

University of Groningen

## Laryngeal contrast and phonetic voicing

Jansen, Wouter

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*  
2004

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. [Thesis fully internal (DIV), University of Groningen]. s.n.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## Chapter 3

# Laryngeal neutralisation

To phonologists familiar with the Germanic or Slavonic languages, the principal example of laryngeal neutralisation is likely to be the ‘dynamic’ process of *final laryngeal neutralisation*, generally referred to as *final devoicing*. This phenomenon is illustrated with examples from Dutch in 4. In Dutch, a lexical contrast between word-final fortis and lenis obstruents is realised in the regular past tense (as discussed above) and before certain vowel-initial suffixes such as the // -ən plural suffix, but not in unsuffixed forms or before a variety of other suffixes, including the diminutive suffix. Similar processes are found in Frisian (Tiersma, 1985), German, and many of the Slavonic and Turkic languages. Terms such as *final devoicing* or (German) *Auslautverhärtung* originate in the realisation of the neutralised series as mostly voiceless in citation forms and the identification of devoicing with fortition (i.e., a change from [-tense] to [+tense]).

#### (4) Final neutralisation in Dutch

UR	Plural	Citation	Diminutive	Gloss
/xrap/	[χrapə(n)]	[χrap]	[χrapjə]	joke
/krab/	[krabə(n)]	[krap]	[krapjə]	crab
/χra:t/	[χra:tə(n)]	[χra:t]	[χra:tjə]	fishbone
/χra:d/	[χra:də(n)]	[χra:t]	[χra:tjə]	degree

However, laryngeal neutralisation also occurs in the form of ‘static’ constraints on lexical items. In Thai, which maintains a three term laryngeal contrast in word-initial stops, only a single series of stops (commonly described as phonetically plain voiceless) is present word finally at the lexical level. In other words, Thai lacks alternations between non-neutralised and neutralised forms of the type illustrated in (4). Similar lexical constraints against the marking of laryngeal contrast word finally occur in several other (south)east Asian languages including Khmer (or Cambodian: Mon-Khmer), Manipuri (or Meithei), Tibetan

(both Tibeto-Burman) as well as Korean (Henderson, 1952; Bhat & Ningomba, 1997; Chang & Shefts, 1964; Cho, 1990a/1999). Furthermore, even languages that impose few or no restrictions on the marking of laryngeal contrast on single obstruents often restrict obstruent clusters to a single laryngeal specification, or suspend the marking laryngeal contrast on such sequences altogether. Finally, in many languages that mark laryngeal contrast on stops across contexts, the scope for marking the same contrast on fricatives is often more limited. For instance, Frisian lacks a contrast between fortis and lenis obstruents word initially, whilst the North Germanic languages neutralise the distinction across the board for sibilants, or, if /v/-type sounds are treated as sonorants (as in Vanvik 1972), for all fricatives.

Although these lexical, ‘static’, constraints on the occurrence of laryngeal contrast tend to receive less attention in the generative literature than their dynamic counterparts, they sometimes play an important role in theory formation. For example, arguments for and against syllable-driven analyses of (final) laryngeal neutralisation are ultimately based on the behaviour of word internal non-alternating obstruent + sonorant clusters.

The goal of this chapter is to present a number of observations about neutralisation phenomena and to assess these observations in the light of formalist and functionalist, cue-driven ideas about neutralisation phenomena. Although voicing assimilation rather than laryngeal neutralisation is the focus of the experimental part of this study, the survey presented here is important as a backdrop to the investigation of Hungarian RVA in chapter 6 and the results of the Dutch experiment reported in chapter 7. The conclusions of this chapter are also central to the critique of current autosegmental models of assimilation and neutralisation that is the topic of chapter 8.

Section 3.1 starts with the recognition that theories of laryngeal neutralisation consist of a component describing the nature of the assimilation process, and a second component detailing the set of contexts in which neutralisation occurs. In this regard, *context* should be understood in a broad sense so as to include the phonetic features of the obstruents targeted by neutralisation as well as prosody. This section proceeds to outline the two rivaling conceptions of the first component that are relevant to the wider purposes of this study and contrasts the relatively fragmented formalist view of the second component with the potential of a unified cue-based model.

The remaining sections examine to what extent the contrasting views of each component are supported by the available data. Thus, section 3.2 finds some evidence for the idea that laryngeal neutralisation results in phonetic underspecification and discusses experiments suggesting that laryngeal neutralisation is often phonetically incomplete. Sections 3.3 and 3.4 assess the effects of obstruent features on the stability of laryngeal contrast, discussing the evidence

for an asymmetry between plosives and fricatives but not between voicing and aspirating languages. There is some evidence that the instability of laryngeal contrasts in fricative inventories is due to the poor realisation of place cues in voiced fricatives and are therefore amenable to an explanation in terms of perceptibility. Next, 3.5 reviews the arguments levelled by Steriade (1997) against syllable-driven models of final neutralisation, which crucially fail in their predictions about the behaviour of nonfinal obstruent + sonorant clusters. By contrast, cue-based models give an accurate account of the observations that neutralisation in these clusters occurs in tautosyllabic as well as heterosyllabic sequences and does not necessarily coincide with word-final neutralisation.

Section 3.6 then tries to extend the cue-based models to the relatively uniform behaviour of laryngeal distinctions in obstruents appearing in (stable) word-initial and (unstable) word-final positions across left and right-hand phonetic environments. The main argument for this idea is that it is likely that prosodic strengthening and weakening have a levelling effect on the perceptibility of laryngeal contrast at word edges. If this is indeed the case, the extended cue-based model is preferable to both word-based formalist models and the *Paradigm Uniformity* of Steriade (1997) account on grounds of generality. Section 3.7 finally, presents a summary of the chapter and highlights some of the issues that are left unexplored.

### 3.1 Theories of (final) laryngeal neutralisation

Theories of laryngeal neutralisation consist of two broad components. The first specifies the nature of the neutralisation itself, i.e., its input (in the case of dynamic neutralisation) and the phonological and phonetic status of its output. The second component defines the phonetic environments which are likely to trigger neutralisation in substantive phonetic and/or prosodic and/or morphosyntactic terms. In this regard, the term *environments* should be taken in a broad sense, so as to include the phonetic features of sounds targeted by neutralisation. For instance, fricatives are more vulnerable to neutralisation than stops, and this effect of obstruent manner should be predicted by any complete theory of laryngeal neutralisation. The range of theories and models that have been proposed in the literature is far too broad to attempt a comprehensive survey here. Instead this section outlines a number of assumptions and predictions that are common to (semi-)formalist approaches, and contrasts these with the cue-based theory of laryngeal neutralisation proposed by Steriade (1997)

### 3.1.1 The nature of neutralisation processes

A pervasive assumption about the nature of final laryngeal neutralisation in obstruent inventories split by a fortis-lenis-type contrast is that it is a *fortition* process converting [-tense] obstruents into their [+tense] counterparts. In such theories, the [-tense] final obstruent of /krab/ in 4 changes into its [+tense] counterpart in the diminutive and citation forms, whereas neutralisation does not apply to the underlying /p/ of /xrɔp/, which already is [+tense].

In a more general sense too, accounts of laryngeal neutralisation often rest on the assumptions (a) that neutralisation is fundamentally asymmetric (b) that the output of neutralisation is identical to a series that is lexically present. These assumptions often play a role in (generative) analyses of final neutralisation in languages with richer laryngeal systems (e.g., Thai), which tend to identify the result of neutralisation with the lexical series realised as plain voiceless. They are often visible in accounts of other forms dynamic neutralisation: postnasal voicing, for instance, is often assumed to result in obstruents that are identical to actively voiced (lenis) stops (cf. Pater 1996/1999). Finally, assumptions (a) and (b) tend to be (implicitly) applied in the representation of lexical neutralisation patterns. and to the treatment of lexical neutralisation For example, English [s] + obstruent clusters are analysed as lexically [+tense] by Iverson & Salmons (1999) (see chapter 8). The key prediction of a fortition analysis of laryngeal neutralisation is that neutralised obstruents should behave as fortis obstruents both phonetically and phonologically.

A number of researchers propose an alternative view of laryngeal neutralisation that harks back to older, ‘archiphonemic’, and therefore symmetric conceptions of laryngeal neutralisation. According to this alternative view, neutralised obstruents form a class that is distinct from both their fortis and lenis counterparts and is therefore predicted to behave in a distinct fashion both phonetically and phonologically. Consequently, all lexically contrastive series are affected in the case of a dynamic neutralisation process: both the final /b/ of /krab/ and the final /p/ of /xrɔp/ in 4 transform into a third class of obstruents in the citation and diminutive forms. For convenience I will refer to this class of obstruents as [0tense], and denote individual members using capitals (/P,T,K,S/), following structuralist notational conventions.

In recent work, this symmetric conception of (laryngeal) neutralisation is typically tied to the concept of *phonetic* or *surface* underspecification in the sense of Keating (1988) or Pierrehumbert & Beckman (1988). Whilst formally similar to older archiphonemic analysis, the notion of phonetic underspecification critically refers to the phonetic behaviour of a sound rather than the (mere) suspension of phonological contrast (cf. chapter 1). This means that for a sound to be analysed as phonetically underspecified for a certain feature, it should behave passively with regard to all phonetic correlates of that feature. For example,

a [0tense] obstruent should be completely passively voiced in the sense defined in chapter 2, and as such be distinct from [+tense] obstruents that are actively devoiced as well as from [-tense] obstruents that are actively voiced.<sup>1</sup> Under-specification models of laryngeal neutralisation are usually associated with phonetic studies (Hsu, 1996; Ernestus, 2000) or functional models (Steriade, 1997), but do not strictly speaking pre-empt the choice for a theory of neutralisation contexts.

### 3.1.2 Syllable-driven approaches to laryngeal neutralisation

It has been common in the phonological literature to characterise the environments in which laryngeal neutralisation processes tend (not) to occur in terms of syllable structure. Two broad varieties of syllabic theory can be distinguished: one that tries to capture the set of contexts which are prone to neutralisation (treating sites resistant to neutralisation as the elsewhere environment); and one that tries to capture the set of environments in which laryngeal contrast is relatively robust (treating sites that are prone to neutralisation as the elsewhere case).

The first variety, exemplified by Mascaró (1987/1995), Cho (1990a/1999) treats final neutralisation as an operation targeting obstruents that appear in a syllable rhyme or (if it is recognised as separate entity) coda, but ignoring all other contexts. Thus, this approach tries to define in positive terms where laryngeal contrast is neutralised and treats the set of contexts where contrast is maintained as the elsewhere environment. Word-final obstruents are subject to the process because they are assumed to be parsed as rhyme constituents.

The second approach, adopted by, e.g., Gussman (1992) and central to the work of Linda Lombardi, instead tries to define the set of contexts in which contrast is *maintained* in positive terms, and specifies the set of contexts in which neutralisation operates as the complement of this environment. Under this analysis, word-final obstruents are targeted by neutralisation because they appear in an elsewhere context. For example, the (parametric) *Laryngeal Constraint* of Lombardi (1994, 1995a,b) states that laryngeal features are only licensed when they appear in the configuration in figure 3.1: the LAR class node that obligatorily dominates the substantive laryngeal features [voice], [asp], and [gl] should be dominated by a *root* node adjacent to a tautosyllabic sonorant (the formulation in Lombardi 1994 et seq. usually omits the root node). In languages with the Laryngeal Constraint switched on, lexically present LAR nodes appearing

<sup>1</sup>Whilst the [0tense] category identified by Ernestus (2000) and others and the lenis plosives of aspirating languages as analysed in chapter 2 both lack targets for voicing they are still distinct in that the latter but not the former category has targets for the other correlates of tense. Further research is needed to establish whether speakers and listeners can indeed distinguish between these two categories.

in any configuration that is different (such as word-final obstruents) are deleted, leaving a laryngeally unmarked obstruent, which in Lombardi's framework is equivalent to [+tense] (see chapter 8 for details).

