Voicing and the Skeleton

Márton Sóskuthy

The main goal of this paper is to show that the behaviour of voicing in obstruents can better be accounted for if we assume that voicing is represented by a specific skeletal configuration rather than by phonological features. It will be shown that previous analyses of voice phenomena run into serious problems, especially when it comes to the representation of voice assimilation and intervocalic voicing. The analysis outlined in the final section of this paper gives a representationally motivated account of voice phenomena in CVCV and manages to unite the effects of voice assimilation, word-final devoicing and intervocalic voicing. The representation proposed here also makes a number of predicitions, which may well prove to be correct: (i) languages with prevoicing in initial stops necessarily have regressive voice assimilation (ii) progressive assimilation and coda devoicing appear only in languages with a contrast based on different phonetic properties (e.g. aspiration).

1. Introduction

Ever since the publication of SPE, there has been a marked tendency in phonological thinking to reduce the number of features employed in the description of speech sounds and to devise a more articulate theory of suprasegmental structure. As a result, some traditional features such as [long], [syllabic], and [stress] have been relegated to skeletal and suprasegmental structure. The main benefit of this approach is that it gives a plausible explanation for why certain phonetic properties of sounds (e.g. syllabicity) seem to be so closely tied up with prosodic structure and that it allows us to make a clear distinction between quality and quantity. One set of features whose melodic status has recently been challenged are MANNER FEATURES (for detailed discussion see Szigetvári to appear), as opposed to PLACE and LARYNGEAL features, which arguably have a genuine featural representation. This line of research seems to me as a very promising enterprise, as it helps to build a more streamlined representational vocabulary and it explains the differential behaviour of distinct classes of features.

In this paper, I will argue that a similar alternative is available for the representation of voicing in obstruents. More specifically, I will demonstrate that the feature [voice] (regardless

Proceedings of ConSOLE XV, 2008, 247-270
http://www.sole.leidenuniv.nl

of whether it is interpreted as binary or unary) is inadequate for the characterisation of voicing and that the phenomena traditionally associated with [voice] can better be explained by means of skeletal configurations. The paper is structured as follows. In section 2, I will set the scene for the discussion in the rest of the paper by defining the range of phenomena—phonological and phonetic as well—which are relevant to the representation of voicing. In section 3, I will show that previous analyses of voice phenomena are seriously flawed and that the reason why they fail on both theoretical and empirical grounds is that they assume a featural representation for voicing. In section 4, I will give a brief outline of the main assumptions of CVCV, since the following section draws heavily on concepts advocated in this theoretical framework. Finally, in section 5, I will present my analysis and discuss some of its implications.

2. The preliminaries

The main goal of this section is to define the range of phenomena which are unambiguously related to voicing. First of all, I will provide a list and a short description of those processes that will be referred to as voice phenomena throughout this paper. Then, I will show that there are two different interpretations of voicing in the literature—a broad and a narrow one—and that the narrow one can make better predictions with regard to voice phenomena in different languages. Finally, I will argue that although voice effects are phonetically motivated, they have to be accounted for in the domain of phonology.

2.1. Voice phenomena

As the present paper intends to provide a unified solution to a number of different phenomena related to voicing, I will first have to clarify which phonological events are relevant from the perspective of the analysis to be presented. There are four different patterns that will be referred to under the label 'voice phenomena' in the rest of the paper: VOICE ASSIMILATION, WORD-FINAL DEVOICING, SYLLABLE-FINAL DEVOICING (or CODA DEVOICING) and INTERVOCALIC VOICING.[1] All of these phenomena can be interpreted as cases of neutralisation, although the range of environments in which they occur and their effects are rather diverse. In what follows, I will point out the defining properties of these processes and show how they can interact with each other in different languages.

VOICE ASSIMILATION is a cross-linguistically well-attested process, whereby one of two adjacent obstruents loses its original specification for voice and assumes that of the other one. This can be formulated as the following SPE-type rule:

(1) $[+\text{obstruent}] \rightarrow [\alpha \text{ voice}] / \_\_ \begin{bmatrix} +\text{obstruent} \\ \alpha \text{ voice} \end{bmatrix}$

[1]These patterns may manifest themselves in the form of phonologically conditioned alternations, distributional restrictions (i.e. even though the language does not have the morphological mechanisms which would create the relevant environment for a neutralisation pattern to show itself in the form of an alternation, it cannot maintain a voice contrast in weak positions, e.g. Thai (Lombardi 1995)) and diachronic processes.

It should be pointed out that the above rule exemplifies REGRESSIVE voice assimilation, in which case it is the second obstruent that determines the voicing of the cluster. Although PROGRESSIVE voice assimilation is also attested in a number of languages, it appears that these two phenomena are rather different in nature. Crucially, progressive voice assimilation is (i) much rarer than its regressive counterpart, (ii) very often morphologically conditioned (e.g. the English plural suffix) and (iii) more closely connected to aspiration than to actual voicing (see van Rooy & Wissing 2001:312ff; for a discussion of what is meant by 'actual voicing' see 2.2). A few examples of regressive voice assimilation taken from Hungarian are shown below in (2):

(2)

keré[k]	'wheel'	keré[gb]en	'in the wheel'
ala[p]	'base'	ala[bz]at	'plinth'
veré[b]	'sparrow'	veré[pt]ől	'from the sparrow'
gő[z]	'steam'	gő[sf]ürdő	'steam-bath'

WORD-FINAL DEVOICING is another frequent and well-documented phenomenon connected to voicing. As is suggested by the name of this process, obstruents can never surface as voiced at the end of the word in languages with word-final devoicing. We can express this generalisation as follows:

(3) $[+\text{obstruent}] \rightarrow [-\text{voice}] \,/\, \_\_ \,\#$

Word-final devoicing differs from voice assimilation in two important respects: it is restricted to the end of the word and the affected consonant loses its voice specification entirely—that is to say it does not assimilate to the following consonant. This complete loss of contrasts along a given dimension in a certain position is called REDUCTIVE NEUTRALISATION in Trubetzkoyan terms, as opposed to ASSIMILATIVE NEUTRALISATION, which is exemplified by voice assimilation. Dutch is a typical example of languages which suppress voicing in word-final position:

(4)

hui[z]en	'houses'	hui[s]	'house'
kra[b]en	'to scratch'	kra[p]	'scratch'
lie[v]eling	'(my) love'	lie[f]	'lovely'
le[z]en	'to read'	lee[s]	'I read'

SYLLABLE-FINAL DEVOICING resembles word-final devoicing in quite a number of respects: it is also an example of reductive neutralisation and the environment in which it operates overlaps with the environment of the latter. Actually, the structural description of word-final devoicing is a proper subset of the structural description of syllable-final devoicing, which can immediately be seen if one compares the rule in (3) with the one below:

(5) $[+\text{obstruent}] \rightarrow [-\text{voice}] \,/\, \_\_ \left\{ \begin{array}{c} \text{C} \\ \# \end{array} \right\}$

However, syllable-final devoicing seems to have a number of properties which clearly mark it off from its word-final congener and make it rather similar to progressive voice assimilation. We will return to this issue in section 5.6, where it will be discussed in a more detailed fashion. For the meantime, here are a few examples of this process taken from German:

(6)	run[d]e	'round pl.'	Run[tg]ang	'stroll'
	lö[z]en	'to loosen'	lö[sl]ich	'soluble'
	We[g]e	'way dat.'	We[k]	'way nom.'
	Grä[z]er	'grass pl.'	Gra[s]	'grass sg.'

