June 1997

Phonetics in Phonology: The Case of Laryngeal Neutralization

Donca Steriade, *UCLA*

steriade@humnet.ucla.edu

0. Introduction

The study identifies the factors responsible for the loss of laryngeal contrasts and the reflexes of these factors in individual grammars. The main result reported is that the site of laryngeal neutralization can be uniformly identified by reference to phonetic implementation factors. Many of these factors are perceptual: laryngeal categories are neutralized in positions where the *cues* to the relevant contrast would be diminished or obtainable only at the cost of additional articulatory maneuvers. Conversely, laryngeal contrasts are permitted (or *licensed*) in positions that are high on a scale of perceptibility. It is argued here that the main factor involved in neutralization and licensing is the distribution of cues to the relevant contrasts. This hypothesis, referred to as Licensing by Cue, is compared here to the idea of Licensing by Prosody (Ito 1986, 1989, Goldsmith 1990, Rubach 1990, Lombardi 1991, 1995) according to which the distribution of features in general - and of laryngeal features in particular - is controlled by their prosodic position. The general idea pursued here is that phonological grammars incorporate knowledge of the conditions under which feature contrasts are physically implemented[1]. The focus in this study is on the empirical evidence supporting such a view.

0. 1. Licensing: by cue or by prosody

An example that clarifies the difference between Licensing by Cue and Licensing by Prosody is that of voicing neutralization in word-final and pre-obstruent position. Consider one such case:

(1) One voicing neutralization pattern (Polish, Lithuanian, Slavic, Sanskrit)

a. Obstruents are distinctively voiced or voiceless before vocoids and consonantal sonorants.
b. Obstruents are neutralized (devoiced) word finally.
c. Obstruents are neutralized before any obstruent: they surface assimilated in voicing to the following obstruent.

In these languages, obstruents followed by vowels or consonantal sonorants are frequently located in onset - e.g. *a.ba, a.bra, a.pa, a.pra* - and thus it is tempting to characterize the position of licensing in (1.a) as the onset and the positions of neutralization (1.b-c) as the coda. The grammatical statements in (2) - representatives of Licensing-by-Prosody thinking - reflect this postulated correlation between syllabically defined positions and sites of licensing or neutralization:

[1]See Ohala 1983 and Westbury and Keating 1985 for explorations of the link betwen neutralization and articulatory difficulty. Kingston (1985, 1990) has drawn our attention to the phonological consequences of perceptual factors in the analysis of laryngeal features. The present study continues Kingston's line of work and focusses more narrowly on the grammatical description of the link between phonetic implementation and contrast maintenance.

(2) a. [Voice] is unlicensed in the coda, licensed in onset. (Goldsmith 1990, Rubach 1990)
b. [Voice] is licensed in a segment by a following *tautosyllabic* sonorant. (Lombardi 1995)

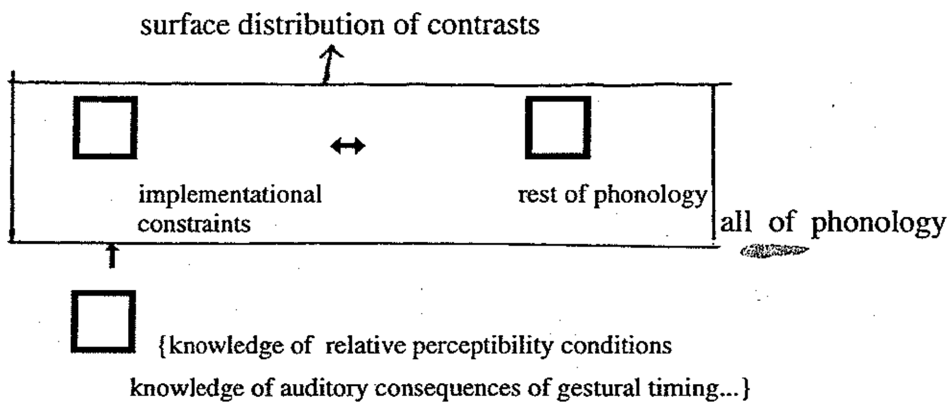
The pattern (1) is open however to a different interpretation: at least one of the major cues to the distinction between voiced and voiceless obstruents is the voice onset time (VOT) value observable on a following segment (Lisker 1957 ; Lisker and Abramson 1964; Keating 1984). Different VOT values - indicating different [voice] categories in the preceding obstruent - can be observed on a following vowel or sonorant but not on obstruents. Therefore pre-obstruent obstruents necessarily lack at least this one bit of information about their laryngeal category. In word-final position the situation is comparable: simplifying a bit, we can identify the word final site with the utterance final position. Clearly here too a distinctively voiced or voiceless obstruent will necessarily lack its VOT cue. The suggestion pursued in this study is that absence of a major cue - or articulatory difficulties in implementing it - represent the main factor responsible for this and other types of neutralization. Unlike the statements in (2), this line of analysis promises to explain the grammar of neutralization, by showing how independently known facts about the perception and production of speech interact with grammatical conditions to yield sound patterns.