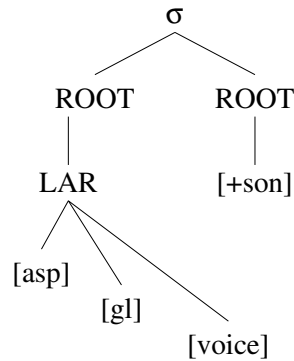


Figure 3.1: Licensing configuration for laryngeal contrast according to the *Laryngeal Constraint* (adapted from Lombardi 1994 with ROOT nodes added)

The differences between the two broad types of syllabic theory reside in matters of overall model architecture and in (alleged) formal efficiency or elegance. Generally speaking, the *licensing* approach envisaged by Lombardi (1994, 1995a,b) is part of a recent trend in generative phonology towards filter-cum-repair-rule or fully declarative models, whilst the ‘coda neutralisation’ theory stems from older rewrite rule-based phonological frameworks. In broad terms, the prediction of any syllabic theory is that the full set of neutralisation sites in a given language can be defined as a natural class in terms of syllable structure, and syllabic theories can only be distinguished from word or cue-based alternatives if the set of predicted neutralisation sites is in some way different from those emerging from morphosyntactic (or higher-level prosodic) or cueing considerations. Consequently, nearly everything in the defense of a syllabic theory of laryngeal neutralisation contexts hinges on an independently motivated theory of syllabification.

Models that combine a fortition approach with a syllable-driven account of neutralisation contexts are fairly frequent in the literature and may perhaps be regarded as the, ‘default’, or zero hypothesis, choice for generative models. The survey of voicing assimilation and laryngeal neutralisation phenomena presented by Mascaró & Wetzels (2001) is a case in point as it questions neither the validity of a fortition analysis nor the claim that final neutralisation can be syllable-driven, whereas it defends the case for a theoretically separate phe-

nomenon of *word-final* neutralisation. However, claims regarding the nature of laryngeal neutralisation and hypotheses about contexts in which neutralisation occurs are separate components of a theory and can therefore be combined freely. Thus, Trommelen & Zonneveld (1979) propose a fortition account of Dutch final neutralisation claiming that the process occurs finally at the end of words, or more precisely, at the end of constituents separated by analytical morphological boundaries. Gussman (1992) on the other hand analyses Polish neutralisation as an essentially symmetric process deleting [-voice] (i.e., [+tense]) as well as [+voice] ([-tense]) specifications operating in a set of contexts defined in terms of syllable structure.<sup>2</sup>

### 3.1.3 Cue-based approaches to laryngeal neutralisation

Steriade (1997) proposes a cue-based theory of neutralisation contexts as an alternative to, and improvement over, syllable-driven models, concentrating on the (universal aspects of the) role of phonetic context in determining laryngeal neutralisation in stops. The prediction of this theory, in line with cue-based approaches more generally (c.f. chapter 1), is that neutralisation of laryngeal contrast is more likely in environments where it is relatively hard to recover from the speech signal. The general thrust of Steriade's approach is consistent with both 'diachronic' and 'synchronic' theories of the relation between perception and phonology, but the specific model she proposes belongs to the 'synchronic school'. Consequently, Steriade (1997) derives hard synchronic implications from relative perceptibility such that the availability of laryngeal contrast in a context with perceptibility  $x$  implies the existence of laryngeal contrast in all phonetic contexts where its perceptibility is greater than  $x$ .

Early on in the paper, Steriade (1997) suggests the hierarchy in (5) to capture the relative perceptibility of fortis-lenis type laryngeal contrast in stops, where 'V' represents a vowel, '>' indicates greater perceptibility, and '||' is used to represent a physical pause. This hierarchy is expanded and refined in the course of the argument, but it suffices to illustrate the basic idea of Steriade's theory. It states, for example, that *ceteris paribus*, the fortis-lenis contrast is more perceptible between a vowel and a following sonorant (including vowels) than between a vowel and another obstruent. According to the 'diachronic' version of functional phonology this means that learners of a language that has contrasts in both contexts are most likely to misperceive it as non-existent in the latter and neutralise it in their own grammars. In Steriade's 'synchronic' model it means that constraints on laryngeal contrast in V\_[-son] contexts outrank those that ban contrast in V\_[+son] environments, and the prediction follows that a language

<sup>2</sup>Gussman uses a 'late' default rule to insert [-voice] on [0tense] obstruents in environments where they are phonetically voiceless.



that maintains contrast in V<sub>-</sub>[-son] also maintains it in V<sub>-</sub>[-son] (and in the intervening environment V<sub>-</sub>||).

- (5) Perceptibility hierarchy for laryngeal contrast in stops, according to (Steriade 1997:12)

$$V_{-}[+son] > V_{-}|| > V_{-}[-son] > \{[-son]_{-}[-son], [-son]_{-}||, ||_{-}[-son]\}$$

The hierarchy in (5) is mainly based on the *number* of cues that are potentially available in each context. Thus, in V<sub>-</sub>[+son] plosive *onset* cues (preceding vowel duration, V-to-C formant transitions, preaspiration or preglottalisation), *internal* cues (presence and quality of voicing, duration), as well as *offset* cues (duration and relative amplitude of the release burst, postaspiration or postglottalisation, C-to-V formant transitions) are available to signal the distinction between fortis and lenis plosives. The V<sub>-</sub>[-son] context is somewhat poorer in cueing potential in lacking at least the V-to-C formant transitions and depending on manner of articulation of [-son] and the amount of coarticulation, the release and postaspiration cues too (a refined version of the hierarchy would need to separate the various options here). It might be surmised from this that the fortis-lenis distinction is less salient in V<sub>-</sub>[-son] than in V<sub>-</sub>[+son] environments. In a number of instances the rankings on Steriade's perceptibility hierarchy are directly backed up by experimental evidence. For example, it Raphael (1981) shows how plosive onset cues are perceptually less salient than offset cues: in case of conflicting information, listeners give priority to the latter. This means that laryngeal contrast in stops lacking offset cues is less perceptible in those that do not, even if the different *numbers* of cues themselves should be demonstrated to have no effect on perceptibility.

Although inferences made on the basis of numbers of (potential) cues are not ultimately valid indicators of relative perceptibility, the set of individual rankings that make up the hierarchy in (5) all amount to testable hypotheses. For instance, the claim that a given laryngeal contrast, realised by a particular set of phonetic cues, is less perceptible (universally and/or to native speakers of a particular language) in the context [-son]<sub>-</sub>[-son] than in V<sub>-</sub>[-son], can be tested using the type of experimental methodology employed by, e.g., Mielke (2001). Thus, both the perceptibility hierarchy in (5) and the predictions derived from it can be falsified independently of each other. On the other hand, as long as they are interpreted as autonomous grammatical mechanisms, the Laryngeal Constraint and similar devices remain stipulations, even if to some extent they share the spirit of the hierarchy in (5) (i.e., in recognising that laryngeal contrast is relatively stable in the presence of a following sonorant).

Moreover, a cue-based theory of laryngeal neutralisation bears the promise of unifying the description of all laryngeal neutralisation asymmetries in terms

of a single mechanism. This would be an achievement beyond anything formalist models seem to be able to deliver. Recall from chapter 1 that the relative perceptibility of a contrast is not just a function of the immediate phonetic environment of the sound bearing it, but also of its own acoustic properties, prosody, and phonetic knowledge of the listener (as gained from exposure during language acquisition). For example, the asymmetry between word-initial and word-final contexts in terms of their propensity for laryngeal neutralisation might be related to the phonetics of initial strengthening and its (presumed) effects on relative perceptibility (c.f. section 3.6). Similarly, given the observation that place cues are harder to detect in voiced than in voiceless sibilants, a cue-based theory predicts that contrastively voiced sibilant series are likely to have fewer distinctive places of articulation than their voiceless counterparts and more likely to be absent altogether from a given language (see section 3.4). In other words, it is predicted that laryngeal contrast is relatively unstable in sibilant fricatives. This prediction is a simple instantiation of the basic hypothesis that neutralisation is a function of relative perceptibility and does not require any additional mechanisms. Although there are (semi-)formalist models (Steriade 1992, 1993; Vaux 1998: see chapter 8) that address the relative instability of laryngeal contrast in (sibilant) fricatives, I am not aware of any such models that employ the same formal mechanism to account for segment internal and context biases in laryngeal neutralisation.

## 3.2 The phonetics of laryngeal neutralisation

This section summarises a two sets of observations concerning the phonetics of laryngeal neutralisation. First it reviews phonetic evidence indicating that laryngeal neutralisation can lead to phonetic underspecification of [tense]. It then considers a second body of phonetic data, which suggests that processes traditionally described as categorical neutralisation processes are in fact phonetically incomplete. The first type of evidence data is problematic for a fortition analysis of laryngeal neutralisation because it requires neutralised obstruents to be classified as distinct from fortis as well as lenis obstruents. The second type of evidence may eventually undermine a more fundamental assumption of generative analyses, namely the idea that every surface form is derived from a single underlying representation.

### 3.2.1 Laryngeal neutralisation as phonetic underspecification

Whereas fortition accounts and more generally lexical feature accounts of laryngeal neutralisation predict that neutralised obstruents should be phonologically and phonetically identical, phonetic underspecification models predict that pho-

netically speaking they are distinct from all lexically contrastive series. One proponent of the phonetic underspecification approach to laryngeal neutralisation is Ernestus (2000), who argues that the behaviour of word-final obstruents in Dutch is consistent with the absence of voicing targets. For example, Dutch final obstruents are voiceless utterance finally but often audibly voiced before vowel-initial enclitics, i.e., at weak prosodic boundaries before unstressed syllables (cf. Gussenhoven 1986): [hɛbɪk] for /hɛb + ɪk/ *have I* as opposed to [ɪkhɛp] *I have*, whilst /datɪk/ *that I* is often realised as [dadɪk]. As noted in chapter 2, fortis plosives occurring in word internal intervocalic environments are generally mostly voiceless in Dutch. Ernestus (2000) reports a listening experiment showing that phonetically trained native speakers of Dutch are able to draw a three way distinction among word-medial fortis and lenis and word-final plosives, which suggests that observed voicing distinctions between unambiguously fortis obstruents and word-final obstruents indeed reflect a difference in underlying targets. In addition, the fact that word-final stops, as opposed to fortis-initial plosives, are realised with a negligible amount of glottal abduction in Dutch (Yoshioka et al., 1982) is also consonant with the idea that they have no phonetic targets for [tense].<sup>3</sup> On similar grounds Hsu (1996) argues that laryngeal neutralisation of (the three term distinction in) Taiwanese stops results in phonetic underspecification.