Finally, INTERVOCALIC VOICING is also one of the processes which affect voicing, although in this case the distinction is neutralised to the seemingly more marked value of the feature [voice]. Languages with intervocalic voicing cannot have plain voiceless[2] obstruents between two vowels:

(7) [+obstruent] → [+voice] / V_V

It is necessary to point out that intervocalic processes whereby voiceless obstruents become sonorants (e.g. flapping in English) do not count as genuine instances of voicing from the point of view of the present paper. The reason for this is that sonorants lack active voicing, which means that they are not specified for the feature responsible for voicing in obstruents.[3] As the emergence of sonorants from underlying voiceless obstruents in intervocalic position does not produce active voicing, the feature [voice] seems not to be involved in this alternation—therefore, such processes are completely irrelevant to the present discussion, since they do not tell us anything about the representation of voicing in obstruents.

One might also wonder whether this process should be understood as an instance of assimilative or reductive neutralisation. Although traditional analyses formalised in the spirit of SPE might claim that the obstruent actually assimilates to the flanking vowels, such an approach seems rather dubious in view of what we have just established about the difference between passive voicing in sonorants (including vowels, of course) and active voicing in obstruents. More specifically, it is impossible to say that the obstruent takes on the voice specification of any of the neighbouring vowels, as these cannot convincingly be argued to be specified for this feature. Hence, one is forced to stipulate that this process is a case of reductive neutralisation. However, it will be shown in section 3.4 that this conclusion is also not plausible in frameworks where voicing is represented by a privative feature (and, more generally, in any framework where voicing in obstruents is considered to be marked *per se*). To conclude the discussion of intervocalic voicing, a few examples from the evolution of Spanish are given below:

[2]It is important to note that voiceless aspirates normally remain unaffected by this process, German and Korean being two prime examples in point (these two languages are usually argued to have a contrast based on aspiration instead of voicing).

[3]For more on passive or spontaneous voicing in sonorants see Harris (1994:135).

(8)

LATIN		SPANISH	
lupu	>	lobo	'wolf'
vi:ta	>	vida	'life'
fi:ku	>	higo	'fig'

It remains to be seen how these phenomena can combine with each other in different languages. First of all, voice assimilation can appear on its own (e.g. in Hungarian) or accompanied by word-final devoicing (e.g. in Dutch). Syllable-final devoicing is obviously almost entirely incompatible with voice assimilation since there is a considerable amount of overlap between their domains (i.e. both affect obstruents in the coda). There is, however, one case where a language with voice assimilation can, in fact, have reductive neutralisation in the coda: when the following consonant is a heterosyllabic sonorant.[4] Since word-final devoicing and syllable-final devoicing seem to be in a subset-superset relationship, it is not surprising that syllable-final devoicing implies word-final devoicing. On the other hand, there is another, much more interesting implicational relationship, stated explicitly by Steriade (1997): 'voicing neutralisation never occurs finally without also occurring in pre-obstruent position' (ibid. 7). This actually means that a language with word-final devoicing must also have either syllable-final devoicing (e.g. German) or voice assimilation (e.g. Russian). Finally, intervocalic voicing may freely combine with each of the processes mentioned above, which is just what one would expect, given that its structural description does not overlap with that of the others at all.

2.2. Two interpretations of voicing

The quotation below is a typical textbook definition of the feature [voice]:

> Voiced/voiceless: [+/−voice]. Sounds produced with vibrations of the vocal cords are voiced; voiceless sounds are produced with a glottal opening so wide that it will prevent vocal vibration if air flows through it. (Durand 1990:54)

This definition clearly exhibits a bias towards articulation as opposed to acoustic characteristics, which is far from surprising given that the original feature system of SPE was based on articulatory notions. Although the representation of laryngeal activity proposed in Halle & Stevens (1971) offers a somewhat more sophisticated solution to the encoding of voicing, the essence of the analysis remains the same: voicing equals vocal cord vibration.[5] The crucial problem with this assumption is that it disregards the fact that the voice value of stop consonants seems to depend more on the relative timing of laryngeal and supralaryngeal events than the presence or absence of phonation *per se*.

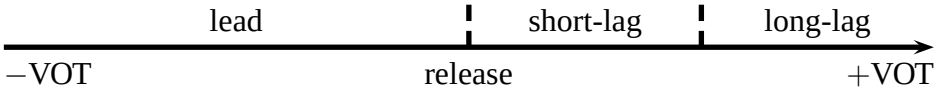
There is, however, an alternative to the interpretation of the feature [voice] which captures the temporeal aspects of voicing better than the traditional articulatory definition: the dimension of

[4] An example in point is Dutch, where syllable-final obstruents cannot surface as voiced before a heterosyllabic sonorant.

[5] For an overview see Durand (1990:54-57).

Voice Onset Time (VOT).[6] According to this assumption, voicing is a function of the amount of time between the onset of voicing and the release of the stop. This can be illustrated as follows:

(9)



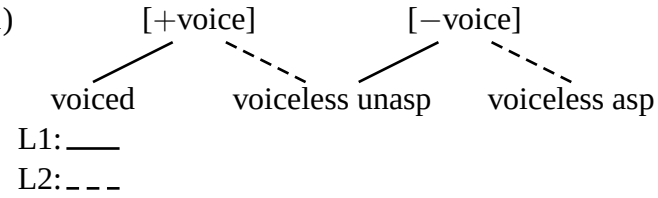
(10)

lead VOT	=	voiced	
short-lag VOT	=	voiceless unaspirated	
long-lag VOT	=	voiceless aspirated	

The chart in (9) shows how the dimension of VOT can be divided into three distinct intervals. If the onset of voicing precedes the release phase of the stop there is closure voicing which makes the stop likely to be perceived as voiced (e.g. [b]). If phonation starts shortly after the release the consonant is perceived as voiceless unaspirated (e.g. [p]). Finally, if there is a considerable lapse between the release and the onset of voicing the consonant is interpreted as voiceless aspriated (e.g. [p^h]). One of the main advantages of this approach is that VOT can be made use of both as the articulatory and the auditory correlate of the underlying feature [voice].

However, there is little agreement on how VOT and the feature [voice] should be matched up. One possible approach is to claim that the underlying feature can correspond to different VOT values in different languages or even in different environments (Keating 1984). This is shown in (11) below:

(11)



According to the diagram above, certain languages like French, Hungarian or Russian (L1) realise [+voice] obstruents as voiced and [−voice] obstruents as plain voiceless, whereas other languages like German, English or Mandarin Chinese (L2) interpret [+voice] as voicelessness without aspiration and [−voice] as voicelessness with aspiration. The crucial insight behind this analysis is that, whatever the actual phonetic interpretation of the underlying feature is, its phonological behaviour will be very similar in all of the languages which possess it. Unfortunately, however, this empirical claim seems to be rather questionable, as will be shown in 5.6.

Others like van Rooy & Wissing (2001), Ringen & Helgason (2004) and Petrova et al. (2006) argue for a narrow interpretation of [voice], according to which only languages with lead VOT values in initial voiced obstruents have the feature [voice]. This suggests that [voice] will be

[6] This interpretation of voice figures prominently in the following works: Lisker & Abramson (1971); Keating (1984); van Rooy & Wissing (2001); de Carvalho (2002); Ringen & Helgason (2004); Petrova et al. (2006); for a somehat more detailed discussion of VOT, see Ladefoged (2001:119-121).

phonologically inert in fortis-lenis systems like German or English. If this prediction is right, these languages will not show the processes characteristic of languages with an active feature [voice] — which may not be such an outlandish claim given that neither of these languages has regressive voice assimilation. In this paper, I will adopt the latter interpretation, since I believe that it mirrors certain finer distinctions between laryngeal features which are lost if one goes for the broad interpretation.

