continuant 'soft g' [ɣ ~ ɰ], which played an important role in Rischel's (1970a) version of the analysis (see Section 2.4.4). For convenience, I use the common notation for referring to onset and coda, e.g. /b-, -b/, as a shorthand for strong and weak position.

Another pattern further complicates the traditional analysis of /g/ in particular. Some morphemes show alternation between final [ɪ̯ ʊ̯] but *not* [k], in accordance with the vowel quality restrictions discussed in Section 2.4.4; this is taken as evidence of underlying /g/. This happens in e.g. nouns with morphologically determined umlaut, as in (6), where the singular has a back vowel and the plural has a front vowel. As with the previously mentioned morphological processes, this is irregular and unproductive.

(6)	[pɔːˀʊ̯]	*bog*	'book'
	[ˈpøːˀɪ̯ɐ]	*bøger*	'books'

A similar process is sometimes observed in compounding processes, where vowel reduction sometimes causes a change in vowel quality in the head. In a few cases, this changes the vowel in a monosyllabic head from a front vowel to a back vowel, specifically [æ ⟶ ɑ], thereby changing the context determining the /g/ allophone. An example can be seen in (7).[3]

(7)	[flæːˀɪ̯]	*flag*	'flag'
	[ˈflɑʊ̯stɑŋˀ]	*flagstang*	'flagpole'

This process is also unproductive, and appears to be waning. As an example, the word [smæːˀɪ̯] *smag* 'taste' used to show this alternation; Brink et al. (1991) describe [ˈsmɑʊ̯løːˀs] as a very conservative pronun-

[3]Recall from Section 2.4.4 that in the context of this rule, [ɑ] is not consistently treated as a back vowel. This is likely due to the recent loss of the three-way distinction between /æ a ɑ/ (Juul et al. 2016).

ciation of *smagløs* 'tasteless', which has arguably been entirely lost in favor of [ˈsmæɪ̯løːˀs].[4]

3.3 Problems with the traditional analysis

In this section, I discuss three major problems with the traditional analysis of stop gradation: 1) It leads to an abundance of intractable neutralizations, 2) some proposed allophones of the same phoneme do not share any discernible phonological features or phonetic characteristics, and 3) the morphophonological evidence in favor of the analysis (as presented above and in Section 2.4.4) is insufficient. I suggest that these deficiencies together render the grammar proposed by the traditional analysis unlearnable.

3.3.1 Intractable neutralizations

We saw above that, following the traditional analysis, a weak allophone [ɪ̯] is phonologically ambiguous, and derives from either /g j/. Similarly, weak [ʊ̯] is ambiguous, and derives from any of /b g v/. This often leads to the neutralization of phonological contrast. These neutralizations can be disambiguated in some morphological contexts, but a large number of lexical items crucially do not participate in any of the morphological alternations that would permit disambiguation. This is particularly problematic for the proposed /g/ phoneme. (8) shows examples of words with weak [ɪ̯] for which the underlying form cannot be determined.

(8)	[ˈkʰæːɪ]	*kage*	'cake'
	[ˈmæːɪ]	*mage*	'mate'
	[ˈlæːɪ]	*lage*	'brine'
	[ˈlɛːɪ]	*læge*	'doctor'
	[ˈʋɛːɪ]	*væge*	'wick'

[4]Note that this particular morpheme *does* show alternation with [k]: the infinitive of the verb form is [ˈsmæːɪ] *smage* 'to taste' and the past tense is [ˈsmɑktə] *smagte* 'tasted'.

(9) shows similar examples with weak [ʊ̯], where the final example covers two homophones that diverge in spelling.

(9)	[lʌʊ̯]	*lov*	'law'
	[ˈkʰʌʊ̯lə]	*kogle*	'cone'
	[hɑʊ̯ˀl]	*hagl*	'hail'
	[ˈkʰʁɑːʊ]	*krage ~ krave*	'crow ~ collar'

This is an analytical problem, assuming that the sound component of lexical entries is composed of phonemes (which is a very common assumption; see e.g. Kenstowicz 1994: 69ff. for discussion of this). If there are no linguistic means of determining the underlying form of a word, it logically follows that speaker–listeners are also unable to arrive at an underlying form.

There are several phonological theories which can help shed a light on this problem with the traditional analysis, and which may help us approach a solution. Below, I discuss how Natural Phonology and Bidirectional Phonetics and Phonology can illuminate the problem.

The framework of Natural Phonology posits that the phonological grammar of a language is ideally maximally uniform, transparent, and biunique (e.g. Stampe 1969; Galéas 2001). A phonological grammar is perfectly uniform when each phoneme has just one allophone, and perfectly transparent when each allophone can represent just one phoneme. In each case, invariance in the relationship between phonemes and allophones is a sign of naturalness. Biuniqueness refers to invariance in both directions: a phoneme has just one allophone, and that allophone can only be derived from one phoneme. In the traditional analysis of Danish, most consonant phonemes have multiple realizations, and several allophones have multiple phonological sources. As such, it scores low on the parameters of both uniformity and transparency (Ács et al. 2008; Ács and Jørgensen 2016). In fact, positional biuniqueness is only observed for six proposed consonant phonemes: /f s m n l h/. Meanwhile, /b d g p t k r v j/ are *not* positionally biunique, as was shown in Figure 3.1.

This issue is not just theoretical, but also poses an acquisition problem. The traditional analysis can be used to derive surface forms from underlying forms, but not to arrive at underlying forms if given only surface forms. In other words, the grammar only works top–down, as Basbøll (2015) also admits. For learners, however, learning a grammar proceeds from the surface forms: they only have a bottom–up approach available.

Eliasson (1997) identifies this discrepancy as a general trend in 20th century theoretical linguistics, and argues that it renders many analyses cognitively implausible. Consider American English tapping; there is a well-defined rule whereby /d t/ both reduce to [ɾ] medially in unstressed syllables. Only the output of this rule [ɾ] is available to language learners. Unless there are alternations which can serve to disambiguate /d t/, the learner has no way of recovering the intended phoneme (Smith 1991). Eliasson (1992, 1997) discusses an example from Swedish, where /h/ was historically lost pre-consonantally in words like [ˈjɛlpːa] *hjälpa* 'to help'. The (archaic) strong past tense of the word is [halp] *halp* 'helped', with /h/ retained but no /j/.[5] Eliasson argues that the historical /h/ is not recoverable in the infinitive *hjälpa*, because the [j ~ h] alternation is lexically isolated. Instead, he argues, although *hjälpa* and *halp* are etymologically related, they must be stored as suppletive allomorphs by speakers. I use the term *suppletive* in this sense extensively below.

Work within the Bidirectional Phonetics and Phonology (BiPhon) model can shed further light on the acquisition problem. A key assumption of the BiPhon model is that the same grammar is used for production and perception, allowing the model to account for both acquisition (Boersma 2011) and mechanisms of language change (Boersma and Hamann 2008; Hamann 2009). The model was originally implemented in Optimality Theory, but recent research implements BiPhon using artifical neural networks (e.g. Boersma et al. 2020). The BiPhon model assumes that learners encounter pairs of *phonetic form* and *semantic content*, and must construct the intermediate levels of *surface form*, *underlying form*, and *morphology*. BiPhon has previously been used to account for a phonological phenomenon with many parallels to Danish consonant gradation, namely French liaison (Boersma and Leussen 2017).

[5]This form has now been replaced by the regular [ˈjɛlptə] *hjälpte* 'helped'.

In the 'traditional' generative account of French liaison (e.g. Schane 1968; Selkirk 1972), a word like *bon* 'good' is assumed to be underlyingly represented as /bɔn/. *Bon* takes the masculine and feminine agreement suffixes /+Ø/ and /+ə/, respectively. Underlying /n/ surfaces as nasalization on the preceding vowel before an underlying consonant, and as [n] before an underlying vowel. Schwa is deleted late in the derivation, and never surfaces; its presence in the underlying form, however, ensures that the feminine form surfaces with [n]. This results in the surface patterns in (10) (from Boersma and Leussen 2017: 352ff.).

(10)	[bɔ̃.ma.ʁi]	/bɔn+Ø#maʁi/	*bon*$_{M}$ *mari*$_{M}$	'husband'
	[bɔn.vwa.tyʁ]	/bɔn+ə#vwatyʁ/	*bonne*$_{F}$ *voiture*$_{F}$	'car'
	[bɔ.nak.tœʁ]	/bɔn+Ø#aktœʁ/	*bon*$_{M}$ *acteur*$_{M}$	'actor'

Boersma and Leussen (ibid.) ran a computer simulation of the acquisition of liaison by so-called 'virtual learners'. They find that virtual learners generally resist establishing a single underlying form for the root. Most virtual learners instead establish suppletive allomorphs. Instead of linking [bɔn] and [bɔ̃] to the same underlying root /bɔn/, they link two underlying forms /bɔn/ and /bɔ̃/ to the same semantic content – likely because the traditional analysis is excessively abstract.[6] This is in line with the results of a production experiment by Sampson (2001), which shows that the pattern in (10) is not particularly productive, and speakers resist extending liaison beyond a small set of frequently occurring lexical items. A similar computer simulation of the traditional account of Danish consonant gradation would likely yield similar results: learners would reject the analysis as excessively abstract, and instead establish suppletive allomorphs. This is of course an empirical question, and one that will hopefully be answered with future research within the BiPhon framework.

[6] For further discussion of this, see Kiparsky (1968/1982) and Selkirk and Vergnaud (1973).

3.3.2 Lack of similarities between allophones

As argued in Section 3.5 below, Modern Standard Danish consonant gradation is the end result of a series of sound changes that are individually well-motivated. From a synchronic point of view, the result is a set of strong and weak surface allophones that in some cases no longer share any phonological features or phonetic properties. This arguably inhibits acquisition.

Any phonological theory that assumes a link between phonology and phonetics will have a hard time assigning a voiceless stop and a semivowel to the same phoneme. Basbøll (2005: 109ff.) represents phonemes and position-specific allophones of Danish using binary distinctive features, which he argues should be grounded in phonetics. The discussion here will mostly rely on Basbøll's rather unconventional set of features. He represents /b d g/ as [+stop, -spread glottis] with the place features [+labial], [+alveolar], and [+velar], respectively. In order to derive the weak realizations, he proposes the rule in (11).

(11) [+stop, -spread glottis] ⟶ [+vocoid] / weak position

The rule in (11) is problematic for at least two reasons. 1) Oral stops and vocoids are essentially maximally different in terms of degree of constriction, sonority sequencing, and (in this case) even voicing. 2) The place features for the weak allophones are not predictable from the strong allophones. In Basbøll's framework, the process of /b/ ⟶ [ʊ̯] entails a change from [+labial] to [+labial, +velar], with no way to explain the addition of [+velar]. Similarly, /g/ ⟶ [ʊ̯] entails the addition of an unexplainable [+labial] feature. Most problematically, /g/ ⟶ [ɪ̯], in terms of Basbøll's distinctive features, translates into [+stop, -spread glottis, +velar] ⟶ [+vocoid, +palatal]. These two representations do not share a single feature.

This issue may be exacerbated by Basbøll's unconventional feature set. Historically, palatal and velar consonants have commonly been assumed to share either the feature specification [+dorsal] (e.g. Chomsky and Halle 1968), or the node DORSAL (e.g. Sagey 1986). This has been the topic of much discussion, though (see e.g. Hall 1997). According to Hall (2007), there is now broad consensus that palatals are CORONAL. Hume (1992) argues that a logical consequence of this is that high front vowels like [i e] (and hence also semivowels like [i̯]) are also CORONAL.

Basbøll's formal apparatus is similar to that of Chomsky and Halle (1968), in that distinctive features are phonetically grounded, but the rule system in itself is not constrained by phonetics. In other words, any operation is allowed. Chomsky and Halle (ibid.: 400) explicitly recognize this as a problem with their framework, and attempts to solve this problem has guided much of phonological theory since, as pointed out by Reiss (2018: 426–427):

> "This call for a theory of markedness in generative phonology is perhaps responsible for inspiring most work in phonology for the last five decades, from the universal processes of Natural Phonology to the universal markedness constraints of Optimality Theory."

In Optimality Theory, where markedness is one of the key guiding principles, probably no ranking of constraints could account for an output realization which does not share a single property with the input.

Before the completion of two recent sound changes, the loss of voicing in /b d g/ (see Section 3.5.3) and the loss of the the soft g, the problem would have been much less severe. If the strong allophones of /b d g/ had been voiced, all allophones would at least share the feature [+voice].[7] Similarly, as we saw in Section 2.4.4, the soft g [ɣ ~ ɰ̟] played a central role in Rischel's (1970a) analysis of /g/. The loss of [ɰ̟] in Modern Standard Danish is detrimental to the traditional analysis, as there is no longer an intermediate step between [k] and [i̯ ʊ̯].

If strong /g/ was voiced, and [ɰ̟] remained in the system, Rischel's analysis would describe a perfectly reasonable synchronic gradation process, with the steps [g] → [ɰ̟] → [i̯ ʊ̯]. Given that all stops in Modern Standard Danish are voiceless, the differences between [k] and [ɰ̟] are quite significant; with the subsequent loss of [ɰ̟], the proposed gradation process from [k] → [i̯ ʊ̯] simply skips too many stages to be plausible as a synchronic process. Consider Hayes' (2009: 54–55) Criterion of Phonetic Similarity:

[7] Recall from Section 2.4.3.1 that [+voice] entails closure voicing for Basbøll, although many phonologists conceptualize the feature in a more abstract way.

Table 3.1: *Feature value changes in the traditional analysis, using Basbøll's (2005) feature set.*

Strong realization		**Weak realization**		**Feature**	**Unchanged**
IPA	**Features**	**IPA**	**Features**	**changes**	**features**
[pʰ tʰ kʰ]	+stop, +spr.gl., (place)	[p t k]	+stop, (place)	spr.gl.	stop
[p]	+stop, +labial	[ʊ̯]	+vocoid, +labial, +velar	stop, vocoid, velar	labial
[t]	+stop, +alveolar	[ɣ̯]	+vocoid, +alveolar, +velar	stop, vocoid, velar	alveolar
[k]	+stop, +velar	[ɪ̯]	+vocoid, +palatal	stop, vocoid, velar, palatal	–
[k]	+stop, +velar	[ʊ̯]	+vocoid, +labial, +velar	stop, vocoid, labial	velar
[ʋ]	+vocoid, +approx., +labial	[ʊ̯]	+vocoid, +labial, +velar	approx., velar	vocoid, labial
[j]	+vocoid, +approx., +palatal	[ɪ̯]	+vocoid, +palatal	approx.	vocoid, palatal

> "It is possible during language change that two allophones drift too far apart to count anymore as variants of the same basic linguistic unit."

In accordance with this principle, it is not plausible to propose that voiceless stops and semivowels are realizations of the same phoneme solely because they *used to* share phonetic content.

Table 3.1 shows a formalization of the phonological distance between the weak and strong realizations for the Danish consonants involved in gradation (excluding /r/) in terms of Basbøll's features. Basbøll conceives of distinctive features as strictly binary, and all his features are defined such that only the positive-valued pole is required to be phonetically homogeneous. For example, phonemes represented as [+alveolar] constitute a class of sounds with an alveolar place of articulation, but phonemes represented as [-alveolar] does not necessarily constitute a particular class of sounds. He assumes that all phonemes are specified as + or - for all distinctive features. As such, [p] is technically specified as [+stop, +labial, -alveolar, -palatal, -velar, -pharyngeal, -fricative, -approximant, -vocoid, -spread glottis], etc. Some features logically imply others: [+vocoid] logically implies [+sonorant], which in turn logically implies [+voiced] (see e.g. Basbøll 1994). Similar implicational relationships hold for vocalic place features. As is usually done by Basbøll, only the informative non-redundant positive-valued features are included in Table 3.1.

It is worth briefly returning to the point that the lack of phonetic and phonological similarities between allophones is not a problem for all theories of phonology. Chomsky and Halle's (1968) features were phonetically grounded, but they famously allowed for rules of the type A ⟶ B / C, where any sound can plausibly be replaced with any other in any possible environment, and processes with phonetic grounding are not required nor favored. In other words, the grammar did not favor natural rules over 'crazy rules' (Bach and Harms 1972).

Much work in phonology since has been preoccupied with constraining the grammar's generative capacity; for a few examples, consider Postal's (1968) Natural Condition, and the markedness constraints of Optimality Theory (Prince and Smolensky 1993/2004). Hale and Reiss (2000, 2008), however, in their 'substance-free' approach to phonology, explicitly cite unconstrained generative capacity as an advantage, and argue that accounting for naturalness falls outside the scope of phonology. In such an approach, an operation like /g/ ⟶ [ɹ̯] is fine, because there is no requirement that rules be natural. Others have argued that phonological features themselves are substance-free and emergent, and that language learners do not construct features on the basis of phonetic similarity, but rather on the basis of evidence such as contrast (e.g. Mielke 2008; Dresher 2009; Iosad 2017). In such an approach, [k] and [ɹ̯] may well be allophones of the same phoneme, as long as there is phonological evidence to group them together. It is outside the scope of this chapter to argue against these positions, and

I will assume below the (arguably mainstream) position that phonetic substance *does* play some role in phonological representations.

As pointed out in BiPhon (see Section 3.3.1), language learners initially only have access to pairs of phonetic form and meaning. In order to establish a phoneme inventory, learners need evidence of which allophones and phonemes belong together. It should be clear from this section that phonetic evidence is scarce or completely lacking for several of the phoneme–allophone pairings proposed in the traditional analysis (see Figure 3.1). As shown in the next section, this is also the case for morphophonological evidence.

3.3.3 Insufficient morphophonological evidence

Given the large number of intractable neutralizations following from the traditional analysis, the burden of phonetic and morphophonological proof in favor of the analysis is especially heavy. We saw in the previous section that there is little to no phonetic evidence supporting the traditional analysis. As I will show in this section, the morphophonological evidence in support of the analysis is also rather weak, and found only in a small subset of the vocabulary. Evidence comes from strong verb declinations of the form [ˈkʰɔːʊ] *koge* 'to boil' ~ [ˈkʰʌktə] *kogte* 'boiled', and from derivational morphology in Latinate words, including alternations like [fonoˈloːˀ] *fonolog* 'phonologist' ~ [fonoloˈkiːˀ] *fonologi* 'phonology', where the former ostensibly has zero-realized /g/. There are two important issues with this line of evidence: 1) The relevant morphological alternations are all irregular and unproductive, and 2) a large portion of the alternations are (presumably) acquired quite late, i.e. at a point in acquisition when the core phonological system should already be in place.

It is difficult to gauge the exact timeline of phonological acquisition for Danish children from the literature. Heger (1979) summarizes a repetition study which shows that 75% of all Danish children have acquired all consonantal allophones by the age of 5½ years. Clausen and Fox-Boyer (2017) show that the vast majority of Danish children are already able to produce all consonants with the exception of [ɕ] between 2 and 3 years of age. Both of these studies primarily target phonetic knowledge rather than phonological knowledge; Clausen and

Fox-Boyer explicitly count productions as correct even if they are produced in the wrong position. The findings, however, are corroborated by studies from related languages such as (British) English, where Dodd et al. (2003) report that the vast majority of children have acquired all consonants except [ɹ θ ð] at the age of 5½ years. In a study of the acquisition of the strong [-tə] *-te* past tense declination, however, Bleses et al. (2000) find that 8 year old children still make errors in approximately half of all productions. This suggests that this particular morphological pattern is acquired after the phonological system is largely in place.

Around 85% of Danish verbs take the regular, productive past tense suffix [-əɤ] *-ede*, and only 10–15% take [-tə] *-te* (Jacobsen 2019). As such, [-əɤ] *-ede* has high type frequency; [-tə] *-te* has low type frequency, but most strong verbs have rather high token frequency. Table 3.2 shows frequencies of the infinitive and past tense forms of verbs that take the [-tə] *-te* past tense resulting in one of the relevant alternations. These numbers come from two corpora: LANCHART (Language Change in Real Time; Gregersen 2009; Gregersen et al. 2014), which is a huge spoken corpus consisting of almost 2,000 sociolinguistic interviews, 600 of which are transcribed. daTenTen17 by Sketch Engine (see e.g. Kilgarriff et al. 2014) is a very large written corpus collected by a web crawler. This corpus consists of roughly 2 billion tokens from relatively recent and stylistically varied texts. Frequencies from the LANCHART corpus come from a word list compiled by Pharao (2009). It is well-established that inflected forms with high token frequency tend to be treated as unanalyzed chunks during language acquisition, while patterns with high type frequency are treated as productive (e.g. Ambridge et al. 2015). Given the low type frequency of the [-tə] *-te* suffix combined with the relatively high token frequency of inflected forms, it is an unlikely source of productive patterns during acquisition.

Latinate words showing relevant alternations have low type frequency *and* low token frequency,[8] and are often technical terms which are likely acquired late (if they are acquired at all). As with the [-tə] *-te* suffix, the relevant derivational affixes are not productive.

[8]Frequencies of selected words are given in Horslund et al. (2022: 95–96).

Table 3.2: *Frequencies of infinitive and past tense forms of verbs with* [-tə] -te *past tense declination showing the relevant consonant alternations. Occurrences per one million in the LANCHART and daTenTen17 corpora.*

Alternation	**Verb** (infinitive)		**LANCHART**	**daTenTen17**
[p/ʊ̯] ~ [p]	*købe*	'buy'	162.4–115.8	185.8–48.9
	slæbe	'drag'	12.1–4.4	5–1.6
	råbe	'shout'	15.4–19.8	9.5–10.8
	skabe	'create'	14.8–2.7	209.9–23.8
	tabe	'drop'	10.7–15.4	22–26.9
[ɣ̯] ~ Ø	*svede*	'sweat'	1.3–1	1.8–0.7
	lede	'lead'	10.4–2	32.3–7.5
	møde	'meet'	77.2–112	145.6–40.2
	føde	'give birth'	5.7–3	15.8–7.3
	støde	'bump'	2.3–5	6.4–7.8
	bløde	'bleed'	5.7–0.7	33–1.1
	sprede	'spread'	1.7–1	12.1–10.2
	rede	'comb'	10.4–1	13.1–0.2
	træde	'step'	5–14.4	19.8–19.4
	klæde	'dress'	4–1	9.8–2.4
[ɪ̯] ~ [k]	*bage*	'bake'	5.7–4	9.3–3.9
	smage	'taste'	53–71.5	54.2–27
	stege	'fry'	1–1.3	6.1–3.1
[ʊ̯] ~ [k]	*koge*	'boil'	7.4–3.7	9.2–5.2
Ø ~ [k]	*søge*	'seek'	52.7–69.5	78.9–24.8
	sluge	'swallow'	1–0.3	3.9–1.5
	bruge	'use'	286.3–66.4	398.6–72.9

There are no studies of the acquisition of Latinate words in Danish, but research on English shows that knowledge of comparable loanwords is highly socially stratified in 12–15 year old native speakers (Corson 1984). Research on the acquisition of Latinate derivational morphology by speakers of English sheds some light on how these loanwoards may affect the phonological grammar. Latinate words in English are subject to a process of trisyllabic shortening, whereby a long vowel shortens (with concomitant changes in vowel quality) in derivations with three or more syllables, as in (12).

(12)	[sɪˈɹiːn]	*serene*	[sɪˈɹɛnɪti]	*serenity*
	[dɪˈvaɪn]	*divine*	[dɪˈvɪnɪti]	*divinity*
	[pʰɹəˈfeɪn]	*profane*	[pʰɹəˈfænɪti]	*profanity*

Chomsky and Halle (1968) assume that words such as those in (12) share an underlying root, and that the changes in vowel quality are derived by rule. However, several experiments have shown that adult speakers usually do not treat these phonological process as productive (Ohala 1974; Steinberg and Krohn 1975; Jaeger 1984).

Some of the alternations relevant for the traditional analysis are found only in the derivational morphology of Latinate words, as these are the only stress shifting affixes in Danish. As mentioned in Section 2.4.4, Pharao (2004) investigated the generalizability of these alternations in a suffixation experiment with nonsense words. 12 out of 30 participants in his study generalized the alternations. Interpreting these results is not straightforward. The study may be taken as evidence that some speakers organize their phonology as predicted by the traditional analysis, but it may just as well be evidence of morphological schemas that are limited to a subset of the lexicon (see e.g. Bybee and Slobin 1982; Bybee 1985, 2001). Returning briefly to the proposed velar softening rule in English which results in [k ~ s] alternation (see Section 3.1), Pierrehumbert (2006) tested its productivity, and found that speakers generally applied velar softening productively to nonce words, but only if they had other Latinate characteristics and combined with the *-ity* suffix; i.e. /k+ɪti/ ⟶ [sɪti]. However, very few speakers applied velar softening productively in a backformation task; if asked to find the root for a derived Latinate word ending in [sɪti], most speakers assumed the root ended in /s/ rather than /k/.

The Latinate derivations in Danish mostly provide evidence for alternations between aspirated and unaspirated stops, as well as evidence for [k] ~ Ø alternations in a.o. a number of words denoting scientific professions and their associated fields, which alternate between [-ˈloːˀ] *-log* ~ [-loˈkiːˀ] *-logi*, as in (13); these were also discussed in Section 2.4.4, where I suggested that the [-ˈiːˀ] *-i* suffix may have been reanalyzed by speakers as [-ˈkiːˀ] *-gi*, since /g/ never surfaces in the bare roots.

(13)	[tʰeːoˈloːˀ]	*teolog*	‘theologist’
	[tʰeːoloˈkiːˀ]	*teologi*	‘theology’
	[soɕoˈloːˀ]	*sociolog*	‘sociologist’
	[soɕoloˈkiːˀ]	*sociologi*	‘sociology’

Evidence for stop–semivowel alternations in these words is limited to small number of items such as those in (14), and a few similar patterns mentioned in Section 2.4.4.

(14)	[ˈɑpɤ]	*abbed*	‘abbot’
	[ɑpəˈtisə]	*abbedisse*	‘abbess’
	[pʰæɐ̯ˈfiɣ̯ˀ]	*perfid*	‘perfid’
	[pʰæɐ̯fitiˈtʰeːˀt]	*perfiditet*	‘perfidy’

Words like those in (14) are unsurprisingly very infrequent (Horslund et al. 2022: 95–96), and presumably virtually non-existent in child-directed speech. It seems very implausible that this type of vocabulary plays a major role for children in establishing phonemes.

3.4 An alternative analysis

Inspired by the Natural Phonology notions of uniformity, transparency, and biuniqueness, discussed in Section 3.3.1 above, Ács and Jørgensen (2016) proposed a different analysis of the Danish consonant phonemes. Their analysis, shown in Figure 3.2, yields a much higher number of phonemes than the traditional analysis, but it is also maximally biunique in that all phonemes have just one realization. As a result, an unaspirated stop [p] is always analyzed as an allophone of /b/, as opposed to the traditional analysis, where [p] in weak position is analyzed as an allophone of /p/. It also means that most consonant phonemes in Ács and Jørgensen’s (2016) analysis are defectively distributed.

The resulting analysis is very different from the traditional analysis as envisioned by Rischel (1970a) and Grønnum (2005), but similar to Basbøll’s (2005) organization of allophones and phonemes, as discussed in Section 2.4.4. Basbøll, however, has a separate layer of morphophonemes which is not rule-governed, where all roots have one unique

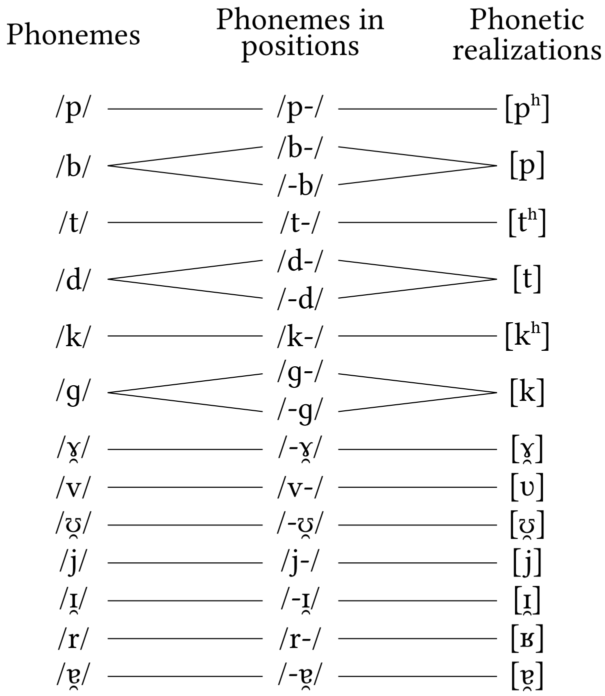


Figure 3.2: *Overview of phonemes and positional allophones according to the maximally biunique analysis of Ács and Jørgensen (2016).*

representation; Ács and Jørgensen instead assume that the burden of accounting for the alternations should be placed squarely in the morphology.

I propose an alternative analyses which differs from Ács and Jørgensen's in two crucial ways: 1) Instead of accounting for the relevant alternations in the morphological domain, they should be accounted for in the lexical domain, where the relevant words are stored with suppletive roots. This is based on the assumption that irregular, unproductive morphology must be rote learned for each lexical item regardless. If they are analyzed as suppletive (in the synchronic, cognitive sense), we can also assume that they do not affect the phonological grammar. 2) A more economical analysis can arguably be achieved by retaining some assumptions from the tradi-

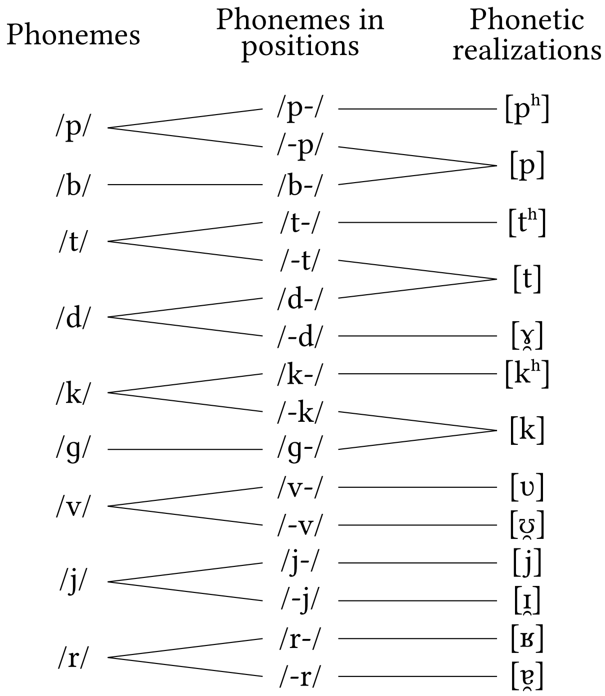


Figure 3.3: *Overview of phonemes and positional allophones according to the alternative analysis suggested here.*

tional analysis, namely by keeping some connections between strong and weak allophones if they are sufficiently well-motivated. I assume that the phonemes /p t k/ are realized as aspirated stops [pʰ tʰ kʰ] in strong position and unaspirated stops [p t k] in weak position. This is well-motivated; a process of final neutralization of laryngeal contrast is very common (Lombardi 1991, 1999; Kehrein and Golston 2004; Blevins 2006; Iverson and Salmons 2006, 2011). In this analysis, only /b g/ are defectively distributed, found only in strong position as [p k], respectively.[9] [ɪ̯ ʋ̯] are always associated with the phonemes whose strong realization they match most closely, i.e. /j v/, respectively.

[9]That is, of the consonants affected by gradation, only /b g/ are defectively distributed; /h ŋ/ would also be considered defectively distributed.

The analysis is shown in Figure 3.3. I argue below that the alternative analysis is more cognitively plausible than the traditional analysis, since 1) it does not pose a neutralization problem, 2) all allophones of the same phoneme share phonetic properties and/or phonological features, and 3) the analysis does not make reference to irregular and unproductive morphology.

3.4.1 No neutralizations

While the analysis proposed in Figure 3.3 is not entirely biunique, it *is* positionally biunique: the underlying representation of an allophone can always be determined with reference to its prosodic position, i.e. strong or weak. This results in a more natural analysis. As discussed in Section 3.3.1 above, the majority of phonemes in the traditional analysis have multiple realizations, and /g/ in particular has several weak realizations. In this alternative analysis, the majority of phonemes still have multiple realizations, but never more than one strong realization or more than one weak realization. Since [ʊ̯ ɪ̯] are always considered realizations of /v j/ in the alternative analysis, /g/ no longer poses a neutralization problem.

In Section 3.3.1, I proposed that the many intractable neutralizations resulting from the traditional analysis result in an acquisition problem. This problem is solved with the alternative analysis, since the underlying forms of all resulting neutralizations can be determined with reference to prosodic structure. Such a system should not pose an acquisition problem; research shows that children by the age of 7½ months are already able to identify word boundaries in fluent speech (Jusczyk and Aslin 1995), suggesting that they are aware of positional information well before they start productively acquiring segmental information.

3.4.2 Shared phonetic and phonological properties

Allophones of the proposed phonemes in the alternative analysis all share phonetic properties, as summarized in Table 3.3. Relying on the distinctive features proposed by Basbøll (2005), they also all share at least one, and generally multiple, phonological features. As discussed

Table 3.3: *Shared phonetic properties between allophones of the same phoneme in the alternative analysis.*

Phoneme	Strong	Weak	Shared phonetic properties	Gradation process
/p t k/	[p^h t^h k^h]	[p t k]	Place and manner unchanged	Deaspiration
/d/	[t]	[ɣ̯]	Alveolar oral consonants	Vocalization
/v/	[ʋ]	[ʊ̯]	Labial voiced oral continuants	Vocalization
/j/	[j]	[ɪ̯]	Palatal voiced oral continuants	Vocalization

in Section 3.3.2 above, a further problem with the traditional analysis is that place features for the weak allophones are always partially unpredictable.

The process of /b/ ⟶ [ʊ̯] involves the addition of [+velar]; the process of /g/ ⟶ [ʊ̯] involves the addition of [+labial], and the process of /g/ ⟶ [ɪ̯] involves a change from [+velar] to [+palatal].[10] All of these changes require explanations which are not given in the traditional analysis. The alternative analysis mostly gets around this issue, although it retains two problematic changes in place features. 1) The process of /v/ ⟶ [ʊ̯] still involves the addition of [+velar]. 2) This is likely also the case for the process of /d/ ⟶ [ɣ̯]. As discussed a.o. in Section 2.2 above, much remains unknown about the exact articulation of [ɣ̯] (Brotherton and Block 2020); it appears to have a coronal component and a dorsal component (Siem 2019), although the exact nature of either component is unclear. A representation containing [+alveolar] and [+velar] seems reasonable based on our existing knowledge, but this is subject to change with further artic-

[10]In a wholly different context, Basbøll (2005: 138ff.) discusses the feature [grave] as encompassing both labial and velar consonants, following Jakobson et al. (1951). He concludes that it plays no role in distinguishing Danish phonemes, although it does serve to explain the distribution of short /a æ/ (see Basbøll 1972).

ulatory research. The rule needed to account for /d/ ⟶ [ɣ̯], disregarding place features, would be [+stop, -spread glottis] ⟶ [+vocoid]; this constitutes a major change in terms of degree of constriction, sonority, and voicing. If further research should show that [ɣ̯] does not in fact have an alveolar component, the analysis would become difficult to maintain.

Basbøll's unusual set of distinctive features exacerbates the problem of missing shared features, as I also touched upon in Section 3.3.2. Most modern approaches to distinctive features (e.g. Chomsky and Halle 1968; Sagey 1986; Broe 1992) do not have a feature [velar], but rather a feature or node [dorsal] ~ DORSAL. This feature or node is shared by velar consonants and vowels.[11] As such, the vocalization of /d v j/ in weak position necessarily implies the addition of DORSAL. The daughter nodes of DORSAL – specifically the value of [back] – can be thought of as either underspecified or inherited. [ɣ̯] is centralized (Juul et al. 2016; see Section 1.4), indicating that [back] is underspecified. The backness of [ʊ̯] may simply be enhancement of the pre-existing LABIAL feature from /v/, since backness and labiality have a similar influence on *F*2 (Flemming 1995: 73); this could also be considered underspecification.

These explanations also hold for the process of /b/ ⟶ [ʊ̯] in the traditional analysis, but not for /g/ ⟶ [ɪ̯ ʊ̯]. As above, the addition of a labial gesture in [ʊ̯] may be considered enhancement of the [+back] feature, but the [+back] feature itself remains unexplained. The varying values for [back] in [ɪ̯ ʊ̯] can neither be considered underspecified nor inherited from /g/. Proponents of the traditional analysis might argue that /g/ is underspecified for [back], and that weak allophones inherit the value for [back] from the preceding vowel. This is a good historical account, but recall changes in the quality of preceding vowels have not always led to corresponding changes in /g/-allophones, such that both [ɪ̯ ʊ̯] are occasionally found after [ɑ] (see Section 2.4.4). This strongly

[11] Although see Steriade (1987), who assumes that DORSAL accounts *only* for vocalic place features, while a separate VELAR node accounts for the consonantal place feature(s). Also recall from Section 3.3.2 that it is a matter of debate whether palatals and high front vowels should be considered DORSAL or CORONAL; this is inconsequential for the alternative analysis, as the process /j/ ⟶ [ɪ̯] does not entail a change in place features.

suggests that the 'allophone selection' is calcified from a previous stage of the language, and not an active synchronic process. As such, using a more conventional set of distinctive features largely solves the problem of missing shared features for the alternative analysis, but *not* for the traditional analysis.

3.4.3 Alternations stored as suppletive allomorphs

The alternative analysis suggests that the alternations found in strong verb conjugations and Latinate derivations are stored as suppletive allomorphs, and do not play a role in phonology. This is in line with the results of Boersma and Leussen's (2017) computer simulation of the acquisition of French liaison (see Section 3.3.1), which provides a comparable example to the strong verb declinations. It is also generally consistent with the results of Pharao's (2004) psycholinguistic experiment which showed that most speakers of Danish do not extend the Latinate derivations to nonce vocabulary. In discussing the results of a similar experiment of trisyllabic shortening in English, Jaeger (1984) suggests that participants who *do* extend the pattern to new vocabulary do so mostly on the basis of orthographic knowledge. This may also be the case for the speakers who extended the gradation patterns in Pharao's study.

This leaves the issue of words showing stylistic alternations between [p ~ ʊ̯]. A few examples are given in (15).

(15)	[ˈkʰøːøp ~ ˈkʰøːʊ]	*købe*	'buy'
	[ˈkʰøptə]	*købte*	'bought'
	[ˈslɛːɛp ~ ˈslɛːʊ]	*slæbe*	'drag'
	[ˈslɛptə]	*slæbte*	'dragged'
	[ˈʁɔːɔp ~ ˈʁɔːʊ]	*råbe*	'shout'
	[ˈʁʌptə]	*råbte*	'shouted'

Some regional varieties of Danish have [ʊ̯] throughout all derivations of these verbs; I assume that the words have underlying /v/ for these speakers of these varieties. Some speakers seemingly have no active alternations between [p] and [ʊ̯], but there is no indication that [ʊ̯] causes comprehension problems for such speakers. This suggests that speakers have two suppletive allomorphs for the relevant roots – one ending in /p/, and one ending in /v/. Speakers differ in which words, if any, can take the /v/ root in production. This is similar to other words showing stylistic alternation in pronunciation, such as the Danish noun *tunnel* 'tunnel', which can be pronounced either [ˈtʰɔnˀl̥ ~ tʰoˈnɛlˀ]. Such idiosyncrasies presumably reflect different underlying representations at the lexical level, and not differences in how the phonological grammar is structured across speakers.

3.4.4 Summary

The alternative analysis proposed here is arguably preferable to the traditional analysis in a number of ways. The alternative analysis is positionally biunique, and does not pose a neutralization problem; all allophones share phonetic and phonological properties; and the analysis does not rely on irregular and unproductive morphological alternations. This is also true for the analysis proposed by Ács and Jørgensen (2016). Their analysis was fully biunique and proposed a larger number of phonemes, most of which were defectively distributed, whereas the alternative analysis proposed here manages with a lower number of phonemes, most of which are not defectively distributed. The alternative analysis is therefore arguably more economical.

3.5 The diachronic trajectory of stop gradation

The traditional analysis may not be a plausible description of the phonological grammar acquired by speakers of Modern Standard Danish, but it does capture a structural generalization which the alternative analysis does not. Whether or not the alternations described in Section 3.2 are relevant for synchronic phonology, they undeniably exist, and there is undeniably some regularity to their occurrence. In this section, I argue that the regularities do not need to be accounted for in a synchronic phonological grammar, as argued in Section 3.3, such a grammar winds up being cognitively implausible. The regularities are rather natural consequences of a number of sound changes which are already reasonably well-described and phonetically well-understood. More specifically, I will argue that the sound changes that produced the current inventory of stops and semivowels in Modern Standard Danish can be understood as reactions to the pressure against obstruent voicing, and/or as maximalization of cues to positional phonological contrasts. As argued by Ohala (e.g. 1990a), this effectively removes the need for an abstract phonological explanation: invoking Occam's razor, Ohala maintains that one phonological fact does not require two explanations.[12]

This idea is the cornerstone of the Evolutionary Phonology framework (e.g. Blevins 2004, 2015). Blevins argues that the explanations for many systematic patterns in synchronic phonology are the result of phonetic pressures operating during previous stages of a language. She proposes a typology of possible sound changes relying on the three-way distinction between CHANGE, CHANCE, and CHOICE, also known as the CCC-model. I introduce the basics of Evolutionary Phonology and the CCC-model below. Subsequently, I cover each of the individual sound changes that led to the current state of affairs (see Section 2.2), and discuss their phonetic bases and how they align with the CCC-model.

3.5.1 Evolutionary Phonology and the CCC-model

A core tenet of Evolutionary Phonology is that current phonological systems are best understood through the sound changes that produced them. This idea was also central to the neogrammarian school of phonology (e.g. Karsten 1894; Baudouin de Courtenay 1895; Jespersen 1924). Our understanding of the phonetic mechanisms underlying systematic sound changes has drastically improved in the last century, and Evolutionary Phonology incorporates this knowledge. The reliance on phonetic explanation means that the

[12] I think this statement is too strong. In this chapter, I argue that speakers build phonological representations with no regard for the language's history; consequently, some phonological patterns may have diachronic explanations *and* be synchronically active. Contrary to Occam's razor, such patterns should be accounted for in both the diachronic and synchronic domains. For a general critique of the reliance on Occam's razor in phonology, see Ploch (2003).

distinction between phonetics and phonology seems to a large extent to be obsolete (see also Ohala 1990b, 2005). Note, however, that phonetic explanation is relegated to the diachronic dimension; the pressures underlying sound change are natural, but synchronic phonology is abstract, non-teleological, and non-optimizing. This leads to synchronic grammars that can appear messy, containing patterns that may be either natural, unnatural, or seemingly random; as Blevins (2004: 84) points out, despite not having "great aesthetic appeal", this is necessary to account for the breadth of phonological data.

Blevins assumes that sound change is listener-oriented (see also Hyman 1976; Ohala 1981); it happens when a speaker produces a sound with a particular phonological representation in mind, and a listener associates it with a different representation. In other words, sound change is rooted in misperception. This seemingly erases the speaker, and hence articulation, from the picture, but this is a little misleading: misperception may well be rooted in articulation. If producing a particular sound is difficult (see Ohala 1983a, 1989), the speaker is more likely to partially miss the articulatory goal, which may cause the listener to perceive the resulting sound differently than the speaker intended. In this case, the speaker may be the catalyst for sound change, but the actual recategorization is still done by the listener, who only has access to the acoustic signal and not to the articulatory mechanism that produces it (Ohala 1996).[13]

CHANGE happens when an intended sound is perceived as another sound due to inherent perceptual similarities. A well-known example is the process [k] → [tʃ] before high front vowels (see e.g. Hock 1991: 71ff.). Dorsal consonants are very prone to consonant–vowel (CV) coarticulation, partially because the tongue body is less finely controlled than the tongue tip and blade (Vilain et al. 1998; Ouni 2014), and partially because the dorsum is the main articulator in vowel production, so dorsal consonants and adjacent vowels are necessarily rather co-dependent, unlike other active articulators which are relatively independent from vowel production. Accordingly, the

[13] As Hamann (2006) points out, BiPhon has an advantage over Evolutionary Phonology in this regard, as BiPhon uses the same grammar for production and perception. This can be modeled explicitly in BiPhon, whereas formal modeling is scarce in Evolutionary Phonology.

precise point of occlusion in [k] is fronted before high front vowels. CV-coarticulation alone cannot account for this change, however, as the point of occlusion in [tʃ] is actually further front than the point of maximal constriction in [i] (Ohala 1992). The change must thus be rooted in perception. The acoustic characteristics of the release burst in fronted [k], particularly when the following vowel has a narrow approximation, are very similar to those of [tʃ] (Guion 1998). As such, [k] → [tʃ] may be conditioned by CV-coarticulation, but the change follows from perceptual similarities caused by CV-coarticulation. CHANGE in Evolutionary Phonology derives from what Ohala (e.g. 1989, 1993) calls *hypo-correction*.

CHANCE happens when the realization of an underlying representation is miscategorized because the listener can assign multiple possible analyses to it. In other words, there is a mismatch between the phonological analyses of the speaker and the listener. Blevins and Garrett (1998; Blevins 2004: ch. 2) give an example from consonant–vowel metathesis with laryngeal segments, as found synchronically in e.g. Cayuga (Foster 1982). For example, /aʔ/ is phonologically ambiguous, because it is likely to be realized with creaky voice throughout and glottal closure on both sides of the vowel, i.e. [ʔa̰ʔ], making it difficult for the listener to decide the 'phonological origin' of /ʔ/; the underlying form could be either /aʔ/ or /ʔa/. CHANCE in Evolutionary Phonology derives from what Ohala calls *hyper-correction*.

Another illuminating example of CHANCE can be found in a development in the transition between Proto-Nordic and Old Norse. In Proto-Nordic, the voiced fricatives *β ð ɣ were allophones of the voiced stops *b d g.[14] In the transition to Old Norse, the voiced fricatives were reanalyzed as allophones of the voiceless fricatives /f θ h/ (Nielsen and Stoklund 2018). [β ð ɣ] were phonologically ambiguous, because they could be analyzed as post-vocalic weakened allophones of either /b d g/ (retaining laryngeal features) or /f θ h/ (retaining manner features). Both analyses are reasonable. /b d g/ are a likely source of [β ð ɣ], since final closure voicing in stops is generally dispreferred,

[14]In Section 3.5.3, I return to the issue of whether these stops were actually voiced. They are generally referred to as such in the Danish historical linguistics tradition (Brøndum-Nielsen 1928–1973; Skautrup 1944–1970; Hansen 1962–1971).

but any increase in aperture makes voicing easier to maintain (Ohala and Riordan 1979; Westbury 1983). Similarly, /f θ h/ are a likely source of [β ð ɣ], as voiced obstruents are generally considered 'weaker' than voiceless obstruents (Anderson and Ewen 1987; Honeybone 2008; this idea is discussed in more detail in Chapter 4). The reanalysis of the voiced fricative allophones in the transition to Old Norse is not due to inherent perceptual similarity, but due to a mismatch between speaker and listener in the phonetics–phonology mapping.

CHOICE is a result of the intrinsic variability of speech. CHOICE may well be a factor in sound changes that are primarily characterized as CHANGE or CHANCE. The sum of the speaker's experience with how an underlying representation is phonetically realized is necessarily different from the listener's, meaning the listener's conception of the 'best exemplar' of a representation will also be slightly different. Over time, this may lead to systematic drift in phonetic realization. This requires Blevins to assume a mechanism whereby listeners' phonological representations are continuously updated by their linguistic experiences; examples of such mechanisms are rich episodic memory of encountered word tokens, as employed in Exemplar Theory (e.g. Goldinger 1996; Bybee 2001, 2006; Pierrehumbert 2001, 2016), or cue constraints/cue connections emerging from linguistic experience as in BiPhon (Boersma 2006, 2009; Chládková 2014; Boersma et al. 2020). Blevins further relies on the hyper–hypo (H&H) theory of Lindblom (1990), where speech is situated on a continuum from hyper-articulated to hypo-articulated, roughly corresponding to very clear speech and very unclear speech; Blevins assumes a direct relationship between frequency and the hyper-to-hypo-articulated continuum, such that changes in relative frequency of words or phonemic categories lead to corresponding changes in articulation.

Consider degemination in high-frequency environments. In a language with a category /tː/, listeners will encounter a lot of variation in its precise realization: true geminates [tː], preaspirated variants [^{h}t], preglottalized variants [ʔt], singleton variants [t], and a lot of variation in phonetic implementation within those broad categories in terms of closure duration, voice onset time, F_0-perturbations, etc. [t] is a hypo-articulated variant likely to be found in high-frequency lexical items. If this is sufficiently common, listeners are likely to reanalyze those items as having an underlying singleton /t/. Such a mechanism may underlie the relatively recent degemination in Danish (see Section 2.2.1), where geminates were found very frequently before schwa in infinitive verbs.

Despite being by definition very frequent, hypo-articulated speech does not always have an evolutionary advantage. Hypo-articulated speech is also more likely to be misperceived, and only correctly perceived tokens can influence the phonological grammar. Wedel (2006) uses computer simulations to show how how very hypo-articulated speech is more likely to result in incorrectly categorized or uncategorizable exemplars, which is an evolutionary disadvantage. Blevins and Wedel (2009) use a similar model to show how lexical competition may inhibit sound change. Boersma and Hamann (2008) and Boersma et al. (2020) have also modelled the evolutionary disadvantage of signals which are difficult to categorize using various computational implementations of BiPhon.

3.5.2 Stop gradation in five diachronic steps

In this section, I argue that stop gradation is the result of a series of related sound changes, all of which can be considered reactions to the pressure against obstruent voicing, and/or increasing the saliency of cues to phonological contrasts. Historically, consonant gradation is one of two main sound changes that resulted in the split between Danish and the other peninsular North Germanic languages, i.e. Norwegian and Swedish, the other being the widespread vowel reduction in unstressed syllables, known in Danish as *infortissvækkelsen* 'infortis weakening'. Recall from Section 2.2.1 that consonant gradation affected both stops and fricatives.

The pressure against obstruent voicing led to vastly different outcomes in strong and weak position in Modern Standard Danish. In strong position, the voicing-based laryngeal contrast in obstruents developed into an aspiration-based one; in other words, strong stops underwent fortition. Voiced fricatives, meanwhile, weakened to approximants. In weak position, voiced stops and fricatives developed into semivowels. I will sketch five diachronic steps that together led to the alternations found synchronically in Modern Standard Danish, and show how each of them are well-motivated with reference to the

CCC-model and known articulatory and perceptual pressures. I argue that Steps 1–4 all constitute CHOICE, as they affect allophone distributions but not contrasts; they are, however, necessary steps to explain the phonological reorganization in Step 5. The relative timing of these steps is supported in part by the existing literature on Danish historical linguistics (Brøndum-Nielsen 1928–1973; Skautrup 1944–1970; Hansen 1962–1971.) This account assumes that Danish used to have a voicing-based contrast, which is somewhat contentious (see Section 2.2); I discuss this further in Section 3.5.3. The five steps are summarized in (16).

(16)
Step 1 Singleton voicing in weak position
[p t k] ⟶ [b d g]
Step 2 Loss of closure in weak position
[b d g] ⟶ [β ð ɣ]
Step 3 Loss of voicing in strong position
[b d g p t k] ⟶ [p t k p^{h} t^{h} k^{h}]
Step 4 Increased aperture in weak position
[β ð ɣ] ⟶ [β̞ ð̞ ɰ]
Step 5 Recategorization of weak allophones
[β̞ ð̞] ⟶ [ʊ̯ ɣ̯]
[ɰ] ⟶ [ʊ̯ ɪ̯]

This discussion is mostly limited to High Copenhagen Danish, since this variety developed into Modern Standard Danish. However, it is worth noting again that the extensive dialect leveling in the Danish speech community is a relatively recent development (Kristiansen 2003a), and the steps described in (16) happened to varying degrees and had varying outcomes in different varieties, as covered in Section 2.5.3.

In the case of the historic velar stop /g/, a further step is worth discussing, namely elision in weak position. There are at least two good reasons to consider elision a natural next step in the stop gradation process: 1) Some regional varieties, in particular the insular varieties of Funen and Lolland-Falster, show complete elision much more extensively than Modern Standard Danish (see Figures 2.1–2.4). 2) There is an increasing tendency in Modern Standard Danish for [ɪ̯] to elide in new contexts (as briefly discussed in Section 2.4.4). In describing the traditional analysis in Section 3.2 above, I noted that /g/ is elided after high vowels, and otherwise realized as [ɪ̯] after front vowels. In recent years, the [ɪ̯] allophone, which is always considered underlying /j/ in the alternative analysis, is increasingly elided after all front vowels. In the pronunciation dictionary of Brink et al. (1991), the word *flag* is described as variably pronounced [flæːˀɪ̯ ~ flæɪ̯ˀ], whereas in Schachtenhaufen's (2020–) more recent pronunciation dictionary, [flæːˀ] is given as the standard pronunciation, with [flæɪ̯ˀ] described as conservative; this development is also mentioned by Grønnum (2005: 295).[15] Elision will not be discussed further below.

3.5.2.1 Singleton voicing in weak position

Step 1 is repeated in (17), which also shows the other assumed positional contrasts at the time.

(17)	*Strong position*	[b d g]
		[p t k]
	Weak position	[p t k] → [b d g]
		[pː tː kː]

This development will have taken place during the Middle Danish period. During this time period, I assume that stops showed a voicing-based contrast in strong position. In weak position, there would have been a distinction between singleton and geminate stops. Earlier still, there were also laryngeal contrasts in both weak singletons and geminates, but these contrasts were lost due to devoicing (see Section 2.2.1). Evidence in favor of a development where the weak singletons were subsequently voiced comes from written sources in the 13th–15th centuries; example spellings include <diyb> from Old Danish *djup* 'deep' and <lægin> from Old Danish *læken* 'the doctor' (Frederiksen 2018).[16] Voicing in weak singletons is the least well-described of the five diachronic steps, but it helps explain both the phonetic mecha-

[15] Transcriptions throughout this chapter have reflected relatively conservative pronunciation where [ɪ̯] is retained after most front vowels.

[16] Note that Danish orthography was not regulated at the time, and it is not possible to evaluate exactly what these orthographic changes reflected (e.g. Jørgensen 2021).

nisms behind stop gradation, and the observed patterns of regional variation, as discussed further in Sections 3.5.3 and 6.8.2.

Several mechanisms could have caused the development of voicing in weak singletons. Continuous voicing requires a transglottal air pressure differential above a certain threshold, which can be difficult to maintain during a stop closure, as discussed further in Section 3.5.2.2. Relatively speaking, conditions for stop voicing are ideal in medial post-vocalic position. The vocal folds are already vibrating when the closure begins, subglottal pressure is high, and supraglottal pressure low. This increases the chances that voicing from the preceding vowel will 'bleed' into the stop closure; this process is also known as passive voicing. The proportion of passive voicing is mediated by closure duration, such that a fully voiced stop is much more likely if closure duration is short (Davidson 2016). In other words, passive voicing will have affected singletons more than geminates. Increased passive voicing in singletons would have had an evolutionary advantage by providing a further cue to the singleton–geminate contrast.

Voicing is less articulatorily natural in final position (Westbury and Keating 1986), but the perceptual advantage of voicing would have been significant. Singleton–geminate contrasts in stops are unstable in final position, since closure duration is not a very salient cue here (Kraehenmann 2001); this is likely why the Norse geminate stops often developed other primary cues, such as preaspiration in Icelandic and Faroese (Page 1997) and preglottalization in West Jutland Danish (Ringgaard 1960), as mentioned in Section 2.2.1. Step 1 will have effectively created a final voicing contrast rather than an unstable final length contrast in Danish. This step is rooted in CHOICE, as it affects the cues to an existing contrast, but does not actually change the contrast.

3.5.2.2 Loss of closure in weak position

Step 2 is repeated in (18), which also shows the other assumed positional contrasts at the time.

(18)	*Strong position*	[b d g]
		[p t k]
	Weak position	[b d g] → [β ð ɣ]
		[pː tː kː]

This development is a natural reaction to the pressure against closure voicing. As mentioned above, voicing can only be maintained with a sufficient transglottal pressure differential; if supraglottal pressure is higher than subglottal pressure, vocal fold vibration ceases by necessity (Ohala and Riordan 1979; Westbury and Keating 1986). Essentially, maintaining vocal fold vibration requires free passage of air through the glottis, which is inhibited eventually if there is no free passage of air through the supraglottal cavities. Put bluntly, this means that closure voicing in stops is generally "unnatural" (Ohala 1983a). This serves to explain a number of typological patterns, namely why languages with voiced stops also always have voiceless stops, but not *vice versa* (Maddieson 1984). It also explains why gaps in voiced stop inventories are found at places of articulation in the back of the oral cavity, which yield only a small cavity between the glottis and the occlusion, and hence a quicker rise in supraglottal air pressure (Ohala 1983a; Hayes and Steriade 2004; Brown 2006). An example of this is found in the history of some traditional Jutlandic varieties of Danish, where the voiced medial and final geminates [bː dː] degeminated, while [gː] devoiced and merged with [kː] (Sørensen 2012; see Section 2.5.3). The literature on this topic is covered in greater detail in Chapter 4.

Some initial closure voicing is common in post-vocalic position; subglottal pressure is initially high and supraglottal pressure low due to influence from the preceding vowel. In medial intervocalic position, this often results in voicing throughout (most of) the closure; in final position, however, Westbury and Keating (1986) hypothesize that the initially good conditions for voicing are quickly counteracted by an increasing inspiratory force.

Languages have two ways of getting rid of final closure voicing: 1) remove the voicing by devoicing the offending stops, or 2) remove the closure by increasing the aperture. Devoicing seems to be the most common solution (e.g. Blevins 2006). A possible explanation for this is that syllable-final segments are generally lengthened, resulting in longer stretches of voicelessness in coda stops (Blevins 2004: 103ff.). This may lead to CHANGE: even if the speaker generally intends to voice final tokens, many of them will be perceived by the listener as voiceless. This solution has the side effect of neutralizing any relevant phonological contrast. Increased aperture, on the other hand, ensures passage of air, which may allow the laryngeal contrast to be maintained. This solution would not have neutralized any phonological contrast in Danish; recall from Section 2.2 that Step 2 occurred as part of a chain shift, where the existing post-vocalic voiced fricatives also developed into approximants.[17]

This serves to explain the increase in aperture syllable-finally, but not in other weak positions, namely onsets in unstressed syllables before neutral vowels. There is a good explanation why weak positions in Danish are treated as a group, namely that many neutral vowels in Danish are (part of) inflectional morphemes, such as the infinitive ending in verbs and definite or plural ending in adjectives, both of which are [-ə], or the present tense ending in verbs and regular plural ending in nouns, both of which as [-ɐ].[18] In other words, onset consonants before neutral vowels are often root-final, and have arguably been resyllabified from syllable-final position.[19] Most likely, the change applied strictly in syllable-final position at first, and then spread to all weak positions as it phonologized. This follows the general life cycle of phonological changes, where phonological rules develop from gradient phonetic phenomena, and then gradually narrow their domain from from phrase level to word level to stem level, before eventually lexicalizing (see e.g. Bermúdez-Otero 2015; Ramsammy 2015).

This account is based on the aerodynamics of articulation, but in accordance with Evolutionary Phonology, the actual change must be rooted in perception. The change in Step 2 consisted of restructuring the primary allophones of an existing set of phonological categories. This is likely rooted in CHOICE: if possible realizations of /b/ consisted of any of [p b β], then [β] resulted in maximal dispersion of the singleton–geminate contrast, as [β] is very unlikely to be perceived as a geminate stop. It is neither CHANGE nor CHANCE yet, as Step 2 does not necessarily lead to any recategorization on the part of the listener.

[17] In fact, it is exceedingly uncommon for phonologically conditioned spirantization to neutralize a contrast (Gurevich 2004).

[18] [ɐ] is often assumed to be derived from underlying /-ər/ (e.g. Basbøll 2005); as mentioned in Section 1.4, it is phonetically indistinguishable from [ʌ] in my speech, but clearly shows 'schwa-like' phonological behavior.

[19] There *are* schwa-final lexical items, and they behave the same. This is perhaps the result of analogy with 'morphological schwas'.

3.5.2.3 Loss of voicing in strong position

Step 3 is repeated in (19), which also shows the other assumed positional contrasts at the time.

(19)	*Strong position*	[b d g] → [p t k]
		[p t k] → [pʰ tʰ kʰ]
	Weak position	[β ð ɣ]
		[pː tː kː]

As mentioned above, Step 3 builds on the somewhat contentious assumption that the laryngeal contrast in Danish used to be voicing-based; this is discussed further in Section 3.5.3. Brink and Lund (2018) date the loss of voicing as sometime before 1700; I am not aware of anyone else who have attempted to date this development.

Step 3 is also a reaction to the pressure against obstruent voicing. In essence, the problems of retaining voicing in final position also hold in initial position. Additionally, in initial position, there is often no preceding vowel, so there is no ensurance that subglottal pressure is initially high during the closure. This may lead to an increasing number of /b d g/ exemplars without closure voicing, which is an obvious problem, since the initial laryngeal contrast in stops carries a very high functional load in Danish. A reaction to this would be a gradual push towards an aspiration-based contrast, since aspirated tokens of /p t k/ would more likely be correctly perceived.[20] Silverman (2004) provides an illuminating exemplar-based account of the timeline of a similar change in American English.

The reactions to the pressure against obstruent voicing in Danish were quite different in strong and weak positions. This is not random; Keating et al. (2004) show a tendency for articulatory strengthening in the beginning of prosodic domains, and a tendency for gestural undershoot at the end of prosodic domains. In other words, initial segments are more likely to be hyper-articulated, while post-vocalic segments are more likely to be hypo-articulated.

[20] Vaux and Samuels (2005) argue that the maximally dispersed two-way laryngeal contrast in stops would be a voiced series and an aspirated series, but an aspiration-based contrast like in Danish likely strikes a better balance between articulatory and perceptual ease.

As with Step 2, I would argue that this change is rooted in CHOICE, as the result is a change in allophone distribution rather than a change in categorization.

3.5.2.4 Increased aperture in weak position

Step 4 is repeated in (20), which also shows the other assumed positional contrasts at the time.

(20)	*Strong position*	[p t k]
		[p^h t^h k^h]
	Weak position	[β ð ɣ] → [β̞ ð̞ ɰ]
		[pː tː kː] → [p t k]

Step 4 is a component of the same chain shift as Steps 1–2. It is difficult to date Step 4 precisely relative to Step 3, since the two developments are relatively independent. The point that I will discuss here is the increased aperture of the original singletons, but note that I assume degemination happened around this time also; these changes are also difficult to date precisely relative to each other. I will not discuss degemination further.

Step 4 is likely a further reaction to the pressure against obstruent voicing. It was noted above that continuous voicing requires free passage of air through the glottis; fricatives *do* allow for the passage of air, but not *freely*. By definition, air passes through a narrow constriction in the oral cavity, which means that supraglottal air pressure does rise over time, but more slowly than in stops. As a result, the constriction in voiced fricatives is generally less narrow than in their voiceless counterparts (see Stevens 1998: 477ff.). Aperture in voiced fricatives is likely to increase over time, as increased aperture produces stable acoustic cues to the laryngeal contrast and eases articulation. As an example, harmonicity in German /v/ is rather high, and comparable to the Dutch approximant /ʋ/ (Hamann and Sennema 2005). On the other hand, maintaining frication during voicing is in itself difficult. Turbulent airflow requires high air pressure behind the oral constriction, which is difficult to achieve when 1) airflow is obstructed at the glottis, and 2) the constriction in itself is relatively open (Ohala 1983a).

Botma and van 't Veer (2013) show that voiced fricatives (in particular non-sibilants) do not obey the general observation that voiced obstruents have voiceless counterparts (see Maddieson 1984). This is particularly common for [β ð], but has also frequently been observed for [ɣ]. They account for this by arguing that these particular sounds are not in fact obstruents, but sonorants – and as such, that their natural state is to be voiced. If true, it seems likely that change in the direction of more 'sonorancy' (i.e. higher aperture) would increase the perceptual saliency of these sounds, producing a change like the one in Step 4.

3.5.2.5 Recategorization of weak allophones

Step 5 is repeated in (21), which also shows the other assumed positional contrasts at the time.

(21)	*Strong position*	[p t k]
		[pʰ tʰ kʰ]
	Weak position	[β̞ ð̞] ⟶ [ʋ̯ ɣ̯], [ɰ] ⟶ [ʋ̯ ɹ̯]
		[p t k]

Following Step 4, the allophones are ripe for recategorization. Recall from Section 2.2 that at this stage, the weak allophones of the historical voiced fricatives and singleton stops are in some cases near-homophonous: the fricatives have either elided or developed into [ʋ̯ ɹ̯], and the stops have developed into [β̞ ð̞ ɰ]. In Step 5, following the alternative synchronic analysis of stop gradation proposed in Section 3.4 above, the weak allophones of /b g/ are reanalyzed as allophones of /v j/. The mechanism behind this is a combination of CHANGE and CHANCE.