To what degree the surface underspecification account of laryngeal neutralisation can be extended beyond the data considered by Hsu (1996) and Ernestus (2000) is an empirical matter. One case where the suspension of laryngeal contrast seems to coincide with the absence of voicing/VOT targets is the realisation of word-initial sibilant + plosive clusters in the aspirating varieties of Germanic. Whereas the [tense] plosives of these languages are realised with aspiration and a long lag VOT word initially, plosives preceded by tautomorphic [s] are normally voiceless unaspirated and have a short lag VOT. In the light of the fact that all the languages and dialects in question suspend laryngeal contrast in word-initial sibilant + plosive clusters ([sp, st, sk, ʃp, ʃt] never contrast with, e.g., [zb, zd, zg, ʒb, ʒd]) might be interpreted as evidence that they are [0tense] rather than [+tense] and phonetically underspecified for [tense] correlates. On this analysis the plain voiceless realisation of sibilant + stop clusters follows directly from passive (de)voicing. As argued in chapter 2 it is plausible that the initial sibilant is subject to passive devoicing because a large amount of glottal abduction represents the ideal configuration for the production of high-intensity noise (measurements reported by Yoshioka et al. 1981 indicate that the initial sibilant is indeed produced with a large amount of glottal abduction). The

<sup>3</sup>Yoshioka et al. (1982) suggest that the lack of glottal abduction in Dutch word-final stops is due to glottalisation as in English. However, in English and other languages with glottalisation, irregular voicing occurs during and in the vicinity of stops affected by the process, whereas no such effect has been reported for Dutch.

voiceless realisation of the initial sibilant removes any source of passive voicing for the adjacent plosive which will therefore be produced without any voicing even in the absence of active devoicing measures. If sibilant + plosive clusters are produced without active (de)voicing measures, their short lag VOT is also predictable on aerodynamic grounds: voicing sets in as soon as the transglottal pressure difference allows for it, i.e., shortly after the oral release of the plosive.

The hypothesis that word-initial sibilant + plosive clusters are [0tense] of course demands that they should be phonetically distinct from both their [+tense] and [-tense] counterparts with regard to correlates of the fortis-lenis distinction, that is, unless it can be shown that a phonetic feature of one of the latter two classes reflects passive behaviour as well (the lack of voicing targets for lenis stops in aspirating languages being a case in point). For example, if sibilant + plosive clusters are [0tense], they should have different effects on the  $F_0$  of a following vowel than the corresponding singleton [+tense] and [-tense] plosives. Unfortunately, data on  $F_0$  trajectories is inconclusive. Caisse (1982) and Ohde (1984) find similar pitch perturbations after fortis stops and [s] + stop clusters, which suggests they should be classified together. On the other hand, for 4 out of the 5 speakers investigated by Kingston & Diehl (1994), the points on the  $F_0$  trajectories after sibilant + plosive clusters are roughly intermediate between the pitch values following singleton fortis and lenis stops, as predicted by the phonetic underspecification hypothesis.

It does seem clear, meanwhile, that not all cases of laryngeal neutralisation lead to phonetic underspecification, since there are several cases in which (lexical) neutralisation produces obstruents that appear to have voicing targets. The word-initial alveolar sibilant of German is realised as [z], despite the fact that it does not contrast with a voiceless fricative at the same place of articulation. A similar phenomenon occurs in eastern dialects of Dutch, whilst the standard variety has [z] in sibilant + [v] clusters (e.g., [zwa:n], *swan*; [zve<sup>h</sup>vən] *to hover, glide*; [zvuu:t], *rind*). An example from outside the Germanic group is Western Aleut, which has a single series of oral stops with word-initial VOTs ranging between 76 and 92 ms (Cho & Ladefoged, 1997), i.e., values well within the > 35 ms bracket that is normally labelled as aspirated or long lag VOT.

It should be clear from the description of obstruent aerodynamics in chapter 2 that neither the voiced realisation of the German and Dutch fricatives nor the long lag VOT of Western Aleut plosives can be attributed to passive voicing. Assuming that neutralised sibilant fricatives are optimised for noise generation, which involves a wide glottal abduction, it is difficult for them to acquire any voicing by passive means, unless they are subject to lenition and occur intervocally. Neither of these two conditions have been described as necessary for the Dutch and German sibilants in question to be pronounced as [z]. With regard to passively devoiced oral stops, aerodynamics dictates that voicing com-

mences as soon as the two basic preconditions (closed vocal folds and sufficient subglottal pressure) are fulfilled. For a plosive-vowel sequence produced on a voiced carrier signal, the time lag from oral release within which these conditions are met is likely to be considerably shorter than the VOTs observed for Western Aleut (c.f. Westbury & Keating 1986; Stevens 1998). Consequently, German and Dutch (dialectal) word-initial [z] must be regarded as specified for phonetic voicing, whilst Western Aleut stops can only be analysed as actively devoiced (aspirated). Without further evidence it is impossible to tell whether these sounds share other phonetic properties normally associated with [-tense] and [+tense] respectively.

It is perhaps important to emphasise that evidence for the phonetic underspecification of neutralised obstruents does not strictly speaking preclude a fortition analysis of laryngeal neutralisation, which must consider phonological as well as phonetic data. A case in which phonological information must be brought to bear on the analysis is the lexical neutralisation of [tense] for velar stops in Dutch. Whilst Dutch opposes /p-b/ and /t-d/ there is only a single velar stop, which is usually transcribed [k]. A reason to represent this stop as lexically fortis across contexts might be that it groups with /p, t/ in the regular past tense paradigm: /ra:k/ + /də/ yields [ra:ktə], *hit (a target)*, just as /ra:p/ + /də/ surfaces as [ra:ptə], *gathered, was picking up*.

However, technically speaking, a fortition analysis of dynamic neutralisation can be maintained even in the absence of this sort of phonological information. In any model that conceives of the mapping between lexical forms and representations at the physical (auditory and articulatory) interface level in derivational terms, it is possible to say that laryngeal neutralisation is a process that converts [-tense] obstruents into their [+tense] counterparts. The distinct phonetic interpretation of neutralised obstruents can then be relegated to a separate mechanism that implements neutralised [+tense] obstruents differently from contrastively [+tense] obstruents. However, for this approach to work, the second mechanism has to operate in exactly the same set of environments as the fortition rule, and to a large degree therefore duplicates it. Consequently, in the absence of evidence that a given set of laryngeally neutralised and phonetically underspecified plosives and fricatives act as fortis obstruents in phonological processes, the phonetically opaque version of the fortition hypothesis cannot be justified.

### 3.2.2 Incomplete laryngeal neutralisation and its implications

Whilst the phonetic data uncovered by Hsu (1996) and Ernestus (2000) undermines the asymmetric aspect of the fortition analysis of (final) laryngeal neutralisation, its *categorical* nature has also been challenged on phonetic grounds. An ever-growing series of production and perception studies reveals that speakers and listeners are able to make subtle but statistically significant phonetic dis-

tinctions between fortis and lenis obstruents in neutralisation contexts, which suggests that laryngeal neutralisation is phonetically incomplete. Mitleb (1981), Port et al. (1981), O'Dell & Port (1983), Charles-Luce (1985), Port & Crawford (1989), Piroth et al. (1991) find evidence that laryngeal neutralisation in German leaves some residual cues to the lexical status of obstruents. Similar results have been reported for Catalan (Dinnsen & Charles-Luce, 1984; Charles-Luce, 1993), Polish (Slowiaczek & Dinnsen, 1985; Slowiaczek & Szymanska, 1989), Romanian (Steriade & Zhang, 2001), and Dutch (Ernestus & Baayen, 2003).

In contrast, Fourakis & Iverson (1984) and Kahlen-Halstenbach (1990), and Jassem & Richter (1989) find phonetically complete neutralisation for German and Polish respectively. In addition, experiments reported by Jongman et al. (1992), Baumann (1995), Ernestus (2000), and Kopkalli (1993) indicate that Dutch and Turkish final neutralisation erases all phonetic distinctions between word-final fortis and lenis obstruents. Finally, an effect of juncture strength on the degree of neutralisation emerges from acoustic data gathered by Piroth et al. (1991), who finds incomplete neutralisation at word boundaries in German, but complete neutralisation before word internal morpheme boundaries.

What is at issue in the incomplete neutralisation debate that was sparked by the earlier of these studies is not so much the (replicability of their) bare results, but rather the nature of phonetic knowledge and the experimental methodology that lends the best insight into that nature. For instance, Fourakis & Iverson (1984) claim that the incomplete neutralisation of final obstruents in German observed by Mitleb (1981), Port et al. (1981), and O'Dell & Port (1983), is due to a special mode of pronunciation that bypasses phonological rules used in normal communication and is directly driven by the orthography. They support this analysis with the results of two experiments, one of which examines the realisation of word-final obstruents in the responses to a reading task. Two out of the 4 subjects involved in this experiment, which is designed to replicate the effects found in the earlier studies, produced residual cues to the underlying distinction between fortis and lenis obstruents. However the same speakers did not distinguish underlying word-final fortis and lenis obstruents in the third person preterite indicative singular forms of strong verbs such as /la:d(ə)n/ (*to load*) in a verb conjugation task that relied on orally presented stimuli. Thus, the pair of experiments conducted by Fourakis & Iverson (1984) appears to establish a connection between orthographically presented stimuli and the incomplete neutralisation effect.

Fourakis and Iverson use the observation that in the reading experiment 3 out of 4 their subjects distinguish the members of the minimal pair <Weck> *breakfast roll (dial.)* vs. <weg> 'away' (by means of vowel duration) as their trump card in this argument. Since <weg> never inflects, learners of German have no auditory evidence that its underlying form is distinct from that of <Weck> and

the only route to a lenis pronunciation of the final velar is through the orthography. Note, however, that this argument only goes through on the assumption that orthographic information has no bearing on underlying phonological forms (and see Giegerich 1994 for evidence that this assumption is not necessarily valid).

Baumann (1995) echoes the idea that incomplete neutralisation reflects a special, spelling driven, mechanism that has little relevance outside the laboratory. She notes a number of further experimental factors that seem to reduce the theoretical import of incomplete neutralisation data. For instance, some studies (Port & O'Dell, 1985; Charles-Luce, 1985) use very low frequency words as stimuli, and many of the test subjects probably were fairly advanced speakers of English, which suggests the possibility of second language interference, especially given the fact that most of the studies listed above were carried out in laboratories in the United States. In addition, the use of minimal pairs may have revealed the purpose of the experiments to the subjects and thereby prompted the observed response patterns.

However, it is hard to find a systematic connection between the possible confounds listed by Baumann (1995) and the designs of the various studies reporting incomplete neutralisation effects. For instance, Catalan orthography does not represent the contrast between word-final lenis and fortis obstruents, yet Dinnsen & Charles-Luce (1984) and Charles-Luce (1993) find that the distinction between the two types may be incompletely neutralised, especially if the words acting as carriers are semantically nonredundant. The Dutch subjects used by Jongman et al. (1992) and Baumann (1995) were probably fairly proficient in one (English) and possibly in a second (French) non-neutralising language. Furthermore, Charles-Luce (1993) argues that in at least one respect the design used by Fourakis & Iverson (1984) was no less 'artificial' or biased than some of the experiments that established incomplete neutralisation: the conjugation task removes all semantic context, which is hardly representative of the average conversational situation.

As the causes of incomplete neutralisation effects remain controversial, there is no agreement about their analysis. One possible account, suggested by Dinnsen & Charles-Luce (1984) among others, effectively treats incomplete neutralisation as the result of gradient phonetic processes rather than a categorical phonological ones. According to this type of account, the final obstruent of, say, Dutch /krab/ would be [-tense] at both the lexical and surface phonological levels, just as its English cognate /kræb/. The Dutch plosive would then be converted into a series of phonetic targets that are very close to those for the final obstruent of /ɣrap/, whereas the phonetic interpretation of English word-final /b/ would maintain a much more perceptible phonetic contrast with /p/. This approach is perfectly feasible within a generative model, even if it may involve the admission that the phonetic (or 'scalar phonological') component of

the grammar is larger than previously thought. From a functionalist perspective on the other hand, it raises the awkward questions why and how speakers maintain and acquire contrasts that appear to be so near the absolute thresholds of perceptibility.

