

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 4

Voicing assimilation

Perhaps the easiest way to introduce *voicing assimilation*, also commonly referred to as *voice assimilation*, is by a brief discussion of the Dutch examples in (8). Dutch has an extremely well-documented process of regressive voicing assimilation that applies at boundaries between words and between stems and non-cohering suffixes (Zwaardemaker & Eijkman, 1928; Cohen et al., 1972; Trommelen & Zonneveld, 1979; Booij, 1995). In many instances where a word-final (or stem-final) obstruent is followed by a lax plosive /b/ or /d/, this obstruent is realised as voiced. Note that word-final obstruents in Dutch are subject to neutralisation of [\pm tense]: as a result RVA applies in equal measure to underlying [+tense] and [-tense] obstruents. For example, the final /s/ of the first member of the compound /vis/ + /di:fjə/ is equally likely to be realised as voiced as the underlying /z/ in /rɛiz/ + /du:l/. Elsewhere, and particularly before fortis obstruents and utterance finally, Dutch final obstruents are often realised as voiceless (barring brief voicing tails of preceding sonorants): [ve^h:k], [zant], [vis], [rɛis].

(8) Regressive voicing assimilation to Dutch lax plosives

UR	Phonetic form	Gloss
/ve:k/ + /di:r/	[ve ^h :gdi ^r]	mollusc
/zand/ + /bank/	[zandbank]	sand bank
/vis/ + /di:fjə/	[vɪzdɪfjə]	common tern (<i>sterna hirundo</i>)
/rɛiz/ + /du:l/	[rɛizdul]	destination

Given the strict distinction between *voicing* and *tense-lax/fortis-lenis* that I have adopted, the term *voicing assimilation* implies that the process only affects this single cue to the tense-lax distinction. However, many descriptions and accounts of voicing assimilation rules assume implicitly or explicitly that they operate directly on [\pm tense], which implies that the whole cue complex associated with it is affected. This assumption is so pervasive in the literature that it

underpins even some of the experimental work attempting to demonstrate that RVA is a non-neutralising, and therefore phonetic process (e.g., [Charles-Luce 1993](#); [Burton & Robblee 1997](#)). It is at least implied by the transcriptions in (8), which follow common practice in representing voicing-assimilated /t/ as [d] rather than as e.g., [t].¹

In the discussion below I will nevertheless retain *voicing assimilation* as a cover term for the range of phenomena that have been so labelled in the literature. The purpose of this terminological compromise is not to confuse the reader but to convey the fact that the whole range of observed assimilation rules is so often viewed as a homogeneous set and attributed to a single underlying mechanism. Strictly speaking, *[tense]-assimilation* would do more justice to the nature of that mechanism, but it is an unfamiliar label and ultimately equally inappropriate as a cover term as *voicing assimilation*.

Voicing assimilation in this broad sense then, is an ubiquitous phenomenon. It occurs in morphological paradigms as well as between words in compounds and phrases, and there are few if any languages that maintain a [±tense] distinction that lack voicing assimilation rules altogether. In addition, many models of laryngeal phonology would also count lexical phonotactic restrictions on [tense] in obstruent sequences as instances of voicing assimilation.

This chapter is an attempt to dissect the set of ‘dynamic’ cases in which the voicing (and other relevant phonetic features) of obstruents vary with a varying context, i.e., the type of phenomenon illustrated in (8) above (lexical constraints on [±tense] are discussed in chapter 3). Section 4.1 immediately below identifies two common approaches to sandhi processes in the literature. The first of treats sandhi phenomena as operations at the (lexical) phonological level and is typically associated with recent generative models of voicing assimilation. The second is a phonetic, or more accurately *articulatory* theory of sandhi, which is embodied by models such as Articulatory Phonology ([Browman & Goldstein 1986](#) et seq.). Although phonological models and articulatory phonetic or ‘gestural’ frameworks are often used to provide alternative accounts of the same data, it would seem that neither of them is dispensable and that they really apply to complementary sets of data rather than overlapping ones. The rest of the section spells out the specific predictions with regard to the phonetic manifestations of voicing assimilation that can be inferred from a phonological approach, and those implied by an articulatory analysis. These predictions are different and form the basis for the experimental hypotheses tested in the next chapters.

The following three sections of this chapter then discuss three forms of voicing assimilation and investigates for each of these which of the two theories

¹For the sake of the argument I have assumed here that neutralised final obstruents in Dutch are properly represented as phonologically and phonetically [+tense]. Although this is a common assumption it may not be ultimately viable. See chapters 3 and 7 for more detailed discussion of final neutralisation and Dutch RVA respectively.

provides the most accurate account of its properties. First, voicing assimilation rules that are found at word-internal morpheme boundaries seem to behave as predicted by a phonological account. Second, in many instances, the progressive devoicing of obstruents at word boundaries appears to result from passive devoicing in the sense of section 2.1 above rather than an operation that ought to be expressed at the phonological level. This is hardly a novel or perhaps even controversial claim (cf. Harms 1973). However, section 4.4 proposes that the third form of voicing assimilation, i.e. regressive voicing assimilation across word boundaries, should be approached as an articulatory process driven by the production of voicing distinctions, or at least as diachronically grounded in such a process. The main argument for this proposal is that RVA across word boundaries is conditioned by the voicing categories of the triggering obstruents, at least in the case of lax plosives and affricates. Since this form of voicing assimilation is often put within the scope of generative models, this proposal is more contentious.

4.1 Modelling voicing assimilation

There are two common but distinct ways of thinking about sandhi phenomena, and hence about voicing assimilation. The first treats sandhi in terms of operations on (lexical) phonological features. In recent generative models this approach is typically implemented by means of autosegmental spreading and/or delinking rules, or (in Optimality Theory) *AGREE*-type constraints. The second approach treats sandhi rules as a product of the processes that govern the temporal co-ordination of articulatory gestures in speech production. This type of approach is sometimes simply referred to as *gestural*, perhaps because *Articulatory Phonology* (Browman & Goldstein 1986 et seq.) is one of the predominant formal frameworks in this area. However, since I do not intend to commit to a specific framework in this regard, and in keeping with the discussion in section 1.3.2 above, I will simply continue to refer to *articulation-driven*, *coarticulation(-based)*, or simply *phonetic* models.

The two approaches identified here are different in more than a conceptual sense: they generate different predictions with regard to the type of sandhi phenomena that can occur, and their phonetic manifestations. Consider first the phonological view of external sandhi. According to this view, any phonological feature may be targeted by sandhi rules, whether the feature in question has a consistent articulatory implementation or not. For example, the feature [\pm sonorant] might be defined phonetically in terms of airflow through the oral tract, spontaneous (de)voicing, or acoustic/perceptual intensity, but it has no consistent articulatory implementation. For instance, the ‘sonorancy’ of nasal consonants is due to the opening of the velopharyngeal port, whilst [l] and glides

are [+sonorant] because of the relatively large size of the passage at the point of maximum oral constriction. Nevertheless, models that adopt [\pm sonorant] as a phonological feature and do not impose any specific restrictions on its availability for sandhi processes are in principle capable of expressing [sonorant] assimilation rules. Such rules would turn sequences of e.g. /d/ + /l/ into [nl] (leftward spread of [+sonorant]) or /p/ + /n/ into [pd] (rightward spread of [-sonorant]).