Some of the perceptual mechanisms responsible for the recategorization were already discussed in Sections 3.4.2 and 3.5.1. As [β] developed a more open approximation, the primary acoustic cue to place of articulation changed from frication noise generated at the lips to formant structure. Lip approximation – especially when accompanied by lip rounding[21] – increases the size of the cavity in front of the tongue constriction, thereby lowering *F*2. The size of this cavity is also increased by raising the tongue body towards the back of the oral cavity. This makes low *F*2 an ambiguous cue: it can be caused by either tongue backing, lip constriction, or both. For this reason, these gestures usually accompany one another (Flemming 1995); most front vowels are unrounded, and most back vowels are rounded (Ladefoged and Maddieson 1996: 292ff.). This may cause CHANGE: a speaker produces [β̞], which is inherently ambiguous. It has low *F*2 due to labial constriction, but this may easily be perceived by the listener as being due to a combination of lip rounding and tongue backing, as in the existing allophone [ʊ̯]. A similar instance of CHANGE is found in some Latin American varieties of Spanish, where [β̞] alternates with [ɰ] after the high back vowel [u] (Mazzaro 2010): low *F*2 from labial constriction is perceived as being caused by velar constriction. Kijak (2017) gives many other examples of labial–velar interactions.

CHANCE is a logical next step. [ʊ̯] is now phonologically ambiguous, as it can be analyzed by the listener as an allophone of either /b/ or /v/. Given the lack of evidence in the input for a /b/ analysis, I suggest that Danish speaker–listeners converged on /v/. As we saw above, however, this does not seem to have been the ultimate consequence in Modern Standard Danish, where there is now only stylistic alternation between [p ~ ʊ̯]. Interestingly, Jørgensen (2021) shows evidence from Renaissance-era orthography (16th–17th centuries) that this sound change had actually progressed much further in previous stages of Danish; for example, Modern Standard Danish [ˈʁɑɪ̯nskæːˀp] *regnskab* ‘account’ was written as <regenschaff>, with <ff> indicating a labial–velar glide at the time. As mentioned above, this change was largely ‘rolled back’; We can only speculate on the causes of this, although a probable cause would be the resulting rampant homophony.

[21] Little is known about how [β̞] was pronounced in earlier stages of Danish, or even how bilabial approximants are produced in general. Ladefoged (1990) points out that sounds described as bilabial fricatives may have either narrowed, vertically compressed lips, or protruded lips, and that the IPA has no way of symbolizing this distinction. See also Martínez-Celdrán (2004), who points out that IPA classification of approximants is generally imprecise – and that many approximants are characterized precisely by their lack of articulatory precision.

In Section 3.5.1 above, I discussed the common change [k] ⟶ [tʃ] before front vowels, and noted that dorsal consonants are particularly prone to CV-coarticulation. This is also true for the velar approximant [ɰ], which is constricted further back after back vowels, and further front after front vowels. As with [β̞], the main acoustic place cues in [ɰ] come from formant structure, and these are significantly affected by the exact tongue position. This may cause CHANGE. Adjacent back vowels cause *F*2 lowering, and they likely also cause assimilatory rounding, since most Danish back vowels are rounded; this makes [ɰ] likely to be perceived by listeners as [ʊ̯]. Adjacent high vowels cause *F*2 raising, making [ɰ] likely to be perceived as [ɪ̯]. As with [β̞], CHANCE is a logical next step, since [ɰ] is now phonologically ambiguous in both of these contexts. With little evidence in favor of a /ɡ/ analysis, I suggest that speaker–listeners converged on analyses of /v/ after back vowels and /j/ after front vowels. In some cases, the phonetic result of this would have been diphthongs with a very short trajectory, such as [iɪ̯ uʊ̯]. These are inherently ambiguous, simply because the perceptual distance between the endpoints is so small that they will likely be perceived and reinterpreted as long monophthongs; see e.g. Lindblom (1986), who argues that optimal diphthongs are those with the longest possible trajectory. An example of this pressure in action is found in the widespread monophthongization of RP English centering diphthongs, such as [ɪə eə ʊə] (Lindsey 2019: ch. 13).

The development of [ð̞] is very intriguing. This sound has also changed its pronunciation, and recently developed a prominent dorsal component (Brink and Lund 2018; see Section 2.2.2), leading to its current semivocalic realization [ɣ̯]. This may be rooted in CHOICE-like change in allophone selection, but there may also be an element of CHANGE at play. Petersen et al. (2021) describe the articulation of [ɣ̯] as raising the entire tongue tip and blade towards the upper teeth, alveolar ridge, and hard palate simultaneously; the approximation, however, remains very open.[22] The tongue tip and blade usually do not play any major role in the production of (semi-)vowels, although they often play a role in the production of so-called 'rhotic vowels' (Lawson et al. 2013; Mielke 2015), which are produced with tongue bunching or

[22]Note that this is a purely introspective description.

retroflexion. The involvement of anterior parts of the tongue have an interesting effect on the resulting formant structure.

*F*2 may reflect either the cavity in front of or behind the dorsal constriction, depending on the respective sizes of these cavities. *F*2 tends to reflect the size of the front cavity when it is sufficiently large, and otherwise mostly reflects the size of the back cavity. Put simply, *F*2 in front vowels mostly reflects the back cavity, *F*2 in back vowels mostly reflects the front cavity, and *vice versa* for *F*3 (Fant 1960). Tongue bunching or retroflexion in the production of American English /ɹ/ may yield spacious cavities both in front of and behind the main constriction, with the front cavity being the largest, and thus reflected in *F*2. Regardless of how /ɹ/ is exactly articulated, the cavities in front of and behind the primary dorsal constriction are both large, yielding low values of both *F*2 and *F*3 (Espy-Wilson et al. 2000; Zhou et al. 2008). Danish [ɣ̞] may be a similar case. There are no detailed studies of the articulation of Danish [ɣ̞], so it remains unclear exactly how the open approximation by anterior parts of the tongue affects the size of the front cavity; this may have a significant influence on the formant structure of the sound. Either way, *F*2 should be low given the very open approximation, and as above, this is a cue that listeners will likely perceive as being related to dorsal approximation. This can, in turn, serve to explain the increasing prominence of the dorsal gesture in [ɣ̞].

Whether or not [ð] ⟶ [ɣ̞] is an instance of CHANGE will depend on how the alternations are analyzed phonologically. In the alternative analysis proposed in Section 3.4, as in the traditional analysis, [ɣ̞] is considered an allophone of /d/. However, for listeners to establish the coupling between [t] and [ɣ̞], they would need positive evidence in favor of such an analysis, which is largely missing from both the phonetics and morphology of Modern Standard Danish. Throughout the course of consonant gradation, the conflicting acoustic evidence from [ɣ̞] may have caused CHANGE, not due to ambiguous evidence about whether [ɣ̞] should be linked to /d/ or something else, but due to the lack of positive evidence linking [ɣ̞] to any other sounds in the inventory. This case is similar to the well-known discussion about whether [h ŋ], which are in complementary distribution, should be considered allophones of one phoneme in English. Most phonologists agree they should not, since complementary distribution alone is not sufficient evidence to establish phonemic categoryhood. Perhaps this is also increasingly the case for [t ɣ̞], as these sounds drift further and further apart; hopefully, future research on the precise articulation and acoustics of [ɣ̞] will shed further light on this issue.

3.5.3 The history of closure voicing

The analysis presented above has relied on the assumption that Danish stops displayed a voicing contrast in an earlier stage of the language. This assumption requires further discussion.

As noted in Section 2.2.1, several scholars have argued that Proto-Germanic did not have distinctive voicing. In fact, Honeybone (2002: 149ff.) suggests that scholars who have "actually considered the problem" agree that the contrast was not voicing-based, and that others tend to use 'voicing' as an abstract catch-all term for two-way laryngeal contrasts.[23] An early opponent to the Proto-Germanic contrast being voicing-based is Alexander (1982), who assumes that the contrast was rather rooted in articulatory force (see 2.3.5). Iverson and Salmons (1995, 2003a) and Honeybone (2002) argue that the contrast was aspiration-based. There are several lines of evidence in favor of this.

Iverson and Salmons (2003a) propose the sound law *Germanic Enhancement*, as defined in (22).

(22) Laryngeally unspecified stop → [glottal width]

The feature [glottal width] in (22) is Avery and Idsardi's (2001) equivalent to [spread glottis]. In accordance with (22), any stop without an underlying laryngeal feature is assigned aspiration during phonetic implementation. As a result of Germanic Enhancement, the bulk of modern Germanic languages have aspirated-based stop contrasts. Iverson and Salmons argue that the result of Germanic Enhancement, namely aspiration of voiceless stops, is a prerequisite for Grimm's

[23] An example of this is Moulton (1954: fn. 7), who explicitly notes that his use of 'voiced' denotes a contrast only, and does not refer to phonetic substance. Perridon (2008), however, is a later counter-example of Honeybone's generalization.

Law, which eventually resulted in the spirantization of these stops in German.

According to Iverson and Salmons (1995: 389), aspiration helps explain spirantization, as aspirated stops usually have shorter and weaker closures. This is not actually straightforward; they cite Kohler (1984) for this, who claims that long closure duration is generally a cue to [+fortis], but that there is a trade-off relation between closure duration and aspiration (see Section 2.4.3.1). Such a trade-off relation has indeed been found for some languages, such as Danish (Fischer-Jørgensen 1954; Puggaard et al. 2019) and Swedish (Löfqvist 1975b); however, closure duration is longer in aspirated stops than unaspirated stops in other languages like German (Braunschweiler 1997; Pohl and Grijzenhout 2010) and English (Luce and Charles-Luce 1985; Byrd 1993). This suggests that closure duration and aspiration are independently controlled. A more likely explanation is that aspirated release is itself a source of assibilation, due to overlap between oral and glottal noise sources, as suggested by Hock (1991: 436).

Another piece of evidence in favor of an aspiration-based contrast in Germanic, pointed out by Honeybone (2002: 150ff.) is that the only laryngeal assimilation process evident in Gothic, the earliest recorded Germanic language, is regressive devoicing. As Iverson and Salmons (1995) show, regressive devoicing is generally found in languages with aspiration-based laryngeal contrasts.

Some Germanic languages do display voicing-based stop contrasts, namely Dutch, Frisian, Afrikaans, and Yiddish (Cohen et al. 1959; Iverson and Salmons 1995). Iverson and Salmons (2003b) assume that stop voicing in these languages is an innovation due to language contact; the 'Netherlandic' languages (Dutch, Frisian, and Afrikaans) developed a voicing-based contrast through contact with Romance languages, and Yiddish developed one through contact with Slavic languages. Stop voicing is also observed more widely in the modern Germanic languages: Swedish displays a contrast between aspirated stops and consistently pre-voiced stops (Helgason and Ringen 2008); closure voicing is found consistently in some varieties of English (e.g. the variety of Northeast England; Harris 1994: 137), and inconsistently in others (Flege 1982); and closure voicing is found in some varieties of German (Braun 1996). As discussed in Section 6.6, there are also relic areas with relatively widespread voicing in Northern Jutland, Denmark.

I do not intend to argue against the perspective that Proto-Germanic had an aspiration-based contrast. It is much beyond the scope of this chapter, and some of the proposed lines of evidence are quite convincing. That leaves a conundrum: in the Danish historical linguistics tradition, earlier stages of Danish are described as having voiced stops (Brøndum-Nielsen 1928–1973; Skautrup 1944–1970; Hansen 1962–1971). This may not be informative; as noted above, this terminology is often used by scholars who make no claims about the phonetic substance of a laryngeal contrast.[24] More importantly, the stop gradation process presented in previous sections is simply more well-motivated and easier to explain if we assume that at least the weak singleton stops were voiced at an early stage. This is especially true for Step 2, whereby /b d g/ weakened to voiced fricatives in weak position. This development is a natural case of lenition if /b d g/ were voiced at the time; on the other hand, a process of [p t k] ⟶ [β ð ɣ] in weak position would be more surprising, especially since there is no evidence of an intermediate step with voiceless fricative alternants.[25] In other words: if we assume that /b d g/ were voiced, we can propose a series of well-motivated changes; if /b d g/ were not voiced, that is much more tricky.

It is possible that Danish historically had a voiced–aspirated contrast, as in Modern Standard Swedish (Helgason and Ringen 2008).[26] This would not change much about the proposed diachronic account here; the motivation for loss of voicing given in Step 3 would be the same, and the development of aspiration could be considered a result of Germanic Enhancement. The development of voicing at some earlier point in time would, however, require an explanation. I consider

[24]Note however that Brink and Lund (2018), who dated the loss of voicing as sometime before 1700, are known to take phonetic substance very seriously.

[25]Some regional varieties do, in fact, show such alternation between /b g/ and [f x], but this is generally only the case in the southern area of Jutland close to the German border (see Section 2.5.3). I discuss this further in Chapter 6, where I also show that this is an area with prominent aspiration in /p t k/.

[26]This could even have been a co-development; after all, Danish and Swedish relatively recently developed from a common language (Gooskens 2020).

that beyond the scope of this chapter, but see Vaux and Samuels (2005) for an argument that voiced–aspirated contrasts are perceptually well-motivated because they are maximally dispersed. It is also possible that voicing was only found in weak position, and that stops in strong position always displayed an aspiration-based contrast. This would also not change much about the proposed diachronic account here, but it would make it more difficult to explain the more widespread voicing found in some traditional regional varieties.

3.6 Conclusion

In this chapter, I have argued that the notorious alternations between voiceless stops and semivowels can no longer be taken as evidence of shared phonological category membership in Modern Standard Danish; instead, the alternations are better understood in the light of the diachronic pressures that produced them.

In the spirit of Ohala (1983b, 1986), I have drawn on many different types of evidence in arguing against the traditional synchronic account of stop–semivowel alternations. From a Natural Phonology perspective, the traditional analysis is highly unnatural; the analysis relies on morphophonological patterns that are unproductive and irregular, have low type frequency, and are acquired late by language-learning children; and the great phonetic distance between allophones will cause problems for any theory of phonological representation that employs a criterion of phonetic similarity. Since the alternations cannot be taken as evidence of phonological category membership, I argue that the morphophonological patterns resulting in alternations must instead be considered suppletive in the cognitive sense. This alternative analysis is arguably not subject to any of the critiques I pose against the traditional analysis.

Historical developments and well-known phonetic pressures operating diachronically are more informative in accounting for the synchronic patterns. There is a long tradition in linguistics of explaining sound patterns with reference to the changes that produced them, leading from the Neogrammarians through the work of Ohala and up to Blevins' theory of Evolutionary Phonology, among many others. I made reference to Evolutionary Phonology's tripartite model of sound change typology, the CCC model, in accounting for the well-established and well-motivated series of sound changes that eventually led to the current state of affairs in Modern Standard Danish. I argue that Danish was subject to a series of sound changes mostly resulting from the pressure against obstruent voicing, which pulled the allophone distributions in strong and weak positions in opposite directions. Eventually, when the historic post-vocalic voiced stops had lenited to approximants, they were (in some cases) recategorized by listeners as allophones of similar existing approximants, due to a combination of inherent phonetic and phonological ambiguity.

Part II

Corpus studies

CHAPTER 4

The rarity of intervocalic voicing

4.1 Introduction

Intervocalic voicing of underlyingly voiceless stops is phonetically well-understood but is a phonological conundrum. Voicing is usually difficult to maintain during closure, leading to the common assumption that the feature [voice] is phonologically marked in stops. Intervocalically, however, the vocal folds are initially adducted and tensed, and subglottal pressure is high, providing ideal conditions for closure voicing (Westbury and Keating 1986). Hence, voicing is often found in this position, even in languages where voicing does not play a role in distinguishing stops (Kaplan 2010). In other words, the markedness of voicing depends on position; voicing requires an effort in initial and final position, while voicelessness requires an effort intervocalically. This distribution of markedness is difficult to account for phonologi-

---

The research reported in this chapter is collaborative work with Camilla Søballe Horslund and Henrik Jørgensen. This chapter is based (largely verbatim) on a published paper (Puggaard-Rode et al. 2022a), and an earlier version of the study was presented at a conference (Puggaard et al. 2020). Audio data are available online in password-protected form (Grønnum 2016); replication data and code are freely available (Puggaard-Rode et al. 2022b).

cally, where voicing is usually associated with a [voice] feature, i.e. a more complex and marked structure.

As discussed in Section 2.3.2, there is no closure voicing in absolute initial position in Modern Standard Danish, and only negligible voicing in final position. Voicing in intervocalic position is less well-understood. There are essentially two different positions in the literature on the status of intervocalic voicing in Modern Standard Danish. One position holds that stops are systematically voiced in medial position (Abrahams 1949; Fischer-Jørgensen 1954, 1980; Spore 1965; Keating et al. 1983; Kingston and Diehl 1994); the other holds that Danish stops are systematically voiceless in all positions (Jessen 1999, 2001; Beckman et al. 2013). The lack of concrete knowledge about intervocalic voicing is a major gap in the phonetics and phonology of Danish stops. On the phonetics side, the glottal activity during stop production is well-described (Frøkjær-Jensen et al. 1971; Fischer-Jørgensen and Hirose 1974; Hutters 1985; see Section 2.3.3), as are accompanying F_0-perturbations in following sonorant sounds (Fischer-Jørgensen 1969; Jeel 1975; Petersen 1983; see Section 2.3.6). On the phonology side, intervocalic voicing has played a pivotal role in discussions of the underlying representation of Danish stops, and both aforementioned positions have been taken as evidence in favor of underlying representations (e.g. Kingston and Diehl 1994; Beckman et al. 2013; see Section 2.4.3.1). The status of intervocalic voicing could be a skeleton key to both 1) contextualizing existing studies of the phonetics of Danish stops, and 2) evaluating proposals regarding the representation of the laryngeal contrast in Danish.

The problem is intriguing beyond just the Danish context. Much of our knowledge about intervocalic voicing comes from aerodynamic models (such as Westbury and Keating 1986); excluding English, there are few quantitative studies of intervocalic voicing in 'aspiration languages', and particularly few where the data has high ecological validity. As laid out in Section 2.3, the breadth of our existing knowledge about Danish stop production is quite impressive; concrete knowledge about intervocalic voicing may help tie together our existing knowledge of voice onset time, glottal activity, and F_0-perturbations. The relationship between the resulting articulatory actions and acoustic cues are central to our understanding of how laryngeal contrast should be phonologically represented.

In this chapter, I present the first empirical study of intervocalic stop voicing in Danish. The study is based on the existing DanPASS corpus (Danish Phonetically Annotated Spontaneous Speech; Grønnum 2009, 2016; see Section 4.5.1 below). In Section 2.3, I summarize a number of observations about the production of Danish stops which make intervocalic voicing patterns difficult to predict. The unaspirated stops /b d g/ are produced with longer closure duration and greater muscular tension than the aspirated stops /p t k/ (Fischer-Jørgensen 1954; Fischer-Jørgensen and Hirose 1974; see Section 2.3.5). From this perspective, /p t k/ should actually be more conducive to voicing. It is sometimes claimed that these differences are too small to be of significance (e.g. Grønnum 2005), and that both sets of stops are phonetically lenis, which suggests that both sets are equally likely to be voiced intervocalically. On the other hand, glottographic and electromyographic (EMG) investigations have shown that both stop types are characterized by a glottal opening gesture during the closure when produced between vowels in careful speech, and that this gesture lasts longer and is of greater magnitude in /p t k/ (Frøkjær-Jensen et al. 1971; Fischer-Jørgensen and Hirose 1974; Hutters 1985; see Section 2.3.3). This suggests that /b d g/ should be most conducive to voicing, but also that voicing is actively blocked in both sets.

The results show that intervocalic voicing is very rare in /p t k/. Although much more frequent in /b d g/, intervocalic voicing occurs in less than half of /b d g/ tokens. This rarity of intervocalic voicing is in essence the opposite conundrum of the one I mentioned at the beginning of the chapter. Voicing is natural in this position, so its rarity can only be accounted for with reference to some mechanism that blocks it. The occurrence of intervocalic voicing is generally correlated with other variables that are associated with phonetic lenition; it occurs more frequently in quick speech, in morphological affixes, before neutral vowels, and in unstressed syllables. This suggests that intervocalic voicing is in itself lenition phenomenon. I suggest that this lenition is best modeled as gesture reduction: Danish has phonologized glottal spreading gestures in all stops, which usually block voicing; however, in environments that are generally prone to lenition, this gesture is lost at a relatively high rate. This happens more frequently in /b d g/, where the gesture has less of a critical function.

The chapter is structured as follows: Section 4.2 provides some background on closure voicing in phonetics and phonology with particular focus on intervocalic position, and Section 4.3 briefly recaps some important facts about the phonetics and phonology of Danish stops from Chapter 2. In Section 4.4, I summarize the research questions and motivate all independent variables of the study. In Section 4.5, I provide an overview of the methods: I introduce the corpus used for the study, and explain how the data were processed. In Section 4.6, I provide an exploratory analysis of the data. In Section 4.7, I describe the selection of a logistic mixed-effects regression model, and give the results of that model. In Section 4.8, I discuss the research questions in light of the results, and in Section 4.9, I briefly summarize the chapter.

4.2 Closure voicing and [voice]

I already alluded to the phonetic pressure against closure voicing in Section 3.5. Here, I will cover the relevant literature in more detail, and discuss the role that this pressure plays in discussions of the underlying representation of laryngeal contrast.

Closure voicing in stops is "unnatural" (Ohala 1983a). A sufficient transglottal pressure differential is required to maintain vocal fold vibration. As air continually flows from the lungs while both the oral and nasal cavities are sealed, supraglottal air pressure rises quickly. This effectively means that it is impossible to maintain closure voicing for a long duration of time. If the size of the supraglottal cavity remained constant during a stop closure, it would only be possibility to maintain a sufficient transglottal pressure differential for roughly 5–10 ms (Ohala and Riordan 1979). This is not actually the case, though; the vocal tract automatically enlarges, primarily due to compliance of the soft tissue making up the inner walls of the cavity. This should allow for approximately 60–70 ms of closure voicing for a male speaker (Westbury 1983), varying depending on e.g. the point of occlusion. Voicing is maintained longest for a fronted occlusion (e.g. bilabial) due to the large cavity between the glottis and the supraglottal occlusion, which yields a slower build-up of air pressure and crucially yields a larger total area of soft, expandable cavity walls. Keating (1984b) shows that bilabial stops naturally retain voicing for roughly 30% longer than velar stops. When some languages show yet longer closure voicing, it is due to active vocal tract enlargement during the occlusion, implemented by e.g. lowering the jaw or raising the velum.

Westbury and Keating (1986) investigate the issue of articulatory naturalness in detail, using a model of breath-stream control with the vocal folds appropriately adducted and tensed for voicing. They show that syllable-initial closure voicing is articulatorily unnatural, since subglottal and supraglottal air pressure rises roughly synchronously, unless the vocal folds are initially fully abducted to allow for preparatory build-up of subglottal air pressure. Closure voicing is also unnatural syllable-finally; Westbury and Keating hypothesize that this is due to an inspiratory force that gradually but quickly counteracts the initially high subglottal pressure from the preceding vowel. However, it is natural for intervocalic stop closures to be voiced throughout most of their duration due to the high initial subglottal pressure following the preceding vowel.

Articulatory naturalness sometimes translates directly into typological patterns in phonology, but this is not always the case. In accordance with articulatory naturalness, there is a strong implicational hierarchy regarding voiced stops in phonological inventories. In almost all cases, languages with voiced stops also have voiceless stops (e.g. Ohala 1983a; Maddieson 1984). Furthermore, final obstruent devoicing is a very common typological pattern, partially because syllable-final segments tend to be lengthened, resulting in longer stretches of voicelessness in coda stops, and as such a lesser chance of closure voicing being interpreted as an important phonological cue (e.g. Blevins 2004: 103ff.). However, in spite of their unnatural status, syllable-initial voiced stops are actually quite common. Furthermore, voicing is most natural in medial position, but languages with no laryngeal distinction in stops generally have voiceless stops in all positions (Keating et al. 1983). This illustrates an important point: there is more to phonological patterns than ease of articulation.

Below, I characterize three approaches to the representation of laryngeal contrasts in the phonological literature, as well as the predictions they make with regards to intervocalic closure voicing. Some of these were also introduced in Section 2.4.3.1, but the discussion here will be somewhat broader, and not limited to the representation of contrast in Danish. There is a huge literature on the representation of laryngeal contrast, so some approaches will necessarily be missed here, while others may be grouped together even if they differ in some respect. I will refer to these approaches as 'concrete [voice]' approaches, 'abstract [voice]' approaches, and 'gesture-based' approaches.

The phonological feature [voice] has been conceptualized in different ways. It sometimes refers narrowly to the presence of voicing during closure, which is what I refer to as concrete [voice]. This is how [voice] is conceptualized in the laryngeal feature geometry of Lombardi (1995), the 'laryngeal realism' approach of Iverson and Salmons (e.g. 1995), and the 'laryngeal dimensions' model of Avery and Idsardi (2001). These are approaches that assume a direct relationship between physical laryngeal constellations and phonological laryngeal features. Such approaches usually assume that languages with aspiration-based contrasts employ an active feature like [spread glottis] or [glottal width] to distinguish laryngeal contrasts. It is common to assume that sonorant sounds are unmarked for [voice], since vocal fold vibration is the natural state of affairs in these sounds (Lombardi 1995). This creates a problem in determining the phonological origin of intervocalic closure voicing; surrounding vowels are unmarked for [voice], so voicing cannot spread from those. One possible solution to this is a non-laryngeal [spontaneous voice] feature node, which can spread from sonorants to obstruents, as proposed by Rice and Avery's (1989). Another solution is to simply relegate intervocalic voicing to phonetic implementation, placing it outside the purview of phonology. This would predict that laryngeally unmarked intervocalic stops always follow the phonetically natural pattern.

Jessen and Ringen (2002) and Beckman et al. (2013) argue that the intervocalic behavior of stops is relevant for determining whether [voice] or [spread glottis] are active in a language. Beckman et al.

show that Russian /b d g/ are voiced throughout the closure in intervocalic position with very few exceptions, while German /b d g/ are variably voiced in intervocalic position (roughly 60% of tokens are voiced throughout). They take the consistent voicing in Russian as evidence for an active [voice] feature, and the variable voicing in German as evidence for a gradient phonetic process of passive voicing. Following Chomsky and Halle (1968), they assume that at some stage in the phonological derivation, segments are assigned numerically-valued features; the degree of intervocalic voicing in a [spread glottis] language will depend on the value assigned to [spread glottis] at this late stage in the derivation.[1] The findings of Beckman et al. can only be taken as evidence for underlying features if one assumes a transparent relationship between phonology and phonetic implementation; see e.g. Keating (1984a) for a general critique of this stance.

In abstract [voice] approaches, [voice] in stops does not necessarily imply closure voicing. Chomsky and Halle (1968) and Keating (1984a) both assume that [voice] can be cued with either closure voicing or voicing onset approximately at the time of release, depending on which contrast the language in question employs. Kingston and Diehl (1994) similarly assume a feature [voice] that does not always entail closure voicing in stops. This argument partially relies on the finding that [voice]-induced F_0-perturbations behave similarly regardless of how the feature is phonetically implemented. In their account, the feature [voice] lowers F_0 on the following vowel.[2] Kingston and Diehl recognize that there is a discrepancy between initial and intervocalic position when it comes to the naturalness of closure voicing; they argue that an 'automatic phonetics' will output voiceless initial stops and voiced intervocalic stops, and that a 'controlled phonetics' is necessary to divert from that pattern.

If we assume a relationship between phonology and phonetics, then there should be a correspondence between which patterns are unmarked in phonology and phonetics. Given the aerodynamic account of stop voicing presented here, this entails that an unmarked

[1] See Kirby and Ladd (2018) for a critical discussion of the predictions that follow from this account, in particular as relates to laryngeally induced F_0-perturbations.

[2] Recall from Section 2.4.3.1 that the cause of this pattern is disputed; Hanson (2009) argues that F_0 is raised locally by voiceless stops.

stop should be voiceless initially and voiced intervocalically. It also entails that phonetic reduction is positionally defined: devoicing of [voice] stops is lenition syllable-initially and syllable-finally, whereas voicing of stops without [voice] is lenition intervocalically.[3] This is difficult to account for in a feature-based framework but seems to hold up for intervocalic position, where voicing of stops without underlying [voice] is crosslinguistically common (Kaplan 2010).[4] In an optimality-theoretic analysis of this problem, Smith (2008) proposes constraints that militate against voiced obstruents in onset position and voiceless obstruents in intervocalic position, which compete with faithfulness constraints (see also Hayes 1999). Katz (2016) points out some typological shortcomings of this account: Smith's account predicts languages which neutralize a laryngeal contrast in initial position due to devoicing, but no such language is attested. Likewise, Smith's account predicts languages where a laryngeal contrast is systematically neutralized intervocalically due to voicing; such languages are also surprisingly rare, and Katz proposes reanalyses of the attested languages.

Gesture-based approaches to phonological representation can straightforwardly account for these positional markedness relations. One such approach is Articulatory Phonology (Browman and Goldstein 1986, 1992). In Articulatory Phonology, articulatory gestures are taken as the primary units of phonological representation rather than segments or features. A consequence is that the duration and magnitude of glottal gestures can be represented separately from other gestures that make up traditional segments. The unmarked state of the glottis is adducted and tensed, which will not cause voicing initially but will cause voicing intervocalically, as discussed above.

Some predictions about the patterning of intervocalic closure voicing follow from these different conceptualizations of laryngeal representation. In concrete [voice] approaches, closure voicing is a necessary and sufficient criterion for [voice] and a different feature like [spread glottis] is needed to account for aspiration. Concrete

[3]See Steriade (2009) for a discussion of positional markedness and laryngeal contrasts focusing on final position.

[4]It does not, however, hold up for initial position; there is no common process of initial devoicing, at least not resulting in positional neutralization (Katz 2016).

[voice] approaches predict essentially categorical voicing of all intervocalic stops in 'true voice' languages; [voice] ensures voicing in one set of stops, and there are no available phonological mechanisms to counteract voicing in the other (unmarked) set of stops.[5] In 'aspiration languages', we predict varying degrees of intervocalic voicing of unmarked stops, and very little voicing in [spread glottis] stops (following Beckman et al. 2013). In abstract [voice] approaches, where [voice] can have different phonetic interpretations, it is less straightforward to predict intervocalic behavior. Following Kingston and Diehl (1994), a 'controlled phonetics' is necessary to divert from the natural pattern of intervocalic voicing. A gesture-based approach such as Articulatory Phonology also predicts the natural pattern of intervocalic stop voicing if no underlying glottal gestures are present; however, Articulatory Phonology allows a great deal of flexibility in how glottal gestures are represented, making it a very powerful representational framework.

4.3 Predictions for Danish

In this section, I will recap some relevant phonetic facts and phonological arguments concerning voicing and laryngeal representation in Danish stops, and discuss how they relate to the predictions posed in the preceding section. In the following, when I refer to /b d ɡ/ and /p t k/, I refer to the surface contrast: /b d ɡ/ are stops that would be unaspirated in distinct speech, and /p t k/ are stops that would be aspirated in distinct speech. I will refer to the two series as *laryngeal categories.*

I discussed in Section 2.3.5 how Fischer-Jørgensen (1972b) has argued on the basis of closure duration and EMG studies that /b d ɡ/ are fortis, and /p t k/ are lenis. It has later been argued (by e.g. Grønnum 2005) that the difference in closure duration and articulatory tension between the laryngeal categories is insignificant, and both are lenis (see Section 2.3.5). This has affected transcription practice, such that /b d ɡ/ are usually narrowly transcribed as [b̥ d̥ ɡ̊] and /p t k/ as [b̥ʰ d̥ˢ ɡ̊ʰ].

[5]Recall that languages displaying this pattern are actually quite rare (Katz 2016).

The terms fortis and lenis are used in distinct ways in the phonetic and phonological literature. One use is as an arbitrary label for stop contrasts in languages which do not depend on closure voicing. Fortis–lenis has often been used in this sense when discussing Germanic languages, where the historic voiced–voiceless distinction has a diverse set of phonetic reflexes in the modern languages (Kohler 1984; Henton et al. 1992). Another use of fortis–lenis is as a phonetically substantial phonological feature referring to force of articulation. This feature may correlate with pulmonic force, muscular tension, closure duration, or indeed closure voicing (see Jaeger 1983 and references therein). Either use of the terminology is usually too imprecise for phonetic or phonological description.

Articulatory studies of carefully read speech have shown that intervocalically before stressed syllables, /b p/ are both accompanied by a glottal opening gesture in Danish (Frøkjær-Jensen et al. 1971), and EMG studies confirm that the posterior crico-arytenoid muscles are active in achieving this gesture (Fischer-Jørgensen and Hirose 1974; Hutters 1985); see Section 2.3.3 for more details. Similar studies of English found no such gesture during /b/ (Sawashima 1970; Hirose and Gay 1972); in Icelandic, which has a contrast between unaspirated and pre-aspirated stops in intervocalic position, both series of stops have a significant glottal spreading gesture (Pétursson 1976). Hutters (1985) proposes that the glottal spreading gesture in Danish is a measure taken to reinforce voicelessness in /b/, although she does not resolve this question; Möbius (2004) has shown that a glottal spreading gesture maintains voicelessness in German intervocalic stops, and Pape and Jesus (2014) have shown the same for European Portuguese and Italian.

Iverson and Salmons (1995) and Basbøll (2005) assume that the laryngeal contrast in Danish is managed with [spread glottis] (see Section 2.4.3.1). Motivation for this comes from the process of progressive sonorant devoicing, which has recently been shown by Juul et al. (2019) to be much less categorical than usually assumed (see Section 2.3.1 and Puggaard-Rode et al. forthc.). Kingston and Diehl (1994) assume that Danish stops are distinguished by an abstract [voice] feature. Motivation for this comes from the finding that /b d g/ trigger local F_0-lowering. Recall from Section 2.3.3 that Petersen (1983) does indeed find such an effect, but he crucially also finds that both laryngeal series trigger local F_0-raising relative to nasals. As Goldstein and Browman (1986) point out, this is consistent with an account where F_0-perturbations are the direct result of glottal aperture, something that Kingston and Diehl (1994) explicitly reject. Nevertheless, Kingston and Diehl's dichotomy between automatic and controlled phonetics can potentially account for both the presence and absence of closure voicing in Danish intervocalic stops.

As mentioned in Section 4.1 above, a number of facts about Danish stops make it difficult to predict the relative likelihood of intervocalic voicing. First of all, most of the relevant literature assumes that intervocalic voicing is categorical or near-categorical. Muscular tension is overall low in Danish stops, which should increase the chances of voicing, but all Danish stops are also accompanied by a glottal spreading gesture, which should decrease the chances of voicing. Closure duration is shorter and muscular tension weaker in the production of /p t k/ relative to /b d g/, but the glottal spreading gesture in /p t k/ has a greater magnitude and longer duration.

The results of this study allow us to compare some of the predictions from different approaches to phonological laryngeal representation, as laid out in Section 4.2. If [spread glottis] is indeed the only active laryngeal feature in Danish, we would predict at most variable voicing in /b d g/, and very negligible voicing in /p t k/ (following Beckman et al. 2013). If the laryngeal contrast is maintained with phonologized glottal gestures (as in Articulatory Phonology), it follows that the two series have underlying glottal spreading gestures of different magnitudes, both of which are expected to counteract voicing. A gestural account predicts that lenition leads to a reduction in the magnitude of these gestures, potentially causing voicing in both laryngeal series, but more readily in /b d g/. Neither of the two featural accounts (i.e. the abstract and concrete [voice] approaches) make any clear predictions about lenition and voicing.

4.4 Research questions and predictors

I set out with the following research questions in mind:

(1) Is there a difference in how frequently members of the two laryngeal series are voiced intervocalically?

The known facts about Danish stop production point in different directions. If the vocal folds were in a neutral, adducted position during the closure, one would expect a higher likelihood of continuous closure voicing in /p t k/, since they have shorter closure duration and a lower degree of articulatory force (Fischer-Jørgensen 1972b). However, for both series of stops, although to varying degrees, the vocal folds are actively spread during the closure. As such, I hypothesize that /b d g/ are voiced more frequently than /p t k/. This seems intuitively obvious and is explicitly predicted from both a concrete [voice] account and a gesture-based account of the contrast.

(2) Is closure voicing in intervocalic stops a lenition phenomenon?

From an aerodynamic perspective, voicing is natural in intervocalic stops, but there is evidence that voicing is actively blocked in all Danish stops. I test whether intervocalic voicing is more common in environments where lenition is generally expected, which would be predicted from gesture-based underlying representations.

(3) What factors predict closure voicing, and how large are their relative effects?

In addition to phonological laryngeal category and lenition, a host of other phonetic and extraphonetic factors are known to or can be expected to affect the probabilistic occurrence of consonant voicing (as established by e.g. Shih and Möbius 1998; Möbius 2004; Strycharczuk 2012). I aim to take as many of these as possible into account in order to explore their relative influence in Danish. These factors are presented in detail below.

4.4.1 Potential predictors

The detailed annotations of the DanPASS corpus (see Section 4.5.1) make it possible to test how a large number of (mostly categorical) predictors affect the rate of closure voicing. These predictors relate to segmental, prosodic, morphosyntactic, and other factors, which are discussed in the following sections.

4.4.1.1 Segmental predictors

The stops are coded according to LARYNGEAL CATEGORY and PLACE OF ARTICULATION. There is really no theory-neutral way to refer to the two laryngeal categories. Here, the terms 'aspirated' and 'unaspirated' are used as short-hand terms for the surface contrast between /p t k/ and /b d g/ in distinct speech, as mentioned in Section 4.3.

Place of articulation is expected to influence the likelihood of voicing, such that the chances of voicing are lower in occlusions further back in the oral cavity. This is aerodynamically motivated (see Section 4.2 for more details), and is reflected typologically: voiced velar stops are less common than alveolar ones, which are in turn less common than bilabial ones (Gamkrelidze 1975; Ohala 1992; Hayes and Steriade 2004; Brown 2006). Since bilabial and alveolar occlusions are physically quite close, and velar occlusions are significantly further back, a place effect should be most noticeable in velar stops.

The quality of surrounding vowels is expected to influence the likelihood of closure voicing; recall that Danish has an exceptionally complex vowel system (see Section 1.4). HIGH VOWELS are expected to decrease the chances of voicing, since high vowels have a tighter constriction in the oral cavity, making them less sonorous and more likely to devoice (e.g. Mortensen 2012). High vowel devoicing is caused by a constriction in the oral cavity, which makes it difficult to maintain voicing over time due to rising supraglottal air pressure. A preceding high vowel should decrease the chances of voicing more than a following high vowel. The following are considered high vowels: [i y ɪ ʏ e ø u ʊ o]. Note that these transcriptions are adapted to Danish (see Section 1.4); many of these vowels are higher than their conventional IPA counterparts, and they all have a mean $F1$ < 400 Hz in Modern Standard Danish (Juul et al. 2016).

In locating intervocalic stops, SEMIVOWELS were also considered vowels. I assume that semivowels occurring immediately before the intervocalic stop decrease the chances of voicing, simply because these are less sonorous than nuclear vowels (e.g. Parker 2002). The sounds in question are [ɪ̯ ʊ̯ ɐ̯ ɤ̯], as well as their syllabic counterparts [ɪ ʊ ɐ ɤ].

As discussed above, intervocalic voicing in Danish may be a lenition phenomenon resulting from passive voicing lasting throughout the closure. Therefore, voicing is expected to be more likely in environments that are associated with lenition. Surrounding NEUTRAL VOWELS should increase the chances of voicing, since the Danish neutral vowels [ə ɐ] generally occur in prosodically weak syllables (e.g. Basbøll 2005), where lenition is strongly expected. Vowel neutrality is strongly negatively correlated with stress: as a general rule, syllables with neutral vowels are always unstressed, but not all unstressed syllables have a neutral vowel. Preceding and following neutral vowels were coded separately, but their influence on closure voicing is expected to be essentially the same.

4.4.1.2 Prosodic predictors

STRESS generally reduces the chances of lenition phenomena occurring, so it is also expected to reduce the chances of closure voicing. If the preceding syllable is stressed, this is expected to increase the chances of voicing, as it is unlikely for two syllables in a row to be stressed. These two variables are coded independently, and one is not predictable from the other; there may be multiple consecutive unstressed syllables.

Adjacent STØD is expected to reduce the chances of voicing, no matter whether on the preceding syllable or the syllable in question.[6] Stød is phonetically akin to creaky voice; it is cued with low pitch and relatively aperiodic voicing during the final part of a long sonorant rhyme (Grønnum and Basbøll 2001, 2007). Stød is produced with laryngeal contraction regulated by vocalis and lateral crico-arytenoid

[6] At the morpheme level, primary stress is a phonological prerequisite for stød. In compounds, however, primary stress generally falls on the first member, while the second member has stød; some derivational processes also behave this way, and stød interacts with inflectional morphology in complex ways (e.g. Basbøll 2003). Furthermore, morpheme level stress is not necessarily realized at the sentence level, and morphemes can lose stress at the sentence level while retaining stød. As such, stød and stress are far from perfectly correlated in the DanPASS corpus; in fact, a small majority of syllables with stød are unstressed.

muscles (Fischer-Jørgensen 1987, 1989). Recall from Westbury and Keating (1986; see Section 4.2) that closure voicing is natural intervocalically, assuming that the vocal fold configuration is amenable to voicing; this is the case for vowels with modal voicing and less so for vowels with stød. As such, stød on the preceding syllable is expected to decrease the chances of continuous voicing. Although stød mainly affects the final part of syllables, it is also cued syllable-initially with many of the same articulatory and acoustic correlates as stress: increased airflow, pharyngeal pressure, intensity, pitch, and articulatory force (Smith 1944; Fischer-Jørgensen 1987, 1989). As such, stød on the syllable itself is also expected to decrease the chances of continuous voicing, but less so than stød on the preceding syllable.

4.4.1.3 Morphosyntactic predictors

The type of MORPHOLOGICAL BOUNDARY at which the intervocalic stop occurs was coded. These include word boundaries, boundaries between roots and (derivational and inflectional) affixes, boundaries between separate parts of compounds, as well as *none*, if the intervocalic stop occurred morpheme-internally. Prefixes in Danish are exclusively derivational; suffixes are mostly inflectional, but can also be derivational. As consonants tend to be strong domain-initially (e.g. Keating et al. 2004), it would be more optimal to code the individual syllables for their position in a prosodic hierarchy (e.g. Nespor and Vogel 1986), but such a coding cannot be easily extracted from the existing DanPASS transcriptions. I hypothesize that the morphological boundary predictor is hierarchical in its influence on closure voicing, as it has been shown that intergestural articulatory timing is more stable within-word and within-morpheme than across words and morphemes (Byrd et al. 2000; Cho 2001). I therefore assume that morpheme-internal stops have higher likelihood of voicing than word-internal stops at morphological boundaries; and these in turn have higher likelihood of voicing than stops at word boundaries. Among morphological boundaries, the following hierarchy of morpheme boundary types is assumed: inflectional > derivational > compound.[7]

There are reasons to assume that stops at inflectional morpheme boundaries might be voiced at much higher rates than stops in other positions: inflectional suffixes are always unstressed and always have neutral vowels in Danish. Following a usage-based framework such as Exemplar Theory (e.g. Bybee 2001), stops in inflectional affixes may also be voiced more often simply because language users encounter voicing more often in affixes, and as such it is weighted higher in the underlying representation of affixes at a morpheme-specific level. Several phonological frameworks assume that morphology is invisible to phonetic interpretation and would thus predict that morpheme-specific underlying representations are impossible; this is the case in e.g. Lexical Phonology (Kiparsky 1985). However, recent studies show that specific morphemes can exhibit phonetic patterns that are not predictable from their phonemic makeup. Plag et al. (2017; see also Tomaschek et al. 2021) have found that the English 'homophonous' *s*-suffixes (third person singular present tense, plural, etc.) differ systematically in phonetic realization; consider e.g. the suffixes in the present tense verb [pʰɛts] 'pet+s' and in the plural noun [pʰɛts] 'pet+s'. The present tense ending is systematically longer than the plural ending. Heegård (2013) found similar results for variable rates of schwa-deletion in homophonous [-tə] *-te* suffixes in Danish; [peˈstɛmˀtə] *bestemte* can either be a past tense verb meaning 'decided' or a definite past participle meaning 'certain', both of which derive from the verb [peˈstɛmˀm̩] *bestemme* 'decide'; schwa-deletion is systematically more likely in the past tense form, meaning the two are only near-homophonous.

Additionally, words were coded as being members of either a closed or an open word class. Words from closed classes are often function words, and these often show significant phonetic reduction (e.g. Bell et al. 2003; Schachtenhaufen 2013).

[7] An alternative approach would be to use Basbøll's (2005: 351) complex hierarchy of graded productivity of morphological endings. However, Basbøll's hierarchy only covers inflectional endings, and the added complexity of Basbøll's hierarchy would potentially make the statistical results very difficult to interpret.

4.4.1.4 Other predictors

In addition to the predictors mentioned above, a lexical frequency measure was included. High lexical frequency is known to cause phonetic reduction, both in the course of language change (e.g. Hooper 1976; Bybee 2000a) and synchronically (Bybee 2000b; Pierrehumbert 2001; Pluymaekers et al. 2005), and has been shown to specifically increase voicing assimilation in Dutch (Ernestus et al. 2006). Although the speech in DanPASS is spontaneous, it is also nested within specific experiments, where lexical frequencies can differ substantially from the ambient language (see Section 4.5.1). Since contextual probability has also been shown to increase phonetic reduction (Jurafsky et al. 2001, 2002), LOCAL LEXICAL FREQUENCY was coded, i.e. lexical frequency in the DanPASS corpus itself. This measure is available in the online version of DanPASS (Grønnum 2016). This was compared to a more general measure of lexical frequency based on the much larger LANCHART corpus (which was also used in Section 3.3.3). Due to the experimental nature of DanPASS (the map task in particular, see Section 4.5.1), many frequent words in DanPASS do not occur in LANCHART at all. This means that modeling with general (LANCHART) frequencies rather than local (DanPASS) frequencies would require excluding just over 300 items, around 8% of the total number of tokens, particularly in the morphological compound category. The two frequency measures are further strongly correlated (r = .78). Given the strong correlation and the disadvantages of using general frequencies, only local frequency is used in the statistical modeling.

INDIVIDUAL WORDS were also coded, since Pierrehumbert (2002) mentions a number of cases where word-specific phonetic encoding goes beyond simple lexical frequency and contextual predictability; this relates directly to the discussion of Exemplar Theory in Section 4.4.1.3. Word-specific effects are not explored in any detail.

A local measure of speech rate was also included. LOCAL SPEECH RATE should affect the chances of voicing for aerodynamic reasons: unless inhibited, post-vocalic voicing should automatically continue for a certain amount of time during a stop closure (see Section 4.2). A higher speech rate also causes decreased closure duration (as demonstrated for Danish by Andersen 1981a), which increases the chances that voicing continues throughout the closure phase. Local speech rate is measured here as the combined duration in seconds of the two syllables flanking the intervocalic stop.[8]

Finally, a few extralinguistic factors pertaining to the speakers were coded. SEX has been shown to influence the degree of closure voicing, such that men are more likely than women to produce fully voiced stops (Swartz 1992; Ryalls et al. 1997). This could be aerodynamically motivated; on average, men have larger supralaryngeal cavities than women, causing a slower rise in supraglottal air pressure during closure. An alternative explanation for the same outcome is that women generally speak more 'clearly' than men (as demonstrated by e.g. Ferguson 2004 for vowel intelligibility), and show less of a tendency for lenition.[9] I am not aware of studies connecting AGE with closure voicing directly, but it has been shown that speech rate decreases with age (Seifert 2009), suggesting that lenition will also decrease with age.

Finally, the INDIVIDUAL SPEAKERS were coded. Sonderegger et al. (2020) recently showed that the implementation of closure voicing in Glasgow Scots is highly speaker-specific even when controlling for a large number of other factors, and Tanner et al. (2020) found similar results for Japanese. Speaker-specific effects are not explored in any detail.

The potential predictors and the directionality of their expected influence on closure voicing are summarized in Table 4.1.

---

[8] This is admittedly a rough measure of speech rate, chosen mostly out of convenience (it was easy to extract from the existing data). It is not unheard of, though; Bohn (2013) also measures duration of target syllables in his study comparing Danish infant directed speech and adult directed speech. The measure is presumably not too rough for the purpose of the current study, as the result show a very strong effect of speech rate.

[9] The latter explanation is almost certainly an effect of gender rather than biological sex.

Table 4.1: *Potential predictors and the expected directionality of their influence on closure voicing.*

Variable	**Predicted likelihood of voicing**	**Notes**
LARYNGEAL CATEGORY	unaspirated > aspirated	
PLACE OF ARTICULATION	bilabial > alveolar > velar	strongest effect for velar stops
ADJACENT SEMIVOWEL	decreased	
ADJACENT HIGH VOWEL	decreased	strongest effect preceding the stop
ADJACENT NEUTRAL VOWEL	increased	
STRESS	unstressed > stressed	
PRECEDING STRESS	stressed > unstressed	
ADJACENT STØD	decreased	strongest effect preceding the stop
MORPHOLOGICAL BOUNDARY	internal (no boundary) > inflectional > derivation > compound > word	
WORD CLASS TYPE	closed > open	
LOCAL LEXICAL FREQUENCY	increases with frequency	
LOCAL SPEECH RATE	increases with speech rate	
LEXICAL ITEM	random	
SEX	men > women	
AGE	decreased with age	
INDIVIDUAL SPEAKER	random	

4.5 Methods

4.5.1 The DanPASS corpus

The DanPASS corpus (Grønnum 2009, 2016) is used to answer the research questions posed in the previous section. The corpus consists of native speakers of Danish solving a number of unscripted tasks, either alone or in pairs. An original motivation behind creating the corpus was to counteract the bias for highly controlled scripted speech in studies of Danish phonetics. The recordings in the DanPASS corpus are unquestionably also laboratory speech, but they are much more spontaneous than what was previously the standard. Grønnum (2009) rightly points out that formal laboratory speech is well-suited for some phonetic studies; some phenomena are rare enough that they can be difficult to find sufficient examples of even in very large corpora, and sometimes it can be important to carefully control for interacting phenomena. For a phenomenon such as intervocalic closure voicing, where the triggering environment is very frequent, and informal speech can perhaps in itself be considered a triggering environment, basing the analysis on informal speech is crucial. The corpus has already been the basis for major contributions to our understanding of Danish speech, in the areas of consonant reduction (Pharao 2009), and phonetic reduction in general (Schachtenhaufen 2013), as well as intonation and prosody (Tøndering 2003, 2008; Grønnum and Tøndering 2007). As mentioned in Section 2.3.1, it is also basis for the most extensive investigation of voice onset time in Danish stops (Mortensen and Tøndering 2013).

The full DanPASS corpus consists of monologues recorded in 1996, and dialogues recorded in 2004. While the dialogues probably constitute a more natural speech setting, they are also somewhat more challenging to analyze. For this reason, the current study only makes use of the monologues. Monologues were recorded from 18 speakers: 13 men and 5 women. The speakers were between 20–68 years old, with a mean age of 29. Overall, the monologues constitute 171 minutes of speech, with a mean duration of $9m27s$ of speech per speaker (range $6m13s$–$15m49s$). Technical details are reported by Grønnum (2009). The speakers were recorded performing three different tasks. *Network* is a description of various shapes and colors, based on a design by Swerts and Collier (1992). *City* is a description of a number of routes through a drawn city map, based on a design by Swerts (1994). *House* is a description of how to build a house model using a number of buildings blocks, based on a design by Terken (1984).

The recordings are accompanied by quite detailed annotations in Praat (Boersma 2001; Boersma and Weenink 2021). Segmentations are made at the levels of prosodic phrase, word, and syllable, whenever these could be segmented with reasonable certainty. While the syllable is a well-defined phonological unit in Danish (see e.g. Basbøll 2005), it is often difficult to find well-defined phonetic units corresponding to those (Schachtenhaufen 2010a, 2010b), particularly because syllabic sonorants and consecutive syllables consisting of homorganic vowels are abundant in Danish speech due to schwa assimilation. The recordings are annotated orthographically, phonemically, and phonetically. They are also coded for morphology and accompanied with parts-of-speech tags and annotations for pitch movements and stress. The phonetic transcriptions use Grønnum's (1998) standards for transcribing Danish and are generally rather narrow except where stops are concerned. [p t k] are used where aspirated stops would be expected in distinct speech, and [b d g] where unaspirated stops would be expected in distinct speech, regardless of phonetic implementation.[10] In other words, closure voicing during stops is ignored in the transcriptions. Grønnum (2009) does not motivate this, and perhaps as a result, previous studies using DanPASS to examine e.g. stops (Pharao 2009, 2011; Mortensen and Tøndering 2013; Schachtenhaufen 2013) have also ignored the distinction between stops with and without closure voicing.

4.5.2 Acoustic analysis

I used a Praat script (written by Dirk Jan Vet; see Puggaard-Rode et al. 2022b) to locate intervocalic stops in the DanPASS monologues, i.e. stops that do not occur initially in a prosodic phrase and are flanked on both sides by (semi-)vowels. Approximants were included because there are well-defined phonological processes whereby they syllabify and become phonetic nuclear vowels (Brink and Lund 1975; Basbøll 2005; Schachtenhaufen 2010b). The DanPASS transcriptions are segmented at the level of phonetic syllable, and only stop-initial intervals were included in the study. This is in line with the studies of glottal activity cited in Sections 2.3.3 and 4.3 above, which also focus on syllable-initial intervocalic stops. This decision makes the coding and interpretation of predictor variables considerably easier. For each stop, the surrounding syllables are isolated, i.e. the stop-initial syllable and the preceding syllable. The script creates a sound file and TextGrid file containing all such syllables from the DanPASS monologues. This sound file lasts just under 24 minutes and contains a total of 3,744 inter-

[10] Except in the case of tapped realizations of /t ~ d/, which are transcribed with with [ɾ].

Table 4.2: *Intervocalic stops in the DanPASS monologues by phoneme.*

Phoneme	Number	By-speaker range
/b/	189	3–25
/d/	1,278	28–167
/g/	752	26–65
/p/	327	8–32
/t/	431	16–39
/k/	767	24–67
Total	3,744	117–341

vocalic stops, with an average of 204.7 stops per speaker (range 117–341). They are broken down by phoneme in Table 4.2.

The variance in lexical frequency is rather extreme. There are 303 unique lexical items in the data, with an average of 12.4 observations per item, albeit a median of just 2 observations (range 1–303). 125 lexical items occur just once, while the 10 most frequent items occur a total of 2,025 times, i.e. they make up more than half of all tokens.

For each of the intervocalic stops, I manually checked if it was voiced throughout the closure. This was done by visually inspecting the waveform: constant periodicity up to the burst was taken as continuous closure voicing.[11] This method was relatively straightforward to implement, although it is certainly a simplification of the complexity in the phonetic signal. Figure 4.1 shows waveforms of stops from both laryngeal series that show continuous voicing and interrupted voicing, respectively.

Recent comparable studies by e.g. Davidson (2016), Sonderegger et al. (2020) and Tanner et al. (2020) all use a three-way distinction between *voiceless*, *partially voiced*, and *fully voiced*. However, none of these studies focus particularly on intervocalic stops.[12] There are two main reasons for not adopting a three-way distinction in this study: 1) Multi-valued categorical dependent variables are much more

[11]Whenever stops from the /p t k/ series were fully voiced, they typically also had breathy voiced release. As a result, intervocalic voicing is very unlikely to lead to neutralization.

[12]Davidson (2016) focuses exclusively on phrase-medial position, so it is likely that many stops in that study were also intervocalic.

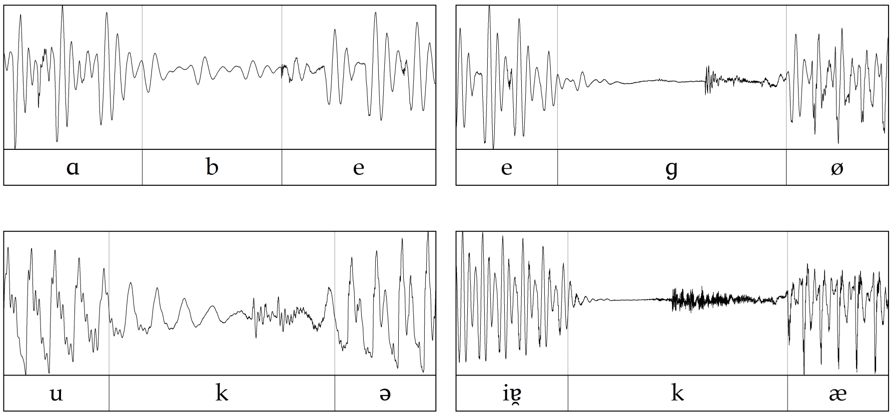


Figure 4.1: *Waveforms exemplifying: 1) A fully voiced token of* /b/ *in the phrase* [fʁɑ beˈkønˀl̩sn̩] fr(a be)gyndelsen *'from the start'. 2) A mostly voiceless token of* /g/ *in the same phrase as 1,* <fra b(egy)ndelsen>. *3) A fully voiced token of* /k/ *from the phrase* [tu gʱə] d(u ka)n *'you can'. 4) A mostly voiceless token of* /k/ *from the word* [fiɐ̯ˈkʰænˀt] f(irka)nt *'square'.*

difficult to model than binary variables, and 2) fully voiceless intervocalic stops are known to be uncommon, so a *voiceless* category would likely have added little explanatory value. In intervocalic position, passive voicing essentially affects the first part of the following closure, regardless of the laryngeal category of the stop. This has been shown for at least Standard Chinese (Shih and Möbius 1998), German (Möbius 2004), and American English (Davidson 2016) aspirated stops, and for voiceless stops in several other languages (Shih et al. 1999). In an unpublished conference paper, Puggaard et al. (2019) showed that in Danish carefully read lab speech, there was no significant difference between the two laryngeal series in the relative duration of voicing during closure in stressed intervocalic position; both /b/ and /p/ were voiced for approximately one fourth of their closure. We compared this to Dutch, a so-called 'true voicing language', where the majority of intervocalic /b/ tokens were voiced throughout their closure duration.

This was a further argument for considering 'fully voiced' to only mean continuous voicing throughout the closure here.

Ideally, I would be working with a continuous measure of closure voicing, possibly measuring both intensity and relative duration of voicing. It is quite possible that true effects of some lower-level variables on voicing are masked in this study because of the relatively rough voicing measure. However, this would require much more fine-grained segmentation of the sound files.

4.5.3 Statistical analysis

All statistics used in the current study were calculated using the R statistical environment (R Core Team 2021; RStudio Team 2022) with a number of add-on packages.[13] I am interested in both exploratory data analysis and confirmatory statistics, although the analysis is not confirmatory in the strict sense of Baayen et al. (2017; see Section 7.7). The precise methods for the statistical analyses are described in Sections 4.6 and 4.7.1 below, respectively.

4.6 Exploratory analysis

In this section, I take a closer look at the data and explore correlations between the individual predictors and the presence of intervocalic voicing. This is a useful first look at patterns in the data, but in the next section, I proceed to build a regression model which provides a better picture of the complexities found in the data.

[13]Logistic mixed effects model were fitted using the `lme4` package (Bates et al. 2015b, 2021). The `car` package used for calculating variance inflation factors (Fox and Weisberg 2019; Fox et al. 2021), the `MuMIn` package for calculating model effect size (Barton 2020), and the `moments` package for checking distributions of continuous variables (Komsta and Novomestky 2015). The `ggplot2` package was used for generic visualizations (Wickham 2016; Wickham et al. 2021) and the `sjPlot` package for visualizing model coefficients (Lüdecke 2021). More details can be found in Puggaard-Rode et al. (2022b).

4.6.1 Categorical predictors

Table 4.3 and Figure 4.2 show the proportions of voiced tokens for each level of each of the categorical variables. The majority of categorical variables show at least some degree of correlation with closure voicing in the direction predicted in Table 4.1 above.

4.6.1.1 Segmental predictors

LARYNGEAL CATEGORY shows a clear correlation in the expected direction. As predicted above, intervocalic voicing is quite rare in /p t k/, where it was only found in 5% of all tokens. Intervocalic voicing is more common in /b d g/, where it was found in 38% of all tokens. Hence, voicing is not the norm for /b d g/, even though it is sometimes described in the literature as being essentially categorical. In total, continuous closure voicing is found in 24.6% of all intervocalic stops in the corpus.

PLACE OF ARTICULATION does not pattern as predicted from our aerodynamically motivated expectations; as expected, bilabials are voiced more often than velars, but unexpectedly, alveolars are voiced at a much higher rate than either of the other places of articulation. Presumably, there are non-aerodynamic reasons for this. Alveolar stops are generally more frequent than other places of articulation (see Table 4.2), and they are found at a higher rate in function words. While the transcriptions do in principle indicate tapped realizations of the alveolar stops, this is likely somewhat inconsistent, such that some realizations transcribed as alveolar stops are in fact taps [ɾ]; these are presumably always voiced.

PRECEDING SEMIVOWELS, as expected, are less likely than nuclear vowels to correlate with voicing in the following stop.

The behavior of HIGH VOWELS goes against the predictions; high vowels were expected to decrease the chances of voicing, in particular preceding the stop. High vowels preceding the stop show a weak correlation in the expected direction, and high vowels in the same syllable in fact correlate positively with voicing. This is contrary to the aerodynamically motivated predictions but could have a number of other explanations: high vowels are found in a number of very frequent function words, and the syllabic semivowels [ɪ ʊ] are both included in

Table 4.3: *Table of proportion of fully voiced tokens for each level of each categorical variable. Variables marked + show correlations in agreement with the hypotheses in Table 4.1, and ones marked – show disagreement with the hypotheses.*

Variable	Level	% voiced	no. voiced	
LARYNGEAL CATEGORY	Aspirated	5.05	77	+
	Unaspirated	38	844	
PLACE OF ARTICULATION	Bilabial	17.25	89	–
	Alveolar	35.75	611	
	Velar	14.55	221	
PRECEDING SEMIVOWEL	Absent	25.81	832	+
	Present	17.08	89	
HIGH VOWEL	Absent	22.38	584	–
	Present	26.69	337	
PRECEDING HIGH VOWEL	Absent	25.18	748	+
	Present	22.38	173	
NEUTRAL VOWEL	Absent	22.31	747	+
	Present	43.94	174	
PRECEDING NEUTRAL VOWEL	Absent	26	612	–
	Present	22.23	309	
STRESS	Absent	26.13	712	+
	Present	17.36	129	
PRECEDING STRESS	Absent	25.31	637	–
	Present	23.15	284	
STØD	Absent	26.39	792	+
	Present	17.36	129	
PRECEDING STØD	Absent	25.26	914	+
	Present	5.56	7	
MORPHOLOGICAL BOUNDARY	Internal	36.12	95	–
	Inflection	68.75	110	
	Derivation	38.81	26	
	Compound	9.76	73	
	Word	24.61	617	
WORD CLASS TYPE	Open	19.33	407	+
	Closed	31.38	514	
SEX	Female	20.96	192	+
	Male	25.75	729	

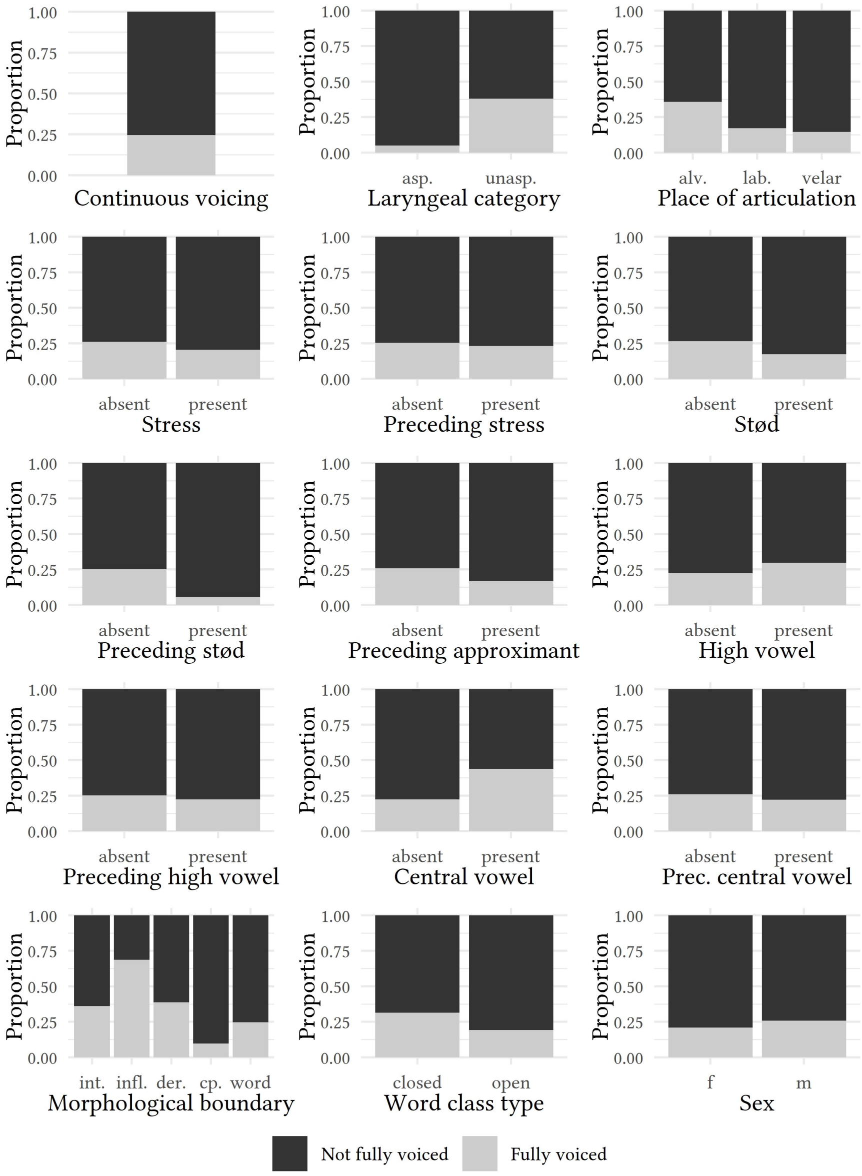


Figure 4.2: *Stacked bar plots showing the proportions of tokens with and without continuous voicing for each level of each categorical variable. (Morphological boundary levels: internal, inflectional, derivational, compound, word).*

this group. As such, there are predictable reasons why syllables with high vowels might frequently undergo phonetic reduction.