0.2. Phonetics in phonology: the downward arrow and alternatives

In flow-chart synopses of grammatical organization, the phonological component is frequently depicted as linked by a downward pointing arrow to a level of phonetic representation, the latter to be fed to a component of phonetic implementation (e.g. Kenstowicz and Kisseberth 1978:7, fig. 1.1; Mohanan's 1995:27 reconstruction of the view presented in the Sound Pattern of English). Although the specifics of such flow charts are seldom made precise, the downward arrow from phonology to phonetics seems to mean this: the phonological component consists of various entities and conditions (the feature set, the OCP, sonority sequencing conditions, the crossing line condition, etc.) whose interaction determines which contrasts a language will have and where. The phonetic implementation component contains laws that map phonological representations onto articulatory instructions, and laws that compute the acoustic and perceptual consequences of articulatory gestures. The downward arrow connecting phonology to phonetics means that the decision to have a contrast and have it in a specific position is taken in phonology. It cannot be affected by "downstream factors", i.e. by physical conditions under which the contrast will be implemented. Phonetic implementation has to live with prior decisions taken in the phonology.

The view presented here is that phonological patterns can be understood only in the context of a different relation between grammar and implementation. The diagram in (3) illustrates this: I assume that speakers possess knowledge of the relative ease with which different types of featural contrasts can be implemented. For instance, it seems reasonable to attribute to speakers awareness of the fact that a *k/g* contrast is more easily detectable in intervocalic position than in inter-obstruent position. Similarly, that the same *k/g* contrast *can* in fact be conveyed in inter-obstruent position (e.g. as in *askta* vs. *asgta*), but only at the cost of additional articulatory effort. Knowledge of this sort enters the grammar in the form of implementational constraints. The interaction of these conditions with the rest of grammar determines whether the language maintains a given contrast in a given position. A similar conception is presented in studies by Flemming (1995), Jun (1995), Kirchner (1997) and Silverman (1995). The important issue of projecting phonetically based constraints from observed data is discussed by Hayes (1996).

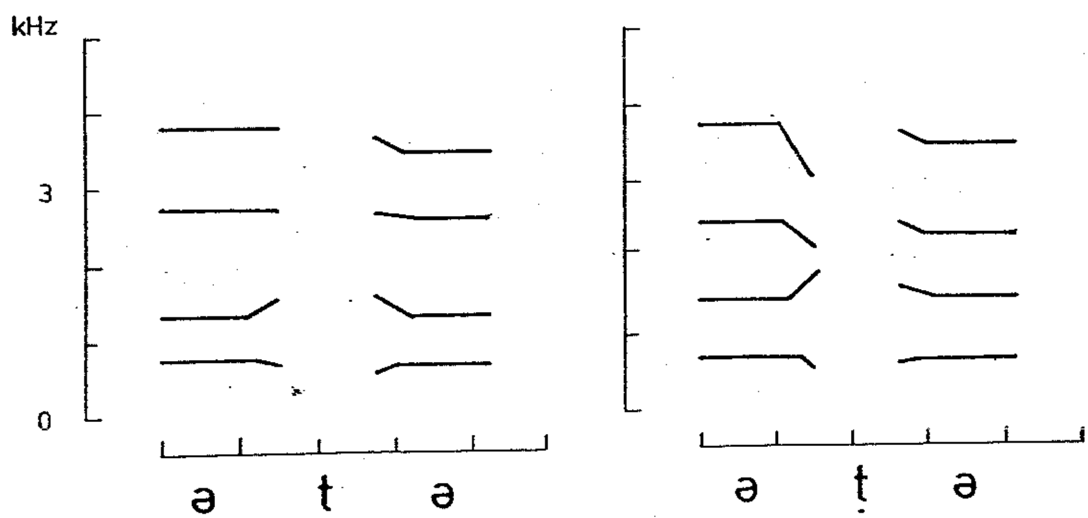
(3)