By the definition provided in section 1.3.2, phonological processes are phonetically discrete. A second prediction of the phonological approach to sandhi processes is therefore that they are phonetically discrete. This means that there are no intermediate phonetic realisations between sequences that have not undergone a sandhi process and those that have, or only a finite number of such realisations. For example, phonological analyses of voicing assimilation generally generate two possible phonological surface forms and therefore two phonetic categories for underlying sequences such as /k/ + /b/ ([-voice][+voice] in underlying phonological representation). The first of these emerges when for some reason voicing assimilation fails to apply and can be symbolised as [kb] ([-voice][+voice] on the surface); the second represents the case of ‘total’ voicing assimilation, [gb] ([+voice][+voice]). Recent autosegmental models are sometimes capable of generating one or more surface forms that are intermediate between these extremes and as a result they provide some handle on cases of ‘subphonemic’, or partial assimilation (cf. Hayes 1992). However, the number of intermediate categories available is usually very small and always finite, and as a consequence phonological approaches predict that even partial assimilation creates discrete phonetic categories.

An articulation-driven approach to sandhi rules, on the other hand, predicts that only the articulatory gestures involved in the production of those sounds are available to sandhi processes. This means, for example, that the phonological feature [\pm sonorant] as such cannot play a role in assimilation rules. Instead, sandhi effects of the sounds classified by this feature are predicted to pattern along articulatory lines, which entails that a [n], but not a [l] should be able to turn a preceding [d] into a nasal because the former but not the latter is accompanied by the requisite velopharyngeal opening movement.

Second, since articulatory control (and physical articulation) operate on phonetic scales that can be regarded as continuous for all practical purposes, a phonetic approach predicts that sandhi phenomena are gradient at the phonetic level. This entails that assimilation rules do not produce discrete phonetic categories, but a continuous range of forms between unassimilated and assimilated sound sequences.

Proponents of both approaches would sometimes seem to imply otherwise, but it seems fairly clear that neither of them is capable of accounting for all documented sandhi phenomena. For example, Sardinian has a rule whereby

word-final /s/ and /r/ are neutralised to [s] if the following word starts with a voiceless stop, and to [l] if it starts with another consonant (Bolognesi 1998; Ladd & Scobbie forthcoming: examples in 9)

- (9) Neutralisation of word final /s/ and /r/ in Sardinian
- | | | |
|---------------------|-----------------|-----------------|
| /sos/ + /puddos/ | [sɔspudːɔzɔ] | ‘the chickens’ |
| /tres/ + /manos/ | [trɛlmanɔzɔ] | ‘three hands’ |
| /battor/ + /frades/ | [batːɔlfradɛzɛ] | ‘four brothers’ |

This rule cannot be modelled as a coarticulation process because the surface sequences cannot be described in terms of articulatory interference between the sounds in the underlying forms. For example, there is nothing in the articulation of [m] (which consists mainly of a full bilabial constriction accompanied by an opening of the velopharyngeal passage) that would force a preceding [s] to approximate the tongue configuration of a lateral approximant or to lose its wide glottal abduction.

On the other hand, a variety of external sandhi phenomena in a number of languages, and in particular place assimilation rules, have been shown to operate in a gradient fashion and are therefore impossible to describe in phonological terms (e.g. Barry 1992; Nolan 1992). Consequently, the most viable model of the phonology-phonetics interface seems to be one that can accommodate both phonological and coarticulatory sandhi rules (cf. Zsiga 1997).

The remainder of this section attempts to flesh out corresponding sets of predictions regarding the phonetic manifestations of voicing assimilation when analysed as a phonological process and when analysed as a coarticulatory. These predictions form an important basis for the rest of this study, which argues that voicing assimilation occurs both as a(n incompletely neutralising) phonological rule and as a coarticulation process.

4.1.1 Phonological analyses of voicing assimilation

Many recent generative accounts of voicing assimilation phenomena assume that the rules in question apply to the phonological feature that represents the lexical contrast between tense and lax obstruents. Here and below, I will focus on this type of account in discussing phonological approaches to voicing assimilation, but note that the predictions in (10) hold for any model which assumes that voicing assimilation entails a transfer of (lexical) phonological identity from the assimilation trigger to the assimilation target.

The mechanics of the phonological, or perhaps more accurately, *lexical feature*, conception of voicing assimilation adopted by the generative accounts referred to in the previous paragraph are illustrated in figure 4.1, using the autosegmental notation adopted by virtually all recent generative models of laryngeal phonology. The lexical contrast between tense and lax obstruents is represented

by a $[\pm\text{tense}]$ feature, or some formal equivalent such as $[\pm\text{voice}]$. This feature may or may not be dominated a *Laryngeal* (LAR) class node. Voicing assimilation is modelled by spreading the lexically contrastive feature or the dominating LAR node from the assimilation trigger, in this instance the obstruent represented by $Root_2$ to a target, here symbolised by $Root_1$. Under the most common version of this analysis, the original LAR node and $[\pm\text{tense}]$ specification of $Root_1$ are delinked and removed from the representation so that trigger and target share a single lexical laryngeal feature. This spreading-cum-delinking analysis is formally equivalent to more old-fashioned polarity switching ‘agreement’ rules of the $[\text{+tense}] \rightarrow [-\text{tense}]/_[-\text{tense}]$ variety.

Under a less common but, in the light of data reported below, more sophisticated approach the underlying LAR node and $[\pm\text{tense}]$ specification of $Root_1$ are preserved, and the result of the spreading operation is a ‘mixed’, doubly articulated or contour segment carrying two $[\text{tense}]$ features with distinct values (cf. Hayes 1992).

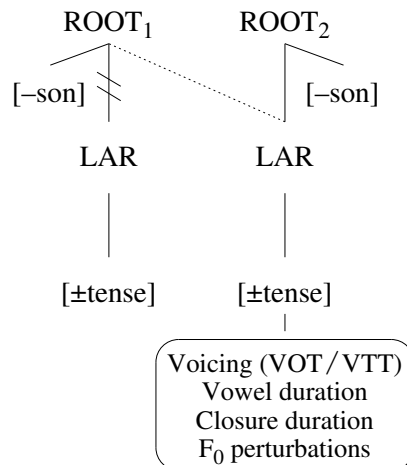


Figure 4.1: An autosegmental lexical feature analysis of voicing assimilation and its phonetic manifestations.

