Open Library of Humanities

The rarity of intervocalic voicing of stops in Danish spontaneous speech

Rasmus Puggaard-Rode, Leiden University Centre for Linguistics, The Netherlands, r.p.hansen@hum.leidenuniv.nl

Camilla Søballe Horslund, School of Communication and Culture, Aarhus University, Denmark, horslund@cc.au.dk

Henrik Jørgensen, School of Communication and Culture, Aarhus University, Denmark, norhj@cc.au.dk

Previous studies of the phonetics of Danish stops have neglected closure voicing. Danish is an aspiration language, but the aspirated stops /p t k/ are produced with shorter closure duration and less articulatory effort than the unaspirated stops /b d g/. Furthermore, all Danish stops are characterized by some degree of glottal spreading during the closure. In this study, we use a corpus of Danish spontaneous speech (DanPASS) to investigate the intervocalic voicing—its distribution across the two laryngeal categories, whether it patterns as a lenition phenomenon, and whether the aerodynamic environment predicts its distribution. We find that intervocalic voicing is not the norm for either set of stops and is particularly rare in /p t k/. Voiced tokens are mostly found in environments associated with lenition. We suggest that the glottal spreading gesture found in all Danish stops is a phonological mechanism blocking voicing, which is probabilistically lost in spontaneous speech. This predicts our results better than relying on laryngeal features like [voice] or [spread glottis]. The study fills a gap in our knowledge of Danish phonetics and phonology, and is also one of the most extensive corpus studies of intervocalic stop voicing in an 'aspiration language.'

1. Introduction

All things considered, the phonetics of Danish stops are extremely well-described, largely due to the keen understanding shown early on in the writings of Otto Jespersen (e.g., 1899, 1906) and the gargantuan effort of Eli Fischer-Jørgensen and other researchers at the now-defunct Institute of Phonetics in Copenhagen. We have a good understanding of the timing of closure (Fischer-Jørgensen, 1954, 1972; Andersen, 1981) and release (ibid.; Fischer-Jørgensen, 1979; Mortensen & Tøndering, 2013), the spectral characteristics of releases (Fischer-Jørgensen, 1954), stop-induced F_0-perturbations (Fischer-Jørgensen, 1968; Jeel, 1975; Petersen, 1983), and muscular and laryngeal activity in stop production (Frøkjær-Jensen, Ludvigsen, & Rischel, 1971; Fischer-Jørgensen & Hirose, 1974; Hutters, 1985). One area where our understanding lags behind is closure voicing. It is generally agreed upon that Danish is an 'aspiration language' rather than a 'true voicing language'; the two-way laryngeal contrast in stops relies on aspiration, not closure voicing. Note that we aim to keep discussions of phonetics and phonology as separate as possible. 'Voicing,' 'closure voicing,' and similar terms are used interchangeably throughout the text to refer to the phonetic implementation of voicing only, while [voice] is used for the phonological feature.

There is no closure voicing in absolute initial position in Danish, and negligible voicing in final position. Voicing is less well-understood in intervocalic position. An oft-repeated claim has it that medial stop allophones in Danish are almost always voiced (Abrahams, 1949; Fischer-Jørgensen, 1954, 1980; Spore, 1965; Keating, Linker, & Huffman, 1983; Kingston & Diehl, 1994), and Fischer-Jørgensen (1954) writes that the first portion of intervocalic stops is generally voiced, as has been reported for many languages; although see also Jessen (2001), who assumes that Danish stops are systematically voiceless, seemingly in all positions. However, no empirical studies of closure voicing in Danish exist.

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 general 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 & Keating, 1986). Hence, voicing is often found in this position, even in languages where [voice] may not be phonologically active in 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 phonologically, where voicing generally corresponds to a [voice] feature, i.e., a more complex and marked structure.

In this paper, we present an empirical study of intervocalic stop voicing in Danish, based on an existing corpus of spontaneous speech (the Danish Phonetically Annotated Spontaneous Speech corpus, or DanPASS; Grønnum, 2009, 2016). This is an interesting venture due to a

number of observations about the production of Danish stops: The unaspirated set /b d g/ are produced with longer closure duration and greater muscular tension in the articulators than the aspirated set /p t k/ (Fischer-Jørgensen, 1954; Fischer-Jørgensen & Hirose, 1974). 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 investigations have shown that both stop types are characterized by a glottal opening gesture during the closure intervocalically in careful speech, but that this gesture lasts longer and is of greater magnitude in the aspirated set /p t k/ (Frøkjær-Jensen et al., 1971; Fischer-Jørgensen & Hirose, 1974; Hutters, 1985). This suggests that /b d g/ should be most conducive to voicing, and that voicing is actively blocked in both sets.

Our 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 all /b d g/ tokens. This rarity of intervocalic voicing is in essence the opposite conundrum of what we discussed above. Voicing *is* natural in this position, so its rarity can only be accounted for with reference to some mechanism that blocks voicing. The occurrence of intervocalic voicing is generally correlated with other variables that we associate 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 in itself is a lenition phenomenon. We suggest that this lenition is best modeled as gesture reduction: Danish has phonologized glottal opening gestures in all stops, which usually blocks voicing, but in some environments, this gesture can be lost. This happens more frequently with /b d g/, where the gesture has less of a critical function.

The paper is structured as follows: In the following subsections, we provide a background of closure voicing in phonetics and phonology with special focus on intervocalic position, and give an overview of the phonetics and phonology of Danish stops. In Section 2, we summarize our research questions and motivate all our independent variables. In Section 3, we provide an overview of our methods: We introduce the corpus we use and our data treatment. In Section 4, we provide an exploratory analysis of the data. In Section 5, we describe the selection of a logistic mixed-effects regression model and the results of that model. In Section 6, we discuss our research questions in light of our results, and in Section 7, we briefly summarize the findings.

1.1. Closure voicing and [voice] in stops

Closure voicing in stops is relatively 'unnatural' (e.g., Ohala, 1983). In order to maintain vocal fold vibration, sufficient transglottal pressure drop is required. As air continually flows from the lungs, supraglottal air pressure will quickly rise if both the oral and nasal cavities are sealed off. This means that it is impossible to maintain closure voicing for a long duration of time.

Ohala and Riordan (1979) claim that sufficient transglottal pressure drop can be maintained only for roughly 5–10 ms if the size of the supraglottal cavity remains constant. The size generally does not remain constant, though, as the vocal tract automatically enlarges due primarily 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 (ibid., 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 glottis and point of occlusion, which yields a slower build-up of air pressure and crucially yields a larger total area of soft, expandable cavity walls. Keating (1984a) 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 an active process of vocal tract enlargement during the occlusion, such as jaw lowering or velum raising.

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-initially, closure voicing is articulatorily unnatural, since subglottal and supraglottal air pressure will rise roughly synchronously unless the vocal folds are initially fully abducted to allow for a build-up of subglottal air pressure. Closure voicing is also unnatural syllable-finally due to an inspiratory force that gradually but quickly counteracts the initially high subglottal pressure from the preceding vowel. However, it *is* natural for a significant portion—usually most—of intervocalic stop closures to be voiced due to the high initial subglottal pressure following the preceding vowel.

Articulatory naturalness does not always translate directly into attested typological patterns. On the one hand, 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, 1983; 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, p. 103ff.). On the other hand, in spite of their unnatural status, syllable-initial voiced stops are actually quite common. Furthermore, voicing is most articulatory 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 phonetic and phonological patterning than ease of articulation.

Below, we will characterize three general approaches to the representation of laryngeal contrasts in the phonological literature, and the predictions they make with regards to intervocalic stop voicing. There is a huge literature on the topic, so some approaches will necessarily be missed, while others may be grouped together even if they differ in some respect. We 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 quite narrowly to the presence of voicing during closure, which is what we refer to as concrete [voice]. This is how [voice] is conceptualized in the laryngeal feature geometry of Lombardi (1995) and the 'laryngeal realism' approach of Iverson and Salmons (e.g., 1995). These are approaches that assume a direct relationship between different physical laryngeal constellations and phonological laryngeal features. Such approaches usually assume that languages with aspiration contrasts employ an active feature [spread glottis] to distinguish between laryngeal stop series. 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 it cannot spread from those. One possible solution to this is Rice and Avery's (1989) proposal of a non-laryngeal [spontaneous voice] feature node, which can spread from sonorants to obstruents. Another solution is to simply relegate intervocalic voicing to phonetic implementation, placing it outside the purview of phonology. This would predict that intervocalic stops that are unmarked for laryngeal features always follow the phonetically natural pattern.

Jessen and Ringen (2002) and Beckman, Jessen, and Ringen (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 their closure intervocalically with very few exceptions, while German /b d g/ are variably voiced intervocalically (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 level in the phonological derivation, segments are assigned numerically-valued features; the degree of intervocalic voicing in a [spread glottis] language will depend on the values assigned to [spread glottis] at this later stage.[1] It should be noted that the findings of Beckman et al. can only be taken as evidence for underlying features if one assumes a transparent relationship between phonology and phonetics; see e.g. Keating (1984b) for a general critique of this stance.

In abstract [voice] approaches, the feature needs not directly refer to closure voicing. Chomsky and Halle (1968) and Keating (1984b) both assume that [voice] can refer to either stops with closure voicing or stops where voicing begins 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. This argument partially relies on the

---

[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.

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 naturalness of closure voicing; an 'automatic phonetics' will output initial stops with no closure voicing and intervocalic stops with closure voicing, while a 'controlled phonetics' is necessary to divert from that pattern.

If we expect a direct relationship between phonetics and phonology, then there should be a correspondence between phonologically and phonetically unmarked material. Given the aerodynamic account of stop voicing given above, this means that a phonologically unmarked stop should be voiceless initially and voiced intervocalically. It also means that phonetic reduction will be positionally defined: Devoicing of [voice] stops is a lenition phenomenon syllable-initially, whereas voicing of stops without [voice] is a lenition phenomenon 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). In an optimality-theoretic analysis of this problem, Smith (2008) proposes constraints militating against voiced obstruents in onset position and voiceless obstruents in intervocalic position, which compete with faithfulness constraints (see also Hayes, 1999).[4]

Gesture-based approaches of phonological representation can straightforwardly account for these positional markedness relations. One such approach is Articulatory Phonology (Browman & Goldstein, 1986, 1992). In Articulatory Phonology, articulatory gestures are taken as the primary units of phonological representation rather than segments or features. A consequence of this 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 result in voicing intervocalically, as per the discussion above.

These are a few predictions about the patterning of intervocalic stop voicing based on 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 represent aspiration. From a concrete [voice] perspective, we would predict essentially categorical intervocalic voicing of all stops in 'true voice' languages, since [voice] ensures voicing in one category, and there are no available phonological mechanisms to counteract voicing in the other (unmarked) category. We would predict varying degrees of intervocalic voicing of unmarked stops in 'aspiration' languages, and very little voicing in [spread glottis]

---

[2] The cause of this pattern is disputed. According to Kingston and Diehl (1994), F_0 is lowered by stops with [voice], and according to Hanson (2009), F_0 is raised locally by voiceless stops.

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

[4] Katz (2016) points out some typological shortcomings of this account.

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, but 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. Below, we will discuss the literature on voicing and laryngeal representation in Danish stops, and how this relates to these predictions.