0.3. An example of cue licensing: retroflexion

Some reason to explore the view in (3) can be provided by juxtaposing a phonetic fact and a phonological observation. The phonetic fact can be inferred from a pair of stylized spectrograms (formant transitions into and out of the apical stops of Gujarati: Dave 1977): observe that only the V-C transitions differentiate *t* and *ṭ*. The C-V transitions are essentially identical in the two cases.

(4) Stylized spectrograms of Gujarati apical (retroflex and alveolar) stops (Dave 1977: 11)



The phonological fact is an implicational law emerging from an extensive survey of apical systems (Steriade 1995; see also Hamilton 1996) : if a language neutralizes the contrast between alveolars and retroflexes then it does so first in contexts where the helpful V-C transitions are missing, i.e. in stops that are either word-initial or post-consonantal. Conversely, if a language does allow the *t/ṭ* contrast, it allows it in postvocalic position. Frequently, this is the only position where such a contrast is permitted. We can attach a causal interpretation to the connection between this phonological generalization and the representative spectrograms in (3): in contexts where the retroflex-alveolar contrast is hard to perceive, it is categorically suppressed, *because* it would be difficult to implement there.

Note that this case, unlike that of voicing neutralization in (1), is unambiguous in regards to the role of prosody in neutralization: there is clearly no connection between the syllabic or word-position of the apical and its ability to carry distinctive retroflexion. Neutralized word initial or postconsonantal apicals are onsets but then so is the distinctively retroflex intervocalic *ṭ* in *VṭV* sequences. In this case then, there is a directly observable connection between the distribution of cues to a contrast, the phonetic implementation fact, and the phonological distribution of the contrast. We shall see that exactly the same connection can be uncovered in the case of the laryngeal features.

In the case of retroflexion, the implementational constraint is the ban on inter-apical contrast in contexts lacking V-C transitions. The implementational constraint reflects directly knowledge of the conditions of physical implementation of the contrast: in fact the knowledge and the constraint are not easily separable and may turn out to be identical.

To the extent that implementational facts are constant cross-linguistically, the typology of neutralization will possess certain invariant properties, such as the implicational law on apical neutralization mentioned above. Markedness theory is then, in part, the study of such constant implementational factors. To the extent however that the conditions of phonetic implementation differ from language to language - or from feature to feature - the facts of neutralization will differ too, at least at the observational level. Thus the optimal contexts for the perception of the *t/ṭ* contrast are not necessarily the optimal contexts for the perception of the *t/d* or *t/tʃ* contrast. This difference in the perceptibility of contrasts across contexts can be tied to language-specific or feature-specific differences in implementation. To understand markedness and phonological typology one must understand the implementational conditions that shape individual systems of contrast.

0.4 Cues

We consider now the facts that stand behind statements of relative perceptibility. The example considered is that of voicing.