As predicted, NEUTRAL VOWELS in tautosyllabic position correlate positively with the presence of closure voicing. However, against expectations, neutral vowels in the preceding syllable show a slight correlation with the absence of closure voicing.

4.6.1.2 Prosodic predictors

As predicted, voicing is more common in UNSTRESSED than STRESSED syllables. Surprisingly, the presence of stress on the preceding syllable shows a (very weak) correlation in the unexpected direction. Also as predicted, voicing is less common in syllables with STØD and is exceedingly uncommon following syllables with stød.

4.6.1.3 Morphosyntactic predictors

The predictions regarding MORPHOLOGICAL BOUNDARY TYPE mostly do not pan out. By far the most voiced stops are at inflectional boundaries, with derivational morphemes and morpheme-internal stops being voiced at approximately the same rate. Stops at word boundaries, by far the most common category, show intervocalic voicing at around chance rate, i.e. the same rate as the data set at large. Finally, stops at compound boundaries are rarely voiced. Given the complexity of this factor, I will refrain from interpreting these results further until I present the results of the regression model.

As predicted, WORD CLASS TYPE interacts with closure voicing, such that members of the closed classes are voiced at a higher rate than members of open classes.

4.6.1.4 Other predictors

SEX correlates with voicing in the predicted direction, such that male speakers produce more voiced stops than female speakers.

4.6.2 Continuous predictors

Having discussed all categorical predictors, I now turn to the continuous ones. Figure 4.3 visualizes the proportion of stops with and without continuous voicing with density plots.

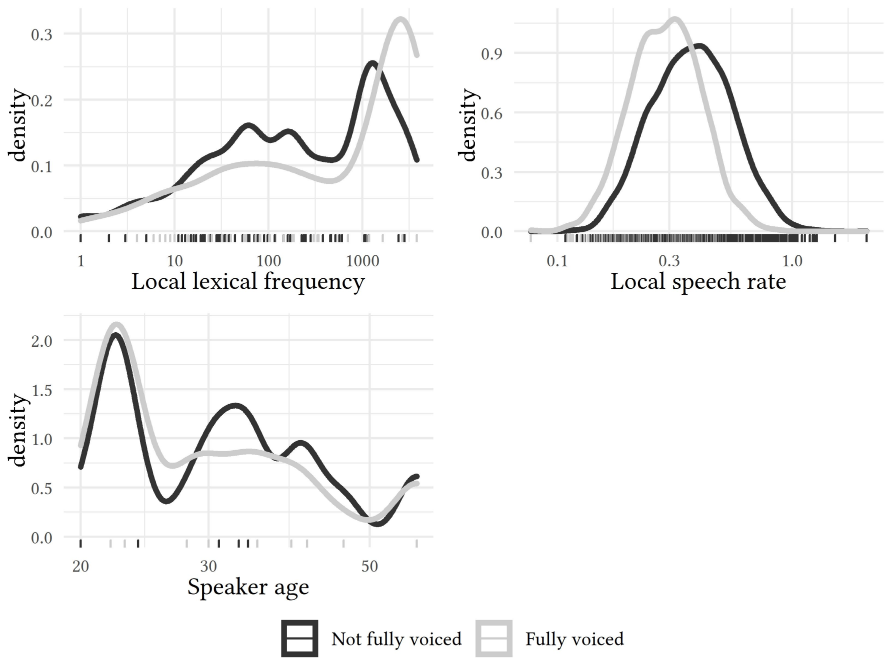


Figure 4.3: *Density plots showing the tokens with and without continuous voicing relative to continuous variables on a log-scale.*

It is clearly (and logically) the case that the most FREQUENT words also account for most tokens, both voiced and voiceless. It is also clearly the case that the words with very high frequency show a higher proportion of voiced tokens, and that the words with medium frequency, particularly between 50–500, show a higher proportion of voiceless tokens.

As predicted, SPEECH RATE correlates with voicing, such that voiceless tokens are more common during slow speech, and voiced tokens are more common during quick speech; recall that speech rate is coded as the duration of the syllables flanking the stop, so a low value (somewhat counter-intuitively) equals high speech rate. In both lexical frequency and speech rate, the distribution of fully voiced tokens is visibly more peaked than tokens which are not fully voiced.

We also see a correlation in the expected direction between age and voicing. Most speakers in the corpus are younger than 25 years old, so it follows naturally that most tokens, both voiced and voiceless, are also produced by this age group. It is, however, also the case that speakers in their thirties and forties produce a relatively higher proportion of voiceless stops.[14]

Having examined the correlations that are found in the data at face value, I now move on to analyzing the data with mixed-effects regression modeling.

4.7 Mixed-effects model

4.7.1 Model selection

The data come from a corpus that is not collected for studying intervocalic voicing, and I am interested in many independent variables. Given the lack of experimental control and the largely exploratory nature of the study, the data is presumably not structured in a way that allows for a maximal random effects structure; this is a common problem with mixed-effects models in linguistics (Meteyard and Davies 2020). This loss in optimal data structure comes with a corresponding gain in ecological validity, which is highly necessary when discussing potential lenition phenomena. There has been a lot of discussion of how to handle this issue in linguistics, with suggestions ranging from maximizing random effects (Barr et al. 2013) to balancing statistical power and Type I error by including only random effects that contribute sufficiently to the model's predictive power (Bates et al. 2015a; Matuschek et al. 2017). These papers generally assume 1) experimental data, and 2) a continuous dependent variable (i.e. linear mixed-effects models). This is important to keep in mind, since experimental data is generally more balanced than corpus data, and linear models are overall more likely to converge than logistic models with binary response variables (Seedorff et al. 2019). I opt for a data-driven

[14] The examples from above the age of 50 all come from a single speaker, so these can safely be ignored.

model selection procedure, inspired by the heuristics proposed by Sonderegger (2022).

The raw values of all continuous variables are positively skewed, so they were log-transformed in order to reach a normal distribution, and standardized to aid interpretation of the model.[15]

The categorical variables are contrast coded (see Schad et al. 2020). Sum contrasts were used for the binary variables. Variables corresponding to articulatory features are all coded as +½ ('present') and -½ ('absent'). laryngeal category is coded as -½ unaspirated, +½ aspirated; sex is coded as -½ female, +½ male; word class type is coded as -½ open, +½ closed. For the three-level variable place of articulation, two theoretically-guided Helmert contrasts were coded: one to test the distinction between velars and non-velars, and one to test the distinction between alveolars and labials, as in (4).

(4) *Velar contrast*: -⅓ bilabial, -⅓ alveolar, +⅔ velar
Bilabials vs. *alveolars*: +½ alveolar, -½ bilabial

The five-level morphological boundary variable is quite complicated. Here, four theoretically-guided Helmert contrasts were used: 1) *internal contrast*, testing the distinction between morpheme-internal and non-morpheme-internal; 2) *affix contrast*, testing the distinction between affix boundaries and other boundaries; 3) *affix type contrast*, testing the distinction between derivational affix boundaries and inflectional affix boundaries; and 4) *compound contrast*, testing the distinction between word boundaries and compound boundaries.

(5) *Internal contrast*: +⅘ internal, -⅕ inflectional, -⅕ derivational, -⅕ compound, -⅕ word
Affix contrast: +½ inflectional, +½ derivational, -½ compound, -½ word
Affix type contrast: +½ derivational, -½ inflectional
Compound contrast: +½ compound, -½ word

[15] Continuous variables were standardized by subtracting the mean and dividing by two standard deviations, following Gelman and Hill (2006).

The data were modeled using logistic mixed-effects regression.[16] The model selection procedure followed two steps: 1) Fixed effects selection with minimal random effects, and 2) pruning of the maximal random effects structure to achieve convergence with (almost) non-singular fit.

FIXED EFFECTS SELECTION. All independent variables which were theoretically motivated in Section 4.4 above are included in the model. There are no theoretical motivations for including interactions. However, we saw in Section 4.6 that the proportion of fully voiced /p t k/ tokens is near-floor, and this could be masking true effects in the data. For this reason, all possible interactions with LARYNGEAL CATEGORY were tested in a random intercepts-only model, in case some effects could be found only in /b d g/. Only significant interactions were kept.

RANDOM EFFECTS SELECTION. All meaningful by-speaker and by-item random slopes were subsequently added to the model; SEX and AGE can of course not vary by-speaker, and all by-item slopes for phonological or morphosyntactic variables are at least potentially problematic. I used strictly uncorrelated random effects; this leads to much higher convergence rates in logistic models, and Seedorff et al. (2019) show that it does not inflate Type I error rates even if the random effects are correlated in the underlying data (although it has a slight adverse effect on statistical power). This model converges with a singular fit, in this case because the model estimates zero-variances within some random slopes. In other words, the variance explained by these random slopes is not found to be different from that explained by random noise in the data. This is a symptom that the model is overparametrized, but should have no influence on the interpretation of the corresponding fixed effects (Brauer and Curtin 2018). All random slopes with estimated zero variances were removed, with the exception of LARYNGEAL CATEGORY. LARYNGEAL CATEGORY was kept since this is a variable of key interest in the study. This means that the resulting

[16] The model was fitted using the `glmer` function in `lme4`, using bound optimization by quadratic approximation (the `bobyqa` optimizer), with the maximal number of iterations increased from the default 10^5 to 10^6. These low-level mechanical details should have no effect on the results, but could be important for reproducibility. See fn. 13 for more details on the R packages used and Puggaard-Rode et al. (2022b) for code and data.

Table 4.4: *Summary of the final model.*

Simple fixed effects	Intercept, laryngeal category, place of articulation (velar contrast, bilabials vs. alveolars), preceding approximant, preceding high vowel, high vowel, preceding neutral vowel, neutral vowel, preceding stress, stress, preceding stød, stød, morphological boundary (internal contrast, affix contrast, affix type contrast, compound contrast), word class type, local lexical frequency, local speech rate, sex, age
Interactions with laryngeal category	Preceding approximant, preceding stress, local speech rate
By-speaker random effects	Intercept, laryngeal category (zero variance), velar contrast, high vowel, stress, stød, internal contrast, affix type contrast, compound contrast, local speech rate
Removed due to zero variance	Bilabials vs. alveolars, preceding approximant, preceding high vowel, preceding neutral vowel, neutral vowel, preceding stress, preceding stød, affix contrast, word class type, local lexical frequency
By-item random effects	Intercept, age, sex
Removed due to zero variance	Local speech rate

model is probably slightly overparametrized, since it is highly unlikely that there is *no* by-speaker variance for laryngeal category in the underlying data, but a reasonable interpretation is that the by-speaker variance for laryngeal category is very close to the random variation for laryngeal category in general. The entire model selection procedure is documented in Puggaard-Rode et al. (2022b), and the final model is summarized in Table 4.4.

None of the included independent variables show problematic collinearity; the variance inflation factor (*VIF*) is below 1.5 for all variables except those appearing in interaction effects.

The coefficients of a generalized linear model correspond to log-odds. These are suitable for regression modeling, as they are unbounded and normally distributed. Odds and odds ratios (*OR*s), on the other hand, are easier to interpret. In order to aid interpretability, I report both the model coefficients and standard error in the log-odds scale, and odds (ratio), which are computed by exponentiating the coefficients. The odds for the intercept can straightforwardly be interpreted as the odds of closure voicing with all other variables kept at zero. Since all variables are either contrast-coded or standardized, the *OR*s can be interpreted straightforwardly as the change in probability associated with that variable (see Sonderegger 2022: ch. 6). Odds and *OR*s are given as fractions; if $OR > 1$, the odds of voicing are higher in the variable level corresponding to + in the contrast coding, and if $OR < 1$, the odds of voicing are higher in the variable level corresponding to - . For the standardized continuous variables, *OR*s refer to the change in predicted likelihood of voicing associated with an increase of 1 standard deviation.

4.7.2 Results

The results of the logistic mixed-effects regression model described above is summarized in Table 4.5. No random effects table is included here, but it can be found in Puggaard-Rode et al. (2022b). The model has a reasonably high marginal effect size of $\Delta R^2 = 0.5$ and conditional effect size of $\Delta R^2 = 0.63$; this is the variance explained by the fixed effects alone and all effects combined, respectively.[17]

In some cases, the results of the mixed-effects model tell quite a different story from the exploratory analysis presented in Section 4.6. In these cases, the results of the mixed effects model should be taken as the best description of the data. The odds for the intercept means that the relative likelihood of a stop being fully voiced is predicted as 11.34 times lower than not being fully voiced if all other variables are ignored.

The significant effects overwhelmingly pattern as predicted. For the following categorical variables, this means that their effects

[17] See Nakagawa et al. (2017) for details of how this is calculated for generalized linear mixed-effects models.

Table 4.5: *Summary of logistic mixed-effects regression model. + indicates agreement with the hypotheses in Table 4.1, and – indicates disagreement with the hypotheses; no symbol indicates a null result. If nothing else is indicated,* OR *< 1 means that the odds of voicing is higher in the absence of a feature, and* OR *> 1 means that the odds are increased in the presence of that feature.*

Variable	**Odds (ratio)**	**coef**	**SE**	***z***	***p***		
(intercept)	1 : 11.34	-2.43	0.42	-5.72	<.001	***	
LARYNGEAL CAT., -asp +unasp	20.15 : 1	3	0.39	7.76	<.001	***	+
PLACE, velar contrast (+velar)	1 : 3.57	-1.27	0.39	7.76	<.001	***	+
PLACE, -bilabial +alveolar	1.16 : 1	0.15	0.34	0.44	0.66		
PRECEDING SEMIVOWEL	1.61 : 1	0.49	0.3	1.62	0.1		
PRECEDING HIGH VOWEL	1.1 : 1	0.1	0.16	0.59	0.55		
HIGH VOWEL	1 : 1.06	-0.06	0.25	-0.24	0.81		
PRECEDING NEUTRAL VOWEL	2.42 : 1	0.28	0.14	1.95	0.05	.	
NEUTRAL VOWEL	1.88 : 1	0.63	0.23	2.77	<.01	**	+
PRECEDING STRESS	3.7 : 1	1.31	0.24	5.41	<.001	***	+
STRESS	1 : 1.94	-0.66	0.21	-3.11	<.01	**	+
PRECEDING STØD	1 : 9.53	-2.25	0.52	-4.36	<.001	***	+
STØD	2.11 : 1	0.75	0.23	3.18	<.01	**	–
BND., internal contrast (+int)	1.28 : 1	0.25	0.32	0.78	0.44		
BND., affix contrast (+affix)	4.79 : 1	1.57	0.36	4.33	<.001	***	(+)
BND., affix type contrast (+inf)	3.13 : 1	1.14	0.61	1.87	0.06	.	
BND., comp. contrast (+cp)	1.57 : 1	0.45	0.41	1.1	0.27		
WORD CLASS, -open +closed	1 : 1.12	-0.11	0.31	-0.37	0.71		
LOCAL SPEECH RATE	1 : 18.29	-2.91	0.27	-10.86	<.001	***	+
LOCAL LEXICAL FREQUENCY	1.85 : 1	0.61	0.27	2.25	0.02	*	+
SEX, -f +m	1.7 : 1	0.53	0.44	1.22	0.22		
AGE	1 : 3.09	-1.13	0.41	-2.74	<.01	**	+
LAR.CAT. : PRECEDING GLIDE	1 : 2.59	-0.95	0.56	-1.71	0.09	.	
LAR.CAT. : PRECEDING STRESS	1 : 3.7	-1.31	0.47	-2.76	<.01	**	
LAR.CAT. : LOCAL SPEECH RATE	6.5 : 1	1.87	0.51	3.67	<.001	***	

are significant in the same (expected) direction as we saw in the exploratory analysis: LARYNGEAL CATEGORY, NEUTRAL VOWEL, STRESS, and PRECEDING STØD. The effect of laryngeal category is very strong, with the unaspirated set being more than 20 times more likely to be voiced intervocalically. The odds of voicing are approximately doubled in syllables with neutral vowels as well as in unstressed syllables, and the odds are around 10 times lower immediately following syllables with stød.

The PLACE OF ARTICULATION variable patterns differently from what we saw in the exploratory analysis. The model finds that voicing in bilabials and alveolars is around four times more likely than in velar stops, but there is no significant difference between bilabials and alveolars. This is in line with aerodynamically motivated predictions. Recall that alveolars were overall voiced at a much higher rate than other places of articulation; this effect disappears in a model that also takes e.g. stress, lexical item, and morphological structure into account.

There is a fairly strong effect of PRECEDING STRESS in the expected direction; voicing is around four times more likely following stressed syllables. This is interesting, as there was essentially no correlation between preceding stress and voicing in the exploratory analysis.

The STØD variable patterns in the opposite direction of the predictions and what we saw in the exploratory analysis. Closure voicing is found to be around twice as likely in syllables with stød. I return to this in the discussion in Section 4.8.3 below.

Only one of the contrasts for MORPHOLOGICAL BOUNDARY TYPE is found to have a significant effect on closure voicing: affix-initial stops are voiced at a much higher rate (around four times) than stops at other kinds of morphological boundary. There are good reasons to expect this at face value: /p t k/ are rarely found in affixes and never in inflectional affixes, affixes are almost never stressed, and affixes often have neutral vowels. However, these are all variables that are controlled for independently in the model, and because of this, I predicted that morpheme-internal stops would be voiced at a higher rate than affixes. I return to this in Section 4.8.2.

Other categorical variables – PRECEDING SEMIVOWEL, PRECEDING HIGH VOWEL, HIGH VOWEL, PRECEDING CENTRAL VOWEL, WORD CLASS TYPE, and SEX – have no significant influence on voicing in the model,

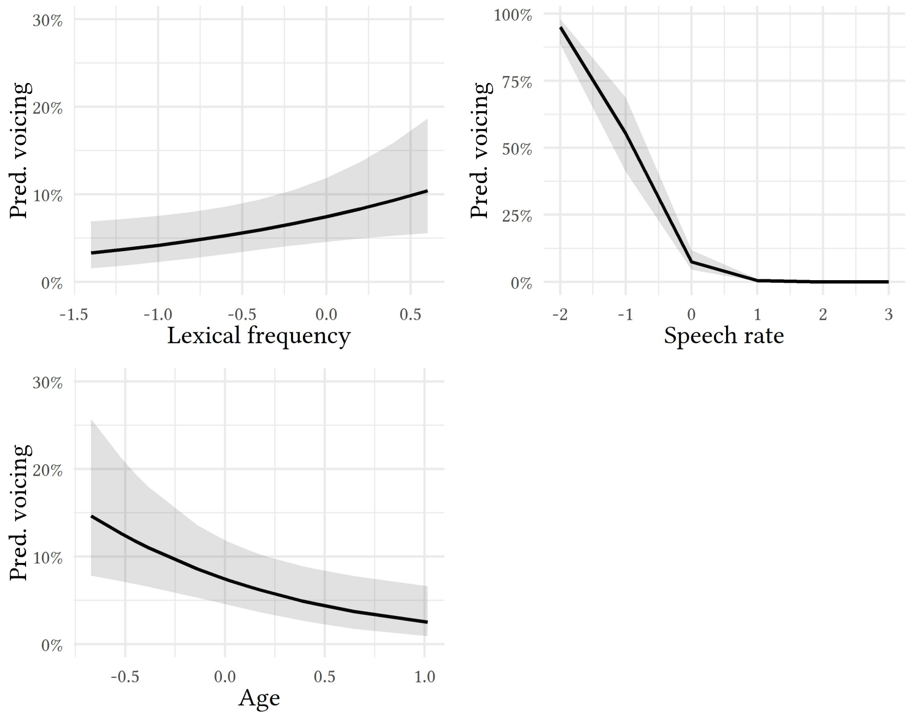


Figure 4.4: *Plots showing the likelihood of fully voiced stops of continuous variables as predicted from the mixed-effects model. The x-axes are standardized. Note that y-axis scales differ; due to the very high likelihood of voicing in very quick speech, keeping the scales identical would blur the effect in other variables.*

even though in some cases, there seemed to be clear correlations in the exploratory analysis. In all cases, we must assume that the correlation we saw at in the exploratory analysis can be better explained by other (potentially random) variables in the data.

The influence of continuous predictors is visualized in Figure 4.4. There is a clear increase in the predicted likelihood of voicing as lexical frequency increases, and a clear decrease in the predicted likelihood of voicing as age increases. The local speech rate variable in particular has an extremely strong influence on voicing, such that quicker speech leads to higher rates intervocalic voicing. In fact, the

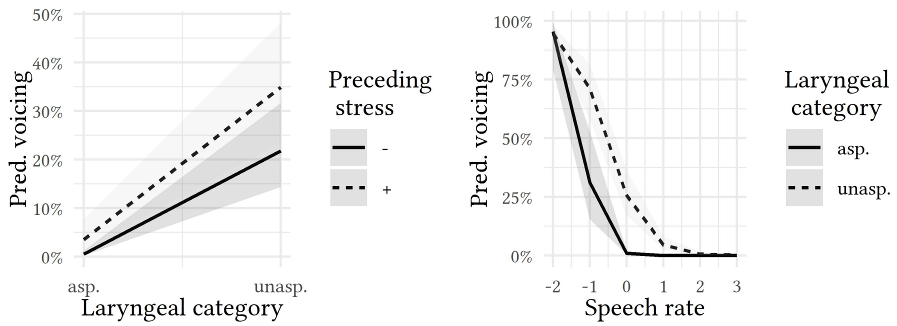


Figure 4.5: *Plots showing the likelihood of fully voiced stops of interaction effects as predicted from the mixed-effects model. The x-axes are standardized. Note that y-axis scales differ; due to the very high likelihood of voicing in very quick speech, keeping scales identical would blur the effect in other variables.*

predicted likelihood of voicing is near-ceiling for the quickest tokens, and near-floor for a large portion of the slower tokens.

Figure 4.5 shows the predicted significant interaction effects. The interaction effect between LARYNGEAL CATEGORY and PRECEDING STRESS patterns as predicted: there is a fairly marginal difference in predicted voicing after stressed syllables in /p t k/, whereas the effect is much more pronounced in /b d g/. The interaction effect between LARYNGEAL CATEGORY and LOCAL SPEECH RATE is similar: both laryngeal categories show near-ceiling voicing in the fastest tokens and near-floor voicing in the slowest tokens, but near-floor voicing is predicted in faster speech in /p t k/ relative to /b d g/.

4.8 Discussion

In this section, I discuss the results in relation to the research questions posed in Section 4.4.

4.8.1 Closure voicing and laryngeal category

The strongest predictor of closure voicing is laryngeal category. There are two main findings: 1) /p t k/ are voiced only very rarely, and much more rarely than /b d g/, and 2) /b d g/ are voiced commonly, albeit still at lower than chance rate. The three major accounts of laryngeal representation in Danish stops (and beyond) that I presented in Section 4.3 all straightforwardly predict the first finding, and all have mechanisms that can account for the second finding.

Abstract [voice] approaches predict the first finding: [+voice] stops are naturally voiced at a higher rate than [+voice] stops. With regards to the second finding, in Kingston and Diehl's (1994) abstract account of [voice], a controlled mechanism could be postulated that actively blocks voicing in [+voice] stops. Such a mechanism seems counter-intuitive, but it is already independently needed for Icelandic, where intervocalic voicing of unaspirated stops is seemingly even more rare than in Danish (Pétursson 1976).

Concrete [voice] approaches also predict the first finding, but not necessarily the second finding. [spread glottis] generally blocks voicing, while unmarked stops are expected to be voiceless in initial position and voiced in intervocalic position.. In Beckman et al.'s (2013) account of [spread glottis], they assume that active privative features are reinterpreted as numerically valued features at some stage in the derivation. Since [spread glottis] is the active laryngeal feature in e.g. German, Danish, and Icelandic, /p t k/ are assigned high numeric values for [spread glottis], while /b d g/ are assigned lower values. They suggest that German /b d g/ are assigned [1sg], which allows for passive intervocalic voicing, and that Danish /b d g/ are assigned [5sg], which blocks passive voicing. This predicts the results quite well. Note, however, that other proponents of [spread glottis] in Danish (like Iverson and Salmons 1995 and Basbøll 2005) do not necessarily assume this mechanism; without such a mechanism, we would simply expect the unmarked /b d g/ to be near-categorically voiced, since this is the unmarked realization of stops in intervocalic position.

Ultimately, I believe the best explanation of the results is one that incorporates our existing knowledge of glottal activity in Danish stops from research by Frøkjær-Jensen et al. (1971), Fischer-Jørgensen and

Hirose (1974), and Hutters (1985). Recall from Section 4.3 that /p t k/ have shorter closure duration and are produced with less muscular tension than /b d g/. Either both series are phonetically lenis, or /b d g/ are in fact fortis. The shorter closure duration and lower muscular tension of /p t k/ would predict more closure voicing in the aspirated series if the vocal folds were properly adduced and tensed for voicing. In careful speech, however, all Danish stops are accompanied by a glottal spreading gesture, presumably to enforce voicelessness. The glottal gestures are different in magnitude across the laryngeal series. /p t k/ have a glottal spreading gesture of great magnitude that lasts throughout the closure and into the release, whereas /b d g/ have a smaller glottal opening gesture that peaks during the closure. Maintaining glottal spreading in /p t k/ is prioritized, because it is required for the aspirated release, which is the primary cue to the laryngeal contrast. Voicing is also actively blocked in /b d g/ through glottal spreading, but in these sounds, it is not crucial for maintaining the contrast. The differences in magnitude of the glottal gesture can explain both findings: the differences between the two series, and the fact that the majority of stops in spontaneous speech are not voiced throughout. Such fine-grained differences in duration and magnitude of gestures can be straightforwardly encoded in the gestural scores of Articulatory Phonology.

The results can be accounted for by all three major accounts of laryngeal representation, but not all theories predict the results equally well. Recall from Section 4.2 that an abstract [voice] account did not allow us to make any specific predictions. A concrete [voice] account only predicts the result with the added machinery of gradient phonetic interpretation of feature values. A gesture-based account predicts the results well with no additional machinery: the necessary 'ingredients', so to speak, are already built into the representational grammar.[18]

On a final note, recall that Schachtenhaufen (2022) has recently suggested abandoning the transcription standard using [b̥ d̥ g̊ b̥^{h} d̥^{s} g̊^{h}] in favor of [p t k p^{h} ts k^{h}], since fortis–lenis is not traditionally

[18] This is, of course, a direct result of the generative capacity of Articulatory Phonology being very powerful; this is an advantage here, but certainly also has drawbacks.

indicated in IPA, and IPA guidelines suggest using [b̥]-style transcription to indicate devoicing of sounds that are usually voiced. The study presented here further cements that Danish /b d g/ are *not* usually voiced: not only are /b d g/ categorically voiceless in most positions, voicing is also regularly blocked in the one syllabic position where it would actually be phonetically natural. I am therefore strongly in favor of Schachtenhaufen's proposal.[19]

4.8.2 Closure voicing as lenition

Closure voicing is mainly found in environments where phonetic lenition is expected: its occurrence increases with speech rate, and it is found at higher rates in unstressed syllables, in syllables with schwa, and in affixes. Based on these results, it seems sensible to consider intervocalic closure voicing a lenition phenomenon in itself.

This has some interesting phonological consequences. As discussed in Section 4.2, it is often difficult to account for intervocalic allophonic voicing with reference to a [voice] feature. In phonological representational frameworks relying on privative features, voiceless unaspirated stops are generally considered unmarked, i.e. they carry no laryngeal features. Similarly, voicing is generally not considered phonologically marked in sonorant sounds (e.g. Lombardi 1995). As such, [voice] is not specified in sonorant sounds and cannot spread from them. Besides, rather than being lenition, the addition of a [voice] feature to a stop entails an increase in markedness and a more complex underlying structure. Rice and Avery's (1989) non-laryngeal [spontaneous voice] feature, which can spread from adjacent sonorant sounds, may be able to represent the process; it does not, however, capture the lenition aspect, as it still entails the addition of phonological material. It also does not capture the probabilistic nature of the process' distribution. We can approach a statistical model of when continuous voicing is more or less likely to occur, but this does not allow us to predict its occurrence with any certainty.

The question remains: *why* is closure voicing a lenition phenomenon in intervocalic stops? A gesture-based approach to

[19] At least as pertains to voicing; see Section 1.4 for a comparison of the transcription standards used here and those proposed by Schachtenhaufen (2022).

laryngeal representation can account for this. I propose that closure voicing in /b d g/ follows from the loss of the glottal spreading gesture that is usually found in these stops. When the vocal fold configuration is optimal for voicing and subglottal pressure is high, some amount of closure voicing is natural and requires no extra effort (Westbury and Keating 1986; see Section 4.2). This vocal fold configuration is required for producing vowels both before and after intervocalic stops, so maintaining it throughout the closure will require the least articulatory effort. In contexts where gestural undershoot is generally expected, it is unsurprising that we also see the loss of a non-distinctive glottal spreading gesture (as in /b d g/), and to a much lesser extent, the loss of a distinctive glottal spreading gesture (as in /p t k/).

This type of lenition is not predicted from either of the featural representational accounts discussed above. If /b d g/ are abstractly specified as [voice], we would not expect lenition to be required for phonetic voicing; in fact, Kingston and Diehl (1994) explicitly use the presence of intervocalic voicing as an argument for why Danish has contrastive [voice]. If /b d g/ receive some value for [spread glottis] late in the phonological derivation, there is no explicit mechanism for reducing this number in environments prone to lenition. However, intervocalic voicing as a consequence of lenition follows directly from the established facts about glottal activity, and as such, can also follow from a representation relying on gestural scores. Recall from Section 4.3 that only a gesture-based account of underlying representation leads to specific predictions about lenition, namely that lenition would cause reduction in the timing and magnitude of associated glottal gestures. This is in line with the results presented here. The difference in lenition rates across the two laryngeal series follows directly from the difference in magnitude of the underlying gestures. This account also correctly predicts that voicing-as-lenition is only found intervocalically; the loss of a glottal opening gesture would not result in voicing in initial position, where voicing requires effort.

In Table 4.6, I summarize the predictions following from different theoretical approaches, and whether or not support for these predictions was found in the current study.

Table 4.6: *Summary of predictions from different theoretical approaches. + indicates that predictions are in agreement with the findings of this chapter.*

	Danish /b d g/		**Danish /p t k/**		**Lenition**	
Approach	*Prediction*		*Prediction*		*Prediction*	
Concrete [voice]	Variable voicing	+	Negligible voicing	+	No predictions	
Abstract [voice]	All outcomes possible		All outcomes possible		No predictions	
Gestures	Limited voicing	+	Negligible voicing	+	Voicing in both series	+
					More voicing in /b d g/	+

4.8.3 The relative predictive power of variables

LARYNGEAL CATEGORY is a very strong predictor of voicing, as are a number of variables associated with lenition. Particularly strong lenition variables are LOCAL SPEECH RATE, PRECEDING STRESS, and affix boundaries, but overall, the majority of lenition variables have a significant influence on voicing in the expected direction. It is interesting that affix-initial stops have a particularly high likelihood of voicing. In Section 4.4.1.3, I hinted that this may have an exemplar theoretic explanation: affixes are so often encountered with closure voicing that it has seeped into the underlying representations at the morpheme-level in a way that is not predictable at the phoneme-level. This is obviously controversial, in large part because it is impossible to represent in many modular approaches to grammar (where phonetic information is invisible to morphology), and it is a very different conception of phonological representation than those discussed above. This remains an interesting problem for further research.

Many of the other variables that were expected to influence closure voicing are aerodynamic in nature, and these generally do not have an observable effect on closure voicing in the data. This may be either because these variables truly do not influence closure voicing, or because the influence of these variables is more gradient in nature. It is possible that aerodynamic variables affect the relative proportion of closure voicing within those closures that I simply categorize as 'not fully voiced'. This can only be tested in a study with more fine-grained coding of voicing.

I had a number of predictions for how the tongue position before and after the occlusion would affect the rate of closure voicing, which can be mostly be summarized as follows: a narrower constriction in the oral cavity before and after the occlusion was expected to decrease the chances of closure voicing, because such sounds are sometimes taken to be less sonorous (e.g. Parker 2002), and voicing follows more naturally from sounds with higher sonority; in fact, Chomsky and Halle (1968) define their distinctive feature [sonorant] exclusively with reference to whether voicing follows naturally from the vocal tract configuration. However, none of these predictions hold up; no effect of tongue body position was found except for the point of occlusion itself.

Place of articulation has a strong effect on voicing, and this has an aerodynamic explanation. The supralaryngeal cavity is relatively small during a velar occlusion, which provides little opportunity for passive expansion, and as such, velar stops are voiced at a lower rate. Alveolar and bilabial occlusions are more amenable to voicing, and the difference in size between the resulting cavities is negligible, which may be why they do not differ significantly in their amenability to voicing.

The influence of stød on the potential for closure voicing can also be thought of as an aerodynamic effect. The naturalness of intervocalic closure voicing crucially depends on high subglottal pressure at the time of occlusion and on the vocal fold configuration being amenable to voicing. Closure voicing following stød is very rare – this was a strong effect in spite of the total number of relevant tokens being quite small – presumably because laryngeal contraction in the production of stød causes a vocal fold configuration that is less amenable to voicing than that of modally voiced vowels. Tautosyllabic stød was found to increase the chances of voicing, which is surprising, given that stød has many of the same syllable-initial cues as stress. However, another initial articulatory correlate of stød reported by Fischer-Jørgensen (1987, 1989) is increased subglottal pressure, which in itself increases the likelihood of voicing.[20] This may serve to explain why tautosyllabic stød empirically shows a negative correlation with voicing (see Table 4.3), but correlates positively with voicing in a model that also controls for stress (see Table 4.5).

4.9 Conclusion

In this chapter, I have investigated intervocalic stop voicing in a corpus of spontaneous Danish speech. Although Danish stops are generally well-described, most of what has previously been written about voicing has been speculative. I have shown that intervocalic voicing is very rare in /p t k/ and occurs at lower than chance rate in /b d g/. In modeling the data, I controlled for a number of aerodynamically motivated predictors, most of which appear to have little influence on the occurrence of closure voicing. However, closure voicing was found at relatively high rates in environments where lenition is expected, i.e. quick speech, unstressed syllables, before neutral vowels, and in morphological affixes. This supports an analysis of intervocalic voicing as a lenition phenomenon. These findings can be accounted for with reference to previous articulatory studies showing that both laryngeal series of Danish stops are produced with glottal spreading gestures that counteract voicing, although these gestures differ in timing, magnitude, and functional load. Intervocalic voicing can be modeled as the loss of this gesture. The gesture is lost at a higher rate in /b d g/, where it is shorter, of smaller magnitude, and does not serve a critical distinctive function. There is a very extensive literature on the representation of laryngeal contrast, and I have necessarily discussed only a few perspectives here. If intervocalic voicing is indeed a lenition phenomenon, I suggest that this is best represented in a phonological representational framework which can directly incorporate the timing and magnitude of articulatory gestures, such as Articulatory Phonology.

[20] Although bear in mind that subglottal pressure was measured for only one participant, and no words with initial oral stops were measured, so this explanation must be taken with a grain of salt.

Few corpus studies of intervocalic voicing are available, and as such, it is difficult to compare these results to other 'aspiration languages' (or 'true voice languages' for that matter). This means that more studies are necessary, detailing how different variables influence the probabilistic occurrence of closure voicing in stops in other languages. This will help determine which effects should be associated with phonetic implementation only, and which should be considered grammatically encoded.

CHAPTER 5

Time-varying spectral characteristics of stop releases

5.1 Introduction

The aspirated alveolar stop /t/ in Standard Danish is usually strongly affricated. This was already pointed out by Otto Jespersen (1897–1899: 355). He maintained that /t/ was best described as an aspirated stop, but assumed that Danish was undergoing a sound change whereby all aspirated stops would eventually become affricates, as had happened in some varieties of German a millennium earlier with the Second Consonant Shift. Jespersen assumed that /t/ was most advanced in this sound change, followed by /k/, and finally /p/. Today, more than a century after Jespersen's observations, the affrication of /t/ is taken for granted in the literature; it has been established several times over, and has been shown to be exceptionless (see Section 2.3.4). While it is cross-linguistically common for the initial burst noise of stops to have a similar frequency range to fricatives at the same place of articulation,

A revised paper corresponding to this chapter has been published (Puggaard-Rode 2022b). Audio data are available online in password-protected form (Grønnum 2016); replication data and code are freely available (Puggaard-Rode 2022a).

this usually makes up a comparatively small portion of stop releases in other languages. Brink and Lund (1975) tracked the development of /t/-affrication across more than a century of recordings of Copenhagen Danish, and showed that it went from a widespread phenomenon in the mid-19th century to an exceptionless phenomenon in the mid-20th century.

As discussed in Section 2.3.4, the prominent affrication in /t/ has led to a variety of different phonetic transcription strategies. In very narrow transcription, it is often assumed that /t/ in simple onset is best represented as /d/ with some 'garnish': [d̥^{s}] (e.g. Basbøll 1968, 2005; Grønnum 1998), [d̥^{sh}] (e.g. Petersen 1983), and [d^{s}h] (Brink and Lund 1975) are all used in the literature, under the assumption that the only meaningful difference between /d t/ is the release. More broad transcriptions include [ts^{h}] (Basbøll and Wagner 1985), and [t^{s}] (e.g. Grønnum 1998), the latter of which has emerged as the standard. More recently, Schachtenhaufen (2022) has proposed that the sound is a true affricate and should be transcribed as [ts].

Fischer-Jørgensen (1972d) shows that having the right noise profile during the release is a crucial cue to the perception of the laryngeal contrast in stops at all places of articulation, which suggests that /t/ is not so special after all. While there is consensus about the affrication in /t/, possible affrication patterns in /p k/ have never been investigated. On the one hand, since /p t k/ show class behavior in other matters (e.g. phonotactics; see Section 2.4.2), we might also expect them to show class behavior in phonetic implementation; on the other hand, Chodroff and Wilson (2018) recently found only moderate signs of class behavior in the realization of place cues in American English /p t k/. The most straightforward explanation for the lack of interest in affrication patterns in Danish /p k/ is that it is not particularly salient (if it is there at all); perhaps this is simply because coronal frication is more salient than labial and dorsal frication. This is a reasonable assumption, which can help account for why most affricates cross-linguistically are coronal (Ladefoged and Maddieson 1996).

One goal of this chapter is to investigate Jespersen's prediction a century later: are Danish aspirated stops changing into affricates? This is not straightforward: the boundary between an aspirated stop and an affricated one is fuzzy, as is boundary between an affricated stop and a proper affricate. I approach the question by looking holistically and dynamically at time-varying spectral characteristics throughout stop releases, and how they vary, using the DanPASS corpus (see Section 4.5.1). I focus on the following questions, which are more readily answerable than the question of whether or not the sounds in question are affricates:

(1) How do the spectral characteristics of Danish stop releases vary across time?

(2) How are the time-varying characteristics of Danish stop releases affected by different phonetic contexts? An example could be coarticulation effects following from features of the following vowel, like backness, height, and rounding, all of which affect the size and shape of the vocal tract.

When analyzing the dynamics of spectral characteristics, researchers usually resort to using a small number of discrete measurements aimed at capturing as much of the relevant spectral information as possible. For vowels and sonorant consonants, an example is formants; for obstruent consonants, examples are spectral moments or coefficients of discrete cosine transformations of the spectrum. A second goal of this chapter is to demonstrate function-on-scalar regression (FOSR; Reiss et al. 2010; Greven and Scheipl 2017a; Bauer et al. 2018) as a method for taking the entire spectrum into account when analyzing sources of phonetic variance. Rather than relying on discrete measurements, FOSR allows for the use of complete spectra as response variables. FOSR gives a clear and easily interpretable overview of the influence of various factors on time-varying spectral characteristics, and does so with minimal reduction of the information in the acoustic signal. Other recent studies have compared full (temporally static) spectra in order to illuminate differences between palatalized and non-palatalized consonants using smoothing spline ANOVA (Iskarous and Kavitskaya 2018) and generalized additive models (Nance and Kirkham 2020); in Section 6.7, I use functional principal component analysis to analyze the main sources of variance in spectra of stop releases. Functional regression models have been used in the analysis of phonetic data previously (e.g. Pouplier et al. 2014, 2017; Cederbaum et al. 2016; Carignan et al. 2020;

Volkmann et al. 2021). However, to the extent of my knowledge, this is the first study to use FOSR to analyze speech spectra.[1]

Section 5.2 of this chapter discusses the acoustic characteristics of aspirated stops and affricates, and the available heuristics (or lack thereof) for determining whether a sound is phonetically one or the other. Section 5.3 discusses available methods for measuring frication and some of the problems associated with these, and Section 5.4 presents FOSR and other smoothing-based approaches to dynamic data analysis as possible solutions to these problems. Section 5.5 presents the methods used in this study in detail, and Section 5.6 shows the results. In Section 5.7, I discuss the hypotheses presented above on the basis of the results, and discuss opportunities and limitations of FOSR as used here. Section 5.8 briefly concludes the chapter.

5.2 Aspirated stops, affricates, and the middle ground

The production of both stop consonants and affricates has been modeled thoroughly in the work of Fant (1960) and Stevens (e.g. 1993a, 1993b, 1998: chs. 7–8). A shared component of both types of sound is a complete occlusion somewhere in the oral cavity, which allows intraoral air pressure to build up. Another shared component is a release phase, in which this pressure is released, resulting in a rapid sequence of acoustic events, including an initial brief transient followed by frication. The transient shows a fairly even distribution of noise throughout the spectrum. Frication noise is subsequently generated at or near the point of occlusion; due to the high pressure behind the constriction and the narrow gap in the oral cavity, the escaping air becomes turbulent and excites the area around the constriction. The nature of this noise gradually changes as the approximation gradually widens. In aspirated stops, air will continue to escape through the open glottis for some time after the release, and turbulence

[1]Wood (2017a: 390ff.) proposes similar models for the analysis of other types of spectra (infrared spectra and protein mass spectra), but in both cases, the spectra are independent variables. In the studies reported here, spectra are the dependent variables.

noise generated at the area around the vocal folds continually excites the vocal tract.

The energy distribution of the turbulent frication noise depends on the nature of the obstruction (Shadle 1991). In labials, since there is no cavity in front of the obstruction, the frication noise is generated directly at the lips, causing a fairly even distribution of noise throughout the spectrum, with a slight linear drop in amplitude at increasing frequencies. In alveolars, the turbulent air stream impinges on the teeth immediately in front of the constriction, meaning there is only a very small cavity anterior to the constriction, causing high resonance frequencies around 5 kHz to be excited. In velars, the turbulent air stream impinges on the hard palate at an oblique angle, before being filtered through a sizeable front cavity, causing relatively low resonance frequencies somewhat below 2 kHz; note, however, that the exact point of occlusion in velars is variable and depends on surrounding vowel(s), since the tongue body is less precisely controlled than the tip and blade (Ouni 2014), and the tongue body is itself more directly involved in the production of vowels than the tip and blade. A more fronted obstruction will cause the air stream to more directly impinge on the hard palate, causing higher resonance frequencies.

During aspiration, low-frequency noise is generated as the airstream passing through the glottis impinges on the vocal folds, epiglottis, and surfaces directly above the glottis; this turbulence noise further excites the natural resonances of the oral cavity, which of course largely depend on e.g. the position of the tongue. The aspiration noise is present throughout the release, but is initially dominated by frication. As the obstruction above the glottis opens, aspiration noise will gradually overtake frication noise in prominence (Hanson and Stevens 2003).

In voiceless unaspirated stops, the frication phase is very brief, but it is an important cue to place of articulation. There are two primary place cues in stops: the spectral characteristics of the initial frication phase (e.g. Stevens 1971; Stevens and Blumstein 1978; Blumstein and Stevens 1979, 1980), and the transitions of formants as the articulators move from occlusion to vowel (Kewley-Port 1982, 1983; Kewley-Port et al. 1983; Stevens et al. 1999). In aspirated stops, formant transitions are relatively weak, because movement of the articulators typically happens before the onset of voicing. This makes frication as a place cue all the more important in aspirated stops. Frication is also usually a stronger cue in aspirated stops: since the glottis is spread during at least part of the closure, there is a greater build-up of supraglottal air pressure, causing quicker releases and greater burst intensities than in unaspirated stops (see e.g. Löfqvist 1975a, 1980; Jaeger 1983). Long voicing lag can in itself lead to affrication in certain environments: when devoiced, high front vowels can be acoustically similar to fricatives (Mortensen 2012). This can lead to the common sound change whereby /k/ ⟶ /tʃ/ before /i/ (Hock 1991; Ohala 1992), as observed in e.g. Slavic, Indo-Iranian, and Middle Chinese (Guion 1998 and references therein), and the common phonological process where /t/ is realized as an affricate or fricative before /i/, as observed in e.g. Finnish and Korean (Kim 2001; Hall and Hamann 2006; Hall et al. 2006).

The timing of gestures in Danish aspirated stops is different from comparable Germanic languages, as discussed throughout Section 2.3. In Icelandic and Swedish, peak glottal opening is achieved relatively early during the closure of aspirated stops (Pétursson 1976; Löfqvist 1980); in English and German as well, the glottis is typically fully spread sometime before the stop release (Sawashima 1970; Hoole et al. 1984). Furthermore, closures in aspirated stops are typically longer than in unaspirated stops (Lisker 1957; Löfqvist 1976; Stathopoulos and Weismer 1983; Braunschweiler 1997). This ensures that supraglottal air pressure is high at the time of the release. In Danish, however, peak glottal opening is typically just after the stop is released (Frøkjær-Jensen et al. 1971), and closure duration is shortest in aspirated stops (Fischer-Jørgensen 1969, 1972b). Taken together, these two facts about Danish aspirated stops – late peak glottal opening, and relatively short closure duration – mean that there are fewer mechanisms in place to ensure high supraglottal air pressure at the time of release, and accordingly, less guarantee of a prominent burst.[2] This can motivate why a constriction would be retained for relatively long in Danish

[2]This exposition suggests that Danish stops are outliers in Germanic, but in fact, all languages which have been examined in detail have idiosyncrasies in their stop articulation. If anything, it should indicate that oral and glottal gestures are largely independently controlled, and that individual languages have a lot of freedom in how phonetic categories are implemented.

stop releases. Functionally, it can also explain the 'need' for affricated releases in Danish: if the place cues of the burst are not otherwise so prominent, they can be strengthened by retaining a constriction after the release.

There are no clear heuristics to decide whether a particular speech sound is an affricated aspirated stop or an affricate – at least not from the acoustic signal alone. In phonology, a decision may be reached on the basis of behavior. Affricates are often assumed to contain a feature like [stop] as well as one usually used in the representation of fricatives, such as [strident] (e.g. Jakobson et al. 1951) or [continuant] (e.g. Lombardi 1990);[3] see Lin (2011) for an overview of how affricates have been modeled in phonological theory. If an occlusive with a lot of frication behaves like an aspirated stop to all extents and purposes, it should probably be considered an aspirated stop at the phonological level; there will be no need to posit a [continuant] feature. If it patterns with fricatives, or shows other forms of exceptional behavior, those would be grounds for considering it an affricate at the phonological level.

On these grounds, Standard Danish /t/ should certainly be considered an aspirated stop. The phonotactic behavior of /t/ is similar to that of other stops (Vestergaard 1967), and /t/ shows the same patterns of positional allophony as /p k/, with truncated release after /s/ and in weak position (see Chapter 3), and loss of release syllable-finally (although optional release phrase-finally; Grønnum 2005: 49). Furthermore, when loan words with alveolar affricates are nativized and adapted to Danish phonology, the affricate is generally reanalyzed as /s/ rather than /t/, as in the examples in (3);[4] etymologies are from DSL (2018).

[3]In binary feature accounts, affricates are often represented with both [-continuant] and [+continuant] (e.g. Sagey 1986).

[4]A counterexample is *tzatziki*, which is nativized as [tʰætˈsiki] (DSL 2018); here, the first /ts/ is reanalyzed as /t/, and the second as ambisyllabic /t.s/.

(3)	[sɑːˀ]	*tsar*	‘czar’	from Russain [tsarʲ]
	[suˈkʰiːni]	*zucchini*	‘zucchini’	from Italian [tsukˈkino]
	[sɛn]	*zen*	‘zen’	from Japanese [dzen]
	[ˈsyʁ̯ek]	*Zürich*	‘Zurich’	from German [ˈtsyːʁɪç]
	[suˈnɑːmi]	*tsunami*	‘tsunami’	from Japanese [tsɨnami]

In a study of Danish speakers’ productive acquisition of Standard Chinese coronal obstruents (Puggaard 2020c), it was further shown that the most common error in the production of (non-aspirated) /ts/ is realizing it with no closure phase, i.e. similar or identical to /s/. Native speakers of Danish do not map Standard Chinese /ts/ to their native /t/ phoneme. They do, however, tend to map Standard Chinese /tsʰ/ to their native /t/ phoneme, further cementing that both affrication and aspiration are crucial cues to Danish /t/.

From a phonetic perspective, Stevens (1993a) defines affricates as sounds which have two separate constrictions formed by the primary articulator. The anterior constriction forms a complete closure, while the posterior one forms a close approximation. In affricates, frication noise is generated at this posterior constriction, while in stops, frication noise is generated directly at the point of occlusion. This distinction is difficult to extend to acoustics or to gauge impressionistically. In practice, most decisions about stop–affricate category membership is likely based on intuition; a sound is categorized as an affricate if frication lasts for more than a certain proportion of the release. It is therefore not a goal of this chapter to determine whether /p t k/ are phonetic affricates in Danish; such a decision can only be made with targeted articulatory studies comparing Danish with other languages with clear-cut stop–affricate distinctions. This is rather an exploratory study aimed at better understanding the distribution of spectral properties in Danish stop releases.

5.3 Measuring frication

It has long been established that frication at different places of articulation (whether in fricatives, stop releases, or otherwise) has distinct spectral properties (see Kopp and Green 1946). A classic method for differentiating places of articulation in frication is locating peaks and valleys in spectral energy distribution, essentially by 'eyeballing' spectrograms (e.g. Hughes and Halle 1956; Strevens 1960).

Forrest et al. (1988) popularized treating the spectrum as a probability mass function, and analyzing it by calculating four moments: 1) the 'mean frequency', also known as center of gravity (COG); 2) standard deviation (SD), 3) skewness, and 4) kurtosis. COG reflects the mean distribution of energy across the spectrum; SD reflects how much the energy deviates from the mean; skewness reflects how much the energy distribution is skewed relative to the mean, and in which direction; kurtosis reflects the peakedness of the energy distribution. Forrest et al. found that spectral moments distinguish fairly well between places of articulation in stop bursts, and that particularly COG, skewness, and kurtosis distinguish fairly well between places of articulation in alveolar and post-alveolar fricatives; Stoel-Gammon et al. (1994), on the other hand, found that SD is particularly stable in determining the difference between dental and alveolar stop bursts. The results of subsequent studies have overall not been particularly stable (see e.g. Shadle and Mair 1996), but COG remains a very popular measure in the analysis of spectral properties of fricatives, often without taking into account other moments; an example is Gordon et al. (2002). This is problematic, since spectra often correspond to functions that are far from normally distributed. The mean value from a non-normal distribution does not give a clear picture of the shape of the distribution, and spectra with quite different shapes may have very similar COG.

A number of other measures have been proposed for analyzing frication, mainly for determining the precise place of articulation in fricatives. Jongman et al. (2000) find that the different places of articulation in English fricatives are distinguished fairly well using the average location of the spectral peak. Koenig et al. (2013) show that the mid-frequency spectral peak, i.e. the frequency with the highest amplitude within a 3–7 kHz band, captures the fairly subtle difference between labialized and non-labialized alveolar fricatives in adolescents.

Another proposed method is using cepstral coefficients derived from a discrete cosine transform of the spectrum (DCT; Watson and Harrington 1999). DCT reduces the high dimensionality of the spectrum to (typically) four discrete values, corresponding to the amplitude of half-cycle cosine waves derived from the spectrum. DCT0 reflects the mean amplitude of the spectrum; DCT1 reflects the linear slope; DCT2 reflects the curvature; and DCT3 reflects the amplitude at higher frequencies. In a comparison of /ʃ ç/ in different varieties of German, Jannedy and Weirich (2017) show that DCT-based classification more closely approximates the perception of these sounds than classification based on spectral moments, and DCT coefficients have been shown to outperform spectral moments in classification of place of articulation in both voiceless stops (Bunnell et al. 2004) and fricatives (Spinu and Lilley 2016). While DCT coefficients give a fuller picture of spectral shape than spectral moments, they are also more difficult to interpret.

Measurements such as the ones discussed above are often taken at static or normalized points in time, such as the midpoint (or some pre-determined range around the midpoint) of fricatives or affricates (for examples, see e.g. Jongman et al. 2000; Liu and Jongman 2013). Mücke et al. (2014) refer to these points in time as 'magic moments'. Magic moments give us a limited picture of the acoustic nature of sounds; affricates are inherently dynamic, and Reidy (2016a) shows that even sibilant coronal fricatives vary dynamically throughout their time course in language-specific ways. Spectral properties of stop releases are usually measured only at the burst, which in aspirated stops corresponds to a relatively small initial portion of the release (see e.g. Chodroff and Wilson 2014).

Summing up, most approaches to quantifying frication reduce the complex time-varying information in spectra to something more manageable. This is very reasonable, because 1) many popular statistical methods in linguistics cannot handle variables with high dimensionality, and 2) it is a goal in itself to propose the simplest possible model of how language works with the highest possible explanatory value. With regards to 1), statistical models which can take complex dynamic information into account are increasingly being used, as discussed in the next section; this chapter demonstrates how FOSR can be used to model time-varying spectral information with little reduction of dimensionality. With regards to 2), deciding on a model of language which balances simplicity and explanatory value can simply not be done without first testing very complex models. Studies mentioned above have demonstrated how some patterns can only be uncovered by increasing dimensionality. For example, Reidy (2016a) shows that the language-specific nature of spectral dynamics in fricatives only becomes apparent when measuring spectral properties at several timepoints, and Jannedy and Weirich (2017) show that the spectral differences between [ʃ ç] in German are more readily apparent when using a measure that takes more of the spectrum into account (i.e. using DCT coefficients rather than moments).

5.4 Smoothing approaches to analyzing dynamic data

In the past years, following Baayen's (2008) popularization of mixed-effects regression models in linguistics, the field has seen a rapid increase in the use of sophisticated statistical techniques. A problem with linear models is the analysis of dynamically varying data, in particular data from time series. If some measure varies as a function of time, then a linear model by necessity assumes that the variation follows a straight line. As Sóskuthy (2017) demonstrates for formants, this is a poor assumption: variance as a function of time is often non-linear. A solution to this problem is using smoothed curves. Given a number of data points associated with e.g. a time series, a smoothing function can be used to approximate a continuous curve corresponding to the data's non-linear variation as a function of time (see de Boor 2001; Wood et al. 2016). Smoothing involves reducing the observations to a number of basis functions (or 'knots'), and using a penalizing smoothing parameter to determine the wiggliness of the resulting curve (see Gubian et al. 2015). Combining too many basis functions with a low smoothing penalty will lead to overfitting, resulting in curves that include irrelevant information; conversely, combining too few basis functions with a high smoothing penalty will likely lead to underfitting, resulting in curves that omit relevant information.

Generalized additive (mixed) models (GAMMs) have rapidly become very popular in linguistics (see e.g. Wood 2017a; Wieling 2018;

van Rij et al. 2020a; Sóskuthy 2021). These are similar to linear mixed-effects models, but allow for the inclusion of smooth effects. They are typically used for time series analysis, but have also been used to analyze the dynamics of e.g. EEG registration (Baayen et al. 2018; Voeten 2020: ch. 5), geo-linguistic variation (e.g. Wieling et al. 2011, 2014; see also Chapter 6), and speech spectra (Nance and Kirkham 2020).

Functional data analysis (FDA; Ramsay and Silverman 2005; Ramsay et al. 2009; Gubian et al. 2015; Pouplier et al. 2017) has overall been less influential in linguistics. FDA refers to a family of statistical methods which extend existing methods to account for functional data. In practice, this means that smoothed curves can be used as input variables in statistical models, in addition to discrete values. An example of this is the functional extension of principal component analysis (FPCA), which can be used to determine the primary sources of variance in curves. For example, Gubian et al. (2015) use FPCA to jointly analyze how $F1$ and $F2$ pattern in the realization of diphthongs and hiatuses in Spanish, respectively. I return to FPCA in Section 6.7, where I use the method to determine the primary modes of variation in noisy spectra.

Functional regression models are suitable when one or more of the analyzed variables are of a functional nature. If an independent variable is functional and the response variable is constant over the functional domain, this can be modeled with scalar-on-function regression; if the response variable is functional and all independent variables are constant over the functional domain, this is suitably modeled with function-on-scalar regression (Bauer et al. 2018). There are several approaches to modeling function-on-scalar data (an overview is given in Greven and Scheipl 2017b: 110ff.). Here, I will focus on the implementation presented by Scheipl et al. (2015, 2016), Greven and Scheipl (2017a) and Bauer et al. (2018). For the mathematically inclined, the model can be summarized as in (4), from Bauer et al. (2018: 353).

(4) $$g(\mathbb{E}(Y_i(t)|\chi_i, E_i(t))) = \beta_0(t) + \sum_{r=1}^{R} f_r(\chi_{ri}, t) + E_i(t)$$

$g(\cdot)$ is a pre-specified link function mapping the predictor to the functional domain. The expected value $\mathbb{E}(\cdot)$ of each observation *i = 1, …, n* of the response variable *Y* as a function of *t* conditional on a set of covariates χ and residual functional error $E(t)$ corresponds to a functional intercept $\beta_0(t)$, as well as *R* covariate effects $f_r(\cdot)$, each of which form a subset χ_r of the full covariate set and may vary over the functional domain *t*, and the residual functional error $E(t)$.

Functional regression models and GAMMs are conceptually very similar. GAMMs are often fitted using the R package `mgcv` (Wood 2017a, 2021), which allows for significant flexibility in the selection of spline bases (Wood 2017a: ch. 5), residual error distributions (Wood et al. 2016), and smoothing parameter estimation methods (Wood 2011; Wood et al. 2015), and handling of autocorrelated residuals (Baayen et al. 2018), and can handle very large data sets (Wood et al. 2017). Wood (2017a: 290ff.) gives a number of examples of how functional regression models can be implemented in `mgcv`. Perhaps for this reason, the framework for functional regression modeling I adhere to here is sometimes referred to as (generalized) functional additive mixed modeling (Scheipl et al. 2015, 2016). A disadvantage of GAMMs is that they cannot take functional response variables. If I wanted to model spectral variance with GAMMs, I would have to use an amplitude measure as the response variable, and model the variation in amplitude across the time and frequency domains. This is conceptually not very satisfactory: the spectral shape is our variable of interest, not the individual amplitude levels.

Functional additive regression models are implemented in the `pffr` function of the R package `refund` (Goldsmith et al. 2021). This function uses the `mgcv` computation engine, and inherits the same flexibility as GAMMs fitted with `mgcv`, but allows for dependent and independent variables to be functional. The syntax is also similar to `mgcv`, except there are several more term constructors for including various kinds of variables; most of these are not discussed here. A fully reproducible example of the model fitting and selection procedure using `pffr` is given in Puggaard-Rode (2022a), where I also decompose the code.

Functional regression models are usually high-dimensional and the number of underlying data points is often very high. This can make traditional significance tests unreliable, as these are highly affected by sample size (see e.g. Kühberger et al. 2015 and references therein). Wood (2013) proposes an *F*-test for calculating significance of non-linear variables in GAMMs, and the results of this test are also reported in the output of `pffr`; however, researchers should exercise caution in interpreting these results, as even tiny effects will appear highly significant if the sample size is sufficiently large. This is also the case for likelihood ratio tests of nested models. For this reason, I do not report *p*-values in this chapter. This issue is not specific to FOSR models; if the same data was fitted to a GAMM, these concerns would still hold.

In any case, *p*-values and associated measures of non-linear effects can only tell us if there *is* an effect, they cannot tell us much about the nature of that effect. A more suitable way to explore non-linear effects in exploratory studies is to visualize them. If the goal is hypothesis testing, Bauer et al. (2018) propose several different solutions. Marra and Wood (2012) propose a method for calculating confidence intervals of non-linear effects; this method can be used to quantify and visualize the uncertainty associated with non-linear fitted effects along the functional grid. Bauer et al. (2018) propose a more precise bootstrap-based method for calculating confidence intervals, but this precision comes with a significant computational cost.

5.5 Methods and materials

5.5.1 Acoustic analysis

As in the study of intervocalic voicing presented in Chapter 4, this study relies on the monologues from the DanPASS corpus (Grønnum 2009; see Section 4.5.1). The initial acoustic analysis was done in Praat (Boersma 2001; Boersma and Weenink 2021). An automated script was used to locate all aspirated stops (i.e. members of /p t k/) in simple onset position in the DanPASS monologues, and combine them into a single sound file with a subset of the original annotations. The script is a modified version of the one used in Chapter 4 written by Dirk Jan Vet (see Puggaard-Rode et al. 2022b). This located a total of 2,539 stops. The release phases of the stops were segmented primarily on the basis of the waveform, with the burst used to demarcate the beginning of the release and the first signs of periodicity used to demarcate the end

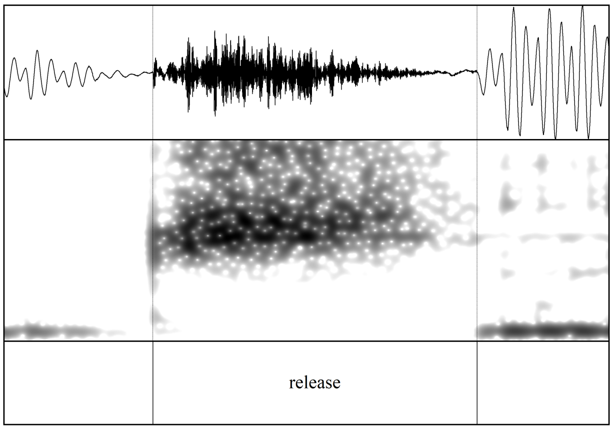


Figure 5.1: *Example of a segmented* /t/ *release.*

of the release (following Francis et al. 2003). If multiple bursts were present, the final one was chosen (following Cho and Ladefoged 1999). This process was partially automated with a Praat script searching for sudden increases in amplitude, but the results required extensive manual correction. An example of a segmented /t/ release is shown in Figure 5.1. 205 tokens were excluded during this process if there was no discernible closure phase. The distribution of stops by phoneme is shown in Table 5.1, along with the mean duration of the release for stressed and unstressed tokens, which is equivalent to positive voice onset time (VOT).

In some cases, the mean VOT values differ quite dramatically from those reported by Mortensen and Tøndering (2013) on the basis of DanPASS (see Section 2.3.1). This is likely because Mortensen and Tøndering follow Fischer-Jørgensen and Hutters (1981) in considering the onset of higher formants to be the relevant landmark for measuring VOT rather than the first signs of periodicity; this strategy leads to higher overall values, particularly for /k/. The VOT measurements are discussed in more detail, and compared to similar measurements from traditional regional varieties of Danish, in Section 6.5.2.1.

A Praat script was subsequently used to extract the release duration and information about the phonetic context for each token. The

Table 5.1: *Number of aspirated stops included in the study, along with mean VOT values. First and third quantiles are given in parentheses.*

Phoneme	**Number**	**Mean duration (stressed), ms**	**Mean duration (unstressed), ms**
/p/	642	57 (42–70)	41 (27–50)
/t/	850	79 (60–92)	68 (53–79)
/k/	842	59 (43–69)	46 (36–54)

phonetic context is coded four binary variables concerning vowel HEIGHT, BACKNESS, and ROUNDING, as well as STRESS. For this purpose, as motivated in Section 4.4.1.1, [i y u ɪ ʏ ʊ e ø o] are all defined as high vowels. [u ʊ o ʌ ɔ ɑ ɒ] are the relevant back vowels, and [y u ʊ ø o œ ɔ ɶ ɒ] are the relevant rounded vowels.

Each release was divided into 20 equally long time steps. This is too coarse-grained to tease apart very dynamic sequences, such as the segue from initial transient to frication, but should be fine-grained enough to capture gross changes in affrication and aspiration. The recordings were filtered to include a frequency range between 500–12,000 Hz. Frequencies below 500 Hz were removed to avoid a potential influence of intrusive voicing or low frequency ambient noise. Frequencies above 12 kHz were removed because they rarely play a role in speech. In fact, 12 kHz is a relatively high cut-off point compared to other comparable studies; this is motivated a study on sociolinguistic variation in Danish /t/ which showed that mean COG for fronted realizations of /t/ can go above 6 kHz, suggesting that very high frequencies may occasionally play a role in /t/ releases (Pharao and Maegaard 2017). For each time step, the four first spectral moments were also extracted; the spectral moments are not used in the analysis, but are available alongside the other data used for the analysis (Puggaard-Rode 2022a).