2.3. Voice phenomena – Phonology or Phonetics?

There seems to be wide agreement on the assumption that voice phenomena are phonetically motivated. For instance, voice assimilation easily lends itself to an analysis which makes use of the phonetic notion of COARTICULATION. This has led several researchers to incorporate these insights into their proposed systems as OT constraints such as AGREE (Lombardi 1999; Petrova et al. 2006), thereby giving their analysis a firm grounding in phonetics. However, I believe that this approach is an untenable one for several reasons.

First of all, phonology and phonetics are very often treated as two separate modules with non-overlapping domains (e.g. Prince & Smolensky 1993). It is not clear how phonetics could reach into phonology from down below when it is assumed that independent modules cannot feed information back into other modules which are activated earlier in the course of the derivation (e.g. phonology cannot perform syntactic operations). Therefore, the phonetically grounded approach needs a highly sophisticated and intricate theory of the phonology-phonetics interface, which is, however, usually lacking.[7]

Moreover, the present analysis is to be couched in the framework of strict CV phonology or CVCV, which explicitly rejects the possibility of having an autonomous phonetic interpretation module. Scheer (1998) claims that there are no 'purely phonetic' effects, since there is a biunique relationship between phonological structures and their phonetic manifestation (ibid. 141–142). In effect, this means that no phonological representation can correspond to two different surface forms, and no single surface form can be encoded in two different ways underlyingly, which implies that phonetic implementation is entirely automatic and universal. If the phonetic module plays no distinct role in the derivation, it cannot be argued to provide the motivation for phonological processes or constraints: phonology has to account for everything. Although this model may be somewhat too restricted, the main insight that underlies it seems to be quite appealing: phonological problems should not be ignored simply because they may have something to do with phonetics.

3. Problems with the feature [voice]

In the following section, I will provide a critique of previous analyses of voice phenomena which assume either a unary or a binary feature for the representation of voicing. First, I will show that analysing contrasts along the dimension of voicing as equipollent in nature makes

[7] One exception to this is Steriade's (1997) manuscript, in which she proposes a very explicit model to handle the issue of how phonetics and phonology communicate with each other.

no predictions as to the direction of voice neutralisation. Second, I will try to elucidate the anomalous relationship between voicing on the one hand and manner and place features on the other. Then, I will show that previous accounts of voice assimilation are all arguably arbitrary from a representational point of view. Lastly, the solution to the encoding of voicing chosen in Government Phonology will be examined and shown to be inadequate for several reasons.

3.1. Binary [±voice]

In the descriptive, SPE-type rules of section 2.1, [voice] was treated as a simple binary feature. Although some authors argue that it is indeed necessary to maintain a binary analysis of voicing distinctions (Wetzels & Mascaró 2001), it seems quite clear that a unary solution should be preferred if one wants to make any predictions concerning the naturalness of different voice phenomena. For example, the representation of voicing should make it possible to determine whether coda voicing is more marked than coda devoicing. Unfortunately, a binary feature [±voice] seems to be unable to do this.

Let us see whether the formulations of voice phenomena as given in 2.1, repeated here for convenience' sake, tell us anything about the naturalness of these processes:

(12) $[+\text{obstruent}] \rightarrow [\alpha \text{ voice}] \,/\, \_\_ \left[\begin{array}{l} +\text{obstruent} \\ \alpha \text{ voice} \end{array} \right]$

(13) $[+\text{obstruent}] \rightarrow [+\text{voice}] \,/\, \text{V}\_\_\text{V}$

(14) $[+\text{obstruent}] \rightarrow [-\text{voice}] \,/\, \_\_ \,\#$

(15) $[+\text{obstruent}] \rightarrow [-\text{voice}] \,/\, \_\_ \left\{ \begin{array}{l} \text{C} \\ \# \end{array} \right\}$

The non-arbitrariness of voice assimiliation is mirrored by the fact that the structural description (SD) and the structural change (SC) of (12) is linked by the same feature [α voice], although there is nothing in the SPE system that would formally necessitate such a relationship (i.e. there may be some rules whose SD and SC do not overlap at all). Intervocalic voicing, as formulated in (13), may also be motivated to some extent, given that vowels are argued to be specified as [+voice] in SPE.[8] However, devoicing appears as a completely arbitrary process in (14) and (15): a change from [−voice] to [+voice] would be equally natural in the binary notation, where plus and minus specifications can be switched without restriction. This prediction is obviously wrong, since no language has final voicing, whereas final devoicing is a completely natural and wide-spread process. The unary approach could clearly achieve better results in this case, since the delinking of a feature in final position (final devoicing) is obviously more natural than the

[8] Although it has already been established in 2.1 that we now see a crucial difference between the active voicing of obstruents and the passive voicing present in sonorants.

spontaneous addition of a feature in the same position (final voicing).

3.2. Voicing and other features

The set of distinctive features in generative phonology can be divided into larger classes on the basis of the relations among different features. Generally speaking, there seem to be three major groups, whose existence was already alluded to in SPE:[9] MANNER FEATURES, PLACE FEATURES and LARYNGEAL FEATURES (see also Clements 1985). Szigetvári (to appear) further elaborates this classification by demonstrating that place and laryngeal features show a rather uniform behaviour as opposed to manner features, which suggests that the former two might actually constitute a larger class together.[10] He assigns this distinction a formal status by arguing that the representation of manner contrasts is crucially different from that of place and laryngeal features. In what follows, I will show that although laryngeal features and voicing in particular do pattern with place features in certain respects, it appears that it is necessary to make a representational distinction between these two sets of features too in the spirit of Szigetvári's proposal.

However, before moving on to the discussion of why voicing is different from both place and manner features, I would like to point out the main advantages of expressing differential phonological behaviour by representational means. It seems obvious that distinctive features are no more than descriptive categories: they denote natural classes of speech sounds established on the basis of phonological behaviour and phonetic properties. For example, actively voiced consonants arguably belong to the same set of sounds, which one might label [+voice], since they all instantiate regressive spreading of voicing in a considerable proportion of the languages of the world. Nevertheless, distinctive features alone fail to provide an explanation for the phonological processes in which they are involved. It does not matter whether the class of voiced segments is characterised as [+voice], [−voice], [+red] or [−Alfred the Great]: the feature itself is just a label, which makes no predictions about the behaviour of the class of segments it stands for.

On the other hand, if we give up the assumption that phonetic properties and paradigmatic relations among sounds can only be expressed by distinctive features, we may get one step closer to explanatory adequacy (for a similar approach, see van der Hulst 1996). For instance, Jensen (1994) and Szigetvári (to appear) propose that stopness should be represented by skeletal relations rather than features. Although space restrictions preclude any detailed discussion of their approach, one important aspect of their analyses is worth mentioning: they both explain the fact that manner features do not normally participate in assimilation by pointing out that skeletal relations are incapable of spreading. In this case, the difference between the representation of manner features on the one hand and place and laryngeal features on the other aptly mirrors the differences in the behaviour of these classes of features.

As it has already been mentioned, there are some crucial differences between the behaviour

[9] Although this classification was assigned no formal status there.

[10] In feature geometry, this is usually achieved by linking manner features (or at least some of them) directly to the root node, while laryngeal and place features are grouped together under separate nodes (e.g. Clements & Hume 1995).

of voicing and that of place features. This is well exemplified by the differential treatment of laryngeal features in Szigetvári (to appear) and de Carvalho (2002). The diagram below shows the domain of the three classes of features established above (i.e. whether they can appear only in consonants or in consonants and vowels as well) as is defined in the two different models:

(16)

	Place	Laryngeal	Manner	Def. of voice
Szigetvári (to appear)	*C's & V's*	*C's & V's*	*only C's*	F_0
de Carvalho (2002)	*C's & V's*	*only C's*	*only C's*	VOT

According to Szigetvári (to appear), laryngeal features pattern with place features, since these two classes show a similar affiliation towards consonants and vowels. However, de Carvalho (2002) claims just the opposite: in his view, the class of laryngeal features patterns with that of manner features, as both of them are incompatible with vowels. The reason for this apparent contradiction is that the two authors use different definitions of voicing: Szigetvári assumes that the invariant cue to voicing is a change in fundamental frequency (see also Brockhaus 1995:119-123), whereas Carvalho holds that underlying voice distinctions are cued mainly by VOT values. It is clear that the reason why they can have such largely differing views on the behaviour of voicing is that laryngeal features do not unambiguously pattern with either place or manner features.