1.2. Danish stops

Standard Danish has six phonemic stops, /b d g p t k/. A common analysis of Danish holds that there are strong and weak syllabic positions: Strong position refers to onsets before full vowels, while weak position refers to codas and onsets before neutral vowels (e.g., Jakobson, Fant, & Halle, 1951).[5] This determines the positional allophone of many consonants, including the stops: /b d g/ are voiceless unaspirated in strong position and manifest as approximants or zero in weak position. /p t k/ are voiceless aspirated in strong position and voiceless unaspirated in weak position. This analysis goes back to e.g., Uldall (1936) and Jakobson et al. (1951) and is more fully fleshed out by e.g., Rischel (1970) and Basbøll (2005). A detailed account of the proposal is beyond the scope of this paper, but we argue elsewhere that it is outdated in some regards (Horslund, Jørgensen, & Puggaard, 2021; Horslund, Puggaard-Rode, & Jørgensen, 2022). As such, in strong position, the laryngeal contrast is based on aspiration, whereas in weak position, only /p t k/ are actually realized as stops. When we refer to /b d g/ and /p t k/ in the following, we do not refer to this rather abstract analysis, but rather to something closer to the surface contrast: /b d g/ refer to stops that would be unaspirated in distinct speech, and /p t k/ refer to stops that would be aspirated in distinct speech. We will refer to the two series as *laryngeal categories*.

Similar to some traditions of English transcription, /b d g/ are in narrow transcription usually given as [b̥ d̥ g̥], indicating that they are voiceless but phonetically lenis. /p t k/ are usually transcribed phonetically as [pʰ tˢ kʰ], with the superscript *s* indicating salient affrication of /t/. In neat transcription, they are sometimes transcribed as [b̥ʰ d̥ˢ g̥ʰ] (Grønnum, 1998), indicating that although aspirated, both sets are equally phonetically lenis. The terms fortis–lenis do not have clearly defined phonetic correlates, so this deserves further unpacking.

---

[5] Neutral vowel here refers to schwa as well as approximants that are syllabic due to schwa-assimilation, as well as unstressed [i] in some morphological contexts.

The terms fortis and lenis are used in quite distinct ways in the phonetic and phonological literature. One use is as an essentially arbitrary label for stop contrasts in languages where the said contrast does not depend on 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, Ladefoged, & Maddieson, 1992). Another use is as a phonetically substantial phonological feature referring to force of articulation. This in turn may correlate with pulmonic force, muscular tenseness of the articulators, closure duration, or indeed closure voicing (Jaeger, 1983, and references therein). Either use of the terminology is usually too imprecise for a phonetic or phonological description of a group of sounds. Fischer-Jørgensen (1972) suggests that Danish /b d g/ are in fact fortis and /p t k/ lenis, since the closure duration is longest for /b d g/ and results from electromyographic investigations show higher organic pressure for /b/ than /p/ (Fischer-Jørgensen & Hirose, 1974).

Regarding the [b̥]-style notation of lenis voiceless stops, this is only briefly mentioned in the *Handbook of the International Phonetic Association* (International Phonetic Association, 1999, p. 24): "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 certainly does not apply to Danish, where stops from both laryngeal series are voiceless (nor is it clear that it applies to English stops, for that matter). The IPA has no way of indicating a fortis–lenis distinction, so it is surprising that indicating lenis should take precedence over transparent representation of voicing in the transcription of /b d g/ for Grønnum and Basbøll, *especially* considering their claim that Danish stops do not actually show a fortis–lenis contrast.

Overall, little has been written about closure voicing in Danish stops, and to our knowledge, no quantitative studies have been made of the topic. In essence, what we know from the existing literature is that all stops show some degree of voicing during the first portion of the closure when they occur between other voiced sounds (Fischer-Jørgensen, 1954), and that intervocalic word-medial stops are continuously voiced categorically or near-categorically (Abrahams, 1949; Fischer-Jørgensen, 1954, 1979, 1980; Spore, 1965; Keating et al., 1983; but see also Jessen, 2001, and Beckman et al., 2013, who assume that Danish stops are categorically voiceless). It is generally assumed that /b d g/ were voiced in previous stages of the language, and according to Brink and Lund (2018), this was lost sometime before 1700.

Articulatory studies of carefully read speech have shown that intervocalically before stressed syllables, both /b/ and /p/ are accompanied by a glottal opening gesture during the closure (Frøkjær-Jensen et al., 1971; Hutters, 1985), although the opening gesture varies significantly in magnitude and timing. Similar studies of English found no such gesture during /b/ (Sawashima, 1970; Hirose & Gay, 1972); in Icelandic, which has a contrast between unaspirated and pre-aspirated stops intervocalically, both series of stops have a significant glottal opening gesture (Pétursson, 1976). Frøkjær-Jensen et al. (1971) hypothesized that the gesture in Danish /b/ is

an artifact of the articulatory transition from vowel to consonant, while in /p/ it is "effectuated by neural commands" (ibid: 134). However, electromyographic studies by Fischer-Jørgensen and Hirose (1974) and Hutters (1985) show that the posterior crico-arytenoid muscles are active in achieving the glottal opening during /b/.[6] Hutters (1985) proposes that the intervocalic glottal opening gesture is a measure taken to reinforce voicelessness in /b/, although she leaves the question relatively open; more recently, 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.

From a phonological perspective, Iverson and Salmons (1995) and Basbøll (2005) assume that [spread glottis] is the laryngeal feature that distinguishes between the two sets of stops, and that [voice] plays no role in distinguishing between Danish stops. [spread glottis] is taken to be an active feature, which causes devoicing of following sonorants. This process has usually been taken for granted, but a recent paper by Juul, Pharao, and Thøgersen (2019) shows that devoicing of sonorants following aspirated stops is much less categorical in Danish than usually assumed.

Kingston and Diehl (1994) assume that [voice] (in the abstract sense discussed in Section 1.1 above) distinguishes between the two sets. The account then holds that [+ voice] is only implemented as actual voicing in Danish when voicing is phonetically natural (i.e., intervocalically). A key motivation for this representation is that F_0 is lowered by [+ voice] stops. This is not exactly straightforward in Danish; Fischer-Jørgensen (1968) finds no evidence for this, while Jeel (1975) and Petersen (1983) both do. However, while Petersen does find that the different stop series trigger different F_0-perturbations, he crucially finds that both series trigger high initial F_0 in following vowels relative to nasals. As Goldstein and Browman (1986) point out, this is consistent with an account where F_0-perturbations follow directly from glottal aperture, something that Kingston and Diehl (1994) explicitly reject. Nevertheless, Kingston and Diehl's dichotomy between automatic and controlled phonetics (see Section 1.1) can potentially account for both the presence and absence of closure voicing in stops in Danish.

As mentioned in Section 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 seems to assume that intervocalic voicing is categorical or near-categorical. Some say that muscular tension is overall low in Danish stops, increasing the chances of voicing; but all Danish stops are also characterized by a glottal opening gesture, decreasing 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 /p t k/ also have a glottal opening gesture of greater magnitude.

---

[6] In Frøkjær-Jensen, Ludvigsen, and Rischel (1973), which is a reprint of Frøkjær-Jensen et al. (1971), they recognize the forthcoming work of Fischer-Jørgensen and Hirose (1974) and note that it throws doubt on their explanation of the findings.

The results may allow us to compare some of the predictions from different approaches to phonological laryngeal specification. If [spread glottis] is indeed the only active laryngeal feature, we would predict at most variable voicing in /b d g/, and little voicing in /p t k/ (following Beckman et al., 2013). If the laryngeal contrast is maintained with phonologized glottal gestures (as in Articulatory Phonology), we would assume that the two series have underlying glottal opening gestures of different magnitudes, both of which are expected to counteract voicing. This gestural account predicts that lenition leads to a reduction in the magnitude of these gestures, potentially causing voicing in either 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.

2. Research questions and potential predictors

This paper is partly hypothesis testing, and partly exploratory in nature. We set out with the following research questions (RQ) in mind:

RQ1: 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 (Fischer-Jørgensen, 1972) and have been alleged to be lenis. However, there is evidence that the vocal folds are not in a neutral position; for both series of stops, although to varying degrees, the vocal folds are actively spread during the closure. Our primary hypothesis is 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.

RQ2: Can closure voicing in Danish stops be considered a lenition phenomenon?

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

RQ3: 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 & Möbius, 1998; Möbius, 2004; Strycharczuk, 2012). We aim to take as many of these into account as possible in order to explore their relative influence in Danish. These factors are presented in detail below.

2.1. Potential predictors

The detailed annotations of the DanPASS corpus (see Section 3.1) allowed us to test how a large number of mostly categorical predictors may affect closure voicing. These predictors relate to segmental, prosodic, morphosyntactic, and other factors, which are discussed in the following subsections. Variables are capitalized when they are first mentioned.

2.1.1. Segmental predictors

We coded the stops themselves according to Laryngeal Category and Place of Articulation. There is really no theory-neutral way to refer to the two laryngeal stop series. Here, we use 'aspirated' and 'unaspirated' as short-hand terms for the surface contrast between /p t k/ and /b d g/ in distinct speech, as discussed in more detail in Section 1.2.

We expect place of articulation to influence the likelihood of voicing, such that occlusions further back in the oral cavity reduce the chance of voicing. This is aerodynamically motivated (see Section 1.1 above 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). Since bilabial and alveolar occlusions are physically quite close, and velar occlusions are significantly further back, we assume that a place effect will be most noticeable for velar stops.

The quality of surrounding vowels is expected to have an influence on the likelihood of closure voicing; note that Danish has an exceptionally complex vowel system (see Grønnum, 1995). We expected surrounding High Vowels 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, 2014). High vowel devoicing happens because a constriction in the oral cavity makes it difficult to keep the subglottal air pressure high enough to maintain voicing over time; this means that a preceding high vowel should decrease the odds 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 (Grønnum, 1998); many of these vowels are higher than their conventional IPA counterparts, and they all have mean F1 < 400 Hz. in modern Standard Danish (Juul, Pharao, & Thøgersen, 2016).

In locating intervocalic stops, Approximants were also considered vowels. We assume that approximants occurring immediately before the intervocalic stop decrease the chances of voicing, simply because approximants are less sonorous than nuclear vowels (e.g., Parker, 2002). The approximants in question are [j ɹ� ʋ ɒ̯ ɣ].[7] These are all frequently syllabic due to processes of schwa assimilation.

---

[7] [ɣ] is usually transcribed with [ð], in spite of the sound being highly vocalic (Juul et al., 2016; Brotherton & Block, 2020). Here, we follow Schachtenhaufen (2020–) in using a vocalic transcription instead. The sound is generally considered the weak allophone of /d/.

As discussed above, there is reason to assume that intervocalic voicing in Danish is a lenition phenomenon resulting from voicing continuing from the preceding vowel lasting throughout the closure. Therefore, we expect voicing to be more likely in environments that are generally associated with weakening. We expected surrounding Neutral Vowels to increase the chances of voicing, since the Danish neutral vowels [ə ɐ] generally occur in prosodically weak syllables (e.g., Basbøll, 2005), where we strongly expect lenition. 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 we expect them to have roughly the same influence on closure voicing.

2.1.2. Prosodic predictors

We expected Stress on the syllable in question to reduce the chances of voicing, since stress generally reduces the chances of lenition phenomena occurring. If the preceding syllable has stress, we expect this to increase the chances of voicing, as it is unlikely for two syllables in a row to carry stress.

We expected the prosodic laryngealization phenomenon Stød to reduce the chance of voicing when adjacent, no matter whether on the preceding syllable or the syllable in question.[8] Stød is akin to creaky voice, i.e., low pitch and relatively aperiodic voicing, occurring on the final part of a long sonorant rhyme (Grønnum & Basbøll, 2001, 2007). Stød is produced with laryngeal contraction, due in particular to activity of the vocalis and lateral crico-arytenoid muscles (Fischer-Jørgensen, 1987, 1989). Recall from Westbury and Keating (1986; see Section 1.1) that closure voicing is natural intervocalically, assuming the vocal fold configuration is amenable to voicing; this is the case for vowels with modal voicing and less so the case for vowels with stød. As such, we expect stød on the preceding syllable to decrease the chances of continuous voicing. Although stød mainly affects the final part of syllables with a long sonorant rhyme, it is also cued 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.