For two reasons, it is hard to pin down the predictions of recent lexical feature analyses regarding the phonetic manifestation of voicing assimilation. First, the lattice-like nature of autosegmental representations defines a greater number of possible structures for the same feature alphabet and number of segments (i.e., root nodes) than the feature bundle model on which SPE was built. This is already evident in the fact that there is a choice of applying spreading with

delinking, or without it. Furthermore, there are two ways to represent sequences with homogeneous specifications for [tense] such as [gb]: the first is a sequence of two root nodes each carrying its own LAR node with [-tense] feature; the second consists of two root nodes sharing a single LAR node and [-tense] specification. Like the difference between singly and doubly articulated segments employed by Hayes (1992), the difference between these two structures can in principle be employed to represent a distinction between underlyingly homogeneous [tense] sequences such as /g/ + /b/ and underlyingly heterogeneous but surface-assimilated clusters such as /k/ + /b/ realised as [gb].

The second reason it is difficult to establish the predictions of recent lexical feature analyses is simply that the relation between autosegmental phonological representations and linguistic phonetic forms (i.e., the data structures required at the interfaces with peripheral processing) is not spelt out, or at least not fully. There are two general ways of thinking about this relation. According to the first, autosegmental feature models occupy the phonology component of modular models of the phonology-phonetics interface. The second sees autosegmental representation itself as (the backbone) of linguistic phonetic representation. The latter view is the one adopted by Harris & Lindsey (1995) and Avery & Idsardi (2001), albeit in different ways.

I will not try to flesh out all the possible predictions about the phonetics of voicing assimilation that can be derived from autosegmental representation and different takes on its relation to linguistic phonetic representation. For the purpose of deriving phonetic predictions I will rather take the sort of spreading-cum-delinking rule illustrated in figure 4.1 as a convenient shorthand for the traditional polarity-switching rules [-tense] → [+tense]_+[tense], [+tense] → [-tense]_-[-tense], and (combined) [-αtense] → [αtense]_-[αtense]. In addition, I will assume that [tense] is a proper phonological feature, and therefore that voicing assimilation operates in the phonological component of a modular model of the phonology-phonetics interface. Note that prominent representatives of autosegmental lexical feature analyses of voicing assimilation such as Lombardi (1994, 1995a,b, 1996) and Iverson & Salmons (1995, 1999) leave this option completely open. Moreover, the finer distinctions that can be encoded by autosegmental lattices are generally ignored by optimality-theoretic models of laryngeal phonology that use *AGREE*-type constraints (e.g., Lombardi 1995c, 1997, 1999; Grijzenhout & Krämer 1998; Borowsky 2000: see 8.2.5). These accounts effectively treat all surface [αtense][αtense] sequences identically regardless of whether they are derived from homogeneous or heterogeneous underlying clusters.

Given the general absence of explicit models of phonetic interpretation in the (recent) generative literature, I feel that my ‘phonemic’ reading of its phonetic implications does not do it any gross injustice. This opinion is bolstered by oc-

casual explicit assertions that voicing assimilation processes are neutralising (e.g., Siptár & Törkenczy 2000), or that optional and/or non-categorical phenomena lie without the scope of generative models (Lombardi, 1999). However, any reader who derives specific phonetic predictions from an autosegmental (or any other) model that do not agree with those I will extract in the next few paragraphs is invited to treat the latter as a straw man. This reader is also invited to test his or her own predictions against the data reported in chapters 5, 6, and 7. However, any autosegmental model of voicing assimilation will also have to address the more fundamental issues raised in chapters 1 and 8.

Now, the key property of a polarity-switching lexical feature analysis of voicing assimilation is that sequences that are subject to assimilation surface as phonologically identical to underlyingly homogeneous sequences. Thus, this type of analysis draws no structural distinction between underlying /g/ + /b/ on the one hand and /k/ + /b/ after regressive assimilation on the other: both are represented as identical [-tense][-tense] sequences. Two key predictions follow from this property (cf. 10a and 10b below). First, the fact that underlying [α tense][α tense] sequences and assimilated clusters are phonologically identical means that they are phonetically identical as well, and therefore that voicing assimilation is a phonetically neutralising process. Phonetic interpretation has no access to information about the underlying status of the obstruents in a cluster and can therefore assign only a single phonetic form to any [α tense][α tense] sequence in any (wider) context. Thus, underlyingly homogeneous /g/ + /b/ and regressively assimilated /k/ + /b/ are predicted to surface as identical [gb].

- (10) Predictions of a (polarity-switching) lexical feature analysis of voicing assimilation
- a. Voicing assimilation processes are phonetically neutralising (polarity-switching analyses only)
 - b. All locally relevant cues to [tense] participate in voicing assimilation processes
 - c. The ability of a given tense or lax obstruent to trigger or undergo voicing assimilation is unrelated to the phonetic manifestation of [tense] in that obstruent

Second, the neutralising nature of voicing assimilation under a lexical feature analysis entails that the process affects all cues to [tense] that are relevant in the environment of the assimilation target. Strictly speaking, this means that *voicing assimilation* is a misnomer. For instance, if the /g/ in a /g/ + /b/ sequence is normally distinct from /k/ in /k/ + /p/ clusters in terms of preceding vowel duration, closure duration and voicing, then assimilated /k/ in /k/ + /b/ is predicted to share all of these characteristics. In a sense therefore, all cues signalling the fortis-lenis distinction are predicted to ‘spread with’

[tense] where voicing assimilation applies. This is symbolised by the cluster of phonetic features suspended from the spread $[\pm\text{tense}]$ feature of Root_2 in figure 4.1. Note, incidentally, that this second prediction might also be assigned to a spreading-without-delinking analysis. On this view, voicing assimilation does not completely neutralise [tense] distinctions in target obstruents, but still affects the whole cluster of relevant cues. So prediction (10b) follows from a wider range of lexical feature accounts than (10a).

If the fortis-lenis contrast is universally represented with a single formal feature, e.g., $[\pm\text{tense}]$, a third prediction that follows from a lexical feature analysis is that voicing assimilation is not phonetically conditioned. For example, if the passively and actively voiced lenis stops /b, d, g/ of English and Yiddish respectively are represented by the same feature and value [-tense], both languages are equally likely to have a rule spreading this feature value, despite the phonetic differences between their lax stops (assuming that [tense] spreading rules are a necessary part of the universally available inventory of phonological rules). Along similar lines, a lexical feature analysis built on a single universal feature [tense] (or some equivalent) predicts that if a language has passively voiced lenis stops but actively voiced lenis fricatives, they should behave alike in terms of their assimilatory behaviour. Only if there are two or more distinct features in the universally available pool to represent tense-lax distinctions can these and similar predictions be avoided, because this allows for the (crude) matching of phonological with phonetic categories, and thus for the phonology to have access to phonetic information. Models that take this approach are discussed in chapter 8.3.