An alternative approach, favoured by, e.g., Ernestus & Baayen (2003), views incomplete neutralisation effects as products of intraparadigmatic interference in speech production and perception. This approach is founded on a view of lexical organisation and lexical access that is different from the one adopted in much generative work. According to the latter, there is little redundancy in the lexicon, most roots are represented only once, and morphologically complex forms are derived on-the-fly from simplex forms. According to the former view, the mental lexicon stores both bare roots and many morphologically complex forms (e.g., Bybee 2001). The activation of a specific instance of a given root during lexical access is then assumed to co-activate (to a somewhat lesser extent) the phonological representation of related forms. Where there are phonological differences among the various manifestations of the root in question, this in turn leads to interference in the production and perception of specific instances of that root because all activated entries (can) feed into the production and perception systems.

On this second type of account, the residual phonetic contrast between the final labial stops of [kɪɒp] and [χɪɒp] would arise from interference from the full-blown phonetic contrast between the labial stops in the corresponding plural forms [kɪɒbə(n)] and [χɪɒpə(n)]. In other words, the pronunciation of the final labial stop in [kɪɒp] would tend to be slightly more [b]-like than that in [χɪɒp] because in producing the former, speakers would coactivate the lexical entry for [kɪɒbə(n)], whereas the final stop of the latter would be ‘biased’ by the /p/ of [χɪɒpə(n)].

Generative phonological theory can accommodate accounts of incomplete neutralisation that are based on intraparadigmatic interference, but only as grammar-external, ‘performance’ mechanisms, since it does not allow for notions such as (spreading) activation. It would be possible to say, for example, that at some abstract level speakers ‘know’ that the pairs [kɪɒp]-[kɪɒbə(n)] and [χɪɒp]-[χɪɒpə(n)] each derive from a single underlying root, but that the way this knowledge is implemented in the minds/brains of speakers gives rise to intraparadigmatic interference and, consequently, incomplete neutralisation effects. However, this argument only holds as long as there is an independent argument for differentiating competence from performance; and the latter argument would seem to be weakened by the viability of non-modular models of the phonetics-phonology interface (cf. 1.3.2).

From a functionalist perspective, accounts of incomplete neutralisation that are based on interference from alternate representations are more attractive than



an account based on phonetic rules that are acquired directly from auditory information. Under an interference account, speakers would not need to extract any subtle contrast between the labials in [kɪap] and [χɪap] from the signal in order to be able to produce it, at least as long as they do extract the contrast between the corresponding plural forms. Acquisition of the latter contrast would automatically trigger interference and hence incomplete neutralisation. Note however, that this line of reasoning does not in itself provide functional grounding of lexical redundancy and coactivation.

### 3.3 Obstruent feature asymmetries: voicing vs. aspirating languages

A cue-based theory predicts that the probabilities of neutralisation affecting two particular forms of laryngeal contrast differ if the sets of phonetic cues associated with each contrast differ in overall relative salience. For example, it might be that the phonetic difference between (actively devoiced) plain voiceless stops such as [p, t, c, k, q] and their ejective counterparts [p', t', c', k' q'] is less salient to listeners than the difference between the former series and their aspirated counterparts [p<sup>h</sup>, t<sup>h</sup>, c<sup>h</sup>, k<sup>h</sup>, q<sup>h</sup>]. If this is indeed the case, a cue-based theory would predict that the voiceless-ejective contrast is more prone to neutralisation than the voiceless-aspiration contrast.

It is beyond doubt that asymmetries of this sort do not appear in categorical fashion: laryngeal contrasts of all phonetic types are more or less subject to (dynamic) neutralisation (e.g., Lombardi 1994). Meanwhile, there is no clear evidence, let alone statistically reliable generalisations, about different tendencies towards neutralisation for phonetically different types of contrast. There is some anecdotal data suggesting that closure voicing contrast is more robust in languages that cross-classify voicing and aspiration, such as Sanskrit, Bangla, or Khasi which are sometimes described as *deaspirating* before obstruents and/or word finally (cf. Kostic & Das 1972; Nagajara 1990; Steriade 1997), rather than *devoicing*. In other words the suggestion is that neutralisation leads to an opposition between e.g., [p, t, k] vs. [b, d, g] rather than [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>] and (say) [p<sup>h̥</sup>, t<sup>h̥</sup>, k<sup>h̥</sup>]. However, recall from the previous section that Thai drops its prevoiced rather than aspirated series before sonorants, and given the lack of (typological) data on the production and perception of four-term systems mixing voicing and aspiration, it is utterly impossible to draw any conclusions from impressionistic and anecdotal data like this.

The only reliable generalisation that can be made about laryngeal neutralisation in the two types of fortis-lenis languages that are the focus of this study is that (dynamical) final neutralisation occurs frequently in both voicing and aspirating languages. The Germanic group itself exhibits the full paradigm of

voicing languages with final neutralisation (standard Dutch, Frisian), aspirating languages with neutralisation (standard German), voicing languages without neutralisation (Yiddish), and aspirating languages without neutralisation (standard varieties of English, Norwegian, Swedish). Outside Germanic, dynamic final neutralisation is common in Slavonic (which is generally voicing) and Turkic (which is generally aspirating: c.f. [Johanson & Csató 1998](#)). Interestingly, incomplete final neutralisation effects have been found for both voicing (Polish: [Slowiazcek & Szymanska 1989](#)) and aspirating (German: [Mitleb 1981](#) and later studies cited in 3.2.2) languages, and the same applies to experiments showing phonetically complete neutralisation (voicing Dutch vs. aspirating Turkish: [Baumann 1995](#); [Kopkalli 1993](#)). A final indication that (the development of) final neutralisation is at least not highly sensitive to the voicing-aspirating distinction (or vice versa) is that voicing dialects of aspirating standard languages (Scottish English and Rhineland German) seem to follow the standard in terms of final neutralisation, and the same applies to aspirating dialects of voicing standards (north-eastern varieties of Dutch).

The safest approach to (final) laryngeal neutralisation phenomena is therefore to analyse them along the same lines as the ‘Germanic past tense paradigm’ discussed in chapter 4, by treating them as independent from the voicing-aspirating distinction. In the absence of data about the relative perceptual efficacy of voicing vs. aspirating fortis-lenis distinctions (or rather the phonetic contrasts between final stops: see chapter 2) there is no ground for establishing whether this assumption clashes with a cue-based model.

### 3.4 Obstruent feature asymmetries: plosives vs. fricatives

Much research on the marking of laryngeal contrast in obstruent inventories notes a relative scarcity of voiced obstruents. [Maddieson \(1984\)](#) observes that 291 out of 317 languages (91.8%) in the 1984 version of the UCLA Phonetic Segment Inventories Database (henceforth UPSID<sub>317</sub>) have a plain voiceless stop series whereas 212 (66.9%) have a series of plain voiced stops, where a *series* is defined as the presence of at least one stop of a particular type in an inventory. An even stronger preference can be observed in the same database for voiceless over voiced fricatives and especially voiceless over voiced sibilants. [Balise & Diehl \(1994\)](#) find that 74.5% of all sibilants in UPSID<sub>317</sub> are voiceless as opposed to 61.9% of nonsibilant fricatives and 61.3% of stops (the overall figure for fricatives is 68.9%). They report that a similar generalisation emerges from the survey by [Ruhlen \(1975\)](#) of 706 languages.

In the light of a potentially critical lack of phonetic detail in UPSID<sub>317</sub> (see below in this section), it is probably safer to consider the frequencies of certain

types of *contrast* instead of the frequencies of phonetic categories themselves. This approach yields an even stronger asymmetry in laryngeal marking between fricatives and stops. For example, the number of languages that have a contrast between unaffricated plosives labelled by UPSID<sub>317</sub> as ‘voiced’ and/or ‘voiceless’ and/or ‘voiceless unaspirated’, i.e., some sort of VOT distinction, in at least one of the labial, coronal (dental through retroflex, excluding postalveolars) or dorsal (palatal through uvular, including postalveolars) regions, is 236 (74.4%). By contrast, the number of languages that has a contrast between fricatives labelled as ‘voiceless’ and ‘voiced’ in one of three regions that can be referred to as *anterior nonsibilant*, *sibilant*, and *posterior nonsibilant* is 119. This amounts to 40.5 % of the languages that have fricatives and 37.5% of the total number of languages in UPSID<sub>317</sub>. The number of languages that have at least 1 sibilant labelled as ‘voiceless’ and 1 sibilant that is labelled as ‘voiced’ is even smaller at 101 (34.4 and 31.9%). This suggests that laryngeal contrast (supported by voicing distinctions) is rarer or less stable in fricatives than in plain stops.<sup>4</sup>

In addition, there is evidence of a (thus far) more anecdotal nature for the relative instability of laryngeal contrast in fricative inventories. For example, whereas the plosive series of all Germanic languages are largely split by a fortislenis contrast, the North Germanic group lacks laryngeal distinctions between sibilants, and between fricative series altogether if weak (sonorant-like) sounds such as [v/v, j, ʝ] are excluded. A lexical constraint against the marking of laryngeal contrast in English fricative + sonorant clusters, discussed in section 3.5 below, also evinces the relatively weak capacity of fricatives to support (voicing-based) laryngeal contrast. A more restricted version of this constraint is active in Dutch, which has only [s] before /m, n, l/ and as mentioned in 3.2.1 above,

<sup>4</sup>For the purpose of these counts I used the label *anterior nonsibilant* for fricatives classified by UPSID<sub>317</sub> as ‘nonsibilant’ and bilabial through retroflex in terms of articulation place, but excluding postalveolar sounds; *posterior nonsibilants* refers to ‘nonsibilant’ postalveolar and palatal sounds through to uvulars. Note that there is evidence for the treatment (in UPSID<sub>317</sub> and here) of sibilancy and place of articulation as separate dimensions in perceptual phonetic space (Choo & Huckvale, 1998).

Pharyngeals, epiglottals and glottals were excluded from both plosive and fricative counts as were geminates and sounds marked with the superscripts 2 (unassimilated loans), 3 (posited underlying segment), 4 (segment possibly derivable from others), or 5 (particularly vague or contradictory description) in Maddieson (1984), but note that this has little effect on the resulting numbers and percentages. The languages in UPSID<sub>317</sub> (Burmese, Karen, Mazahua) that have aspirated fricatives invariably have ‘voiceless’ and ‘voiced’ fricatives at the same place of articulation, and the inclusion of this third phonetic category therefore has no bearing on the results. Finally, the definition of ‘aspiration’/‘voicing’ contrast in terms of broad place of articulation used here is more restrictive than the definition implied by the criteria for a *series* in Maddieson (1984). According to the latter, an inventory with three plain plosives labelled as [b, t, k] (e.g., Seneca) would exhibit laryngeal contrast. Under the former definition this inventory does not show laryngeal contrast because there is no pair of plosives with distinct voicing in at least one of the labial, coronal or dorsal regions.

[z] preceding /v/.

Ohala (1983) and many others have sought to attribute the apparent bias against voicing distinctions in obstruent systems, and fricative inventories in particular, to the relative amount of articulatory effort involved in the production of actively voiced obstruents. This hypothesis recognises that, under circumstances defined in chapter 2, obstruent voicing requires some form of active enhancement whereas voicelessness can be realised without articulatory intervention. From this it is inferred that voiced obstruents are more costly in terms of articulatory effort, and in a functional model built on effort avoidance it follows that voiced obstruents tend to be rejected in favour of easier to produce voiceless obstruents. Taking this line of reasoning a step further, Ohala (1983) and Vallée et al. (2002) suggest that voiced fricatives are disfavoured more strongly than (pre)voiced plosives because the precise co-ordination between glottal and supraglottal constrictions that is critical in the production of the former requires more effort than articulation of the latter (cf. chapter 2).