2.1.3. Morphosyntactic predictors

We coded the type of Morphological Boundary at which the intervocalic stop occurred. These include word boundaries, boundaries between roots and (derivational and inflectional) affixes,

---

[8] 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 (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 our data; in fact, a small majority of syllables with stød are unstressed.

boundaries between separate parts of compounds, as well as no boundary if the intervocalic stop occurred morpheme-internally. It should be noted here that prefixes in Danish are exclusively derivational, while suffixes are mostly inflectional but can also be derivational. As consonants tend to be strong domain-initially (e.g., Keating, Cho, Fougeron, & Hsu, 2004), it would be more optimal to have the individual syllables coded for their position in a prosodic hierarchy (e.g., Nespor & Vogel, 1986), but such a coding cannot be easily extracted from the existing DanPASS transcriptions. We hypothesize that our 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, Kaun, Narayanan, & Saltzman, 2000; Cho, 2001). We 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, we assume the following hierarchy of morpheme boundary types: inflectional > derivational > compound.[9] There are reasons to assume that stops at inflectional morpheme boundaries might be voiced at much higher rates than stops in other positions: They are always unstressed, and they always have neutral vowels in Danish. Following a usage-based framework such as Exemplar Theory (e.g., Bybee, 2001), inflectional affixes may also be voiced intervocalically more often simply because language users encounter voicing more often in affixes, and as such it is weighted as more likely in the underlying representation of these, at a morpheme-specific level. Several phonological frameworks (e.g., Lexical Phonology; Kiparsky, 1985) assume that morphology is invisible to phonetic interpretation and would thus predict morpheme-specific underlying representations to be impossible. However, recent studies show that specific morphemes *can* exhibit phonetic patterns that are not predictable from their phonemic makeup. Plag, Lohmann, Hedia, and Zimmermann (2020) and Tomaschek, Plag, Ernestus, and Baayen (2021) have found that the English 'homophonous' 's'-suffixes (third person singular present tense, plural, etc.) differ systematically in phonetic realization; an example would be the suffixes in the present tense verb 'pet-s' and in the plural noun 'pet-s.' Heegård (2013) found similar results for variable rates of schwa-deletion in homophonous -/tə/ suffixes in Danish.

Additionally, we also coded words for being a member of either a Closed or an Open word class. Words from closed classes are often function words, and it is well-known that these often show significant phonetic reduction (e.g., Bell et al., 2003; Schachtenhaufen, 2013).

2.1.4. Other predictors

In addition to the predictors already mentioned above, we included a lexical frequency measure. Lexical frequency is known to cause phonetic reduction, both in the course of language change

---

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

(e.g., Hooper, 1976; Bybee, 2000a) and synchronically (e.g., Bybee, 2000b; Pierrehumbert, 2001; Pluymaekers, Ernestus, & Baayen, 2005) and has been shown to specifically increase voicing assimilation in Dutch (Ernestus, Lahey, Verhees, & Baayen, 2006). Although the speech in DanPASS is spontaneous, it remains nested within specific experiments, where specific lexical frequencies can be quite different from what is found in language use in general (see Section 3.1). Since contextual probability has likewise been shown to increase phonetic reduction (Jurafsky, Bell, & Girand, 2001; Jurafsky, Bell, Gregory, & Raymond, 2002), we coded 'Local' Lexical Frequency, i.e., lexical frequency in the DanPASS corpus itself, which is available in the online version of DanPASS (Grønnum, 2016). We compared this with a more general measure of lexical frequency based on the much larger LANCHART corpus (Language Change in Real Time; see Pharao, 2009, p. 145ff.), which includes just over 3 million words. However, due to the experimental nature of DanPASS (the map task in particular; see Section 3.1), many of the DanPASS words do not occur in LANCHART. This means that modeling with general (LANCHART) frequencies rather than local (DanPASS) frequencies would require us to exclude just over 300 items, around 8% of our 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, we focus only on local frequency in our modeling.

We also included a local measure of speech rate. 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 1.1). A higher speech rate also causes a shorter occlusion (as demonstrated for Danish by Andersen, 1981), 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.[10]

We also coded the Individual Words, 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 2.1.3. We do not explore word-specific effects in any detail.

Finally, we coded for a few extralinguistic factors pertaining to the speakers. Sex has been shown to have an influence on closure voicing, such that men are more likely than women to produce fully voiced stops (Swartz, 1992; Ryalls, Zipprer, & Baldauff, 1997). This could be aerodynamically motivated; on average, men have larger supralaryngeal cavities than women, making them likely to maintain closure voicing over longer stretches of time. An alternative explanation for the same outcome is that women generally speak more 'clearly' than men (as

---

[10] This is admittedly a rough measure of speech rate, chosen mostly out of convenience (it was easy to extract from the existing data frame). 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. We find a very strong effect of speech rate, so we assume it is not too rough for our purposes.

demonstrated by e.g., Ferguson, 2004 for vowel intelligibility), and show less of a tendency for lenition; note that this is likely an effect of gender rather than biological sex. We are 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.

We also coded the Individual Speakers. Sonderegger, Stuart-Smith, Knowles, Macdonald, and Rathcke (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, Sonderegger, and Stuart-Smith (2020) find similar results for Japanese. We do not explore speaker-specific effects in any detail.

The potential predictors and the directionality of their expected influence on closure voicing are summarized in **Table 1**.

Variable	Predicted likelihood of voicing	Notes
Laryngeal Category	unaspirated > aspirated	
Place of Articulation	bilabial > alveolar > velar	strongest effect for velar stops
Adjacent Approximant	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 > derivational > 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	decreases with age	
Individual Speaker	random	

Table 1: Potential predictors and the expected directionality of their influence on closure voicing.

3. Methods

3.1. Corpus

In order to answer our research questions, we used the DanPASS corpus (Grønnum, 2009, 2016). The corpus consists of native speakers of Danish solving a number of unscripted tasks, either alone or in pairs. An original motivation behind establishing the corpus was to counteract the bias for highly controlled scripted speech in phonetic investigations of Danish. The recordings in the DanPASS corpus are unquestionably also laboratory speech, but they are of a much less formal nature than what was previously the standard. Grønnum (2009) rightly points out that highly 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 voicing of plosives, where the phonological 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), phonetic reduction in general (Schachtenhaufen, 2013), as well as intonation and prosody (Tøndering, 2003, 2008; Grønnum & Tøndering, 2007). It also served as the basis for the most thorough investigation of (positive) voice onset time in Danish stops (Mortensen & Tøndering, 2013). A complete list of publications using DanPASS can be found in Grønnum (2016).

The full DanPASS corpus consists of a number of monologues recorded in 1996, and a number of 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 and 68 years of age, with a mean age of 29 years. Overall, the monologues constitute 171 minutes of speech, with a mean duration of 9m27s of speech per speaker (range 6m13s – 15m49s). Technical details about the recordings can be found in Grønnum (2009). The speakers were recorded performing three different tasks: *Network* is a description of various shapes and various 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 & Weenink, 2019). 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 quite neatly defined phonological unit in Danish (see e.g., Basbøll, 2005), it is often difficult to find neatly defined phonetic units corresponding to those (Schachtenhaufen, 2010), particularly because schwa-assimilation processes cause syllabic sonorants and consecutive syllables consisting of homorganic vowels to be abundant in Danish speech. 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; here, [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.[11] This decision effectively means that closure voicing during stops is ignored in the transcription. Grønnum (2009) does not motivate this, and perhaps as a result of this, later studies using DanPASS to examine variation in e.g., stops (Pharao, 2009, 2011; Mortensen & Tøndering, 2013; Schachtenhaufen, 2013) also ignore the distinction between stops with and without closure voicing.

3.2. Acoustic analysis

We used a Praat script to find stops that occur intervocalically in the DanPASS monologues, i.e., stops that do not occur initially in a prosodic phrase and are flanked on both sides by either vowels or central approximants. Approximants were included because there are well-defined phonological processes whereby they syllabify (Basbøll, 2005) and thus become phonetic nuclear vowels. The DanPASS transcriptions are segmented at the level of phonetic syllable, and we included only intervals that were stop-initial. This is in line with the studies of glottal activity cited in Section 1.2 above, which also focus on syllable-initial intervocalic stops, and this decision makes the coding and interpretation of predictor variables considerably easier. For each of these stops, the surrounding syllables are isolated, i.e., the stop-initial syllable and the preceding syllable. The script then 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 intervocalic stops, with an average of 204.7 stops per speaker (range 117–341). They are broken down by phonemic stop category in **Table 2**.

Phoneme	Number	Range per speaker
/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

Table 2: Intervocalic stops in the DanPASS monologues by phonemic category.

---

[11] However, tapped realizations of /t ~ d/ are marked as such, i.e., with [ɾ].

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). The variance in lexical frequency is rather extreme. There are 125 lexical items which occur just once, while the 10 most frequent items occur a total 2,025 times.

For each of the intervocalic stops, we manually checked if it was voiced throughout the closure. This was done on the basis of visual inspection of the waveform: Constant periodicity up to the burst was taken as continuous closure voicing. Whenever stops from the /p t k/ series were fully voiced, they typically also had breathy voiced release. This method proved to be relatively straightforward to implement, although it is certainly a simplification of the complexity in the phonetic signal. **Figures 1–4** show waveforms of stops from both laryngeal series that show continuous voicing and interrupted voicing, respectively.
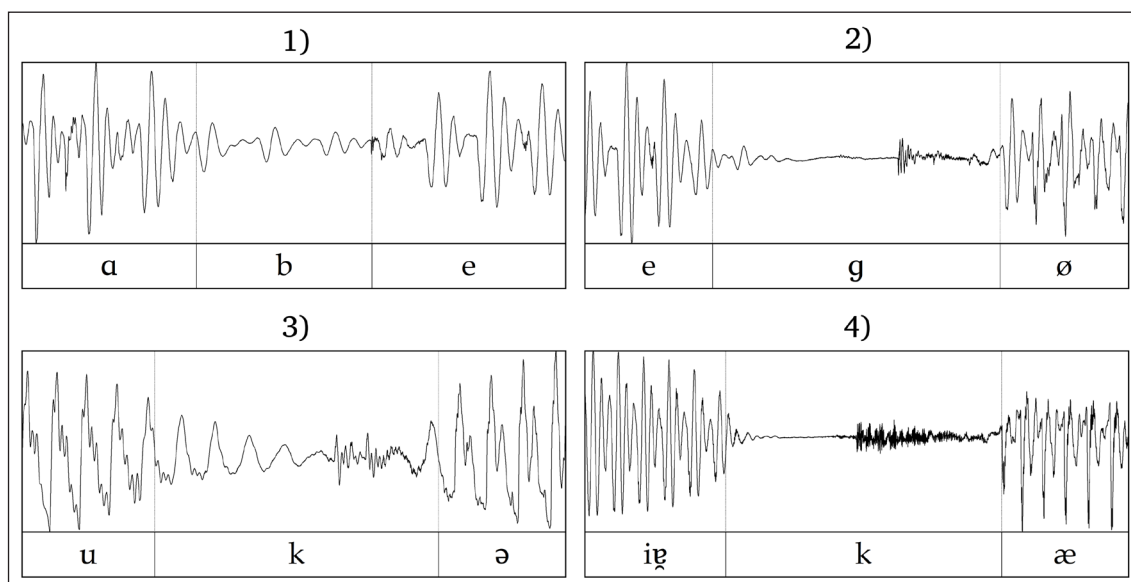
Figures 1–4: Waveforms exemplifying: 1) A fully voiced token of /b/ in the phrase <fr(a be)ˈgyndelsen> 'from the start.' 2) A mostly voiceless token of /g/ in the same phrase as 1, <fra b(eˈgy)ndelsen>. 3) A fully voiced token of /k/ from the phrase <d(u ka)n> 'you can'. 4) A mostly voiceless token of /k/ from the word <ˈf(irka)nt> 'square.'