It may also be interesting to see how voice phenomena fit into the classification of lenition processes offered by Ségéral & Scheer (1999) and Szigetvári (to appear). Both of these articles point out that there are two different kinds of lenition: C-lenition, which results in the loss of place and laryngeal contrasts in the coda and V-lenition, which brings about a rise in the sonority of intervocalic segments. This is summarised below:

(17) C-lenition: affects place & laryngeal specification _ {C, #}
V-lenition: affects manner specification V_V

As I have shown above, Szigetvári expresses this difference by arguing that manner distinctions are encoded by skeletal relations while place and laryngeal features are represented by phonological primes. The only problem with this approach is that voicing seems to undergo both types of lenition: it is lost in the coda (C-lenition) and it appears spontaneously in intervocalic position (V-lenition) — that is to say it patterns with both place and manner features. This can be seen in (18) below:

(18) loss of voice in _{C, #} → C-lenition (cf. place features)
voicing in V_V → V-lenition (cf. manner features)

This ambiguous behaviour suggests that the representation of voicing may be different from that of either place or manner of articulation.

3.3. Arbitrariness of spreading in voice assimilation

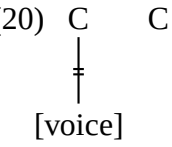
There are quite a number of contrasting accounts of voice assimilation, but they all seem to share one feature: the spreading of voicing is not motivated in the phonological component itself. Although recent analyses are crucially right in that they make it possible to derive certain conditions on spreading from more general principles of the grammar, they are incapable of expressing the naturalness of spreading itself. In the following paragraphs, I will briefly review some of the approaches to voice assimilation and point out their shortcomings.

First of all, it has already been established in 3.1 that binary features do not make any real predictions as to which rules will qualify as natural and which of them will never appear in natural languages. The same goes for SPE-type rules: there is no phonological motivation for their existence, since a rewrite rule is no more than the formal expression of an empirical observation. There is only one respect in which SPE-type rules approach explanatory adequacy: the general rule template of assimilation seems more natural than that of dissimilation, since it contains one symbol less. This is illustrated in (19):

(19) assimilation: $[F] \rightarrow [\alpha F] / \_\_ [\alpha F]$
dissimilation: $[F] \rightarrow [-\alpha F] / \_\_ [\alpha F]$

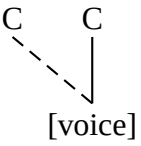
However, the actual value of this observation is rendered rather questionable by the fact the feature-counting evaluation metric of SPE is not an organic part of the phonological theory presented there and does not actively participate in the description of phonological processes (i.e. it will never block rules that contain too many features).

Lombardi (1995) presents a theory of laryngeal neutralisation based on the notions of delinking and spreading. In her view the 'spreading of voicelessness' is a natural consequence of the diminished autosegmental licensing potential of the coda (cf. Goldsmith 1990:123ff): the autosegment responsible for voicing is not licensed in the coda, so it is delinked.[11] The autosegmental chart in (20) below illustrates this:

(20)


However, there is no licensing constraint that would force out the spreading of the feature [voice]. In fact, it seems somewhat counter-intuitive to claim that an inherently weak position like the coda has to be linked to a feature licensed by the following onset. The spreading of voicing is shown below:

[11]This, of course, raises the question of why [voice] does not become delinked in the coda when it is also attached to the onset (i.e. when the feature [voice] is shared by both consonants). The answer to this may be that the feature is actually licensed by the onset, which makes delinking unnecessary.

(21) 

Thus, the spreading of [voice] in standard autosegmental analyses of voice assimilation is no more than a stipulative and parametric rule which fails to capture the near-universal nature of voice assimilation.

An OT analysis of voice phenomena is presented in Lombardi (1999), in which she accounts for assimilation by using the phonetically grounded constraint AGREE, whose status has been discussed in 2.3. Even if we disregard the difficulties that arise when phonetics is permitted to reach into the domain of phonology, it is clear that the option of using constraints which embody generalisations external to phonology is unavailable in a representationally oriented model like CVCV, where phonological processes are the direct result of the interaction of different representational units. Therefore, we can conclude that although the constraint AGREE does serve as the motivation for voice assimilation, the constraint itself is arbitrary, at least from a representational point of view.

Finally, Government Phonology (GP) also cannot account for what leads to the spreading of **L**, the representative of active voicing in this framework. Harris (1994) proposes a definition of spreading which rules that the source of this process has to be a licensor and the target a licensed position (ibid. 167). In effect, this means that GP does predict *where* voice assimilation will take place, but it cannot explain *why* it should happen. This is also reflected by the coda parameter proposed in Brockhaus (1995:136), which is supposed to account for the presence or absence of voice assimilation in different languages:

(22) **Coda Parameter II**
A coda position may share the LARYNGEAL node of its governing onset: ON/OFF

Once again, this parameter does not mirror the unmarked nature of regressive voice assimilation. In fact, the solution of GP is not much different from that of Lombardi (1995) — which is not surprising in view of the fact that both of them rely heavily on the notion of autosegmental licensing.

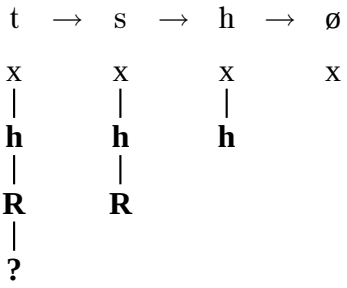
3.4. GP and the element L

Since the present paper intends to provide an alternative to the representation of voice within GP, it is necessary to show how this framework fares with respect to the description of voice phenomena. To begin with, Kaye et al. (1990) propose an individual element for the representation of voicing which they refer to as **L**. Later analyses of voice phenomena (e.g. Harris 1994 and Brockhaus 1995) have all been based on this element, which means that the representation of voicing has not changed much since the birth of GP. However, it will be shown below that the

elemental representation of voicing has a considerable number of drawbacks which are often overlooked in GP analyses, but which should by no means be completely disregarded.

However, before showing why GP cannot come to grips with voice phenomena, I will briefly discuss the definition of markedness offered in Harris (1994) and Harris & Lindsey (1995). Harris claims that phonological COMPLEXITY can be directly read off from GP representations, by counting the number of elements that are involved in the composition of a given speech sound, which implies that markedness correlates with elemental complexity in this particular framework. An important consequence of this is that lenition can now be made sense of as elemental simplification, which manifests itself in the form of delinking. For example, the lenition trajectory from [t] to nothing through [s] and [h] can be illustrated as follows (Harris 1994:124):

(23)



As can be seen in (23), at each consecutive step of the lenition trajectory one of the elements is lost, until the skeletal slot becomes devoid of all melodic content (the exact identity of the elements involved in this particular case of lenition need not concern us here).

It has been shown in 2.1 that devoicing and intervocalic voicing can and should be interpreted as cases of reductive neutralisation, which basically corresponds to lenition in the present context; the question is whether the notation introduced above can account for both of these types of voice neutralisation. Devoicing and intervocalic voicing allow us to make the following generalisations: voicelessness is unmarked in the coda (cf. C-lenition) and voicing is unmarked intervocalically (cf. V-lenition). This suggests that the analysis of devoicing as elemental loss is crucially right. However, intervocalic voicing cannot be analysed along the same lines, since the obstruent targeted by this process seems to acquire an additional prime (**L**) instead of losing one, which clearly contradicts Harris' definition of lenition. The only way out of this situation would be to stipulate that intervocalic voicing is realised by a completely different mechanism and it does not involve element addition. This, however, would leave us with two distinct representations for the same phonetic property, which is clearly an undesirable state of affairs (cf. 2.3).