Balise & Diehl (1994) on the other hand, propose that the typological bias against (voicing-supported) laryngeal contrast in fricative systems derives at least in part from the fact that the presence of voicing interferes with the perception of place cues in fricatives, at least for speakers of English. According to this theory, voicing-based laryngeal contrasts in fricative inventories tend to be avoided (or neutralised diachronically) because it is relatively hard to recover their place cues. It is supported by two types of observation: (1) the presence of voicing in a fricative reduces the amplitude of friction noise, which is an important carrier of place information; (2) studies of consonant confusions indicate that across various signal-to-noise ratios, voiceless fricatives are identified correctly more often than their voiced counterparts. In addition, Balise & Diehl (1994) report data suggesting that the latter effect is stronger for sibilants than for nonsibilants, which might account for the (slightly) greater tendency towards laryngeal neutralisation in sibilant inventories.<sup>5</sup>

The findings reported by Balise & Diehl (1994) and their (admittedly tentative) interpretation of those findings are highly interesting from the perspective of a cue-based theory of laryngeal neutralisation because they suggest that laryngeal neutralisation asymmetries related to adjacent sounds and the effects of phonetic features of the target obstruents themselves can be explained in terms of a single factor: relative perceptibility. In other words, the propensity for laryngeal neutralisation in pre-obstruent contexts (see section 3.5 below) and the observation that fricative inventories are prone to laryngeal neutralisation might be accounted for by the low perceptibility of lexical contrasts in both phonetic

---

<sup>5</sup>Interestingly, the perception data indicates that confusion along the place axis is far greater than along the voicing axis. For example, [v] is more likely to be confused with [ð] than with [f]. This observation would seem to argue against the proposal by Steriade (1992, 1993) that fricatives are less suited for the expression of laryngeal contrast.

‘contexts’. This goal seems unattainable to formalist models, which seem to be forced to rely on separate mechanisms to account for context effects (e.g., the Laryngeal Constraint in figure 3.1) and asymmetries between different obstruent types (i.e., if they deal with observations of the latter type at all: cf. chapter 8).

This is not to say that there is anything near a sufficiently fleshed out perceptual account of plosive-fricative asymmetries in laryngeal neutralisation. Such an account would have to be part of a broader theory of obstruent internal phonetic effects on laryngeal neutralisation, which would require a great amount of additional data. For instance, as indicated in the previous section, it is not clear whether laryngeal contrasts in stops with different phonetic expressions exhibit neutralisation asymmetries. Second, the low perceptibility of place contrast in the voiced fricative series highlighted by Balise & Diehl (1994) might account for neutralisation of voicing-assisted laryngeal contrast in fricative inventories, but it does not necessarily explain the higher incidence of neutralisation in fricative inventories than in stop inventories. It has been suggested that voicing affects the perception of place contrast in stops as well (e.g., Boersma 1998), and according to the theory described here that would mean that stop inventories would also gravitate towards an exclusively voiceless state. Thirdly, because it essentially treats laryngeal neutralisation of fricative oppositions as an epiphenomenon of place neutralisation in the voiced series, Balise and Diehl’s theory predicts that the process always yields a voiceless (actively devoiced) fricative. This prediction is contradicted by the voicing of word-initial alveolar sibilants in German and similar cases of neutralised fricative voicing (see sections 3.2.1 and 3.5).

The biggest caveat associated with both phonetic-substance based theories of neutralisation in obstruent inventories mentioned in this section, however, is the reliability of the typological generalisations they seek to account for, and especially of those relating to plosives. UPSID<sub>317</sub> and similar databases, as many of their sources, partially or wholly conflate the distinction between voicing as a phonetic feature and the phonological contrasts it is used to support. Thus, UPSID<sub>317</sub> reduces the four way phonetic voicing distinction recognised in this study among actively devoiced aspirated (fortis), actively devoiced unaspirated (fortis), passively voiced (lenis or neutralised) and actively voiced (lenis) to a three term taxonomy of voiceless aspirated, plain voiceless, and plain voiced, and does so in a way that leaves the true phonetic voicing of plosives partially opaque. For example, it represents the fortis-lenis distinction in both German and Bulgarian plosives as plain voiced vs. voiceless aspirated, even though it is clear that German is an aspirating language (Moulton 1962; Jessen 1998 and a host of references in the latter) and Bulgarian a voicing language (Ternes & Vladimirova-Buhtz, 1999). This means that the fact that both languages exhibit a VOT-based (fortis-lenis) distinction is correctly encoded, but the difference

between German (word-initial) [b, d, g] and Bulgarian [b, d, g(, b<sup>j</sup>, d<sup>j</sup>, g<sup>j</sup>)] is obscured, as is the phonetic distinction between German (word-initial, prestress) [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>] and Bulgarian [p, t, k(, p<sup>j</sup>, t<sup>j</sup>, k<sup>j</sup>)]. Consequently, generalisations drawn from UPSID<sub>317</sub> about the relative frequency of ‘voiced’ stops are in part about passively voiced lenis stops, and the total set of ‘voiceless aspirated’ stops includes at least some (actively devoiced) plain voiceless items.<sup>6</sup>

It is tempting to attach a greater amount of phonetic realism to the labels *voiceless fricative* and *voiced fricative* in UPSID<sub>317</sub>. First, it seems that the phonetic typology of laryngeal contrast is much simpler in fricatives than in stops, and this constrains the space for error (i.e., given the set of labels used by UPSID<sub>317</sub> to mark laryngeal distinctions in fricatives). In Germanic, fortis-lenis distinctions between fricatives are always supported by voicing, regardless of the use of voicing/VOT to cue stops. Furthermore, the number of fricative inventories divided by more than 2 contrastive ‘laryngeal’ dimensions seems genuinely low in comparison to the number of stop inventories with this property. This may well be another indication of the relatively restricted phonetic means of expressing laryngeal contrast in fricatives: note that the continuous high airflow across an oral constriction required for the production of fricative noise puts inherent limitations on the number of laryngeal actions and configurations that are available. In other words, it seems likely that in many instances where UPSID<sub>317</sub> represents a fricative as voiced, it is indeed produced with some amount of active voicing. I have certainly not found any glaring problems of the German/Bulgarian type mentioned in the previous paragraph. However, the possibility remains that some of the fricatives represented by UPSID<sub>317</sub> as ‘voiced’ or ‘voiceless’ should be reanalysed (e.g., as breathy voiced or voiceless depressor fricatives: cf. [Downing & Gick 2001](#) and chapter 2). Consequently, any generalisations drawn from UPSID<sub>317</sub> about the behaviour of voiced or voiceless fricatives should be treated with caution, if not the same amount of caution that should be observed when dealing with generalisations about VOT/voicing in stops.

I strongly suspect that the lack of phonetic discrimination with respect to obstruent voicing in UPSID<sub>317</sub> is related to the ambivalence of terms such as *voicing* and *aspiration* in phonetic descriptions. Regardless of their ultimate source, the phonetic approximations in UPSID<sub>317</sub> and elsewhere are sometimes insufficient to test theories grounded in phonetic substance. For instance the effort-based theory proposed by [Ohala \(1983\)](#) and others makes predictions about the phonetic voicing of preferred obstruents and can therefore not be tested against a database that occasionally glosses over the difference between actively voiced

<sup>6</sup>Interestingly, a more recent edition of UPSID containing 451 languages (UPSID<sub>451</sub>) reclassifies the fortis stops of both languages as plain voiceless, whilst maintaining a voiced-voiceless aspirated system for Norwegian.

and passively voiced (lenis) obstruents, or the distinction between plain voiceless and voiceless aspirated obstruents. The difference between the passively voiced word-initial [b̥, d̥, ɡ̊] of German and Bulgarian [b, d, ɡ(, bʰ, dʰ, ɡʰ)] is critical to the effort-based theory since it predicts that the former but not the latter are disfavoured by languages. The distinction between passively voiced and *actively devoiced* obstruents is equally critical under an effort-based theory, which predicts that, as far as voicing is concerned, the passively voiced lenis stops of aspirating languages are preferred over the actively devoiced stops of voicing languages. Similarly, the theory espoused by [Balise & Diehl \(1994\)](#) relies crucially on the presence of phonetic voicing, not on other (clusters of) phonetic features or structural properties. However, since distinctions between active and passive (de)voicing are not represented (consistently) in UPSID<sub>317</sub>, it does not allow theories based on phonetic voicing to be tested. Finally note that, conversely, the significance of the high frequency of ‘plain voiceless obstruents’ in UPSID<sub>317</sub> is severely reduced by the lack of phonetic specificity of the label.<sup>7</sup>

### 3.5 Context asymmetries: right and left-adjacent sounds

The behaviour of laryngeal contrast in consonant clusters, and in particular in obstruent + sonorant clusters, provides the critical data for comparing syllabic theories and cue-based accounts of laryngeal neutralisation (strictly word-based theories have nothing to say about word-initial and medial clusters). The former predict that phonetically similar sequences exhibit differences in neutralisation if they are syllabified differently. For instance, for  $C_1C_2$  sequences where  $C_1$  is an obstruent and  $C_2$  a sonorant, syllable-driven theories predict that laryngeal neutralisation is much more likely to affect the obstruent if  $C_1$  and  $C_2$  are heterosyllabic than if they group together in the same syllable. Linear models

<sup>7</sup>[Westbury & Keating \(1986\)](#) list a number of additional gaps and indeterminacies in the available data that make it hard to test effort-based models of laryngeal neutralisation. The assignment of a very specific physical (or perceptual) interpretation to phonetic symbols in the absence of instrumental data is problematic for substance-based theories more generally. For example, the common use of the symbols [ʌ] and [ɔ] to refer to the vowels in (American) English <done> and <dawn> might be interpreted by assigning these vowels a similar acoustic or perceptual vowel height ( $F_1$ ) and backness ( $F_2$  or  $F_2-F_1$ ), because the symbols are paired as open-mid and back in the IPA (cardinal) vowel chart. [Schwartz et al. \(1997a,b\)](#) essentially generalise this procedure to the set of vowel symbols used in UPSID<sub>317</sub>. Yet acoustic data indicates that the practical use of IPA symbols provides only very rough approximations of the orientation of vowels in  $F_1$ - $F_2$  space. Measurements by [Peterson & Barney \(1952\)](#) show that American English [ʌ] is both considerably lower and fronter (has higher values for both  $F_1$  and  $F_2$ ) than [ɔ]: in terms of height it groups with [æ] and [a] rather than with [ɔ]. Therefore, evidence from UPSID<sub>317</sub> or similar sources for dispersion or symmetry ([Boersma 1998](#): chapter 16) in vowel inventories, can only be interpreted in relatively global or abstract (phonological) fashion.

on the other hand predict that, all else (e.g., position in the word) being equal, there should be no such difference. This section does not offer the sort of data survey that would allow anything near a definitive conclusion in this matter. It does argue that the available evidence favours linear and cue-based rather than syllabic accounts, specifically because laryngeal distinctions may be maintained in heterosyllabic obstruent-sonorant sequences whilst they may be suspended in tautosyllabic clusters.