4.1.2 Articulation-driven voicing assimilation

According to the logic of articulation-driven models, the phonetic effects of a sandhi rule are purely determined by the articulatory properties of the sounds involved. In order to establish any predictions about the phonetic characteristics of an articulation-based form of voicing assimilation, it is therefore necessary to establish which reflexes of [tense] are ‘available to’ the coarticulation mechanism. It is argued below that the main phonetic features of [tense] that are able to trigger articulatory adjustment in a neighbouring sound are the articulations involved in the production of voicing contrasts.

The key idea behind this theory is probably best illustrated with the behaviour of actively voiced lax stops. Recall from 2.1 above that actively voiced stops are likely to be produced with an array of enhancing gestures aimed at enlarging the size of the oral cavity behind the occlusion, and (to a lesser extent) lowering the threshold transglottal pressure difference for vocal fold vibration. Under virtually all models of articulatory implementation this entails that these gestures are subject to coarticulation. As reviewed in chapter 1, speakers are as-

sumed to set certain limits on the extent of coarticulation that depend in part on their communicative needs and are in part likely to be language-specific, but the underlying mechanism (negotiation between subsequent targets involving the same or mechanically linked articulators) is invariably supposed to be universal. Consequently, the voicing-enhancing gestures associated with an actively voiced stop are predicted to ‘spill over’ into neighbouring sounds, and mainly the preceding one(s) if there is indeed a bias towards anticipatory over perseveratory coarticulation (cf. Farnetani 1997). Because the gestures that spill over are designed to control the voicing of the trigger obstruent, they will also have an effect on the voicing of neighbouring sounds. For example, if a word-initial /b/ is produced with an advanced tongue root and lowered larynx in order to facilitate voicing, a coarticulated preceding /k/ will be subject to some degree of larynx lowering and tongue root advancement, too. If the effect is sufficiently strong it will improve the conditions for voicing during the /k/ and thereby lengthen the voice tail into this obstruent from a preceding vowel.

Along the same lines, the voicing control measures associated with actively *devoiced* obstruents are predicted to influence the voicing of neighbouring obstruents. Passively voiced plosives, on the other hand can have no coarticulatory effect on the voicing control of neighbouring obstruents: if the word-initial /b/ referred to above is produced without any articulated moves designed for voicing control, there is simply nothing to spill over into flanking sounds. Instead, the voicing control gestures of flanking sounds will have relatively free reign in a passively voiced sound since there is no local voicing control to counterbalance the spill over.

Thus, a coarticulation-based account of voicing assimilation derives the process as the inevitable by-product of two mechanisms that are plausible on independent grounds. As discussed in chapter 1, coarticulation (in a narrow phonetic sense) is normally conceived as a truly universal mechanism that governs the production of any continuous sequence of sounds produced by any speaker at any given time. Although there are differences of opinion about how the mechanism should be modelled exactly, its reflexes in a wide variety of contexts are well-documented, and consequently its existence is motivated independently from any observations concerning (de)voicing in obstruent clusters. Similarly, as pointed out in 2.1 above, the existence of secondary articulations aimed at voicing control is arguable on independent (aerodynamic) grounds, and is not invoked only to deal with voicing patterns in obstruent clusters.

In other words, according to a coarticulation-based theory, RVA at word boundaries occurs automatically as part of a much more general process of sound co-production every time an actively (de)voiced obstruent is juxtaposed with another obstruent: *a specific voicing coarticulation rule does not have to be postulated*. In this respect, a coarticulation-based approach to word bound-

aries is different from a lexical feature analysis. Rules of [\pm tense] agreement or spreading-and-delinking adhere to more general formal templates, but formal grammars in which every phonological feature agrees or spreads in every possible context have to my knowledge never been proposed. Consequently, [tense] spreading rules have to be specified explicitly for every language which possesses RVA at word boundaries.

This study is certainly not the first to look at voicing assimilation as a coarticulation process: notably [Slis \(1985, 1987\)](#) and more recently [Ernestus \(2000\)](#) have made similar proposals. They are supported in part by a series of studies of glottal coarticulation ([Löfqvist, 1980, 1981](#); [Löfqvist & Yoshioka, 1980, 1984](#); [Yoshioka et al., 1981, 1982](#)), which demonstrate that the glottal abduction gestures associated with contrastively voiceless stops and fricatives vary according to the phonetic context. For example, the results of [Yoshioka et al. \(1981\)](#) indicate that when flanked by sonorants or a pause, English word-final /s/ and word-initial /k/ are accompanied by an opening and closing movement of the glottis with a distinct peak. But when word-final /s/ is combined with word-initial /k/ the results is not two discrete peaks separated by a brief closure of the glottis, but a composite gesture without full medial closure, and in some cases only a faint trace of two separate peaks. In the light of acoustic studies reporting incompletely neutralising RVA, it is interesting that this composite gesture remains distinct in timing, shape, and size from the large single peak found in word-initial /s/ + /k/ sequences. Moreover, the posterior cricoarytenoid muscle, which is largely responsible for glottal abduction, shows two distinct activity peaks in the production of the former but not the latter.

However, none of these studies spell out the precise typological implications of a coarticulation-based theory of voicing assimilation. Some of these predictions appear in (11). First, a coarticulation-based approach predicts that only the cues to [tense] that ‘can be coarticulated’ should participate in voicing assimilation. As explained above this set includes active (de)voicing because it is supported directly by articulatory gestures. In addition, it includes any phonetic features that are mechanically dependent on voicing gestures such as the duration of frication in fricatives, which varies as a function of the size of (peak) glottal abduction (see 2.3.2 above). The set of features that can be coarticulated potentially also includes F_0 and F_1 , but as pointed out in 2.2.3 the articulatory basis of these cues remains unclear.

By contrast, assimilatory effects on the *durational* correlates of [tense] would seem to be excluded by a coarticulation model. The work of [Kluender et al. \(1988\)](#) rules out a mechanic link between voicing and obstruent or preceding vowel duration (cf. 2.2.3), and consequently the coarticulatory spill over of voicing control moves should not be able to influence durational parameters apart from frication duration in fricatives (and perhaps stop releases: see 2.3).

Second, on most accounts, segmental duration does not correspond to ‘substantive’ articulatory gestures but to the *scaling* of such gestures in phonetic space and/or time. This scaling can be modelled ‘extrinsically’ by directly assigning (relative) durations to phonetic targets, or ‘intrinsically’. Intrinsic timing control consists of aligning specific articulatory or auditory landmarks of subsequent sounds with each other (e.g., the onset of a nasal with the oral release of a preceding plosive) and specifying the excursion size and/or speed of the gestures used in their production (Browman & Goldstein 1986 et seq.; Huffman 1993; Kirchner 1998).² But under neither of these analyses does segmental duration ‘spread’ or ‘spill over’ into neighbouring sounds. For example, modelling increased duration by assigning greater excursion sizes to the relevant articulators may in some cases force a greater amount of coarticulation on neighbouring sounds (to compensate for longer trajectories between subsequent articulatory targets), but the resulting adjustments would not lead to lengthening of those neighbouring sounds.