Another problem with the spontaneous appearance of **L** in intervocalic position is that it is not clear where this element comes from.[12] This may not be so embarrassing for a theory which allows the insertion of default feature values under certain conditions, but GP explicitly rejects this possibility. The following quotation from Kaye et al. (1990) is very revealing as to the range of possible phonological processes in GP: 'there is a direct relation between the phonological process and the context in which it occurs' (ibid. 194), that is to say the only operations allowed

[12]Obviously enough, it cannot originate in the neighbouring vowels, since they do not have active voicing.

in this framework are spreading and delinking. Elements emerging out of nothing are not countenanced by GP, which suggests that intervocalic voicing may need a different interpretation in this framework.

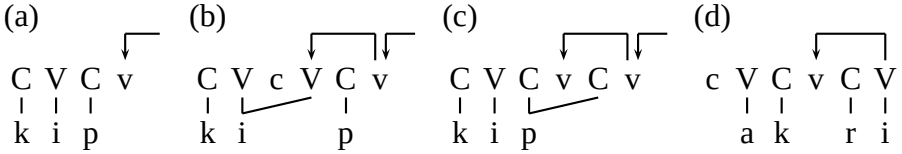
Lastly, the behaviour of **L** also seems to be in sharp contradiction with the AUTONOMOUS INTERPRETATION HYPOTHESIS of GP (Harris & Lindsey 1995). The point of this principle is that the primes of phonological representations should all enjoy 'stand-alone phonetic interpretability' (Harris & Lindsey 1995:34). Both Szigetvári (1996) and de Carvalho (2002) point out that **L** has no such individual realisation in GP (as opposed to **A**, for instance, which is realised as [a] when it stands on its own).

I believe that all of the problems outlined above may be solved if we give up the assumption that voicing is represented by an element. A claim like this, of course, has a considerable number of consequences, which should all be tested so as to see whether a solution along these lines is superior to previous analyses. This task will be undertaken in section 5 — but before that, we need to see what the most important assertions made in the framework of CVCV are, since the actual analysis will take these for granted.

4. What is CVCV?

In this section I will very briefly review the main assumptions of CVCV, a relatively recent upshot of standard GP (Kaye et al. 1985, 1990; Harris 1994) based on Jean Lowenstamm's seminal article (Lowenstamm 1996) which carries some ideas already present in standard GP to the extreme. All representations in CVCV are entirely lateral in the sense that 'phonological constituent structure [is] reduced to a strict sequence of non-branching Onsets and non-branching Nuclei' (Scheer 2004:40). By way of illustration, let us see how the hypothetical words *kip, kiip, kipp* and *akri* can be represented (lowercase **v** represents a mute vowel and lowercase **c** a mute consonant):

(24)



The reader will have noticed that the representation of long vowels and consonant clusters requires the postulation of empty positions and additional syllables (or rather CV units), which might be seen as a serious shortcoming of the theory. However, empty positions do not come for free in this framework: they must be subject to GOVERNMENT, a certain type of skeletal relation between two V slots (this is indicated by the arrows above the skeleton).[13] For further discussion of these issues, I refer the reader to Scheer (2004).

[13] Apparently, empty C slots do not need government to be able surface as mute; Szigetvári (to appear) claims that the reason for this is that the defining property of consonants is muteness, as opposed to vowels, which are characterised by high sonority. Therefore, it is natural for a consonant to surface as mute without the involvement of any extra mechanism, whereas vowels will always need government—which is a disruptive force by definition—to be muted.

There is one further property of CVCV and Government Phonology in general which I need to mention here: the outstanding role that they assign to representation in phonology. GP seems to be strongly committed to the idea that explanatory adequacy can only be achieved through representational means. There are no rules or constraints in GP, which means that problems that are solved procedurally elsewhere often receive a representational reinterpretation in this framework. As the scope of this paper does not allow us to delve any deeper into this issue, I refer the interested reader to an article in which this line of reasoning can easily be detected: Lowenstamm (1999).

5. Voicing and the Skeleton

In the remainder of the paper I will present my alternative to the representation of voicing and see how it can cope with the voice phenomena described in 2.1. The main assumptions of the analysis will be outlined in 5.1 and further elaborated in 5.3. In 5.2 and 5.3, I will test my hypothesis against the empirical observations made in 2.1 and show that it can give a unified explanation for the characteristic properties of voice phenomena. In 5.4 and 5.5, I will examine some of the more questionable details of the analysis and provide a number of further arguments in support of the solution advocated here. In 5.6, I will show that this model makes a number of empirical predictions which have received little or no attention in the literature, but which appear to be crucially right.

5.1. The proposal

As may be seen from the discussion above, the solution to the representation of voicing has to be:

(i) representational
(ii) economical
(iii) phonetically motivated
(iv) empirically motivated.

Being exclusively representational it should not make use of any rules or arbitrary constraints. It has to be economical, that is it cannot give a disjunctive account of voice phenomena (e.g. we cannot assume that voice assimilation is a phonological process and intervocalic voicing a phonetic one[14]). Furthermore, it has to be demonstrated that it mirrors the actual phonetic reality of voicing (phonetic motivation). Finally, it cannot contradict the empirical observations which relate to voicing (empirical motivation).

With these provisos in mind, I now turn to the actual analysis of the representation of voicing. I propose that the phonological identity of voicing is a specific SKELETAL CONFIGURATION:

[14] As is done in Szigetvári (1996), for instance.

(25) **Voicing in obstruents** (first definition)
(a) obstruents are voiced between two ungoverned V slots
(b) obstruents are voiceless elsewhere.

In the first place, it should be noted that this generalisation holds only at the level of abstract phonological structure as conceived of in CVCV, which means that it is not necessarily surface true in a strict phonetic sense. The importance of this reservation is easier to understand if we consider the relationship between CV phonological representations and their phonetic realisation. For instance, a word with the surface pattern CVCCV (e.g. [tɔrkɔ]) is represented as CVCvCV (where the lowercase **v** once again denotes a governed vowel which does not surface phonetically) which means that all of its consonants except for the first one are flanked by vowels on both sides.[15] Therefore, the crucial difference between voiced and voiceless obstruents as defined in (25) cannot be that the former are in intervocalic position as opposed to the latter, which are not: all consonants are enclosed between vowels in CVCV. What will really matter is whether these vowels are governed or not. We will see in 4.3 that this assumption has an important bearing on the encoding of voicing in consonant clusters.

5.2. Intervocalic voicing and word-final devoicing

The immediate advantage of this analysis is that it makes it unnecessary to burden the grammar with any special provisions in order to account for intervocalic voicing and word-final devoicing. In (26) below we can see the representation of a voiced obstruent in intervocalic position:

(26)

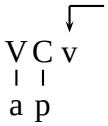

It should be remarked that an underlying structure like (26) will never contain a voiceless consonant on the surface. The reason for this is that the particular skeletal configuration it illustrates is directly responsible for voicing: the consonant is between two ungoverned vowels. If any of the vowels were subject to government the consonant would not be in intervocalic position anymore and would surface as voiceless. Thus, one of the main benefits of this analysis is that it establishes a non-arbitrary formal relationship between the environment of intervocalic voicing and the process itself.