There can be little doubt that laryngeal neutralisation is common in obstruent clusters. Examining a sample of 104 languages, Greenberg (1978) found a strong preference for clusters that are homogeneous in terms of “voicing”, which might be treated as an indication that there is a typological tendency not to allow marking of laryngeal contrast on individual members of obstruent clusters. There is an additional preference for such clusters to be phonetically voiceless.<sup>8</sup> The Germanic languages fit this typological pattern by suspending the marking of laryngeal contrast on most obstruent clusters in monomorphemic forms. For instance, with the exception of Yiddish (Birnbbaum, 1979) and to a lesser extent English, Germanic does not seem to allow laryngeal contrast between the individual members of tautomorphemic clusters (oppositions between e.g., /bt/ and /pt/ or /pt/ and /pd/) and bars all contrast between word-initial sibilant + obstruent clusters. The exceptions to the ban on cluster-internal contrast in English involve words with the Latinate prefix <ab> (<absent, absurd, abstract>) and a few other forms (e.g., <magpie>) which seem likely to have been reanalysed as monomorphemic, as well as clusters that might be argued to stem from synchronic vowel elision (notably <medicine>). Zonneveld (1983) shows how in Dutch, laryngeal contrast between clusters is extremely limited word medially and finally, where the overwhelming majority of obstruent sequences is phonetically voiceless. Exceptions such as /ma:ɣd/ ‘virgin’ vs. /maxt/ *power, might*, and /vɔxt/ ‘moisture’ vs. /vo:ɣd/, *guardian, custodian*, are far and few between. This generalisation extends to German (Brockhaus, 1995; Wiese, 1996), Norwegian (Kristoffersen, 2000), and to a lesser extent English.

In addition, there is at least one case in which laryngeal neutralisation in obstruent sequences conforms to the predictions of a strict ‘implicational’ interpretation of the hierarchy in (5): it appears that languages with word-final, and therefore utterance-final (prepausal), neutralisation of singleton obstruents never retain laryngeal contrast in the prefinal positions of word-final obstruent sequences.<sup>9</sup> As illustrated in (6), the lexical distinctions between the final obstruents of the Dutch stems in the leftmost columns of (4) are neutralised not only utterance finally but also before participial -/d/: /ɣə/ + /krab/ + /d/ (*has*)

<sup>8</sup>Unfortunately the data presented by Greenberg (1978) do not allow for any robust generalisations about the marking of laryngeal contrast in obstruent clusters vis-à-vis other contexts.

<sup>9</sup>On the treatment of *word-final* contexts in a cue-based theory, see section 3.6 below.



*scratched*, and /ɣə/ + /ɣrɑp/ + /d/ (*has*) *joked*, surface with phonetically identical [pt] clusters rather than as [χəkrɑbt], [χəχrɑpt]. Similar generalisations apply to German and Frisian as well as to Lithuanian (Steriade, 1997). Whilst the presence of the pattern in (4) thus appears to imply the pattern in (6) below, the Norwegian ‘regressive devoicing’ data discussed in chapter 4 strongly suggests that the reverse is not true: underlying /tryg/ *secure* surfaces as [try kt] when suffixed with /t/, but retains its lenis features in unsuffixed forms: [tryg].<sup>10</sup>

(6) Neutralisation in Dutch participles

UR	Participle	Gloss
/xrap/	[χəχrɑpt]	joke
/krɑb/	[χəkrɑpt]	scratch (an itch)
/ɣrɑz/	[χəχrɑ:st]	graze (of animals)
/krɑs/	[χəkrɑst]	scrape, scratch (a smooth surface)

Note that whereas the implicational relationship between the patterns in (4) and (6) follows automatically from (a specific interpretation of) the hierarchy in (5), any (word-based) formalist model trying to account for the pattern in (6) using an assimilation rule would have to stipulate that this rule is always present in languages with final neutralisation. However, the behaviour of obstruent sequences does not critically distinguish cue-based theories from *syllabic* accounts, since both approaches predict that prefinal positions in such clusters are prime neutralisation sites. Laryngeal contrast in obstruents is relatively difficult to perceive before another obstruent because it removes the carrier for a some important cues, notably formant perturbations and positive VOT distinctions. Depending on the amount of coarticulation between obstruents in a given language, a following obstruent may also partially or wholly obscure the release bursts of plosives (Henderson & Repp, 1982) and the constriction duration of both plosives and obstruents. In languages where this is the case, laryngeal contrast would be even harder to perceive than in languages that clearly delimit obstruents in sequence. Syllabic accounts predict that prefinal positions in obstruent sequences are neutralisation targets because they are in a coda position or otherwise structurally different from the sort of onsets that typically preserve laryngeal contrast. This applies to word-initial obstruent clusters as well as to medial and final sequences since they are often treated as heterosyllabic (Kaye, 1992; Harris, 1994) or (mainly because they violate sonority sequencing) otherwise different from normal complex onsets (e.g., obstruent + sonorant sequences: cf. Blevins 1995). Thus, the fact that speakers of Dutch pronounce the acronym <ABVA> as [ɑpfa] (e.g., Booij 1995) does not necessarily mean that laryngeal neutralisation in this language is constrained by syllabic structure.

<sup>10</sup>Both the Dutch and the Norwegian patterns are assumed to be neutralising here.

Moreover, cue-based and syllabic theories make similar predictions for laryngeal marking in the final positions of obstruent sequences and the behaviour as clusters of a whole. Virtually all syllabification algorithms parse the final elements of obstruent clusters and the corresponding singleton obstruents (e.g., the alveolar plosives in English <top> and <stop>) identically and so it follows that they should behave identically in terms of laryngeal marking. Similarly, if sequences of obstruents as a whole are allowed to inherit the properties of their final members (cf. the account of the Dutch past tense rule presented by Lombardi 1994), e.g., English /stert/ should be as likely to contrast with /zdert/ as /dert/ with /tert/ (i.e., <Tate>, proper name). In cue-based models, cluster-final obstruents would have the benefit of the relatively salient offset cues, whilst the marking of contrast at the level of clusters as a whole could in principle involve the same set of cues that is involved in the realisation of [tense] in singleton obstruents. Consequently, both syllabic and cue-based models would seem to need extra apparatus to distinguish clusters as a whole and final elements on the one hand from prefinal positions.<sup>11</sup>

All this means that the key set of cases that allows syllabic and cue-based theories to be tested against each other, consists of obstruent + sonorant clusters, which can have different parses, depending on the consonants and language involved. Syllabic accounts predict that if the obstruent and sonorant in such sequences group together as a complex onset (or are assigned some equivalent structure), the obstruent is less likely to be targeted by laryngeal neutralisation than if the sequence is split by a syllable boundary, i.e., if the obstruent is in a coda position or part of some equivalent structure. For example, syllabic accounts predict that neutralisation is more likely to affect the contrast between the medial stops in English /mægnəm/, /hæknɪ/, /kædmɪəm/, /pʌtnɪ/ (area of London), /bræd lɪ/ (proper name), [mɒtɪ], than the medial stops in e.g., /koubrə/, /kouprə/, /pɛtrɪ/, /skwɒdrɪ/, /proʊgræm/, /ækɪd/. Since plosive + /r/ clusters are allowed word initially in English they are often treated as (at least potentially) tautosyllabic wherever (else) they occur, whilst plosive + nasal clusters do not occur word initially and are therefore generally treated as heterosyllabic where they occur word medially. By contrast, cue-based accounts would predict that to the extent that the cues to [±tense] are equally salient in both contexts, medial plosive + /r/ and plosive + nasal sequences are equally likely targets for laryngeal neutralisation.

Steriade (1997) makes two sets of observations that contradict the predictions of syllable-based approaches. First, in languages with word-final laryngeal neutralisation arguably heterosyllabic obstruent + sonorant clusters may preserve laryngeal contrast. Lithuanian is one of her main examples. This lan-

<sup>11</sup>On the problems posed by the Dutch regular past tense paradigm for syllable-driven models, and, to a lesser extent, cue-based approaches, see section 3.7 below

guage suspends all laryngeal contrast word finally, but preserves it word medially before obstruents, witness forms such as [silpnas], *weak* vs. [skobnis], *table*. Word-initial labial plosive + nasal sequences do not occur in Lithuanian, which is an argument for treating them as heterosyllabic. Likewise, some German speakers pronounce the plosives in words like <Adler>, *eagle*, and <ordnen>, *to put in order*, as [a:dlər] and [ɔrdnən], thus maintaining a contrast with the fortis plosives in words such as <partner>, [partnər], *partner* (see Brockhaus 1995 for extensive discussion of the small set of words capable of maintaining this contrast). Note that the speakers who use these forms do neutralise the opposition between fortis and lenis contrasts word finally. Alveolar stop + nasal/lateral clusters do not occur in German other than in medial position, and so again there seems little reason to treat them as tautosyllabic.

Second, laryngeal neutralisation may occur in obstruent + sonorant clusters that are best analysed as tautosyllabic. The behaviour of English fricative + sonorant clusters exemplifies this phenomenon. Whereas /f, v/, /s, z/ and marginally /θ, ð/ and /ʃ, ʒ/, contrast word initially before vowels, a single series of voiceless fricatives occurs before /r, l, w, j/: cf. (7a). Plosives on the other hand, retain the fortis-lenis distinction in this environment (7b). Word medially before sonorants, the contrast between fortis and lenis fricatives is at best limited. A near-minimal pair like <chevron>-<saffron>, [ʃɛv.ɪən], [sæf.ɪən] establishes a contrast for the labiodental place of articulation but a similar pair for the opposition /s/-/z/ is hard to find with alveolar sibilants showing a marked tendency to voicing before medial sonorants (7c). On the other hand, only voiceless dentals and postalveolars occur word medially before a sonorant. As the obstruent clusters illustrated here can occur word initially, there seems little reason to treat the laryngeal neutralisation of English fricatives before sonorants as syllable-final.<sup>12</sup>

There is hardly any more reason to treat the English data as exceptional or a phenomenon restricted to fricatives. Steriade (1997) lists several languages which maintain laryngeal oppositions word initially between prevocalic stops but not between stops followed by a sonorant consonant, including the Mon-Khmer languages Pacoh and Sre. Thai maintains a three term opposition between prevoiced, short lag and long lag stops prevocalically but only contrasts short lag and long lag stops before /r, l, w/ (Noss, 1964).

It is perhaps perhaps ironic that German, the textbook example of a language with ‘final devoicing’ should provide a further instance of this pattern. A number of High German dialects maintains a contrast between tense and lax plosives prevocalically, but suspends it before (tautosyllabic) liquids: interestingly,

<sup>12</sup>In American (as well as several other) varieties of English alveolar plosives followed by a high back rounded vowel have no intervening palatal glide. Coronal fricative + high back vowel sequences lack a palatal glide even for many British speakers, but they are included for completeness. On the realisation of word-medial fricatives before obstruents, see chapter 7.

the single series of plosives that occur in this context are generally described as ‘voiceless lenis’ (i.e. voiceless unaspirated). The phenomenon is known as *binnenhochdeutsche Konsonantenschwächung* in German dialectology and seems fairly well-documented. The Darmstadt (Rhenish Franconian) dialect described by Keller (1961) is a good example. In this dialect, Middle High German /p/ and /b/ have merged before sonorants to yield pronunciations such as [b̥]latz, *square* (Standard German [p<sup>h</sup>lats]), and A[b̥]ril, *April* (Keller 1961:171).<sup>13</sup>

(7) Laryngeal contrast in (British) English obstruent + sonorant clusters (Jones, 1977; Wells, 2000)

a. Suspension of contrast in word-initial fricative + sonorant fricatives

Orthography	Pronunciation	Orthography	Pronunciation
thrive	[θɹaɪv]	thwart	[θwɔ:t]
fright	[fɹaɪt]	swan	[swɒn]
flight	[flaɪt]	Thule	[θju:lɪ]
slight	[slaɪt]	suit	[sju:t]
refuse (v.)	[ɹɛfju:z]		

b. Contrast is retained in word-initial plosive + sonorant clusters

Orthography	Pronunciation	Orthography	Pronunciation
plight	[plɑɪt]	blight	[bɹɑɪt]
try	[tɹaɪ]	dry	[dɹaɪ]
Punic	[p <sup>h</sup> ju:nɪk]	bugle	[bju:glɪ]
twelve	[twɛlv]	dwelt	[dweɪl]
cute	[k <sup>h</sup> ju:t]	gules (red)	[gju:lz]
tune	[t <sup>h</sup> ju:n]	dune	[dju:n]

c. Realisation of medial alveolar sibilants followed by a sonorant

Orthography	Pronunciation	Orthography	Pronunciation
osmosis	[ɒzməʊsɪs]	Bosnia	[bɒznɪə]
Oslo	[ɒsləʊ][ɒzləʊ]	gosling	[gɒzɪŋ]
Israel	[ɪzrɛɪ]	Bosworth	[bɒzwɜθ]