Multitaper spectra for each time step were generated in R (R Core Team 2021; RStudio Team 2022).[5] Compared to spectra computed using fast Fourier transformation (FFT), such as those computed in Praat, multitaper spectra provide a lower variance spectral estimate which make them suitable for spectra that are noisy and highly dynamic (Blacklock 2004; Reidy 2015). 3 tokens of /k/ were excluded because the total duration of their release was below 10 ms, and the algorithm used to generate the spectra does not work for sound files shorter than 0.5 ms. The dependent variables in the statistical models are the multitaper spectra; each of these consists of a vector of amplitude values along the frequency domain. The frequency ranges differ in size depending on the duration of the time step; longer time steps result in more fine-grained spectra, and thus smaller frequency ranges. Within each spectrum, the amplitude measurements were standardized,[6] since plenty of non-linguistic factors can lead to deviations in overall amplitude level. Note that the multitaper spectral analysis returns intensity values in watt per square meter (W/m^2) rather than amplitude values in the more common logarithmic decibel (dB) scale. I use the W/m^2 scale for this study, as statistical results proved similar across scales, but visualizations are more readily interpretable when using the W/m^2 scale. Only the frequency range between 500–10,000 Hz was used for the statistical analysis of /t/ spectra, and 500–8,000 Hz for /p k/, since the minor activity above these limits seemed to be essentially random noise, and interfered with the clarity of the results.

[5] This was done using the add-on packages `tuneR` (Ligges 2021) and `multitaper` (Rahim 2014; Rahim and Burr 2020), with convenience functions based on code from Reidy (2013, 2016b).

[6] The amplitude measurements were standardized by subtracting the mean and dividing by two standard deviations, following Gelman and Hill (2006).

5.5.2 Statistical analysis

All statistics were calculated in R (R Core Team 2021; RStudio Team 2022) with a number of add-on packages.[7] Separate FOSR models were fitted for each stop with multitaper spectra as the dependent variables. The spectra are smoothed using P-splines with the number of basis functions for the global intercept set as the mean number of amplitude observations per spectrum (corresponding to 32 for /t/, 19 for /k/, and 17 for /p/). This seems to strike a good balance between signal and noise. For the functional responses, 6 basis functions were used for the /t/ model and 5 for the /k/ and /p/ models, guided by the selection procedure proposed by Pya and Wood (2016). P-splines are useful for data sampled on uneven grids (Wood 2017b). Normalized time is included as a non-linear independent variable, smoothed with thin plate regression splines (Wood 2003) with 16 basis functions to ensure high granularity in the temporal dimension. Smoothing penalization parameters were automatically selected using fast restricted maximum likelihood estimation (fREML; Wood 2011). The residuals for the models are reasonably normally distributed,[8] although for the /p/ model, they are somewhat leptokurtic (kurtosis = 5.45); however, Gaussian models with a high number of observations should be quite robust to violations of normality (e.g. Knief and Forstmeier 2021).

A major advantage of GAMMs is the ability to account for autocorrelated residual error (Baayen et al. 2018; Wieling 2018); for example, measurements taken at adjacent steps in a time series are likely to be correlated simply because they are adjacent, which adds unwanted structure to the model residuals. This also applies to adjacent amplitude values in the frequency domain. One way to correct for this is by setting a ρ-parameter, often corresponding to the autocorrelation at 'lag-1', i.e. the mean correlation between adjacent measurements. This correction, called an AR(1) model, can also be included in FOSR models. AR(1) models are included in all models with ρ set at 0.1 below the

[7]As mentioned above, `refund` (Goldsmith et al. 2021) was used to fit FOSR models. `mgcv` (Wood 2017a, 2021), `itsadug` (van Rij et al. 2020b), and `moments` (Komsta and Novomestky 2015) were used for health checks of the resulting models. `ggplot2` (Wickham 2016; Wickham et al. 2021) was used for visualizations, with added convenience functions from `FoSIntro` (Bauer et al. 2018; Bauer 2021).

[8]See the supplementary data (Puggaard-Rode 2022a) for various residual plots.

lag-1 autocorrelation in a corresponding model with no correction.[9] Autocorrelation along the functional domain in the AR(1)-corrected models is moderate and short-range, and autocorrelation along the temporal domain is relatively moderate and short-range even without correction. Another method for accounting for autocorrelated errors is the use of functional random intercepts, with smoothing parameters set using splines based on functional principal components (Greven and Scheipl 2017a; Bauer et al. 2018; for an introduction to the latter concept, see Section 6.7.1). Pouplier et al. (2017) argue in favor of the latter approach because 1) the influence of random effects can then be more readily decomposed, and 2) the basis for the correction is computed directly from the data, while the parameter setting used for AR1-correction is necessarily somewhat *ad hoc*. The latter approach can also be implemented in `pffr`, but at a significant computational cost.

The models further include by-category smooths for a number of independent binary variables: speaker SEX, following vowel HEIGHT, BACKNESS, and ROUNDING, as well as STRESS. The influence of speaker sex on the spectral profile has not been discussed much above, but is also included here, since previous studies have shown a gender effect on the spectral profile of fricatives (e.g. Stuart-Smith 2007). I am interested only in how these variables affect the time-varying characteristics of spectra, so no main effects were included for these variables. The binary variables are contrast coded (see Schad et al. 2020 and Section 4.5.3), such that absence of the feature in question is coded numerically as -½ and the presence as +½; the SEX variable is (randomly) coded as -½ female, +½ male. Contrast coding categorical variables is similar to centralizing continuous variables, and ensures that the global intercept corresponds to a weighted global mean, which makes the final results much easier to interpret. For each of these

[9] The reason for setting ρ lower than the autocorrelation at lag-1 is that all models show some degree of negative autocorrelation at higher lags, which is exacerbated when ρ is increased; see more details in Puggaard-Rode (2022a).

effects, by-speaker random slopes are also included (except for SEX, which logically cannot vary by-speaker).[10]

As discussed in Section 5.4, I do not report *p*-values for the FOSR models, as they likely reflect the number of observations rather than practical significance. Instead, I explore the model fits through two types of plots: 1) *Spectrum intercepts*, which visualize the functional intercepts of the models, corresponding to an average release spectrum when all other variables are kept at zero. These are not very telling in themselves, but are important for interpreting other effects. The spectrum intercepts are plotted with 95% confidence intervals, computed in the manner proposed by Marra and Wood (2012). 2) Spectro-temporal fits, which visualize the spectrum across time. The interpretation of these is similar to spectrograms; they are 'flipped' spectra, with normalized time along the x-axes, frequency along the y-axes, and greyscale shading indicating differences in fitted amplitude along the time–frequency domains. These visualizations reflect the effect size of different variables. The plots of the main effect of time are computed by combining the functional intercept with the fitted effect of time; the plots of other variables are computed by combining the functional intercept, the fitted (main) effect of time, and the fitted time-varying effect of the variable in question. This means that if the model finds no noticeable effect of time, there will be no noticeable change along the horizontal dimension; if there is no noticeable effect of a particular variable, the plot associated with this variable will be similar or identical to the plotted main effect of time. Since these plots are two-dimensional, visualizing 95% confidence intervals would require separate plots for the upper and lower limits; in order to keep the number of visualizations manageable, I do not include these here, but they can be found online in Puggaard-Rode (2022a). These plots demonstrate the uncertainty associated with each fitted effect. I will refer to these plots only when they show that a variable is associated with a great deal of uncertainty.

[10]Using factor smooths instead of random slopes would have given a more thorough estimation of the by-speaker variation in the data (Baayen et al. 2018; Wieling 2018; Sóskuthy 2021), but unfortunately these cannot currently be fitted with data along uneven grids.

Another way to get an indication of the fitting–complexity trade-off of including an individual variable is by comparing the minimized smoothing parameter selection scores (fREML scores) of a nested model without that variable (van Rij 2016; van Rij et al. 2020a). fREML scores are conceptually similar to information criteria like the Akaike Information Criterion (AIC): a lower fREML score indicates a better model fit.[11] For each variable in each model, I report the increase in fREML score of a nested model 1) without that variable and its associated random slope, and 2) without its associated random slope only. If a variable is associated with a large fREML decrease, this means that including the variable results in a much better model fit, i.e. the variable is very influential. This gives an indication of the relative effect size of each variable, and (in the case of random effects) how much of this can be accounted for with by-speaker variation. Note, however, that it is only meaningful to compare fREML scores within the same model, and not across models; fREML scores can be taken as a proxy for relative effect size, not for statistical significance.

5.6 Results

The results of the three different models will be presented in separate sections below, starting with the model for /t/.

5.6.1 /t/

The model of /t/ releases has a high effect size of R^2 = 0.54. The functional intercept (see Figure 5.2) shows an energy peak around 3.5–5 kHz, with comparatively little energy elsewhere, particularly above 8 kHz. Recall that the intercept summarizes the grand weighted mean over a dynamic series of events, so it is not in itself very meaningful. In the spectro-temporal fits, any changes on the horizontal dimension are a result of spectral characteristics changing as a function of time.

The /t/ model shows a strong main effect of time in the expected direction, as shown in Figure 5.3. Initially, energy is skewed towards the higher end of the spectrum, with fairly strong energy around the

[11] AIC does not provide a reliable test for smooth variables (van Rij 2016).

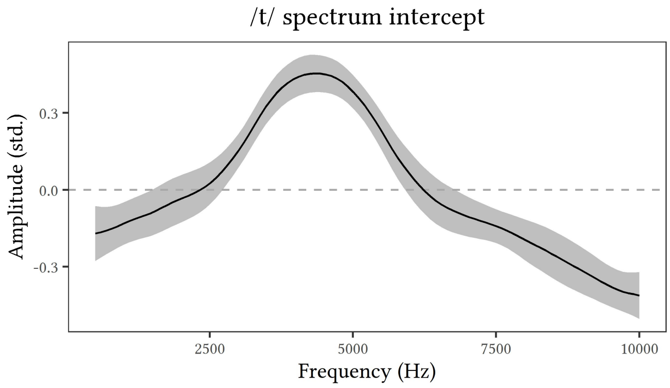


Figure 5.2: *Functional intercept for the model of* /t/ *releases.*

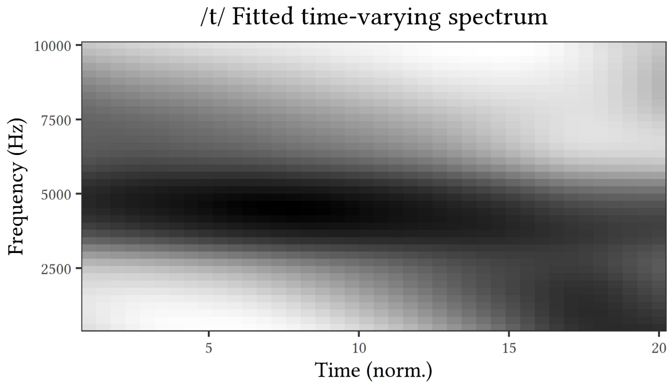


Figure 5.3: *Fitted time-varying spectrum of* /t/ *(main effect of time).*

intercept but also reasonably equal distribution of energy in the 5.5–8 kHz range. Increased energy above the main peak gradually tapers off, and in the final three-fourths of releases, energy is broadly distributed below 5 kHz, including at the lowest frequencies visualized (500 Hz).

Spectro-temporal fits for each direction of the individual variables are shown in Figure 5.4. Table 5.2 shows the reduction in fREML score associated with each variable. Figure 5.4 reflects a residual issue with this modeling technique. In contexts where we expect reduced affrication and earlier onset of aspiration, as in e.g. non-high vowels

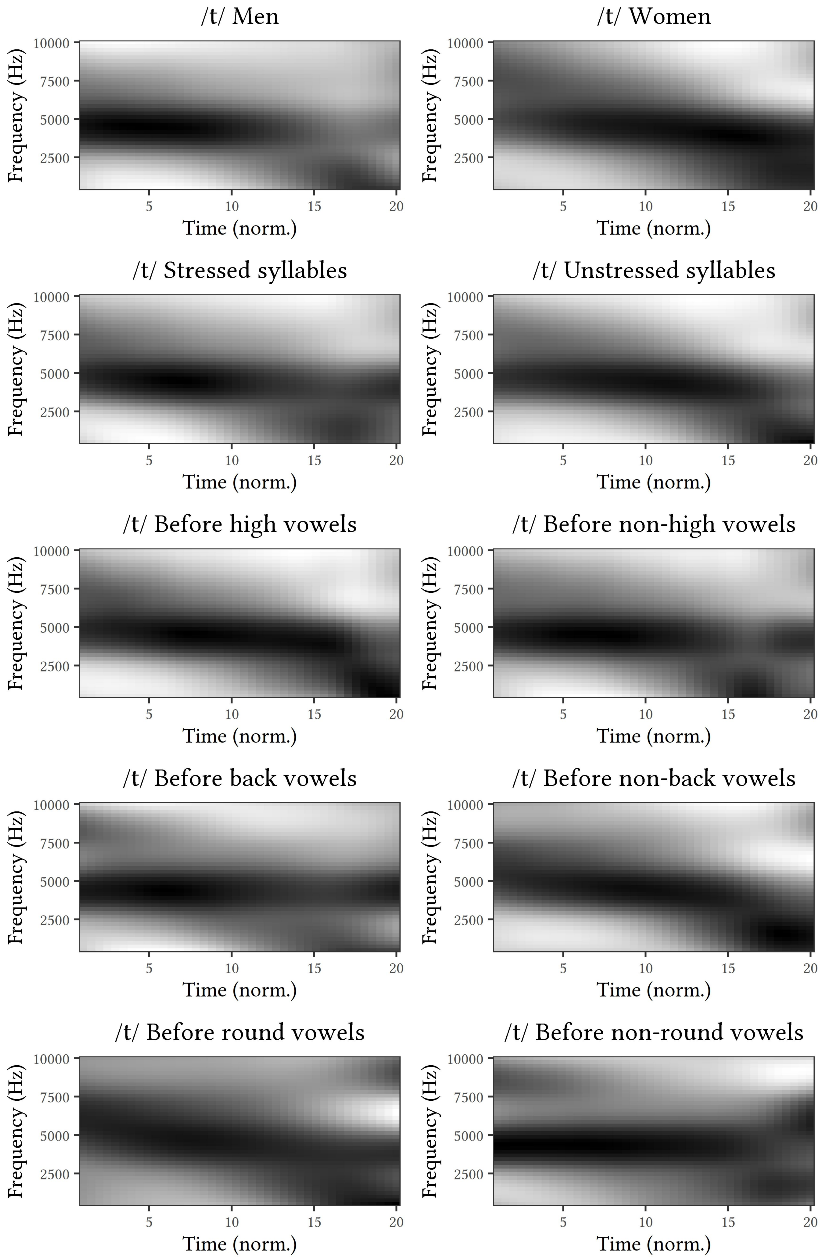


Figure 5.4: *Spectro-temporal fits of* /t/ *for each direction of the individual variables.*

Table 5.2: *fREML score reduction associated with each variable in the* /t/ *model.*

Variable	fREML reduction
SEX	1,539
STRESS	1,875
STRESS (random slope only)	1,546
HIGH VOWEL	2,264
HIGH VOWEL (random slope only)	1,384
BACK VOWEL	913
BACK VOWEL (random slope only)	648
ROUND VOWEL	286
ROUND VOWEL (random slope only)	174

relative to high vowels, the figures show a relatively early increase in energy at low frequencies, but also tend to show a sudden final increase in energy at higher frequencies. There is no linguistic reason to expect this, and it is consistent across models; I assume that this is a technical issue that does not reflect the data or the linguistic reality.

Overall, men show relatively little energy above the peak in the intercept spectrum, and lower frequencies (indicative of aspiration)[12] begin dominating relatively early. Women show strong initial energy in frequencies above 5 kHz, and although lower frequencies come into play late in the release, frequencies up to 5 kHz are excited throughout the release. The effect of SEX is strong and associated with a large reduction in fREML score.

Lower frequencies start dominating towards the end of the release in unstressed syllables, and much earlier in stressed syllables. STRESS is an influential variable, although much of its influence is due to the by-speaker random slope. Lower frequencies also dominate relatively late before high vowels, and frequencies above 6 kHz are also more excited at the beginning of the release in this context. This is a very

[12] As mentioned in Section 5.2, during aspiration, low-frequency noise is generated at or near the glottis, and the turbulent airstream excites the resonant frequencies of the oral cavity. The dominance relationship between these sources may differ, but in both cases, the primary frequencies being excited are well below those excited during alveolar frication.

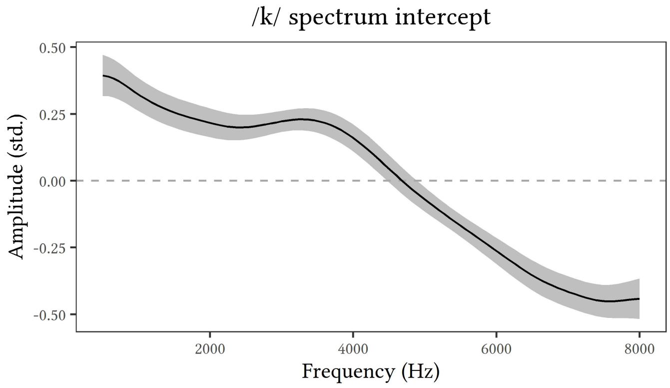


Figure 5.5: *Functional intercept for the model of* /k/ *releases.*

strong effect, which is relatively stable across speakers (i.e. the random slope contributes fairly little.) Lower frequencies dominate relatively early before back vowels and round vowels. In both of these contexts, there is also a coarticulatory effect at the start of the release: relatively high frequencies are excited before round and non-back vowels. These variables are less influential, with the fitting–complexity trade-off being relatively poor for the ROUND VOWEL variable in particular.

It is interesting that none of these variables are particularly influential around the middle portion of the release; they may affect whether particularly high frequencies are excited around the start of the release, and whether/when lower frequencies begin to dominate near the end of the release, but high energy in frequencies around 3.5–5 kHz in the middle of the release is a consistent feature across all variables.

5.6.2 /k/

The model of /k/ releases has a high effect size of $R^2 = 0.57$. Recall that the frequency range for the models of /k/ and /p/ does not extend above 8 kHz. The functional intercept (see Figure 5.5) shows almost evenly distributed energy below 4 kHz, with small peaks around 500 Hz and

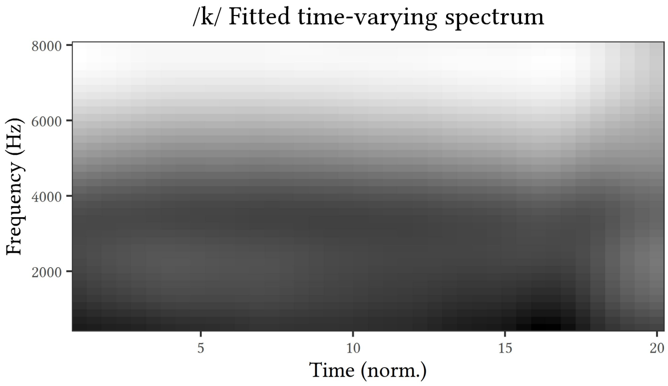


Figure 5.6: *Fitted time-varying spectrum of* /k/ *(main effect of time).*

Table 5.3: *fREML score reduction associated with each variable in the* /k/ *model.*

Variable	fREML reduction
SEX	1,708
STRESS	577
STRESS (random slope only)	507
HIGH VOWEL	6,082
HIGH VOWEL (random slope only)	4,785
BACK VOWEL	17,829
BACK VOWEL (random slope only)	13,502
ROUND VOWEL	3,620
ROUND VOWEL (random slope only)	3,231

just below 4 kHz, and linearly decreasing energy between approx. 4–7 kHz.

There is no strong main effect of time; there is little variance in the time domain in Figure 5.6, and the variance that we do see is associated with significant uncertainty (as evidenced by the 95% confidence intervals shown in Puggaard-Rode 2022a). Spectro-temporal fits for each direction of the individual variables are shown in Figure 5.7. Table 5.3 shows the reduction in fREML score associated with each variable.

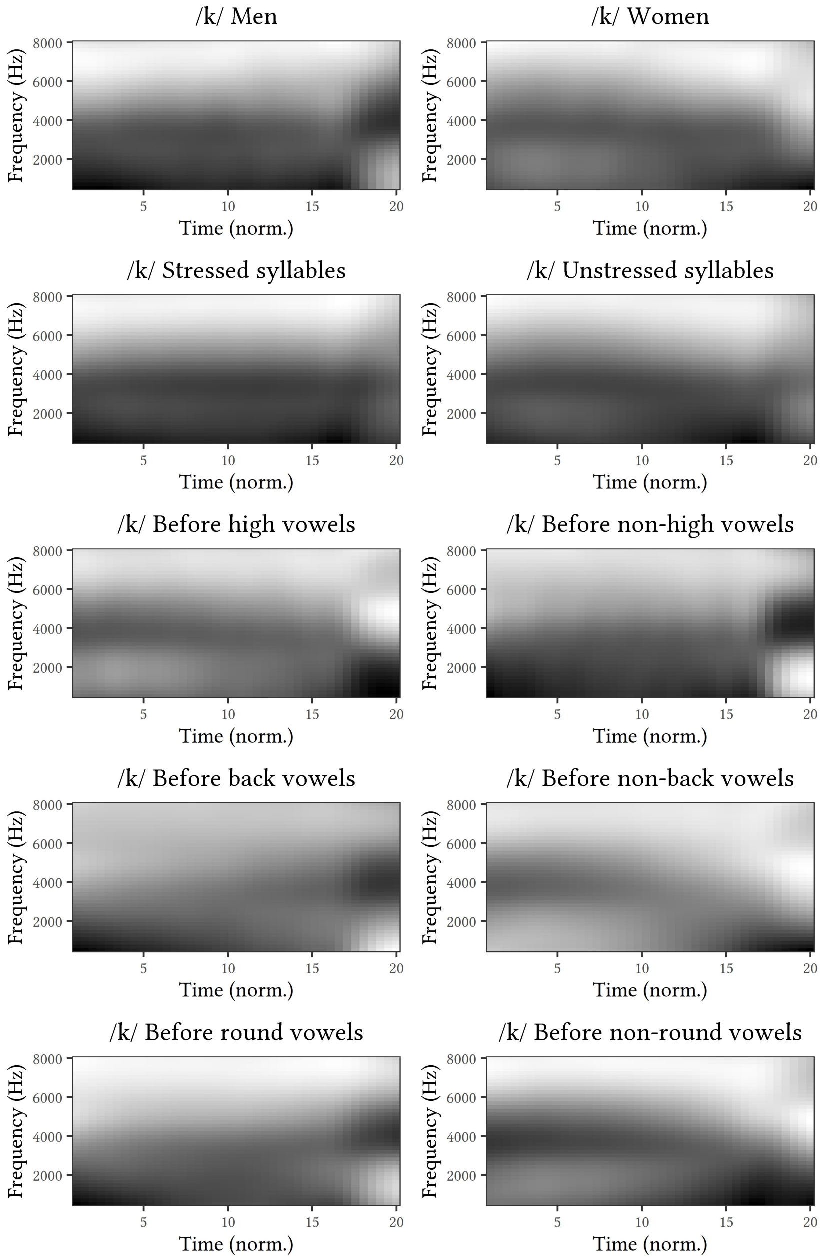


Figure 5.7: *Spectro-temporal fits of* /k/ *for each direction of the individual variables.*

There is a noticeable sex difference, although the associated reduction in fREML is modest. There is little energy at lower frequencies during the first half of the release for female speakers, and more activity at frequencies above 4 kHz. Lower frequencies becomes dominant in the last quarter of the release for female speakers, whereas for male speakers, they are seemingly dominant throughout the release.

As expected, phonetic context effects have a clear influence on the /k/ spectral trajectory, particularly those effects that reflect properties of the following vowel. Stressed syllables have somewhat more energy at the lower band around 500–1,000 Hz, while unstressed syllables have more energy at the higher band around 3.5–4 kHz, although lower frequencies gradually become dominant in the latter half of the release. The size of this effect is modest, and mostly comes down to by-speaker variation; it is also associated with significant uncertainty, as evidenced by 95% confidence intervals (see Puggaard-Rode 2022a).

Before high vowels, there is a lot of high frequency energy between 3–5 kHz during the first half of the release, with more diffuse distribution of energy before the onset of low-frequency noise towards the end of the release; low frequency energy overall dominates releases before non-high vowels. This variable is associated with a large fREML reduction. Non-back vowels and non-round vowels show roughly the same patterns as high vowels, although with slightly varying temporal alignment. The BACK VOWEL variable in particular is associated with a very large fREML reduction. High frequency noise lasts somewhat longer for non-round vowels than non-back vowels. The fREML reduction associated with the ROUND VOWEL variable is also relatively large, although largely a result of by-speaker variation.

5.6.3 /p/

The model of /p/ releases has a very high effect size of R^2 = 0.71. The functional intercept (see Figure 5.8) shows most energy in the lowest frequencies around 500 Hz, with energy gradually reducing at higher frequencies. Assuming that the more diffuse distribution of noise towards the end of the release is not linguistically relevant, there is only a very marginal main effect of time (see Figure 5.9).

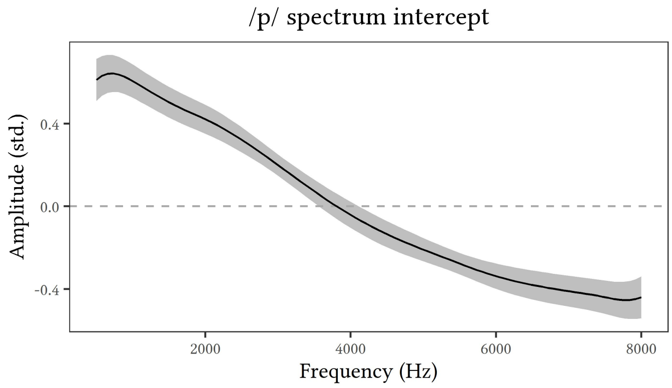


Figure 5.8: *Functional intercept for the model of* /p/ *releases.*

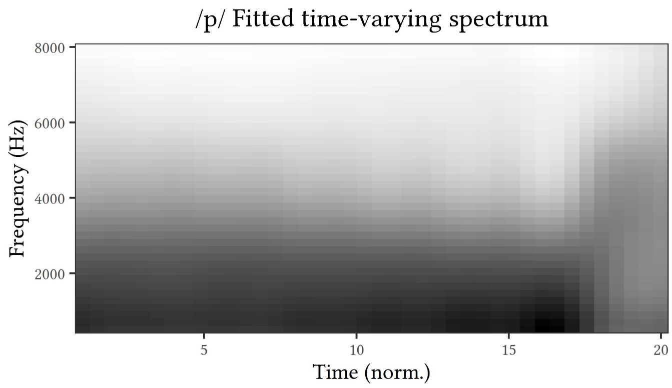


Figure 5.9: *Fitted time-varying spectrum of* /p/ *(main effect of time).*

There are clearer by-variable time-varying characteristics of /p/, as shown in Figure 5.10. Table 5.4 shows the reduction in fREML score associated with each variable. Compared to the other models, random slopes account for a large proportion of the variance in /p/ releases. There are modest signs of higher frequencies being excited more in the first half of releases produced by women, but not by men. The SEX effect is, however, quite weak, and associated with a great deal of uncertainty, as evidenced by 95% confidence intervals (Puggaard-Rode 2022a).

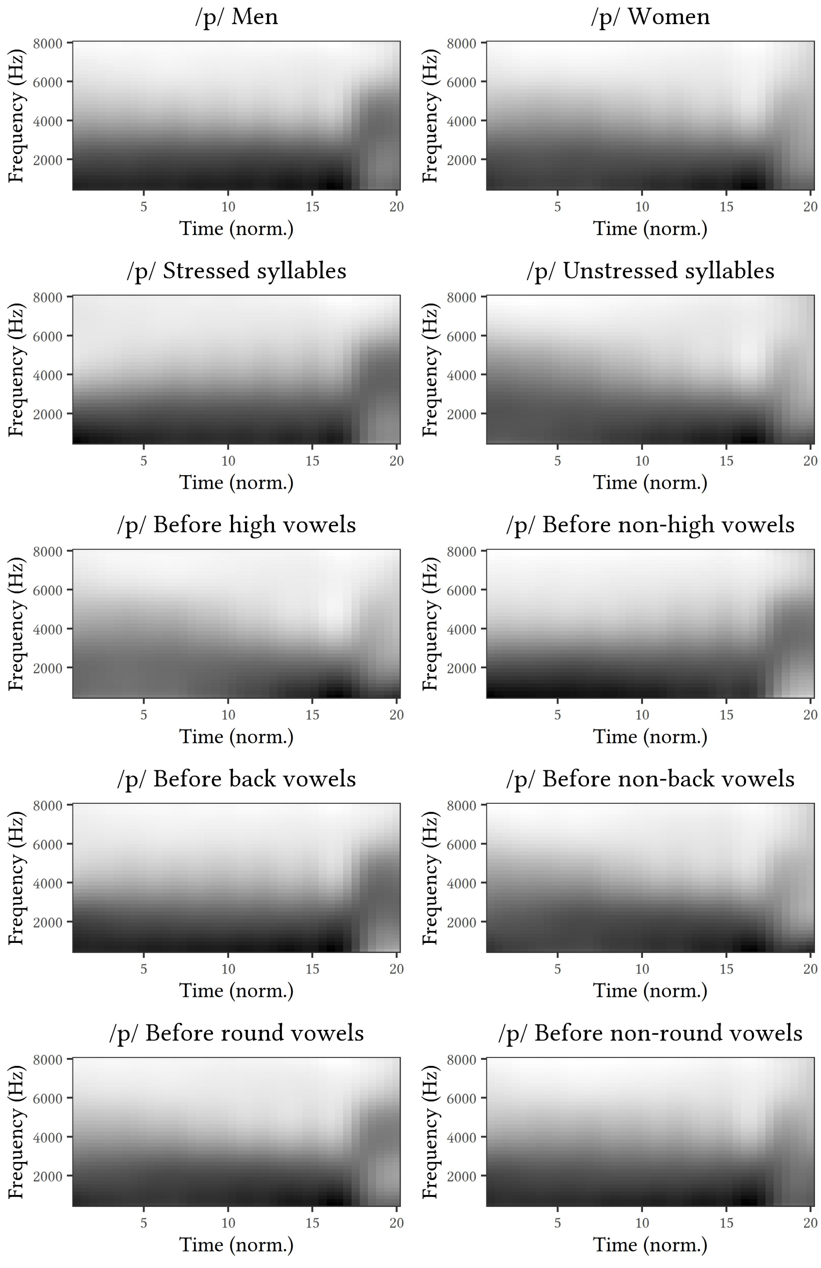


Figure 5.10: *Spectro-temporal fits of* /p/ *for each direction of the individual variables.*

Table 5.4: *fREML score reduction associated with each variable in the* /p/ *model.*

Variable	fREML reduction
SEX	189
STRESS	1,531
STRESS (random slope only)	1,248
HIGH VOWEL	1,123
HIGH VOWEL (random slope only)	1,035
BACK VOWEL	911
BACK VOWEL (random slope only)	804
ROUND VOWEL	494
ROUND VOWEL (random slope only)	460

During the first portion of the release, unstressed tokens have a broader distribution of energy throughout the spectrum, and more energy at higher frequencies (above approx. 5 kHz). The STRESS variable is quite strong, and relative to other variables, quite robust across speakers. A similar pattern is found before high vowels, with lower frequencies dominating relatively late in the release. To a lesser extent, the same pattern is found before non-back vowels. Both of these effects are associated with large fREML reductions, but largely due to by-speaker variation. There is no obvious influence of round vowels, and this variable is associated with significant uncertainty, as evidenced by 95% confidence intervals. The modest fREML reduction associated with the ROUND VOWEL variable is almost exclusively due to by-speaker variation.

5.7 Discussion

5.7.1 Contextual variation in stop releases

In Section 5.6 above, I described the patterns of energy distribution that are visible in the spectro-temporal fits in prose. In this section, I aim to provide a link between those representations and the articulatory mechanisms that presumably underlie them. This discussion is necessarily somewhat speculative, but relies on established knowledge about the articulation–acoustics link, and about the articulation of Danish specifically.

While all stops show diffuse patterns of energy distribution towards the end of the release, only /t/ clearly shows a strong main effect of time, with a gradual downward trend in energy distribution over time. During the first half of the release, high frequencies are excited, often above and beyond what is necessarily expected for an alveolar constriction. During the second half of the release, lower frequency energy consistent with a glottal noise source gradually becomes dominant. As mentioned in Section 5.2 above, there is reason to assume that oral air pressure is not particularly high at the time of release in Danish aspirated stops, which provides both an aerodynamic reason and a functional–phonological motivation for why the constriction is maintained somewhat longer than in comparable 'aspiration languages': there is no high air pressure to ensure that the constriction is quickly released, and to ensure a salient burst. Nevertheless, contrary to the general conception in literature, alveolar constriction usually does not dominate the entire release.

The relative timing of the shift in dominance from an alveolar noise source to a glottal one is largely determined by contextual factors, in particular stress and vowel height. Speaker sex also plays a role. Stop releases in stressed syllables show a larger proportion of aspiration. In other words, phonetic reduction mainly targets the aspiration in /t/ releases, not the frication. Features of the following vowel affect the relative timing of the dominance shift much more than they affect the distribution of energy during the first half of the release, although high and round vowels do show coarticulatory effects lasting throughout the release. The linguistic upshot is that lengthy alveolar frication is an invariant feature of /t/ releases in Modern Standard Danish, but the proportion of alveolar frication varies; some degree of aspiration is almost always observed.

Stevens' (1998) model of velar stop releases suggested that the velar frication excites low resonance frequencies mostly below 2 kHz. The results here, however, show two primary patterns of energy distribution: much higher resonance frequencies around 4 kHz, or resonance frequencies centered around the lower end of the spectrum. I presume that the former represents a velar noise source – likely fronted, since a fronted velar constriction leads to a shorter distance between the constriction and the hard palate, which the turbulent air stream partially impinges on – and that the latter corresponds primarily to a glottal noise source. However, it may be difficult to tease apart a noise source in the back portion of the velum and a glottal noise source. The dominant noise source is mostly contextually determined. The main effect of time is marginal, although low-frequency aspiration is overall dominant during the final portion of the release. Before high vowels and non-back vowels in particular, noise at higher frequencies is dominant during the first part of the release. If the following vowel is high, the tongue dorsum logically remains fairly close to the velum throughout the release, causing a dominant dorsal noise source, the characteristics of which vary on the basis of other vowel features. The point of occlusion varies by backness of the following vowel, such that the outgoing air impinges more directly on the hard palate before front vowels, causing more salient noise at higher frequencies. The energy from the palatal noise source is dampened by lip rounding, which increases the size of the oral cavity. The linguistic upshot is that coarticulation has a major influence on spectral characteristics throughout /k/ releases; this is in line with the general observation that the point of occlusion in /k/ is prone to coarticulatory variation (e.g. Ouni 2014).

/p/ releases also vary in whether there is a primary glottal noise source (a strong energy peak at lower frequencies), or whether there is a primary labial noise source (no strong energy peak at lower frequencies). There is no strong main effect of time. In unstressed syllables, before high vowels, and to some extent before non-back vowels, energy is more broadly distributed throughout the spectrum, indicating a dominant labial noise source. /p/ releases vary relatively little compared to /t k/, and much of the variance found in the data is the result of by-speaker variation.

These results confirm the observation that /t/-affrication in Modern Standard Danish is invariant. Generally, however, /t/ affrication does not last throughout the release; aspiration is also an important component of /t/ releases, especially in stressed position. There is also a frication component in /p k/ releases, but under many conditions, these releases are dominated by a glottal noise source. During a /t/ release, the outgoing air impinges on a hard surface – the teeth – immediately downstream of the preceding occlusion. This is not the case for either /p/ or /k/; the lips constitute a soft surface, and the hard palate is further removed from the velar occlusion. As such, it is well-understood why an alveolar noise source dominates a glottal one more readily than corresponding bilabial or velar noise sources.

5.7.2 Function-on-scalar regression and the spectrum

This chapter has introduced the use of FOSR in the analysis of speech spectra and their variance as a function of time. This method shows a lot of promise. It allows us to get around the problem of choosing one or a few discrete measures to represent the spectrum, all of which come with their own set of methodological problems. In a sense, analyzing these models is similar to the classical technique of 'eyeballing' spectrograms, but in a way that allows the user to more efficiently and reliably find systematic patterns of variation in the data, to tease apart various influences on the results, and to filter out by-speaker variation. Some lingering issues remain with the method; some specific to this study, and some inherent to the field. I will briefly address a few of these.

As with any kind of quantitative phonetic study, there are significant researcher degrees of freedom involved in FOSR modeling of spectra (see Roettger 2019). Token selection, spectral estimation, smoothing procedure, low-level software implementation, as well as several other factors all have a potentially non-trivial influence on the results. There is no easy remedy to this, but transparent reporting and motivation of all these choices goes a long way. I have aimed to do that here, and the actual code used to implement the analysis is available in annotated form (Puggaard-Rode 2022a).

FOSR modeling of spectra quickly leads to highly multidimensional data, especially if the temporal dimension is also taken into account, and this makes the use of traditional methods for significance testing problematic. I do not consider this to be an issue in the current study. For one, the study is largely exploratory, and the research questions are not necessarily suitable for null hypothesis significance testing. With that said, there are methods for testing the stability of the results. This includes the 95% confidence intervals proposed by Marra and Wood (2012), which I have occasionally referred to here, and include in the online appendix to this chapter (Puggaard-Rode 2022a); this method is implemented for FOSR visualization in the `FoSIntro` package in R (Bauer 2021). Additionally, there are functional implementations of discriminant analysis and regression trees which may be used to explore the generalizability of results, and fully Bayesian implementation of the analysis would make it possible to readily quantify the uncertainty related to the results (see e.g. Vasishth et al. 2018b). This will hopefully be explored in future research, but is beyond the scope of the current study. The prospects of hypothesis testing in FOSR models is explored in a recent dissertation by Biswas (2022).

The implementation of FOSR in this study shares a problem with analyses based on e.g. spectral moments, mid-frequency peaks, and DCT: the Hz-based frequency scale and the W/m^2-based amplitude scale are 'physicalist' in nature, in that they represent the behavior of vibrations in the air, and not how these vibrations are perceived by the human ear (Plummer and Reidy 2018). I use the Hz scale here because it results in a model output which is more immediately interpretable for readers with experience with analyzing spectrograms; I use the W/m^2 scale because it results in more clearly interpretable patterns in the fitted models. It is, however, worth exploring in future studies how the results would be affected by combining perceptually motivated scales, such as the equivalent rectangular bandwidth (ERB) scale and the decibel scale.[13]

The most serious lingering issue is the diffuse patterns sometimes seen in the final time steps of the spectro-temporal visualizations. These cannot be considered linguistically meaningful; there is no linguistic reason why high frequencies above 4 kHz would suddenly be excited immediately before the onset of voicing in a stop–vowel sequence. I can see three possible explanations for this: 1) the spectral

[13] Alternatively, the positions of knots used for smoothing could be placed according to a (semi-)logarithmic scale, e.g. giving the model higher granularity in frequency regions where humans have greater perceptual acuity. This could potentially achieve a similar effect while keeping the 'physicalist' scales. In this study, the knots are equidistantly spaced, but `mgcv` and consequently `pffr` allow the user to specify knot locations freely.

characteristics of aspiration are highly variable, making it impossible for the model to make precise predictions, 2) the pseudo-centralization of categorical variables sometimes causes the model to infer patterns that are not meaningful for one value of variables, or 3) it is caused by phase variation. Regarding 2), consider /k/ before high and non-high vowels: the model finds a strong increase in low frequency energy in the final time steps before high vowels, which is linguistically meaningful, as the glottal noise source becomes dominant immediately before the onset of voicing. The model finds a corresponding increase in high frequencies and decrease in low frequencies in the final time steps before non-high vowels, which is *not* linguistically meaningful, but is the direct opposite of the meaningful finding before high vowels. A possible solution would be to fit the model without contrast-coded categorical variables, but this would make it impossible to interpret models' intercepts and main effects of time, which I believe would seriously harm the interpretability of the findings. Regarding 3), phase variation is a practical problem in functional data analysis, where lateral displacement in input curves can cause results to be blurred and distorted. Managing phase variation in the analysis of functional data is an area of active research (Marron et al. 2015; Bauer et al. 2021)

5.8 Conclusion

The study presented in this chapter is, to the extent of my knowledge, the first to use function-on-scalar regression to analyze sound spectra. This method forgoes the need to boil down the complex, multi-dimensional information in the spectrum to a few discrete values, and it forgoes the need to rely on 'magic moments' in time. By plotting the fit of a FOSR model, we can explore the systematic influences of different variables on the spectrum with visualizations that should be intuitively familiar to anyone used to working with spectrograms. I showed how this tool can be fruitfully applied in the analysis of Danish stop releases, how their spectral characteristics vary vary over time, and how they are affected by their phonetic environments.

The analysis finds that /t/, as expected from the literature, is invariably affricated, but also that the spectrum is very dynamic throughout /t/ releases, with dominant affrication gradually being replaced by dominant aspiration. Affrication dominates the majority of the spectrum, and much of the aspiration is lost in unstressed syllables. Coarticulatory context effects may affect the entirety of /t/ releases, and not just the final portion. Coarticulatory context effects greatly influence the spectra of /k/ releases, particularly in the first portion of the release. The acoustic characteristics of /p/ releases show a lot of by-speaker variation, but also coarticulatory context effects, mainly in the first half of the release.

CHAPTER 6

Regional variation in stops

6.1 Introduction

An overt feature of varieties of Jutlandic Danish is the use of a variant of /t/ known colloquially as *tørt t*, ‘dry t’. As we saw in the previous chapter, the Standard Danish variant of /t/ invariably has an affricated release; the dry t does not. There are different opinions in the literature as to which varieties use the dry t, ranging from “all the country’s dialects” (Brink and Lund 1975: 353), western Jutlandic (Petersen et al. 2021: 156ff.), conservative Standard Danish (Grønnum 2005: 51), and northern Jutlandic (Petersen 2009b); see Section 2.5.3. I have previously shown in a pilot study that variation in the realization of /t/ goes beyond just affrication, and is difficult to delimit geographically: the dry t also has shorter voice onset time (VOT) than affricated variants, and a short variant of /t/ with little affrication is found in both the

Parts of this chapter report collaborative work with Yonatan Goldshtein. Section 6.2.2 is partially based on Goldshtein and Puggaard (2019), and Section 6.5 is a revised version of Puggaard (2021). The account presented here has been developed throughout a number of presentations (Puggaard 2019b, 2019c, 2020a, 2020b; Puggaard and Goldshtein 2019, 2020). The audio data is freely available (DS 1971–1976), as are replication data and code (Puggaard-Rode 2022a).

north and the center of the Jutland peninsula (Puggaard 2018a). In this chapter, I expand on that study with much more data, a more suitable statistical methodology, and a broader focus on geographical variation in stop realization. I focus specifically on VOT and spectral characteristics, comparing VOT measurements and spectra from speakers throughout Jutland. The data come from a large corpus of legacy recordings which preserves an older stage of regional variation in Danish (Andersen 1981b; Pedersen 1983). Parts of the corpus have been used as a source for the Dictionary of Insular Dialects (ØMO 1992–; cf. Gudiksen and Hovmark 2008), but the recordings from Jutland have never before been used systematically for research.

A section of this chapter is dedicated to introducing this corpus, *Dialektsamlingen* 'the dialect collection' (DS 1971–1976). This means that I will briefly divert from the dissertation's main focus in the first sections of this chapter. I do this for two reasons: First, an understanding of the dialect situation in Denmark and the historical context for the recordings is important both in order to understand the results of this chapter, and in order to understand why a study like this is needed in the first place. Second, this is the first introduction to the corpus written in English.

There are many descriptions of Danish dialects, including partial dictionaries, grammars, (morpho-)phonological descriptions, and topical descriptions of individual dialects (for an overview, see Ejskjær 1993). There are also holistic descriptions of the Danish dialect landscape which define geographical boundaries between dialects, primarily on the basis of isogloss bundles (Bennike and Kristensen 1898–1912; Brøndum-Nielsen 1927; Skautrup 1944–1970: IV). With few exceptions, however, the descriptive work has lain dormant since the 1970s, leaving much of the existing work theoretically dated (see Section 6.2.1).[1] As a consequence, methodological progress in acoustic phonetics has barely improved our knowledge of regional phonetic variation (although see Ejstrup and Hansen 2003; Ejstrup 2010). Our knowledge of phonological variation is rich if spotty, while our

[1]The lexicographic work, however, is still very much ongoing, centered around *Jysk Ordbog* 'Jutlandic dictionary' (Skautrup et al. 1970–; Hansen 2020), and *Ømålsordbogen* 'Dictionary of insular dialects' (ØMO 1992–.)

knowledge of subphonemic systems is much poorer and mostly limited to impressionistic description. Similarly, the recent great strides in statistics for corpus phonology have not been applied to regional variation in Danish.

The initial hypothesis of the research reported here, following Puggaard (2018a), is that the dry t is found in large parts of Jutland, and that cues to place contrasts and laryngeal contrasts show geographical variation. A number of theoretically motivated hypothesis follow.

Chodroff and colleagues (Chodroff and Wilson 2017; Chodroff et al. 2019) have recently shown that variation in VOT tends to co-vary across laryngeal settings and places of articulation. With this in mind, I hypothesize that variation is not limited to /t/, but rather that all stops follow similar patterns of variation. Early findings in VOT research showed that voiced, voiceless, and aspirated stops are cross-linguistically consistent categories (e.g. Lisker and Abramson 1964). Later research has not been able to support this neat, constrained three-way division. As VOT is measured in more languages, typological variation in VOT increasingly looks continuous rather than categorical (Cho and Ladefoged 1999; Ladd 2011), suggesting that the only principal limits on variation in VOT comes from limits on perceptual acuity. The Jutlandic data were gathered in a relatively small geographical area shared by one language community with the same set of distinctive stops, so it is an interesting test case for limits of variation in VOT. I use generalized additive mixed modeling to investigate which aspects of the observed variation are attributable to geography, with the assumption that this relationship is non-linear (Wieling et al. 2011, 2014); using this method also makes it possible to test a number of other hypotheses about the influence of contextual variables on VOT.

Chodroff and Wilson (2018) do not find strong signs of covariation in stop release characteristics across places of articulation, but the study presented in Chapter 5 makes it interesting to test whether there are differences in affrication patterns of /p t k/ across varieties, and whether the Jutlandic varieties pattern similarly to Modern Standard Danish. Comparing spectral characteristics across varieties is much less straightforward than comparing VOT. Using function-on-scalar regression as in Chapter 5 is technically possible, but the added complexity of non-linear geographical variation would make the results prohibitively difficult to interpret. I solve this by focusing on the entire spectrum, but only at the midpoint of stop releases, i.e. removing variation in time from the equation. I use a combination of functional principal components analysis (FPCA) and generalized additive mixed modeling to investigate the main sources of variance in release spectra, and how these relate to geography. To the extent of my knowledge, this is the first study to use FPCA to explore spectral variance. As with the model of VOT, this also allows for testing hypotheses about the influence of contextual variables on spectral shape, and whether the influence of contextual variables is similar in Modern Standard Danish and in the traditional Jutlandic varieties.

Throughout this dissertation, a lot of weight has been placed on the relationship between (known) data about articulatory mechanisms, (new) corpus data revealing acoustic patterns, and how the articulation–acoustics link may come to affect phonological systems. I return to all the previously covered topics in this chapter. I briefly discuss some data on intervocalic voicing in Jutlandic varieties, and compare it to the results in Chapter 4; I discuss variation in the spectral characteristics in stop releases, and compare it to the results in Chapter 5; and I relate these results to the well-known outcomes of stop gradation in these varieties, comparing it to the Modern Standard Danish patterns discussed in Chapter 3. I argue that the default in Jutlandic varieties is to differ from Modern Standard Danish on various parameters: VOT is usually shorter in /p t k/, voicing is more widespread in /b d g/, and (particularly) /t/-affrication is less widespread. When varieties deviate from the default pattern, it requires an explanation; I suggest such explanations for the relatively long VOT in southern Jutlandic, and the /t/-affrication in the area around Djursland.

In Section 6.2, I give a brief survey of the field of Danish dialectology, and a longer introduction to the corpus used throughout the chapter. In Section 6.3, I survey the major dialect areas of Jutland, and in Section 6.4, I discuss the relationship between linguistic variation and geography in traditional dialectology and modern dialectometry. Section 6.5 describes a corpus study of regional variation in VOT in the corpus recordings. In Section 6.6, I present a few non-systematic observations on closure voicing in the corpus. Section 6.7 describes a corpus study of variation in spectral characteristics at the midpoint of aspirated stop releases, based on a subset of the material used for the VOT study. Finally, in Section 6.8, I discuss how the results of the chapter compare with our knowledge of Modern Standard Danish and how they relate to categorical phonology, and speculate on the sources of the observed variation.

6.2 Documentation of Danish dialects

In this section, I give a brief historical account of dialectology in Denmark (Section 6.2.1), and a detailed introduction to the collection of recordings at the dialect research centres in Aarhus and Copenhagen, including but not limited to the DS (1971–1976) corpus, which is the source of the data for the subsequent corpus studies (Section 6.2.2).

6.2.1 Dialect descriptions

There is a long tradition of dialectological research in Denmark. Description of regional varieties has been practised unsystematically for several centuries, and dialectology as a serious scientific discipline has been around since the mid-19th century. For political reasons, coverage of the southern Jutlandic varieties of Danish spoken near the unstable border to Germany were of particular interest; an important early example is Lyngby's (1858) highly detailed grammar. The tendency for meticulous annotation and long word lists increased as phonetics gained popularity as a discipline in the late 19th century; Jensen's (1897–1902) description of the morphology and phonology of Northern Jutlandic is a noteworthy example of this trend.

Towards the end of the 19th century, dialectology throughout Europe took a turn towards more comparative and geographically oriented research (see Section 6.4). In Denmark, this a.o. resulted in a detailed collection of dialect maps by Bennike and Kristensen (1898–1912). The authors were teachers at a *højskole* – a type of boarding school for (mostly young) adults from all over the country – and used their students as informants for the project. Another important publication in this vein is Brøndum-Nielsen's (1927) comparative survey of Danish dialects.

Structuralism (particularly glossematics; see Hjelmslev 1943) became dominant in Danish dialectology in the 20th century. This led to many detailed descriptions of individual dialects, which can be quite difficult to follow for readers unfamiliar with the framework. Key publications include Bjerrum (1944), Jensen (1944), and Andersen (1959). Glossematics was openly uninterested in phonetic substance (Hjelmslev 1943: 46), which is also reflected in these descriptions. A notable exception to the glossematic dominance of this time is found in the work of Skautrup (e.g. 1930, 1944–1970).

In recent years, most work in Danish dialectology has been lexically oriented. There are still two dialect research centres in operation – the Peter Skautrup Centre for Jutlandic Dialect Research in Aarhus, and the Department of Dialect Research in Copenhagen – and both are occupied with long-ongoing dictionary projects (Skautrup et al. 1970–, ØMO 1992–). Otherwise, traditional dialectology has been mostly replaced by variationist sociolinguistics, focusing mainly on urban varieties (e.g. Gregersen and Pedersen 1991) or the waning status of traditional dialects (e.g. Kristensen 1977; Kristiansen 1998; Pedersen 2003; Maegaard and Monka 2019).

For more extensive overviews on Danish dialectology, see Ejskjær (1993) and Hovdhaugen et al. (2000: chs. 4.5.5.1 and 5.6.4.1).

6.2.2 Audio recordings

For as long as it has been possible, the Danish dialect research centers have been recording speakers of traditional dialects for research and posterity (Andersen 1981b; Pedersen 1983). Their combined collections include recordings from as far back as 1934, with speakers born as early as 1849. The collections contain more than 2,100 hours of audio, and recordings from 1,080 different parishes spread throughout the country. Figure 6.1 shows the number of recordings made per year, and Figure 6.2 shows the geographical distribution of recording locations. The corpus is eclectic; it contains sociolinguistic interviews, radio interviews, poetry readings, read word lists, free interaction, musical performances, etc., at very inconsistent levels of audio quality. The

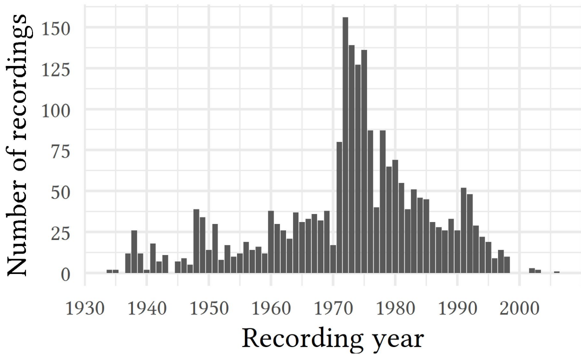


Figure 6.1: *Number of recordings made per year in the combined corpora of dialect recordings.*

corpus maintained by the Peter Skautrup Center for Jutlandic Dialect Research is digitized and freely available on site. These recordings were originally transferred from tape recordings to CD in the 1990s, and later digitized from the CD recordings as MPEG-1 files with a bit rate of 320 kb/s. This format is practical for storage, but the resulting loss of detail is non-trivial (e.g. Bounds et al. 2011); for reference, uncompressed CD quality corresponds to a bit rate of 1411 kb/s. The corpus maintained by the Department of Dialect Research at the University of Copenhagen has not been digitized, but is available on CD on site.

In 1971, the research centers received a grant for a five-year project which aimed at (and almost achieved) gathering tape recordings from every fourth parish in the country (Andersen 1981b; see Figure 6.3). At this time, they already had a broad network of informants, and portable tape recorders had become more readily available. A total of 525 recording sessions took place between 1971–1976, and these recordings are much more uniform than the collection at large. They mostly consist of sociolinguistic interviews with a single informant at their place of residence. Most speakers conform to the NORM (non-mobile older rural males) often found in dialectological studies (Chambers and Trudgill 1998), although the proportion of female informants is higher than commonly found in traditional dialectology. The ages and birth years of the informants in the 1971–1976 recordings are

Figure 6.2: *The geographical coverage of the combined corpora of dialect recordings.*

Figure 6.3: *The geographical coverage of dialect recordings made in 1971–1976 and digitized by the Royal Danish Library.*

Figure 6.4: *Summary of the ages and birth years of informants in the 1971–1976 recordings.*

shown in Figure 6.4. Approximately two thirds of the recordings have one male informant, one fourth has one female informant, and the rest are group interviews with two or more informants.

Early descriptions of the recordings (Andersen 1981b; Pedersen 1983) clearly state that their primary purpose was to serve as material for dialect dictionaries (Skautrup et al. 1970–; ØMO 1992–).[2] A positive effect of this is that topics generally revolve around old cultural customs and methods of agriculture, food preparation, etc. As Hay and Foulkes (2016) show, speech about older events also tends to elicit conservative pronunciation. A negative effect is that phonetics research clearly did not factor into the considerations at the time; little effort is made to avoid background noise and, the speech of informants often overlaps with discourse particles and questions from the interviewers.

A considerable advantage of these recordings is that the original tapes have all been digitized by the Royal Danish Library and are freely

[2]In the end, the recordings have been used only sparingly for the dictionary of Insular Danish (Gudiksen and Hovmark 2008), and the recordings have not been used systematically for research on Jutlandic varieties until very recently (Puggaard 2018a; Goldshtein and Ahlgren 2021).

available online in high quality.[3] These audio files are of much higher quality than those directly available from the dialect research centers.

This section has given an overview of one particular corpus of recordings of regional varieties, but recordings have also been made for several other projects. For example, Ejstrup (2009) gathered materials in connection with his dissertation on regional variation in vowels, and material was more recently gathered in connection with the *Dialekt i periferien* 'Dialect in the Periphery' project (Maegaard and Monka 2019).

6.3 The dialects of Jutland

Modern Standard Danish is now the primary means of communication throughout Denmark (Kristiansen 1998; Pedersen 2003), but Kristiansen (2003a) judges that the majority of the speech community spoke other regional varieties until the 1960s. In the late 1960s, Skautrup (1944–1970: IV:96ff.) wrote that the dialects were in poor condition, and that the most likely features to survive were phonetic ones which would not affect mutual intelligibility. Skautrup noted at the time that this development was quite advanced on the island of Zealand, where the standard variety had mostly spread from Copenhagen to the rest of the island. The development was less advanced in the Jutland peninsula and the smaller islands, but still well underway.

In the 19th century, an obligatory education system was introduced and agrarian reforms led to increased mobility both in cities and rural areas (Skautrup 1944–1970; Kristiansen 2003b). This led to disruption in the traditional regional varieties and the rise of the current standard language, based on the traditional High Copenhagen variety (Kristiansen 2003a). In the mid-20th century, dialect leveling was accelerated through the spread of national broadcasting in Standard Danish, and through government policies enforcing the use of Standard Danish in the education system (Kristiansen 1990); a more detailed overview of these developments is given in Section 1.3. A recent research project shows that dialects are alive and well in parts of southern Denmark (Monka and Hovmark 2016; Monka 2019),

[3] See the bibliography entry for DS (1971–1976).

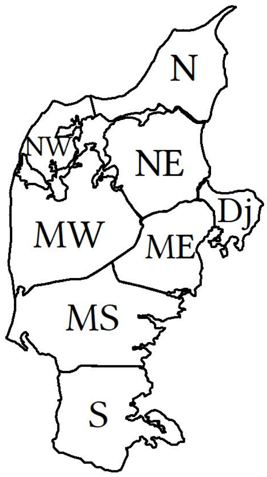


Figure 6.5: *Traditional dialect areas of Jutland as defined by Skautrup et al. (1970–: K.03). N = Northern; NW = North-Western; NE = North-Eastern; Dj = Djursland; MW = Mid-Western; ME = Mid-Eastern; MS = Mid-Southern; S = Southern.*

but that same project also finds complete leveling in other regions that have traditionally been strongly associated with dialect use (Stæhr and Larsen 2019). In yet other regions, regionalized versions of the standard language have replaced traditional dialects (Mortensen 2019). Dialect features may coexist with standard features but take on different social functions that are less geographically delimited than they were in the past (Scheuer et al. 2019).

Figure 6.5 shows the major traditional dialect areas of Jutland as defined by Skautrup et al. (1970–: map K.03).[4] Skautrup (1937, 1944–1970: IV:97) bemoans the fact that there has been relatively little discussion of the basis of these divisions. He maintains that there are no sharp borders between Jutlandic dialect areas. Lines between areas are essentially drawn in transition areas between dialect 'cores' (see

[4]Dialects from this map are written with initial capital letters throughout the chapter. Note that Skautrup et al. (1970–) refer to the two southernmost dialects as *sønderjysk* and *sydjysk* respectively; as there are no fitting English translations for these terms, Southern and Mid-Southern are used here.

Aakjær 1925). Skautrup judges that Danish dialects are mostly defined on the basis of isophones in the form of common phonological developments. Differences in morphology and lexicon also play a role, but less so.

Skautrup (1944–1970: IV:97ff.) provides the clearest diagnostics for the dialect areas. The most important diagnostic is *artikelgrænsen* 'the article boundary', which is responsible for the relatively strict division between eastern and western dialects. In eastern dialects, as in Modern Standard Danish, definiteness in nouns is marked with a suffix [-ən], while in western dialects, it is marked with a phrase-initial article [ə]. This is an exception to the generalization that the defining isoglosses in Danish dialectology are phonological, but note that the article boundary also has major implications for both sentence prosody and segmental phonology (Skautrup 1952). In Section 2.5.3, I discussed a number of further phonological boundaries which involve stops in particular.

6.4 Language variation and geography

Dialectology in Europe became very preoccupied with geography in the late 19th century. Rather than focusing on individual dialects, scholars started drawing detailed maps of distributions of features or lexical items; dialect atlases were drawn of several countries, including Germany (Wenker and Wrede 1895), France (Gilliéron and Edmont 1902–1910), and Denmark (Bennike and Kristensen 1898–1912). In the wake of this work, a debate ensued about whether individual dialects exist at all, or whether geographical variation is of a purely continuous nature (Paris 1888; Gauchat 1903). The conclusion seems to be that although the geographical distribution of features can be chaotic, there are usually areas with adjacent bundles of important isoglosses, and other areas not crossed by significant isoglosses (i.e. dialects do exist). The field of dialect geography has yielded much rich descriptive work, but a common feature of studies from this era is that geography (in a pre-theoretical sense) is typically the only predictor of language variation (Chambers 2000; Britain 2010). Perhaps as a counterreaction, early variationist sociolinguistics (e.g. Labov 1963) was relatively uninterested in geography, with a major exception being the work of Trudgill (e.g. 1974).

Research into the relationship between geography and language variation remains active in the data-driven field of dialectometry (Séguy 1973; see Wieling and Nerbonne 2015 for a recent overview). Dialectometry has made large strides towards estimating the geographical basis of language variation using aggregated features and modern statistical and computational methods. An explicit goal is to estimate how much variation can be explained by geographical distance. By aggregating pronunciations of a large number of words in a single analysis of variation in northern Dutch rather than focusing on well-known loci of variation, Nerbonne and Heeringa (2007) find that geographical distance accounts for more than half of the variation found in their data, making it the most influential predictor. While dialectometry often works with simple Euclidian space ('distance as the crow flies'), the framework also allows for more socially meaningful measures of space, as in e.g. Gooskens' (2005) study of variation in Norwegian using traveling time rather than geographical distance as the main predictor.

In this chapter, I simply model geography in terms of longitude and latitude. As such, I implicitly make the assumption that there are no obvious differences between natural geography and human geography (for more on this notion, see Britain 2010). This is not a good assumption, but it is a highly practical one; quantifying human geography is in itself a difficult task, and more so quantifying human geography as it looked a century ago. This is also part of the motivation for limiting the area of study in this chapter to the Jutland peninsula. This area is relatively densely populated, flat, and has few major obstructions to human movement. Denmark is an island nation, so using Euclidian space as a predictor of variation in the whole country would be much more problematic.

6.5 Variation in voice onset time

In this section, I present a quantitative study of variation in VOT in the traditional Jutlandic varieties. For general background on VOT and information on previous studies of VOT in Danish, see Section 2.3.1. In Section 6.5.1, I present the strategy for selecting the recordings and tokens used in the study, and cover the acoustic and statistical methodology. Finally, in Section 6.5.2, I provide some descriptive statistics, and present the results of a generalized additive mixed model of the data.

6.5.1 Methods

6.5.1.1 Selection of recordings

All recordings used in the current study come from the DS (1971–1976) corpus presented in section 6.2.2 above. DS contains recordings from 230 parishes in Jutland. 17 of these parishes were excluded from the current study. There were three reasons for exclusion: 1) The audio quality was too poor, 2) the recording was too short to include a sufficient number of stops, and 3) group interviews were excluded unless they contained long stretches of speech from a single informant. This only affects areas with a reasonably high density of recordings. If multiple recordings were available from a single parish, one recording was chosen on the basis of either the dialect authenticity judgments made by the original interviewers[5] or audio quality. The audio quality is similar across recordings, and good enough that VOT could be delimited without any major difficulties.

The geographical coverage is shown in Figure 6.6; the peninsula is generally densely covered, although coverage is a little sparse in parts of the Mid-Eastern and Mid-Western dialect areas. The informants' median year of birth is 1896 (range: 1871–1927), and their mean age at the time of recording was 77.4 years (ranging from 45–101 years); the age distribution in these recordings is very similar to that seen in Figure 6.4 above. For thirteen of the informants, no year of birth is reported; I presume that these fall within the reported range. Most recording sessions consist of multiple tapes; whenever possible, the second tape was used for the analysis, so that the informant would have had time to accommodate to the presence of a recording device. All metadata and coordinates, including information about where to

[5]Most recordings in the DS corpus have an accompanying note by the original interviewers evaluating the dialect 'purity' of the informant. This notion is problematized by Goldshtein and Ahlgren (2021).

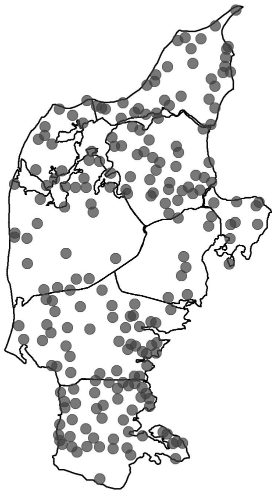


Figure 6.6: *Geographical coverage in the current study.*

Table 6.1: *Distribution of informants across dialect areas by speaker sex.*

Dialect area	**Informants**	**Male**	**Female**
Southern	48	58.3% (*n*=28)	41.7% (*n*=20)
Mid-Southern	40	70% (*n*=28)	30% (*n*=12)
Mid-Eastern	9	66.7% (*n*=6)	33.3% (*n*=3)
Mid-Western	25	92% (*n*=23)	8% (*n*=2)
Djursland	9	66.7% (*n*=6)	33.3% (*n*=3)
North-Eastern	35	85.7% (*n*=30)	14.3% (*n*=5)
North-Western	12	91.7% (*n*=11)	8.3% (*n*=1)
Northern	35	94.1% (*n*=32)	8.6% (*n*=3)
Total	213	77% (*n*=164)	23% (*n*=49)

find to the original recordings, are available in an online appendix (Puggaard-Rode 2022a).

The distribution of informants across dialect areas by speaker sex is seen in Table 6.1. SEX has been shown to affect VOT (e.g. Swartz 1992), particularly among elderly speakers (Torre and Barlow 2009), and by including sex in a statistical model of the data, I will test if the current study can lend credence to those findings. Since speakers were explicitly chosen from a relatively uniform background (non-mobile, rural, previously employed in agriculture) there is little point in attempting to quantify social factors like class.