Word-final devoicing also gains a natural interpretation if we assume that voicing is encoded along the lines of (25). An important assumption of GP (dating back to Kaye 1990) is that words that are consonant-final on the surface end in an empty nucleus which has to be muted by means of a special type of government.[16] As final empty nuclei (FEN) are supposedly always subject to government, word-final obstruents are predicted never to surface as voiced. This is shown in (27):

[15] The status of word-initial consonants in this model will be discussed in 4.4.

[16] For a discussion of the theoretical status and relevance of final empty nuclei, see Scheer (2004, 2006:103ff).

(27)



Clearly enough, the predictions of the present model with regard to the voice phenomena discussed in this section seem to be somewhat too strong. After all, both intervocalic voiceless obstruents and word-final voiced obstruents are amply attested in the languages of the world. If, however, we disregard these seeming contradictions for the moment we may make a very important observation in connection with the representations in (26) and (27): they aptly mirror the fact that voicelessness is marked intervocalically and voicing word-finally.

It seems to me that there is something inherently appealing in this way of approaching markedness and neutralisation. As it has already been mentioned, a representation which mirrors the behaviour of different phonological units must be seen as superior to a representation which serves only to distinguish between natural classes. This implies that unmarked phenomena should have simpler representations than marked ones (this is the main insight underlying Harris' conception of complexity). Therefore, the foundation stone of any analysis which involves markedness considerations should be the unmarked case; everything else has to be built on top of that. Marked configurations have to be incorporated into the representation by means of additional mechanisms. In this particular case, this means that we first have to account for why voicing is marked in one position and unmarked in another. It is only after this that we can turn to structures that violate the generalisations about the positional markedness of voicing. Of course, the model should not be allowed to undergenerate either—in 4.4–4.6 I will show that the present model lives up to this expectation. But before that, I will have to attempt to find a solution to another important—and apparently unmarked—facet of voicing, namely voice assimilation.

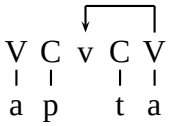
5.3. Voice assimilation

It is not immediately obvious how the present model could handle voice assimilation given that it excludes the possibility of analysing assimilation as spreading, which clearly contradicts our intuitions. This seems to be highly problematic indeed in view of the fact that this process is no less widespread than intervocalic voicing or word-final devoicing. In order to see how this contradiction can be resolved, we first have to consider the status of consonant clusters and word-internal empty nuclei in CVCV.[17]

In CVCV there is no constituent corresponding to the traditional coda, which means that heterosyllabic consonant clusters have to be represented with a governed empty nucleus sitting in the middle. An example of this can be seen in (28):

[17] In this paper, I will only consider coda-onset clusters, since GP does not allow obstruent clusters to be dominated by a single constituent. However, it should be noted that voice agreement restrictions also seem to hold inside branching constituents (Kehrein & Golston 2004).

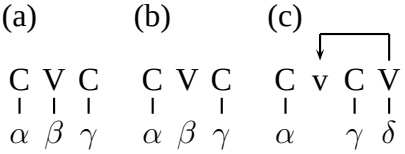
(28)



On the basis of this, one would predict that consonant clusters must be uniformly voiceless, since superficially adjacent consonants are always separated by a governed V slot. This prediction is clearly wrong, which means that either (25) has to be modified or we have to reinterpret the relationship between empty nuclei and government. Since I believe that the explanatory power of the present model would be significantly reduced if we made any changes to the definition of voicing in (25), I will opt for the latter solution.

It should be clear from the preceding discussion that there are at least two different kinds of nuclei in CVCV: governed empty nuclei, which are never realised phonetically and ungoverned nuclei, which always make it to the surface. Actually, there is also a third type which we have not mentioned yet: nuclei with alternating vowels, which can be realised phonetically but may also be muted by government. Scheer (2004) argues that this three-way classification of nuclei is rooted in the lexicon. In the case of full vowels (i.e. vowels that are never governed) the melody is underlyingly associated to the V-slot, while alternating vowels have floating melodic elements which may be associated to the skeleton provided that the vowel remains ungoverned. Non-alternating governed nuclei (i.e. those that never surface; for example, the vowel between the two consonants in (27)) have no melodic content at all. This is summarised in (29), where (a) is a full vowel, (b) an alternating one and (c) a mute one:

(29) (a) (b) (c)



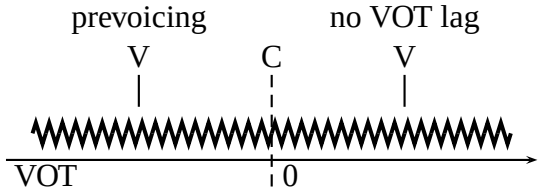
There is, however, an important question in connection with this classification: how can we ensure that a mute vowel trapped between two consonants (as in [apta]) will always be governed? Obviously, the only way to solve this problem is to stipulate that an empty vocalic position always has to be subject to government. This requirement has no grounding in the phonological component: there is no theory-internal reason for assuming that governed empty nuclei are less marked than ungoverned ones. In fact, empty nuclei did have a phonetic manifestation in standard GP: they could surface as schwa or as [ɨ], the cold vowel of Kaye et al. (1985). Apparently, the only reason for the existence of this stipulation is that there seems to be no other way to incorporate the traditional coda into this framework.

Therefore, I will assume that there is no such restriction and that a word-internal empty nucleus may be either governed or ungoverned. This means that this kind of government has to be specified in the lexicon—which is clearly not such an unorthodox view given that Kaye et al. (1990) originally specified government as a relation 'defined at the level of lexical representation' (ibid. 221). Now, if the empty nucleus can remain ungoverned it has to have an individual phonetic manifestation. Once again, there is some reason to believe that this observation is crucially right, since certain aspects of the actual phonetic realisation of a speech sound depend

on whether it is associated with an onset or a nucleus (e.g. the elements **U** and **I** are realised as glides when dominated by an onset, and as vowels when dominated by a nucleus). In this respect, the V-slot itself (henceforth **V**) seems to behave rather like an element. The question is what the actual phonetic realisation of **V** is. Clearly enough, the phonetic event corresponding to an empty nucleus has to be common to all vowels, since **V** is present in every vocalic expression. One such defining property of vowels is periodicity, which makes it reasonable to suppose that the phonetic identity of **V** is PERIODIC SOURCE or PHONATION. This conception of the central characteristic of nuclei is not entirely new to the field of phonology: the component |V| in Dependency Phonology was defined as 'relatively periodic' (van der Hulst 1996:328), which clearly resembles our formulation of **V**.

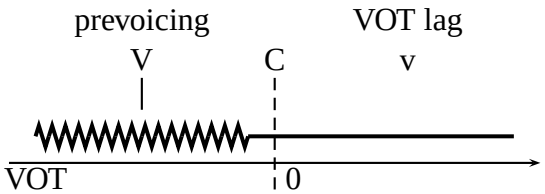
At this point we may return to our original topic, that is the representation of voicing in obstruents. The advantages of interpreting phonation as the phonetic manifestation of **V** will become more obvious if we consider the possibility of relating the presence or absence of a phonetically realised nucleus to VOT values.[18] The following schematic diagram serves to illustrate how such a relationship should be conceived of in the case of an intervocalic voiced obstruent:

(30) VOICED OBSTRUENT (as in [aba]):



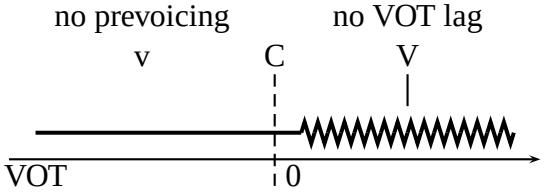
In (30) both nuclei are realised phonetically which results in the obstruent being voiced throughout the closure and the release phase. It is important to see that voicing does not originate in the consonant itself: it comes from the neighbouring vowels, which is in line with the assumption presented in (25). The next two diagrams illustrate two voiceless consonants, one of which is followed and the other preceded by a mute empty nucleus:

(31) VOICELESS OBSTRUENT I (as in [ap]):



[18] As it has already been mentioned, VOT values may be taken to be the invariant cues to voicing. However, it should be remarked that there are quite a number of additional phonetic events which may also help to cue voicing; for an overview of the cues to voicing, see Raphael (2005:189-193) and Raphael et al. (2003:173).