One can classify the three contexts mentioned in (1) according to the acoustic correlates to voicing available in each one of them. At least some of these acoustic properties have been shown to influence the perception of voicing categories (Raphael 1981) and thus are cues to voicing; others are potential cues. The classification in (5) below characterizes the distribution of cues that *would* obtain if distinctive voiced and voiceless obstruents occurred in all three contexts in (1). The point is to show that even if the voicing contrast had been maintained word finally and before obstruents - as it is for instance in certain lexical classes in English (cf. *mob* vs. *mop*; *mo[b]ster* vs. *qui[p]ster*) - it would nonetheless be harder to reliably identify there. I propose then a correlation between positions of poor perceptibility and sites of neutralization. I add to the contexts mentioned earlier a few others, which reinforce this correlation. Cues to voicing other than the VOT and their distribution are discussed, among others, by Wang (1959), Summerfield and Haggard (1977), Wolf (1978), Barry (1979), Repp (1979), Lisker (1986), Raphael (1981), Port and Dalby (1982), Westbury and Keating (1986), Hillenbrand et al. (1992), Kingston and Diehl (1994, 1995).

(5) Hypothesized distribution of cues to the [voice] category of a stop depending on context :
(distinctively voiced and voiceless C's are assumed to occur in all contexts listed).

i. <u>possible cues to voicing for C after V and before sonorant</u>: e.g. *abra, aba, apra, apa*

closure voicing, closure duration; V_1 duration; F_0 and F_1 values in V_1; burst duration and amplitude; VOT value; F_0 and F_1 values at the onset of voicing in V_2.

ii. <u>possible cues to voicing word initially or after an obstruent and before a sonorant</u>:
e.g. *bra, ba, pra, pa*; and *asbra, asba, aspra, aspa*

closure voicing, closure duration (for post C obstruents only); burst duration and amplitude; VOT value; F_0 and F_1 values at the onset of voicing in the following V.

iii. <u>possible cues to voicing for C after V at end of the word</u>: e.g. *ab, ap*

closure voicing, closure duration; V duration; F_0 and F_1 values in V ; burst duration and amplitude.

iv. possible cues to voicing for C after V and before obstruent: e.g. ***absa, apsa***

closure voicing, closure duration; V_1 duration; F_0 and F_1 values in V_1.

v. possible cues to voicing for C between obstruents: e.g. ***asbta, aspta***

closure voicing, closure duration.

vi. possible cues to voicing for C after an obstruent at the end of the word: ***asb, asp***

closure voicing, closure duration.

vii. possible cues to voicing for C before obstruent word initially: ***bsa, psa***

closure voicing, closure duration.

The reader will observe that as we go down the list of contexts in (5), the set of typically available cues to voicing progressively diminish. The positions where the identification of voicing categories emerges as the most difficult (5.iv - vii) are in fact positions where such contrasts have seldom been documented (Greenberg 1978:253): in particular, the cases in (3.v-vii) are highly significant because they involve obstruent clusters that are rather well attested, yet only one language - Khasi (Henderson 1976, Nagaraja 1985) - is known to allow distinctively voiced obstruents in sequences like *bsa*. The typical absence of distinctive voicing in these positions has been the subject of separate stipulation in recent work (Cho 1990; Lombardi 1991, 1995). The alternative view presented here is that a single factor - relative poverty of cues - induces neutralization in all the contexts listed in (5.ii-vii): the difference between contexts is not of kind but of degree.

Consider now the somewhat more informative context in (5.iii): V_Obstruent. In this position the voicing of an obstruent can in principle be identified more reliably on the basis of the duration and F_0-F_1 values of the preceding vowel: indeed a small number of languages do maintain a voicing contrast morpheme-internally in this position. Among them are Maithili (Yadav 1984), Lamani (Trail 1979), Shilha (Applegate 1958) and various Arabic dialects (Syrian: Cowell 1964, Eastern Arabic: Rice and Sa'id 1979, Moroccan: Harrell 1962, Heath 1987, Iraqi: Abeer Alwan p.c.). These languages do not preserve the voicing contrast in the #_Obstruent, Obstruent _ #, or in inter-obstruent contexts (corresponding to (3.iv-vi) but they do maintain it when the obstruent is either left or right adjacent to a vowel.