6.5.1.2 Selection of tokens

As in previous chapters, I distinguish between two laryngeal categories, /b d g/ and /p t k/. This distinction is contrastive in all dialects, although the implementation of the contrast differs.

For each speaker, all stop releases were segmented until the 50th member of /p t k/ had been located. As a result, more tokens of /p t k/ are segmented than of /b d g/. The motivation for this is both practical and theoretical; I have more concrete hypotheses about variation in /p t k/, and only /p t k/ are relevant in the study of affrication patterns in Section 6.7 below. Furthermore, since these recordings are not transcribed, segmentation is very time-demanding. For this reason, /b d g/ are prioritized less.

Only stops in simple onset position were segmented, although palatalized tokens and tokens in stop + /j/ clusters were both treated as simple onsets. Palatalized tokens were included because several dialects show allophonic palatalization of /k/ and /g/ (Bennike and Kristensen 1898–1912: 84ff.); tokens in stop + /j/ clusters were included, because the implementation of these appeared phonetically identical to phonologically palatalized tokens. There were different criteria for the inclusion of the two laryngeal categories. /b d g/ were very often weakened to fricatives or fully voiced when appearing in function words (prepositions, pronouns, and high-frequency adverbs). This would often make segmentation difficult or impossible. For this reason, function words with /b d g/ were excluded unless they were either stressed or initial in an intonational phrase, since gestures are enhanced in these prosodic environments (Steriade 1994). All instances of the pronoun [te] *det* 'it, that' were excluded due to its extremely high frequency (Puggaard 2019a). Function words with initial /p t k/ were not excluded, since they display less weakening than /b d g/ and there are fewer high-frequency function words beginning with /p t k/ than /b d g/. This likely means that this study somewhat underestimates the actual difference between the laryngeal series. This discrepancy is

Table 6.2: *Distribution of stops used in the study by phoneme.*

Consonant	Number
/b/	2,212
/d/	2,369
/g/	2,273
/p/	1,386
/t/	5,169
/k/	4,095
Total	17,504

important to keep in mind, but it is not too concerning, since I am generally more concerned with /p t k/.

The distribution of stops used in the study is shown in Table 6.2. While /b d g/ are reasonably evenly distributed across places of articulation, /p t k/ are more skewed, with relatively few instances of /p/. This was also the case in the study of VOT in Standard Danish by Mortensen and Tøndering (2013), and to a lesser extent in the DanPASS-based studies in Chapters 4–5. This is presumably due to /p/-initial words being rare in Proto-Germanic (see Section 2.2.1).

6.5.1.3 Acoustic analysis

Stop releases were segmented manually in Praat (Boersma 2001; Boersma and Weenink 2021).[6] The beginning of a stop release was demarcated at the burst, which was identified from the waveform. Whenever multiple bursts were visible on the waveform, the final one was chosen (following Cho and Ladefoged 1999: 215).[7] The end of a

[6] There are semi-automatic methods of measuring VOT, such as the AutoVOT software (Sonderegger and Keshet 2012); however, this method relies on training data, and due to the highly variable nature of stop implementation in the Jutlandic data, it was not feasible to provide suitable training data.

[7] This phenomenon is not well-understood. It often comes up in studies of individuals with speech disorders, but it is also relatively common in the general population, particularly in velar stops (see Parveen and Goberman 2012 and references therein). The choice of burst has a non-trivial influence on resulting VOT measurements (Gráczi and Kohári 2014), but this influence is presumably similar across varieties, since there are no indications that the presence of multiple bursts is a regional feature

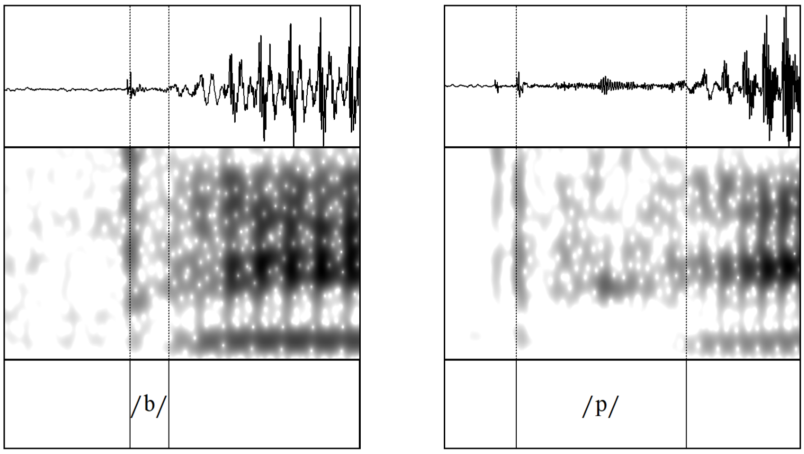


Figure 6.7: *Demarcation of VOT exemplified for unaspirated* /b/ *and aspirated* /p/.

stop was demarcated at voicing onset, identified from the onset of periodicity in the waveform, in accordance with Francis et al.’s (2003) findings that this landmark is most similar to physiological measurements of the onset of vocal fold vibration. This choice of landmark leads to relatively low VOT values, and inflates the differences between dialects as reported in this chapter and the studies of Standard Danish summarized in Section 2.3.1, where the onset of higher formants is typically used to demarcate the end of a stop. It is, however, the same method as were used to delimit stop releases in Modern Standard Danish in Chapter 5, which makes the VOT measurements reported in that chapter comparable to those reported here. The demarcation is exemplified in Figure 6.7. The VOT values for all tokens were extracted using a Praat script, which is freely available in an online appendix along with all TextGrids and measurements (Puggaard-Rode 2022a).

This study only looks at voicing lag, i.e. positive VOT. Measuring pre-voicing in running speech comes with a number of difficulties, particularly in rapid speech. In intervocalic stops, the first portion of the closure is essentially always voiced, due to passive voicing from the preceding vowel; in rapid speech, passive voicing may continue throughout most or all of the closure (Davidson 2016; see Chapter 4).

There is no logical place to start measuring pre-voicing when voicing is continuous; its duration would essentially be a measure of closure duration. When a stop is pre-voiced, there is often still a brief pause in voicing around the release resulting in a short voicing lag. Other recent large-scale studies of VOT have also relied exclusively on voicing lag (e.g. Stuart-Smith et al. 2015; Chodroff and Wilson 2017; Chodroff et al. 2019).

Each token was coded for a number of phonetic environmental predictors, which previous studies indicate influence VOT, and which serve as linear predictors in the statistical model. This allows us to separate the influence of geography from the influence of environmental predictors, and potentially allows us to lend further credence to previous findings. The predictors are:

HEIGHT of the following vowel, which has been shown to influence VOT by e.g. Fischer-Jørgensen (1980), Higgins et al. (1998), and Berry and Moyle (2011); cf. Mortensen and Tøndering (2013), who only found an influence on /b d g/ in Modern Standard Danish. In spite of the great complexity in Modern Standard Danish vowel height distinctions (see Section 1.4), I limited this variable to three levels based on auditory impression. This decision follows Mortensen and Tøndering's (2013) study, which found roughly the same results using a simplified three-way classification of vowel height and Grønnum's (2005: 105) physiological four-way classification of vowel aperture.[8] Danish dialects show a great degree of variability in vowel implementation (Ejstrup and Hansen 2003; Ejstrup 2010), so coding more levels of vowel height would be either too impressionistic or much too time-demanding. Previous studies show that higher vowels increase VOT.

ROUNDNESS of the following vowel, which has been shown to influence VOT in interaction with place of articulation. Fischer-Jørgensen (1972c) found that bilabials have longer VOT before rounded vowels in Standard Danish, while other stops have longer VOT before unrounded vowels.

[8]Note that there are major differences between Grønnum's (2005) vowel height levels, which are based on acoustics, and her vowel aperture levels, which are based on articulation.

BACKNESS of the following vowel, which has also been shown to influence VOT in interaction with place of articulation. Gósy (2001) found that bilabials have longer VOT before back vowels in Hungarian, while other stops have longer VOT before front vowels. Vowels are coded as back or non-back.

PALATALIZATION, which is hypothesized to increase VOT, since palatalized stops are more complex. It is coded as a binary distinction on the basis of auditory impression; no distinction is made between allophonic palatalization and biphonemic sequences of stop + /j/.

STRESS, which has been shown to increase VOT (e.g. Lisker and Abramson 1967; see Section 5.5.1). Stress is coded as a binary distinction.

PLACE OF ARTICULATION and LARYNGEAL CATEGORY. The laryngeal distinction, however it is phonetically implemented, is trivially expected to account for most of the variation in the data. The literature further suggests that the place distinction will influence VOT. A decent rule of thumb is that an occlusion further back in the oral cavity increases VOT, i.e. bilabial < alveolar < velar, although Lisker and Abramson (1964) and Cho and Ladefoged (1999) also find a number of languages not following this pattern. Studies of VOT in Modern Standard Danish have generally found longer VOT for /t/ than /k/ (Fischer-Jørgensen 1954, 1980; Mortensen and Tøndering 2013; see Section 2.3.1). This is also the case with my measurements from DanPASS, as reported in Table 5.1 above. This may be due to the salient affrication of /t/ ensuring a longer release.

Speech rate has also been shown to influence VOT (Andersen 1981a); measuring speech rate of these recordings is far from straightforward, due to the lack of systematic transcriptions of the data, the presence of both informant(s) and interviewer(s), and the general problem with delimiting Danish phonetic syllables (Schachtenhaufen 2010a), particularly in some varieties of Jutlandic (Hansen 1978, 1981). Allen et al. (2003) report that speech rate only partially accounts for idiolectal differences in VOT, which suggests that modeling the individual informant with random effects should largely account for global speech rate effects.

6.5.1.4 Statistical modeling

In order to model the relationship between VOT and geography, the data was fitted to a generalized additive mixed model (GAMM), which can model a potentially non-linear effect of geographical area. Furthermore, descriptive statistics are provided based on the dialect areas. Statistics were calculated in R (R Core Team 2021; RStudio Team 2022) using a number of add-on packages.[9] All R code is freely available (Puggaard-Rode 2022a).

GAMMs were already discussed as one of the smoothing-based approaches to dynamic data analysis in Section 5.4. They are a method of non-linear statistical analysis that is well-suited for data that varies dynamically across time or space (see Wood 2017a for a general introduction, and Sóskuthy 2017 and Wieling 2018 for linguistics-themed introductions). While a linear analysis of e.g. vowel formants across time will have to either measure formants at a chosen landmark or normalize across time steps, a GAMM-based analysis can take into account a full formant trajectory, as demonstrated by Sóskuthy (2017). Similarly, rather than normalizing across dialect areas, a GAMM-based analysis can take into account the full scope of geographical variation in a given area (see also e.g. Wieling et al. 2011, 2014).

The GAMM has VOT as its dependent variable. Regional variation is included in the model through thin plate regression spline smooths (Wood 2003) for the interaction between longitude and latitude; one smooth variable models the main effect of geography, and individual smooth variables model how the two laryngeal categories vary as a function of geography. The contextual phonetic variables discussed above are all included in the model as fixed effects: HEIGHT, ROUNDNESS, BACKNESS, PALATALIZATION, STRESS, SEX, PLACE OF ARTICULATION, and LARYNGEAL CATEGORY.

The fixed effects are contrast coded (Schad et al. 2020; see Sections 4.5.3 and 5.5.2). Binary variables are coded such that the direction with highest expected VOT is positive, and lowest expected VOT is

[9]The following packages were used: `dplyr` (Wickham et al. 2022) for data management; `mgcv` (Wood 2017a, 2021) for fitting GAMMs; `itsadug` (van Rij et al. 2020b) for likelihood ratio tests; `mgcViz` (Fasiolo et al. 2020, 2021) for three-dimensional map-based visualization of GAMMs, and `ggplot2` (Wickham 2016; Wickham et al. 2021) for other visualizations.

Table 6.3: *Overview of contrast coding for categorical variables.*

Variable	Contrast
ROUNDNESS	-½ round, +½ non-round
BACKNESS	-½ back, +½ non-back
PALATALIZATION	-½ non-palatal, +½ palatal
STRESS	-½ unstressed, +½ stressed
SEX	-½ female, +½ male
LARYNGEAL CATEGORY	-½ /b d g/, +½ /p t k/
HEIGHT	-⅓ low, -⅓ mid, +⅔ high
	-½ low, +½ mid
PLACE OF ARTICULATION	-⅔ bilabial, +⅓ alveolar, +⅓ velar
	-½ alveolar, +½ velar

negative. For ternary variables, two theoretically motivated contrasts were coded. For HEIGHT, I test the distinction between 1) high and non-high vowels, and 2) mid and low vowels. For PLACE OF ARTICULATION, I test the distinction between 1) labials and non-labials, and 2) alveolars and velars. The contrast codes are summarized in Table 6.3.

As mentioned in Section 6.5.1.3, we have some specific hypotheses about the interactions between PLACE OF ARTICULATION : BACKNESS and PLACE OF ARTICULATION : ROUNDNESS, respectively. These only pertain to /b p/, so they are included in the GAMM for the labial–non-labial only. The model is further fitted with by-speaker random slopes corresponding to each of the fixed effects, with the exception of SEX.

The model is run with fast restricted maximum likelihood estimation (fREML) with discretized values for covariates to decrease computing load (Wood et al. 2017), assuming the scaled-*t* error distribution to account for heavy-tailed residuals. In order to estimate the influence of the geographical variables, a nested model without the effect of area was fitted, and minimized smoothing parameter scores (fREML scores; see Section 5.5.2) of the two models are compared using a likelihood ratio test (van Rij 2016).

The model described here differs from the one included in a previously published paper (Puggaard 2021) in a number of ways. In the previous paper, categorical variables were not contrast coded. Instead of having separate variables for PLACE OF ARTICULATION

Table 6.4: *Mean VOT in ms for each phoneme by dialect area (see Figure 6.5 for abbreviations; MSD = Modern Standard Danish). First and third quantile in parentheses.*

Area	**/b/**	**/p/**	**/d/**	**/t/**	**/g/**	**/k/**
S	9.3 (5–12)	42.7 (28–55)	14 (8–18)	53.7 (40–66)	16.7 (11–22)	54.5 (41–66)
MS	8.1 (4–10)	38.5 (26–48)	12.4 (6–16)	52.3 (38–64)	13.1 (5–19)	45.9 (34–57)
ME	5.7 (3–7)	41.5 (23–53)	11.4 (5–15)	51.7 (31–70)	12.7 (6–17)	49.9 (37–62)
Dj	9.2 (4–13)	46.2 (30–60)	14.5 (9–19)	53.8 (39–67)	14.4 (7–20)	48.3 (35–60)
MW	5.8 (3–7)	32 (19–44)	10.2 (4–13)	40.1 (28–50)	10.8 (3–15)	39.8 (27–51)
NE	6.8 (3–10)	30.6 (17–40)	12.7 (5–17)	42.5 (29–53)	10.2 (4–14)	41.5 (27–52)
NW	8.2 (4–11)	31.8 (20–42)	12.9 (6–16)	36.3 (24–46)	13.3 (6–20)	42.9 (31–53)
N	6.9 (3–10)	30.6 (18–40)	11.5 (5–16)	42.2 (27–53)	11.8 (4–16)	41.8 (29–52)
MSD	–	44.3 (29–56)	–	69.8 (54–81)	–	50.1 (38–59)

and LARYNGEAL CATEGORY, the stops were coded with a categorical PHONEME variable with the levels /b d g p t k/. The model reported here furthermore has a much more elaborate random effects structure, and includes two interaction variables. These changes should make the results more interpretable and reliable. There are no substantial concomitant changes to the interpretation of the results.

6.5.2 Results

6.5.2.1 Descriptive statistics

Table 6.4 gives VOT values for the different phonemes as grouped by dialect area and for Modern Standard Danish /p t k/ (see Section 5.5.1 above). Mortensen and Tøndering (2013) measured the VOT of

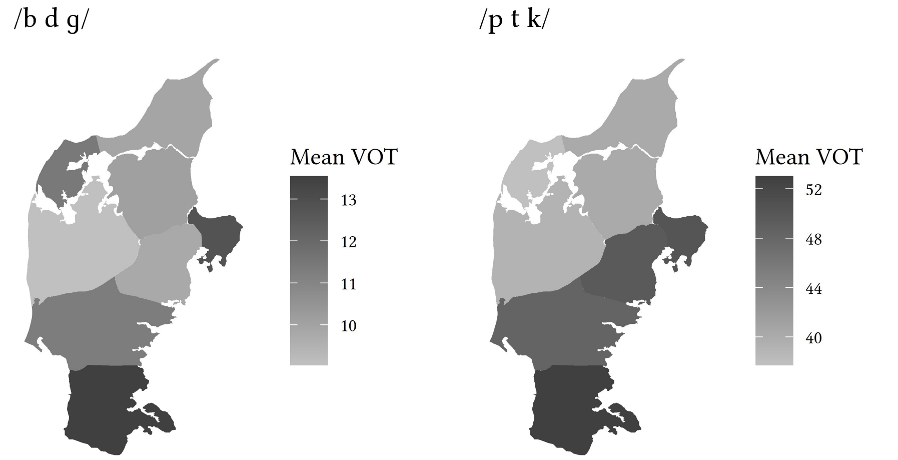


Figure 6.8: *Mean VOT values for the two laryngeal categories by dialect area projected onto maps.*

/b d g/ in the DanPASS corpus, but given the differences in methodology between that study and the present one, I will not be making comparisons with their results here. The results for the two laryngeal categories are projected onto maps in Figure 6.8, and the results for individual phonemes are projected onto maps in Figure 6.9. With only a few counterexamples, VOT is shorter in the traditional regional varieties than in Modern Standard Danish. Modern Standard Danish /t/ is much longer on average than /t/ in all Jutlandic varieties.

For /p t k/, dialect area clearly influences VOT. The pattern is roughly similar for all stops. It is most pronounced for /t/ and least for /k/. The dialect areas seem to form clusters: essentially, south(-eastern) varieties have longer VOT, and north(-western) varieties have shorter VOT. Interestingly, in most cases, the minimum gap in VOT between a member of either dialect cluster seems to be approximately 10 ms, which was found by Blumstein et al. (2005) to be the lower limit of what the human neural system can perceive. This suggests that the limits on regional variation in VOT is only constrained by limits on perception. The only thing approaching a pattern in /b d g/ is the relatively high values across the board in the Southern and Djursland varieties.

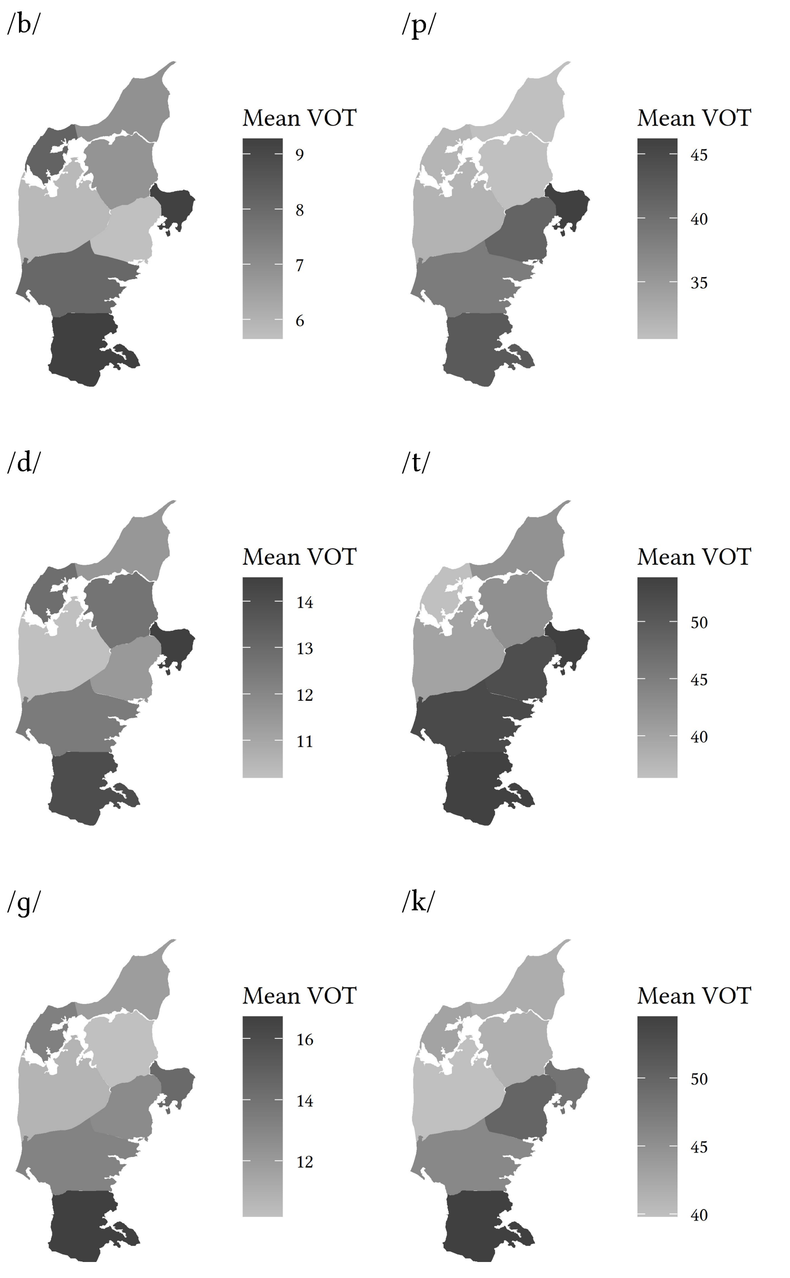


Figure 6.9: *Mean VOT values for individual phonemes by dialect area projected onto maps.*

Table 6.5: *Parametric coefficients of generalized additive mixed model.*

	estimate	SE	*t*	*p*	
intercept	31.03	0.6	51.36	<.001	***
LARYNGEAL CATEGORY	31.18	0.63	49.54	<.001	***
HEIGHT: high *vs.* non-high	2.7	0.28	9.74	<.001	***
HEIGHT: low *vs.* mid	-0.07	0.24	-0.31	0.76	
ROUNDNESS	2.86	0.33	8.7	<.001	***
BACKNESS	2.49	0.38	6.64	<.001	***
PALATALIZATION	9.39	0.94	10	<.001	***
STRESS	2.6	0.22	11.96	<.001	***
PLACE: labial *vs.* non-labial	5.56	0.29	18.9	<.001	***
PLACE: alveolar *vs.* velar	0.82	0.38	2.15	0.03	*
SEX	0.03	0.45	0.07	0.94	
ROUNDNESS : labial contrast	1.6	0.66	2.42	0.02	*
BACKNESS : labial contrast	3.04	0.74	4.11	<.001	***

6.5.2.2 Generalized additive mixed model

This section presents the results of the GAMM. A likelihood ratio test found that a model including the geographical variable performs significantly better than a nested model without the geographical variable, with χ^2 (9) = 72.8, p < .001. With a high effect size of R^2 = .66, the model explains the data quite well. The parametric coefficients and estimated significance of smooth terms are given in Tables 6.5 and 6.6. Table 6.5 summarizes the influence and significance of the linear predictors, and Table 6.6 summarizes the non-linear influence of geography. I unpack this information below, starting with the linear predictors. The intercept (β_0 = 31.01) indicates that the mean VOT in the data set is 31 ms when all other variables are kept at zero. As this averages over both laryngeal categories, the intercept is not particularly meaningful.

LARYNGEAL CATEGORY. There is a (trivially) strong effect of laryngeal category on VOT, such that /p t k/ are longer than /b d g/. This is by far the strongest categorical predictor.

HEIGHT. The data support the hypothesis that VOT is increased when a high vowel follows, but does not support a more complex effect of vowel height.

Table 6.6: *Approximate significance of smooth terms modeling geographical variation. (edf = estimated degrees of freedom, ref.df = referential degrees of freedom).*

	edf	**ref.df**	***F***	***p***	
lon,lat	10.98	12.96	3	<0.001	***
lon,lat : /b d g/	2	2	1.51	0.22	
lon,lat : /p t k/	2	2	11.6	<.001	***

ROUNDNESS. The hypothesis was that following vowel roundness influences VOT in interaction with place of articulation, such that VOT is longer before unrounded vowels, unless the stop is bilabial. This is supported by the analysis, which shows that VOT is generally longer before unrounded vowels, and more so if the stop is not bilabial.

BACKNESS. As with ROUNDNESS, the hypothesis was that following vowel backness influences VOT in interaction with place of articulation, such that VOT is longer before non-back vowels, unless the stop is bilabial. This is supported by the analysis, which shows that VOT is generally longer before non-back vowels, particularly if the stop is non-labial.

PALATALIZATION. The data strongly support the hypothesis that palatalization increases VOT.

STRESS. The data support the hypothesis that stress increases VOT. This effect is stable, although the magnitude is surprisingly small.

PLACE OF ARTICULATION. The data support the hypothesis that VOT is modulated by place of articulation. There is strong support for the generalization that labials are shortest, and more moderate support for alveolar < velar. This is rather different from Modern Standard Danish, where the results reported in Section 5.5.1 and earlier studies of VOT (Section 2.3.1) all find much higher VOT in /t/ than /k/. If this pattern in Modern Standard Danish is due to the salient affrication in /t/, this suggests that /t/-affrication will be much less common in the Jutlandic varieties. The descriptive statistics suggest that some areas show a stable pattern of alveolar < velar.

SEX. The model shows no signs of a stable correspondence between sex and VOT.

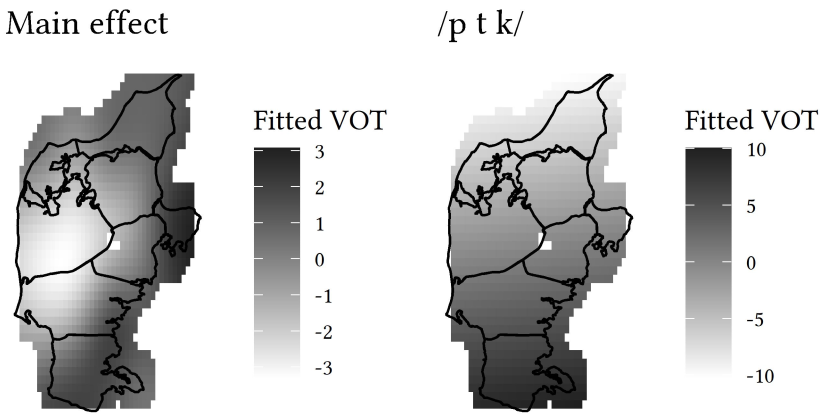


Figure 6.10: *Fitted VOT values attributable to area. Darker shading indicates higher fitted values. Black lines indicate traditional dialect boundaries.*

GEOGRAPHY. The main purpose of this study is to examine the non-linear influence of geography on VOT. There is a strong main effect, suggesting that a primary geographical pattern is shared by all stops. Additionally, /p t k/ show further patterns of geographical variation. Table 6.6 only tells us that significant differences exist; in order to inspect these differences, we need to visualize the fitted values. In Figure 6.10, the fitted main effect of geography on VOT is projected onto a map (left), as is the fitted effect of VOT in /p t k/ (right). There is relatively high VOT in a small area around Djursland, and in the Southern variety, as well as the eastern part of the Mid-Southern variety. Fairly high VOT is also found in a small area covering parts of the Northern and North-Eastern varieties. A large area in the mid-western part of the peninsula has very short VOT. The main effect of geography is highly non-linear, and with a few exceptions does not follow the major traditional dialect areas.[10] The effect attributed to /p t k/ shows a strong, if much simpler, continuous effect of increasing VOT going north–south.

[10]The completely white squares in the middle of both maps are due to the scarcity of recordings in this area, and have nothing to do with the results.

The plot legends in Figure 6.10 refer to the variance in VOT that the model attributes to coordinates relative to the intercept β_0. The magnitude of this variance is of course bigger in /p t k/, as VOT is overall higher in these sounds; the *p*-value for this effect is lower, which is likely because the model's confidence in the geographical fit is lower. However, both results are quite strong. The results are discussed further in Section 6.8 below.

6.5.3 Interim conclusions

This study of VOT in Jutlandic varieties of Danish follows up on Puggaard (2018a), which showed that a variant of /t/ with short VOT is not limited to any specific part of the peninsula, but is found to different extents throughout. I have provided stronger support for that finding, and shown that variation is not limited to /t/, but reflects more general patterns in stop phonetics. Shorter variants of stops than in Modern Standard Danish are nearly consistent throughout Jutland. The longest VOT is found in southern Jutland, parts of mid-eastern Jutland, and Djursland. Very short VOT is found in the centre of the peninsula, and (specifically for /p t k/) in the far north. There is no good explanation for why VOT values are so short in the center of the peninsula; we must assume that it was simply a feature of the traditional local variety. Variation follows consistent but complicated geographical patterns, and the GAMM results suggest that there are multiple continua of variation. The study further finds support for most claims in the literature about how environmental phonetic variables influence VOT.

6.6 Some observations on closure voicing

In Chapter 4, I discussed the extent of intervocalic stop voicing in Modern Standard Danish in the light of previous studies of glottal activity in stop production. Unsurprisingly, there are no corresponding studies of glottal activity in the traditional regional varieties. We only know that glottal spreading is needed to maintain aspiration, and on the basis of the results presented in the previous section, we know that aspirated releases are generally shorter in traditional

Jutlandic varieties than in Modern Standard Danish. This implies that the glottal opening gesture in /p t k/ either peaks earlier or has a smaller magnitude. The results do not allow us to compare /b d g/ in Jutlandic varieties and Modern Standard Danish, since there are no studies of Modern Standard Danish /b d g/ with comparable methodology. It seems a reasonable assumption, however, that the glottal spreading gesture in /b d g/ would also be smaller (or non-existent), in order to achieve a more dispersed contrast. I will not present an empirical study of closure voicing in regional varieties here, but simply discuss some of my observations from working with the corpus, and how they relate to the situation in Modern Standard Danish.

In Modern Standard Danish, intervocalic voicing is common in inflectional suffixes, but otherwise uncommon. It is especially rare in stressed syllables, and after syllables with stød. I argued in Chapter 4 that this relative rarity is caused by a glottal spreading gesture during the closure of /b d g/, as reported by e.g. Hutters (1985). If this glottal gesture is smaller or non-existent in other varieties of Danish, the predicted voicing patterns would be different, as discussed in Section 4.2: intervocalic voicing should overall be more likely, and should be found in more contexts. A smaller or non-existent glottal opening gesture will not result in closure voicing in absolute initial position, as initial closure voicing does not follow automatically from a narrow glottis, but requires separate articulatory effort.

While working with the Jutlandic recordings, I observed that continuous voicing in intervocalic (or intersonorant) position was much more common than in Modern Standard Danish. This was especially the case in the northern part of the peninsula. For most Northern Jutlandic Danish speakers in the corpus, intersonorant voicing is indeed categorical or near-categorical (as has sometimes been incorrectly claimed for Modern Standard Danish). This is not only the case in contexts where voicing is relatively common in Modern Standard Danish, but also e.g. in initial stops in stressed syllables and after syllables with stød.

Consider the phrase in (1), which is spoken by an informant from Bindslev.

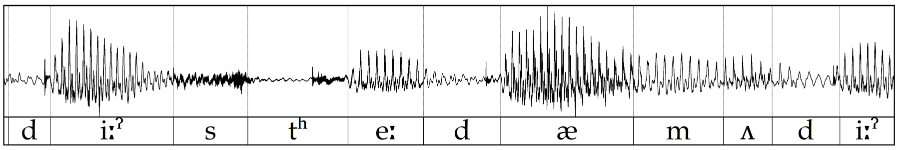


Figure 6.11: *Waveform showing three instances of fully voiced intersonorant* [d] *initially in stressed syllables; spoken by informant from Bindslev; see* (1).

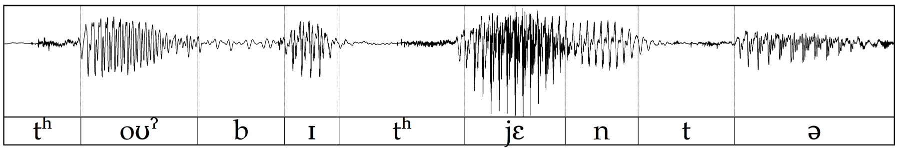


Figure 6.12: *Waveform showing* [b] *with creaky voice after a vowel with stød; spoken by informant from Asdal; see* (2).

(1) ʌ ˈçɛlˀp ʌ ˈpʰluh noʊn ˈdiːˀs tʰeː ˈdæm ʌ ˈdiːˀs tʰe ə ˈseːˀl
and help to pluck some partly for them and partly for to sell
'and help pluck some, in part for them, and in part to sell'

All tokens of [d] in (1) are fully continuously voiced, in spite of being initial in stressed syllables. This is evident from Figure 6.11, which shows the waveform of the relevant parts of (1).

The phrase in (2), which is spoken by an informant from Asdal, shows creaky voice continuing throughout the closure of /b/, which follows a stressed vowel with stød, as evident from Figure 6.12.

(2) ˈtʰoʊ̯ˀ bɪˈtʰjɛntə
'two officers'

This pattern is common in Northern Jutlandic varieties, despite being exceedingly rare in Modern Standard Danish.

As argued in Section 4.2, closure voicing is articulatorily natural in intervocalic position, and as such, does not require a phonological

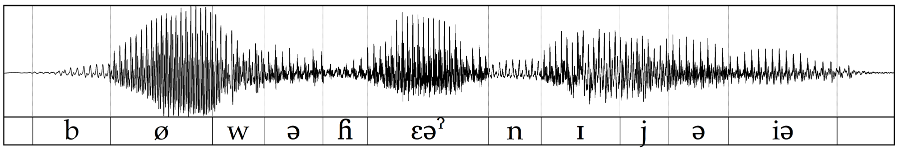


Figure 6.13: *Waveform showing post-pausal, word-initial, and stressed* [b] *with closure voicing; spoken by informant from Elling; see* (3).

account.[11] It is noteworthy, however, that voicing is usually maintained even after stød, which should significantly impact the 'naturalness' of voicing. This clearly suggests that these speakers of Northern Jutlandic varieties lack the glottal spreading gesture found in Modern Standard Danish intervocalic /b d g/: while voicing is not usually found initially, there is no effort to block voicing in other positions.

Closure voicing in initial position, however, requires articulatory effort. I have found no examples in the Northern Jutlandic data of voiced /d g/ in absolute initial position, but some speakers realize /b/ with closure voicing near-consistently, including in absolute initial position. This is the case for (at least) the informants from Elling, Volstrup, Voer, and Bording.

The phrase in (3), which is spoken by an informant from Elling, shows closure voicing in /b/, which is spoken in a word-initial stressed syllable after a pause; the notation (.) indicates a pause. This is evident from Figure 6.13.

(3) ˈlilə (.) ˈhɑ̃õ (.) ˈbøwə ɦɛəˀˈnɪjə iə (.) ˈstʁɑjby
little (.) harbor (.) built here.down in (.) Strandby
'little harbor built down here in Strandby'

When post-pausal voicing is only found in /b/, this is likely because the supraglottal cavity behind the closure is particularly large in /b/ (see Section 4.2). This corresponds to a well-known typological observation: if languages have just one voiced stop, it is bilabial, as demonstrated by a number of Mayan languages, including Tzeltal and Huatesco

[11] The arguments given in that section all extend to intersonorant position.

(Gamkrelidze 1975). This leaves two possible explanations for the voicing pattern in those Northern Jutlandic speakers who display /b/ voicing: 1) Some articulatory effort is put into voicing of /b d g/, and it is sufficient for achieving voicing in /b/ but insufficient in /d g/. 2) No articulatory effort is put into voicing of /d g/. The actual explanation is likely a combination of the two.

The lone pre-voicing in /b/ may either signal that closure voicing was gaining ground in Northern Jutlandic, or it may be the last vestige of a system with more closure voicing. The latter option seems more plausible: as discussed in Section 3.5.3, there is reason to believe that voicing generally used to be more widespread in earlier stages of Danish. The fact that Northern Jutlandic had more widespread closure voicing relatively recently also helps to explain the short VOT in /p t k/ in this variety.

6.7 Variation in spectral characteristics

In Chapter 5, the spectral characteristics of stop releases are discussed in some detail. As pent-up air behind a supraglottal closure is released, frication noise is generated which reflects properties of the primary constriction in the oral cavity. As the constriction opens, low-frequency aspiration noise generated near the glottis gradually becomes dominant, and the turbulent airstream excites the natural frequencies of the oral cavity. In Chapter 5, I explored how the distribution of noise in Modern Standard Danish stops gradually changes throughout the release. As discussed above, there is reason to believe that spectral characteristics during stop releases in some Jutlandic varieties are either different from Modern Standard Danish, or change at a different rate throughout the release. This is particularly to be expected in /t/, which is prominently affricated in Modern Standard Danish.

I used function-on-scalar regression models to explore Modern Standard Danish stop releases in Chapter 5. Ideally, similar models would be used to test how time-varying spectral characteristics vary throughout the Jutland peninsula. This is mathematically possible; the software used to fit these models does not 'care' how complex and multidimensional the underlying data structure is, although computing time may drastically increase with more complex models. However, a statistical model including interrelationships between amplitude, frequency, time, longitude, and latitude would be very tricky to interpret and visualize. For this reason, and in spite of the counterarguments I gave in Section 5.3, this exploration of variation in spectral characteristics makes use of 'magic moments' (Mücke et al. 2014). The full spectrum is analyzed, but in order to model spectral characteristics in an interpretable manner, the complexity of the underlying data must first be reduced.

In the following section, I discuss methods of dimension reduction in statistics, focusing specifically on functional principal component analysis. In Section 6.7.2 below, I present the token selection procedure and statistical methods used in the current study, and in Section 6.7.3, I discuss variation in the spectral characteristics of stop releases in Jutlandic varieties.

6.7.1 Clustering and dimension reduction

Phonological contrasts are regulated with many phonetic cues of varying importance. For various reasons, phoneticians have often had to focus on one cue at a time. Some reasons are theoretical: similar to how some phonologists have attempted to boil phonological grammars down to as few features as possible, some phoneticians have tried to boil phonetic contrasts down to as few invariant acoustic cues as possible (e.g. Stevens and Blumstein 1981). Other reasons are practical; for example, many popular statistical methods in linguistics require a single continuous or categorical dependent variable (see Sections 5.3–5.4).

There are several statistical methods available for reducing a large number of variables to a smaller, more manageable number. Some linguists have used these methods for classification of languages and varieties. For example, Dunn et al. (2005) use hierarchical cluster analysis to arrive at phylogenetic relationships between languages on the basis of large number of data points describing different aspects of language structure; Nerbonne (2009) uses multidimensional scaling to classify dialects on the basis of pronunciation differences in cognate lexical items (see also Wieling 2012; Prokić and Nerbonne 2013).

A common approach to reducing the complexity of multidimensional data is principal component analysis (PCA). In an early example of PCA applied to acoustic data, amplitudes for Dutch vowels were measured in 18 frequency bins, and PCA was used to reduce the original 18 dimensions to 10 principal components (PCs) with no loss in explanatory value (Plomp et al. 1967; Pols et al. 1969; Klein et al. 1970). Instead of being classified on the basis of 18 binned amplitude measurements, each observation can now be classified on the basis of coordinates in a ten-dimensional space. These coordinates are called PC scores, abbreviated s_{n}. The 10 PCs are ranked by how much variance in the original dimensions they account for, and the studies show that the 3–4 PCs with the highest explanatory value can account for the vast majority of information in the original measurements. Classification of the vowels based on these 3–4 PCs is strongly correlated with formant-based classification of the vowels.

PC scores can serve as variables in regression models. For example, if there are several continuous predictor variables in a study, PCA can reduce them to a smaller number, ideally making the results more easy to interpret, and overcoming the problem of fitting models with potentially correlated variables. If a study has a high number of continuous response variables, requiring many separate regression models to be fitted, PCA can be used to reduce the number of necessary models.

Similar to the functional extension of regression we saw in Chapter 5, there is also a functional extension of PCA (FPCA; Castro et al. 1986; Ramsay and Silverman 2005). PCA is used to reduce a large number of continuous variables, and FPCA is used to reduce the complex information found in functions (such as curves) to a small number of continuous variables. This has an obvious appeal in phonetic research, where we often work with variables that change as a function of time, such as pitch (Zellers et al. 2010). FPCA can even be applied to multiple curves simultaneously; Gubian et al. (2015) use FPCA to jointly analyze the patterns of $F1$ and $F2$ in the realization of diphthongs and hiatuses in Spanish, and Gubian et al. (2019) use a similar technique to track an ongoing vowel merger in New Zealand English. In the first study, the authors find that the majority of variation in the formant contour pairs can be expressed with just one PC. A linear regression with s_1 (scores for the first PC) as its dependent variable can predict the diphthong–hiatus contrast in Spanish quite well.

Unlike most previous studies applying FPCA to phonetic data, I do not analyze how some variable changes as a function of time, but rather amplitude as a function of frequency (i.e. spectra). Each PC is associated with an aspect of variance in the input spectra, and the influence of each PC can be visualized by showing how it deviates from the mean spectrum. As I show in Section 6.7.3.1, it is also possible to largely reconstruct any of the original spectra *amp(f)* by weighting the mean spectrum in the data *μ(f)* by the PC scores of *amp(f)*.

6.7.2 Methods

6.7.2.1 Token selection and acoustic analysis

The tokens used in the analysis of spectral characteristics are the same as those used in the analysis of VOT in Section 6.5, although some tokens were excluded here. Similar to the study in Chapter 5, this study is limited to /p t k/.

As mentioned above, this study relies on 'magic moments' in the sense of Mücke et al. (2014). A Praat script was used to extract a 5 ms sound file from the midpoint of each annotated stop release. A bandpass filter was applied to remove frequencies below 500 Hz and above 12.5 kHz. The results in Chapter 5 suggest that the noise distribution at release midpoint will reflect different sources in affricated and non-affricated stops. 5 ms is presumably long enough that the derived spectra are not overly affected by random noise, and short enough that the derived spectra do not span across multiple phases of the release. Stops with VOT < 10 ms were excluded from the study, since 5 ms taken from the midpoint of such short releases likely do span across multiple phases. The distribution of stops by phoneme, as well as the number of excluded stops by phoneme, are given in Table 6.7.

Multitaper spectra were generated in R for each sound file, using the same procedure as described in Section 5.5.1, fn. 13. Within each spectrum, the intensity values (in W/m^2) were log-transformed and standardized (following Gelman and Hill 2006) to keep all measurements on the same scale. Since the values are standardized, simple

Table 6.7: *Distribution of stops used in the study by phoneme.*

Consonant	Number	Excluded
/p/	1,295	91
/t/	5,103	66
/k/	4,042	53
Total	10,440	210

log transformation is equivalent to converting W/m^2 values to dB.[12] Measurements at frequencies > 8 kHz were removed from the analysis, as they did not contribute much to the models other than noise.

The variables used in the analysis of VOT are all included in this study as well; the potential influence of these variables on spectral characteristics of stop releases are discussed in Chapter 5.

6.7.2.2 Statistical modeling

Since a model including both geographical coordinates and dynamic spectral characteristics would be prohibitively difficult to interpret, I first applied FPCA to the spectra within each stop phoneme. Scores for the most influential PCs subsequently serve as dependent variables in (potentially) spatial GAMMs. Statistics were calculated in the R environment (R Core Team 2021; RStudio Team 2022) using a number of add-on packages.[13]

Unlike the study of affrication in Chapter 5, the sound recordings used to compute multitaper spectra for this study were all of equal duration, which means that all spectra have the same number of frequency bins. In other words, the data are regularly sampled and densely distributed. Dense data are often presumed to be relatively noiseless (Gajardo et al. 2021b), but stop release spectra are inher-

[12]Recall that I did not log-transform the intensity values for the study in Chapter 5. My impression from exploratory analysis is that using a logarithmic scale is too fine-grained for fitting FOSR models, and using the W/m^2 scale is too coarse-grained for FPCA of the spectrum, although differences across scales are in both cases relatively minor. It remains a problem for future research to establish best practices.

[13]In addition to the packages used in the study of VOT (see fn. 9), `fdapace` (Gajardo et al. 2021a) was used to implement FPCA.

ently noisy, in spite of multitaper spectra being much less noise-prone than spectra computed from Fourier transformation. For this reason, the spectra are smoothed using a local linear smoother (see Zhang and Wang 2016), with smoothing parameters set automatically using generalized cross-validation (GCV). This procedure is used to select a combination of parameters which yields a good compromise between overfitting and underfitting.[14] For each stop, I computed as many PCs as necessary to account for 95% of the variance in the data. However, for each stop, I only analyzed the three most influential PCs, as the variance accounted for by lower PCs was negligible and difficult to interpret.

For each stop, s_{1-3} served as dependent variables in separate GAMMs. The model selection procedure used here was rather different from previously discussed models. I have had specific theoretical reasons for including all independent variables in previously discussed models, so variables were included regardless of whether their inclusion improved the model fit. In Chapter 5, I motivated how a number of variables may influence the shape of release spectra, and these were all included in the FOSR models. Those reasons also apply to the spectra studied here, but I cannot evaluate how (or if) they will affect specific PC scores computed from spectra. For this reason, I used a step-up model selection procedure, starting with a minimal model including only the geographical variable and by-speaker random intercept. For each stop, I added variables and associated by-speaker random slopes (except in the case of SEX) in the order of their effect size in Chapter 5, i.e. their associated reduction in fREML score. For example, for the /k/ PCs, the order was first BACKNESS, HEIGHT, ROUNDNESS, SEX, STRESS (and finally PALATALIZATION and VOICE ONSET TIME, which were not included in the models in Chapter 5). Variables were kept only if the cost of fitting a more complex model was sufficiently outweighed by an improved model fit; in practice, I ran likelihood ratio tests to test if variables caused a significantly reduced fREML score. Finally, I compared the final model with a nested model

[14]This approach differs from Gubian et al. (2015), who recommend using B-spline smoothing with parameters guided by a combination of GCV and domain-specific knowledge. They implement FPCA using a different R package, `fda` (Ramsay et al. 2009), which requires input functions to be pre-smoothed by the user.

which did not include the geographical variable. When I report *p*-values for the geographical variable below, they are computed from these likelihood ratio tests (see van Rij 2016).

As with the model presented in Section 6.5.1.4, all models are run with fREML with discretized values for covariates. Unlike the spatial GAMM discussed in Section 6.5.1.4, these GAMMs are run with the presumption of a Gaussian error distribution, and the resulting residuals are reasonably normally distributed (for more details, see Puggaard-Rode 2022a). The categorical variables inherit the same contrast coding from the VOT study; see Table 6.3. VOT is standardized within each phoneme, following the procedure of Gelman and Hill (2006).

6.7.3 Results

In this section, I will present the results of the analyses described above in turn, starting with /t/. The results for /t/ spectra will be described in more detail than /k/ and /p/, as I will use /t/ to exemplify some applications of FPCA.

6.7.3.1 /t/

Five PCs are needed to account for 95% of the variance in /t/ spectra. PC1 accounts for 58.4% of the variance, PC2 accounts for 18.2%, and PC3 accounts for 9.3%. The influence of the first three PCs is shown in Figure 6.14. The top left plot shows the variance captured by the respective PCs. The other plots all show the mean spectrum as well as the shape of the mean spectrum when weighted by PC1–3, respectively. This gives an indication of what spectra with a relatively high score s_n and a relatively low s_n look like. The 'low' scores correspond to the mean spectrum weighted by the first quantile of all scores, and the 'high' scores correspond to the mean weighted by the third quantile of all scores.

The mean spectrum of /t/ shows little energy at lower frequencies, relatively high amplitude at frequencies between 1.5–4.5 kHz, a peak just below 2.5 kHz, and gradual loss of energy at higher frequencies. The main source of variance in the data (PC1) is the location and magnitude of the primary peak. High s_1 corresponds to a lower, more

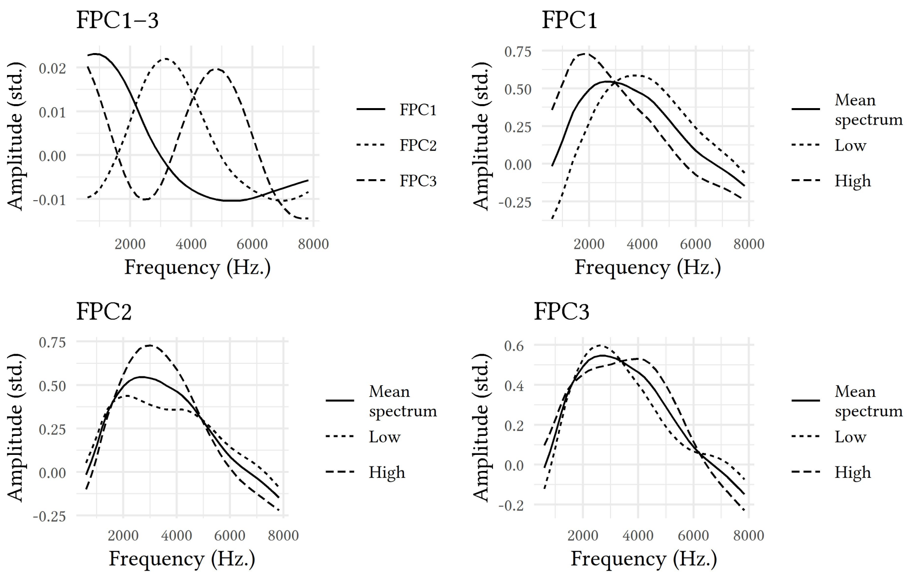


Figure 6.14: *Plots showing the variance in /t/ spectra captured by PC1–3 (top left) and the effect of weighting the mean by high and low scores of PC1–3.*

prominent peak and more energy at lower frequencies, while low s_1 corresponds to a higher peak and less energy at lower frequencies. Another important source of variance in the data (PC2) is in the magnitude of the peak. High s_2 corresponds to a more prominent peak, and correspondingly somewhat less energy at higher frequencies, while low s_2 corresponds to a less prominent peak, and somewhat more energy at higher frequencies. PC3 clearly accounts for less variance, and it mostly corresponds to the peakedness of the energy distribution. High s_3 corresponds to a more restricted and somewhat more prominent peak in the same location as in the mean spectrum, and low s_3 corresponds to a more broad distribution of energy.

It may be conceptually helpful to compare these results to spectral moments (see Section 5.3). The mean spectrum shows significant negative skew. PC1 interacts with both the mean of the distribution (i.e. COG) and skewness; higher s_1 corresponds to lower COG and more negative skew. PC2 mostly interacts with kurtosis, with high s_2 corre-

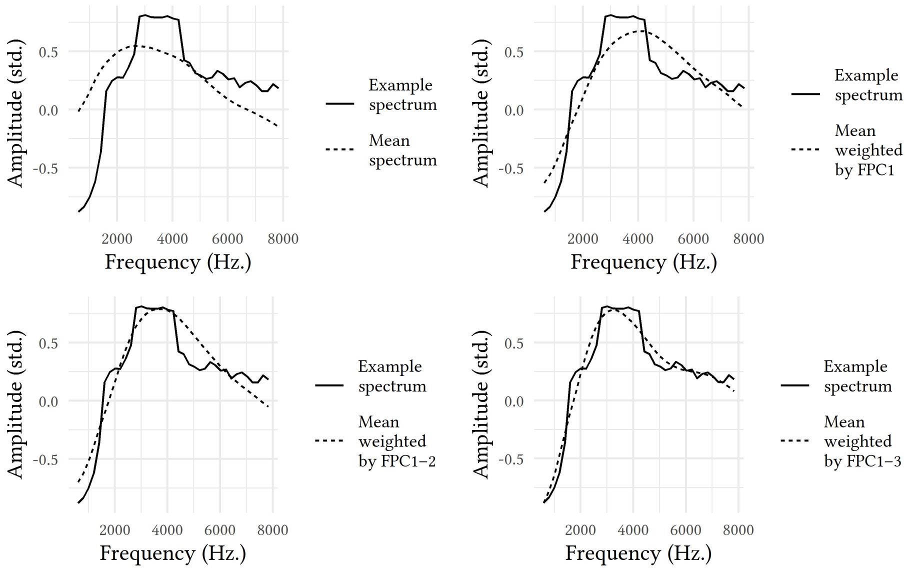


Figure 6.15: *Plots showing a random function* amp(f) *and the mean function* μ(f)*, and the effect of weighing* μ(f) *by* amp(f)*'s scores for FPC1–3.*

sponding to a more peaked distribution. PC3 interacts with all of the four first spectral moments, with high s_3 simultaneously corresponding to a higher COG, higher SD, less skew, and lower kurtosis.

For each input spectrum and each PC, a score is computed. These scores can be used to roughly reconstruct any input curve from the mean, using the equation in (4) (adapted from Gubian et al. 2015: 21).

(4) $$amp(f) \approx \mu(f) + s_1 \cdot PC1(f) + s_2 \cdot PC2(f) + \ldots + s_n \cdot PC\mathrm{n}(f)$$

The equation in (4) states that any input spectrum *amp(f)* can be approximated by weighting the mean spectrum *μ(f)* by *amp(f)*'s scores $s_{1 \ldots n}$ multiplied by the PC functions *PC*1...n*(f)*. Figure 6.15 shows how this is implemented for a random spectrum *amp(f)* in the data, which has relatively low scores for PC1 and PC3, and a relatively high score for PC2. Note that PC1–3 together account for just 85.9% of variance in the data, so weighting the mean by PC1–3 for any curve will only yield a rough approximation of *amp(f)*; note also that *amp(f)* as seen in Figure 6.15 has not yet been smoothed. In the top left figure, we see *amp(f)* and $\mu(f)$. They have quite different shapes. In the top right figure, we see $\mu(f)$ weighted by $s_1 \cdot PC1(f)$ (in this case, s_1 = -27). This serves to move the primary peak in the weighted mean to a higher frequency, increase the energy somewhat at higher frequencies, and drastically decrease the energy at lower frequencies. In the bottom left figure, $\mu(f)$ is further weighted by $s_2 \cdot PC2(f)$ (in this case, s_2 = 7). This has the effect of increasing the energy at the peak. In the bottom right figure, $\mu(f)$ is further weighted by $s_3 \cdot PC3(f)$ (in this case, s_3 = -9.2). This has the effect of narrowing the main peak and skewing it towards somewhat lower frequencies, and aligning both the lowest and highest frequencies more with *amp(f)*. Considering that *amp(f)* in Figure 6.15 is not pre-smoothed, the weighted mean approximates the original function quite well; weighting the mean by the scores of lower-ranked PCs would, of course, yield an even closer approximation.

Figure 6.16 shows the mean PC scores by dialect area projected onto maps. It is immediately noteworthy that Djursland behaves differently from other dialect areas, in having either very low or very high PC scores. s_2 and s_3 do not differ much by dialect area, but s_1 shows quite different patterns in Mid-Eastern and Djursland Danish relative to the rest of the peninsula. The upshot is that mid-point spectra from /t/ releases in Mid-Eastern and Djursland Danish show amplitude peaks at higher frequencies, which suggests that these are areas with /t/-affrication. The upshot of the exceptional s_2–s_3 levels in Djursland is that this variety shows less of a narrow peak around 2.5 kHz in /t/ spectra, and generally a broader distribution of energy in the spectrum.

A GAMM was fitted with s_1 as the dependent variables and all candidate independent variables except STRESS. This model has a medium strong effect size of R^2 = .39. A likelihood ratio test found that a model which included the geographical variable performed significantly better than a nested model without this variable, with χ^2 (3) = 4.1, p = 0.043. The influence of the geographical variable is mapped in Figure 6.17; these patterns are similar to the corresponding map in Figure 6.16. They also show particularly low fitted s_1 in Djursland and the surrounding area. All other independent variables also significantly influence s_1 (see Table 6.8). s_1 is lower – i.e. there is

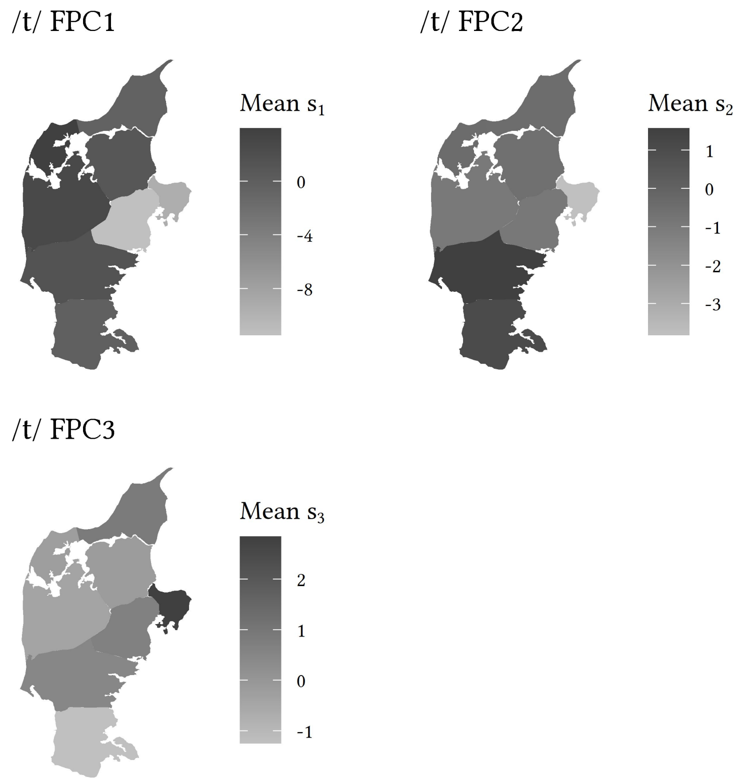


Figure 6.16: *Mean PC scores for /t/ spectra by dialect area projected onto maps.*

less energy at higher frequencies – before non-high, and particularly low, vowels; before rounded vowels; before back vowels; and in non-palatalized /t/ tokens. Men have much higher s_1 than women. Finally, there is an inverse relationship between VOT and s_1, such that lower VOT corresponds to higher s_1. These are all expected patterns, and similar to the results in Chapter 5.

Fewer variables contribute to the GAMM modeling s_2: HEIGHT, ROUNDNESS, STRESS, and BACKNESS. The geographical variable does not significantly improve the fit of this model. The final model has a medium strong effect size, with R^2 = .29. Only two variables significantly influence s_2, namely BACKNESS, with $\hat{\beta}$ = 1.97, SE = 0.7, t = 2.82,

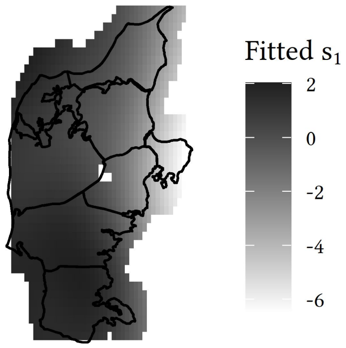


Figure 6.17: *Fitted* s_1 *values for /t/ attributable to area. Darker shading indicates higher fitted values. Black lines indicate traditional dialect boundaries.*

Table 6.8: *Parametric coefficients of generalized additive mixed model with* s_1 *of /t/ spectra as the independent variable.*

	estimate	**SE**	***t***	***p***	
intercept	-1.42	1.38	1.03	0.3	
HEIGHT: high *vs.* non-high	-3.65	0.74	-4.96	<.001	***
HEIGHT: low *vs.* mid	-4.38	0.91	-4.84	<.001	***
ROUNDNESS	3.47	0.97	3.58	<.001	***
BACKNESS	-3.91	1.21	-3.22	<.01	**
PALATALIZATION	-6.79	2.03	-3.35	<.001	***
SEX	6.26	1.87	3.35	<.001	***
VOICE ONSET TIME	-2.03	0.69	-2.95	<.01	**

$p < .01$, and ROUNDNESS, with $\hat{\beta} = 2.21$, $SE = 0.61$, $t = 3.6$, $p < .001$. In other words, high s_2, associated with an especially prominent energy peak around 2.5 kHz, is especially found before non-back and round vowels.

All variables contribute to the GAMM modeling s_3; the geographical variable does not significantly improve the fit of this model. The resulting model has a medium small effect size of $R^2 = .25$. The only variable that significantly influences s_3 is ROUNDNESS, with $\hat{\beta} = 2.62$, $SE = 0.45$, $t = -5.84$, $p < .001$. In the case of both s_2 and s_3, the few significant variables can certainly not fully

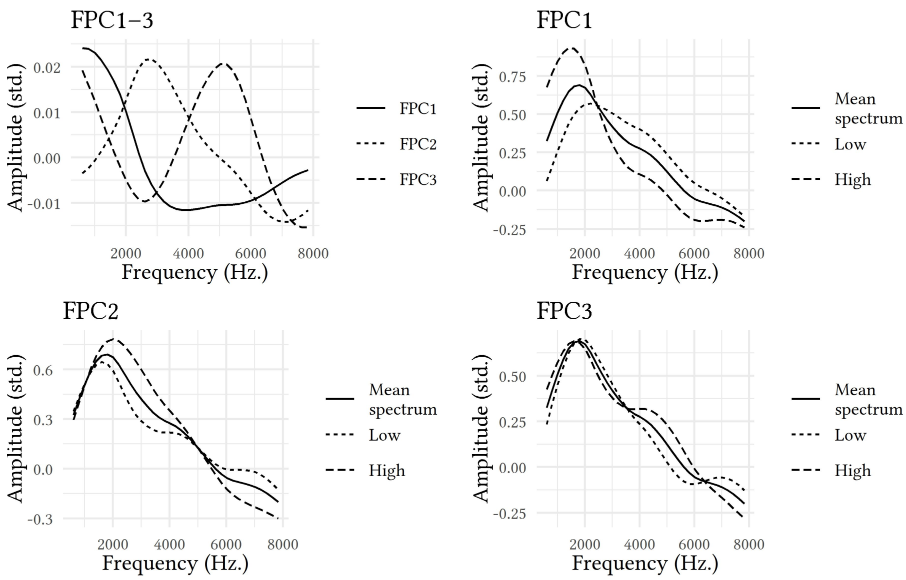


Figure 6.18: *Plots showing the variance in* /k/ *spectra captured by PC1–3 (top left) and the effect of weighting the mean by high and low scores of PC1–3.*

account for all variance in the data; lower-ranked PCs are often more difficult to interpret, and in turn, more difficult to account for.

6.7.3.2 /k/

Five PCs are needed to account for 95% of variance in the /k/ spectra. PC1 accounts for 54.4% of the variance, PC2 accounts for 19.9%, and PC3 accounts for 9.6%. The influence of the first three PCs are shown in Figure 6.18, which is structured like Figure 6.14 above.

PC1–3 look strikingly similar for /t/ and /k/, but it should be kept in mind that the mean /k/ spectrum looks quite different from the mean /t/ spectrum. The main peak of the mean /k/ spectrum falls just below 2 kHz; energy gradually decreases at higher frequencies, and there is very little energy above 6 kHz. PC1 corresponds to more energy at lower frequencies, and less energy at higher frequencies. This means that a token with high s_1 has more energy at very low frequencies relative to the mean, a more prominent peak at slightly

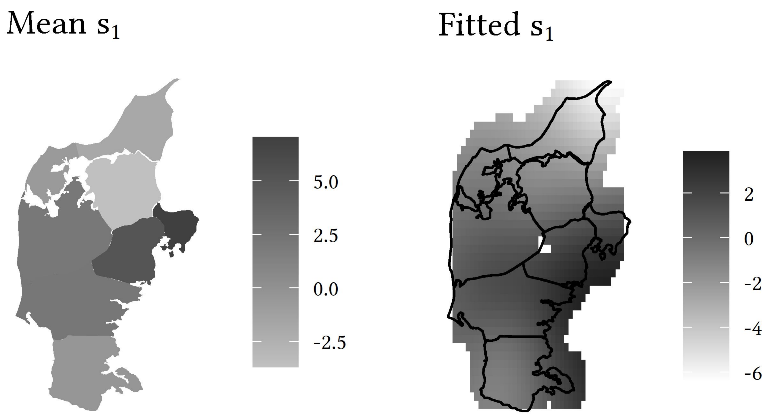


Figure 6.19: *Mean and fitted* s_1 *values for* /k/ *projected onto maps. Darker shading indicates higher* s_1 *values. Black lines indicate traditional dialect boundaries.*

lower frequencies, and much less energy at frequencies above 2.5 kHz. Low s_1 corresponds to a less prominent peak located somewhat higher than the mean, and more equally distributed energy at higher frequencies. PC1 seemingly interacts with all four spectral moments, and low s_1 shows characteristics of velar affrication. PC2 corresponds to more energy at higher frequencies, particularly around 2.5–3 kHz. Tokens with high s_2 have a more prominent peak at slightly higher frequencies than the mean, and generally more energy at frequencies between 1.5–4.5 kHz. PC3 corresponds to more energy both at the lowest frequencies and at frequencies well above the peak. The most prominent difference in a token with high s_3 is a somewhat more even distribution of energy above the peak and below 6 kHz.

A GAMM modeling s_1 was fitted with all candidate independent variables except SEX. This model has a medium strong effect size of $R^2 = .33$. A likelihood ratio test found that the geographical variable significantly improves the fit of such a model, with $\chi^2\,(3) = 4.61$, $p = 0.026$. Figure 6.19 shows two maps: on the left, mean s_1 by dialect area is projected onto a map, and on the right, fitted s_1 values from the GAMM projected onto a map This map shows particularly high s_1 in the area around Djursland and Mid-Eastern Jutland, and particularly

Table 6.9: *Parametric coefficients of generalized additive mixed model with* s_1 *of* /k/ *spectra as the independent variable.*

	estimate	**SE**	***t***	***p***	
intercept	-3.12	0.96	-3.25	<.01	**
HEIGHT: high *vs.* non-high	-4.21	0.86	-4.88	<.001	***
HEIGHT: low *vs.* mid	-5.1	0.97	-5.24	<.001	***
ROUNDNESS	5.23	1.35	3.89	<.001	***
BACKNESS	-5.51	1.26	-4.38	<.001	***
PALATALIZATION	-8.26	1.76	-4.7	<.001	***
STRESS	-1.4	0.64	-2.2	0.028	*
VOICE ONSET TIME	-1.67	0.69	-2.41	0.016	*

low s_1 in the Southern and South-Eastern areas. This is also essentially what the GAMM finds.