In sum, the first set of observations establishes that the occurrence of word-final laryngeal neutralisation does not entail neutralisation in environments that can be regarded as syllable-final or non-syllable-initial, whilst the second set of observations shows that laryngeal neutralisation may occur before sonorants irrespective of syllabic structure. Moreover, the Lithuanian and English fricative data establish the double dissociation of word-final and pre-sonorant laryngeal neutralisation: in the former language word-final neutralisation exists in the absence of neutralisation before sonorants, whilst the latter exhibits pre-sonorant neutralisation but maintains the contrast between fortis and lenis fricatives word

<sup>13</sup>Thanks to Wiebke Brockhaus for pointing me to the literature on this topic.

finally.<sup>14</sup> Since the two phenomena can occur independently of each other, their cooccurrence is predicted as one of 4 possible patterns and thereby ceases to be a convincing argument for a syllabic analysis of neutralisation. In other words, the observations that Dutch suspends laryngeal contrast both word finally and word medially before nasals (the medial stops in orthographic <partner>, *partner* and <ordner> appear to be pronounced identically) can be plausibly construed as the (chance) cooccurrence of two independent processes rather than reflections of the same process.

### 3.6 The word-initial vs. word-final asymmetry

Word-initial and word-final contexts have such different propensities for laryngeal neutralisation that the observation is rarely made in these terms, [Westbury & Keating \(1986\)](#) being a notable exception. Usually it is simply implied by statements about (syllable-)final neutralisation, and illustrated with paradigms such as (4) Yet it remains a striking observation that many, if not most, of the languages that neutralise laryngeal contrast word finally, preserve it word initially as well as medially between vowels. Initial neutralisation occurs trivially in languages that lack laryngeal contrast altogether, but it appears to be rare in languages that maintain a contrast elsewhere. [Steriade \(1997\)](#) even seems to claim that word-initial laryngeal neutralisation invariably implies final neutralisation. [Westbury & Keating \(1986\)](#) list Cuna, Efik, Ewondo, and Tamil as languages that lack laryngeal contrast word initially (and finally) but retain it medially.<sup>15</sup> [Steriade \(1997\)](#) adds Lac Simon, an Algonquian language, and Totontepec, which belong to the Mixtecan group. I can add the fricatives of Frisian, which support laryngeal contrast medially between sonorants and vowels, but not word initially or finally ([Tiersma, 1985](#)). It would seem therefore, that laryngeal contrast is only marginally less stable word initially than medially (between sonorants).

By contrast, the list of languages that maintain some form of contrast initially (and medially), but not finally remains considerably longer, and although a typological survey of laryngeal neutralisation phenomena is long overdue (cf. [Brockhaus 1995](#)), this is an indication that laryngeal contrast is more stable word

<sup>14</sup>This dissociation contradicts the strict implicational interpretation of the perceptibility hierarchy in (5) which predicts that laryngeal contrast is always more stable before sonorants than utterance finally.

<sup>15</sup>Tamil is somewhat problematic as an example of this phenomenon. Whether or not it has contrasts initially depends on what status is assigned to the voiced stops it has borrowed from Sanskrit. Similarly the status of the medial contrast referred to by [Westbury & Keating \(1986\)](#) rests on the synchronic analysis of medial stop gemination and voicing (by Caldwell's Law). Thus, another author ([Steever 1990:239](#)) is able to state that in Tamil "[v]oiced stops contrast with voiceless stops only in initial position. . ."

initially than finally. This set of languages includes Dutch, Frisian (plosives), and German within Germanic, many of the Slavonic languages, including Bulgarian, Polish, Russian, Slovak (Rubach, 1993), Lithuanian, many of the Turkic languages, including Turkish and Turkmen (Johanson & Csató, 1998; Clark, 1998), Thai, Vietnamese, Zoque, and Basque (Westbury & Keating, 1986).

As pointed out above, syllable-driven models of neutralisation account for the initial-final asymmetry by virtue of the fact that syllabification algorithms typically parse word initial (pre-sonorant) obstruents as onsets and word-final obstruents as codas (or in some other way different from prevocalic obstruents). Other formalist models treat the phenomenon in terms of morphological or suprasyllabic prosodic structure (Trommelen & Zonneveld, 1979; Rubach, 1993; Brockhaus, 1995). I have referred to both of the latter as *word-based* since, if prosodic domains are involved, they are usually of the kind that interacts with word-level (analytical, non-cohering) morphology. Thus, the neutralisation of the lexical contrast between the final stops in Dutch /ro:d/ *red*, and /vərye:t/ *forget*, before the suffix /axtəχ/ *-ish, -like* (cf. [ro:<sup>w</sup>taxtəχ] and [vərye<sup>l</sup>:taxtəχ]) can be attributed to morphological structure which indicates that a word (level) boundary separates the suffix from its hosts, or to the fact that host and suffix are parsed as independent prosodic domains. Instances of neutralisation in word internal contexts are clearly outside the scope of either of the two word-based approaches, unless morphological or prosodic structure is used in a wildly diacritic fashion. This is especially true of instances of neutralisation before sonorant consonants, as they cannot be analysed as (static) assimilation of [tense]

For cue-based models on the other hand, word-based asymmetries might seem to be a problem. It is impossible to attribute strong tendencies for laryngeal contrasts to be suspended word finally but to be preserved initially to the influence of adjacent sounds. Unlike *utterance-final* (prepausal) or *utterance-initial* (postpausal), *word-initial* and *word-final* are not phonetic contexts in the way that  $\_ [l]$  is a phonetic context or  $\_ [+son]$  a range of phonetic contexts. Word-final obstruents can precede any sound that is found word initially, whilst their word-initial counterparts can be preceded by anything found word finally, yet word final laryngeal neutralisation tends to behave uniformly across right-hand phonetic contexts, and the preservation of word-initial laryngeal contrast rarely if ever depends on the nature of the preceding sound.<sup>16</sup>

Steriade (1997) concludes that the initial-final asymmetry can not be driven by phonetic context alone. She therefore adopts a *Paradigm Uniformity* constraint to account for the uniform behaviour of word-final neutralisation across contexts, and a positional faithfulness constraint to shield initial contrasts from

<sup>16</sup>Especially under the assumption that Dutch post-obstruent (fricative) devoicing and a similar process noted in Basque by Hualde (1991) are phonetic devoicing rather than phonological neutralisation processes.

neutralisation. Paradigm Uniformity (Kenstowicz, 1995) is a family of output-output faithfulness constraints demanding that all the phonological or phonetic forms of a morpheme across a certain paradigm should be identical to the form of an ‘attractor’ morpheme that occurs at a single designated point in the paradigm. Constraints of this type are used in optimality-theoretic models of paradigmatic levelling and other ‘morpheme invariance’ effects. The formal property that sets Paradigm Uniformity apart from most other forms of output-output faithfulness is that it evaluates *n-tuplets* instead of *pairs* of forms, where *n* equals the number of positions in the paradigm.

Paradigm Uniformity is brought to bear on the analysis of final neutralisation by extending the notion of *paradigm* to the full set of possible phonetic contexts in which a morpheme can occur, and by employing citation forms of words as attractors. Obstruents at the end of citation forms, i.e., utterance-final obstruents, have a consistent right-hand phonetic context (silence), which might be argued to have relatively poor cueing potential. Although the number of available cues is fairly large, consisting of both onset cues (preceding vowel duration, V-to-C formant transitions, timing and nature of voicing offset), internal cues (closure duration), and some offset cues (amplitude and duration of release bursts), Steriade claims that the absence of several highly salient offset cues (timing and nature of voicing onset C-to-V formant transitions) nevertheless ranks V\_|| contexts (universally) below all \_[+son] environments. Consequently, citation form-final obstruents are comparatively vulnerable to final neutralisation, and their attractor status generalises the neutralisation resulting from this vulnerability to all word-final obstruents, some of which occur in contexts with better cueing potential (i.e., before sonorants).<sup>17</sup>

The *Paradigm Uniformity (right edge)* constraint employed by Steriade (1997) is relativised to apply to word-final contexts only so as to exempt word-initial laryngeal distinctions from the effects of cueing limitations on citation form initial (postpausal) obstruents. Nevertheless, initial contrasts need an extra boost to attain uniform and approximately equal stability as word internal intersonorant distinctions: the two environments in question both support the highly salient offset cues as well as internal cues, but the latter adds a stable set of onset cues, which is more or less unavailable to the latter, depending on the preceding context. If neutralisation is derived from cueing potential alone it would follow that the marking of initial laryngeal distinctions is left-hand context sensitive, and less stable than medially between sonorants to the degree that

<sup>17</sup>A detailed review of the mechanics of the OT model constructed by Steriade (1997) is of no concern here, but note that it is inessential to designate the citation form as paradigmatic attractor: (high-ranked) *Paradigm Uniformity (right edge)* filters out every nonuniform paradigm regardless of which context is granted attractor status. This is fortunate since there is no independent evidence to motivate a psychologically special status for citation forms as opposed to, say, other relatively hyperarticulated forms of a word.

they are less perceptible. This is not borne out by the available data, and Steriade (1997) obtains uniformity and stability in initial environments by means of a positional faithfulness constraint *Preserve [voice] in #\_* which demands that laryngeal distinctions be preserved word initially. *Preserve [voice] in #\_* is highly similar to positional faithfulness constraints on the marking of laryngeal contrast employed by Grijzenhout & Krämer (1998) and Lombardi (1999) (see chapter 8)

The adoption of this constraint highlights both the problems that Steriade's model as it stands suffers from and, to my mind at least, its true potential. The principal problem is that the introduction of *Preserve [voice] in #\_* means that the uniformity in neutralisation in word-initial and word-final contexts is accounted for in terms of two separate, formally and functionally unrelated, devices. The good news is that in principle it is possible to unify the analysis of word edge effects on neutralisation as a cue-based phenomenon. A crucial assumption underpinning Steriade's model is that the perceptibility of a contrast is solely a function of the sounds adjacent to the carrier, but recall from the discussion of neutralisation in chapter 1 that prosody is also likely to be an important influence on relative perceptibility. Recall too, that a wealth of evidence, in particular from work on articulatory strengthening, indicates that prosody has markedly asymmetric effects at the word level, 'strengthening' word-initial sounds (even in the absence of lexical stress) and weakening (unstressed) segments elsewhere, including word finally. Crucially, although it may affect word sandhi phenomena, prosody itself does not vary with the left-hand and right-hand phonetic context of a word: word-initial segments are strengthened and word-final ones weakened whether they are preceded or followed by an obstruent, sonorant or physical pause.