Recent 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 difficult to model than binary variables, and 2) fully voiceless intervocalic stops

---

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

are known to be uncommon, so a 'voiceless' category would likely have added little explanatory value. In intervocalic position, voicing from the vowel essentially always continues into 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 & Möbius, 1998), German (Möbius, 2004), and American English (Davidson, 2016) aspirated stops and for voiceless stops in several other languages (Shih, Möbius, & Narasimhan, 1999). In an unpublished conference paper, Puggaard, Grijzenhout, and Botma (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 intervocalically; both /b/ and /p/ were voiced for the first approximately 25% of their closure duration. They 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 an argument for considering 'true voicing' to only mean *continuous voicing throughout the closure* in this study. Incidentally, the above-mentioned studies of 'aspiration languages' found that the voiceless unaspirated German /b d ɡ/ are most often fully voiced intervocalically, while around half of the Standard Chinese (voiceless unaspirated) /b d ɡ/ tokens are voiced throughout intervocalically.

Ideally, we would be working with a continuous measure of closure voicing, possibly measuring both intensity and relative duration of voicing. However, this would require much more fine-grained segmentation of the sound files, and we did not have the resources to add these to the existing annotations. It is quite possible that true effects of some lower-level variables on voicing are masked in this study because of our relatively rough voicing measure.

3.3. Statistical analysis

All statistics used in the current study were calculated using the R statistical environment (R Core Team, 2021; RStudio Team, 2022) and a number of add-on packages.[13] We are interested in both exploratory analyses and confirmatory analyses. The precise methods for our statistical analyses are described in Sections 4 and 5.1 below, respectively.

4. Exploratory analysis

In this section, before proceeding to building a regression model, we take a closer look at the data and explore correlations between the individual predictors and the presence of intervocalic voicing.

---

[13] We fitted logistic mixed effects models using the *lme4* package (Bates, Mächler, Walker, & Bolker, 2015, 2021). We also used the *car* package for calculating variance inflation factors (Fox & Weisberg, 2019; Fox, Weisberg, & Price, 2021), the *MuMIn* package for calculating model effect size (Barton, 2020), and the *moments* package for checking distributions of continuous variables (Komsta & Novomestky, 2015). We used the *ggplot2* package 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. (2022).

4.1. Categorical predictors

Table 3 and **Figure 5** show the proportions of voiced tokens for each level of each of our categorical variables. The majority of categorical variables show at least some degree of correlation with closure voicing in the direction we predicted in **Table 1** above.

Variable	Level	% voiced	Number voiced	
Laryngeal category	Aspirated Unaspirated	5.05 38	77 844	√
Place of articulation	Bilabial Alveolar Velar	17.25 35.75 14.55	89 611 221	÷
Preceding approximant	Absent Present	25.81 17.08	832 89	√
High vowel	Absent Present	22.38 29.69	584 337	÷
Preceding high vowel	Absent Present	25.18 22.38	748 173	√
Neutral vowel	Absent Present	22.31 43.94	747 174	√
Preceding neutral vowel	Absent Present	26 22.23	612 309	÷
Stress	Absent Present	26.13 20.51	712 209	√
Preceding stress	Absent Present	25.31 23.15	637 284	÷
Stød	Absent Present	26.39 17.36	792 129	√
Preceding stød	Absent Present	25.26 5.56	914 7	√
Morphological boundary	Internal Inflection Derivation Compound Word	36.12 68.75 38.81 9.76 24.61	95 110 26 73 617	÷
Word class type	Open Closed	19.33 31.38	407 514	√
Sex	Female Male	20.96 25.75	192 729	√

Table 3: Table of proportion of fully voiced tokens for each level of each categorical variable. Variables marked √ show correlations in agreement with our hypotheses in Table 1, and ones marked ÷ show disagreement with our hypotheses.
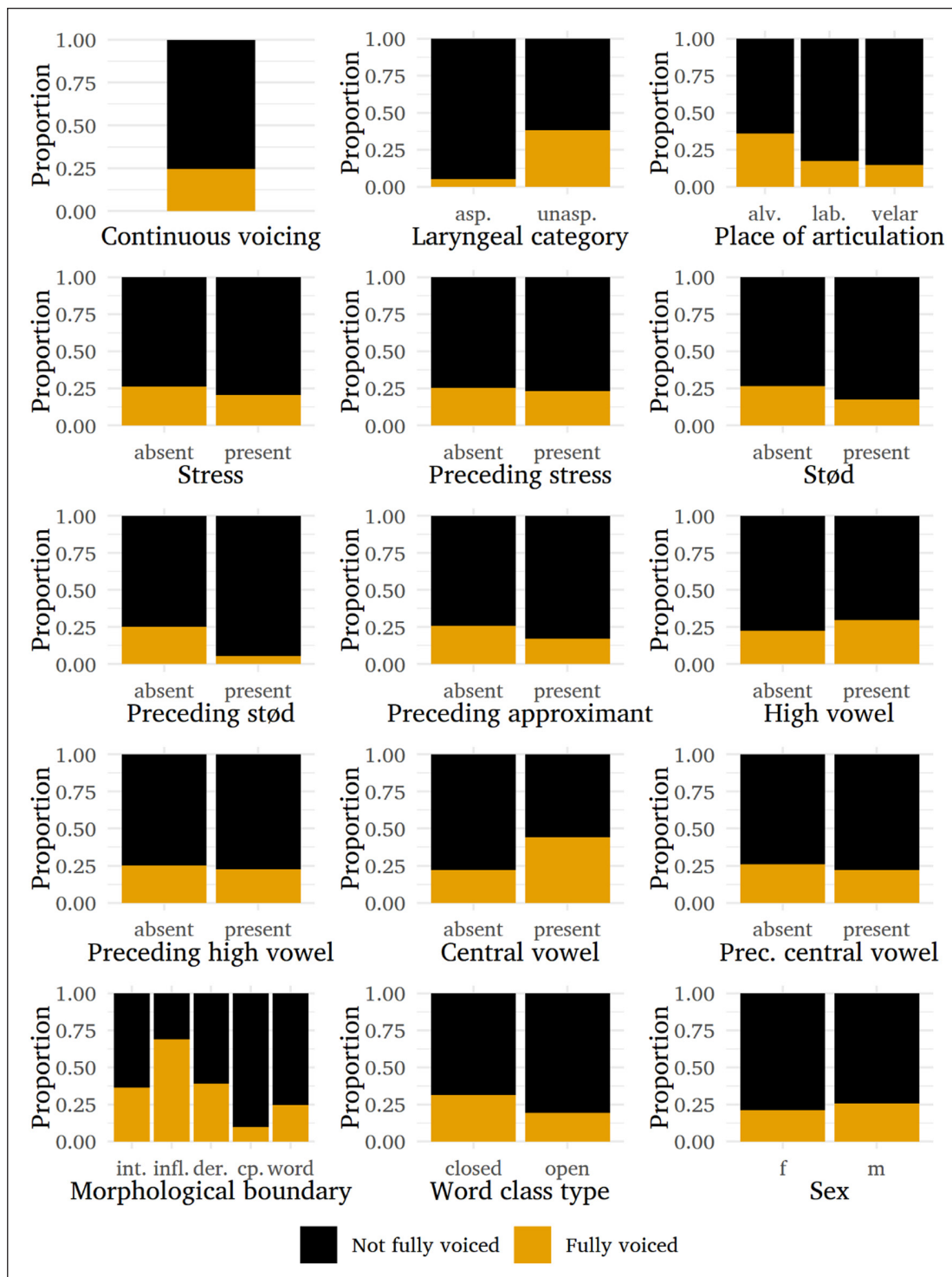
Figure 5: Stacked bar plots showing the proportions of tokens with and without continuous voicing for each level of each categorical variable. (Morphological boundary levels = internal, inflectional, derivational, compound, word).

4.1.1. Segmental predictors

Laryngeal Category shows a clear correlation in the expected direction. As we 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 this is sometimes described 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. Presumably, there are non-aerodynamic reasons for this. Alveolar stops are generally more frequent than other places of articulation, and they are found at a higher rate in function words. While the transcriptions do in principle indicate tapped realizations of the alveolar stops, it is also likely to be somewhat inconsistent, such that some realizations that are transcribed as alveolar stops are in fact alveolar taps [ɾ]; these are of course always voiced.

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

The behavior of High Vowels goes against our predictions; we expected high vowels to decrease the chances of voicing, in particular preceding the stop. In fact, high vowels preceding the stop just show a weak correlation in the expected direction, and high vowels in the same syllable correlate positively with voicing. This is contrary to our aerodynamically motivated predictions but could have a number of other explanations: High vowels are found in a number of very frequent function words; [ɪ ʊ] are both included in this group, and they are derived from underlying sequences of the approximants [j̞ w̞] assimilating with schwa. As such, there are predictable reasons why we might expect syllables with high vowels to 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.

4.1.2. 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.

4.1.3. Morphosyntactic predictors

Our 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 as a whole. Finally, stops at compound boundaries are rarely voiced. Given the complexity of this factor, we will hold off on interpreting these results further until we 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.

4.1.4. Other predictors

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

4.2. Continuous predictors

Having discussed all categorical predictors, we now turn to the continuous ones. **Figure 6** visualizes the proportion of stops with and without continuous voicing with density plots.
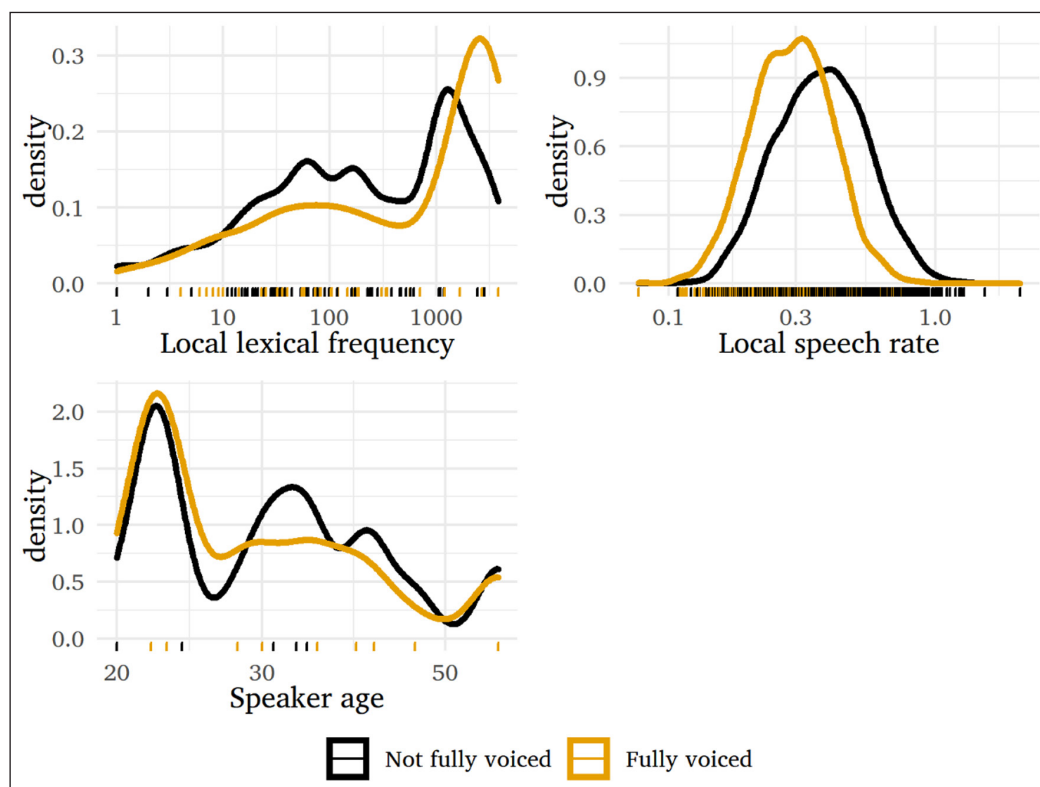
Figure 6. Density plots showing the tokens with and without continuous voicing relative to continuous variables on a log-scale.