All other variables also significantly influence s_1, as seen in the summary in Table 6.9. These results show that phonetic context has a similar influence on s_1 in /k/ spectra and /t/ spectra, and the resulting patterns are similar to what was found for Modern Standard Danish in Chapter 5. s_1 is higher, i.e. energy is more concentrated at lower frequencies, before non-high vowels, particularly low vowels; before rounded vowels; before back vowels; in non-palatalized tokens; and in unstressed syllables. There is an inverse relationship between VOT and s_1, such that lower VOT correlates with higher s_1. These are all contexts where we would not expect affrication.

A separate GAMM was fitted modeling s_2, where only HEIGHT, BACKNESS, PALATALIZATION, and SEX contribute. The geographical variable was excluded from this model, as it did not significantly improve the model fit. The model has a medium small effect size of $R^2 = .244$. Other than the intercept β_0, only two variables significantly influence s_2: BACKNESS, with $\hat{\beta} = 4.66$, $SE = 0.63$, $t = 7.43$, $p < .001$, and PALATALIZATION, with $\hat{\beta} = 4.17$, $SE = 1.05$, $t = 3.96$, $p < .001$. Unsurprisingly, high s_2, corresponding to more energy at higher frequencies, is found before non-back vowels and in palatalized /k/ tokens.

Finally, a GAMM modeling s_3 was fitted. Only HEIGHT, BACKNESS, PALATALIZATION and VOICE ONSET TIME contribute to this model. A likelihood ratio test found that the geographical variable does not

Table 6.10: *Parametric coefficients of generalized additive mixed model with* s_3 *of* /k/ *spectra as the independent variable.*

	estimate	SE	*t*	*p*	
intercept	-0.88	0.44	-2	0.045	*
HEIGHT: high *vs.* non-high	-0.93	0.4	-2.33	0.02	*
HEIGHT: low *vs.* mid	-0.63	0.37	-1.72	0.086	.
BACKNESS	-1.14	0.44	-2.6	<.01	**
PALATALIZATION	-1.89	0.71	-2.65	<.01	**
VOICE ONSET TIME	-0.68	0.31	-2.15	0.031	*

significantly improve the model fit. The final model has a medium small effect size of $R^2 = .237$. Several of the variables show a significant influence on s_3; the results are summarized in Table 6.10. The upshot is that high s_3, which corresponds to a broader distribution of energy at higher frequencies, is found before non-high vowels, before back vowels, in non-palatalized tokens, and in tokens with shorter VOT. This is tricky to interpret. A possible explanation is that PC3, more so than the higher PCs, signifies a bimodal energy distribution in the spectrum; in this case, high s_3 may be indicative of a (weak) formant structure midway through the release, which we may well expect in these particular contexts. (PC3 in /t/ spectra may have a similar explanation, but this would not particularly help explain the results of the statistical model of s_3 in /t/.)

6.7.3.3 /p/

Five PCs are needed to account for 95% of variance in the /p/ spectra. PC1 accounts for 54.3% of the variance, PC2 accounts for 20.6%, and PC3 accounts for 9.5%. The influence of the first three PCs is shown in Figure 6.20, which is structured in the same way as Figures 6.14 and 6.18.

The mean /p/ spectrum has a fair amount of energy at lower frequencies, peaks at around 1.5 kHz, and shows linear reduction in energy at higher frequencies. PC1 is associated with much more energy at lower frequencies, and somewhat less energy at frequencies above 2.5 kHz. A token with high s_1 has a low and prominent peak relative to the mean. PC2 is associated with more energy between 2–4 kHz.

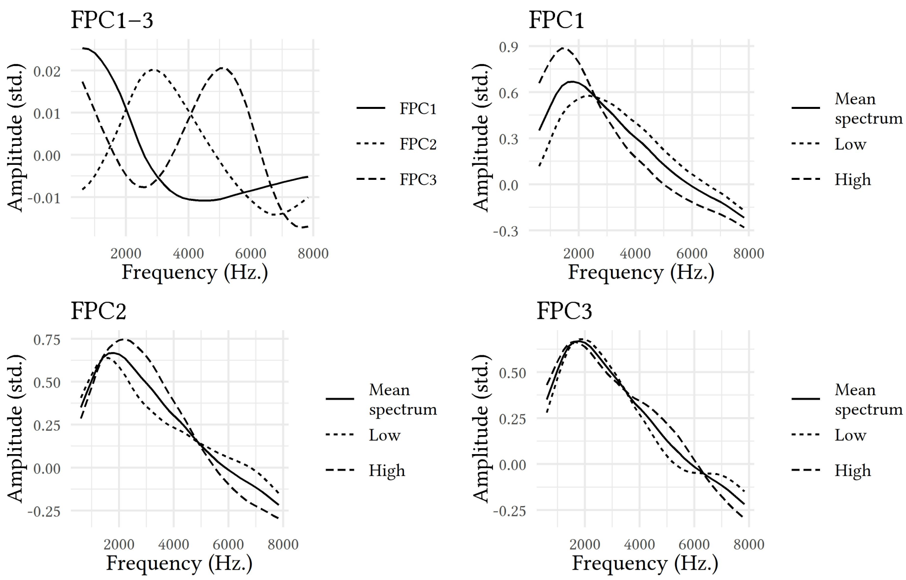


Figure 6.20: *Plots showing the variance in* /p/ *spectra captured by PC1–3 (top left) and the effect of weighting the mean by high and low scores of PC1–3.*

A token with high s_2 has a higher and more prominent peak than the mean, and generally more energy at mid-high frequencies. PC3 is associated with more energy both at the lowest frequencies and around 5 kHz. High s_3 seems to add a second, less prominent peak at higher frequencies around 4–5 kHz; I suggested that a similar pattern in /k/ spectra could be due to weak formant activity during the stop release.

Geography is not a significant predictor of any of the /p/ PCs.

In the GAMM modeling s_1, all variables except BACKNESS and VOICE ONSET TIME contribute. The model has a medium low effect size with $R^2 = .281$. Only two variables significantly influence s_1, namely HEIGHT (high *vs.* non-high) with $\hat{\beta} = -4.24$, $SE = 1.26$, $t = -3.36$, $p < .001$, and SEX with $\hat{\beta} = 4.45$, $SE = 1.76$, $t = 2.53$, $p = 0.012$. The upshot is that s_1 is higher, i.e. energy in the spectrum peaks at lower frequencies and has a narrower distribution, before non-high vowels and in tokens from male speakers.

Table 6.11: *Parametric coefficients of generalized additive mixed model with* s_2 *of* /p/ *spectra as the independent variable.*

	estimate	**SE**	***t***	***p***	
intercept	0.45	0.58	0.77	0.44	
HEIGHT: high *vs.* non-high	3.25	0.9	3.62	<.001	***
HEIGHT: low *vs.* mid	1.96	0.97	2.02	0.043	*
BACKNESS	3.92	0.84	4.67	<.001	***
STRESS	0.68	0.65	1.06	0.292	
SEX	-2.57	1.1	-2.35	0.019	*

In the GAMM modeling s_2, HEIGHT, BACKNESS, STRESS, and SEX were the only variables to contribute. This model has a medium high effect size of R^2 = .329. Several variables significantly influence s_2; the model is summarized in Table 6.11. The results show that s_2 – associated with an energy peak at a higher frequency than the mean, and more evenly distributed energy at higher frequencies – is found before mid and particularly high vowels, before non-back vowels, and in female speakers. This is unsurprising, and in line with the results from Chapter 5.

Finally, the GAMM modeling s_3 is not very informative; only the HEIGHT variable contributes, and the model has a rather low effect size of R^2 = .178. The 'high *vs.* non-high' contrast is significant, with $\hat{\beta}$ = -1.73, *SE* = 0.53, t = -3.27, p = 0.001. This indicates that the somewhat more bimodal spectral energy distribution associated with s_3 is found particularly before non-high vowels; this is sensible, but given the low effect size, this model should be taken with a grain of salt.

6.7.4 Interim conclusions

In this section, I have applied FPCA to multitaper spectra computed from the midpoint of stop releases. This method was used to summarize main sources of variance in the spectra, which were then analyzed with regression models. It was shown that midpoint spectra in /p t k/ vary in largely analogous ways, the two most immediately interpretable ones being 1) the location of the main energy peak, and 2) the prominence of that peak. These sources of variance are to a high extent determined by phonetic context in predictable ways, similar to what we saw for Modern Standard Danish in Chapter 5. For example, the main peak is usually higher before high, non-back, and non-round vowels, in palatalized tokens, in stressed syllables, and in female speakers. It is an important finding in itself that the results are largely in line with expectations, since this is the first study to use FPCA to analyze variance in speech spectra.

In /t k/, the primary source of variance s_1 displayed regional variation. In /t/, s_1 was particularly low in Djursland and the area immediately to the north-west of Djursland, indicating that /t/ in this area is associated with a main energy peak at relatively high frequencies at the release midpoint; in other words, /t/-affrication is relatively common in this area, but otherwise not common in the traditional varieties of Jutland. The pattern is different in /k/; here, s_1 is high in Djursland and surrounding areas, but very low in the Northern and parts of the North-Eastern dialect areas. Very low s_1 in /k/ corresponds to a peak at high frequencies relative to the mean, but also a broader distribution of energy in high frequencies, which is consistent with velar friction. As such, regional patterns of affrication at the release midpoints are different in /t/ and /k/; there are no signs of regional patterns of affrication in /p/. These results, as well as the results for VOT and the observations on closure voicing, are discussed in more detail below.

6.8 Discussion

In this section, I will discuss the results presented above from three different perspectives. In Section 6.8.1, I discuss how the different patterns of VOT, spectral energy distribution, and closure voicing in Jutlandic varieties compare to Modern Standard Danish. In Section 6.8.2, I briefly discuss possible parallels between the phonetic findings in this chapter and regional variation in stop gradation (see Chapter 3). Finally, in Section 6.8.3, I hypothesize about possible sources of the observed variation; I argue that the variation can shed a light on both lesser understood aspects of the early stages of stop gradation, and on sociolinguistic aspects of the spread of the standard language.

6.8.1 Comparison with Modern Standard Danish

In this chapter, I have reported the results of corpus studies on VOT and spectral characteristics in the stops of Jutlandic varieties of Danish. I have also provided some preliminary data on closure voicing in these varieties. These topics have all been discussed for Modern Standard Danish earlier in the dissertation, so it is worth comparing Jutlandic varieties and Modern Standard Danish.

In Section 2.3.1, I mentioned several previous studies of VOT in Modern Standard Danish. Those studies are not directly comparable to the one presented here due to differences in methodology. However, even without direct comparison, it is obvious that the difference in VOT between /b d g/ and /p t k/ is more extensive in Modern Standard Danish than in the Jutlandic varieties. This signals that the laryngeal contrast is cued differently across varieties. The northern varieties as well as Mid-Western show particularly small differences in VOT between /b d g/ and /p t k/; in northern varieties, we further see closure voicing in contexts where it would be rare or simply non-existent in Modern Standard Danish.

In Chapter 4, I discussed how a reported glottal spreading gesture in Modern Standard Danish /b d g/ (Frøkjær-Jensen et al. 1971; Jeel 1975; Hutters 1985) may account for the rarity of intervocalic voicing in these sounds. Some speakers of (particularly northern varieties of) Jutlandic have continuous voicing in stressed syllables and following stød, which strongly implies that these speakers do not have such a phonologized glottal spreading gesture. Likewise, comparison with the DanPASS corpus (see Table 6.4) shows that VOT in /p t k/ is generally shorter in Jutlandic varieties than in Modern Standard Danish. This is also particularly the case for northern varieties. This implies that the glottal spreading gesture in /p t k/ in these varieties is timed differently and/or is of a smaller magnitude than in Modern Standard Danish.

With regards to spectral characteristics of stop releases, the results in Section 6.7 can not be directly compared to those in Chapter 5, but it can be inferred that phonetic and phonological environment affects stop releases in similar ways in Jutlandic varieties and Modern Standard Danish. The mean spectral energy distribution at /t/ release midpoints is, however, quite different from Modern Standard Danish in most of the traditional Jutlandic varieties. In Modern Standard Danish, the /t/ release midpoint shows spectral properties that are consistent with alveolar frication, regardless of phonetic and phonological context. The mean energy peak in the Jutlandic varieties is much lower, suggesting that affrication is *not* an invariant feature of /t/ in these varieties, except perhaps in Djursland and immediately surrounding areas. As in Modern Standard Danish, /k/ spectra in Jutlandic varieties show strong coarticulation effects. In addition, they also show regional variation, with energy peaks at relatively high frequencies (consistent with affrication) being common in the far north-west of the peninsula.

The underlying articulatory mechanisms of these seemingly unrelated findings are likely related. If we take Northern Jutlandic as an example, this variety shows low VOT across the board, widespread intervocalic voicing, and no /t/-affrication. These findings can all be explained if we assume a glottal spreading gesture that peaks early in /p t k/, and no glottal spreading gesture in /b d g/. The early-peaking gesture in /p t k/ ensures short VOT, but it also ensures higher intraoral air pressure at the time of the release. This in turn ensures a prominent burst, and removes the 'need' for /t/-affrication. The lack of a glottal spreading gesture in /b d g/ is functionally motivated by the need to maximally disperse the laryngeal contrast; this in turn ensures short VOT, and a greater propensity for intervocalic voicing. These hypotheses are in principle empirical questions, but they may not be practically testable given the current status of these varieties (see Section 6.3).

6.8.2 Parallels between phonetics and phonology

VOT was found to be low in the center of the peninsula, and VOT in /p t k/ in particular was found to be quite low in the far north; VOT was rather high around Djursland, and VOT in /p t k/ in particular was high in the far south. These results show parallels to regional variation in stop gradation in Jutland (following Bennike and Kristensen 1898–1912; see Sections 2.4.4 and Chapter 3 for more discussion of stop gradation, and Section 2.4.4 for maps showing the various outcomes of some of these processes).

In the center of the peninsula, stop gradation generally resulted in voiced fricatives across the board, although Old Danish /t/ evolved further into a (post-)alveolar approximant [j ~ ɹ] in some areas. In the far north of the peninsula, stop gradation often progressed further, with Old Danish /t/ being lost completely, and Old Danish /k/ developing into [j] after front vowels.[15] In other words, areas with low VOT also showed significant reduction of Old Danish stops, resulting in voiced fricatives or sonorants. In southern parts of the peninsula, however, Old Danish /p k/ developed into voiceless fricatives; and in the far south (e.g. in Viöl, present-day Germany), /p t/ were both retained as stops in most contexts (Bjerrum 1944). In areas with high VOT, reduction was more constrained, and generally resulted in voiceless obstruents.

I suggested above that regional differences in VOT are the result of differences in the timing and magnitude of phonologized glottal spreading gestures. The differences in stop gradation follow directly from this observation: /b d g/ are more likely to alternate with voiceless fricatives in areas with a glottal spreading gesture, and more likely to alternate with voiced fricatives in an area without such a gesture. Further reduction to voiced approximants is unsurprisingly more common in areas where voiced fricatives developed in the first place.

6.8.3 The spread of aspiration and affrication

In Section 3.5.3, I argued that /b d g/ in Danish were likely voiced by the time the first steps of stop gradation took place; this is rather difficult to verify. It is easier to verify that /t/-affrication is a relatively recent development in Standard Danish (Brink and Lund 1975).

If we assume that /b d g/ were voiced in earlier stages of the language, and /t/ was not affricated, then the widespread Jutlandic pattern of short VOT and non-affricated /t/ requires no explanation. However, two findings from this chapter do need to be explained: 1) Why do southern Jutlandic varieties have relatively high VOT across the board, and 2) why does Djursland and surrounding areas to the south have relatively high VOT and /t/-affrication? In other

[15] The situation is somewhat more complex in the area around Djursland, where Old Danish /p/ was retained as a stop, but /t/ developed into an approximant.

words, we do not need to explain why many areas are different from Modern Standard Danish, but rather why a few areas share similarities with Modern Standard Danish. I argue in this section that southern Jutlandic traditionally did *not* have a voicing-based contrast, and that the area around Djursland was relatively early in adapting prominent aspiration and /t/-affrication from Modern Standard Danish.

Step 2 of stop gradation is repeated in (5):

(5) /b d g/ → [β ð ɣ] / weak position

As mentioned in the previous section, a more general version of (5), viz. loss of closure in weak position, is found in almost all varieties of Danish. In most varieties, the result was indeed [β ð ɣ], and these often weakened further into approximants or other sonorants. In southern Jutland, however, /b g/ developed into [f χ] in weak position. This suggests that Step 2 of stop gradation resulted in spirantization of /b d g/ without changing their voicing status. Just as the result of stop gradation in Modern Standard Danish is difficult to explain if /b d g/ were voiceless at the time of Step 2, it is difficult to explain why the outcome in southern Jutlandic varieties would include [f χ] *unless* /b g/ were voiceless at the time. This is supported by the relatively high VOT found in southern Jutlandic varieties in the 20th century. A possible reason why the historical southern Jutlandic varieties did not have voiced stops is areal influence from German.

The situation is different in the area surrounding Djursland, where Step 2 generally resulted in voiced fricatives or glides, as in Modern Standard Danish.[16] This suggests that in this area, unlike in southern Jutland, /b d g/ were voiced at the time of Step 2. In this area, lengthy VOT is a more recent innovation. Since /t/-affrication is also found in this area, both features were likely borrowed from Modern Standard Danish relatively recently. This calls for a sociolinguistic explanation.

[16]In fact, Bennike and Kristensen (1898–1912: maps 49–52) report a lot of variation in this area; Old Danish /p/ is mostly realized as [v], but a small area north-west of Djursland retains a stop which the authors transcribe as [b], but which may well be voiceless; Old Danish [t] has a lot of reflexes in this general area, including [ð ɹ j ɣ]; Old Danish [k] is consistently realized as [ɣ]. See Section 2.5.3 for further details.

The findings are in line with the *cascade model* of interdialectal influence (Labov 2003), which is a more general version of Trudgill's (1974, 1983) *gravity model.* Both models predict that change does not spread purely geographically, but rather spreads between population centers in a manner that is predictable from a combination of population size and geographical distance. From that perspective, high VOT and /t/-affrication in and around Djursland, particularly south of Djursland along the east coast, may be due to Modern Standard Danish influence first reaching Aarhus (the largest city of Jutland) and spreading from there to other major cities. Such an effect would of course be somewhat obscured in this study, as the DS recordings all feature speakers from rural areas, but the cascade model predicts that rural areas sufficiently close to cities will be affected by changes originating in those cities relatively early. There is no indication of a similar effect in the northern part of the peninsula, where the two largest urban areas at the turn of the 20th century (Aalborg and Randers) show no signs of high VOT or /t/-affrication.[17] Locations of major cities in Jutland at the turn of the 20th century are shown in Figure 6.21.

It has been argued that a simpler and more effective way to account for aggregated linguistic variation is to view variation simply as a sublinear function of geographical distance (Séguy 1971; Nerbonne and Heeringa 2007; Nerbonne 2009).[18] Geographical distance in itself, however, can be a poor predictor in a country such as Denmark, where a significant portion of the population live on islands.[19] The cascade model seems to work quite well for our purposes, since Aarhus – a coastal city quite far removed from Copenhagen – *does* appear to be ground zero of the innovations of long VOT and /t/-affrication in Jutland.

[17] Likewise, the city of Esbjerg is in an area with rather low VOT and no signs of /t/-affrication. Esbjerg is a special case, as it was a very young city at the turn of the 20th century, settled only a few decades earlier (Matthiessen 1985).

[18] Others have criticized the cascade model for taking into account only population size and density, and not other social factors such as prestige and age (e.g. Bailey et al. 1993; Boberg 2000; Horvath and Horvath 2001).

[19] A possible solution is calculating traveling time instead of raw geographical distance, following Gooskens (2005).

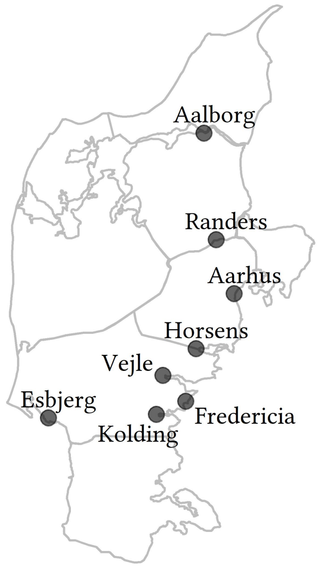


Figure 6.21: *Cities with population sizes above 10,000 in the year 1901. Data from Matthiessen (1985).*

To recap, I have argued that long overall VOT in southern Jutland, especially in /p t k/, is an old feature of these varieties, and that this is reflected in how stop gradation progressed in this area. Conversely, I have argued that long overall VOT and /t/-affrication in the area surrounding Djursland is a recent innovation which spread directly from Modern Standard Danish to the highly urbanized central–southern east coast of Jutland, most clearly targeting the area around the largest city, viz. Aarhus.

6.9 Conclusions

In this chapter, I have provided the first English-language introduction to the DS (1971–1976) legacy corpus of Danish dialect recordings from the 1970s. I further used the recordings from Jutland for two large-scale quantitative studies of phonetic variation. It is the first time that this incredible resource has been used systematically for research.

In the first study, I compare the VOT of /b d g p t k/ from 213 recordings across the peninsula. The data are analyzed using a spatial generalized additive mixed model which also accounts for a number of contextual variables that are known to affect VOT. In the second study, I compare spectral characteristics of stop release midpoints of /p t k/ in the same recordings. Here, I use functional principal component analysis to summarize the main sources of variance in release spectra; principal component scores then serve as dependent variables in (potentially) spatial generalized additive mixed models. Finally, I provided some preliminary data on closure voicing in the northern varieties in particular.

The results of the first study show that VOT is generally rather low in the Jutlandic varieties relative to Modern Standard Danish, and that there are complex patterns of regional variation. The second study shows that the main source of variation in /p t k/ release midpoint spectra is the frequency and prominence of the primary energy peak; in /t k/, this source of variance is partially dependent on geography, although in different ways. In Djursland, for example, /t/ spectra show a prominent peak at relatively high frequencies, while /k/ spectra show a prominent peak at relatively low frequencies; this indicates that /t/ is affricated, whereas /k/ is not.

I have suggested in earlier chapters that Danish stops used to have a voicing-based laryngeal contrast, and we know from previous research that /t/-affrication in Standard Danish is a relatively recent development. On the basis of this, I suggested that only the relatively long VOT in some areas and the /t/-affrication in some areas require an explanation; short VOT and unaffricated /t/ are the default in Jutlandic Danish. On the basis of earlier phonological developments, I suggest that the relatively long VOT in the Southern variety has long been a feature of that dialect. I further suggest that the relatively long VOT and /t/-affrication in Djursland and surrounding areas are borrowed features from the Copenhagen area, which spread directly to the major city of Aarhus and from there to nearby cities and the surrounding rural areas.

CHAPTER 7

Conclusion

7.1 Towards a synthesis

This dissertation has had two main goals: to provide a complete overview of the existing literature on Danish stops, and to fill some specific gaps in our knowledge. In this chapter, I aim to synthesize the resulting research, and pinpoint some areas for future research. In what follows, I assess the impact of the dissertation from five different perspectives: history, phonetics, phonology, variation, and data analysis. The first four of these link back to the topical subdivision made in the literature review in Chapter 2; data analysis is a more general perspective, but it has been equally impactful for the work presented in the dissertation. In Section 7.2, I give a concise chapter-by-chapter summary of the dissertation, and in the sections that follow, I discuss each of these five perspectives in turn.

7.2 Chapter-by-chapter summary

In Chapter 1, I present and motivate the subject area of the dissertation. I also provide a general overview of the Danish language, and introduce the transcription conventions used throughout. The rest of the dissertation is divided into two parts: Part I gives a 'big picture' perspective on Danish stops, and Part II consists of corpus studies on more narrow topics.

Chapter 2 provides a comprehensive overview of the existing literature on Danish stops. As mentioned above, this overview is divided into four perspectives: history, phonetics, phonology, and variation. This chapter mostly sticks to summarizing the literature, but also provides some critical discussion of abstract underlying consonant clusters proposed in the literature, as well as the predictions that follow from some approaches to the representation of laryngeal contrast.

Chapter 3 deals with Danish consonant gradation, the phenomenon which is a.o. responsible for alternations between unaspirated stops [p t k] and semivowels [ʊ̯ ɤ̯ ɪ̯]. The traditional phonological analysis of Danish consonants holds that these are positional allophones of the phonemes /b d g/. The chapter isolates a number of problems with this analysis: it results in rampant neutralization, it is phonetically 'unnatural', and the morphophonological evidence in favor of the analysis is weak, since it relies exclusively on irregular and unproductive morphology. An alternative analysis is proposed, where [ʊ̯ ɪ̯] are always considered allophones of /v j/. The traditional analysis does account for some structural generalizations that are not accounted for in the alternative analysis; it is argued that these are best understood as the result of well-known phonetic pressures operating during earlier stages of the language's history.

Chapter 4 presents a corpus study of intervocalic closure voicing in Danish stops – the first empirical study of this topic. Some previous research assumes that intervocalic voicing is essentially categorical, while other research assumes that intervocalic voicing is essentially non-existent; both of these positions have come up in discussions of the phonological representation of laryngeal contrast in Danish and beyond. The study finds that intervocalic voicing is very negligible in Danish /p t k/, and quite rare in /b d g/, presumably because /b d g/ are produced with an active glottal spreading gesture which serves to block voicing. In statistically modeling the data, I tested how a range of different variables affect the rate of closure voicing, and concluded that it is first and foremost a lenition phenomenon. The results are predicted by a gesture-based approach to phonological representation (such as Articulatory Phonology); other approaches which have made explicit predictions about Danish stops either incorrectly predict, or are unable to predict, aspects of the results.

Chapter 5 presents an exploratory corpus study of the spectral characteristics of Danish aspirated stop releases, focusing on how they change through time, and how they are affected by phonetic context. It is often brought up in the literature that /t/ is highly affricated, but previous descriptions have all been impressionistic or introspective. /p k/ releases are rarely mentioned. The results indeed show that /t/ is affricated across all phonetic contexts, but also that affrication is gradually replaced by aspiration towards the end of releases; the proportion of affrication and aspiration is modulated by phonetic context. /k/ releases are highly affected by phonetic context, presumably since the precise place of articulation in velar consonants is very variable. /p/ releases are also affected by phonetic context, but they are more prone to interspeaker variation than contextually determined variation. The main contribution of Chapter 5 is methodological. The data are fitted to function-on-scalar regression models using the entire spectrum as the response variable. This is a method which arguably overcomes a more general problem faced by researchers in analyzing the complex multidimensional information in noisy spectra, and which allows for intuitive visualization of how the spectrum changes through time under different conditions.

Chapter 6 combines two corpus studies of stop realization in the traditional regional varieties of Danish spoken on the Jutland peninsula. One study investigates variation in voice onset time, and the other investigates variation in the spectral characteristics of aspirated stop releases. The chapter also briefly discusses closure voicing in northern Jutlandic varieties, where voicing is much more widespread than in Modern Standard Danish. Voice onset time is generally shorter in traditional Jutlandic varieties than in Modern Standard Danish, and shows complex regional patterning, which is accounted for with a combination of social and historical factors. /t/ releases in particular also vary in their release characteristics; the vast majority of the peninsula has not adapted the prominent /t/-affrication of Modern Standard Danish discussed in Chapter 5. This chapter models some complex dynamic variables: the regional variable is approached with generalized additive mixed models, and the spectral variable is approached with functional principal component analysis, which is adapted to solve some of the same problems as function-on-scalar regression in Chapter 5.

7.3 History

Inspired by Evolutionary Phonology, the dissertation has suggested phonetic accounts of the series of sound changes that eventually resulted in stop gradation, i.e. the process whereby voiceless unaspirated stops [p t k] in some contexts came to alternate with semivowels [ʊ̯ ɣ̯ ɪ̯] in Modern Standard Danish. Evolutionary Phonology builds on the neogrammarian insight that previous stages of languages were subject to the same phonetic pressures as synchronic languages; in the past century, our understanding of those phonetic pressures has drastically improved. While the account in Section 3.5 concerns language history, much of the explanation comes from modern experimental phonetics.

I argue in Section 3.5 that Danish had voiced stops in a previous stage of the language. This is not a common assumption in Germanic historical linguistics, but it makes it much easier to account for subsequent sound changes. I proposed that the sounds which currently alternate between [p t k]–[ʊ̯ ɣ̯ ɪ̯] were previously voiced stops [b d g] in all positions; the voiced stops gradually weakened in post-vocalic position, and later lost their voicing in pre-vocalic position. In other words, there was a decrease in sonority in pre-vocalic position, and a drastic increase in sonority in post-vocalic position. Both changes can be considered reactions to the well-known articulatory pressure against obstruent voicing.

Closure voicing at an earlier stage was arguably a prerequisite for the post-vocalic lenition patterns. This idea finds support in data from regional varieties of Danish. The varieties traditionally spoken in northern Jutland have short voice onset time across the board and much more widespread closure voicing than Modern Standard Danish; they also display extensive lenition in post-vocalic position.

Meanwhile, varieties traditionally spoken in southern Jutland have long voice onset time across the board and generally avoid closure voicing; here, stops historically weakened to voiceless fricatives in post-vocalic position (if they weakened at all). In other words, consonant gradation affected all varieties, but varieties with more widespread voicing were subject to lenition in a more 'sonorous' direction. Since consonant gradation in Modern Standard Danish resulted in drastically higher sonority, this variety likely also had more widespread closure voicing during the early stages of gradation.

7.4 Phonetics

This dissertation has shown that intervocalic voicing is relatively rare in Modern Standard Danish spontaneous speech, and has provided multidimensional analyses of how the sound spectrum changes over time in different phonetic contexts during aspirated stop releases. The studies of intervocalic voicing and spectral changes both focus on the acoustic signal, but explanations of the data rely on the articulation–acoustics link, and draw extensively on previous studies of e.g. glottal activity during stop production.

Although intervocalic voicing in Danish has never been studied empirically before, it has often been discussed in the literature, and has been used in support of (sometimes opposing) proposals about the representation of laryngeal contrast in Danish and beyond. The study reported in Chapter 4 shows that intervocalic voicing is relatively rare even in /b d g/. I tested how a range of factors affect the likelihood of voicing, and found that voicing patterns as a lenition phenomenon. Earlier articulatory research has shown that /b d g/ are produced with an active glottal spreading gesture, and this gesture has been argued to enforce voicelessness. Building on these previous studies, Chapter 4 argues that voicelessness is the default setting in Danish stops across all contexts; when they are sometimes voiced intervocalically, this is due to reduction of the glottal spreading gesture in lenition-prone environments.

The study of intervocalic voicing shows that aerodynamic variables (such as vocalic environment) have little effect on whether stop releases are voiced throughout. This is somewhat unexpected, and may be due to the binary voicing variable being too coarse-grained. I assume that aerodynamic variables do affect the relative proportion of closure voicing, and that this could in principle be tested in a study with a continuous measure of voicing. The statistical model in such a study would be similar to the one reported in Chapter 4 (although the regression would be linear and not logistic), but the data processing would be significantly more time-demanding.

In the terminology of Kingston and Diehl (1994), who propose a dichotomy between 'automatic' and 'controlled' phonetics, the patterns of intervocalic voicing are surely controlled. Generally speaking, voicing is phonetically natural in intervocalic position. Voicelessness in Modern Standard Danish is managed with a small glottal spreading gesture which is not found in comparable languages, and voicing is found in higher rates in other regional varieties of Danish. These observations strongly suggest that when voicing is blocked intervocalically in Modern Standard Danish /b d g/, this has to reflect cognitively controlled, learned behavior.

The dissertation also includes the first acoustic study of affrication in Danish aspirated stops in Chapter 5; previous discussions of this topic have been impressionistic. The results show that the distribution and prominence of place cues throughout releases vary across different places of articulation. Furthermore, /p t k/ respond quite differently to phonetic context. Little attention has been paid to /p k/ previously. /p/ mostly shows evidence of vowel–consonant coarticulation during the beginning of the release, and the energy distribution in the spectrum is highly speaker-specific. /k/ shows prominent signs of vowel–consonant coarticulation throughout the release. /t/ is, as often assumed, invariably affricated; however, unlike the impression given by most previous descriptions, affrication almost always gives way to aspiration well before the end of the release.

Affrication is likely controlled separately from aspiration to some extent, but the two are not entirely independent. Compared to aspirated stops in other languages, Danish /p t k/ are produced with short closure duration and with a glottal spreading gesture which peaks late. These factors may contribute to the prominent /t/-affrication. Longer closure duration and early glottal spreading serves to ensure high intraoral air pressure, which in turn ensures a prominent burst; if these mechanisms are not available in Danish, an affricated release may indeed be the result. As such, /t/-affrication can partially be explained as a physical reaction to differences in glottal behavior. Like most sound patterns, however, /t/-affrication likely follows from a combination of physical pressures (automatic phonetics) and learned behavior (controlled phonetics); targeted articulatory studies would be necessary to determine the balance between the two.

In Chapter 5, I also provided measurements of voice onset time from aspirated stops in stressed and unstressed syllables. These measurements were not the focus of the chapter and were not discussed much, but they are worth briefly returning to. Previous studies of Danish voice onset time have focused on stressed syllables only. The measurements in Chapter 5 unsurprisingly show that voice onset time is much shorter in unstressed syllables. This further supports the claim that the magnitude of glottal gestures is reduced in unstressed syllables. This mechanism also causes a higher rate of intervocalic voicing in /b d g/ when they are unstressed. The magnitude of glottal spreading in /b d g/ is modest, so the gesture may be elided entirely in unstressed position, paving the way for intervocalic voicing; the magnitude is greater in /p t k/, so reducing this gesture will usually only result in a shorter aspiration phase.

7.5 Phonology

It has been discussed at several points in the dissertation how different approaches to phonology make different predictions about the phonetics–phonology interface. In some frameworks, phonetic patterns are not assumed to influence phonology at all; in others, phonology and phonetics are intimately connected. In what follows, I will proceed from the assumption that they are closely connected, and that phonetic evidence is useful in discussions of phonological representation.

Chapter 3 argued against the traditional analysis of Danish consonants, which assumes that [ʊ̯ ɪ̯] are both possible positional allophones of /g/, and that /b/ in some contexts displays stylistic alternations between [ʊ̯ ~ p]. I proposed an alternative analyses where [ʊ̯] is always an allophone of /v/, and [ɪ̯] is always an allophone of /j/. The traditional analysis requires the /g/ phoneme to be underspecified with regards to both place, manner, and voicing, and also poses synchronically active alternations which are difficult to derive. Perhaps most gravely, it relies exclusively on alternations found in irregular and unproductive morphology; even if we accept that language learners have no problem establishing the unnatural phonological processes, the evidence they encounter in favor of such an analysis is scarce.

The discussion in Chapter 3 is largely theoretical, but it begs several questions which are in principle testable and empirical, particularly with regards to whether speakers categorize [k ʊ̯ ɪ̯] as the same phoneme. There are psycholinguistic experiments which are well-suited for gauging this type of categorization, including concept formation experiments (e.g. Ohala 1983b), acceptability judgment experiments (e.g. Ohala and Ohala 1986), and perceptual similarity experiments (e.g. Flege et al. 1994). I discussed in Section 3.3.1 how computer simulations in the framework of Bidirectional Phonetics and Phonology could be used to test the feasibility of establishing the traditional analysis on the basis of the data presented to language learners, and building on ongoing work in BiPhon on the phonology–orthography interface may help to determine whether speakers' knowledge of orthography affect how they analyze the alternations (Hamann and Colombo 2017).[1]

In the alternative analysis of stop gradation, I remained agnostic as to whether [ɣ̯] can be considered an allophone of /d/, as proposed in the traditional analysis. The evidence in favor of such an analysis is limited, but unlike the proposed /g/ phoneme, it does not result in problematic neutralizations. [ɣ̯] has to be established by learners as either 1) an allophone of /d/, or 2) a separate phoneme /ɣ̯/. Due to the lack of morphophonological evidence in favor of the first analysis, it can only be maintained if it is phonetically grounded;

[1] The proposals for future research given here were included in a funding application with Camilla Søballe Horslund as principal investigator, and should be attributed to her.

whether or not this is the case remains an open question, which may be resolved with future articulatory studies. Several sources have mentioned a secondary dorsal or pharyngeal gesture in the articulation of [ɣ̯], and Siem (2019) showed that this gesture is sometimes more prominent than the coronal gesture. If [ɣ̯] is indeed primarily dorsal with a secondary coronal gesture, this would be evidence in favor of a separate /ɣ̯/ phoneme. The nature of the relevant gestures can be examined with e.g. ultrasound tongue imaging, which is a relatively inexpensive method often used to investigate unusual tongue shapes (e.g. Lawson et al. 2013; Mielke 2015), or with real-time magnetic resonance imaging, which is much more costly but also captures more precise spatial information (Carignan et al. 2020).

In Chapter 4, I discussed how different approaches to the underlying representation of laryngeal contrast make different predictions about Danish intervocalic stop voicing. Representational approaches with more phonetic integration generally fare best. Approaches which assume an abstract [voice] feature make no useful predictions about intervocalic voicing; approaches which assume a [spread glottis] feature can explain why intervocalic voicing is very rare in /p t k/, and perhaps also why it is fairly rare in /b d g/; approaches which use articulatory gestures as phonological primitives can straightforwardly account for the results of the corpus study, including why voicing in /b d g/ patterns as lenition.

7.6 Variation

Variation is most explicitly discussed in Chapter 6, which covers regional variation, but it has been relevant in all chapters. For example, speakers of Modern Standard Danish vary in whether they have active [p ~ ʊ̯] alternations, and if so, which lexical items allow for the alternation; this was relevant for the discussion in Chapter 3. Chapter 4 finds that speaker age is a significant predictor of intervocalic stop voicing, with voicing being less common among older speakers. Chapter 5 finds that speaker sex is often a strong predictor of the spectral composition of stop releases, with female speakers having more affricated releases. Additionally, all statistical models have included random variables filtering out speaker-specific information, due to the baseline expectation that results will show some degree of systematic interspeaker variation.

The patterns of regional variation are too complex to recap here, but it was shown that there is systematic variation in both voice onset time and release characteristics, and seemingly also closure voicing. Some of this comes down to differences in laryngeal activity during stop closures, but there are likely also other systematic articulatory differences at play. Danish stops show regional variation not just in phonetics, but in history and phonology as well: regional varieties have been subject to different historical developments, which sometimes led to differences in phonology.

The phonetic results in Chapter 6 are useful in determining whether the phonetic results in earlier chapters reflect automatic or controlled behavior. Voice onset time in /p t k/ shows complex regional variation, which indicates that the duration of aspiration is cognitively controlled in a highly granular way. The differences in constraints on voicing in some varieties cement that intervocalic voicing patterns are to some extent controlled. The spectral characteristics of stop releases also show variation, particularly in /t/. As discussed in Section 7.4 above, this may be partially due to differences in glottal behavior, but the control of glottal behavior and affrication is presumably partially independent.

An open question concerns the current status of variation in the Jutland peninsula. While some studies have argued that non-standard Danish dialects are essentially extinct, more recent studies show that they are still spoken to varying extents in some of the areas that were explicitly discussed in Chapter 6, also by younger speakers. My intuition is that the unaffricated 'dry t' is still used in northern Jutland, but it would require targeted fieldwork to uncover whether this is the case, whether its use is socially stratified, and whether it is treated as an overt regional feature.

7.7 Data analysis

Many scholars are currently grappling with how to solve deep-seated structural problems in quantitative research. In this section, I briefly outline some of these currents, and reflect on how I have tried to take them into account in my approach to data analysis.

Quantitative research is facing a replication crisis. This was famously pointed out in a large-scale replication study of 100 major psychological findings, where the authors were able to replicate less than half (Open Science Collaboration 2015). The problem is not isolated to psychological research, and has also been discussed for several other scientific disciplines, including linguistics (Roettger and Baer-Henney 2019). It is a gnarly issue, perhaps especially prevalent in confirmatory experimental research, where a significant p-value at $\alpha = 0.05$ has often been a *de facto* threshold for publication. This potentially rather lenient 'filter' is not a guarantee that a study will replicate (Vasishth et al. 2018a), especially since the coveted $p < .05$ can almost always be achieved with some flexibility in data collection and analysis (Simmons et al. 2011; see Stefan and Schönbrodt 2022 for a recent overview of so-called 'p-hacking' strategies). Current academic incentive structures are favorable towards publishing high volumes of novel significant findings (Smaldino and McElreath 2016); in addition to promoting questionable research practices in general, other adverse consequences of this include a bias in the publication record against findings that are not considered 'impactful', such as null results and replication studies (Roettger 2021), and a tendency to reframe exploratory studies as confirmatory during the publication process (Roettger et al. 2019).[2]

Researchers are increasingly taking steps to mitigate these problems, both at the individual and institutional levels. In linguistics, preregistering research (i.e. specifying procedures for data selection and analysis prior to collection) is gaining popularity (Roettger 2021); linguists are increasingly adapting Bayesian approaches to data analysis, which counteracts the dependence on p-values to quantify significance (Nicenboim and Vasishth 2016; Vasishth et al. 2018b);

[2]Strictly speaking, studies are only confirmatory if a single statistical model, which was motivated prior to data collection, is fitted to the data (Baayen et al. 2017).

meta-analytic methods are adopted to quantify the robustness of established findings (e.g. Nicenboim et al. 2018; Cristia et al. 2020), and large-scale international collaborations also seek to test how robust our methods and results are, including the ManyBabies Consortium (e.g. 2020) working on first language acquisition in infancy, and the currently ongoing Many Speech Analyses project. At a more institutional level, journals are increasingly adapting policies mandating open sharing of data and code for quantitative studies, which can drastically improve analytic reproducibility (e.g. Hardwicke et al. 2018).

Not all of these solutions are approachable in early career research; preregistration, large-scale collaborations, and focus on less 'impactful' research are obviously more readily available to established researchers in permanent positions, in large part due to the aforementioned incentive structures. Luckily, Roettger (2019) also proposes some best practices that are in principle available to anyone. These can largely be summarized under the umbrella of *transparency*.

Throughout the dissertation, I have sought to be as transparent as possible with regards to analytical decision-making. In describing the corpus studies in Chapters 4–6, I go into detail about each step of the analytical process, including token selection, acoustic analysis, technical details about statistical models, and interpretation of results. I do not go into much detail about how analyses are implemented in practice, i.e. how they are coded, since these details are liable to become obsolete with software changes. This information, however, is included in online appendices in the *DataverseNL* repository (Puggaard-Rode et al. 2022b; Puggaard-Rode 2022a). These include all materials needed for analytical reproducibility.[3] All code is written in open-source software, viz. Praat (Boersma 2001; Boersma and Weenink 2021) and R (R Core Team 2021; RStudio Team 2022). All R code is further written in Markdown format, meaning raw code and output are interspersed with text, where I aim to illustrate and motivate each analytical decision (providing sources if necessary), and describe the purpose of each

[3]The audio recordings themselves are not included. I do, however, provide enough information for the individual recordings to be findable for anyone with access to the corpora.

command used in the code. In principle, this allows anyone to evaluate the viability of the analyses. Additionally, it may serve as inspiration for others who need to carry out similar analyses.

The quantitative studies in this dissertation are largely exploratory; I have not aimed at answering simple yes/no questions. While I do report *p*-values in most chapters, I have tried not to rely on them as a proxy for practical significance; instead, in most cases, I have discussed coefficients and the certainty attributed to them. Ideally, confirmatory studies should be carried out to test whether the observations made in the dissertation replicate. In the case of Chapters 4–5, this would involve collecting new semi-spontaneous speech data with a more balanced group of participants with regards to age and gender, and setting well-motivated criteria for statistical modelling and sample size prior to the analysis. This would be relatively straightforward for the intervocalic voicing results, but less so for the affrication results; the observations made in Chapter 5 may not be suitable for null hypothesis significance testing. In the case of regional variation, it would be possible to use different portions of the same recordings as in Chapter 6, since only a small subset of the corpus is used in the analysis; gathering new data of the traditional variants is probably not possible.

Logistic mixed-effects regression models (see Chapter 4) are well-known in linguistic research, and generalized additive mixed models (see Chapter 6) have become very widely used in the past half decade, although their use in spatial variation has been relatively limited. These methods require no further discussion. On the other hand, functional principal component analysis (see Section 6.7) has only been sparsely used in linguistics, and to my knowledge has not previously been used in the analysis of spectral variance. Function-on-scalar regression (Chapter 5) is a completely new method in linguistics. The use of functional data analysis in analyzing speech spectra is a major contribution of this dissertation. I believe these methods can be used to overcome many problems researchers currently face with selecting variables to reflect the complex multidimensional information of the spectrum. Measures derived from the spectrum are used for many kinds of research question, so the potential of these methods go well beyond analyzing stop releases. An important avenue for further research is to develop best practices for selecting, fine-tuning, and interpreting these models, and to determine how they are best used for hypothesis testing.

Bibliography

For sources which are available online, I strive to provide information on how to find them; whenever possible, with digital object identifiers (DOI), handles (HDL), or persistent identifiers from JStor or ProQuest. For web sources, I give a full link. For other sources available online, I do not provide links in the printed version of the dissertation, but the bibliography entries themselves are hyperlinked in the digital version. These links were all active when the dissertation was submitted, but are liable to become obsolete.

A few lengthy and commonly occurring names of journals, proceedings, and publishers are abbreviated in the reference list. These are: ARIPUC (*Annual Report of the Institute of Phonetics, University of Copenhagen*); CUP (Cambridge University Press); ICPhS (*Proceedings of the International Congress of the Phonetic Sciences*); JASA (*Journal of the Acoustical Society of America*); and OUP (Oxford University Press).

Aakjær, Svend (1925). Tingsted og Maalkærne. *Frem A* 1, 571–575.

Abrahams, Henrik (1949). *Études phonétiques sur les tendances évolutives des occlusives germaniques.* Aarhus Universitetsforlag.

Ács, Péter, Katalin Fenyvesi, and Henrik Jørgensen (2008). Dansk fonologi og morfologi set i lyset af de såkaldte naturlighedsteorier. *Nydanske Sprogstudier* 36, 10–37. DOI: 10.7146/nys.v36i36.13466.

Ács, Péter and Henrik Jørgensen (2016). Hvorfor er dansk vanskeligt? Danske konsonantsegmenters to ansigter. *Skandinavisztikaj Füzetek* 10, 89–100.

Ag, Astrid (2010). *Sprogbrug og identitetsarbejde hos senmoderne storbypiger.* University of Copenhagen. (Københavnerstudier i tosprogethed 53).

Alexander, Gerda Isolde (1982). *Fortis and lenis in Germanic*. PhD dissertation, University of Alberta. Published 1983 by Peter Lang (European University Studies 734).

Allen, J. Sean, Joanne L. Miller, and David DeSteno (2003). Individual talker differences in voice-onset-time. *JASA* 113 (1), 544–552. doi: 10.1121/1.1528172.

Ambridge, Ben, Evan Kidd, Caroline F. Rowland, and Anna L. Theakston (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42 (2), 239–273. doi: 10.1017/S030500091400049X.

Andersen, Børge (1959). *Rønnemålet. En strukturallingvistisk analyse af udtryksplanet i en bornhomsk dialekt*. Schultz (Udvalg for Folkemaals Publikationer A18).

Andersen, Henning (1972). Diphthongization. *Language* 48 (1), 11–50. doi: 10.2307/412489.

Andersen, Poul (1955). "Klusilspring" i danske Dialekter. *Nordisk Tidsskrift for Tale og Stemme* 15 (1), 70–77.

Andersen, Preben (1981a). The effect of increased speaking rate and (intended) loudness on the glottal behaviour in stop consonant production as exemplified by Danish *p*. *ARIPUC* 15, 103–146.

Andersen, Torben Arboe (1981b). Dialektbånd og databehandling. *Ord & Sag* 1, 11–18.

Anderson, John M. and Colin J. Ewen (1987). *Principles of Dependency Phonology*. CUP (Cambridge Studies in Linguistics 47). doi: 10.1017/CBO9780511753442.

Avery, Peter and William J. Idsardi (2001). Laryngeal dimensions. Completion and enhancement. In: *Distinctive feature theory*. Ed. by Tracy Alan Hall. Mouton de Gruyter (Phonology and Phonetics 2), 41–70. doi: 10.1515/9783110886672.41.

Avery, Peter and Keren Rice (1989). Segment structure and coronal underspecification. *Phonology* 6 (2), 179–200. doi: 10.1017/S0952675700001007.

Baayen, R. Harald, Shravan Vasishth, Reinhold Kliegl, and Douglas Bates (2017). The cave of shadows. Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language* 94, 206–234. doi: 10.1016/j.jml.2016.11.006.

Baayen, R. Harald (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. CUP. doi: 10.1017/CBO9780511801686.

Baayen, R. Harald, Jacolien van Rij, Cecile de Cat, and Simon N. Wood (2018). Autocorrelated errors in experimental data in the language sciences. Some solutions offered by generalized additive mixed models. In: *Mixed-effects regression models in linguistics*. Ed. by Dirk Speelman, Kris Heylen, and Dirk Geeraerts. Springer (Quantitative Methods in the Humanities and Social Sciences), 49–69. doi: 10.1007/978-3-319-69830-4_4.

Bach, Emmon and Robert T. Harms (1972). How do languages get crazy rules? In: *Linguistic change and generative theory*. Ed. by Robert P. Stockwell and Ronald K.S. Macaulay. Indiana University Press (Indiana University Studies in the History and Theory of Linguistics), 1–21.

Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand (1993). Some patterns of linguistic diffusion. *Language Variation and Change* 5 (3), 359–390. DOI: 10.1017/S095439450000154X.

Bandle, Oskar, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, eds. (2002–2005). *The Nordic languages. An international handbook of the North Germanic languages.* Walter de Gruyter (Handbooks of Linguistics and Communication Science 22). DOI: 10.1515/9783110197051.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily (2013). Random effects structure for confirmatory hypothesis testing. Keep it maximal. *Journal of Memory and Language* 68 (3), 255–278. DOI: 10.1016/j.jml.2012.11.001.

Barton, Kamil (2020). `MuMIn`. Multi-model inference. (R package version 1.43.17).

Basbøll, Hans (1968). The phoneme system of Advanced Standard Copenhagen. *ARIPUC* 3, 33–54.

Basbøll, Hans (1972). Some conditioning phonological factors for the pronunciation of short vowels in Danish with special reference to syllabification. *ARIPUC* 6, 185–210.

Basbøll, Hans (1975). On the phonological interpretation of the falling diphthongs in Danish. *ARIPUC* 9, 49–108.

Basbøll, Hans (1977). The structure of the syllable and a proposed hierarchy of distinctive features. In: *Phonologica 1976.* Ed. by Wolfgang U. Dressler and Oskar E. Pfeiffer. Universität Innsbruck (Innsbrucker Beiträge zur Sprachwissenschaft 19), 143–148.

Basbøll, Hans (1985). Stød in Modern Danish. *Folia Linguistica* 19 (1/2), 1–50. DOI: 10.1515/flin.1985.19.1-2.1.

Basbøll, Hans (1994). How to derive the sonority syllable from the prototypical peak. *Acta Linguistica Hafniensia* 27 (1), 51–65. DOI: 10.1080/03740463.1994.10420393.

Basbøll, Hans (1999). Syllables in Danish. In: *The syllable. Views and facts.* Ed. by Harry van der Hulst and Nancy Ritter. Walter de Gruyter (Studies in Generative Grammar 45), 69–92. DOI: 10.1515/9783110806793.69.

Basbøll, Hans (2003). Prosody, productivity and word structure. The stød pattern of Modern Danish. *Nordic Journal of Linguistics* 26 (1), 5–44. DOI: 10.1017/S033258650300101X.

Basbøll, Hans (2005). *The phonology of Danish.* OUP. (The Phonology of the World's Languages).

Basbøll, Hans (2015). Morphophonology. In: *International encyclopedia of the social and behavioral sciences.* Ed. by James D. Wright. 2nd ed. Vol. Linguistics B. Elsevier, 826–833. DOI: 10.1016/B978-0-08-097086-8.53036-0.

Basbøll, Hans (2018). Danish imperative formation. A problem for the phonology/morphology interface. *Scandinavian Philology* 16 (2), 219–244. DOI: 10.21638/11701/spbu21.2018.203.

Basbøll, Hans and Johannes Wagner (1985). *Kontrastive Phonologie des Deutschen und Dänischen. Segmentale Wortphhonologie und -phonetik.* Max Niemeyer (Linguistische Arbeiten 160). DOI: 10.1515/9783111358345.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth, and R. Harald Baayen (2015a). Parsimonious mixed models. (Unpublished manuscript). arXiv: 1506.04967v2.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steven Walker (2015b). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67 (1), 1–48. doi: 10.18637/jss.v067.i01.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steven Walker (2021). `lme4`. Linear mixed-effects models using "Eigen" and S4. (R package version 1.1-27.1).

Baudouin de Courtenay, Jan (1895). *Versuch einer Theorie phonetischer Alternationen. Ein Capitel aus der Psychophonetik.* Karl J. Trübner. doi: 10.1515/9783111472317.

Bauer, Alexander (2021). `FoSIntro`. Convenience functions for function-on-scalar regression. (R package version 1.0.3, available on GitHub).

Bauer, Alexander, Fabian Scheipl, Helmut Küchenhoff, and Alice-Agnes Gabriel (2018). An introduction to semiparametric function-on-scalar regression. *Statistical Modelling* 18 (3), 346–364. doi: 10.1177/1471082X17748034.

Bauer, Alexander, Fabian Scheipl, Helmut Küchenhoff, and Alice-Agnes Gabriel (2021). Registration for incomplete non-Gaussian functional data. (Unpublished manuscript, in review with *Journal of Computational and Graphical Statistics*). arXiv: 2108.05634.

Beckman, Jill, Michael Jessen, and Catherine Ringen (2013). Empirical evidence for laryngeal features. Aspirating vs. true voice languages. *Journal of Linguistics* 49 (2), 259–284. doi: 10.1017/S0022226712000424.

Beguš, Gašper (2022). Distinguishing cognitive from historical influences in phonology. *Language* 98 (1), 1–34. doi: 10.1353/lan.2021.0084.

Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory, and Daniel Gildea (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *JASA* 113 (2), 1001–1024. doi: 10.1121/1.1534836.

Bennike, Valdemar and Marius Kristensen (1898–1912). *Kort over de danske folkemål med forklaringer.* Gyldendalske.

Bermúdez-Otero, Ricardo (2015). Amphichronic explanation and the life cycle of phonological processes. In: *The Oxford handbook of historical phonology.* Ed. by Patrick Honeybone and Joseph C. Salmons. OUP, 374–399. doi: 10.1093/oxfordhb/9780199232819.013.014.

Berry, Jeff and Maura Moyle (2011). Covariation among vowel height effects on acoustic measures. *JASA* 130 (5), 365–371. doi: 10.1121/1.3651095.

Birkmann, Thomas (2002). From Ancient Nordic to Old Nordic. Definition and delimitation of the period. In: *The Nordic languages. An international handbook of the history of the North Germanic languages.* Ed. by Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman. Vol. 1. Walter de Gruyter (Handbooks of Linguistics and Communication Science 22.1), 691–698. doi: 10.1515/9783110197051-079.

Biswas, Mityl (2022). *Hypothesis testing in function on scalar regression.* PhD dissertation, North Carolina State University.

Bjerrum, Anders (1944). *Fjoldemålets lydsystem.* Ejnar Munksgaard.

Blacklock, Oliver S. (2004). *Characteristics of variation in production of normal and disordered fricatives, using reduced-variance spectral methods.* PhD dissertation, University of Southampton.

Bleses, Dorthe (2000). Transparens og produktivitet i danske børns tilegnelse af verbers præteritum. *Nydanske studier & Almen kommunikationsteori* 26/27, 165–196. DOI: 10.7146/nys.v26i26-27.13414.

Blevins, Juliette (2004). *Evolutionary Phonology. The emergence of sound patterns.* CUP. DOI: 10.1017/CBO9780511486357.

Blevins, Juliette (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32 (2), 117–166. DOI: 10.1515/TL.2006.009.

Blevins, Juliette (2015). Evolutionary Phonology. A holistic approach to sound change typology. In: *The Oxford handbook of historical phonology.* Ed. by Patrick Honeybone and Joseph C. Salmons. OUP (Oxford Handbooks Online), 485–500. DOI: 10.1093/oxfordhb/9780199232819.013.006.

Blevins, Juliette and Andrew Garrett (1998). The origins of consonant–vowel metathesis. *Language* 74 (3), 508–556. DOI: 10.1353/lan.1998.0012.

Blevins, Juliette and Andrew B. Wedel (2009). Inhibited sound change. An evolutionary approach to lexical competition. *Diachronica* 26 (2), 143–183. DOI: 10.1075/dia.26.2.01ble.

Blumstein, Sheila E., Emily B. Myers, and Jesse Rissman (2005). The perception of voice onset time. An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience* 17 (9), 1353–1366. DOI: 10.1162/0898929054985473.

Blumstein, Sheila E. and Kenneth N. Stevens (1979). Acoustic invariance in speech production. Evidence from measurements of the spectral characteristics of stop consonants. *JASA* 66 (4), 1001–1017. DOI: 10.1121/1.383319.

Blumstein, Sheila E. and Kenneth N. Stevens (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *JASA* 67 (2), 648–662. DOI: 10.1121/1.383890.

Boberg, Charles (2000). Geolinguistic diffusion and the U.S.–Canada border. *Language Variation and Change* 12 (1), 1–24. DOI: 10.1017/S0954394500121015.

Boeck, Simon Skovgaard (2018). Udtale: Ældre nydansk. In: *Dansk sproghistorie 2. Ord for ord for ord.* Ed. by Ebba Hjorth. Aarhus Universitetsforlag, 179–195.

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5 (9/10), 341–345.

Boersma, Paul (2006). Prototypicality judgements as inverted perception. In: *Gradience in grammar. Generative perspectives.* Ed. by Gisbert Fanselow, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky. OUP, 168–184. DOI: 10.1093/acprof:oso/9780199274796.001.0001.

Boersma, Paul (2009). Cue constraints and their interactions in phonological perception and production. In: *Phonology in perception.* Ed. by Paul Boersma and Silke Hamann. Mouton de Gruyter (Phonology and Phonetics 15), 55–109. DOI: 10.1515/9783110219234.

Boersma, Paul (2011). A programme for bidirectional phonology and phonetics and their acquisition and evolution. In: *Bidirectional Optimality Theory.* Ed. by Anton

Benz and Jason Mattausch. John Benjamins (Linguistik Aktuell/Linguistics Today 180), 33–72. DOI: 10.1075/la.180.02boe.

Boersma, Paul, Titia Benders, and Klaas Seinhorst (2020). Neural network models for phonology and phonetics. *Journal of Language Modeling* 8 (1), 103–177. DOI: 10.15398/jlm.v8i1.224.

Boersma, Paul and Silke Hamann (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25 (2), 217–270. DOI: 10.1017/S0952675708001474.

Boersma, Paul and Jan-Willem van Leussen (2017). Efficient evaluation and learning in multilevel parallel constraint grammars. *Linguistic Inquiry* 48 (3), 349–388. DOI: 10.1162/ling_a_00247.

Boersma, Paul and David Weenink (2021). Praat. Doing phonetics by computer. (Version 6.2.04). URL: fon.hum.uva.nl.

Bohn, Ocke-Schwen (2013). Acoustic characteristics of Danish infant directed speech. *Proceedings of Meetings on Acoustics* 19. DOI: 10.1121/1.4798488.

de Boor, Carl (2001). *A practical guide to splines*. 2nd ed. Springer (Applied Mathematical Sciences 27).

Botma, Bert and Marijn van 't Veer (2013). A fraction too much friction. The phonological status of voiced fricatives. *Linguistics in the Netherlands* 30 (1), 46–60. DOI: 10.1075/avt.30.04bot.

Bounds, Paulina, Naomi Palosaari, and William A. Kretzschmar Jr. (2011). Issues in using legacy data. In: *Sociophonetics. A student's guide*. Ed. by Marianna Di Paolo and Malcah Yaeger-Dror. Routledge, 46–57.

Brauer, Markus and John J. Curtin (2018). Linear mixed-effects models and the analysis of nonindependent data. A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods* 23 (3), 389–411. DOI: 10.1037/met0000159.

Braun, Angelika (1996). Zur regionalen Distribution von VOT im Deutschen. In: *Untersuchungen zu Stimme und Sprache*. Ed. by Angelika Braun. Franz Steiner (Zeitschrift für Dialektologie und Lingvistik), 19–32.

Braunmüller, Kurt (2002). The Ancient-Nordic linguistic system from a typological point of view. Phonology, graphemics, morphology, syntax, and word order. In: *The Nordic languages. An international handbook of the history of the North Germanic languages*. Ed. by Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman. Vol. 1. Walter de Gruyter (Handbooks of Linguistics and Communication Science 22.1), 649–656. DOI: 10.1515/9783110197051-074.

Braunschweiler, Norbert (1997). Integrated cues of voicing and vowel length in German. A production study. *Language and Speech* 40 (4), 353–376. DOI: 10.1177/002383099704000403.

Brink, Lars (2013). Nyere danske lydlove. Udtaleændringer siden 1955-generationen. *Mål og Mæle* 36 (1), 24–30.

Brink, Lars and Jørn Lund (1975). *Dansk rigsmål. Lydudviklingen siden 1840 med særligt henblik på sociolekterne i København*. Gyldendal.

Brink, Lars and Jørn Lund (2018). Udtale: Yngre nydansk. In: *Dansk sproghistorie 2. Ord for ord for ord.* Ed. by Ebba Hjorth. Aarhus Universitetsforlag, 197–228.

Brink, Lars, Jørn Lund, Steffen Heger, and Jens Normann Jørgensen (1991). *Den store danske udtaleordbog*. Munksgaard (Munksgaards Ordbøger).

Britain, David (2010). Conceptualizations of geographic space in linguistics. In: *Language and space. An international handbook of linguistic variation. Volume 2: Language mapping*. Ed. by Alfred Lameli, Roland Kehrein, and Stefan Rabanus. De Gruyter Mouton (Handbooks of Linguistics and Communication Science 30/2), 69–102. DOI: 10.1515/9783110219166.1.69.

Broe, Michael B. (1992). An introduction to feature geometry. In: *Gesture, segment, prosody*. Ed. by Gerard J. Docherty and D. Robert Ladd. CUP (Papers in Laboratory Phonology 2), 149–165. DOI: 10.1017/CBO9780511519918.007.

Brotherton, Chloe and Aleese Block (2020). Soft d in Danish. Acoustic characteristics and issues in transcription. *Proceedings of the Linguistic Society of America* 5, 792–797. DOI: 10.3765/plsa.v5i1.4739.

Browman, Catherine P. and Louis M. Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook* 3, 219–252. DOI: 10.1017/S0952675700000658.

Browman, Catherine P. and Louis M. Goldstein (1992). Articulatory Phonology. An overview. *Phonetica* 49 (3/4), 155–180. DOI: 10.1159/000261913.

Brown, Jason (2006). An evolutionary perspective on stop inventories. *Theoretical Linguistics* 32 (2), 175–184. DOI: 10.1515/TL.2006.011.

Brøndum-Nielsen, Johs. (1927). *Dialekter og dialektforskning*. J.H. Schulz.

Brøndum-Nielsen, Johs. (1928–1973). *Gammeldansk grammatik i sproghistorisk fremstilling*. Schultz.

Bunnell, H. Timothy, James Polikoff, and Jane McNicholas (2004). Spectral moment vs. bark cepstral analysis of children's word-initial voiceless stops. In: *8th International Conference on Spoken Language Processing*. International Speech Communication Association (Interspeech). DOI: 10.21437/Interspeech.2004-81.

Bybee, Joan L. (1985). *Morphology. A study of the relation between meaning and form.* John Benjamins (Typological Studies in Language 9). DOI: 10.1075/tsl.9.

Bybee, Joan L. (2000a). Lexicalization of sound change and alternating environments. In: *Acquisition and the lexicon.* Ed. by Michael B. Broe and Janet B. Pierrehumbert. CUP (Papers in Laboratory Phonology 5), 250–268. Reprinted 2007 in *Frequency of use and the organization of language.* OUP, 216–234. DOI: 10.1093/acprof:oso/9780195301571.003.0010.

Bybee, Joan L. (2000b). The phonology of the lexicon. Evidence from lexical diffusion. In: *Usage-based models of language.* Ed. by Michael Barlow and Suzanne Kemmer. Center for the Study of Language and Information Publications, 65–86. Reprinted 2007 in *Frequency of use and the organization of language.* OUP, 199–215. DOI: 10.1093/acprof:oso/9780195301571.003.0009.

Bybee, Joan L. (2001). *Phonology and language use.* CUP (Cambridge Studies in Linguistics 94). DOI: 10.1017/CBO9780511612886.

Bybee, Joan L. (2006). From usage to grammar. The mind's response to repetition. *Language* 82 (4), 711–733. DOI: 10.1353/lan.2006.0186.