An even more favorable environment for voicing identification is that of postvocalic, word final stops (e.g. *mob* vs. *mop*): final stops possess all cues to voicing that pre-obstruent stops do, plus a longer preceding vowel and the higher probability of an audible burst, whose amplitude and duration may be an additional voicing cue (Raphael 1981). Any bursts that pre-obstruent stops may have will possess significantly less acoustic salience than word-final ones (Henderson and Repp 1982) and may therefore be counted as unlikely cues to voicing. Related to this is the fact that voicing neutralization never occurs finally without also occurring in pre-obstruent position. In addition, the two contexts (before _# and before an obstruent) differ as follows: the perception of voicing in a sequence of intervocalic obstruents VO_1O_2V (O= obstruent) is likely to be influenced by the strong cues to voicing present in O_2. It appears that, all else equal, the cues present in the burst and C-V transitions have primacy over those carried by the V-C transition (cf. Ohala 1990 for place, Raphael 1981 and Slis 1986 for voicing): therefore the categorization of O_1 with respect to voice is likely to be influenced by that of O_2, the better cued obstruent in the cluster. In contrast, a word final obstruent can be identified as voiced or voiceless without comparable interference. This too contributes to explaining why neutralization in the _# context is less likely than - and therefore implies - neutralization in the _O context.

If the facts reviewed so far bear on the incidence of voicing neutralization, as claimed here, we expect that the voicing contrast will be maintained in some context as a direct function of the cues available there: all else equal, the better the cue package, the greater the likelihood of contrast preservation. This type of link between the relative likelihood of F neutralization and the relative perceptibility of F in a given context will be documented here and extended to contexts and features not yet discussed; the evidence will also show that the sites of neutralization have no uniform characterization in terms of prosodic (esp. syllabic) organization. For the moment, I provide initial evidence for correlation claimed using the data in (4), which illustrates the range of contexts in which voicing neutralization is attested.

(6) **Patterns of [voice] neutralization** (O = obstruent, R = sonorant, incl. vowel)

fewer cues <--> **more cues**

	#_ O, O_# e.g. **bsa** vs. **psa**	R_O e.g. **absa** vs. **apsa**	R_# e.g. **ab** vs. **ap**	_R e.g. **ba** vs. **pa**	R_R e.g. **aba** vs. **apa**
Totontepec Mixe (Crawford 1964)	no voice contrast	no voice contrast	no voice contrast	no voice contrast	**contrast**
Lithuanian (Senn 1966)	no voice contrast	no voice contrast	no voice contrast	**contrast**	**contrast**
French (Dell 1995)	no voice contrast	no voice contrast	**contrast**	**contrast**	**contrast**
Shilha (Applegate 1958)	no voice contrast	**contrast**	**contrast**	**contrast**	**contrast**
Khasi (Nagaraja 1985)	**contrast**	(sequence missing)	**contrast**	**contrast**	**contrast**

As usual, the significant part about a chart like (6) lies in the missing patterns: no language surveyed maintains the voicing contrast in a less informative context, unless it also does so in the more informative contexts identified in (5). Thus, using T as a symbol for a voice-neutralized obstruent, no system known to me neutralizes word finally after a vowel without also neutralizing medially in the V_obstruent context.