It is clearly (and logically) the case that there are most tokens of the most Frequent words in both the voiced and voiceless group. It is also clearly the case that the words with very high frequency show a higher proportion of voiced tokens, and similarly that the words with medium frequency, particularly between 50–500, show a higher proportion of voiceless tokens.

As predicted, Speech Rate clearly 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 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 empirical data, we will now move on to analyzing the data with mixed-effects regression modeling.

5. Confirmatory analysis

5.1. Model selection

Our data comes from a corpus that was not collected for our purposes, and we are interested in quite many independent variables. Given the lack of experimental control and the partly exploratory nature of the study, our data is presumably not structured in a way that allows us to retain a maximal random effects structure; this is a common problem with mixed-effects models in linguistics (Meteyard & Davis, 2019). This loss in optimal data structure is also 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 opinions ranging from maximizing random effects (Barr, Levy, Scheepers, & Tily, 2013) to balancing statistical power and Type I errors by including only random effects that contribute sufficiently to the model's predictive power (Bates, Kliegl, Vasishth, & Baayen, 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). These papers generally assume 1) experimental data, and 2) a continuous dependent variable (i.e., linear mixed-effects models). This is important because experimental data is ideally more balanced than ours, and linear models are overall more likely to converge than logistic models with binary dependent variables (Seedorff, Oleson, & McMurray, 2019). We opt for a data-driven model selection procedure, largely inspired by the heuristics proposed by Sonderegger (2022).

---

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

The raw values of all our 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, Vasishth, Hohenstein, & Kliegl, 2020 for an introduction to this). We coded sum contrasts 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, we coded two theoretically-guided Helmert contrasts: One to test the distinction between velars and non-velars, and one to test the distinction between alveolars and labials:

1.   Velar contrast: –⅓ bilabial, –⅓ alveolar, +⅔ velar
       Bilabials versus alveolars: +½ alveolar, –½ bilabial

The five-level Morphological Boundary variable is rather complicated. Here we coded four theoretically-guided Helmert contrasts: 1) Internal Contrast, testing the distinction between morpheme-internal and non-morpheme-internal; 2) Affix Contrast, testing the distinction between affix-boundaries and non-affix-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.

2.   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

The data is modeled using logistic mixed-effects regression.[16] The model selection procedure followed two main 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 were theoretically motivated in Section 2 above, and are all included in the model. We have no theoretical motivation for including interactions. However, we saw in Section 4 that voicing in /p t k/ is near-floor, and this could be masking true effects in the data. For this reason, we tested all possible interactions with Laryngeal Category in a random intercepts-only model, in case some effects could be found only in /b d g/. Only significant interactions were kept.

---

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

[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 note 12 for more details on the R packages used and Puggaard-Rode et al. (2022) for code and data.

Random effects selection: All meaningful by-speaker and by-item random slopes were then 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. We used strictly uncorrelated random effects; this leads to much higher convergence rates in logistic models, and Seedorff et al. (2019) show in simulations that this 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, which in our case means that 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 & Curtin, 2018). Accordingly, we removed all random slopes with estimated zero variances except Laryngeal Category, since this is a variable of key interest in our study. This means that the resulting model is probably slightly overparametrized, since it is highly unlikely that there is actually *no* by-speaker variance for laryngeal category, but a reasonable interpretation is that the by-speaker variance for laryngeal category is very close to the random variation for laryngeal category in the data. The entire model selection process is documented in Puggaard-Rode, Horslund, Jørgensen, and Vet (2022), and the final model is summarized in **Table 4**.

Simple fixed effects	Intercept, Laryngeal Category, Place of Articulation (velar contrast, bilabials versus 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 versus 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)

Table 4: Summary of the final model.

None of the included independent variables shows 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 ratio (OR), on the other hand, are easier to interpret. In order to aid interpretability, we report both the model coefficients and standard error in the log-odds scale, and odds (ratio), which corresponds to exponentiated 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 ORs can be interpreted straightforwardly as the change in probability associated with that variable (see Sonderegger, in press, chapter 6). Odds and ORs are given as fractions – if OR > 1, the odds of voicing are higher in the variable level corresponding to + in the contrast coding; if OR < 1, the odds of voicing are higher in the variable level corresponding to –. For the standardized continuous variables, OR refers to the change in predicted likelihood of voicing associated with an increase of 1 standard deviation.

5.2. Results

The results of the logistic mixed-effects regression model described above is summarized in **Table 5**; we do not include a random effects table here, but it can be found in Puggaard-Rode et al. (2022). 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 the fixed and random effects combined, respectively (see Nakagawa, Johnson, & Schielzeth, 2017 for details of how this is calculated for generalized linear mixed-effects models).

Variable	Odds (ratio)	coef (log-odds)	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	0.3	–4.22	<.001	***	√
Place, – bilabial, + alveolar	1.16 : 1	0.15	0.34	0.44	0.66		
Preceding Approximant	1.63 : 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	1.32 : 1	0.28	0.14	1.95	0.05	.	

(Contd.)

Variable	Odds (ratio)	coef (log-odds)	SE	z	p		
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 (+infl.)	3.13 : 1	1.14	0.61	1.87	0.06	.	
Bnd: Compound Contrast (+cp.)	1.57 : 1	0.45	0.41	1.1	0.27		
Word Class (– open, +closed)	1.57 : 1	–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	***	

Table 5: Summary of logistic mixed-effects regression model. √ indicates agreement with our hypotheses in Table 1, and ÷ indicates disagreement with our 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 phonetic feature, and OR > 1 means that the odds are increased in the presence of said feature.

In some cases, the results of the mixed effects model tell quite a different story than the exploratory analysis presented in Section 4. In these cases, the results of the mixed effects model should be taken as the best possible 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 controlled for (i.e., kept at their average).

The significant variables overwhelmingly pattern as predicted. For the following categorical variables, this means that they 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 probability of voicing is approximately doubled in

syllables with neutral vowels as well as in unstressed syllables, and the probability is 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 stops with a fronted occlusion, i.e., 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 our 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 and lexical item into account.

We actually see a fairly strong effect of Preceding Stress in the expected direction; voicing is around four times more likely following stressed syllables. This is interesting, because in the exploratory analysis there was essentially no correlation between preceding stress and voicing.

Unexpectedly, the Stød variable patterns in the opposite direction of our predictions and what we saw in the exploratory analysis. Closure voicing is found to be around twice as likely in syllables with stød. We return to this in the discussion in Section 6.3 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 morphological boundaries. 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 we control for independently in the model, and because of this, we predicted that word-internal stops would be voiced at a higher rate than affixes. We return to this in the discussion.

Other categorical variables—Preceding Approximant, Preceding High Vowel, High Vowel, Preceding Central Vowel, Word Class Type, and Sex—have no significant influence on voicing in the model, although 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 face value can be better explained by other (potentially random) variables in the data.

The influence of continuous predictors is visualized in **Figure 7**. 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 more intervocalic voicing. In fact, the predicted likelihood of voicing is near ceiling for the quickest tokens, and near floor for a large portion of the slower tokens.
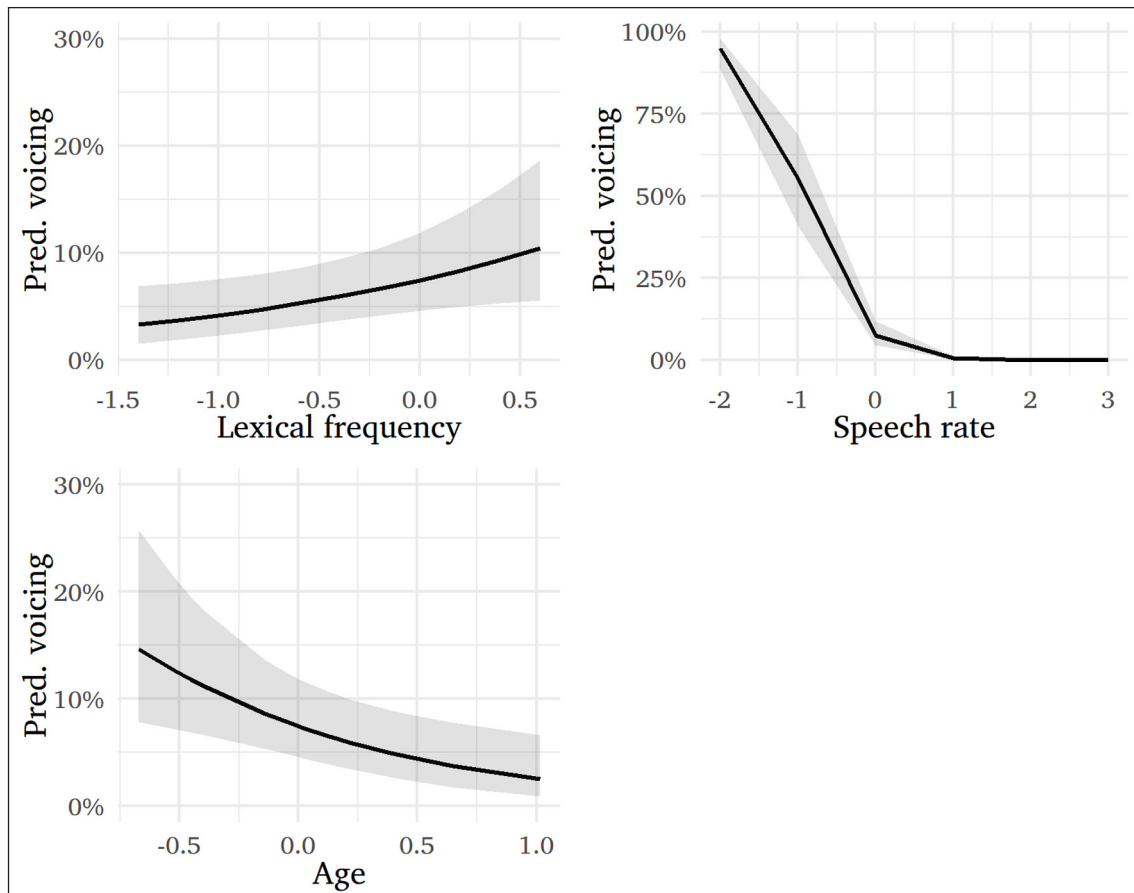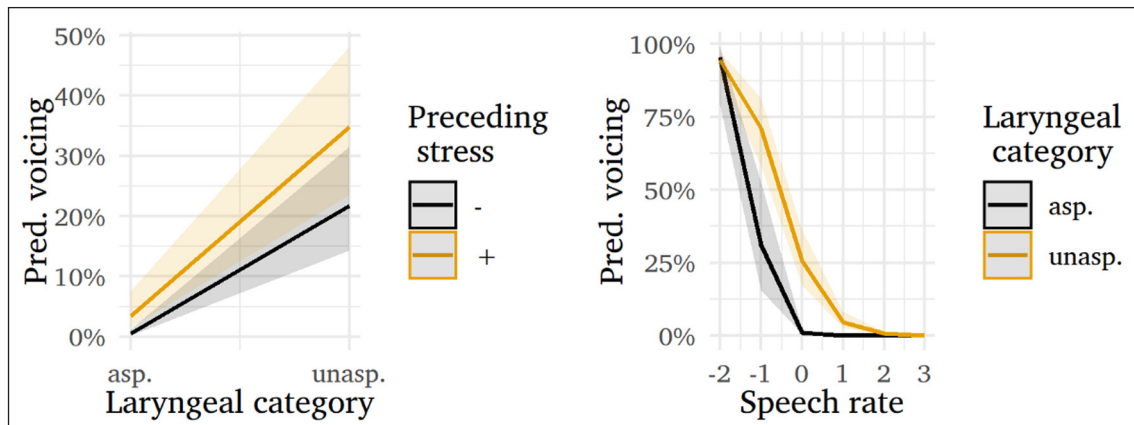
Figure 7: Plots showing the likelihood of fully voiced stops of continuous variables as predicted from the mixed-effects model. The x-axes are standardized units. Note that y-axes differ due to the very high likelihood of voicing in very quick speech, so keeping y-axes identical would blur the effect in other variables.