(32) VOICELESS OBSTRUENT II (as in [pa]):



Again, what we see is that the voicelessness of these consonants has nothing to do with their internal featural content. Instead, voicing appears to be a function of their vocalic context: if the following vowel is mute there is a VOT lag, and if the preceding vowel is mute there can be no prevoicing. Both the absence of prevoicing and a lag in the onset of voicing exclude the possibility of interpreting the consonant as voiced. Of course, the main advantage of this analysis is that there is a very obvious connection between the diagrams in (30)–(32) and the definition of voicing in (25). Thus, we can see that there is indeed a direct relationship between the phonological representation and VOT values, which is summed up below:

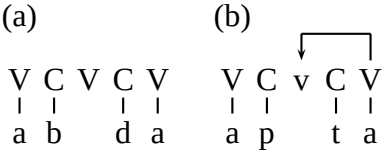
(33)

	PREVOICING	VOT LAG	VOICING
V C V	+	–	voiced
v C V	–	–	voiceless
V C v	+	+	voiceless

Let us briefly summarise what we have seen so far in this section. I have shown that word-internal empty nuclei need not always be governed. Moreover, I have identified phonation as the acoustic and articulatory correlate of **V** and the stand-alone phonetic interpretation of ungoverned empty nuclei. Finally, I have demonstrated that there is a close relationship between the presence or absence of phonetically realised nuclei on the one hand and VOT values on the other. In the remainder of this section I will show how these three seemingly unconnected observations may help us understand how voice assimilation can be analysed in the model advocated here.

The chunks of phonological representations in (34) illustrate two structures in which the consonants agree in voicing:

(34) (a) (b)



In (34a) both consonants surface as voiced, since they are both flanked by ungoverned vowels, which means that they satisfy (25a). On the other hand, both consonants are voiceless in (34b), since the nucleus that sits between them is governed by the vowel following the second consonant.

It is important to see that this representation incorporates all the observations made earlier in this section. First of all, the two structures differ only in that the second has lexically specified government, which the first one obviously lacks. Then, the vowel in the middle of (34a) is realised phonetically as phonation, and thus serves to maintain the voicing in both obstruents. Finally, the diagram in (30) can be taken to illustrate both of the obstruents in (34a), while (31) corresponds to the first obstruent in (34b) and (32) to the second one.

The representation of voice assimilation outlined above has two further important aspects, which should be discussed here. First, it seems that two adjacent obstruents can never have different voice values, since the voicing of both consonants is determined by whether the vowel in the middle surfaces phonetically or not. It is impossible to envisage a situation in which they could contrast with each other in terms of voicing, since the representation clearly excludes this possibility. Second, the voicing of the cluster depends on whether the vowel after the second obstruent is a governor or not, which follows from the fact that the direction of government is always right to left in CVCV. Therefore, the predominantly regressive nature of voice assimilation can actually be derived from a more general principle of CVCV.

The careful reader will have noticed that the predictions about the universality and the direction of voicing seem to be contradicted by a considerable number of languages. For instance, German obstruent clusters may differ along the dimension of voicing (e.g. Run[tg]ang) and Swedish has bidirectional voice assimilation. In 5.6 I will show that there is a possible explanation for the existence of such seeming counterexamples. For the meantime, I would like to quote the title of one of my sources (van Rooy & Wissing 2001), which may help to convince the reader about the plausibility of the predictions of my model: 'Distinctive [voice] implies regressive voicing assimilation'.[19]

Finally, let us take a look at how the new insights gained in this section may help us in rephrasing the definition of voicing in (25) in a formally more adequate way. As it has already been shown, the voicing of obstruent clusters is determined by whether the vowel following the second consonant is a governor or not. This is the observation that serves as the basis of the following definition:

(35) **Lexical Governor**
A vowel specified as a governor in the lexical representation

The definition of voicing can thus be reinterpreted as follows:

(36) **Voicing in $\mathbf{C_1V_1}$**
(a) C_1 is voiced if V_1 is not a lexical governor (*abda*-type)
(b) C_1 is voiceless if V_1 is a lexical governor (*apta*-type)

[19] Although it should be noted that there exists another article (Ringen & Helgason 2004) which bears the title 'Distinctive [voice] does not imply regressive voicing assimilation', which makes the previous argument rather less forceful.

5.4. Word-initial and word-final voiced obstruents

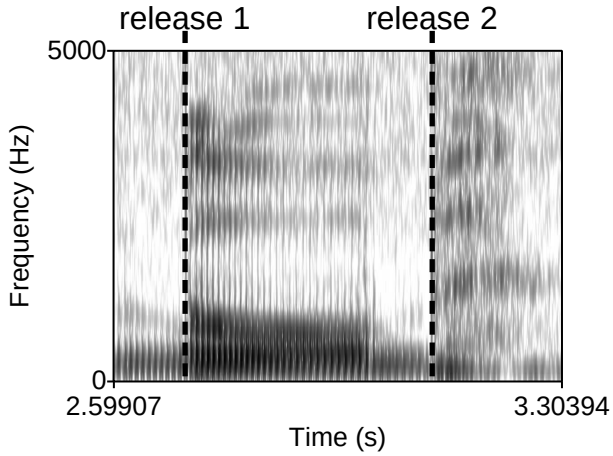
The representation of word-initial and word-final voiced obstruents is rather troublesome from the perspective of the present analysis, since neither of these consonants is situated between two ungoverned vowels. The only way out of this situation is to posit that word-initial consonants are actually preceded by an empty nucleus which may be ungoverned, and that word-final consonants may be followed by an ungoverned FEN. Although this step may give the impression of being completely arbitrary, it is actually far less so than one would expect.

In the first place, both final empty nuclei and initial empty CV slots have been argued for on independent grounds by Kaye (1990) and Lowenstamm (1999) respectively. Therefore, it seems natural that the points made with regard to the possibility of having ungoverned empty nuclei should be extended to the beginning and the end of the word as well. The following two diagrams illustrate the representations I propose for voiced consonants in initial and final position:

(37) (a) (b)

Moreover, there are a number of indications that the present approach is far closer to the actual phonetic data than its apparent abstractness would suggest. Provided that the phonetic manifestation of empty nuclei is phonation, we can interpret the initial nucleus in (37a) as prevoicing and the final nucleus in (37b) as the kind of audible voiced release that can be heard at the end of utterance-final words after voiced obstruents in French or Hungarian for example. This is also supported by spectrographical evidence from Hungarian:

(38)



The spectrogram of the Hungarian word [buːb] shows two things very clearly: there is a rel-

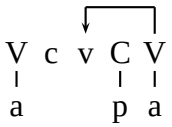
atively long period of phonation before the release of the first consonant and after that of the second one. The presence of periodic source before the word-initial consonant and after the word-final one correlates with the presence of an ungoverned empty nucleus in the representation.

5.5. Voiceless obstruents in intervocalic position

Languages that tolerate voiceless obstruents in intervocalic position also seem to contradict the predictions of the present model. However, I believe that the analysis pursued here is essentially right, since it correctly accounts for the fact that voiceless consonants are marked intervocalically by showing that we need some kind of additional mechanism in order to be able to represent this kind of configuration.