Now if prosodic strengthening and weakening indeed affect the perceptibility of contrast, this would entail a decrease in the perceptibility of word-final laryngeal distinctions *across phonetic contexts provided by adjacent sounds*. In other words, after the effects of prosody are factored in, the [tense] cues of a word-final (or unstressed medial) obstruent preceding a sonorant would be less perceptible than the [tense] cues of the same obstruent preceding the same sonorant in word-initial (or stressed) position. Under a cue-based theory of neutralisation it follows that laryngeal distinctions are less viable word finally than word initially, again across contexts provided by adjacent sounds. Another way of phrasing this is to say that prosody creates a partially uniform phonetic context at word edges and can therefore be expected to act as a leveller of perceptibility differences and hence neutralisation asymmetries arising from other factors.<sup>18</sup>

<sup>18</sup>J. Beckman (1997) suggests that positional faithfulness constraints are ultimately grounded in phonetic factors such as perceptibility. A possible additional source of uniformity in neutralisation at word edges is that the (temporal) organisation of laryngeal cues is generally optimised to work



Whereas the formalist prosodic approaches referred to at the outset of this section claim that neutralisation is directly driven by prosodic phrasing, the cue-based alternative proposed here holds that prosody is simply one of the factors influencing relative perceptibility and affects neutralisation via the same single mechanism that is responsible for neutralisation asymmetries triggered by adjacent sounds and internal phonetic features (e.g., the plosive fricative-asymmetry). In ‘diachronic’ functional models this single mechanism is misperception by learners, in ‘synchronic’ models, a ban on contrasts with a perceptual salience below a specified level. There is one sense in which the cue-driven and some formalist theories converge: to the degree that prosody is indeed a system for signalling grammatical boundaries, the former reconstructs the idea espoused by [Kaye \(1989\)](#) and [Harris \(1994\)](#) that phonological rules exist to assist in word recognition and parsing. But since neither formalist word-based models nor the account by [Steriade \(1997\)](#) seem able to unify the analysis of neutralisation patterns triggered by adjacent sounds, internal features and prosody, a fully cue-based approach represents a considerable improvement, at least in principle.

A considerable amount of new phonetic data is needed, first to establish whether a word-final vs. word-initial neutralisation asymmetry can indeed be predicted on the basis of perceptibility, and second to test whether prosodic effects on perceptibility indeed map into neutralisation asymmetries. At least impressionistically, flapping of alveolar stops in (American) English greatly reduces the amount of perceptual contrast between /t/ and /d/ (so much so that it has been treated as a neutralising process) and this might be counted as preliminary evidence that processes directly driven by (or originating in) prosodic weakening are able to reduce the perceptibility of phonological contrasts. Note, incidentally, that the dialects in question lenite word-final alveolar stops across right-hand phonetic contexts, even where this does not lead to flapping ([Harris, 1994](#)). Furthermore, it is not inconceivable that (if they are indeed not artefacts of ‘spelling pronunciation’) the incomplete neutralisation phenomena discussed in [3.2.2](#) above reflect final weakening, perhaps reinforced by some other mechanism. However, the relative perceptibility of laryngeal contrast in obstruents needs to be examined more systematically, and preferably on the basis of mate-

---

within, rather than across word boundaries. As pointed out in [chapter 2](#), the observation that English fortis plosives are not aspirated (but rather preglottalised) word finally even when they occur before a sonorant, can be interpreted in these terms. From a functional perspective, this type of paradigmatic uniformity in cue organisation might be advantageous in offering a robust cue to laryngeal distinctions across phonetic contexts and as an aid in word segmentation. Note that there is an important and testable difference between this hypothesis and the use of *Paradigm Uniformity* by [Steriade \(1997\)](#). The former predicts paradigm uniformity at the phonetic level in languages that *maintain* laryngeal contrast word finally, and (as a result) at the phonological (neutralisation) level in languages that suspend it. Steriade’s model on the other hand proposes that laryngeal *neutralisation* instead of cue organisation is generalised across the paradigm, and thus predicts that uniformity effects only occur at the phonological level.

rial that does not exhibit highly conspicuous lenition processes.

As far as the interaction between prosody and laryngeal neutralisation is concerned, there does not seem to be a large amount of data either, and much of it relates to laryngeal segments rather than laryngeally marked obstruents. The fact that in English, /h/ only occurs prevocally in word-initial and stressed syllables is well-known (cf. [Harris 1994](#)). I am not aware of any examples in which the fortis-lenis distinction is neutralised outside these contexts other than the Danish pattern where originally lenis stops spirantise to leave a single series of stops behind (which is very common, and may in fact explain the absence of fortis-lenis mergers). Prosody-driven neutralisation in nonfinal contexts certainly does not figure in the survey article of [Mascaró & Wetzels \(2001\)](#). However, on the basis of a survey of languages spoken across Europe, north and central Asia, [Butskhrikidze \(1998\)](#) claims that the occurrence of final neutralisation is highly constrained by lexical stress. Her *Universal 2.1* reads “[i]f the fixed accent is on the final syllable, devoicing of the final obstruents does not occur”. Because there are languages that, by most accounts, have final stress and final neutralisation, e.g., several Turkic languages, this would seem to be a ‘soft’ universal expressing a (statistically reliable) tendency, or a statement about the preconditions for the development of final neutralisation. Unfortunately, as the body of her work has not been translated from the Georgian original, I am currently not in a position to evaluate the data underpinning Butskhrikidze’s claims.

### 3.7 Summary and remaining issues

This chapter presented a number of generalisations about laryngeal neutralisation that form a necessary backdrop for some of the experimental work presented in the next two chapters and especially for chapter 8. In addition, the general thrust of formalist approaches to some of these generalisations was compared to a fully cue-based account. This cue-based account attempts to combine the insights of [Steriade \(1997\)](#) with regard to right and left-adjacent context effects with, first, the suggestion by [Balise & Diehl \(1994\)](#) that laryngeal neutralisation of fricatives may be grounded in perception, and second, the hypothesis that positional faithfulness ultimately derives from the effects of prosodic strengthening and weakening on the relative perceptual salience of phonological distinctions (cf. [J. Beckman 1996, 1997](#)). Thus, this theory is an instantiation of the strongest version of the functionalist hypothesis described in chapter 1 which states that all phonological neutralisation phenomena derive from the low perceptibility of distinctions.

Section 3.2 discussed data pertaining to the phonetic manifestation of neutralisation. Two kinds of phonetic data emerged here: data suggesting that laryngeal neutralisation results in phonetic underspecification, and evidence that

final, ‘dynamic’ neutralisation is sometimes incomplete, leaving residual cues to the underlying values of [tense]. The first observation is troublesome for a fortition analysis of neutralisation in fortis-lenis systems because fortition predicts that the output of neutralisation is an obstruent that is phonetically and phonologically indistinguishable from its unequivocally fortis counterparts.

The next two sections examined the effects of obstruent features on the likelihood that those obstruents are targeted by laryngeal neutralisation. Section 3.3 concluded that on the basis of the available evidence, there is little reason to assume that laryngeal neutralisation behaves differently in voicing and aspirating languages. Section 3.4 discussed the well-known claim that laryngeal contrast is less stable in (sibilant) fricatives than in plosives. This claim is largely supported by the frequencies of laryngeal contrast (as distinct from phonetic voicing) in plosive and fricative inventories in typological databases such as UPSID<sub>317</sub> as well as by the context-sensitive neutralisation of fricatives in languages such as English and Frisian. In addition, section 3.4 reviewed the possibilities of explaining this asymmetry in terms of articulatory effort or perceptibility. However, whilst there is encouraging experimental support for the latter, both theories are hard to test against available typological databases, which tend to partially conflate phonetic voicing and participation in phonological contrasts such as [±tense]. The generalisations established in this section play an important role in the discussion of autosegmental models in chapter 8.

Section 3.5 engaged with the central generalisations underlying the theory of Steriade (1997). Steriade shows that syllable-driven accounts of final neutralisation fail in their core predictions about neutralisation in obstruent + sonorant clusters. First, they predict that word-final neutralisation should always coincide with neutralisation in heterosyllabic obstruent clusters. Second, they predict that neutralisation in word-medial obstruent + sonorant clusters follows syllabification patterns so that neutralisation occurs in heterosyllabic but not in tautosyllabic sequences. Both claims are demonstrably wrong. In fact, the dissociation between laryngeal neutralisation in word-final contexts and nonfinal obstruent + sonorant sequences is even stronger than Steriade (1997) suggests: both processes can occur independently from one another. A cue-based model is better suited to account for this data because it predicts that obstruent + sonorant sequences should behave in identical fashion irrespective of syllabification (and *ceteris paribus*).

Finally 3.6 assessed the possible explanations for the uncontroversial generalisation that word-final neutralisation is much more common than initial neutralisation, which seems to hardly occur at all (unless as part of neutralisation at the inventory level). Morphology or prosody-driven formalist accounts simply claim that the right edge of the word is a weak licenser of laryngeal contrast, whilst Steriade (1997) attempts to explain the initial-final asymmetry

by generalising neutralisation in the weakly-cued *utterance-final* context across all other phonetic contexts. However, since word-level prosody manifests itself as strengthening and weakening, it seems likely to have a levelling effect on context-induced perceptibility differences in word-final (and unstressed) as well as word-initial (and stressed) environments. If this is indeed the case, a cue-based theory derives the neutralisation asymmetry between word-initial and word-final contexts in exactly the same way as the effects of obstruent internal features and adjacent sounds, and this would represent an improvement over both formalist accounts and Steriade's model. Although a great deal of research is needed both to establish the precise predictions of the fully cue-based theory and to test them, I want to stress again that both are possible enterprises. Perceptibility measures can be established using known experimental designs whilst testing them against neutralisation asymmetries does not require any (statistical) methods that are beyond the means of phonologists.

Inevitably, there is a range of interesting and important issues that must remain unexplored here. To conclude this chapter I will briefly mention two topics that would deserve attention in a fuller survey of neutralisation phenomena. First it is often implicitly assumed that dynamic final neutralisation is fully symmetric with regard to manner and place of articulation distinctions. This is certainly the case in Dutch and German, where final neutralisation targets both fricatives and plosives across all places of articulation. However, it is not at all clear to what extent this represents the 'normal', most frequently occurring pattern in languages with established or developing neutralisation rules. Parker (1981), quoting Andersen (1972) and Stevens (1975), observes that in Belorussian final plosives started to neutralise before the corresponding fricatives whilst this pattern was reversed in German. Moreover, final neutralisation in Turkmen, as described by Clark (1998), generally targets stops and affricates but leaves dental nonsibilant fricatives untouched (there is no stable contrast for sibilant fricatives): cf. /du:ð/, [du:ð], *salt* vs. /nɑ:miθ/, [nɑ:miθ], *honour*, *shame*.

Second, in principle, theories about the phonological and phonetic nature of the neutralisation process are independent from theories about the contexts (in a broad sense) in which it is most likely to occur. This means that it is technically speaking possible to maintain a strictly formalist theory of neutralisation contexts whilst adopting a (phonetic) underspecification view of the neutralisation process. To some extent this is the sort of approach pursued by Ernestus (2000). Conversely, a cue-based model does not of itself rule out that laryngeal neutralisation is fortition. However, from a broader functionalist point of view, the claim by Balise & Diehl (1994) that voiceless fricatives are better carriers of fricative place contrasts implies that there is some pressure on phonetic grammars to employ active devoicing in the realisation of neutralised fricatives not to cue [tense], but to enhance the expression of place contrast. This implication

is contradicted by [tense]-neutralised but voiced word-initial [z] in German and in Dutch dialects (which appears in a strengthening context, so that phonetic reduction resulting in passive voicing can be ruled out). But the idea that voiceless and voiceless aspirated (fortis) stops are more salient than (lenis) voiced ones has a long tradition (cf. [Parker 1981](#)) and may go some way in explaining the aspiration of laryngeally neutralised stops of Western Aleut and other languages examined by [Cho & Ladefoged \(1997\)](#).