Note that this property of coarticulation models is attractive on empirical grounds, because relative phonetic length generally does not spill over between neighbouring sounds: phonetically long vowels, for instance, are not necessarily flanked by long consonants. In this respect coarticulation models clearly mirror recent generative treatments of (contrastive) length as prosodic rather than substantive. Such models encode length distinctions above the *Root* level (cf. figure 4.1) as positions on a dedicated timing (moraic or skeletal) tier, which (given certain basic assumptions) means that in contrast to, e.g., [nasal] or place of articulation, length does not spread between sounds. In this sense, the prediction of a coarticulation-based model that RVA should have no effect on the duration of the target obstruent and the preceding vowel parallels the prediction of generative models that phonological geminates do not induce lengthening in neighbouring vowels.

The second prediction of a coarticulation-based approach is that as any form of coarticulation, coarticulatory voicing assimilation rules should operate as gradient rules that are sensitive to speech rate and style to the degree that such phenomena can be understood in terms of global hypoarticulation. This implies that voicing assimilation can occasionally be phonetically neutralising, but only at the end point of a scale that contains many incompletely neutralised cases. A further implication of the phonetic status of voicing assimilation is that the process is automatically blocked by a sufficiently long physical pause: unlike the effects of phonological feature spreading or agreement, which operate on adjacent points in abstract time, the effects of coarticulation decay as a function of real time. As a result, any stop or fricative that is separated from a neighbouring actively (de)voiced obstruent by a (silent) interval of a certain duration will be

²The distinction between extrinsic and intrinsic timing is due to Fowler (1980).

‘out of reach’ of the anticipation or perseveration of the relevant (de)voicing-enhancing articulatory moves.

- (11) Predictions of a coarticulation-based approach to voicing assimilation
- a. Only those cues to [tense] that are subject to coarticulation in obstruent sequences (mainly voicing) participate in voicing assimilation
 - b. Voicing assimilation is a gradient, non-neutralising process
 - c. The ability of a given tense or lax obstruent to trigger or undergo voicing assimilation is a function of the phonetic manifestation of [tense] in that obstruent (and primarily its voicing targets)

I already described the third prediction of the phonetic theory of voicing assimilation proposed in this section with specific reference to the behaviour of lax stops: the ability of a given obstruent to trigger or undergo voicing assimilation is intrinsically related to the phonetic manifestation of [tense] in that obstruent. Given that the articulation of voicing distinctions provides the main set of articulatory gestures that are available for coarticulation, it follows that the capacity of a segment for triggering voicing assimilation can be predicted from its voicing target (which can in turn be read off its behaviour in utterance-initial and postsonorant environments).

This general hypothesis brings several interesting predictions with respect to the patterning of specific classes of sounds in specific languages. For instance, English lax obstruents are predicted to behave in a manner-asymmetric fashion with respect to voicing assimilation: actively voiced /v, z, ʒ/ but not (passively voiced) b, d, g, d͡ʒ are expected to trigger voicing assimilation in a preceding obstruent. Furthermore, both types of fortis obstruent defined above (aspirated and plain voiceless) should in principle trigger voicing assimilation in a preceding obstruent, since both categories are arguably actively devoiced. Consequently (regressive) voicing assimilation to fortis but not lenis stops is predicted to occur in aspirating as well as voicing languages.

A final and potentially important prediction of the present theory of coarticulatory voicing assimilation is that sonorant consonants should never trigger RVA, except in languages where they are contrastively (de)voiced. I noted in 2.1 that it is likely that in languages without contrastive voicing on sonorants, the common voicing of this class of sounds is a result of the voiced carrier sound and passive voicing rather than the presence of an active voicing target. Since the phonetic theory described in this section derives all RVA (at word boundaries at least) from the presence of active (de)voicing targets, it follows that noncontrastively voiced sonorants should not be able to trigger voicing assimilation.³

³Many recent generative models make a similar prediction because they treat sonorants without contrastive voicing as unmarked for the [voice] or an equivalent feature.

4.2 Voicing assimilation in morphological paradigms

The Germanic group of languages provides two excellent examples of voicing assimilation rules that appear to be phonological in nature, insofar as they are not conditioned by the phonetics of tense. The first example, illustrated with well-known examples from English and Dutch in (12) consists of the alternation of a [t]-initial suffix after stems ending in a fortis obstruent with a [d]-initial suffix after (most) sonorant-final and lax obstruent-final stems. Where it is associated with past tense morphology I will refer to it as the ‘Germanic past tense paradigm’ for convenience. What is perhaps the most common analysis of this allomorphic rule treats the [d]-initial variants as underlying and derives the [t]-initial forms by progressive assimilation to stem-final fortis obstruents but the pattern itself does not force this analysis. But irrespective of whether [d] or [t] is regarded as underlying, the essential generalisation remains that the process derives obstruent clusters which are homogeneous in terms of voicing, or perhaps rather the lexical feature [tense]. Crucially for present purposes, it is not phonetically conditioned to the extent that it occurs both in voicing varieties (e.g., standard Dutch, Scottish English, Frisian) and aspirating varieties (e.g., north-eastern dialects of Dutch, standard varieties of English, Swedish, and Norwegian) of Germanic. A highly similar pattern is found in several Turkic languages, where it appears to be equally insensitive to the voicing aspirating distinction, since it occurs both in e.g., (aspirating) Turkish and (voicing) Turkmen.⁴

Quantitative information on the phonetic behaviour of the various incarnations of the pattern in (12) is not available, but it seems fairly safe to assume that it is discrete rather than gradient in all of the languages in which it occurs, and affects all correlates of [tense]. Like final laryngeal neutralisation, it certainly does not exhibit the characteristics that are typical of phonetic rules such as dependency on speech rate and/or style. It is never optional, and in some ways even phonetically abstract. For instance, many Dutch speakers retain a lexical distinction between stem-final [x] and [ɣ] in their marking of the regular past tense (cf. /lɑx/ + /tə/, [lɑχtə], *laughed* vs. /za:ɣ/ + /tə/, [za:ɪdə], *sawed*) without phonetically realising the contrast in any other environment: such speakers produce both /lɑx/ + /ən/ and /za:ɣ/ + /ən/ with a fully voiceless velar or uvular fricative. Thus, the pattern in (12) can be maintained in the absence of any phonetic motivation.