Bybee, Joan L. and Dan I. Slobin (1982). Rules and schemas in the development and use of the English past tense. *Language* 58 (2), 265–289. DOI: 10.2307/414099.

Byrd, Dani (1993). 54,000 American stops. *UCLA Working Papers in Phonetics* 83, 97–116.

Byrd, Dani, Abigail Kaun, Shrikanth Narayanan, and Elliott Saltzman (2000). Phrasal signatures in articulation. In: *Acquisition and the lexicon.* Ed. by Michael B. Broe and Janet B. Pierrehumbert. CUP (Papers in Laboratory Phonology 5), 70–87.

Carignan, Christopher, Phil Hoole, Esther Kunay, Marianne Pouplier, Arun Joseph, Dirk Voit, Jens Frahm, and Jonathan Harrington (2020). Analyzing speech in both time and space. Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI. *Laboratory Phonology* 11 (2). DOI: 10.5334/labphon.214.

Castro, P.E., W.H. Lawton, and E.A. Sylvestre (1986). Principal modes of variation for processes with continuous sample curves. *Technometrics* 28 (4), 329–337. DOI: 10.2307/1268982.

Cederbaum, Joan, Marianne Pouplier, Phil Hoole, and Sonja Greven (2016). Functional linear mixed models for irregularly or sparsely sampled data. *Statistical Modelling* 16 (1), 67–88. DOI: 10.1177/1471082X15617594.

Chambers, J.K. (2000). Region and language variation. *English World-Wide* 21 (2), 169–199. DOI: 10.1075/eww.21.2.02cha.

Chambers, J.K. and Peter Trudgill (1998). *Dialectology.* 2nd ed. CUP (Cambridge Textbooks in Linguistics). DOI: 10.1017/CBO9780511805103.

Chládková, Kateřina (2014). *Finding phonological features in perception.* PhD dissertation, University of Amsterdam. HDL: 11245/1.410161.

Cho, Taehong (2001). Effects of morpheme boundaries on intergestural timing. Evidence from Korean. *Phonetica* 58 (3), 129–162. DOI: 10.1159/000056196.

Cho, Taehong and Peter Ladefoged (1999). Variation and universals in VOT. Evidence from 18 languages. *Journal of Phonetics* 27 (2), 207–229. DOI: 10.1006/jpho.1999.0094.

Chodroff, Eleanor, Alessandra Golden, and Colin Wilson (2019). Covariation of stop voice onset time across languages. Evidence for a universal constraint on phonetic realization. *JASA* 145 (1), 106–115. DOI: 10.1121/1.5088035.

Chodroff, Eleanor and Colin Wilson (2014). Burst spectrum as a cue for the stop voicing contrast in American English. *JASA* 136 (5), 2762–2772. DOI: 10.1121/1.4896470.

Chodroff, Eleanor and Colin Wilson (2017). Structure in talker-specific phonetic realization. Covariation of stop consonant VOT in American English. *Journal of Phonetics* 61, 30–47. DOI: 10.1016/j.wocn.2017.01.001.

Chodroff, Eleanor and Colin Wilson (2018). Predictability of stop consonant phonetics across talkers. Between-category and within-category dependencies among cues for place and voice. *Linguistics Vanguard* 4 (s2). DOI: 10.1515/lingvan-2017-0047.

Chomsky, Noam and Morris Halle (1968). *The sound pattern of English.* Harper & Row (Studies in Language).

Clausen, Margit Carolin and Annette Fox-Boyer (2017). Phonological development of Danish-speaking children. A normative cross-sectional study. *Clinical Linguistics & Phonetics* 31 (6), 440–458. DOI: 10.1080/02699206.2017.1308014.

Clements, George N. (1990). The role of the sonority cycle in core syllabification. In: *Between the grammar and physics of speech*. Ed. by John Kingston and Mary E. Beckman. CUP (Papers in Laboratory Phonology 1), 283–333. DOI: 10.1017/CBO9780511627736.017.

Cohen, A., C.L. Ebeling, K. Fokkema, and A.G.F. van Holk (1959). *Fonologie van het Nederlands en het Fries. Inleiding tot de moderne klankleer*. Nijhoff.

Corson, David (1984). The lexical bar. Lexical change from 12 to 15 years measured by social class, region and ethnicity. *British Educational Research Journal* 10 (2), 115–133. DOI: 10.1080/0141192840100201.

Cristia, Alejandrina, Sho Tsuji, and Christina Bergmann (2020). How to do theory evaluation in the age of cumulative science. (Unpublished manuscript, in review with *Meta-Psychology*).

Davidson, Lisa (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54, 35–50. DOI: 10.1016/j.wocn.2015.09.003.

Dodd, Barbara, Alison Holm, Zhu Hua, and Sharon Crosbie (2003). Phonological development. A normative study of British English-speaking children. *Clinical Linguistics & Phonetics* 17 (8), 617–643. DOI: 10.1080/0269920031000111348.

Dresher, B. Elan (2009). *The contrastive hierarchy in phonology*. CUP (Cambridge Studies in Linguistics 121). DOI: 10.1017/CBO9780511642005.

DS = *Dialektsamlingen* (1971–1976). Corpus of tape recordings collected by fieldworkers from the Peter Skautrup Center for Jutlandic Dialect Research, Aarhus University, and the Center for Dialect Research, University of Copenhagen. Digitized and hosted online by the Royal Danish Library. URL: dansklyd.statsbiblioteket.dk.

DSL = Det Danske Sprog- og Litteraturselskab (2018). *Den danske ordbog*. URL: ordnet.dk/ddo.

Dunn, Michael, Angela Terrill, Ger Reesink, Robert A. Foley, and Stephen C. Levinson (2005). Structural phylogenetics and the reconstruction of ancient language history. *Science* 309 (5743), 2072–2075. DOI: 10.1126/science.1114615.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (2020). *Ethnologue. Languages of the World*. 23rd ed. SIL International. URL: ethnologue.com.

Ejskjær, Inger (1967). *Kortvokalstødet i sjællandsk*. Akademisk Forlag (Udvalg for Folkemaals Publikationer A22).

Ejskjær, Inger (1990). Stød and pitch accents in the Danish dialects. *Acta Linguistica Hafniensia* 22 (1), 49–75. DOI: 10.1080/03740463.1990.10411522.

Ejskjær, Inger (1993). *Danish dialect research*. C.A. Reitzel (Universitets-jubilæets danske samfund 523).

Ejskjær, Inger (2006). Glottal stop (stød, parasitic plosive) and (distinctive) tonal accents in the Danish dialects. In: *Germanic tone accents*. Ed. by Michiel de Vaan. Franz Steiner (Zeitschrift für Dialektologie und Lingvistik 131), 25–33.

Ejstrup, Michael (2009). *Danske talesprog i begyndelsen af det tredje årtusinde. En undersøgelse af danske talesprog i Vest-, Syd- og Østdanmark med fokus på akustisk undersøgelse af vokaler.* PhD dissertation, University of Southern Denmark.

Ejstrup, Michael (2010). På godt forskelligt dansk. *Nydanske Sprogstudier* 39, 93–136. DOI: 10.7146/nys.v39i39.13534.

Ejstrup, Michael and Gert Foget Hansen (2003). Danish vowels in spontaneous speech in three modern regional variants. In: *ICPhS 15.* Ed. by Maria-Josep Solé, Daniel Recasens, and J. Romero. Universitat Autonoma De Barcelona, 2119–2122.

Eliasson, Stig (1992). Cognition in phonological change. In: *Diachrony within synchrony. Language history and cognition.* Ed. by Günter Kellerman and Michael D. Morrissey. Peter Lang (Duisburg Papers on Research in Language and Culture 14), 283–308.

Eliasson, Stig (1997). The cognitive calculus and its function in language. In: *Language change and functional explanations.* Ed. by Jadranka Gvozdanović. De Gruyter Mouton (Trends in Linguistics: Studies and Monographs 98), 53–70. DOI: 10.1515/9783110813753.53.

Ernestus, Mirjam, Mybeth Lahey, Femke Verhees, and R. Harald Baayen (2006). Lexical frequency and voice assimilation. *JASA* 120 (2), 1040–1051. DOI: 10.1121/1.2211548.

Espegaard, Arne (1995). *Nogle nordjyske mål.* Forlaget Vendsyssel.

Espy-Wilson, Carol Y., Suzanne Boyce, Michel Jackson, Shrikanth Narayanan, and Abeer Alwan (2000). Acoustic modeling of American English /r/. *JASA* 108 (1), 343–356. DOI: 10.1121/1.429469.

Fant, C. Gunnar M. (1960). *Acoustic theory of speech production with calculations based on x-ray studies of Russian articulations.* Mouton. DOI: 10.1515/9783110873429.

Fasiolo, Matteo, Raphaël Nedellec, Yannig Goude, and Christian Capezza (2021). `mgcViz`. Visualisations for generalized additive models. (R package version 0.1.9).

Fasiolo, Matteo, Raphaël Nedellec, Yannig Goude, and Simon N. Wood (2020). Scalable visualization methods for modern generalized additive models. *Journal of Computational and Graphical Statistics* 29 (1), 78–86. DOI: 10.1080/10618600.2019.1629942.

Ferguson, Sarah Hargus (2004). Talker differences in clear and conversational speech. Vowel intelligibility for normal-hearing listeners. *JASA* 116 (4), 2365–2373. DOI: 10.1121/1.1788730.

Fischer-Jørgensen, Eli (1952). On the definition of phoneme categories on a distributional basis. *Acta Linguistica* 7 (1/2), 8–39. DOI: 10.1080/03740463.1952.10415400.

Fischer-Jørgensen, Eli (1954). Acoustic analysis of stop consonants. *Le Maître Phonetique* 32 (69), 42–59. JSTOR: 44705403.

Fischer-Jørgensen, Eli (1968). Les occlusives françaises et danoises d'un sujet bilingue. *Word* 24 (1–3), 112–153. DOI: 10.1080/00437956.1968.11435519.

Fischer-Jørgensen, Eli (1969). Voicing, tenseness and aspiration in stop consonants, with special reference to French and Danish. *ARIPUC* 3, 63–114.

Fischer-Jørgensen, Eli (1972a). Identification of unreleased stop consonants. *ARIPUC* 6, 89–103.

Fischer-Jørgensen, Eli (1972b). Kinesthetic judgement of effort in the production of stop consonants. *ARIPUC* 6, 59–73.

Fischer-Jørgensen, Eli (1972c). ptk et bdg français en position intervocalique accentuée. In: *Papers in linguistics and phonetics to the memory of Pierre Delattre.* Ed. by Albert Valdman. Mouton (Janua Linguarum 54), 143–200. DOI: 10.1515/9783110803877-014.

Fischer-Jørgensen, Eli (1972d). Tape cutting experiments with Danish stop consonants in initial position. *ARIPUC* 6, 104–168.

Fischer-Jørgensen, Eli (1979). Temporal relations in consonant–vowel syllables with stop consonants based on Danish material. In: *Frontiers of speech communication research.* Ed. by Björn Lindblom and Sven Öhman. Academic Press, 51–68.

Fischer-Jørgensen, Eli (1980). Temporal relations in Danish tautosyllabic CV sequences with stop consonants. *ARIPUC* 14, 207–261.

Fischer-Jørgensen, Eli (1987). A phonetic study of stød in Standard Danish. *ARIPUC* 21, 55–267.

Fischer-Jørgensen, Eli (1989). Phonetic analysis of the stød in Standard Danish. *Phonetica* 46 (1/2/3), 1–59. DOI: 10.1159/000261828.

Fischer-Jørgensen, Eli and Hajime Hirose (1974). A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production. *Status Report on Speech Research* 39/40, 231–253.

Fischer-Jørgensen, Eli and Birgit Hutters (1981). Aspirated stop consonants before low vowels, a problem of delimitation, its causes and consequences. *ARIPUC* 15, 77–102.

Flege, James Emil (1982). Laryngeal timing and phonation onset in utterance-initial English stops. *Journal of Phonetics* 10 (2), 177–192. DOI: 10.1016/S0095-4470(19)30956-8.

Flege, James Emil, Murray J. Munro, and Robert Allen Fox (1994). Auditory and categorical effects on cross-language vowel perception. *JASA* 95 (6), 3623–3641. DOI: 10.1121/1.409931.

Flemming, Edward S. (1995). *Auditory representations in phonology.* PhD dissertation, University of California, Los Angeles. PROQUEST: 9601362. Published 2002 by Routledge. DOI: 10.4324/9781315054803.

Forrest, Karen, Gary Weismer, Paul Milenkovic, and Ronald N. Dougall (1988). Statistical analysis of word-initial voiceless obstruents. Preliminary data. *JASA* 84 (1), 115–123. DOI: 10.1121/1.396977.

Foster, Michael K. (1982). Alternating weak and strong syllables in Cayuga words. *International Journal of American Linguistics* 48 (1), 59–72. DOI: 10.1086/465713.

Fox, John and Sanford Weisberg (2019). *An R companion to applied regression.* 3rd ed. Sage.

Fox, John, Sanford Weisberg, and Brad Price (2021). `car`. Companion to applied regression. (R package version 3.0-9).

Francis, Alexander L., Valter Ciocca, and Jojo Man Ching Yu (2003). Accuracy and variability of acoustic measures of voicing onset. *JASA* 113 (2), 1025–1032. DOI: 10.1121/1.1536169.

Frederiksen, Britta Olrik (2018). Udtale: Gammeldansk. In: *Dansk sproghistorie 2. Ord for ord for ord*. Ed. by Ebba Hjorth. Aarhus Universitetsforlag, 151–178.

Frøkjær-Jensen, Børge, Carl Ludvigsen, and Jørgen Rischel (1971). A glottographic study of some Danish consonants. In: *Form and substance. Phonetic and linguistic papers presented to Eli Fischer-Jørgensen*. Ed. by L.L. Hammerich, Roman Jakobson, and Eberhard Zwirner. Akademisk Forlag, 123–140.

Frøkjær-Jensen, Børge, Carl Ludvigsen, and Jørgen Rischel (1973). A glottographic study of some Danish consonants. *ARIPUC* 7, 269–295.

Gajardo, Alvaro, Satarupa Bhattacharjee, Cody Carroll, Yaqing Chen, Xiongtao Dai, Jianing Fan, Pantelis Z. Hadjipantelis, Kyunghee Han, Hao Ji, Changbo Zhu, Hans-Georg Müller, and Jane-Ling Wang (2021a). `fdapace`. Functional data analysis and empirical dynamics. (R package version 0.5.8).

Gajardo, Alvaro, Satarupa Bhattacharjee, Cody Carroll, Yaqing Chen, Xiongtao Dai, Jianing Fan, Pantelis Z. Hadjipantelis, Kyunghee Han, Hao Ji, Changbo Zhu, Hans-Georg Müller, and Jane-Ling Wang (2021b). Functional PCA in R. (`fdapace` R package vignette).

Galéas, Grazia Crocco (2001). Uniformity, transparency of encoding, and biuniqueness within the model of Natural Morphology. *Studi Italiani di Linguistica Teorica ed Applicata* 30 (2), 397–405.

Gamkrelidze, Tamaz V. (1975). On the correlation of stops and fricatives in a phonological system. *Lingua* 35 (3-4), 231–261. DOI: 10.1016/0024-3841(75)90060-1.

Garibaldi, Camila Linn and Ocke-Schwen Bohn (2015). Phonetic similarity predicts ultimate attainment quite well. The case of Danish /i, y, u/ and /d, t/ for native speakers of English and of Spanish. In: *ICPhS 18*. University of Glasgow.

Garvin, Karee, Myriam Lapierre, and Sharon Inkelas (2018). A Q-theoretic approach to distinctive subsegmental timing. *Proceedings of the Linguistic Society of America* 3 (9). DOI: 10.3765/plsa.v3i1.4330.

Garvin, Karee, Myriam Lapierre, Martha Schwarz, and Sharon Inkelas (2020). Modeling vowel quantity scales in Q Theory. In: *Proceedings of the 2019 Annual Meeting on Phonology*. Ed. by Hyunah Baek, Chikako Takahashi, and Alex Hong-Lun Yeung. Linguistic Society of America. DOI: 10.3765/amp.v8i0.4685.

Gauchat, Louis (1903). Gibt es Mundartgrenzen? *Archiv für das Studium der neueren Sprachen und Litteraturen* 111 (3/4), 365–403.

Gelman, Andrew and Jennifer Hill (2006). *Data analysis using regression and multilevel/hierarchical models*. CUP (Analytical Methods for Social Research). DOI: 10.1017/CBO9780511790942.

Gilliéron, Jules and Edmond Edmont (1902–1910). *Atlas linguistique de la France*. Champion.

Goad, Heather (2012). sC clusters are (almost always) coda-initial. *The Linguistic Review* 29 (3), 335–373. DOI: 10.1515/tlr-2012-0013.

Goblirsch, Kurt (2018). *Gemination, lenition, and vowel lengthening. On the history of quantity in Germanic.* CUP (Cambridge Studies in Linguistics 157). DOI: 10.1017/9781139540780.

Goldinger, Stephen D. (1996). Words and voices. Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22 (5), 1166–1183. DOI: 10.1037/0278-7393.22.5.1166.

Goldshtein, Yonatan (forthc.). *Outline of a prosodic theory of stød.* (PhD dissertation to be submitted, Aarhus University).

Goldshtein, Yonatan and Liv Moeslund Ahlgren (2021). Ideologies of language and place. Conflicting expectations to dialectal speech between informants and dialectologists. *Journal of Postcolonial Linguistics* 5, 178–203.

Goldshtein, Yonatan and Rasmus Puggaard (2019). Overblik over danske dialektoptagelser. *Ord & Sag* 39, 18–28. HDL: 1887/85527.

Goldsmith, Jeff, Fabian Scheipl, Lei Huang, Julia Wrobel, Chongzhi Di, Jonathan Gellar, Jaroslaw Harezlak, Mathew W. McLean, Bruce Swihart, Luo Xiao, Ciprian Crainiceanu, and Philip T. Reiss (2021). `refund`. Regression with functional data. (R package version 0.1-24).

Goldsmith, John (1976). An overview of Autosegmental Phonology. *Linguistic Analysis* 2 (1), 23–68.

Goldsmith, John (1990). *Autosegmental and metrical phonology.* Basil Blackwell.

Goldstein, Louis M. and Catherine P. Browman (1986). Representation of voicing contrasts using articulatory gestures. *Journal of Phonetics* 14 (2), 339–342. DOI: 10.1016/S0095-4470(19)30662-X.

Gooskens, Charlotte (2005). Traveling time as a predictor of linguistic distance. *Dialectologia et Geolinguistica* 13, 38–62. DOI: 10.1515/dig.2005.2005.13.38.

Gooskens, Charlotte (2006). Linguistic and extra-linguistic predictors of inter-Scandinavian intelligibility. *Linguistics in the Netherlands* 23, 101–113. DOI: 10.1075/avt.23.12goo.

Gooskens, Charlotte (2020). The North Germanic dialect continuum. In: *The Cambridge handbook of Germanic linguistics.* Ed. by Michael T. Putnam and Richard Page. CUP (Cambridge Handbooks in Language and Linguistics), 761–782. DOI: 10.1017/9781108378291.033.

Gooskens, Charlotte, Vincent J. van Heuven, Renée van Bezooijen, and Jos J.A. Pacilly (2010). Is spoken Danish less intelligible than Swedish? *Speech Communication* 52 (11–12), 1022–1037. DOI: 10.1016/j.specom.2010.06.005.

Gordon, Matthew, Paul Barthmaier, and Kathy Sands (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32 (2), 141–174. DOI: 10.1017/S0025100302001020.

Gósy, Mária (2001). The VOT of the Hungarian voiceless plosives in words and in spontaneous speech. *International Journal of Speech Technology* 4 (1), 75–85. DOI: 10.1023/A:1009608900453.

Gráczi, Tekla Etelka and Anna Kohári (2014). Multiple bursts in Hungarian voiceless plosives and VOT measurements. In: *Proceedings of the 10th International Seminar* *on Speech Production*. Ed. by Susanne Fuchs, Martine Grice, Anne Hermes, Leonardo Lancia, and Doris Mücke. University of Cologne, 158–161.

Gregersen, Frans (2009). The data and design of the LANCHART study. *Acta Linguistica Hafniensia* 41 (1), 3–29. DOI: 10.1080/03740460903364003.

Gregersen, Frans, Marie Maegaard, and Nicolai Pharao (2014). The LANCHART corpus. In: *The Oxford handbook of corpus phonology*. Ed. by Jacques Durand, Ulrike Gut, and Gjert Kristoffersen. OUP (Oxford Handbooks Online), 534–545. DOI: 10.1093/oxfordhb/9780199571932.013.007.

Gregersen, Frans and Inge Lise Pedersen, eds. (1991). *The Copenhagen study in urban sociolinguistics*. C.A. Reitzel (Institut for Dansk Dialektforsknings publikationer A30).

Greven, Sonja and Fabian Scheipl (2017a). A general framework for functional regression modelling. *Statistical Modelling* 17 (1–2), 1–35. DOI: 10 . 1177 / 1471082X16681317.

Greven, Sonja and Fabian Scheipl (2017b). Rejoinder. *Statistical Modelling* 17 (1–2), 100–115. DOI: 10.1177/1471082X16689188.

Grønnum, Nina (1995). Danish vowels. Surface contrast versus underlying form. *Phonetica* 52 (3), 215–220. DOI: 10.1159/000262173.

Grønnum, Nina (1998). Illustrations of the IPA. Danish. *Journal of the International Phonetic Association* 28 (1/2), 99–105. DOI: 10.1017/S0025100300006290.

Grønnum, Nina (2005). *Fonetik og fonologi. Almen og dansk*. 3rd ed. Akademisk Forlag.

Grønnum, Nina (2009). A Danish phonetically annotated spontaneous speech corpus (DanPASS). *Speech Communication* 51 (7), 594–603. DOI: 10.1016/j.specom.2008.11.002.

Grønnum, Nina (2016). DanPASS. Danish Phonetically Annotated Speech. URL: danpass.hum.ku.dk.

Grønnum, Nina and Hans Basbøll (2001). Consonant length, stød and morae in Standard Danish. *Phonetica* 58 (4), 230–253. DOI: 10.1159/000046177.

Grønnum, Nina and Hans Basbøll (2003). Stød and length in Standard Danish. Experiments in laboratory phonology. In: *ICPhS 15*. Ed. by Maria-Josep Solé, Daniel Recasens, and J. Romero. Universitat Autonoma De Barcelona, 455–458.

Grønnum, Nina and Hans Basbøll (2007). Danish stød. Phonological and cognitive issues. In: *Experimental approaches to phonology*. Ed. by Maria-Josep Solé, Patrice Speeter Beddor, and Manjari Ohala. OUP, 192–206.

Grønnum, Nina and John Tøndering (2007). Question intonation in non-scripted Danish dialogues. In: *ICPhS 16*. Ed. by Jürgen Trouvain and William J. Barry. Saarland University, 1229–1232.

Grønnum, Nina, Miguel Vazquez-Larruscaín, and Hans Basbøll (2013). Danish stød. Laryngealization or tone. *Phonetica* 70 (1/2), 66–92. DOI: 10.1159/000354640.

Gubian, Michele, Jonathan Harrington, Mary Stevens, Florian Schiel, and Paul Warren (2019). Tracking the New Zealand English NEAR/SQUARE merger using functional principal components analysis. In: *Proceedings of Interspeech*. International Speech Communication Association, 296–300. DOI: 10.21437/Interspeech.2019-2115.

Gubian, Michele, Francisco Torreira, and Lou Boves (2015). Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts. *Journal of Phonetics* 49, 16–40. DOI: 10.1016/j.wocn.2014.10.001.

Gudiksen, Asgerd and Henrik Hovmark (2008). Båndoptagelser som kilde til Ømålsordbogen. *Nordiske Studier i Leksikografi* 9, 173–182.

Guion, Susan Guignard (1998). The role of perception in sound change of velar palatalization. *Phonetica* 55 (1/2), 18–52. DOI: 10.1159/000028423.

Gurevich, Naomi (2004). *Lenition and contrast. The functional consequences of certain phonetically conditioned sound changes.* Routledge. DOI: 10.4324/9780203505052.

Guy, Gregory R. (1991). Explanation in variable phonology. An exponential model of morphological constraints. *Language Variation and Change* 3 (1), 1–22. DOI: 10.1017/S0954394500000429.

Haberland, Hartmut (1994). Danish. In: *The Germanic languages.* Ed. by Ekkehard König and Johan van der Auwera. Routledge (Routledge Language Family Descriptions), 313–348. DOI: 10.4324/9781315812786.

Hale, Mark and Charles Reiss (2000). "Substance abuse" and "dysfunctionalism". Current trends in phonology. *Linguistic Inquiry* 31 (1), 157–169. DOI: 10.1162/002438900554334.

Hale, Mark and Charles Reiss (2008). *The phonological enterprise.* OUP.

Hall, Tracy Alan (1989). Lexical Phonology and the distribution of German [ç] and [x]. *Phonology* 6 (1), 1–17. DOI: 10.1017/S0952675700000920.

Hall, Tracy Alan (1997). *The phonology of coronals.* John Benjamins (Current Issues in Linguistic Theory 149). DOI: 10.1075/cilt.149.

Hall, Tracy Alan (2007). Segmental features. In: *The Cambridge handbook of phonology.* Ed. by Paul de Lacy. CUP, 311–334. DOI: 10.1017/CBO9780511486371.014.

Hall, Tracy Alan and Silke Hamann (2006). Towards a typology of stop assibilation. *Linguistics* 44 (6), 1195–1236. DOI: 10.1515/ling.2006.039.

Hall, Tracy Alan, Silke Hamann, and Marzena Żygis (2006). The phonetic motivation for phonological stop assibilation. *Journal of the International Phonetic Association* 36 (1), 59–81. DOI: 10.1017/S0025100306002453.

Halle, Morris and Kenneth N. Stevens (1971). A note on laryngeal features. *Quarterly Progress Report, Research Laboratory of Electronics, Massachusetts Institute of Technology* 101, 198–213. Reprinted 2003 in *From memory to speech and back.* Mouton de Gruyter (Phonology and Phonetics 3), 45–61. DOI: 10.1515/9783110871258.45.

Hamann, Silke (2006). On the use of Evolutionary Phonology for phonological theory. *Theoretical Linguistics* 32 (2), 197–204. DOI: 10.1515/TL.2006.013.

Hamann, Silke (2009). The learner of a perception grammar as a source of sound change. In: *Phonology in perception.* Ed. by Paul Boersma and Silke Hamann. Mouton de Gruyter (Phonology and Phonetics 15), 111–149. DOI: 10.1515/9783110219234.

Hamann, Silke and Ilaria E. Colombo (2017). A formal account of the interaction of orthography and perception. English intervocalic consonants borrowed into

Italian. *Natural Language and Linguistic Theory* 35 (3), 683–714. DOI: 10.1007/s11049-017-9362-3.
Hamann, Silke and Anke Sennema (2005). Acoustic differences between German and Dutch labiodentals. *ZAS Papers in Linguistics* 42, 33–41. DOI: 10.21248/zaspil.42.2005.272.
Hansen, Erik and Jørn Lund (1983). *Sæt tryk på. Syntaktisk tryk i dansk.* Lærerforeningens Materialeudvalg.
Hansen, Gert Foget (2015). *Stød og stemmekvalitet. En akustisk–fonetisk undersøgelse af ændringer i stemmekvaliteten i forbindelse med stød.* PhD dissertation, University of Copenhagen.
Hansen, Gert Foget and Nicolai Pharao (2010). Prosody in the Copenhagen multiethnolect. In: *Multilingual urban Scandinavia. New linguistic practices.* Ed. by Pia Quist and Bente Ailin Svendsen. Multilingual Matters, 79–95. DOI: 10.21832/9781847693143-011.
Hansen, Inger Schoonderbeek (2020). Jysk Ordbog. Af hvem, til hvem og hvorfor? *Nordiske Studier i Leksikografi* 15, 135–143.
Hansen, Peter Molbæk (1978). Stød and syllabicity in a Jutlandic dialect. *ARIPUC* 12, 15–24.
Hansen, Peter Molbæk (1981). *Udvalgte fonologiske temaer i to Himmerlandske ideolekter tilligemed en kritisk gennemgang af stavelsesproblematikken i jysk dialektologi.* Master's thesis, University of Copenhagen.
Hansen, Aage (1956). *Udtalen i moderne dansk.* Gyldendal.
Hansen, Aage (1962–1971). *Den lydlige udvikling i dansk.* G.E.C. Gad.
Hanson, Helen M. (2009). Effects of obstruent consonants on fundamental frequency at vowel onset in English. *JASA* 125 (1), 425–441. DOI: 10.1121/1.3021306.
Hanson, Helen M. and Kenneth N. Stevens (2003). Models of aspirated stops in English. In: *ICPhS 15.* Ed. by Maria-Josep Solé, Daniel Recasens, and J. Romero. Universitat Autonoma De Barcelona, 783–786.
Hardwicke, Tom E., Maya B. Mathur, Kyle MacDonald, Gustav Nilsonne, George C. Banks, Mallory C. Kidwell, Alicia Hofelich Mohr, Elizabeth Clayton, Erica J. Yoon, Michael Henry Tessler, Richie L. Lenne, Sara Altman, Bria Long, and Michael C. Frank (2018). Data availability, reusability, and analytic reproducibility. Evaluation the impact of a mandatory open data policy at the journal *Cognition. Royal Society Open Science* 5 (8). DOI: 10.1098/rsos.180448.
Harris, John (1994). *English sound structure.* Blackwell.
Hay, Jennifer and Paul Foulkes (2016). The evolution of medial /t/ over real and remembered time. *Language* 92 (2), 298–330. DOI: 10.1353/lan.2016.0036.
Hayes, Bruce (1999). Phonetically driven phonology. The role of Optimality Theory and inductive grounding. In: *Functionalism and formalism in linguistics. Volume I: General papers.* Ed. by Michael Darnell, Edith M. Moravcsik, Michael Noonan, Frederick J. Newmeyer, and Kathleen M. Wheatley. John Benjamins (Studies in Language Companion Series 41), 243–286. DOI: 10.1075/slcs.41.
Hayes, Bruce (2009). *Introductory phonology.* Wiley–Blackwell (Blackwell Textbooks in Linguistics).

Hayes, Bruce and Donca Steriade (2004). Introduction. The phonetic bases of phonological markedness. In: *Phonetically based phonology*. Ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade. CUP, 1–33. DOI: 10.1017/CBO9780511486401.001.
Heegård, Jan (2013). Morphologisation or reduction by context? The *-te* ending on adjectives and preterite verb forms in Standard Copenhagen Danish. *Acta Linguistica Hafniensia* 45 (1), 100–125. DOI: 10.1080/03740463.2013.849901.
Heger, Steffen (1979). Den fonetiske udvikling hos børn. In: *Børnenes sprog. Sprogene omkring børn*. Ed. by Mogens Jansen and Jørn Lund. Hans Reitzel, 13–67.
Heger, Steffen (1981). *Sprog og lyd. Elementær dansk fonetik*. 2nd ed. Akademisk Forlag.
Helgason, Pétur and Catherine Ringen (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36 (4), 607–628. DOI: 10.1016/j.wocn.2008.02.003.
Henton, Caroline, Peter Ladefoged, and Ian Maddieson (1992). Stops in the world's languages. *Phonetica* 49 (2), 65–101. DOI: 10.1159/000261905.
Higgins, Maureen B., Ronald Netsell, and Laura Schulte (1998). Vowel-related differences in laryngeal articulatory and phonatory function. *Journal of Speech, Language, and Hearing Research* 41 (4), 712–724. DOI: 10.1044/jslhr.4104.712.
Hirose, Hajime and Thomas Gay (1972). The activity of the intrinsic laryngeal muscles in voicing control. *Phonetica* 25 (3), 140–164. DOI: 10.1159/000259378.
Hjelmslev, Louis (1936). On the principles of phonematics. In: *ICPhS 2*. Ed. by Daniel Jones and Dennis Butler Fry. CUP, 49–54.
Hjelmslev, Louis (1943). *Omkring sprogteoriens grundlæggelse*. Munksgaard. English translation published 1961 as *Prolegomena to a theory of language*. University of Wisconsin Press.
Hjelmslev, Louis (1951). Grundtræk af det danske udtrykssystem med særligt henblik på stødet. *Selskab for Nordisk Filologi. Årsberetning for* 1948/49/50, 12–21. English translation published 1973 in *Essais linguistiques II*. Nordisk Sprog- og Kulturforlag (Travaux du Cercle linguistique de Copenhague 14), 247–266.
Hjorth, Ebba, ed. (2017–2022). *Dansk sproghistorie*. Aarhus Universitetsforlag.
Hock, Hans Henrich (1991). *Principles of historical linguistics*. 2nd ed. Mouton de Gruyter (Trends in Linguistics: Studies and Monographs 34). DOI: 10.1515/9783110219135.
Holtz, Annie (2021). Ambiguity and variable phonological rules. The challenge of TD deletion in US English. In: *Proceedings of the 29th Conference of the Student Organization of Linguistics in Europe*. Ed. by Annie Holtz, Iva Kovač, Rasmus Puggaard-Rode, and Joanna Wall. Leiden University Centre for Linguistics (ConSOLE XXIX), 98–115.
Honeybone, Patrick (2002). *Germanic obstruent lenition. Some mutual implications of theoretical and historical phonology*. PhD dissertation, University of Newcastle upon Tyne. HDL: 10443/3121.
Honeybone, Patrick (2005). Diachronic evidence in segmental phonology. The case of obstruent laryngeal specifications. In: *The internal organization of phonological segments*. Ed. by Marc van Oostendorp and Jeroen van de Weijer. Mouton

de Gruyter (Studies in Generative Grammar 77), 318–351. DOI: 10 . 1515 / 9783110890402.317.

Honeybone, Patrick (2008). Lenition, weakening and consonantal strength. Tracing concepts through the history of phonology. In: *Lenition and fortition*. Ed. by Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral. Mouton de Gruyter (Studies in Generative Grammar 99), 9–92. DOI: 10.1515/9783110211443.1.9.

Hoole, Phil, Bernd Pompino-Marschall, and M. Dames (1984). Glottal timing in German voiceless occlusives. In: *ICPhS 10*. Ed. by M.P.R. Van den Broecke and A. Cohen. Foris (Netherlands Phonetics Archives 2B), 399–403. DOI: 10.1515/9783110884685-059.

Hooper, Joan Bybee (1976). Word frequency in lexical diffusion and the source of morphophonological change. In: *Current progress in historical linguistics*. Ed. by William M. Christie. North-Holland (North-Holland Linguistic Series 31), 96–105. Reprinted 2007 in *Frequency of use and the organization of language*. OUP, 23–34. DOI: 10.1093/acprof:oso/9780195301571.001.0001.

Horslund, Camilla Søballe, Henrik Jørgensen, and Rasmus Puggaard (2021a). En alternativ, fonetisk baseret fonemanalyse af det danske konsonantsystem. In: *18. møde om udforskningen af dansk sprog*. Ed. by Yonatan Goldshtein, Inger Schoonderbeek Hansen, and Tina Thode Hougaard. Aarhus University, 251–267. HDL: 1887/3210583.

Horslund, Camilla Søballe, Rasmus Puggaard, and Henrik Jørgensen (2020). En alternativ, fonetisk baseret fonemanalyse af det danske konsonantsystem. (Paper presented at *Møde om Udforskningen af det Danske Sprog* 18, Aarhus University).

Horslund, Camilla Søballe, Rasmus Puggaard, and Henrik Jørgensen (2021b). An alternative, phonetically based phoneme analysis of the Danish consonant system. (Paper presented at *International Conference of Nordic and General Linguistics* 12, University of Oslo).

Horslund, Camilla Søballe, Rasmus Puggaard-Rode, and Henrik Jørgensen (2022). A phonetically-based phoneme analysis of the Danish consonant system. *Acta Linguistica Hafniensia* 54 (1), 73–105. DOI: 10.1080/03740463.2021.2022866.

Horvath, Barbara M. and Ronald J. Horvath (2001). A multilocality study of a sound change in progress. The case of /l/ vocalization in New Zealand and Australian English. *Language Variation and Change* 13 (1), 37–57. DOI: 10.1017/S0954394501131029.

House, Arthur S. and Grant Fairbanks (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *JASA* 25 (1), 105–113. DOI: 10.1121/1.1906982.

Hovdhaugen, Even, Fred Karlsson, Carol Henriksen, and Bengt Sigurd (2000). *The history of linguistics in the Nordic countries*. Societas Scientarum Fennica.

Hughes, George W. and Morris Halle (1956). Spectral properties of fricative consonants. *JASA* 28 (2), 303–310. DOI: 10.1121/1.1908271.

Hume, Elizabeth V. (1992). *Front vowels, coronal consonants and their interaction in nonlinear phonology*. PhD dissertation, Cornell University. Published 1994 by Routledge (Outstanding Dissertations in Linguistics). DOI: 10 . 4324 / 9781315049366.

Hutters, Birgit (1978). The glottal gesture in some Danish consonants. Preliminary observations. *ARIPUC* 12, 87–101.

Hutters, Birgit (1984). Vocal fold adjustments in Danish voiceless obstruent production. *ARIPUC* 18, 293–385.

Hutters, Birgit (1985). Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica* 42 (1), 1–24. DOI: 10.1159/000261734.

Hyman, Larry M. (1976). Phonologization. In: *Linguistics studies offered to Joseph Greenberg on the occasion of his 60th birthday*. Ed. by Alphonse Juilland, A.M. Devine, and Laurence D. Stephens. Anma Libri (Studia linguistica et philologica 4), 407–418.

Hyttel-Sørensen, Liva (2017). *"Gangster" or "wannabe". Experimental and ethnographic approaches to a contemporary urban vernacular in Copenhagen*. PhD dissertation, University of Copenhagen.

Inkelas, Sharon and Stephanie S. Shih (2017). Looking into segments. In: *Proceedings of the 2016 Annual Meeting on Phonology*. Ed. by Karen Jesney, Charlie O'Hara, Caitlin Smith, and Rachel Walker. Linguistics Society of America. DOI: 10.3765/amp.v4i0.3996.

Iosad, Pavel (2017). *A substance-free framework for phonology. An analysis of the Breton dialect of Bothoa*. Edinburgh University Press (Edinburgh Studies in Theoretical Linguistics 2). DOI: 10.1515/9781474407380.

IPA = International Phonetic Association (1999). *Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet*. CUP.

Iskarous, Khalil and Darya Kavitskaya (2018). Sound change and the structure of synchronic variability. Phonetic and phonological factors in Slavic palatalization. *Language* 94 (1), 43–83. DOI: 10.1353/lan.2018.0001.

Iverson, Gregory K. and Joseph C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* 12 (3), 369–396. DOI: 10.1017/S0952675700002566.

Iverson, Gregory K. and Joseph C. Salmons (2003a). Laryngeal enhancement in early Germanic. *Phonology* 20 (1), 43–74. DOI: 10.1017/S0952675703004469.

Iverson, Gregory K. and Joseph C. Salmons (2003b). Legacy specification in the laryngeal phonology of Dutch. *Journal of Germanic Linguistics* 15 (1), 1–26. DOI: 10.1017/S1470542703000242.

Iverson, Gregory K. and Joseph C. Salmons (2006). On the typology of final laryngeal neutralization. Evolutionary Phonology and laryngeal realism. *Theoretical Linguistics* 32 (2), 205–216. DOI: 10.1515/TL.2006.014.

Iverson, Gregory K. and Joseph C. Salmons (2011). Final devoicing and final laryngeal neutralization. In: *The Blackwell companion to phonology. Volume III: Phonological processes*. Ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice. Wiley–Blackwell (Blackwell Companions to Linguistics), 1622–1643. DOI: 10.1002/9781444335262.wbctp0069.

Jacobsen, Henrik Galberg (2019). Bøjning: Yngre nydansk. In: *Dansk sproghistorie 3. Bøjning og bygning*. Ed. by Ebba Hjorth. Aarhus Universitetsforlag, 93–111.
Jaeger, Jeri J. (1983). The fortis/lenis question. Evidence from Zapotec and Jawoñ. *Journal of Phonetics* 11 (2), 177–189. DOI: 10.1016/S0095-4470(19)30814-9.
Jaeger, Jeri J. (1984). Assessing the psychological status of the Vowel Shift Rule. *Journal of Psycholinguistic Research* 13 (1), 13–36. DOI: 10.1007/BF01067900.
Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle (1951). *Preliminaries to speech analysis. The distinctive features and their correlates*. MIT Press.
Jannedy, Stefanie and Melanie Weirich (2017). Spectral moments vs discrete cosine transformation coefficients. Evaluation of acoustic measures distinguishing two merging German fricatives. *JASA* 142 (1), 395–405. DOI: 10.1121/1.4991347.
Jeel, Vivi (1975). An investigation of the fundamental frequency of vowels after various consonants, in particular stop consonants. *ARIPUC* 9, 191–211.
Jensen, Ella (1944). *Houlbjergmålet. Bidrag til Beskrivelse af en østjysk Dialekt*. Schultz (Udvalg for Folkemaals Publikationer A6).
Jensen, J.M. (1897–1902). *Et vendelbomåls lyd- og formlære*. Thieles (Universitets-jubilæets danske samfund).
Jespersen, Otto (1890). Danias lydskrift. *Dania. Tidsskrift for Folkemål og Folkeminder* 1, 33–80.
Jespersen, Otto (1897–1899). *Fonetik. En systematisk fremstilling af læren om sproglyd*. Schubotheske. German translation published 1904 in two volumes, *Lehrbuch der Phonetik* and *Phonetische Grundfragen*. B.G. Teubner.
Jespersen, Otto (1906). *Modersmålets fonetik*. Schubotheske.
Jespersen, Otto (1924). *The philosophy of grammar*. George Allen and Unwin Ltd. DOI: 10.4324/9780203716045.
Jessen, Michael (1999). *Phonetics and phonology of tense and lax obstruents in German*. John Benjamins (Studies in Functional and Structural Linguistics 44). DOI: 10.1075/sfsl.44.
Jessen, Michael (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In: *Distinctive feature theory*. Ed. by Tracy Alan Hall. Mouton de Gruyter (Phonology and Phonetics 2), 237–294. DOI: 10.1515/9783110886672.237.
Jessen, Michael and Catherine Ringen (2002). Laryngeal features in German. *Phonology* 19 (2), 189–218. DOI: 10.1017/S0952675702004311.
Jongman, Allard, Ratree Wayland, and Serena Wong (2000). Acoustic characteristics of English fricatives. *JASA* 108 (3), 1252–1263. DOI: 10.1121/1.1288413.
Jurafsky, Daniel, Alan Bell, and Cynthia Girand (2002). The role of the lemma in form variation. In: *Laboratory Phonology 7*. Ed. by Carlos Gussenhoven and Natasha Warner. Mouton de Gruyter (Phonology and Phonetics 4–1), 3–34. DOI: 10.1515/9783110197105.1.3.
Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond (2001). Probabilistic relations between words. Evidence from reduction in lexical production. In: *Frequency and the emergence of linguistic structure*. Ed. by Joan L. Bybee and

Paul J. Hopper. John Benjamins (Typological Studies in Language 45), 229–254. DOI: 10.1075/tsl.45.13jur.

Jusczyk, Peter W. and Richard N. Aslin (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 29 (1), 1–23. DOI: 10.1006/cogp.1995.1010.

Juul, Holger, Nicolai Pharao, and Jacob Thøgersen (2016). Moderne danske vokaler. *Danske Talesprog* 16, 35–72.

Juul, Holger, Nicolai Pharao, and Jacob Thøgersen (2019). Stemthed og afstemthed i danske sonoranter. *Danske Talesprog* 19, 108–130.

Jørgensen, Henrik (2021). Konsekvens i renæssancens retskrivning? In: *Nogle betænkninger om dansk sprog og litteratur. Festskrift til Marita Akhøj Nielsen.* Ed. by Jens Bjerring-Hansen, Simon Skovgaard Boeck, and Eva Skafte Jensen. University Press of Southern Denmark (Universitets-jubilæets danske samfund), 273–294.

Kaplan, Abby (2010). *Phonology shaped by phonetics. The case of intervocalic lenition.* PhD dissertation, University of California, Santa Cruz. DOI: 10.7282/T30G3J2K.

Karsten, Gustaf E. (1894). The psychological basis of phonetic law and analogy. *Publications of the Modern Language Association* 9 (2), 312–341. DOI: 10.2307/456363.

Katz, Jonah (2016). Lenition, perception and neutralisation. *Phonology* 33 (1), 43–85. DOI: 10.1017/S0952675716000038.

Keating, Patricia A. (1984a). Phonetic and phonological representation of stop consonant voicing. *Language* 60 (2), 286–319. DOI: 10.2307/413642.

Keating, Patricia A. (1984b). Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics* 59, 18–28.

Keating, Patricia A., Taehong Cho, Cécile Fougeron, and Chai-Shune Hsu (2004). Domain-initial articulatory strengthening in four languages. In: *Phonetic interpretation.* Ed. by John Local, Richard Ogden, and Rosalind Temple. CUP (Papers in Laboratory Phonology 4), 145–163. DOI: 10.1017/CBO9780511486425.009.

Keating, Patricia A., Wendy Linker, and Marie K. Huffman (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11 (3), 277–290. DOI: 10.1016/S0095-4470(19)30827-7.

Kehrein, Wolfgang and Chris Golston (2004). A prosodic theory of laryngeal contrasts. *Phonology* 21 (3), 325–357. DOI: 10.1017/S0952675704000302.

Kenstowicz, Michael (1994). *Phonology in generative grammar.* Blackwell (Blackwell Textbooks in Linguistics).

Kewley-Port, Diane (1982). Measurements of formant transitions in naturally produced stop consonant–vowel syllables. *JASA* 72 (2), 379–389. DOI: 10.1121/1.388081.

Kewley-Port, Diane (1983). Time-varying features as correlates of place of articulation in stop consonants. *JASA* 73 (1), 322–335. DOI: 10.1121/1.388813.

Kewley-Port, Diane, David B. Pisoni, and Michael Studdert-Kennedy (1983). Perception of static and dynamic cues to place of articulation in initial stop consonants. *JASA* 73 (5), 1779–1793. DOI: 10.1121/1.388813.

Kijak, Artur (2017). *Labial–dorsal interactions. A phonologically based approach.* Wydawnictwo Uniwersytetu Śląskiego (Językoznawstwo Neofilologiczne). HDL: 20.500.12128/11067.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel (2014). The Sketch Engine. Ten years on. *Lexicography Asialex* 1, 7–36. DOI: 10.1007/s40607-014-0009-9.

Kim, Hyunsoon (2001). A phonetically based account of phonological stop assibilation. *Phonology* 18 (1), 81–108. DOI: 10.1017/S095267570100402X.

Kingston, John and Randy L. Diehl (1994). Phonetic knowledge. *Language* 70 (3), 419–454. DOI: 10.1353/lan.1994.0023.

Kiparsky, Paul (1968/1982). How abstract is phonology? In: *Explanation in phonology.* Foris (Publications in Language Sciences 4), 119–164. DOI: 10.1515/9783111666242.119.

Kiparsky, Paul (1985). Some consequences of Lexical Phonology. *Phonology Yearbook* 2 (1), 85–138. DOI: 10.1017/S0952675700000397.

Kirby, James P. and D. Robert Ladd (2016). Effects of obstruent voicing on vowel *F*0. Evidence from "true voicing" languages. *JASA* 140 (4), 2400–2411. DOI: 10.1121/1.4962445.

Kirby, James P. and D. Robert Ladd (2018). Effects of obstruent voicing on vowel F0. Implications for laryngeal realism. *Yearbook of the Poznań Linguistic Meeting* 4, 213–235. DOI: 10.2478/yplm-2018-0009.

Kirchmeier-Andersen, Sabine (2007). Er dansk truet? In: *Ved lejlighed. Festskrift til Niels Davidsen-Nielsen i anledning af 70-års-dagen.* Ed. by Margrethe Heidemann Andersen, Jørgen Nørby Jensen, Marianne Rathje, and Jørgen Schack. Dansk Sprognævn (Dansk Sprognævns skrifter 39), 109–121.

Klein, W., Reinier Plomp, and Louis C.W. Pols (1970). Vowel spectra, vowel spaces, and vowel identification. *JASA* 48 (4/2), 999–1009. DOI: 10.1121/1.1912239.

Kluge, Friedrich (1884). Die germanische Consonantendehnung. *Beiträge zur Geschichte der deutschen Sprache und Litteratur* 9 (2), 149–186. DOI: 10.1515/bgsl.1884.9.2.149.

Knief, Ulrich and Wolfgang Forstmeier (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods* 53 (6), 2576–2590. DOI: 10.3758/s13428-021-01587-5.

Koenig, Laura L., Christine H. Shadle, Jonathan L. Preston, and Christine R. Mooshammer (2013). Toward improved spectral measures of /s/. Results from adolescents. *Journal of Speech, Language, and Hearing Research* 56 (4), 1175–1189. DOI: 10.1044/1092-4388(2012/12-0038).

Kohler, Klaus J. (1984). Phonetic explanation in phonology. The feature fortis/lenis. *Phonetica* 41 (3), 150–174. DOI: 10.1159/000261721.

Komsta, Lukasz and Frederick Novomestky (2015). `moments`. Moments, cumulants, skewness, kurtosis and related tests. (R package version 0.14).

Kopp, G.A. and H.C. Green (1946). Basic phonetic principles of visible speech. *JASA* 18 (1), 74–89. DOI: 10.1121/1.1916345.

Kortlandt, Frederik (1985). Proto-Indo-European glottalic stops. The comparative evidence. *Folia Linguistica Historica* 6 (2), 183–201. DOI: 10.1515/flih.1985.6.2.183.

Kraehenmann, Astrid (2001). Swiss German stops. Geminates all over the word. *Phonology* 18 (1), 109–145. DOI: 10.1017/S0952675701004031.

Kristensen, Kjeld (1977). Variationen i vestjysk stationsmål. *Dialektstudier* 4 (1), 29–109.

Kristiansen, Tore (1990). *Udtalenormering i skolen. Skitse af en ideologisk bastion.* Gyldendal.

Kristiansen, Tore (1998). The role of standard ideology in the disappearance of the traditional Danish dialects. *Folia Linguistica* 32 (1/2), 115–129. DOI: 10.1515/flin.1998.32.1-2.115.

Kristiansen, Tore (2003a). Danish. In: *Germanic standardizations. Past to present.* Ed. by Ana Deumart and Wim Vandenbussche. John Benjamins (Impact. Studies in Language and Society 18), 69–91. DOI: 10.1075/impact.18.04kri.

Kristiansen, Tore (2003b). Language attitudes and language politics in Denmark. *International Journal of the Sociology of Language* 159, 57–71. DOI: 10.1515/ijsl.2003.009.

Krämer, Martin and Draga Zec (2020). Nasal consonants, sonority and syllable phonotactics. The dual nasal hypothesis. *Phonology* 37 (1), 27–63. DOI: 10.1017/S0952675720000032.

Kühberger, Anton, Astrid Fritz, Eva Lermer, and Thomas Scherndl (2015). The significance fallacy in inferential statistics. *BioMed Central Research Notes* 8 (84). DOI: 10.1186/s13104-015-1020-4.

Labov, William (1963). The social motivation of a sound change. *Word* 19 (3), 273–309. DOI: 10.1080/00437956.1963.11659799.

Labov, William (2003). Pursuing the cascade model. In: *Social dialectology. In honour of Peter Trudgill.* Ed. by David Britain and Jenny Cheshire. John Benjamins (Impact. Studies in Language and Society 16), 9–22. DOI: 10.1075/impact.16.03lab.

Ladd, D. Robert (2011). Phonetics in phonology. In: *The handbook of phonological theory.* Ed. by John Goldsmith, Jason Riggle, and Alan C.L. Yu. 2nd ed. Wiley–Blackwell (Blackwell Handbooks in Linguistics), 348–373. DOI: 10.1002/9781444343069.ch11.

Ladd, D. Robert and Stephan Schmid (2018). Obstruent voicing effects on F0, but without voicing. Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics* 71, 229–248. DOI: 10.1016/j.wocn.2018.09.003.

Ladefoged, Peter (1973). The features of the larynx. *Journal of Phonetics* 1 (1), 73–83. DOI: 10.1016/S0095-4470(19)31376-2.

Ladefoged, Peter (1975). *A course in phonetics.* Harcourt Brace Jovanovich.

Ladefoged, Peter (1990). What do we symbolize? Thoughts prompted by bilabial and labiodental fricatives. *Journal of the International Phonetic Association* 20 (2), 32–36. DOI: 10.1017/S0025100300004254.

Ladefoged, Peter and Ian Maddieson (1996). *The sounds of the world's languages.* Blackwell.

Laker, Stephen (2014). The downfall of dental fricatives. Frisian perspectives on a wider Germanic trend. *Amsterdamer Beiträge zur Älteren Germanistik* 73, 261–300. DOI: 10.1163/9789401211918_011.

Lawson, Eleanor, James M. Scobbie, and Jane Stuart-Smith (2013). Bunched /r/ promotes vowel merger to schwar. An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics* 41 (3–4), 198–210. DOI: 10.1016/j.wocn.2013.01.004.

Liberman, Anatoly (2006). Epenthetic consonants and the accentuation of words with old closed vowels in Low German, Dutch, and Danish dialects. In: *Germanic tone accents*. Ed. by Michiel de Vaan. Franz Steiner (Zeitschrift für Dialektologie und Lingvistik 131), 73–83.

Ligges, Uwe (2021). `tuneR`. Analysis of music and speech. (R package version 1.3.3.1).

Lillelund-Holst, Aleksandra Culap and Nicolai Pharao (2014). Sociale betydninger af [t^j] og [s+] blandt unge i København. *Danske Talesprog* 14, 94–123.

Lillelund-Holst, Aleksandra Culap, Nicolai Pharao, and Marie Maegaard (2019). "Not from a 'Danish' home; typical of trying to sound 'tough'". Indexical meanings of variation in /s/ and /t/ in the speech of adolescent girls in Copenhagen. *Revista de Estudos da Linguagem* 27 (4), 1701–1736. DOI: 10.17851/2237-2083.27.4.1701-1736.

Lin, Yen-Hwei (2011). Affricates. In: *The Blackwell companion to phonology. Volume I: General issues and segmental phonology*. Ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice. Wiley–Blackwell (Blackwell Companions to Linguistics), 367–390. DOI: 10.1002/9781444335262.wbctp0016.

Lindblom, Björn (1986). Phonetic universals in vowel systems. In: *Experimental phonology*. Ed. by John J. Ohala and Jeri J. Jaeger. Academic Press, 13–44.

Lindblom, Björn (1990). Explaining phonetic variation. A sketch of the H&H theory. In: *Speech production and speech modelling*. Ed. by William J. Hardcastle and Alain Marchal. Kluwer (Nato Science Series D 55), 403–439. DOI: 10.1007/978-94-009-2037-8_16.

Lindsey, Geoff (2019). *English after RP. Standard British pronunciation today*. Palgrave Macmillan. DOI: 10.1007/978-3-030-04357-5.

Linell, Per (1975). *Psychological reality in phonology*. CUP (Cambridge Studies in Linguistics 25).

Lisker, Leigh (1957). Closure duration and the intervocalic voiced–voiceless distinction in English. *Language* 33 (1), 42–49. DOI: 10.2307/410949.

Lisker, Leigh and Arthur S. Abramson (1964). A cross-language study of voicing in initial stops. Acoustical measurements. *Word* 20 (3), 384–422. DOI: 10.1080/00437956.1964.11659830.

Lisker, Leigh and Arthur S. Abramson (1967). Some effects of context on voice onset time in English stops. *Language and Speech* 10 (1), 10–28. DOI: 10.1177/002383096701000101.

Liu, Jiang and Allard Jongman (2013). American Chinese learners' acquisition of L2 Chinese affricates /ts/ and /tsh/. *Proceedings of Meetings on Acoustics* 18. DOI: 10.1121/1.4798223.

Lodge, Ken (2009). *A critical introduction to phonetics.* Continuum (Continuum Critical Introductions to Linguistics).

Lombardi, Linda (1990). The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* 8 (3), 375–425. DOI: 10.1007/BF00135619.

Lombardi, Linda (1991). *Laryngeal features and laryngeal neutralization.* PhD dissertation, University of Massachusetts. PROQUEST: AAI9132881. Published 1994 by Routledge. DOI: 10.4324/9780429454929.

Lombardi, Linda (1995). Laryngeal features and privativity. *The Linguistic Review* 12 (1), 35–59. DOI: 10.1515/tlir.1995.12.1.35.

Lombardi, Linda (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17 (2), 267–302. DOI: 10.1023/A:1006182130229.

Luce, Paul A. and Jan Charles-Luce (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *JASA* 78 (6), 1949–1957. DOI: 10.1121/1.392651.

Lüdecke, Daniel (2021). `sjPlot`. Data visualization for statistics in social science. (R package version 2.8.10).

Lühr, Rosemarie (1988). *Expressivität und Lausgesetz im Germanischen.* Winter (Monographien zur Sprachwissenschaft 15).

Lyngby, K.J. (1858). *Bidrag til en sønderjysk sproglære. Sprogartens forhold til de øvrige dansk, dens forgreninger, sproglære for Braderup (Kær herred).* Thieles.

Löfqvist, Anders (1975a). A study of subglottal pressure during the production of Swedish stops. *Journal of Phonetics* 3 (3), 175–189. DOI: 10.1016/S0095-4470(19)31366-X.

Löfqvist, Anders (1975b). On the control of aspiration in Swedish. *Phonetics Laboratory/Department of General Linguistics, Lund University Working Papers* 12, 99–106.

Löfqvist, Anders (1976). Closure duration and aspiration for Swedish stops. *Phonetics Laboratory/Department of General Linguistics, Lund University Working Papers* 13, 1–40.

Löfqvist, Anders (1980). Interarticulator programming in stop production. *Journal of Phonetics* 8 (4), 475–490. DOI: 10.1016/S0095-4470(19)31502-5.

Maddieson, Ian (1984). *Patterns of sounds.* CUP (Cambridge Studies in Speech Science and Communication). DOI: 10.1017/CBO9780511753459.

Maddieson, Ian and Karen Emmorey (1984). Is there a valid distinction between voiceless lateral approximants and fricatives? *Phonetica* 41 (4), 181–190. DOI: 10.1159/000261725.

Madsen, Lian Malai (2008). *Fighters and outsiders. Linguistic practices, social identities and social relationships among youth in a martial arts club.* PhD dissertation, University of Copenhagen.

Madsen, Lian Malai (2013). "High" and "low" in urban Danish speech styles. *Language in Society* 42 (2), 115–138. DOI: 10.1017/S0047404513000018.

Maegaard, Marie (2007). *Udtalevariation og -forandring i københavnsk. En etnografisk undersøgelse af sprogbrug, sociale kategorier og social praksis blandt unge på en københavnsk folkeskole.* University of Copenhagen (Danske Talesprog 8).

Maegaard, Marie and Malene Monka (2019). Patterns of dialect use. Language standardization at different rates. In: *Standardization as sociolinguistic change. A transversal study of three traditional dialect areas.* Ed. by Marie Maegaard, Malene Monka, Kristine Køhler Mortensen, and Andreas Candefors Stæhr. Routledge (Routledge Studies in Language Change), 27–46. DOI: 10.4324/9780429467486-2.

ManyBabies Consortium, The (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science* 3 (1), 24–52. DOI: 10.1177/2515245919900809.

Marra, Giampiero and Simon N. Wood (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics* 39 (1), 53–74. DOI: 10.1111/j.1467-9469.2011.00760.x.

Marron, J.S., James O. Ramsay, Laura M. Sangalli, and Anuj Srivastava (2015). Functional data analysis of amplitude and phase variation. *Statistical Science* 30 (4), 468–484. DOI: 10.1214/15-STS524.

Martinet, André (1937). *La phonologie du mot en danois.* C. Klincksiek.

Martínez-Celdrán, Eugenio (2004). Problems in the classification of approximants. *Journal of the International Phonetic Association* 34 (2), 201–210. DOI: 10.1017/S0025100304001732.

Matthiessen, Christian Wichmann (1985). *Danske byers folketal 1801-1981.* Danmarks Statistik (Statistiske Undersøgelser 42).

Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, R. Harald Baayen, and Douglas Bates (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94, 305–315. DOI: 10.1016/j.jml.2017.01.001.

Mazzaro, Natalia (2010). Changing perceptions. The sociophonetic motivations of the labial velar alternation in Spanish. In: *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology.* Ed. by Marta Ortega-Llebaria. Cascadilla Proceedings Project, 128–145.

Meteyard, Lottie and Robert A.I. Davies (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language* 112. DOI: 10.1016/j.jml.2020.104092.

Mielke, Jeff (2008). *The emergence of distinctive features.* OUP. (Oxford Studies in Typology and Linguistic Theory).

Mielke, Jeff (2015). An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *JASA* 137 (5), 2858–2869. DOI: 10.1121/1.4919346.

Monka, Malene (2019). Southern Jutland. Language ideology as a means to slow down dialect leveling. In: *Standardization as sociolinguistic change. A transversal study of three traditional dialect areas.* Ed. by Marie Maegaard, Malene Monka, Kristine Køhler Mortensen, and Andreas Candefors Stæhr. Routledge (Routledge Studies in Language Change), 47–89. DOI: 10.4324/9780429467486-3.

Monka, Malene and Henrik Hovmark (2016). Sprogbrug blandt unge i Bylderup anno 2015. *Danske Talesprog* 16, 73–114.

Mortensen, David R. (2012). The emergence of obstruents after high vowels. *Diachronica* 29 (4), 434–470. DOI: 10.1075/dia.29.4.02mor.

Mortensen, Johannes and John Tøndering (2013). The effect of vowel height on voice onset time in stop consonants in CV sequences in spontaneous Danish. In: *Proceedings of Fonetik 2013. The XXVIth Annual Phonetics Meeting*. Linköping University, 49–52.

Mortensen, Kristine Køhler (2019). Northern Jutland. Local positioning and global orientation. In: *Standardization as sociolinguistic change. A transversal study of three traditional dialect areas*. Ed. by Marie Maegaard, Malene Monka, Kristine Køhler Mortensen, and Andreas Candefors Stæhr. Routledge (Routledge Studies in Language Change), 80–113. DOI: 10.4324/9780429467486-4.

Moulton, William G. (1954). The stops and spirants of early Germanic. *Language* 30 (1), 1–42. DOI: 10.2307/410217.

Mücke, Doris, Martine Grice, and Taehong Cho (2014). More than a magic moment. Paving the way for dynamics of articulation and prosodic structure. *Journal of Phonetics* 44, 1–7. DOI: 10.1016/j.wocn.2014.03.001.

Möbius, Bernd (2004). Corpus-based investigations on the phonetics of consonant voicing. *Folia Linguistica* 38 (1/2), 5–26. DOI: 10.1515/flin.2004.38.1-2.5.

Møller, Janus Spindler (2009). Stereotyping categorisations of speech styles among linguistic minority Danes in Køge. In: *Language attitudes, standardization and language change*. Ed. by Marie Maegaard, Frans Gregersen, Pia Quist, and Jens Normann Jørgensen. Nuvus, 231–254.

Nakagawa, Shinichi, Paul C.D. Johnson, and Holger Schielzeth (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* 34 (134). DOI: 10.1098/rsif.2017.0213.

Nance, Claire and Sam Kirkham (2020). The acoustics of three-way lateral and nasal palatalisation contrasts in Scottish Gaelic. *JASA* 147 (4), 2858–2872. DOI: 10.1121/10.0000998.

Nerbonne, John (2009). Data-driven dialectology. *Language and Linguistics Compass* 3 (1), 175–198. DOI: 10.1111/j.1749-818x.2008.00114.x.

Nerbonne, John and Wilbert Heeringa (2007). Geographic distributions of linguistic variation reflect dynamics of differentiation. In: *Roots. Linguistics in search of its evidential base*. Ed. by Sam Featherston and Wolfgang Sternefeld. Mouton de Gruyter (Studies in Generative Grammar 96), 267–297. DOI: 10 . 1515 / 9783110198621.

Nespor, Marina and Irene Vogel (1986). *Prosodic phonology*. Foris (Studies in Generative Grammar 28). DOI: 10.1515/9783110977790.

Nicenboim, Bruno, Timo B. Roettger, and Shravan Vasishth (2018). Using meta-analysis for evidence synthesis. The case of incomplete neutralization in German. *Journal of Phonetics* 70, 39–55. DOI: 10.1016/j.wocn.2018.06.001.

Nicenboim, Bruno and Shravan Vasishth (2016). Statistical methods for linguistic research. Foundational ideas – Part II. *Language and Linguistics Compass* 10 (11), 591–613. DOI: 10.1111/lnc3.12207.

