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

# CHAPTER 3

# The synchrony and diachrony of stop gradation

## 3.1 Introduction

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

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

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

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

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

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

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

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

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

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

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

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

## 3.2 The alternations

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

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

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

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

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

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

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

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

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

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

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

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

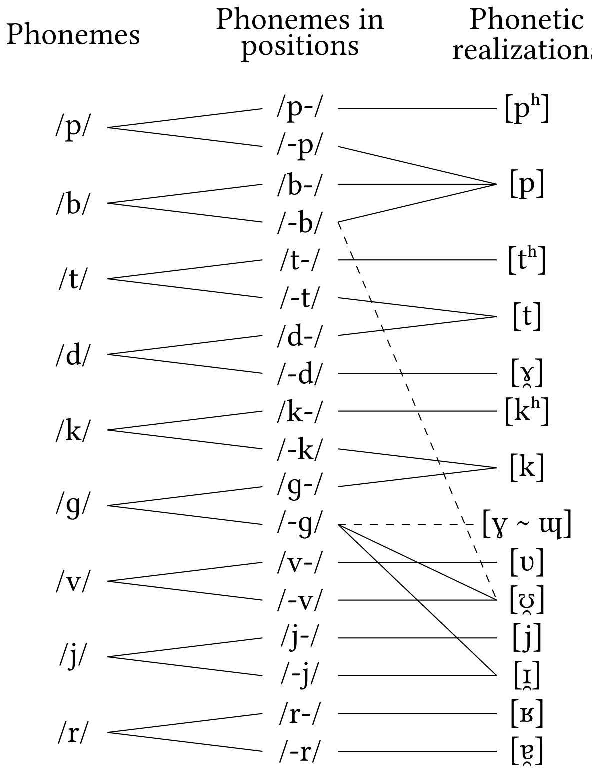


Figure 3.1: *Overview of phonemes and positional allophones according to the traditional analysis.*

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

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

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

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

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

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

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

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

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

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

## 3.3 Problems with the traditional analysis

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

### 3.3.1 Intractable neutralizations

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### 3.3.2 Lack of similarities between allophones

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### 3.3.3 Insufficient morphophonological evidence

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## 3.4 An alternative analysis

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

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

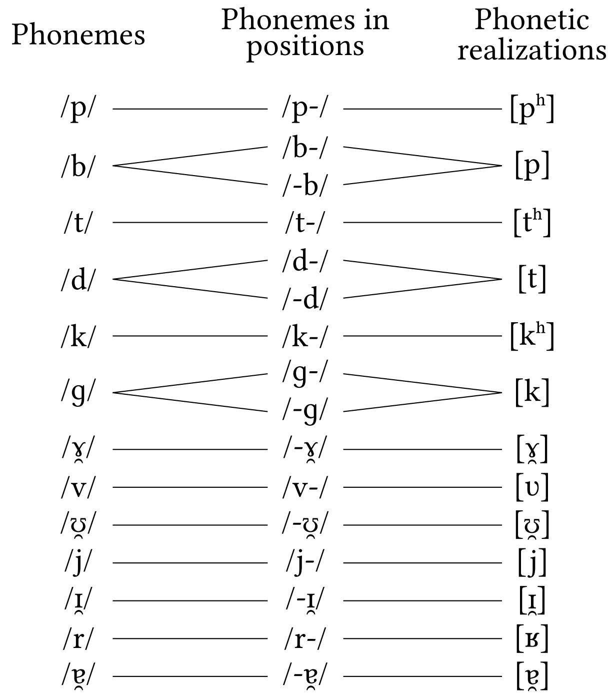


Figure 3.2: *Overview of phonemes and positional allophones according to the maximally biunique analysis of Ács and Jørgensen (2016).*

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

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

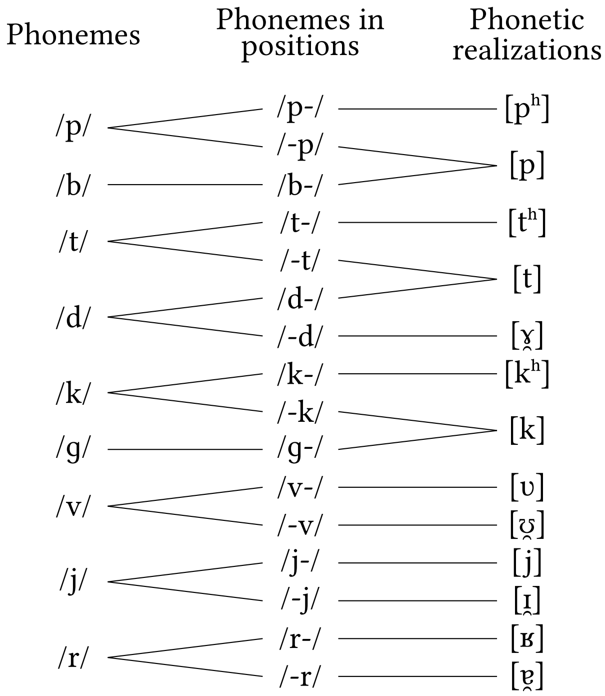


Figure 3.3: *Overview of phonemes and positional allophones according to the alternative analysis suggested here.*

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

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

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

### 3.4.1 No neutralizations

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

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

### 3.4.2 Shared phonetic and phonological properties

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

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

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

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

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

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

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

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

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

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

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

### 3.4.3 Alternations stored as suppletive allomorphs

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

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

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

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

### 3.4.4 Summary

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

## 3.5 The diachronic trajectory of stop gradation

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

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

### 3.5.1 Evolutionary Phonology and the CCC-model

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### 3.5.2 Stop gradation in five diachronic steps

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

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

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

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

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

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

#### 3.5.2.1 Singleton voicing in weak position

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

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

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

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

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

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

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

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

#### 3.5.2.2 Loss of closure in weak position

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

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

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

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

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

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

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

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

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

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

#### 3.5.2.3 Loss of voicing in strong position

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

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

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

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

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

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

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

#### 3.5.2.4 Increased aperture in weak position

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

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

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

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

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

#### 3.5.2.5 Recategorization of weak allophones

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

| | | |
|---|---|---|
| (21) | *Strong position* | [p t k] |
| | | [pʰ tʰ kʰ] |
| | *Weak position* | [β̞ ð̞] → [ʊ̯ ɣ̯], [ɰ] → [ʊ̯ ɹ̯] |
| | | [p t k] |

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

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

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

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

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

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

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

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

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

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

### 3.5.3 The history of closure voicing

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

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

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

(22) Laryngeally unspecified stop → [glottal width]

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

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

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

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

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

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

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

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

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

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

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

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

## 3.6 Conclusion

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

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

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