⁴Needless to say, the pattern in (12) is not limited to past tense paradigms but also governs, e.g., the English regular plural. The Norwegian version of the rule shows an interesting twist in comparison with its cognates. [d(ə)] suffixes appear after vowels, glides and lenis obstruents including /v/-/v/ whereas [t(ə)] is used after fortis obstruents and (other) sonorant consonants (Kristoffersen, 2000). Baitchura (1975) suggests that like most other Turkic languages, Turkmen is an aspirating language, but the grammar by Clark (1998) describes it as voicing.

(12) Progressive assimilation in the ‘Germanic past tense paradigm’

a. Dutch past tense paradigm

UR	Phonetic form	Gloss
/zak/ + /də/	[zaktə]	failed
/vas/ + /də/	[vastə]	washed
/dæb/ + /də/	[dʌbdə]	doubted, wavered
/xra:z/ + /də/	[xra:zdə]	grazed
/krœl/ + /də/	[krʌdə]	curled

b. English past tense paradigm

UR	Phonetic form	Gloss
/lɒp/ + /d/	[lɒpt]	lopped
/lɒk/ + /d/	[lɒkt]	locked
/lɒb/ + /d/	[lɒbd]	lobbed
/lɒg/ + /d/	[lɒgd]	logged
/lo:n/ + /d/	[loʊnd]	loaned

The second assimilatory pattern that is likely to be a case of [tense] assimilation rather than coarticulatory voicing assimilation is illustrated in (13) with the behaviour of the Norwegian adjectival agreement marker /t/, which causes a preceding lax obstruent to fully devoice (Kristoffersen, 2000). It can be accounted for in a straightforward manner by spreading a [+tense] feature backwards from the agreement marker. Similar processes are found in morphosyntactically restricted contexts across the Germanic group including (aspirating) Danish (Panzer, 1981), Swedish (Hellberg, 1974), and (voicing) Yiddish (Birnbaum, 1979; Katz, 1987), and depending on how the pattern in (6) in chapter 3 is analysed, (aspirating) German and (voicing) Dutch as well. Thus, Yiddish realises /red/ + /st/, *you-SG.FAM. speak* as [retst] and /fraib/ + /st/, *you-SG.FAM. write* as [fraipst] (Katz, 1987).⁵

Again, quantitative evidence demonstrating the phonetically discrete behaviour of this type of rule does not seem to be available, but like instances of the Germanic past tense paradigm, it typically displays all the hallmarks of a phonological rule. Interestingly, Kristoffersen (2000) finds that there is at least one lexical exception to the Norwegian rule in (13): /glɔg/, *intelligent* surfaces with a (voiced) lax stop in both its uninflected ([glɔg]) and inflected ([glɔgt]) forms. Note that regressive [tense]/voicing assimilation in morphological paradigms is by no means restricted to [+tense] spreading/devoicing: regressive assimilation to suffixes with tense or lax initial obstruents (including single /d/) occurs in Hungarian (Kenesei et al., 1998; Siptár & Törkenczy, 2000).

⁵The same pattern is found in a small number of (lexicalised) English forms such as the irregular past tense [left] from /li:v/, and [lost] from /lu:z/.

- (13) Norwegian regressive ‘devoicing’ (data from [Kristoffersen 2000](#))

UR	Uninflected	Inflected	Gloss
/tryg/	[tryg]	[try kt]	secure
/grov/	[gro:v]	[grɔft]	coarse
/stiv/	[sti:v]	[sti:ft]	rigid

4.3 Progressive devoicing at word boundaries

Progressive devoicing refers to the situation in which an obstruent that is voiced in (some) other contexts becomes voiceless when preceded by another obstruent. At first sight this might be regarded as a good description of what happens to the underlying /d/ of the ‘Germanic’ past tense suffix in (12), but it has long been recognised that some cases of progressive devoicing are better treated as passive devoicing in the sense of 2.1. For instance, [Harms \(1973\)](#) notes that [mæddɔŋ] with a voiceless second /d/ is a common realisation of English <mad dog>. Similarly, the final obstruents of <bagged> or <lodged> are often phonetically voiceless, even though they remain distinct from the final sounds of <racked> or <botched> (e.g., [Jones 1956](#)). This is illustrated in figure 4.2. A lexical feature analysis of these cases is inappropriate on two grounds. The first is that a [tense] spreading or agreement account predicts that devoicing occurs only after tense obstruents, which is clearly incorrect. As far as English lenis plosives are concerned the second reason is that these sounds are generally voiceless in a wider range of contexts (e.g., utterance initially) that is perfectly predictable once it is assumed that they are only passively voiced (cf. 2.1).

Progressive devoicing at word boundaries is not restricted to passively voiced obstruents such as the lenis stops of English (or German: [Drosdowski & Eisenberg 1995](#)). Actively voiced obstruents are also routinely subject to devoicing when preceded by another obstruent. In spite of the fact that the devoicing of actively voiced obstruents can not be attributed to their lack of voicing targets, a number of the cases in question are probably amenable to an explanation in terms of passive devoicing.

Consider the devoicing of lax stops in Dutch as an example. A production study carried out by [Slis \(1986\)](#) shows that the devoicing of Dutch word-initial lax plosives preceded by another obstruent is sensitive to (nuclear) stress: of the lax plosives in the onset of a stress-bearing syllable 21% is devoiced, but of those in the onset of a stressed syllable no less than 57% is voiceless. One possible account of this process treats it in terms of a rule spreading [+tense] from a preceding obstruent (such an approach is available to the many models that treat the single (neutralised) series of word-final obstruents in Dutch as [+tense]).

Alternatively, Dutch lenis plosive devoicing can be viewed as another ‘non-rule’ in the sense of [Harms \(1973\)](#). The key to this second analysis is the obser-

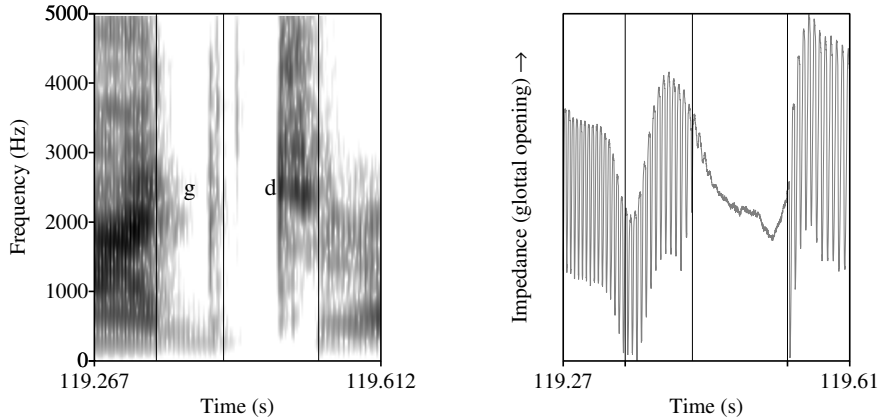


Figure 4.2: Progressive devoicing in English. Broad band spectrogram (left) and EGG trace of English <nagged>, produced in the carrier sentence *with what does — rhyme?* The stem-final velar stop is voiced throughout, but the suffix /d/ is almost completely voiceless.