0.5. Cue weighting

I have described the difference between the contexts in (4) in terms of more vs. fewer cues to voicing. But one may also compare the cues themselves in terms of their quality. For this comparison, we adopt Wolff's (1978) distinction between voicing cues clustered at the onset of voicing (*onset cues*) and cues clustered at the offset of voicing (*offset cues*). Thus in a V_1-O-V_2 sequence, the onset cues involve the transition between the obstruent and V_2, while the offset cues involve the transition between V_1 and the obstruent. We will refer globally to onset and offser cues as *transitional* or *contextual* cues, since they are scattered over the external context in which the consonant occurs. A third type of cue - voicing or lack of it during closure - will be referred to as an *internal cue*, since it resides during the period of oral constriction of the consonant. Several studies of voicing (Raphael 1981, Slis 1986; and data in Duez 1995) suggest (a) that the onset cues have primacy over offset cues, in the sense that they may determine the categorization of the segment in the presence of conflicting information and (b) onset cues may have primacy over the combination of offset and internal cues. Slis (1986) shows that Dutch speakers listening to obstruent clusters differing in voice (e.g. [kd]) perceived more frequently regressively assimilated (e.g. [gd]) than progressively assimilated or unassimilated clusters. In this instance of perceptual assimilation, categorization of the cluster's voicing was more frequently determined by the onset cues of [d], which were able to override the offset cues of [k] and [k']s lack of closure voicing, the internal cue. Comparably, Raphael's (1981) results for English show that when the obstruent contains conflicting cues to voicing, the onset cues dominate the percept. These findings correlate clearly with the fact that the most common environments of voicing neutralization (_# and _O) share the absence of onset cues. The significant fact is that, in contexts where reliable onset cues like VOT exist, the absence of other voicing cues - such as V1 duration, closure duration, or F1, F0 values on V1 - is mostly irrelevant. Thus word initial prevocalic stops are seldom voice-neutralized: that's because the presence of onset cues may compensate for and outweigh the lack of the offset cues. We will assume then that an analysis based on cue licensing will have to incorporate a cue weighting mechanism.

We will also observe that, although infrequent, voice neutralization *is* attested in the #_V context: this relates to the fact that the offset cues (V1 duration, F0, F1 values in V1) are necessarily absent there. As table (6) indicates, however, neutralization in the #_V context occurs only in the languages that neutralize everywhere else, save possibly in the most informative V_V context.

0.6. Cue duration

A further point that will be developed here is that the relative duration of the string over which transitional cues are manifested plays a role in neutralization: we will compare the likelihood of neutralization in sequences where the obstruent is adjacent to a relatively long modal-voiced sonorant with sequences in which the obstruent is adjacent to a very short modal-voiced sonorant: e.g. [litra] vs. [litr]. The data suggests that cue duration also plays a role in identifying neutralization contexts.

0.7. The descriptive system

Before proceeding we must consider briefly the grammatical questions raised by the hypothesis of Licensing by Cue. The simplest implementation of this idea is to characterize in standard segmental terms the contexts where contrastive voicing is more or less likely to be identified. Such descriptions have been used in (6). Based on this list of contexts, a perceptibility scale for voicing may be postulated: this is a series of statements about the relative perceptibility of the voicing contrast depending on context. The sign ▶ used in (7) indicates that voicing in one context is more perceptible than in the context listed to its right. The scale is partial, since not every conceivable context appears on it. We will expand the scale as the evidence is presented.

(7) Scale of obstruent voicing perceptibility according to context

V_[+son] ► V_# ► V_[-son] ► {[-son] __[-son], [-son] __#, #__[-son] }

A central analytical move in this study is the assumption that this and other perceptibility scales project families of corresponding constraints. Corresponding to the scale in (7) we have a family of *voice constraints in (8): there is a constraint of the form *[α voice]/X_Y corresponding to every context or set of contexts occupying a distinct position on the perceptibility scale. The constraints are universally ranked in the order of inverse perceptibility: the lower the context is on the perceptibility scale, the higher ranked the corresponding *[α voice]/X_Y constraint:

(8) Constraints on the distribution of voicing: ranking is universally fixed by alignment to the perceptibility scale in (5).

(i) *αvoice / [-son] __[-son], [-son] __#, #__[-son]
↓
(ii) *αvoice/ V_ [-son]
↓
(iii) *αvoice/ V_ #
↓
(iv) *αvoice/ V_ [+son]

The constraints in (8) represent the speaker's knowledge of the fact that voicing distinctions are harder to implement in certain contexts than in others. In this sense then, the scale in (7) and the constraint family in (8) are two facets of the same thing: (7) is a statement of the perceptibility facts related to voicing, whereas (8) is a model of the speaker's knowledge of these facts. Both (7) and (8) have a large speculative component, since our understanding of both actual perceptibility and of its mental representation is imperfect: but it is clear that at least scales like (7) are empirically verifiable, independently of their use in explaining neutralization patterns. In this sense, the approach to phonology pursued here is deductive (Lindblom 1990): the contents of the grammar are deduced from knowledge of the conditions in which speech is perceived and produced, to the extent that such knowledge is attributable to naive speakers.