Figure 8 shows the predicted significant interaction effects. The interaction effect between Laryngeal Category and Preceding Stress is 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 predicted 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 much faster speech for /p t k/ than /b d g/.

Figure 8: Plots showing the likelihood of fully voiced stops of interaction effects as predicted from the mixed-effects model. The x-axes are standardized units. Note that y-axes differ due to the very high likelihood of voicing in very quick speech, so keeping y-axes identical would blur the effect in other variables.

6. Discussion

In this section, we discuss the results in relation to the research questions we presented in Section 2.

6.1. RQ1: Closure voicing and laryngeal category

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

Abstract [voice] approaches straightforwardly predict the first finding; [–voice] stops are naturally voiced less frequently than [+voice] stops. With regards to the second finding, in Kingston and Diehl's (1994) abstract account of [voice], we could postulate a controlled phonetic mechanism that actively counteracts voicing in [+voice] stops in order to account for the data. Such a mechanism seems intuitively strange but is already independently needed for Icelandic, in which intervocalic voicing of unaspirated stops is seemingly even rarer than in Danish (Pétursson, 1976).

Concrete [voice] approaches also straightforwardly account for the first finding, but not necessarily the second finding. [spread glottis] should block voicing, while unmarked stops are expected to be voiced whenever natural (i.e., intervocalically). In Beckman et al.'s (2013) account of [spread glottis], they assume that there is a point in the phonological derivation

where active privative features are reinterpreted as numerically valued features. Since [spread glottis] is the active laryngeal feature in e.g., German, Danish, and Icelandic, /p t k/ are assigned a high numeric value for [spread glottis], while /b d g/ are assigned lower values. They go on to suggest that German /b d g/ are assigned [1sg], allowing for passive intervocalic voicing, and that Danish /b d g/ are assigned [5sg], blocking passive voicing. This predicts our results quite well. Note, however, that other proponents of [spread glottis] in Danish (like Iverson & 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 this position (see Section 1.1).

In the end, we believe the best explanation of our acoustic results is one that relies on 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 1.2 that /p t k/ have shorter closure duration and are produced with less muscular tension than /b d g/. Either both sets are phonetically lenis, or /b d g/ are in fact the fortis set. The shorter closure duration and lower muscular tension of /p t k/ would predict more closure voicing in this set if the vocal folds were properly adduced and tensed for voicing. In careful speech, however, all Danish stops are usually accompanied by a glottal opening gesture, the purpose of which is presumably to enforce voicelessness. The glottal gestures are different in magnitude across the laryngeal series. /p t k/ have a glottal opening 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 aspirated release, which is the primary cue to the contrast between the two sets. /b d g/ also actively block voicing through glottal spreading, but for these phones it serves little to no role during the release, and it is not crucial to maintaining the contrast. The differences in magnitude of the glottal gesture can explain both findings: the probabilistic 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 relating to laryngeal category can be accounted for by all three major accounts of laryngeal representation, but not all theories predicted the results equally well. Recall from Section 1.1 that an abstract [voice] account did not allow us to make any specific predictions. A concrete [voice] account only predicts the results 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.[17]

---

[17] 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 its disadvantages.

On a final note, Schachtenhaufen (2019) recently suggested abandoning the transcription standard using [b̥ d̥ ɡ̊ b̥ʰ d̥ˢ ɡ̊ʰ] for [p t k pʰ ts kʰ], since fortis/lenis is not traditionally indicated in IPA, and IPA guidelines suggest using [b̥]-style transcription to indicate devoicing of sounds that are usually voiced. This study further cements that Danish /b d ɡ/ are clearly not usually voiced: Not only are /b d ɡ/ categorically voiceless in most positions, voicing is also regularly blocked in the one syllabic position where it would actually be phonetically natural. We are therefore strongly in favor of Schachtenhaufen's proposal.

6.2. RQ2: Closure voicing as phonetic and phonological lenition

Closure voicing is to a large extent found in environments where we expect to find phonetic lenition: Its occurrence increases with speech rate; it is found more frequently in unstressed syllables, in syllables with schwa, and in affixes. On the basis of our results, it seems sensible to consider intervocalic closure voicing a lenition phenomenon in itself.

This has some interesting phonological consequences. As discussed in Section 1.1, it is often difficult to account for phonetic voicing processes with reference to spreading of a [voice] feature. In phonological representational frameworks relying on privative features, the voiceless unaspirated stop is generally considered the unmarked one, i.e., it carries no laryngeal features. Similarly, voicing is generally not considered phonologically marked for sonorant sounds (e.g., Lombardi, 1995). As such, [voice] is not specified for sonorant sounds and cannot spread from them, and the addition of a [voice] feature to a stop appears to be phonological fortition rather than lenition, since material is added, making for a more complex underlying structure. Spreading of Rice and Avery's (1989) non-laryngeal [spontaneous voice] feature node may be able to represent the process, but it does not 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 even if the context is exactly right, it may not occur.

The question remains: *Why* is closure voicing a lenition phenomenon in intervocalic stops? A gesture-based approach to laryngeal representation can account for this. We propose that closure voicing in /b d ɡ/ follows from the loss of the glottal opening gesture that is usually associated with these stops. When the vocal fold configuration is optimal for voicing and subglottal pressure is high, some duration of closure voicing is natural and requires no extra effort (Westbury & Keating, 1986; see Section 1.1). This vocal fold configuration is required for producing vowels both before and after intervocalic stops, so maintaining it throughout the stop will require the least articulatory effort. In contexts where we generally expect gestural undershoot, it is unsurprising that we also see the loss of a non-distinctive glottal opening gesture (as in /b d ɡ/), and to a much lesser extent, the loss of a distinctive glottal opening 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 ɡ/ are abstractly specified as [voice], we would not expect lenition to be a requirement for phonetic voicing—in fact, Kingston and Diehl (1994) explicitly use the presence of intervocalic voicing as an argument for why Danish has [voice]. If /b d ɡ/ receive some value for [spread glottis] late in the phonological derivation, there is no explicit mechanism for reducing this number in the case of 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 as in Articulatory Phonology. Recall from Section 1.2 that only a gesture-based account of underlying representation leads to any specific predictions about lenition, namely that lenition would lead to reduction in the timing and magnitude of associated glottal gestures. This is in line with our results. The difference in lenition rates in 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 a separate, active gesture.

In **Table 6**, we summarize the predictions following from different theoretical approaches, and whether or not we found support for these predictions in the current study.

Approach	Danish /b d ɡ/		Danish /p t k/		Lenition	
	Prediction	Support	Prediction	Support	Prediction	Support
Concrete [voice]	Variable voicing	√	Little voicing	√		
Abstract [voice]	All outcomes possible		All outcomes possible			
Gestures	Little voicing	√	Very little voicing	√	Voicing in both series	√
					More voicing in /b d ɡ/	√

Table 6: Summary of predictions from different theoretical approaches.

6.3. RQ3: The relative predictive power of different variables

We have already discussed the predictive power of some of our variables. Laryngeal Setting was a very strong predictor of voicing, as were 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 no effect was found for morpheme-initial stops, but one was found for stops at affix boundaries. In Section 2.1.3, we 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 frameworks (where phonetic information is invisible to morphology), and it represents a quite different conception of phonological representation than those we have discussed above. The effect of affix boundaries remains an interesting problem for further research.

Many of the other variables that we expected to influence closure voicing were aerodynamic in nature, and fewer of those have an observable effect on closure voicing in our data. This may be either because these variables truly have no effect on closure voicing, or because the influence of these variables is more gradient in nature. It is possible that these variables would affect the relative duration of closure voicing within those stops that we simply categorize as 'not fully voiced.'

We had a number of predictions for how the tongue position before and after the occlusion would affect the prevalence of closure voicing, which mostly come down to this: We expected a narrower constriction in the oral cavity before and after the occlusion to decrease the odds 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 holds up; we found no effect of tongue body position except for point of occlusion in the stop itself.

Place of articulation has quite a strong effect on voicing, and this has an aerodynamic explanation. The supralaryngeal cavity is relatively small during a velar occlusion and provides little opportunity for passive expansion, and as such, velar stops are generally 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 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, one initial articulatory correlate of stød reported by Fischer-Jørgensen (1987, 1989) is increased subglottal pressure (although note that subglottal pressure was measured for only one participant, and no words with initial oral stops were measured). This may serve to explain why tautosyllabic stød empirically shows a negative correlation with voicing (see **Table 3**), but correlates positively with voicing in a model that also controls for stress (see **Table 5**).

7. Conclusion

In this study, we report on the occurrence of intervocalic stop voicing in a corpus of spontaneous Danish speech. Although Danish stops are generally well-described, most of what has been previously written about voicing has been speculative. We show that intervocalic voicing is very rare in /p t k/ and occurs in less than half of the /b d g/ tokens. In our modeling of the data, we 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 generally found at relatively high rates in environments where we also expect lenition, 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 opening gestures that counteract voicing, although these gestures differ in timing, magnitude, and functional importance. Intervocalic voicing can be modeled as the loss of this gesture, which 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 an extremely broad literature on laryngeal features and related phonological representation, and we have necessarily discussed only a few possible viewpoints here. If intervocalic voicing is indeed a lenition phenomenon, we suggest that this is best represented in a phonological representational framework with the capacity to directly incorporate the timing and magnitude of articulatory gestures, such as Articulatory Phonology.

Few corpus studies of intervocalic voicing are available, and as such, it is difficult to compare these results to other aspiration languages (or with 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.

Availability of data and code

Acknowledgements

Funding information

Competing interests

The authors have no competing statements to declare.

References

Abrahams, H. (1949). *Études phonétiques sur les tendances évolutives des occlusives germaniques.* Aarhus: Aarhus Universitetsforlag.

Andersen, P. (1981). The effect of increased speaking rate and (intended) loudness on the glottal behaviour in stop consonant production as exemplified by Danish p. *Annual Report of the Institute of Phonetics, University of Copenhagen,* 15, 103–146.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing. Keep it maximal. *Journal of Memory and Language,* 68(3), 255–278. DOI: https://doi.org/10.1016/j.jml.2012.11.001

Barton, K. (2020). *MuMIn. Multi-model inference.* R package version 1.43.17. Available from CRAN, https://CRAN.R-project.org/package = MuMIn.

Basbøll, H. (2003). Prosody, productivity and word structure. The stød pattern of Modern Danish. *Nordic Journal of Linguistics,* 26(1), 5–44. DOI: https://doi.org/10.1017/S033258650300101X

Basbøll, H. (2005). *The phonology of Danish.* Oxford & New York: Oxford University Press.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Available on ArXiv, ID: 1506.04967v2.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2021). *lme4. Linear mixed-effects models using "Eigen" and S4*. R package version 1.1-27.1. Available from CRAN, https://CRAN.R-project.org/package = lme4.

Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features. Aspirating vs. true voice languages. *Journal of Linguistics*, 49(2), 259–284. DOI: https://doi.org/10.1017/S0022226712000424

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, 113(2), 1001–1024. DOI: https://doi.org/10.1121/1.1534836

Blevins, J. (2004). *Evolutionary Phonology. The emergence of sound patterns*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486357

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, 5(9/10), 341–345. Retrieved from https://www.praat.org.