vation that the prolongation of an obstruent constriction makes the continuation or initiation of voicing progressively more difficult, even if voicing-enhancing gestures are present (cf. 2.1). Thus, the presence of a preceding obstruent imposes adverse aerodynamic conditions for the voicing of a lax obstruent, which means that articulatory moves aimed at voicing control are less likely to sustain (full) voicing than in phonetic contexts that are less hostile to vocal fold vibration. Within certain physical limits, it should be possible to counteract the aerodynamic conditions imposed by a preceding obstruent by expanding the gestures involved in voicing control. However, the general effect of destressing is articulatory weakening rather than strengthening, which means that in unstressed contexts the articulatory gestures involved in the production of active voicing are less likely to be sufficient to counteract the effect of a preceding obstruent. In other words, under a passive devoicing account the effect of stress on the devoicing of lax plosives does not have to be stipulated (as it has to in a [+tense]-based analysis), but follows from the independently motivated effect of prosodic structure on articulation (cf. chapter 1).⁶

The partial devoicing of English lenis fricatives, briefly mentioned in 2.3 is a further likely candidate for analysis in terms of passive devoicing. It is unclear, however, whether the same applies to the devoicing of Dutch lax fricatives. This process, which is illustrated in (14), is highly noticeable and generally regarded

⁶Ernestus (2000) presents an analysis roughly along these lines of the process that devoices the initial /d/ of several Dutch function words.

as a phonological rule, unlike the processes described in the previous paragraphs (Booij, 1995; Lombardi, 1997; Grijzenhout & Krämer, 1998; Ernestus, 2000). This is reflected in the customary transcription of devoiced /v, z/ as [f, s] rather than [v̥, z̥].

An alternative approach is to treat Dutch lax fricatives as lacking a voicing target, analogously to the lax plosives of aspirating languages. This approach would predict that the contrast between fortis and lenis fricatives is consistently realised as [f, s] vs. [v̥, z̥] utterance initially as well as after obstruents. Ernestus (2000) claims that Dutch /v, z/ may be produced with voicing utterance initially, but there seems to be no quantitative data available to assess the extent of this voicing (and compare it with the voicing of e.g., English lax fricatives). Given the account of regressive voicing assimilation at word boundaries developed below, it would also be predicted that Dutch fricatives do not trigger RVA, which according to Slis (1987) they do only very occasionally. Finally, one of the key predictions of a passive devoicing account of Dutch postobstruent fricative devoicing (and every other passive devoicing analysis) is that /v, z/ retain the other phonetic features of [-tense] in contexts where they are devoiced. A lexical feature account on the other hand, predicts that the devoiced fricatives of /rit/ + /zɑŋər/ and /reiz/ + /vərha:l/ are phonetically identical to lexical /f, s/.

(14) Progressive devoicing of lax fricatives in Dutch

UR	Phonetic form	Gloss
/drœk/ + /vat/	[drœkfət]	pressure vessel
/rit/ + /zɑŋər/	[ritsɑŋəɪ]	sedge warbler
/lɔf/ + /zɑŋ/	[lɔfsɑŋ]	eulogy
/klas/ + /vərklɛmɪŋ/	klasfərklɛmɪŋ]	class size reduction
/drɛiv/ + /zɑnd/	[drɛifsɑnt]	quicksand
/reiz/ + /vərha:l/	[reɪsfərha:l]	travel story

4.4 Regressive voicing assimilation at word boundaries

It has been common practice in generative phonology to treat assimilation phenomena of the type described in section 4.2 above on a par with voicing assimilation rules operating word sandhi contexts. This is especially evident in some generative accounts of the typology of voicing assimilation phenomena and in descriptions of assimilation rules that fail to mention whether the processes in question are in any way restricted by the morphology. The attempt by Lombardi (1999) to force the voicing assimilation processes found in Yiddish and Swedish into a single typology defined on purely phonological grounds provides recent examples of this style of analysis. It is striking that Lombardi proceeds with her account despite the fact that by all descriptive accounts (several of which she

cites) voicing assimilation in Swedish is restricted to very specific morphosyntactic paradigms (e.g., [Hellberg 1974](#)), whereas Yiddish has regular regressive voicing assimilation across word boundaries as well as within morphological paradigms (e.g., [Katz 1987](#)).

In the light of a tradition of such analyses, and the lexical feature view of voicing assimilation more generally, it is no surprise to find explicit assertions of the idea that voicing assimilation is not (ever) conditioned by the phonetic realisation of [tense]. Thus, [Keating \(1984\)](#) states that:

[c]luster voicing assimilation is another common phonological rule which appears to apply generally across phonetic categories. Thus Polish has regressive voicing assimilation [...] and a [voice] contrast of {voiced} vs. {vl. unaspirated} stops; Danish however, has progressive 'voicing' assimilation, but an aspiration contrast in initial position [...].([Keating 1984:294](#))

There is evidence however, that in contrast to the rules described in 4.2 above, voicing assimilation under word sandhi is, or at least can be, phonetically conditioned as well as phonetically gradient. The principal observation underpinning the first claim is that lax plosives only seem to trigger RVA across word boundaries if they belong to the actively prevoiced category; the most reliable evidence for the second is provided by quantitative studies of RVA in a number of languages. To some extent this evidence is corroborated by descriptive claims that assimilation across word boundaries is incomplete or optional and dependent on speech style or register and blocked by physical pauses. But as such claims often differ from language to language and linguist to linguist (as is shown by a comparison of, e.g., [Vago 1980](#) and [Siptár & Törkenczy 2000](#)), it is difficult to gauge their articulatory and/or perceptual reality or their generality.