One possible solution to this problem is to claim that voiceless consonants that do not undergo intervocalic voicing have the following rather geminate-like structure:

(39)



However, it has to be admitted that this is just a provisional account — it remains to be seen whether the predictions made by this particular representation are correct or not.

5.6. Coda devoicing

In 5.3, I noted in passing that German final devoicing may be difficult to implement in the present model, since the particular solution to the representation of voicing adopted in this paper does not allow adjacent obstruents to be have different voice specifications. The same objection holds for any language with syllable-final devoicing. As the model has to be able to account for all the phenomena associated with voicing, we have to briefly examine the validity of this counterargument.

The fact that there can be no voiced obstruents in the coda is problematic only inasmuch as we can prove that the language in question has a contrast based on voicing (cf. the narrow interpretation of voicing in 2.2). However, German apparently has no real voicing contrast: all obstruents in this language are either plain or aspirated (although the plain series may be subject to intervocalic voicing). Similarly, Korean and Maidu—two languages which are argued to have syllable-final devoicing in Lombardi (1995) — do not have underlyingly voiced consonants (Korean: plain–aspirated–ejective; Maidu: plain–implosive–glottalised). Although this by no means proves that the model is right, it does not disprove it either. Therefore, the only way to decide whether syllable-final devoicing really threatens to invalidate the present analysis would be to examine all the languages which are reported to have syllable-final devoicing on a case by

case basis and see whether any of them has a real voiced series of obstruents.[20]

6. Conclusion

Let me conclude my analysis with a brief summary of its main findings. First of all, I have given a preliminary classification of voice phenomena, which has served as a basis for all consequent investigations. Then, I have shown that previous accounts of processes related to voicing are incapable of capturing the unmarkedness of voice phenomena, and that the main reason for this is their commitment to feature-based analyses. Finally, I have provided an analysis of voice phenomena based on skeletal configurations, and demonstrated the potential of this approach.

Acknowledgements

My thanks are due to Ádám Nádasdy for his valuable comments on earlier versions of this paper, and for all the help I have received from him as a fledgling phonologist; to Péter Szigetvári and Szilárd Szentgyörgyi for their suggestions and encouragement; and to Anna Csíky for all her support and help with the manuscript.

Márton Sóskuthy
Eötvös Loránd University (ELTE), Budapest
mrhowardppa@yahoo.com

References

Brockhaus, W. (1995). *Final Devoicing in the Phonology of German*. Niemeyer, Tübingen.
de Carvalho, J. B. (2002). What are phonological syllables made of? the voice/length symmetry. Durand, J. & B. Laks (eds.), *Phonetics, Phonology, and Cognition*, Oxford University Press, Oxford, pp. 51–79.
Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2, pp. 225–252.
Clements, G. N. & E. Hume (1995). The internal organization of speech sounds. Goldsmith, J. A. (ed.), *Handbook of Phonological Theory*, Blackwell, Oxford & Cambridge, Mass., pp. 245–306.
Durand, J. (1990). *Generative and Non-Linear Phonology*. Longman, London & New York.
Goldsmith, J. A. (1990). *Autosegmental and Metrical Phonology*. Blackwell, Oxford & Cambridge, Mass.
Halle, M. & K. N. Stevens (1971). A note on laryngeal features. *MIT Quarterly Progress Report* 101, pp. 198–213.
Harris, J. (1994). *English Sound Structure*. Blackwell, Oxford & Cambridge, Mass.
Harris, J. & G. Lindsey (1995). The elements of phonological representation. Durand, J. & F. Katamba (eds.), *Frontiers of Phonology: Atoms, Structures, Derivations*, Longman, Harlow, Essex, pp. 34–79.
van der Hulst, H. (1996). The segmentsyllable connection. Durand, J. & B. Laks (eds.), *Current Trends in Phonology*, ESRI, Salford, Manchester, pp. 326–354.
Jensen, S. (1994). Is ? an element? towards a non-segmental phonology. *SOAS Working Papers in Linguistics and Phonetics* 4, pp. 71–78.

[20] It should be noted that this does not apply to word-final devoicing: there are indeed languages with a real voicing contrast, which cannot have voiced obstruents word-finally. This is also reflected by the representation, which rules out syllable-final devoicing but has nothing against word-final devoicing (cf. 5.2).

Kaye, J. (1990). 'Coda' licensing. *Phonology* 7, pp. 301–330.
Kaye, J., J. Lowenstamm & J.-R. Vergnaud (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* 2, pp. 305–328.
Kaye, J., J. Lowenstamm & J.-R. Vergnaud (1990). Constituent structure and government in phonology. *Phonology* 7, pp. 193–232.
Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language* 60, pp. 286–319.
Kehrein, W. & C. Golston (2004). A prosodic theory of laryngeal contrasts. *Phonology* 21, pp. 1–33.
Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Languages*. Blackwell, Malden, MA & Oxford.
Lisker, L. & A. S. Abramson (1971). Distinctive features and laryngeal control. *Language* 47, pp. 767–785.
Lombardi, L. (1995). Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory* 13, pp. 39–74.
Lombardi, L. (1999). Positional faithfulness and voicing assimilation in optimality theory. *Natural Language and Linguistic Theory* 17, pp. 267–302.
Lowenstamm, J. (1996). CV as the only syllable type. Durand, J. & B. Laks (eds.), *Current Trends in Phonology*, ESRI, Salford, Manchester, pp. 419–441.
Lowenstamm, J. (1999). The beginning of the word. Rennison, J. & K. Kühnhammer (eds.), *Phonologica 1996: Syllables?!*, Thesus, The Hague, pp. 153–166.
Petrova, O., R. Plapp, C. Ringen & S. Szentgyörgyi (2006). Voice and aspiration: evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review* 23, pp. 1–35.
Prince, A. & P. Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar. Manuscript, Rutgers University and University of Colorado at Boulder. Available at ROA.
Raphael, L. J. (2005). Acoustic cues to the perception of segmental phonemes. Pisoni, D. B. & R. E. Remez (eds.), *The Handbook of Speech Perception*, Blackwell, Oxford, pp. 182–206.
Raphael, L. J., G. J. Borden & K. S. Harris (2003). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Lippincott Williams & Wilkins, Baltimore, 4 edn.
Ringen, C. & P. Helgason (2004). Distinctive [voice] does not imply regressive voicing assimilation: evidence from Swedish. *International Journal of English Studies* 2, pp. 53–71.
van Rooy, B. & D. Wissing (2001). Distinctive [voice] implies regressive voicing assimilation. Hall, T. A. (ed.), *Distinctive Feature Theory*, Mouton de Gruyter, Berlin, pp. 295–334.
Scheer, T. (1998). La structure interne des consonnes. Sauzet, P. (ed.), *Langues et Grammaire II-III, Phonologie*, Université Paris 8, Paris, pp. 140–172.
Scheer, T. (2004). *A lateral theory of phonology. Vol 1: What is CVCV, and why should it be?* Mouton de Gruyter, Berlin.
Scheer, T. (2006). How non-phonological information is processed in phonology. Handout for the 2006 EGG Summer School held in Olomouc, Czech Republic.
Ségéral, P. & T. Scheer (1999). The Coda Mirror. Manuscript, Université de Paris 7 and Université de Nice.
Steriade, D. (1997). Phonetics in phonology: the case of laryngeal neutralization. Manuscript, UCLA.
Szigetvári, P. (1996). Laryngeal contrasts and problematic representations. *The Even Yearbook* 2, pp. 97–110.
Szigetvári, P. (to appear). Manner as a skeletal relation. Neubarth, F., M. Pöchtrager & J. Rennison (eds.), *Phonologica 2002: Structure and Melody*, Mouton de Gruyter, Berlin.
Wetzels, W. L. & J. Mascaró (2001). The typology of voicing and devoicing. *Language* 77, pp. 207–244.