Nielsen, Bent Jul (1978). On the phonology of stød and parasite consonant in long vowels in West Jutland dialects. In: *Nordic prosody*. Ed. by Eva Gårding, Gösta Bruce, and Robert Bannert. Lund University (Travaux de l'Institut de Linguistique de Lund 13), 153–160.
Nielsen, Bent Jul (1984). *Bidrag til vendelbomålets fonologi. Del I: Dialekten i Tornby*. Akademisk Forlag.
Nielsen, Gunhild (1947). Glidning og lukning i Rømømaalets høje Vokaler ('Klusilspring'). *Danske Folkemaal* 15, 41–65.
Nielsen, Michael Lerche and Marie Stoklund (2018). Udtale: Runetiden. In: *Dansk sproghistorie 2. Ord for ord for ord*. Ed. by Ebba Hjorth. Aarhus Universitetsforlag, 134–150.
Ohala, John J. (1974). Experimental historical phonology. In: *Historical linguistics, vol. 2. Theory and description in phonology*. Ed. by John M. Anderson and Charles Jones. North-Holland (North-Holland Linguistic Series 12b), 353–389.
Ohala, John J. (1981). The listener as a source of sound change. In: *Papers from the 17th regional meeting of the Chicago Linguistic Society. Volume II: Papers from the parasession on language and behavior*. Ed. by Robert A. Hendrick, Carrie S. Masek, and Mary Frances Miller. Chicago Linguistic Society, 178–203.
Ohala, John J. (1983a). The origin of sound patterns in vocal tract constraints. In: *The production of speech*. Ed. by Peter F. MacNeilage. Springer, 189–216. DOI: 10.1007/978-1-4613-8202-7_9.
Ohala, John J. (1983b). The phonological end justifies any means. In: *ICPhS 13*. Ed. by Shirō Hattori and Kazuko Inoue. Sanseido Shoten, 232–243.
Ohala, John J. (1986). Consumer's guide to evidence in phonology. *Phonology Yearbook* 3, 3–26. DOI: 10.1017/S0952675700000555.
Ohala, John J. (1989). Sound change is drawn from a pool of synchronic variation. In: *Language change. Contributions to the study of its causes*. Ed. by Leiv E. Breivik and Ernst Håkon Jahr. Mouton de Gruyter (Trends in Linguistics: Studies and Monographs 43), 173–198. DOI: 10.1515/9783110853063.173.
Ohala, John J. (1990a). The phonetics and phonology of aspects of assimilation. In: *Between the grammar and physics of speech*. Ed. by John Kingston and Mary E. Beckman. CUP (Papers in Laboratory Phonology 1), 258–275. DOI: 10.1017/CBO9780511627736.014.
Ohala, John J. (1990b). There is no interface between phonology and phonetics. A personal view. *Journal of Phonetics* 18 (2), 153–171. DOI: 10.1016/S0095-4470(19)30399-7.
Ohala, John J. (1992). What's cognitive, what's not, in sound change. In: *Diachrony within synchrony. Language history and cognition*. Ed. by Günter Kellerman and Michael D. Morrissey. Peter Lang (Duisburg Papers on Research in Language and Culture 14), 309–355.
Ohala, John J. (1993). The phonetics of sound change. In: *Historical linguistics. Problems and perspectives*. Ed. by Charles Jones. Longman, 237–278. DOI: 10.4324/9781315845814.

Ohala, John J. (1996). Speech perception is hearing sounds, not tongues. *JASA* 99 (3), 1718–1725. DOI: 10.1121/1.414696.

Ohala, John J. (2005). Phonetic explanations for sound patterns. Implications for grammars of competence. In: *A figure of speech. A festschrift for John Laver*. Ed. by William J. Hardcastle and Janet Mackenzie Beck. Lawrence Erlbaum, 23–38. DOI: 10.4324/9781410611888.

Ohala, John J. and Manjari Ohala (1986). Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In: *Experimental phonology*. Ed. by John J. Ohala and Jeri J. Jaeger. Academic Press, 239–252.

Ohala, John J. and Carol J. Riordan (1979). Passive vocal tract enlargement during voiced stops. In: *Speech communication papers presented at the 97th Meeting of the Acoustical Society of America*. Ed. by Jared J. Wolf and Dennis H. Klatt. Massachusetts Institute of Technology, 89–92.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349 (6251). DOI: 10.1126/science.aac4716.

Ouni, Slim (2014). Tongue control and its implication in pronunciation training. *Computer Assisted Language Learning* 27 (5), 439–453. DOI: 10.1080/09588221.2012.761637.

Page, B. Richard (1997). On the origin of preaspiration in Scandinavian. *American Journal of Germanic Linguistics and Literatures* 9 (2), 167–190. DOI: 10.1017/S1040820700002857.

Pape, Daniel and Luis M.T. Jesus (2014). Production and perception of velar stop (de)voicing in European Portuguese and Italian. *EURASIP Journal on Audio, Speech, and Music Processing*. DOI: 10.1186/1687-4722-2014-6.

Paradis, Carole and Jean-François Prunet (1989). On coronal transparency. *Phonology* 6 (2), 317–348. DOI: 10.1017/S0952675700001056.

Paris, Gaston (1888). Les parlers de France. *Revue des patois gallo-romans* 2, 161–175.

Parker, Steve (2002). *Quantifying the sonority hierarchy*. PhD dissertation, University of Massachusetts, Amherst. DOI: 10.13140/2.1.3546.4005.

Parveen, Sabiha and Alexander M. Goberman (2012). Presence of stop bursts and multiple bursts in younger and older adults. *Asia Pacific Journal of Speech, Language and Hearing* 15 (4), 265–275. DOI: 10.1179/136132812804731811.

Pedersen, Inge Lise (2003). Traditional dialects of Danish and the de-dialectalization 1900–2000. *International Journal of the Sociology of Language* 159, 9–28. DOI: 10.1515/ijsl.2003.012.

Pedersen, Karen Margrethe (1983). Om transskription af båndoptagelser og anvendelse af teksterne til leksikografiske og syntaktiske undersøgelser. Et EDB-projekt. *Folkmålsstudier* 28, 131–138.

Perridon, Harry (2008). Reconstructing the obstruents of Proto-Germanic. In: *Evidence and counter-evidence. Essays in honour of Frederik Kortlandt. Volume 1: Balto-Slavic and Indo-European linguistics*. Ed. by Alexander Lubotsky, Jos Schaeken, and Jeroen Wiedenhof. Rodopi (Studies in Slavic and General Linguistics 32), 415–429. DOI: 10.1163/9789401206358_031.

Petersen, Jan Heegård, Holger Juul, Nicolai Pharao, and Marie Maegaard (2021). *Udtalt. En introduktionsbog til dansk fonetik.* Samfundslitteratur.
Petersen, Niels Reinholt (1978). The influence of aspiration on the fundamental frequency of the following vowel in Danish. Some preliminary observations. *ARIPUC* 12, 91–112.
Petersen, Niels Reinholt (1983). The effect of consonant type on fundamental frequency and larynx height in Danish. *ARIPUC* 17, 55–86.
Petersen, Niels Reinholt (2009a). *Affrikater.* In: *Den store danske.* Gyldendal.
Petersen, Niels Reinholt (2009b). *Aspirater.* In: *Den store danske.* Gyldendal.
Pétursson, Magnús (1976). Aspiration et activité glottale. Examen expérimental à partir de consonnes islandaises. *Phonetica* 33 (3), 169–198. DOI: 10 . 1159 / 000259721.
Pharao, Nicolai (2004). *Inde bag ved ordene. Nogle fonologiske repræsentationer og processers psykologiske realitet for danske talere.* Master's thesis, University of Copenhagen.
Pharao, Nicolai (2009). *Consonant reduction in Copenhagen Danish. A study of linguistic and extra-linguistic factors in phonetic variation and change.* PhD dissertation, University of Copenhagen.
Pharao, Nicolai (2011). Plosive reduction at the group level and in the individual speaker. In: *ICPhS 17*, 1590–1593.
Pharao, Nicolai (2012). Lukkelydsreduktioner i dansk. *Danske Talesprog* 12, 111–126.
Pharao, Nicolai and Gert Foget Hansen (2005). Prosodiske træk i et- og tosprogede unges københavnsk. *Danske Talesprog* 6, 1–50.
Pharao, Nicolai and Marie Maegaard (2017). On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics* 55 (5), 1141–1167. DOI: 10.1515/ling-2017-0023.
Pharao, Nicolai, Marie Maegaard, Janus Spindler Møller, and Tore Kristiansen (2014). Indexical meanings of [s+] among Copenhagen youth. Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43 (1), 1–31. DOI: 10.1017/S0047404513000857.
Pharao, Nicolai, Mia Ridder Malmstedt, and Selina Veng (2017). Hvor svært er det lige? En undersøgelse af genkendelsen af ord med reducerede segmenter. *Nydanske Sprogstudier* 52/53, 259–288. DOI: 10.7146/nys.v1i52-53.26359.
Pierrehumbert, Janet B. (2001). Exemplar dynamics. Word frequency, lenition and contrast. In: *Frequency and the emergence of linguistic structure.* Ed. by Joan L. Bybee and Paul J. Hopper. John Benjamins (Typological Studies in Language 45), 137–157. DOI: 10.1075/tsl.45.08pie.
Pierrehumbert, Janet B. (2002). Word-specific phonetics. In: *Laboratory Phonology 7.* Ed. by Carlos Gussenhoven and Natasha Warner. Mouton de Gruyter (Phonology and Phonetics 4-1), 101–139. DOI: 10.1515/9783110197105.1.101.
Pierrehumbert, Janet B. (2006). The statistical basis of an unnatural alternation. In: *Laboratory Phonology 8.* Ed. by Louis M. Goldstein, Douglas H. Whalen, and Catherine T. Best. Mouton de Gruyter (Phonology and Phonetics 4/2), 83–108. DOI: 10.1515/9783110197211.1.81.

Pierrehumbert, Janet B. (2016). Phonological representation. Beyond abstract versus episodic. *Annual Review of Linguistics* 2, 33–52. DOI: 10.1146/annurev-linguist-030514-125050.

Plag, Ingo (1999). *Morphological productivity. Structural constraints in English derivation*. Mouton de Gruyter (Topics in English Linguistics 28). DOI: 10.1515/9783110802863.

Plag, Ingo, Julia Homann, and Gero Kunter (2017). Homophony and morphology. The acoustics of word-final S in English. *Journal of Linguistics* 53 (1), 181–216. DOI: 10.1017/S0022226715000183.

Ploch, Stefan (2003). Metatheoretical problems in phonology with Occam's Razor and non-*ad-hoc*-ness. In: *Living on the edge. 28 papers in honour of Jonathan Kaye*. Ed. by Stefan Ploch. Mouton de Gruyter (Studies in Generative Grammar 62), 149–202. DOI: 10.1515/9783110890563.149.

Plomp, Reinier, Louis C.W. Pols, and John P. van de Geer (1967). Dimensional analysis of vowel spectra. *JASA* 41 (3), 707–711. DOI: 10.1121/1.1910398.

Plummer, Andrew R. and Patrick F. Reidy (2018). Computing low-dimensional representations of speech from socio-auditory structures for phonetic analysis. *Journal of Phonetics* 71, 355–375. DOI: 10.1016/j.wocn.2018.09.008.

Pluymaekers, Mark, Mirjam Ernestus, and R. Harald Baayen (2005). Lexical frequency and acoustic reduction in spoken Dutch. *JASA* 118 (4), 2561–2569. DOI: 10.1121/1.2011150.

Pohl, Muna and Janet Grijzenhout (2010). Phrase-medial bilabial stops in three West Germanic languages. *Linguistische Berichte* 222, 141–167.

Pols, Louis C.W., L.J.Th. van der Kamp, and Reinier Plomp (1969). Perceptual and physical space of vowel sounds. *JASA* 16 (2/2), 458–467. DOI: 10.1121/1.1911711.

Postal, Paul M. (1968). *Aspects of phonological theory*. Harper & Row.

Pouplier, Marianne, Joan Cederbaum, Phil Hoole, Stefania Marin, and Sonja Greven (2017). Mixed modeling for irregularly sampled and correlated functional data. Speech science applications. *JASA* 142 (2), 935–946. DOI: 10.1121/1.4998555.

Pouplier, Marianne, Phil Hoole, Joan Cederbaum, Sonja Greven, and Manfred Pastätter (2014). Perceptual and articulatory factors in German fricative assimilation. In: *Proceedings of the 10th International Seminar on Speech Production*. Ed. by Susanne Fuchs, Martine Grice, Anne Hermes, Leonardo Lancia, and Doris Mücke, 332–335.

Prince, Alan and Paul Smolensky (1993/2004). *Optimality Theory. Constraint interaction in generative grammar*. Blackwell. DOI: 10.1002/9780470759400.

Prokić, Jelena and John Nerbonne (2013). Analyzing dialects biologically. In: *Classification and evolution in biology, linguistics and the history of science. Concepts, methods, visualization*. Ed. by Heiner Fangerau, Hans Geisler, Thorsten Halling, and William Martin. Franz Steiner (KulturAnamnesen. Schriften zur Geschichte und Philosophie der Medizin und der Naturwissenschaften 5), 147–161. DOI: 10.17169/refubium-21260.

Puggaard, Rasmus (2018a). Realizations of /t/ in Jutlandic dialects of Danish. *Linguistica Lettica* 26, 368–393.

Puggaard, Rasmus (2018b). Telefoniske båndpasfiltres indflydelse på talergenkendelse. Høje frekvenser og dansk /t/. *Nydanske Sprogstudier* 54, 129–155. DOI: 10.7146/nys.v1i54.100580.

Puggaard, Rasmus (2019a). Flexibility of frequent clause openers in talk-in-interaction. *Det* 'it, that' and *så* 'then' in the prefield in Danish. *Nordic Journal of Linguistics* 42 (3), 291–327. DOI: 10.1017/S0332586519000088.

Puggaard, Rasmus (2019b). Microvariation in stop realization as a regional feature of Danish. (Poster presented at *International Conference on Language Variation in Europe* 10, Frisian Academy, Leeuwarden).

Puggaard, Rasmus (2019c). Regional variation in plosive realization in Danish. (Poster presented at *Dag van de Fonetiek*, University of Amsterdam).

Puggaard, Rasmus (2020a). Affrikation af klusiler i jyske dialekter. (Paper presented at *Møde om Udforskningen af det Danske Sprog* 18, Aarhus University).

Puggaard, Rasmus (2020b). Plosives in weak prosodic positions in Jutlandic Danishes. Variation in realization or representation? (Paper presented at *Fonologi i Norden/Phonology in the Nordic Countries* 5, University of South-Eastern Norway, Drammen).

Puggaard, Rasmus (2020c). The productive acquisition of dental obstruents by Danish learners of Chinese. In: *Explorations of Chinese theoretical and applied linguistics.* Ed. by Dongyan Chen and Daniel Bell. Cambridge Scholars, 168–195. HDL: 1887/3146417.

Puggaard, Rasmus (2021). Modeling regional variation in voice onset time of Jutlandic varieties of Danish. In: *Language variation. European perspectives VIII.* Ed. by Hans Van de Velde, Nanna Haug Hilton, and Remco Knooihuizen. John Benjamins (Studies in Language Variation 25), 79–110. DOI: 10.1075/silv.25.04pug.

Puggaard, Rasmus and Yonatan Goldshtein (2019). Large-scale corpus-based analysis of regional varieties of Danish. (Poster presented at *International Conference on Language Variation in Europe* 10, Frisian Academy, Leeuwarden).

Puggaard, Rasmus and Yonatan Goldshtein (2020). Realization and representation of plosives in Jutlandic varieties of Danish. Variation in phonetics predicts variation in phonology. (Paper presented at *Laboratory Phonology* 17, University of British Columbia, Vancouver).

Puggaard, Rasmus, Janet Grijzenhout, and Bert Botma (2019). The acoustics of phonological contrast. Fortis and lenis stops in Danish and Dutch. (Paper presented at *Phonetik und Phonologie Tagung* 15, Heinrich-Heine University, Düsseldorf).

Puggaard, Rasmus, Camilla Søballe Horslund, and Henrik Jørgensen (2020). Intervocalic voicing of Danish stops. (Paper presented at *Phonetics and Phonology in Denmark* 2, Aarhus University).

Puggaard-Rode, Rasmus (2022a). Data and code for Chapters 5 and 6 of 'Stop! Hey, what's that sound?' *DataverseNL.* DOI: 10.34894/FJKFD1.

Puggaard-Rode, Rasmus (2022b). Time-varying spectral characteristics of Danish stop releases analyzed with function-on-scalar regression. *Journal of Phonetics* 95. DOI: 10.1016/j.wocn.2022.101191.

Puggaard-Rode, Rasmus, Bert Botma, and Janet Grijzenhout (forthc.). Towards a quantised representation of phonological stop contrasts. (Accepted for publication in *The primitives of phonological structure*. Ed. by Andrew Nevins, Bert Botma, Marijn van 't Veer, and Marc van Oostendorp. OUP.)

Puggaard-Rode, Rasmus, Camilla Søballe Horslund, and Henrik Jørgensen (2021). Modern Standard Danish stop gradations explained diachronically. (Poster presented at *Edinburgh Symposium on Historical Phonology* 5, University of Edinburgh).

Puggaard-Rode, Rasmus, Camilla Søballe Horslund, and Henrik Jørgensen (2022a). The rarity of intervocalic voicing of stops in Danish spontaneous speech. *Laboratory Phonology* 13 (11). DOI: 10.16995/labphon.6449.

Puggaard-Rode, Rasmus, Camilla Søballe Horslund, Henrik Jørgensen, and Dirk Jan Vet (2022b). Replication data for: The rarity of intervocalic voicing of stops in Danish spontaneous speech. *DataverseNL*. DOI: 10.34894/OSGTR8.

Puggaard-Rode, Rasmus, Henrik Jørgensen, and Camilla Søballe Horslund (2022c). Consonant gradation in Southern Scandinavia. A sound law and its outcome. (Paper presented at *Fonologi i Norden/Phonology in the Nordic Countries* 7, Aarhus University).

Pya, Natalya and Simon N. Wood (2016). A note on basis dimension selection in generalized additive modelling. (Unpublished manuscript). ARXIV: 1602.06696.

Quist, Pia (2000). Ny københavnsk 'multietnolekt'. Om sprogbrug blandt unge i sprogligt og kulturelt heterogene miljøer. *Danske Talesprog* 1, 143–211.

R Core Team (2021). R. A language and environment for statistical computing. (Version 4.1.2). URL: R-project.org.

Rahim, Karim J. (2014). *Applications of multitaper spectral analysis to nonstationary data*. PhD dissertation, Queen's University, Kingston, Ontario. HDL: 1974/12584.

Rahim, Karim J. and Wesley S. Burr (2020). `multitaper`. Spectral analysis tools using the multitaper method. (R package version 1.0-15).

Ramsammy, Michael (2015). The life cycle of phonological processes. Accounting for dialectal microtypologies. *Language and Linguistics Compass* 9 (1), 33–54. DOI: 10.1111/lnc3.12102.

Ramsay, James O., Giles Hooker, and Spencer Graves (2009). *Functional data analysis with R and MATLAB*. Springer (Use R!). DOI: 10.1007/978-0-387-98185-7.

Ramsay, James O. and Bernard W. Silverman (2005). *Functional data analysis*. 2nd ed. Springer (Springer Series in Statistics). DOI: 10.1007/b98888.

Reidy, Patrick F. (2013). An introduction to random processes for the spectral analysis of speech data. *Ohio State University Working Papers in Linguistics* 60, 67–116.

Reidy, Patrick F. (2015). A comparison of spectral estimation methods for the analysis of sibilant fricatives. *JASA: Express Letters* 137 (4), 248–254. DOI: 10.1121/1.4915064.

Reidy, Patrick F. (2016a). Spectral dynamics of sibilant fricatives are contrastive and language specific. *JASA* 140 (4), 2518–2529. DOI: 10.1121/1.4964510.

Reidy, Patrick F. (2016b). `spectRum`. Compute periodogram and multitaper spectral estimates of waveform data. (R convenience functions, available on GitHub).

Reiss, Charles (2018). Substance free phonology. In: *The Routledge handbook of phonological theory*. Ed. by S.J. Hannahs and Anna R.K. Bosch. Routledge (Routledge Handbooks in Linguistics), 425–452. DOI: 10.4324/9781315675428-15.

Reiss, Philip T., Lei Huang, and Maarten Mennes (2010). Fast function-on-scalar regression with penalized basis expansions. *International Journal of Biostatistics* 6 (1). DOI: 10.2202/1557-4679.1246.

Repp, Bruno H., Alvin M. Liberman, Thomas Eccardt, and David Pesetsky (1978). Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance* 4 (4), 621–637. DOI: 10.1037/0096-1523.4.4.621.

Rice, Keren and Peter Avery (1989). On the interaction between sonorancy and voicing. *Toronto Working Papers in Linguistics* 10, 65–82.

van Rij, Jacolien, Nemanja Vaci, Lee H. Wurm, and Laurie Beth Feldman (2020a). Alternative quantitative methods in psycholinguistics. Implications for theory and design. In: *Word knowledge and word usage. A cross-disciplinary guide to the mental lexicon*. Ed. by Vito Pirrelli, Ingo Plag, and Wolfgang U. Dressler. Mouton de Gruyter (Trends in Linguistics: Studies and Monographs 337), 83–126. DOI: 10.1515/9783110440577-003.

van Rij, Jacolien, Martijn Wieling, R. Harald Baayen, and Hedderik van Rijn (2020b). `itsadug`. Interpreting time series and autocorrelated data using GAMMs. (R package version 2.4).

van Rij, Jacolien (2016). Testing for significance. (`itsadug` R package vignette).

Ringe, Don (2006). *From Proto-Indo-European to Proto-Germanic*. OUP (A Linguistic History of English 1). DOI: 10.1093/acprof:oso/9780199284139.001.0001.

Ringgaard, Kristian (1960). *Vestjysk stød*. Aarhus Universitetsforlag.

Ringgaard, Kristian (1971). *Danske dialekter. En kortfattet oversigt*. Akademisk Forlag.

Ringgaard, Kristian (1974). Generative rules for glottalized stops in long vowels in West Jutland dialects. In: *Festskrift til Kristian Hald*. Ed. by Poul Andersen, Christian Lisse, Karl Martin Nielsen, and John Kousgård Sørensen. Museum Tusculanum (Navnestudier 13), 273–288.

Rischel, Jørgen (1970a). Consonant gradation. A problem in Danish phonology and morphology. In: *The Nordic languages and modern linguistics*. Ed. by Hreinn Benediktsson. Vísindafélag Íslendinga, 460–480. Reprinted 2009 in *Sound structure in language*. OUP, 26–43.

Rischel, Jørgen (1970b). Morpheme stress in Danish. *ARIPUC* 4, 111–144.

Rischel, Jørgen (1972). Compound stress in Danish without a cycle. *ARIPUC* 6, 211–228.

Rischel, Jørgen (1983). On unit accentuation in Danish, and the distinction between deep and surface phonology. *Folia Linguistica* 17, 51–98. DOI: 10.1515/flin.1983.17.1-4.51.

Rischel, Jørgen (2000). Review of Nina Grønnum, *Fonetik og fonologi: Almen og dansk*. *Acta Linguistica Hafniensia* 32 (1), 161–185. DOI: 10.1080/03740463.2000.10412307.

Roettger, Timo B. (2019). Researcher degrees of freedom in phonetic research. *Laboratory Phonology* 10 (1). DOI: 10.5334/labphon.147.

Roettger, Timo B. (2021). Preregistration in experimental linguistics. Applications, challenges, and limitations. *Linguistics* 59 (5), 1227–1249. DOI: 10.1515/ling-2019-0048.

Roettger, Timo B. and Dinah Baer-Henney (2019). Toward a replication culture. Speech production research in the classroom. *Phonological Data & Analysis* 1 (4). DOI: 10.3765/pda.v1art4.13.

Roettger, Timo B., Bodo Winter, and R. Harald Baayen (2019). Emergent data analysis in phonetic sciences. Towards pluralism and reproducibility. *Journal of Phonetics* 73, 1–7. DOI: 10.1016/j.wocn.2018.12.001.

RStudio Team (2022). RStudio. Integrated development environment for R. (Version 2021.09.2+382). URL: Rstudio.com.

Ryalls, John, Alison Zipprer, and Penelope Baldauff (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research* 40 (3), 642–645. DOI: 10.1044/jslhr.4003.642.

Sagey, Elizabeth Caroline (1986). *The representation of features and relations in non-linear phonology.* PhD dissertation, Massachusetts Institute of Technology. HDL: 1721.1/15106.

Sampson, Rodney (2001). Liaison, nasal vowels and productivity. *Journal of French Language Studies* 11 (2), 241–258. DOI: 10.1017/S0959269501000254.

Sawashima, Masayuki (1970). Glottal adjustments for English obstruents. *Status Report on Speech Research* 21/22, 187–200.

Schachtenhaufen, Ruben (2010a). Looking for lost syllables in Danish spontaneous speech. In: *Linguistic theory and raw sound.* Ed. by Peter Juel Henrichsen. Samfundslitteratur (Copenhagen Studies in Language 40), 61–85.

Schachtenhaufen, Ruben (2010b). Schwa-assimilation og stavelsesgrænser. *Nydanske Sprogstudier* 39, 64–92. DOI: 10.7146/nys.v39i39.13532.

Schachtenhaufen, Ruben (2012). Reduction of word final schwa in Danish. Rhythmic and syntactic factors. In: *Speech in action.* Ed. by Jan Heegård and Peter Juel Henrichsen. Samfundslitteratur (Copenhagen Studies in Language 42), 53–72.

Schachtenhaufen, Ruben (2013). *Fonetisk reduktion i dansk.* PhD dissertation, Copenhagen Business School.

Schachtenhaufen, Ruben (2020–). *[u]dtaleordbog.dk.* URL: udtaleordbog.dk.

Schachtenhaufen, Ruben (2022). *Ny dansk fonetik.* Books-on-demand.

Schachtenhaufen, Ruben (forthc.). Illustrations of the IPA. Danish. (Submitted to *Journal of the International Phonetic Alphabet*).

Schad, Daniel J., Shravan Vasishth, Sven Hohenstein, and Reinhold Kliegl (2020). How to capitalize on a priori contexts in linear (mixed) models. A tutorial. *Journal of Memory and Language* 110. DOI: 10.1016/j.jml.2019.104038.

Schane, Sanford A. (1968). *French phonology and morphology.* MIT Press.

Scheer, Tobias (2015). How diachronic is synchronic grammar? Crazy rules, regularity, and naturalness. In: *The Oxford handbook of historical phonology.* Ed.

by Patrick Honeybone and Joseph C. Salmons. OUP (Oxford Handbooks Online), 313–336. DOI: 10.1093/oxfordhb/9780199232819.013.003.

Scheipl, Fabian, Jan Gertheiss, and Sonja Greven (2016). Generalized functional additive mixed models. *Electronic Journal of Statistics* 10 (1), 1455–1492. DOI: 10.1214/16-EJS1145.

Scheipl, Fabian, Ana-Maria Staicu, and Sonja Greven (2015). Functional additive mixed models. *Journal of Computational and Graphical Statistics* 24 (2), 477–501. DOI: 10.1080/10618600.2014.901914.

Scheuer, Jann, Anne Larsen, Marie Maegaard, Malene Monka, and Kristine Køhler Mortensen (2019). Language ideologies. A key to understanding language standardization. In: *Standardization as sociolinguistic change. A transversal study of three traditional dialect areas.* Ed. by Marie Maegaard, Malene Monka, Kristine Køhler Mortensen, and Andreas Candefors Stæhr. Routledge (Routledge Studies in Language Change), 189–218. DOI: 10.4324/9780429467486-8.

Schwarz, Martha, Myriam Lapierre, Karee Garvin, and Sharon Inkelas (2019). Recent advances in Q theory. Segment strength. *Proceedings of the Linguistic Society of America* 4 (51). DOI: 10.3765/plsa.v4i1.4550.

Seedorff, Michael, Jacob Oleson, and Bob McMurray (2019). Maybe maximal. Good enough mixed models optimize power while controlling Type I error. (Unpublished manucript).

Séguy, Jean (1971). La relation entre la distance spatiale et la distance lexicale. *Revue de Linguistique Romane* 35 (139–140), 335–357. DOI: 10.5169/seals-399508.

Séguy, Jean (1973). La dialectométrie dans l'Atlas linguistique de Gascogne. *Revue de Linguistique Romane* 37 (145), 1–24. DOI: 10.5169/seals-658403.

Seifert, Jan (2009). Does speech reveal one's age? On the use of gerontolinguistic topics for forensic authorship analysis. In: *Formal linguistics and law*. Ed. by Günther Grewendorf and Monika Rathert. Mouton de Gruyter (Trends in Linguistics: Studies and Monographs 212), 163–180. DOI: 10.1515/9783110218398.2.163.

Selkirk, Elisabeth O. (1972). *The phrase phonology of English and French.* PhD dissertation, Massachusetts Institute of Technology. HDL: 1721.1/14788. Published 1980 by Routledge. DOI: 10.4324/9781315687995.

Selkirk, Elisabeth O. and Jean-Roger Vergnaud (1973). How abstract is French phonology? *Foundations of Language* 10 (2), 249–254. JSTOR: 25000715.

Shadle, Christine H. (1991). The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics* 19 (3–4), 409–424. DOI: 10.1016/S0095-4470(19)30332-8.

Shadle, Christine H. and Sheila J. Mair (1996). Quantifying spectral characteristics of fricatives. In: *Proceedings of the 4th International Conference on Spoken Language Processing*. Ed. by H. Timothy Bunnell and William Idsardi. International Speech Communication Association, 1521–1524. DOI: 10.1109/ICSLP.1996.607906.

Shih, Chilin and Bernd Möbius (1998). Contextual effects on voicing profiles of German and Mandarin consonants. In: *Proceedings of the 3rd ESCA/COCOSDA Workshop on Speech Synthesis*, 81–86.

Shih, Chilin, Bernd Möbius, and Bhuvana Narasimhan (1999). Contextual effects on consonant voicing profiles. A cross-linguistic study. In: *ICPhS 14*. Ed. by John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville, and Ashlee C. Bailey, 989–992.

Shih, Stephanie S. and Sharon Inkelas (2019a). Autosegmental aims in surface-optimizing phonology. *Linguistic Inquiry* 50 (1), 137–196. DOI: 10.1162/ling_a_00304.

Shih, Stephanie S. and Sharon Inkelas (2019b). Subsegments and the emergence of segments. *Proceedings of the Linguistic Society of America* 4 (37). DOI: 10.3765/plsa.v4i1.4541.

Shipp, Thomas (1975). Vertical laryngeal position during continuous and discrete vocal frequency change. *Journal of Speech and Hearing Research* 18 (4), 707–718. DOI: 10.1044/jshr.1804.707.

Siem, Andrea Brink (2019). *An articulatory account of the Danish soft d.* Master's thesis, Aarhus University.

Silverman, Daniel (2004). On the phonetic and cognitive nature of alveolar stop allophony in American English. *Cognitive Linguistics* 15 (1), 69–93. DOI: 10.1515/cogl.2004.002.

Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn (2011). False-positive psychology. Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* 22 (11), 1359–1366. DOI: 10.1177/0956797611417632.

Skautrup, Peter (1928). Klusiler og "yngre" stød i vestjysk. *Acta Philologica Scandinavica* 3, 32–51.

Skautrup, Peter (1930). *Hardiske mål.* Vol. 1. Th. Linds Efterfølger.

Skautrup, Peter (1937). Dialektkriterier og deres værdi. *Sprog og kultur* 6, 59–80.

Skautrup, Peter (1944–1970). *Det danske sprogs historie.* Gyldendal.

Skautrup, Peter (1952). Sammenhænge i dialektal udvikling. In: *Runer og rids. Festskrift til Lis Jakobsen.* Ed. by Jørgen Glahder. Roskilde og Bagger, 95–103.

Skautrup, Peter, Ove Rasmussen, Viggo Sørensen, Torben Arboe, Inger Schoonderbeek Hansen, Nina Grøftehauge, Mette-Marie Møller Svendsen, and Kristoffer Friis Bøegh, eds. (1970–). *Jysk ordbog.* Aarhus University. URL: jyskordbog.dk.

Smaldino, Paul E. and Richard McElreath (2016). The natural selection of bad science. *Royal Society Open Science* 3 (9). DOI: 10.1098/rsos.160384.

Smith, George W. (1991). *Computers and human language.* OUP.

Smith, Jennifer L. (2008). Markedness, faithfulness, positions, and contexts. Lenition and fortition in Optimality Theory. In: *Lenition and fortition.* Ed. by Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral. Mouton de Gruyter (Studies in Generative Grammar 99), 519–560. DOI: 10.1515/9783110211443.3.519.

Smith, Svend (1944). *Bidrag til Løsning af Problemer vedrørende Stødet i dansk Rigssprog. En eksperimentalfonetisk Studie.* Kaifers.

Sobkowiak, Mikołaj (2018). Can phonemes be considered troublemakers? The case of Danish /r/. *Lublin Studies in Modern Languages and Literature* 42 (1), 49–65. DOI: 10.17951/lsmll.2018.42.1.49.

Sonderegger, Morgan (2022). *Regression modeling for linguistic data.* (To be published by MIT Press). DOI: 10.17605/OSF.IO/PNUMG.

Sonderegger, Morgan and Joseph Keshet (2012). Automatic measurement of voice onset time using discriminative structured prediction. *JASA* 132 (6), 3965–3979. DOI: 10.1121/1.4763995.

Sonderegger, Morgan, Jane Stuart-Smith, Thea Knowles, Rachel Macdonald, and Tamara Rathcke (2020). Structured heterogeneity in Scottish stops over the twentieth century. *Language* 96 (1), 94–125. DOI: 10.1353/lan.2020.0003.

Sóskuthy, Márton (2017). Generalised additive mixed models for dynamic analysis in linguistics. A practical introduction. (Unpublished manuscript). ARXIV: 1703.05339.

Sóskuthy, Márton (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics* 84. DOI: 10.1016/j.wocn.2020.101017.

Spinu, Laura and Jason Lilley (2016). A comparison of cepstral coefficients and spectral moments in the classification of Romanian fricatives. *Journal of Phonetics* 57, 40–58. DOI: 10.1016/j.wocn.2016.05.002.

Spore, Palle (1965). *La langue danoise. Phonétique et grammaire contemporaines.* Akademisk Forlag.

Stampe, David (1969). The acquisition of phonetic representation. In: *Papers from the 5th regional meeting of the Chicago Linguistic Society.* Ed. by Robert I. Binnick, Alice Davidson, Georgia M. Green, and Jerry L. Morgan. Chicago Linguistic Society, 443–454.

Stathopoulos, Elaine T. and Gary Weismer (1983). Closure duration of stop consonants. *Journal of Phonetics* 11 (4), 395–400. DOI: 10.1016/S0095-4470(19)30838-1.

Stefan, Angelika M. and Felix D. Schönbrodt (2022). Big little lies. A compendium and simulation of *p*-hacking strategies. (Unpublished manuscript).

Steinberg, Danny D. and Robert K. Krohn (1975). The psychological validity of Chomsky and Halle's Vowel Shift Rule. In: *The transformational–generative paradigm and modern linguistic theory.* Ed. by E.F.K. Koerner. John Benjamins (Current Issues in Linguistic Theory 1), 233–259. DOI: 10.1075/cilt.1.11ste.

Steriade, Donca (1987). Locality conditions and feature geometry. In: *Proceedings of North East Linguistics Society 17, Volume II.* Ed. by Joyce McDonough and Bernadette Plunkett. Graduate Linguistics Student Association, University of Massachusetts, 595–617.

Steriade, Donca (1994). Positional neutralization and the expression of contrast. (Unpublished manuscript).

Steriade, Donca (2009). The phonology of perceptibility effects. The P-Map and its consequences for constraint organization. In: *The nature of the word. Studies in honor of Paul Kiparsky.* MIT Press (Current Studies in Linguistics), 151–197. DOI: 10.7551/mitpress/7894.003.0011.

Stevens, Kenneth N. (1971). Airflow and turbulence noise for fricative and stop consonants. Static considerations. *JASA* 50 (4), 1180–1191. DOI: 10.1121/1.1912751.
Stevens, Kenneth N. (1993a). Modelling affricate consonants. *Speech Communication* 13 (1), 33–43. DOI: 10.1016/0167-6393(93)90057-R.
Stevens, Kenneth N. (1993b). Models for the production and acoustics of stop consonants. *Speech Communication* 13 (3), 367–375. DOI: 10.1016/0167-6393(93)90035-J.
Stevens, Kenneth N. (1998). *Acoustic phonetics*. MIT Press. DOI: 10.7551/mitpress/1072.001.0001.
Stevens, Kenneth N. and Sheila E. Blumstein (1978). Invariant cues for place of articulation in stop consonants. *JASA* 64 (5), 1358–1368. DOI: 10.1121/1.382102.
Stevens, Kenneth N. and Sheila E. Blumstein (1981). The search for invariant acoustic correlates of phonetic features. In: *Perspectives on the study of speech*. Ed. by Peter D. Eimas and Joanne L. Miller. Lawrence Erlbaum, 1–38. DOI: 10.4324/9780203780862.
Stevens, Kenneth N., Sharon Y. Manuel, and Melanie Matthies (1999). Revisiting place of articulation measures for stop consonants. Implications for models of consonant production. In: *ICPhS 14*. Ed. by John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville, and Ashlee C. Bailey. University of California, 1117–1120.
Stoel-Gammon, Carol, Karen A. Williams, and Eugene Buder (1994). Cross-language differences in phonological acquisition. Swedish and American /t/. *Phonetica* 51 (1–3), 146–158. DOI: 10.1159/000261966.
Strevens, Peter (1960). Spectra of fricative noise in human speech. *Language and Speech* 3 (1), 32–49. DOI: 10.1177/002383096000300105.
Strycharczuk, Patrycja (2012). Sonorant transparency and the complexity of voicing in Polish. *Journal of Phonetics* 40 (5), 655–671. DOI: 10.1016/j.wocn.2012.05.006.
Stuart-Smith, Jane (2007). Empirical evidence for gendered speech production. /s/ in Glaswegian. In: *Laboratory Phonology 9*. Ed. by Jennifer Cole and José Ignacio Hualde. Mouton de Gruyter (Phonology and Phonetics 4/3), 65–86.
Stuart-Smith, Jane, Tamara Rathcke, Morgan Sonderegger, and Rachel Macdonald (2013). A real-time study of plosives in Glaswegian using an automatic measurement algorithm. Change or age-grading? In: *Language variation. European perspectives V*. Ed. by Eivind Torgersen, Stian Hårstad, Brit Mæhlum, and Unn Røyneland. John Benjamins (Studies in Language Variation 17), 225–237. DOI: 10.1075/silv.17.17stu.
Stuart-Smith, Jane, Morgan Sonderegger, Tamara Rathcke, and Rachel Macdonald (2015). The private life of stops. VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology* 6 (3–4), 505–548. DOI: 10.1515/lp-2015-0015.
Stæhr, Andreas Candefors (2015). Reflexivity in Facebook interaction. Enregisterment across written and spoken language practices. *Discourse, Context & Media* 8, 30–45. DOI: 10.1016/j.dcm.2015.05.004.
Stæhr, Andreas Candefors and Anne Larsen (2019). Bornholm. The terminal stage of dedialectalization. In: *Standardization as sociolinguistic change. A transversal*

study of three traditional dialect areas. Ed. by Marie Maegaard, Malene Monka, Kristine Køhler Mortensen, and Andreas Candefors Stæhr. Routledge (Routledge Studies in Language Change), 114–144. DOI: 10.4324/9780429467486-5.

Swartz, Bradford L. (1992). Gender difference in voice onset time. *Perceptual and Motor Skills* 75 (5), 983–992. DOI: 10.2466/pms.1992.75.3.983.

Sweet, Henry (1874). On Danish pronunciation. *Transactions of the Philological Society* 15 (1), 94–112. DOI: 10.1111/j.1467-968X.1874.tb00867.x.

Swerts, Marc (1994). *Prosodic features of discourse units*. PhD dissertation, Eindhoven University of Technology. DOI: 10.6100/IR411593.

Swerts, Marc and René Collier (1992). On the controlled elicitation of spontaneous speech. *Speech Communication* 11 (4–5), 463–468. DOI: 10.1016/0167-6393(92)90052-9.

Søndergaard, Bent (1970). Den såkaldte klusilspring, specielt i morsingmål. *Sprog og kultur* 26, 61–68.

Sørensen, Viggo (2012). Lyd og prosodi i de danske dialekter. (Unpublished manuscript).

Tanner, James, Morgan Sonderegger, and Jane Stuart-Smith (2020). Structured speaker variability in Japanese stops. Relationships within versus across cues to stop voicing. *JASA* 148 (2), 793–804. DOI: 10.1121/10.0001734.

Tanner, James, Morgan Sonderegger, and Michael Wagner (2017). Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology* 8 (15). DOI: 10.5334/labphon.96.

Terken, Jacques M.B. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech* 27 (3), 269–289. DOI: 10.1177/002383098402700306.

Thráinsson, Höskuldur (1978). On the phonology of Icelandic preaspiration. *Nordic Journal of Linguistics* 1 (1), 3–54. DOI: 10.1017/S0332586500000196.

Tomaschek, Fabian, Ingo Plag, Mirjam Ernestus, and R. Harald Baayen (2021). Phonetic effects of morphology and context. Modeling the duration of word-final s in English with naïve discriminative learning. *Journal of Linguistics* 57 (1), 123–161. DOI: 10.1017/S0022226719000203.

Torp, Arne (1998). *Nordiske språk i nordisk og germansk perspektiv*. Novus.

Torre, Peter and Jessica A. Barlow (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42 (5), 324–333. DOI: 10.1016/j.jcomdis.2009.03.001.

Trudgill, Peter (1974). Linguistic change and diffusion. Description and explanation in sociolinguistic dialect geography. *Language in Society* 3 (2), 215–246. DOI: 10.1017/S0047404500004358.

Trudgill, Peter (1983). *On dialect. Social and geographical perspectives*. Basil Blackwell.

Tsuchida, Ayako, Abigail C. Cohn, and Masanobu Kumada (2000). Sonorant devoicing and the phonetic realization of [spread glottis] in English. *Working Papers of the Cornell Phonetics Laboratory* 13, 167–181.

Tøndering, John (2003). Intonation contours in Danish spontaneous speech. In: *ICPhS 15*. Ed. by Maria-Josep Solé, Daniel Recasens, and J. Romero. Universitat Autonoma De Barcelona, 1241–1244.

Tøndering, John (2008). *Skitser af prosodi i spontant dansk*. PhD dissertation, University of Copenhagen.

Uldall, Hans Jörgen (1936). The phonematics of Danish. In: *ICPhS 2*. Ed. by Daniel Jones and Dennis Butler Fry. CUP, 54–57.

Vasishth, Shravan, Daniela Mertzen, Lena A. Jäger, and Andrew Gelman (2018a). The statistical significance filter leads to overoptimistic expectations of replicability. *Journal of Memory and Language* 103, 151–175. DOI: 10.1016/j.jml.2018.07.004.

Vasishth, Shravan, Bruno Nicenboim, Mary E. Beckman, Fangfang Li, and Eun Jong Kong (2018b). Bayesian data analysis in the phonetic sciences. A tutorial introduction. *Journal of Phonetics* 71, 147–161. DOI: 10.1016/j.wocn.2018.07.008.

Vaux, Bert and Bridget Samuels (2005). Laryngeal markedness and aspiration. *Phonology* 22 (3), 395–436. DOI: 10.1017/S0952675705000667.

Vázquez-Larruscain, Miguel (2021). Complementary length in Danish. Why not? *Nordlyd* 45, 39–55. DOI: 10.7557/12.6247.

Veirup, Hans (1955). Redegørelse for udviklingen af postvokalisk t i jyske dialekter. (Unpublished manuscript).

Vestergaard, Torben (1967). Initial and final consonant combinations in Danish monosyllables. *Studia Linguistica* 21 (1), 37–66. DOI: 10.1111/j.1467-9582.1967.tb00547.x.

Vilain, Anne, Christian Abry, and Pierre Badin (1998). Coarticulation and degrees of freedom in the elaboration of a new articulatory plant: GENTIANE. In: *5th International Conference on Spoken Language Processing*. International Speech Communication Association.

Voeten, Cesko C. (2020). *The adoption of sound change. Synchronic and diachronic processing of regional variation in Dutch*. PhD dissertation, Leiden University. Published by LOT (Dissertation Series 578). HDL: 1887/137723.

Volkmann, Alexander, Almond Stöcker, Fabian Scheipl, and Sonja Greven (2021). Multivariate functional additive mixed models. *Statistical Modelling*. DOI: 10.1177/1471082X211056158.

Watson, Catherine I. and Jonathan Harrington (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *JASA* 106 (1), 458–468. DOI: 10.1121/1.427069.

Wedel, Andrew B. (2006). Exemplar models, evolution and language change. *The Linguistic Review* 23 (3), 247–274. DOI: 10.1515/TLR.2006.010.

Wenker, Georg and Ferdinand Wrede (1895). *Der Sprachatlas des deutschen Reichs*. Elwert.

Westbury, John R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *JASA* 73 (4), 1322–1336. DOI: 10.1121/1.389236.

Westbury, John R. and Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22 (1), 145–166. DOI: 10.1017/S0022226700010598.

Wickham, Hadley (2016). *ggplot2. Elegant graphics for data analysis.* Springer (Use R). DOI: 10.1007/978-0-387-98141-3.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington (2021). `ggplot2`. Create elegant data visualizations using the grammar of graphics. (R package version 3.3.5).

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller (2022). `dplyr`. A grammar of data manipulation. (R package version 1.0.8).

Wieling, Martijn (2012). *A quantitative approach to social and geographical dialect variation.* PhD dissertation, University of Groningen. (Groningen Dissertations in Linguistics 103).

Wieling, Martijn (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling. A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70, 86–116. DOI: 10.1016/j.wocn.2018.03.002.

Wieling, Martijn, Simonetta Montemagni, John Nerbonne, and R. Harald Baayen (2014). Lexical differences between Tuscan dialects and Standard Italian. Accounting for geographic and sociodemographic variation using generalized additive mixed modeling. *Language* 90 (3), 669–692. DOI: 10.1353/lan.2014.0064.

Wieling, Martijn and John Nerbonne (2015). Advances in dialectometry. *Annual Review of Linguistics* 1, 243–264. DOI: 10.1146/annurev-linguist-030514-124930.

Wieling, Martijn, John Nerbonne, and R. Harald Baayen (2011). Quantitative social dialectology. Explaining linguistic variation geographically and socially. *Plos One* 6 (9). DOI: 10.1371/journal.pone.0023613.

Wood, Simon N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society B* 65 (1), 95–114. DOI: 10.1111/1467-9868.00374.

Wood, Simon N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society B* 73 (1), 3–36. DOI: 10.1111/j.1467-9868.2010.00749.x.

Wood, Simon N. (2013). On *p*-values for smooth components of an extended generalized additive model. *Biometrika* 100 (1), 221–228. DOI: 10.1093/biomet/ass048.

Wood, Simon N. (2017a). *Generalized additive models. An introduction with R.* 2nd ed. CRC Press (Texts in Statistical Science). DOI: 10.1201/9781315370279.

Wood, Simon N. (2017b). P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. *Statistics and Computing* 27 (4), 985–989. DOI: 10.1007/s11222-016-9666-x.

Wood, Simon N. (2021). `mgcv`. Mixed GAM computation vehicle with automatic smoothness estimation. (R package version 1.8-38).

Wood, Simon N., Yannig Goude, and Simon Shaw (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society C* 64 (1), 139–155. DOI: 10.1111/rssc.12068.

Wood, Simon N., Zheyuan Li, Gavin Shaddick, and Nicole H. Augustin (2017). Generalized additive models for gigadata. Modeling the U.K. Black Smoke network daily data. *Journal of the American Statistical Association* 112 (519), 1199–1210. doi: 10.1080/01621459.2016.1195744.

Wood, Simon N., Natalya Pya, and Benjamin Säfken (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association* 111 (516), 1548–1563. doi: 10.1080/01621459.2016.1180986.

Yan, Mengzhu (2016). *Perceptual cue weighting of Chinese /t, t^h, ts, ts^h/ by Danish learners of Chinese and native Chinese listeners.* Master's thesis, Aarhus University.

Yan, Mengzhu and Marjoleine Sloos (2019). Perceptual cue weighting in L1 and L2 Chinese. An investigation of aspiration and frication in Chinese obstruents by Danish learners. *Chinese as a Second Language Research* 8 (1), 57–83. doi: 10.1515/caslar-2019-0003.

Yip, Moira (1991). Coronals, consonant clusters, and the coda condition. In: *The special status of coronals. Internal and external evidence.* Ed. by Carole Paradis and Jean-François Prunet. Academic Press (Phonetics and Phonology 2), 61–78. doi: 10.1016/B978-0-12-544966-3.50010-5.

Zellers, Margaret, Michele Gubian, and Brechtje Post (2010). Redescribing intonational categories with functional data analysis. In: *Proceedings of Interspeech.* International Speech Communication Association, 1141–1144. doi: 10.21437/Interspeech.2010-361.

Zhang, Xiaoke and Jane-Ling Wang (2016). From sparse to dense functional data and beyond. *The Annals of Statistics* 44 (5), 2281–2321. doi: 10.1214/16-AOS1446.

Zhou, Xinhui, Carol Y. Espy-Wilson, Suzanne Boyce, Mark Tiede, Christy Holland, and Ann Choe (2008). A magnetic resonance imaging-based articulatory and acoustic study of "retroflex" and "bunched" American English /r/. *JASA* 123 (6), 4466–4481. doi: 10.1121/1.2902168.

ØMO = *Ømålsordbogen. En sproglig-saglig ordbog over dialekterne på Sjælland, Lolland-Falster, Fyn og omliggende øer* (1992–). Edited by the Center for Dialect Research, University of Copenhagen. Syddansk Universitetsforlag (Universitetsjubilæets danske samfund).

Resume på dansk

Denne afhandling behandler de danske lukkelyde. Moderne rigsdansk har seks kontrastive lukkelyde, /b d g p t k/. I prævokalisk stilling er den laryngale kontrast baseret på aspiration; /b d g/ er ustemte og uaspirerede, mens /p t k/ er ustemte og aspirerede. Det siges af og til at /b d g/ har stemte mediale allofoner. /t/ har klart affrikeret opløsning, og transskiberes som regel fonetisk med [tˢ]. Ifølge den mest udbredte fonologiske analyse af lukkelydene alternerer /b d g/ med semivokaler og nulrealisationer i postvokalisk medial og final stilling. Der er variation på alle disse parametre i de traditionelle dialekter, men de eksisterende beskrivelser af danske dialekter har primært fokuseret på variation i kategoriske fonologiske mønstre.

Jeg introducerer og motiverer afhandlingens emneområde i kapitel 1. Her giver jeg også et generelt overblik over det danske sprog og præsenterer lydskriftskonventionerne som benyttes i afhandlingen. Resten af afhandlingen er fordelt over to dele. Del I giver et bredt overblik over de danske lukkelyde og del II består af korpusundersøgelser som behandler mere afgrænsede emner.

Kapitel 2 er en omfattende gennemgang af den eksisterende akademiske litteratur omhandlende danske lukkelyde, fordelt over fire indfaldsvinkler: historie, fonetik, fonologi og variation. Gennemgangen fokuserer mest på den laryngale kontrast: hvordan den har udviklet sig over tid, dens akustiske og artikulatoriske træk, hvordan den evt. er repræsenteret på et abstrakt fonologisk plan, og hvordan den varierer. Dette kapitel holder sig mest til at opsummere den eksis-

terende litteratur, men indeholder også kritisk diskussion af foreslåede abstrakte underliggende konsonantklynger, og af de konsekvenser som forudsiges af nogle tilgange til den fonologiske repræsentation af laryngal kontrast.

Kapitel 3 behandler et berygtet fænomen i dansk fonologi, nemlig *konsonantsvækkelsen*, som bl.a. er ansvarlig for alterneren mellem uaspirerede lukkelyde [p t k] og semivokaler [ʊ̯ ɣ̯ ɪ̯]. Ifølge traditionelle fonologiske analyser af danske konsonanter er disse stillingsbestemte allofoner af fonemerne /b d g/. Kapitlet isolerer en række problemer med de traditionelle analyser: de resulterer i udbredt kontrastneutralisering, de er fonetisk set unaturlige, og den eneste morfofonologiske evidens for de traditionelle analyser hviler på uregelmæssig, ikke-produktiv morfologi. Kapitlet foreslår en alternativ analyse, hvor [ʊ̯ ɪ̯] altid er allofoner af /v j/. Der argumenteres yderligere for, at de strukturelle generaliseringer som de traditionelle analyser afdækker bedre kan forstås som et historisk resultat af velkendte fonetiske principper, i særdeleshed princippet som regulerer imod stemthed i obstruentlyde.

Kapitel 4 præsenterer resultaterne af en korpusundersøgelse af intervokalisk stemthed i danske lukkelyde. Dette er det første empiriske studie af dansk intervokalisk stemthed, på trods af at mange nært relaterede emner tidligere er blevet undersøgt eksperimentelt. Nogle kilder har omtalt intervokalisk stemthed som tilnærmelsesvist undtagelsesløst, mens andre kilder har omtalt intervokalisk stemthed som grundlæggende ikke-eksisterende; begge disse standpunkter har skabt grundlag for fonologiske analyser af hvordan laryngal kontrast bedst repræsenteres (både i et dansk perspektivt og i et mere generelt perspektiv). Undersøgelsen påviser at intervokalisk stemthed er overordentligt sjældent i /p t k/ og rimeligt begrænset i /b d g/, antageligt fordi stemmeridsen (*glottis*) er let spredt i udtalen af /b d g/, hvilket blokerer stemthed. Jeg undersøger med en logistisk blandet regressionsmodel hvordan en række forskellige variable påvirker sandsynligheden for stemthed, og konkluderer at der først og fremmest er tale om et svækkelsesfænomen. Såkaldte "gestus-baserede" tilgange til fonologisk repræsentation (såsom den *artikulatoriske fonologi*) forudsiger disse resultater, mens andre tilgange forudsiger delvist afvigende resultater.

Kapitel 5 præsenterer en eksplorativ korpusundersøgelse af de spektrale karakteristika i aspirerede lukkelydes opløsning, og hvordan disse karakteristika påvirkes af den fonetiske kontekst. Det nævnes tit i litteraturen at /t/ er stærkt affrikeret, men tidligere beskrivelser har alle været impressionistiske. Opløsningen i /p k/ nævnes sjældent i tidligere kilder, men Otto Jespersen foreslog for mere end hundrede år siden at alle de aspirerede lukkelyde /p t k/ var i gang med at udvikle sig til affrikater, ligesom det skete på tysk i det andet germanske konsonantskifte. Resultaterne af undersøgelsen viser at /t/ er affrikeret på tværs af alle fonetiske kontekster, men også at affrikation gradvist afløses af aspiration i løbet af opløsningen; den relative andel af affrikation og aspiration moduleres af den fonetiske kontekst. Opløsninger i /k/ påvirkes i høj grad af fonetisk kontekst, antageligt fordi det præcise artikulationssted i velære konsonanter i høj grad bestemmes af den fonetiske kontekst. Opløsninger i /p/ påvirkes også af den fonetiske kontekst, men de påvirkes mere af talerbestemt end af kontekstbestemt variation. Disse resultater er interessante i sig selv, men kapitlets primære bidrag må siges at være metodologisk. Data analyseres med såkaldte *function-on-scalar* ("funktion-på-skalær") regressionsmodeller med hele spektret som responsvariabel. Denne metode kan bruges til at løse et klassisk problem i analysen af den komplekse multidimensionelle information i støjfyldte spektra, og metoden tillader for intuitiv visualisering af hvordan spektret ændrer sig over tid i forskellige fonetiske kontekster.

Kapitel 6 kombinerer to korpusundersøgelser af udtalen af lukkelyde i traditionelle jyske dialekter på baggrund af et ældre korpus af båndoptagelser som delvist bevarer et ældre stadie af sproglig variation. Det ene studie udforsker variation i hvornår stemthed begynder i forhold til lukkets opløsning (såkaldt *voice onset time*), og det andet udforsker variation i de spektrale karakteristike i de aspirerede lukkelydes opløsning. Kapitlet diskuterer også kort stemthed i nordjyske dialekter, hvor stemthed tilsyneladende er meget mere udbredt end i moderne rigsdansk. Stemthed begynder typisk hurtigere i jyske dialekter end i moderne rigsdansk, men der er komplekse regionale variationsmønstre, som jeg forklarer med henvisning til en kombination af sociale og historiske faktorer. De spektrale karakteriska i /t/-opløsninger varierer på tværs af den jyske halvø; langt størstedelen af Jylland havde endnu ikke tilpasset den moderne rigsdanske /t/-affrikation som diskuteres i kapitel 5 da optagelserne blev lavet. Kapitlet behandler nogle komplekse dynamiske variable. Den geografiske variabel behandles med generaliserede additive blandede modeller, og den spektrale variabel behandles med funktionel hovedkomponentanalyse, som bruges til at løse nogle af de samme problemer som *function-on-scalar* regression løser i kapitel 5.

I kapitel 7 opsummerer jeg afhandlingens primære påstande. Som i kapitel 2 foregår dette med fokus på lukkelydenes historie, fonetik, fonologi og variation. Til sidst diskuterer jeg mere generelt afhandlingens forhold til nye strømninger i dataanalytisk forskning.

Samenvatting in het Nederlands

Dit proefschrift gaat over plofklanken (plosieven) in het Deens. Het moderne standaard Deens heeft zes contrastieve plofklanken, /b d g p t k/. In een traditionele fonologische analyse worden deze klanken gecategoriseerd op basis van een stemcontrast: /b d g/ zijn stemhebbend en /p t k/ stemloos. De fonetische realisatie van het stemcontrast is gecompliceerd. De realisatie hangt af van de positie van de plofklank in het woord en varieert bovendien per dialect. Bestaande fonologische beschrijvingen schenken doorgaans weinig aandacht aan deze variatie, maar zijn vooral gericht op categorische patronen. In deze beschrijvingen wordt bijvoorbeeld opgemerkt dat /b d g/, wanneer ze voorafgaan aan een klinker, stemloos en ongeaspireerd zijn, en /p t k/ stemloos en geaspireerd; dat /b d g/, wanneer ze zich tussen klinkers bevinden, stemhebbend zijn; en dat /t/ wordt gerealiseerd met sterke affricatie. In sommige bronnen wordt deze klank dan ook getranscribeerd als [ts].

Hoofdstuk 1 geeft een algemeen overzicht van de fonetiek en fonologie van het Deens en van de transcriptieconventies die in het proefschrift worden gebruikt. De rest van het proefschrift bestaat uit twee delen. Deel I behandelt de fonetische en fonologische aspecten van Deense plofklanken. Deel II bespreekt de resultaten van een aantal fonetische onderzoeken die zijn gebaseerd op corpusdata.

Hoofdstuk 2 geeft een gedetailleerde beschrijving van de plofklanken van het Deens. Achtereenvolgens worden de diachrone ontwikkeling, de fonetische eigenschappen, de fonologische distributie en representatie, en de sociale en regionale variatie in de realisatie van plofklanken besproken.

Hoofdstuk 3 gaat dieper in op een berucht aspect van de fonologie van het Deens: de verzwakking (lenitie) van plofklanken. Dit proces zorgt onder andere voor alternanties tussen ongeaspireerde [p t k] en de halfklinkers [ʊ̯ ɣ̯ ɪ̯]. Volgens de traditionele fonologische analyse van Deense medeklinkers zijn deze klanken positionele allofonen van de fonemen /b d g/. Het proefschrift laat echter zien dat deze analyse problematisch is. De analyse heeft als ongewenst effect dat sommige onderliggende contrasten geneutraliseerd worden. Daarnaast zijn de alternanties die worden voorgesteld fonetisch gezien onnatuurlijk, en zijn ze beperkt tot onregelmatige en niet-productieve morfologie. In het hoofdstuk wordt een alternatieve analyse voorgesteld waarin [ʊ̯ ɪ̯] allofonen van /v j/ zijn. Verder wordt betoogd dat de synchrone generalisaties van de traditionele analyse beter kunnen worden gezien als het resultaat van een fonetisch natuurlijke diachrone ontwikkeling, die in de eerste plaats is veroorzaakt door een afkeer van stemhebbende obstruenten.

Hoofdstuk 4 presenteert de resultaten van een corpusonderzoek naar de stemhebbendheid van /b d g/ wanneer deze zich tussen klinkers bevinden – een proces dat *intervocalic voicing* wordt genoemd. Hoewel de fonetiek van het Deens een lange onderzoekstraditie kent, is dit de eerste gedetailleerde studie naar *intervocalic voicing*. In sommige bronnen wordt *intervocalic voicing* beschreven als een categorisch proces, terwijl andere bronnen het bestaan van dit proces ontkennen. Beide standpunten hebben de basis gevormd voor verschillende fonologische representaties van het Deense stemcontrast. De resultaten van het corpusonderzoek laten zien dat /p t k/ tussen klinkers zeer zelden stemhebbend zijn, en dat ook /b d g/ betrekkelijk ongevoelig zijn voor *intervocalic voicing*. Dit komt waarschijnlijk doordat de glottis tijdens de realisatie van /b d g/ een beetje gespreid is, wat stembandtrilling tegengaat.

Hoofdstuk 5 presenteert de resultaten van een corpusonderzoek naar de spectrale kenmerken van Deense geaspireerde plofklanken. In de literatuur is vaak opgemerkt dat /t/ in het Deens sterk geaffriceerd is, maar deze observaties waren tot op heden impressionistisch of introspectief. Vroegere bronnen maken zelden gewag van affricatie van /p/ en /k/, al stelde Otto Jespersen meer dan honderd jaar geleden voor dat de geaspireerde plofklanken van het Deens zich aan het ontwikkelen waren tot affricaten, net zoals in het Duits was gebeurd tijdens de tweede Germaanse medeklinkerverschuiving. De resultaten van het corpusonderzoek laten zien dat /t/ inderdaad in alle fonetische contexten geaffriceerd is, maar ook dat de affricatie tijdens de ruisplof geleidelijk overgaat in aspiratie. De verhouding tussen de mate van affricatie en aspiratie is afhankelijk van de fonetische context, in het bijzonder voor /k/. De verhouding van affricatie en aspiratie is het meest onderhevig aan sprekerspecifieke variatie voor /p/. Hoewel de resultaten van dit onderzoek op zichzelf al interessant zijn, is de belangrijkste bijdrage van dit hoofdstuk methodologisch van aard. De data zijn geanalyseerd met behulp van functie-op-scalair regressiemodellen, met het hele spectrum als responsvariabelen. Deze methode leent zich goed voor de analyse van complexe, multidimensionale informatie in ruisgevulde spectra en geeft een goed beeld van hoe het spectrum in de loop van de tijd verandert in verschillende fonetische contexten.

Hoofdstuk 6 combineert twee corpusonderzoeken naar de realisatie van plofklanken in traditionele dialecten die gesproken worden in Jutland. De onderzoeken zijn gebaseerd op een ouder corpus van bandopnames, die een eerder stadium van taalvariatie laten zien. Het ene onderzoek betreft de voice onset time van plofklanken, d.w.z. de periode tussen het loslaten van de obstructie van de plofklank en het begin van stembandtrilling. Het andere onderzoek betreft variatie in de spectrale kenmerken van geaspireerde plofklanken. Het hoofdstuk bespreekt ook kort de aanwezigheid van stemhebbendheid in dialecten van Noord Jutland. Het lijkt erop dat stemhebbendheid in deze dialecten wijder verspreid is dan in het moderne Standaard Deens. In de plofklanken van Jutlandse dialecten begint stembandtrilling doorgaans later (d.w.z. ze hebben een hogere *voice onset time*) dan in het moderne standaard Deens. De Jutlandse dialecten laten complexe patronen zien, die verklaard kunnen worden door een samenspel van sociale en historische factoren. De mate van affricatie van /t/ is in Jutlandse dialecten onderhevig aan variatie, wat suggereert dat het Jutlands zich ten tijde van de opnames nog niet had aangepast aan het standaard Deens, waar, zoals in Hoofdstuk 5 werd besproken, /t/-affricatie consistent aanwezig is. De statistische analyse in Hoofdstuk 6 maakt gebruik van complexe dynamische variabelen. De regionale variabele wordt behandeld met zogenoemde *generalized additive mixed models* (GAMMs), en de spectrale variabele wordt behandeld met functioneel hoofcomponentenanalyse, die gebruikt wordt om soortgelijke problemen op te lossen als met functie-op-scalair regressie in Hoofdstuk 5 opgelost werden.

Hoofdstuk 7 geeft een samenvatting van de belangrijkste bevindingen van het proefschrift. Net als in Hoofdstuk 2 worden hier achtereenvolgens de diachrone ontwikkeling, de fonetische eigenschappen, de fonologische distributie en representatie, en de sociale en regionale variatie in de realisatie van plofklanken behandeld. Het hoofdstuk sluit af met een korte discussie waarin het perspectief van het proefschrift op data-analyse besproken wordt.

Curriculum vitae

Rasmus Puggaard-Rode was born Rasmus Puggaard Hansen in Varde, Denmark in 1992. He finished his secondary education at Varde Gymnasium in 2011. In 2011–2012, he spent a year working in the technical support department at Vestas Towers. He obtained his BA in Linguistics and China Studies from Aarhus University in 2016. During his BA studies, he spent one semester studying at Peking University as part of the European–Chinese Language and Culture Programme. He obtained his MA in Linguistics from Aarhus University in 2018. During his MA studies, he spent a semester at the Katholieke Universiteit Leuven, attending courses in the ResMA Advanced Studies in Linguistics. In 2018, he started his PhD at the Leiden University Centre for Linguistics. During his PhD studies, he spent two months in 2018 working at Aarhus University, hosted by the Peter Skautrup Centre for Jutlandic Dialect Research, and three months in 2020 working at The Arctic University of Norway in Tromsø, hosted by the Center for Advanced Studies in Theoretical Linguistics. In 2021–2022, he briefly worked part time as lecturer at the University of Amsterdam. In 2022, he accepted a position as a postdoctoral research scientist in the Spoken Language Processing group at the Institute of Phonetics and Speech Processing, Ludwig Maximilian University, Munich.