Boersma, P., & Weenink, D. (2019). *Praat. Doing phonetics by computer*. Version 6.0.55. Available from https://www.praat.org.

Bohn, O.-S. (2013). Acoustic characteristics of Danish infant directed speech. *Proceedings of Meetings on Acoustics*, 19. DOI: https://doi.org/10.1121/1.4798488

Brauer, M.; Curtin, J. J. (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: https://doi.org/10.1037/met0000159

Brink, L., & Lund, J. (2018). Udtale: Yngre nydansk. In E. Hjorth (ed.), *Dansk sproghistorie 2. Ord for ord for ord* (pp. 197–228). Aarhus: Aarhus Universitetsforlag.

Brotherton, C., & Block, A. (2020). Soft d in Danish. Acoustic characteristics and issues in transcription. *Proceedings of the Linguistic Society of America*, 5, 792–797. DOI: https://doi.org/10.3765/plsa.v5i1.4739

Browman, C. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219–252. DOI: https://doi.org/10.1017/S0952675700000658

Browman, C. & Goldstein, L. (1992). Articulatory Phonology. An overview. *Phonetica*, 49(3/4), 155–180. DOI: https://doi.org/10.1159/000261913

Bybee, J. L. (2000a). Lexicalization of sound change and alternating environments. In M. B. Broe & J. B. Pierrehumbert (eds.), *Acquisition and the lexicon* (pp. 250–268). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1093/acprof:oso/9780195301571.003.0010

Bybee, J. L. (2000b). The phonology of the lexicon. Evidence from lexical diffusion. In M. Barlow & S. Kemmer (eds.), *Usage-based models of language* (pp. 65–86). Stanford: Center for the Study of Language and Information Publications. DOI: https://doi.org/10.1093/acprof: oso/9780195301571.003.0009

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511612886

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe & J. B. Pierrehumbert (eds.), *Acquisition and the lexicon* (pp. 70–87). Cambridge: Cambridge University Press. Available from CiteSeerX, https://www.citeseerx.ist. psu.edu.

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing. Evidence from Korean. *Phonetica*, 58(3), 129–162. DOI: https://doi.org/10.1159/000056196

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row. Available from Massachusetts Institute of Technology repository, https://web.mit.edu.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, 54, 35–50. DOI: https://doi.org/10.1016/j.wocn.2015.09.003

Ernestus, M., Lahey, M., Verhees, F., & Baayen, R. H. (2006). Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America*, 120(2), 1040–1051. DOI: https://doi. org/10.1121/1.2211548

Ferguson, S. H. (2004). Talker differences in clear and conversational speech. Vowel intelligibility for normal-hearing listeners. *Journal of the Acoustical Society of America*, 116(4), 2365–2373. DOI: https://doi.org/10.1121/1.1788730

Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Le Maître Phonetique*, 32(69), 42–59. Available from JSTOR, https://www.jstor.org/stable/44705403.

Fischer-Jørgensen, E. (1968). Voicing, tenseness and aspiration in stop consonants, with special reference to French and Danish. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 3, 63–114.

Fischer-Jørgensen, E. (1972). Kinesthetic judgement of effort in the production of stop consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 6, 59–73.

Fischer-Jørgensen, E. (1979). Temporal relations in consonant-vowel syllables with stop consonants based on Danish material. In B. Lindblom & S. Öhman (eds.), *Frontiers of speech communication research* (pp. 51–68). London: Academic Press.

Fischer-Jørgensen, E. (1980). Temporal relations in Danish tautosyllabic CV sequences with stop consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 14, 207–261.

Fischer-Jørgensen, E. (1987). A phonetic study of stød in Standard Danish. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 21, 55–267.

Fischer-Jørgensen, E. (1989). Phonetic analysis of the stød in Standard Danish. *Phonetica*, 46(1/2/3), 1–59. DOI: https://doi.org/10.1159/000261828

Fischer-Jørgensen, E., & Hirose, H. (1974). A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production. *Status Report on Speech Research*, 39/40, 231–253. Retrieved from ERIC, https://eric.ed.gov.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Thousand Oaks: Sage.

Fox, J., Weisberg, S. & Price, B. (2021). *car. Companion to applied regression*. R package version 3.0-12. Available from CRAN, https://CRAN.R-project.org/package=car.

Frøkjær-Jensen, B., Ludvigsen, C., & Rischel, J. (1971). A glottographic study of some Danish consonants. In L. L. Hammerich, R. Jakobson & E. Zwirner (eds.), *Form and substance. Phonetic and linguistic papers presented to Eli Fischer-Jørgensen* (pp. 123–140). Odense: Akademisk Forlag.

Frøkjær-Jensen, B., Ludvigsen, C., & Rischel, J. (1973). A glottographic study of some Danish consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 7, 269–295.

Gamkrelidze, T. V. (1975). On the correlation of stops and fricatives in a phonological system. *Lingua*, 35(3–4), 231–261. DOI: https://doi.org/10.1016/0024-3841(75)90060-1

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511790942

Goldstein, L. M., & Browman, C. P. (1986). Representation of voicing contrasts using articulatory gestures. *Journal of Phonetics*, 14(2), 339–342. DOI: https://doi.org/10.1016/S0095-4470(19)30662-X

Grønnum, N. (1995). Danish vowels. Surface contrast versus underlying form. *Phonetica*, 52(3), 215–220. DOI: https://doi.org/10.1159/000262173

Grønnum, N. (1998). Illustrations of the IPA. Danish. *Journal of the International Phonetic Association*, 28(1/2), 99–105. DOI: https://doi.org/10.1017/S0025100300006290

Grønnum, N. (2005). *Fonetik og fonologi. Almen og dansk* (3rd ed.). Copenhagen: Akademisk Forlag.

Grønnum, N. (2009). A Danish phonetically annotated spontaneous speech corpus (DanPASS). *Speech Communication*, 51(7), 594–603. DOI: https://doi.org/10.1016/j.specom.2008.11.002

Grønnum, N. (2016). *DanPASS. Danish Phonetically Annotated Speech*. https://danpass.hum.ku.dk.

Grønnum, N., & Basbøll, H. (2001). Consonant length, stød and morae in Standard Danish. *Phonetica*, 58(4), 230–253. DOI: https://doi.org/10.1159/000046177

Grønnum, N., & Basbøll, H. (2007). Danish stød. Phonological and cognitive issues. In M.-J. Solé, P. S. Beddor & M. Ohala (eds.), *Experimental approaches to phonology* (pp. 192–206). Oxford: Oxford University Press. Available from ResearchGate, https://www.researchgate.net/publication/292768809.

Grønnum, N., & Tøndering, J. (2007). Question intonation in non-scripted Danish dialogues. In *Proceedings of the XVIth International Congress of Phonetic Sciences* (pp. 1229–1232). Saarbrücken: Saarland University. Retrieved from conference web page, https://icphs2007.de.

Hanson, H. M. (2009). Effects of obstruent consonants on fundamental frequency at vowel onset in English. *Journal of the Acoustical Society of America*, 125(1), 425–441. DOI: https://doi.org/10.1121/1.3021306

Hayes, B. (1999). Phonetically driven phonology. The role of Optimality Theory and inductive grounding. In M. Darnell, E. M. Moravcsik, M. Noonan, F. J. Newmeyer & K. M. Wheatley (eds.), *Functionalism and formalism in linguistics. Volume I: General papers* (pp. 243–286). Amsterdam & Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/slcs.41

Heegård, J. (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: https://doi.org/10.1080/03740463.2013.849901

Henton, C., Ladefoged, P., & Maddieson, I. (1992). Stops in the world's languages. *Phonetica*, 49(2), 65–101. DOI: https://doi.org/10.1159/000261905

Hirose, H., & Gay, T. (1972). The activity of the intrinsic laryngeal muscles in voicing control. *Phonetica*, 25(3), 140–164. DOI: https://doi.org/10.1159/000259378

Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. M. Christie (ed.), *Current progress in historical linguistics* (pp. 96–105). Amsterdam: North-Holland.

Horslund, C. S., Jørgensen, H. & Puggaard, R. (2021). En alternativ, fonetisk baseret fonemanalyse af det danske konsonantsystem. In Y. Goldshtein, I. S. Hansen & T. T. Hougaard (eds.), *18. møde om udforskningen af det danske sprog* (pp. 251–267). Aarhus: Aarhus University. Available from Aarhus University, https://projekter.au.dk/muds/tidligere-muds-rapporter/.

Horslund, C. S., Puggaard, R. & Jørgensen, H. (2022). A phonetically-based phoneme analysis of the Danish consonant system. *Acta Linguistica Hafniensia*. DOI: https://doi.org/10.1080/03740463.2021.2022866

Hutters, B. (1985). Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica*, 42(1), 1–24. DOI: https://doi.org/10.1159/000261734

International Phonetic Association. (1999). *Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Iverson, G. K., & Salmons, J. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, 12, 369–396. DOI: https://doi.org/10.1017/S0952675700002566

Jaeger, J. J. (1983). The fortis/lenis question. Evidence from Zapotec and Jawoñ. *Journal of Phonetics*, 11(2), 177–189. DOI: https://doi.org/10.1016/S0095-4470(19)30814-9

Jakobson, R., Fant, C. G. M., & Halle, M. (1951). *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.

Jeel, V. (1975). An investigation of the fundamental frequency of vowels after various consonants, in particular stop consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 9, 191–211.

Jespersen, O. (1989). *Fonetik. En systematisk fremstilling af læren om sproglyd*. Copenhagen: Det Schubotheske Forlag.

Jespersen, O. (1906). *Modersmålets fonetik*. Copenhagen: Det Schubotheske Forlag.

Jessen, M. (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In T. A. Hall (ed.), *Distinctive feature theory* (pp. 237–294). Berlin & New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110886672.237

Jessen, M. & Ringen, C. (2002). Laryngeal features in German. *Phonology*, 19(2), 189–218. DOI: https://doi.org/10.1017/S0952675702004311

Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In C. Gussenhoven & N. Warner (eds.), *Laboratory Phonology 7* (pp. 3–34). Berlin & New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197105.1.3

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words. Evidence from reduction in lexical production. In J. L. Bybee & P. J. Hopper (eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam & Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/tsl.45.13jur

Juul, H., Pharao, N., & Thøgersen, J. (2016). Moderne danske vokaler. *Danske Talesprog*, 16, 35–72.

Juul, H., Pharao, N., & Thøgersen, J. (2019). Stemthed og afstemthed i danske sonoranter. *Danske Talesprog*, 19, 108–130.

Kaplan, A. (2010). *Phonology shaped by phonetics. The case of intervocalic lenition* (Doctoral dissertation). University of California, Santa Cruz. DOI: https://doi.org/10.7282/T30G3J2K

Katz, J. (2016). Lenition, perception and neutralisation. *Phonology*, 33(1), 43–85. DOI: https://doi.org/10.1017/S0952675716000038

Keating, P. A. (1984a). Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics*, 59, 18–28. Available from eScholarship, https://escholarship.org/uc/item/2497n8jq.