The observation that word-initial lax stops only trigger RVA if they belong to the actively prevoiced category is made by [Kohler \(1979\)](#) as part of a description of the realisation of [tense] in German (dialects) and French. It appears to hold across languages that maintain a distinction between plosives that is (partially) based on VOT distinctions and is beautifully illustrated by the typology of Germanic languages and dialects. All and only the varieties of this group that are described as employing a prevoiced vs. short lag realisation of fortis and lenis plosives are also reported as exhibiting RVA to lenis plosives (cf. [Kohler 1979, 1984](#)): Afrikaans ([Wissing, 1991](#)), (Western and Southern) Dutch (e.g., [Cohen et al. 1972](#)), Scottish English ([Kohler, 1979](#); [Wells, 1982a](#)), (West) Frisian ([Riemersma, 1979](#); [Tiersma, 1985](#)), Rhineland German ([Kohler, 1979](#)), and (Eastern varieties of) Yiddish ([Katz, 1987](#); [Jacobs et al., 1994](#)). The remaining varieties of English and German (that maintain a VOT contrast for their initial plosive series) and the North Germanic languages, all of which belong to

the aspirating type, show no RVA to lenis plosives.⁷

Although the assimilatory behaviour of lax stops indicates that RVA at word boundaries can be phonetically *conditioned*, this does not in itself establish the phonetic (gradient, cue-specific) status of the phenomenon. Direct quantitative evidence of gradience is hard to find, but there is ample evidence for *incomplete neutralisation* in RVA at word boundaries, which is a necessary albeit not a sufficient property of coarticulation rules. This evidence provided by a range of quantitative studies: O. Thorsen (1966) on French, N. Thorsen (1971) on English, Charles-Luce (1993) on Catalan, Burton & Robblee (1997) on Russian, and Barry & Teifour (1999) on Syrian Arabic. These studies investigate the effects of the [\pm tense] specification of a word-initial obstruent (C_2) on the phonetic features of a single preceding (word-final) obstruents (C_1). In spite of their slight methodological differences, all three studies find that assimilation of C_1 to C_2 does not completely erase underlying [tense] distinctions in C_1 position. In other words, in contrast to prediction (10a) of a lexical feature analysis underlying /k/ + /b/ and /g/ + /b/ sequences do not surface as phonetically indistinguishable [gb]. For example, Barry & Teifour (1999) found that the mean duration of the voiced interval of a [+tense] C_1 fricative followed by a [-tense] obstruent (e.g., /s/ + /d/) was shorter (63 ms) than that of a underlyingly [-tense] fricative in the same context (e.g., /z/ + /d/: 88 ms) but considerably longer than that of a [+tense] fricative followed by another [+tense] obstruent (e.g., /s/ + /t/: 11 ms).

In addition, the study of Catalan by Charles-Luce (1993) provides some evidence against prediction (10b) of the phonological approach. Charles-Luce reports that whereas C_1 voicing in Catalan stop + /s/ and stop + /r/ and the duration of the preceding vowel show the expected assimilatory behaviour (albeit not always in a neutralising fashion), this does not extend to the duration of C_1 itself. A lexical feature analysis predicts that C_1 should be longer before /s/ than before /r/, but Charles-Luce (1993) finds exactly the opposite pattern.⁸

⁷Myers (2002) provides quantitative data on English which appear to show that (aspirating) English lax obstruents do have the capacity to trigger RVA. However, my own experimental data, reported in chapter 5, strongly indicate that the overall effect of lax obstruents on the voicing of a preceding obstruent found by Myers is an artefact of a manner-specific effect of English lax fricatives. Slis (1987) found no difference in RVA between speakers of (voicing) southern/western dialects and (aspirating) north-eastern varieties of Dutch. However, he failed to establish whether his north-eastern subjects indeed used long lag VOT tense stops and passively voiced lax stops during the experiment, or whether they shifted their pronunciation towards the voicing standard language.

⁸Charles-Luce (1993) seems to treat /r/ on a par with lax obstruents (which are excluded from her experiment) as a trigger of voicing assimilation, presumably because Catalan word-final obstruents are regularly voiced before sonorants as well as before lax obstruents. There are good reasons to distinguish between voicing assimilation to sonorants and voicing assimilation to lax obstruents (see below), but they do not alter the conclusion that the results obtained by Charles-Luce (1993) run counter to a lexical feature analysis of RVA.

Finally, there are reports of regressive assimilation to non-contrastively voiced sonorants, e.g., in Krakow Polish, in Catalan, and in Frisian (fricatives only). This observation might seem to contradict prediction (11c). There is an important argument to treat this phenomenon as distinct from regressive assimilation to actively voiced lenis obstruents, however. It seems that only word-final obstruents that are subject to (dynamic) laryngeal neutralisation can assimilate to a following sonorant: note that all the languages mentioned above neutralise the fortis-lenis contrast word finally. Moreover, the only dialects of Hungarian which exhibit RVA to sonorants are the ones that also have final neutralisation, e.g., the variety of the West Dunántúl region (Kiss, 2001). Varieties of Hungarian that maintain the contrast between tense and lax obstruents word finally only show regressive assimilation to obstruents.

This apparent generalisation suggests an alternative account of voicing assimilation to sonorants, which relies on the idea that neutralised obstruents lack targets for voicing and other correlates of [tense]. Evidence for this idea was reviewed in 3.2.1 above. If neutralised obstruents indeed lack voicing targets, they should show a greater degree of voicing between a vowel and a following sonorant than actively devoiced obstruents, simply as a result of the passive continuation of voicing into the constriction phase. It could well be this increased amount of voicing (relative to utterance-final and $_{-}$ [+tense] contexts) that is interpreted by linguists as voicing assimilation. It could also become a source of confusion to listeners, who might reanalyse all presonorant obstruents (along with obstruents preceding a lax obstruent) as [-tense] on the surface, at least in theory (which would in turn lead to pronunciations that are likely to be interpreted as assimilation by linguists).

According to the first variant of this explanation, word-final neutralised obstruents before sonorants should have less voicing than those preceding an actively voiced obstruent: the latter would ‘inherit’ some of the active voicing of the following sound in addition to the passive voicing spilling over from the preceding vowel or sonorant consonant. According to the second variant there should be at least a group of languages for which this description holds, i.e., languages such as Dutch, which have final neutralisation, but on most descriptions lack RVA to sonorants. Chapter 7 reports direct evidence for the prediction that neutralised obstruents have more voicing when followed by a sonorant consonant than before a tense obstruent, but less than when followed by an actively voiced lenis stop.⁹

⁹The idea that neutralised stops can undergo passive voicing between voiced sonorants and thus appear to exhibit assimilation of voicing is not new: it is central to the account of English flapping in Harris (1994).

4.5 Summary

The aim of this chapter was to dissect voicing assimilation, a phenomenon that might be regarded by some as one of the better understood phonological processes. Drawing from the theoretical treatment of sandhi processes more in general, I first defined two phonetic templates: one for a (lexical) phonological form of voicing assimilation and one for a purely coarticulatory form. Subsequent chapters then matched descriptions of various assimilation processes against these templates.

The first set of processes I considered involved the voicing assimilation rules found in the morphological paradigms of the Germanic and other languages. These processes appear to be phonological in nature, although as yet no clinching quantitative evidence is available. Second, in many instances, the phonetic characteristics of progressive devoicing rules that operate across word boundaries suggest that they are fully explained in terms of passive devoicing. Third, regressive voicing assimilation across word boundaries is clearly conditioned by the phonetic properties of the trigger obstruents in a way that suggests that they are driven by coarticulation, or at least diachronically rooted in coarticulation. This suggestion is bolstered by quantitative evidence indicating that several RVA processes that operate across word boundaries are incompletely neutralising. The next three chapters represent an attempt to extend the quantitative evidence in this area.