Keating, P. A. (1984b). Phonetic and phonological representation of stop consonant voicing. *Language*, 60(2), 286–319. DOI: https://doi.org/10.2307/413642

Keating, P. A., Cho, T., Fougeron, C., & Hsu, C.-S. (2004). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden & R. Temple (eds.), *Phonetic interpretation* (pp. 145–163). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486425.009

Keating, P. A., Linker, W., & Huffman, M. K. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics*, 11(3), 277–290. DOI: https://doi.org/10.1016/S0095-4470(19)30827-7

Kingston, J., & Diehl, R. L. (1994). Phonetic knowledge. *Language*, 70(3), 419–454. DOI: https://doi.org/10.1353/lan.1994.0023

Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology Yearbook*, 2(1), 85–138. DOI: https://doi.org/10.1017/S0952675700000397

Kirby, J. & Ladd, D. R. (2018). Effects of obstruent voicing on vowel F0. Implications for laryngeal realism. *Yearbook of the Poznań Linguistic Meeting*, 4, 213–235. DOI: https://doi.org/10.2478/yplm-2018-0009

Kohler, K. J. (1984). Phonetic explanation in phonology. The feature fortis/lenis. *Phonetica*, 41(3), 150–174. DOI: https://doi.org/10.1159/000261721

Komsta, L., & Novomestky, F. (2015). *moments. Moments, cumulants, skewness, kurtosis and related tests*. R package version 0.14. Available from CRAN, http://CRAN.R-project.org/package = moments.

Lombardi, L. (1995). Laryngeal features and privativity. *The Linguistic Review*, 12(1), 35–59. DOI: https://doi.org/10.1515/tlir.1995.12.1.35

Lüdecke, D. (2021). *sjPlot. Data visualization for statistics in social science*. R package version 2.8.10. Available from CRAN, http://CRAN.R-project.org/package = sjPlot.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511753459

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, R. H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. DOI: https://doi.org/10.1016/j.jml.2017.01.001

Meteyard, L., & Davies, R. A. I. (2019). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, 112. DOI: https://doi.org/10.1016/j.jml.2020.104092

Möbius, B. (2004). Corpus-based investigations on the phonetics of consonant voicing. *Folia Linguistica*, 38(1/2), 5–26. DOI: https://doi.org/10.1515/flin.2004.38.1-2.5

Mortensen, D. R. (2014). The emergence of obstruents after high vowels. *Diachronica*, 29(4), 434–470. DOI: https://doi.org/10.1075/dia.29.4.02mor

Mortensen, J., & Tøndering, J. (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* (pp. 49–52). Linköping: Linköping University. Available from conference web page, https://old.liu.se/ikk/fonetik2013/proceedings.

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (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: https://doi.org/10.1098/rsif.2017.0213

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris. DOI: https://doi.org/10.1515/9783110977790

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The production of speech* (pp. 189–216). New York: Springer. DOI: https://doi.org/10.1007/978-1-4613-8202-7_9

Ohala, J. J., & Riordan, C. J. (1979). Passive vocal tract enlargement during voiced stops. In J. J. Wolf & D. H. Klatt (eds.), *Speech communication papers presented at the 97th Meeting of the Acoustical Society of America* (pp. 89–92). Cambridge, MA: Massachusetts Institute of Technology. Available from author web page, http://linguistics.berkeley.edu/~ohala.

Pape, D. & Jesus, L. M. T. (2014). Production and perception of velar stop (de)voicing in European Portuguese and Italian. *EURASIP Journal on Audio, Speech, and Music Processing*. DOI: https://doi.org/10.1186/1687-4722-2014-6

Parker, S. (2002). *Quantifying the sonority hierarchy* (Doctoral dissertation). University of Massachusetts, Amherst. Available from ProQuest, ID: AAI3056268.

Petersen, N. R. (1983). The effect of consonant type on fundamental frequency and larynx height in Danish. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 17, 55–86.

Pétursson, M. (1976). Aspiration et activité glottale. Examen expérimental à partir de consonnes islandaises. *Phonetica*, 33(3), 169–198. DOI: https://doi.org/10.1159/000259721

Pharao, N. (2009). *Consonant reduction in Copenhagen Danish. A study of linguistic and extra-linguistic factors in phonetic variation and change* (Doctoral dissertation). University of Copenhagen. Available from DanPASS, https://danpass.hum.ku.dk.

Pharao, N. (2011). Plosive reduction at the group level and in the individual speaker. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1590–1593). Hong Kong. Available from conference web page, https://www.internationalphoneticassociation.org/icphs/icphs2011.

Pierrehumbert, J. B. (2001). Exemplar dynamics. Word frequency, lenition and contrast. In J. L. Bybee & P. J. Hopper (eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). Amsterdam & Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/tsl.45.08pie

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (eds.), *Laboratory Phonology 7* (pp. 101–139). Berlin & New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197105.1.101

Plag, I., Lohmann, A., Hedia, S. B., & Zimmermann, J. (2020). An s is an 's, or is it? Plural and genitive plural are not homophonous. In L. Körtvélyessy & P. Stekauer (eds.), *Complex words. Advances in morphology* (pp. 260–292). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781108780643.015

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118(4), 2561–2569. DOI: https://doi.org/10.1121/1.2011150

Puggaard, R., Grijzenhout, J., & Botma, B. (2019). The acoustics of phonological contrast. Fortis and lenis stops in Danish and Dutch. Paper presented at *Phonetik und Phonologie Tagung* 15. Düsseldorf: Heinrich-Heine University.

Puggaard-Rode, R., Horslund, C. S., Jørgensen, H. & Vet, D. J. (2022). Replication data for: The rarity of intervocalic voicing of stops in Danish spontaneous speech. DataverseNL. DOI: https://doi.org/10.34894/OSGTR8

R Core Team. (2021). *R. A language and environment for statistical computing*. Version 4.1.2. Available from https://R-project.org.

Rice, K., & Avery, P. (1989). On the interaction between sonorancy and voicing. *Toronto Working Papers in Linguistics*, 10, 65–82. Available from journal web page, https://twpl.library.utoronto.ca.

Rischel, J. (1970). Consonant gradation. A problem in Danish phonology and morphology. In H. Benediktsson (ed.), *The Nordic languages and modern linguistics* (pp. 460–480). Reykjavík: Vísindafélag Íslendinga.

RStudio Team. (2022). *RStudio. Integrated development environment for R.* Version 2021.09.2 + 382. Available from http://rstudio.com.

Ryalls, J., Zipprer, A., & Baldauff, P. (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research*, 40, 642–645. DOI: https://doi.org/10.1044/jslhr.4003.642

Sawashima, M. (1970). Glottal adjustments for English obstruents. *Status Report on Speech Research*, 21/22, 187–200. Available from Haskins Laboratories, https://haskinslabs.org.

Schachtenhaufen, R. (2010). Looking for lost syllables in Danish spontaneous speech. In P. J. Henrichsen (ed.), *Linguistic theory and raw sound* (pp. 61–85). Frederiksberg: Samfundslitteratur. Available from Copenhagen Business School repository, https://research.cbs.dk.

Schachtenhaufen, R. (2013). *Fonetisk reduktion i dansk (*Doctoral dissertation). Copenhagen Business School. Available from EconStor, http://hdl.handle.net/10419/208850.

Schachtenhaufen, R. (2019). IPA og IPA – ny dansk lydskriftstandard. Available from author web page, https://schwa.dk.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models. A tutorial. *Journal of Memory and Language*, 110. DOI: https://doi.org/10.1016/j.jml.2019.104038

Seedorff, M., Oleson, J., & McMurray, B. (2019). Maybe maximal. Good enough mixed models optimize power while controlling Type I error. Unpublished manuscript, available on PsyArXiv. DOI: https://doi.org/10.31234/osf.io/xmhfr

Seifert, J. (2009). Does speech reveal one's age? On the use of gerontolinguistic topics for forensic authorship analysis. In G. Grewendorf & M. Rathert (eds.), *Formal linguistics and law* (pp. 163–180). Berlin & New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110218398.2.163

Shih, C., & Möbius, B. (1998). Contextual effects on voicing profiles of German and Mandarin consonants. In *Proceedings of the 3rd ESCA/COCOSDA Workshop on Speech Synthesis* (pp. 81–86). Jenolan Caves House, Blue Mountains. Available from ISCA Archive, https://www.isca-speech.org.

Shih, C., Möbius, B., & Narasimhan, B. (1999). Contextual effects on consonant voicing profiles. A cross-linguistic study. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 989–992). San Francisco. Available from conference web page, https://www.internationalphoneticassociation.org/icphs/icphs1999.

Smith, J. L. (2008). Markedness, faithfulness, positions, and contexts. Lenition and fortition in Optimality Theory. In J. Brandão de Carvalho, T. Scheer & P. Ségéral (eds.), *Lenition and fortition* (pp. 519–560). Berlin & New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110211443.3.519

Smith, S. (1944). *Bidrag til Løsning af Problemer vedrørende Stødet i dansk Rigssprog. En eksperimentalfonetisk Studie.* Copenhagen: Kaifers.

Sonderegger, M., Stuart-Smith, J., Knowles, T., Macdonald, R, & Rathcke, T. (2020). Structured heterogeneity in Scottish stops over the twentieth century. *Language*, 96(1), 94–125. DOI: https://doi.org/10.1353/lan.2020.0003

Sonderegger, M. (in press). *Regression modeling for linguistic data*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.17605/OSF.IO/PNUMG.

Spore, P. (1965). *La langue danoise. Phonétique et grammaire contemporaines*. Copenhagen: Akademisk Forlag.

Steriade, D. (2009). The phonology of perceptibility effects. The P-Map and its consequences for constraint organization. In K. Hanson & S. Inkelas (eds.), *The nature of the word. Studies in honor of Paul Kiparsky* (pp. 151–179). Cambridge, MA & London: MIT Press. DOI: https://doi.org/10.7551/mitpress/7894.003.0011

Strycharczuk, P. (2012). Sonorant transparency and the complexity of voicing in Polish. *Journal of Phonetics*, 40(5), 655–671. DOI: https://doi.org/10.1016/j.wocn.2012.05.006

Swartz, B. L. (1992). Gender difference in voice onset time. *Perceptual and Motor Skills*, 75, 983–992. DOI: https://doi.org/10.2466/pms.1992.75.3.983

Swerts, M. (1994). *Prosodic features of discourse units* (Doctoral dissertation). Eindhoven University of Technology. DOI: https://doi.org/10.6100/IR411593

Swerts, M., & Collier, R. (1992). On the controlled elicitation of spontaneous speech. *Speech Communication*, 11(4–5), 463–468. DOI: https://doi.org/10.1016/0167-6393(92)90052-9

Tanner, J., Sonderegger, M., & Stuart-Smith, J. (2020). Structured speaker variability in Japanese stops. Relationships within versus across cues to stop voicing. *Journal of the Acoustical Society of America*, 148(2), 793–804. DOI: https://doi.org/10.1121/10.0001734

Terken, J. M. B. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech*, 27(3), 269–289. DOI: https://doi.org/10.1177/002383098402700306

Tomaschek, F., Plag, I., Ernestus, M., & Baayen, R. H. (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: https://doi.org/10.1017/S0022226719000203

Tøndering, J. (2003). Intonation contours in Danish spontaneous speech. In M.-J. Solé, D. Recasens & J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1241–1244). Universitat Autonoma de Barcelona. Available from conference web page, https://www.internationalphoneticassociation.org/icphs/icphs2003.

Tøndering, J. (2008). *Skitser af prosodi i spontant dansk* (Doctoral dissertation). University of Copenhagen. Available from University of Copenhagen repository, https://curis.ku.dk/portal.

Uldall, H. J. (1936). The phonematics of Danish. In D. Jones & D. B. Fry (eds.), *Proceedings of the 2nd International Congress of the Phonetic Sciences* (pp. 54–57). Cambridge: Cambridge University Press.

Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America*, 73(4), 1322–1336. DOI: https://doi.org/10.1121/1.389236

Westbury, J. R., & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics,* 22(1), 145–166. DOI: https://doi.org/10.1017/S0022226700010598

Wickham, H. (2016). *ggplot2. Elegant graphics for data analysis*. New York: Springer. DOI: https://doi.org/10.1007/978-0-387-98141-3

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H, & Dunnington, D. (2021). *ggplot2. Create elegant data visualizations using the grammar of graphics*. R package version 3.3.5. Available from CRAN, https://CRAN.R-project.org/package = ggplot2.