One antecedent of the ranking scheme in (8) is Prince and Smolensky's (1993: 135) idea of aligning constraint hierarchies to harmonic scales. The notions of ranking, evaluation and related concepts in the formalization of constraint-based analyses are also adopted from Prince and Smolensky's work (cf. also McCarthy and Prince 1993, 1994 for further developments). Constraint rankings are indicated by the sign >> or, as in (6), by downward arrows (↓) : the upper constraint is more highly ranked, meaning that it will determine which alternative realization of the same input string (which *candidate*) is more highly valued in a given grammar. Lower ranked constraints determine the outcome of such comparisons only when the higher constraints are moot or violated equally by some candidate pair.

It is fundamental in understanding what follows to bear in mind that the perceptibility scale (7) projects the hierarchy in (8). By this I mean that the constraints have no independent status from the scale: if the scale changes, the constraints change correspondingly. This follows from our view that the scale represents facts about perceptibility, while the constraints represent knowledge of these facts. This conception can be verified: we will observe that the perceptibility of laryngeal distinctions depends on inter-gestural timing and the magnitude of glottal gestures, factors which vary from language to language and from context to context. When the oral-glottal timing changes, the ranking of contexts on the perceptibility scale changes too: for instance a preaspirated stop is more perceptible in post-vocalic position, regardless of what follows, whereas a postaspirated stop is more perceptible in pre-vocalic position, regardless of what precedes. Thus the relative ranking of V_ and _V contexts on a perceptibility scale depends on how glottal abduction is timed relative to oral closure in a stop. The result then is that the same laryngeal feature (here aspiration) may be subject to different constraint hierarchies in different languages, precisely because its perceptibility conditions, and therefore the grammatical constraints reflecting them, change when oral-glottal timing changes.

To return to [voice] neutralization, this phenomenon will be modelled as the interaction between faithfulness to input voice values - the *Preserve [voice]* constraint below - and a fixed hierarchy of *voice constraints aligned to the voice perceptibility scale. Some relevant ranking options are shown below. The top constraint appears to be undominated in the three cases shown.

(9)

*voice / [-son] __[-son], [-son] __#, #__[-son]

(i) voice licensed before sonorants	(ii) voice licensed before sonorants and word finally	(iii) voice licensed after V and before sonorants
↙	↓	↘
*voice/ V_ [-son]	*voice/ V_ [-son]	**Preserve [voice]**
↓	↓	↓
*voice/ V_ #	**Preserve [voice]**	*voice/ V_ [-son]
↓	↓	↓
Preserve [voice]	*voice/ V_ #	*voice/ V_ #
↓	↓	↓
*voice/ V_ [+son]	*voice/ V_ [+son]	*voice/ V_ [+son]

An analysis equivalent to (9.i) will be justified for Lithuanian and a number of other Indo-European languages in section 1.1. The hierarchy in (9.ii) is appropriate for Hungarian and Kolami, as seen in section 1.3. The case of (9.iii) is that of the Arabic dialects mentioned earlier, where no voicing neutralization obtains in the usual "coda" contexts. Observe that the fixed hierarchy in (8) precludes the existence of grammars in which voicing is neutralized finally (V_#) but not before obstruents (V_[-son]) and more generally grammars in which voicing is licensed in a less informative context than the ones where it is neutralized. This and other implicational predictions of the analysis appear to be borne out: for instance, all languages where voicing is neutralized word finally also neutralize it before obstruents, initially before obstruents and in inter-obstruent position.

The type of analysis sketched in (8) and (9) can obviously be generalized: for any given feature F, the contexts where F might in principle occur can be arrayed on a perceptibility scale, in which contexts containing more and/or clearer cues to F will rank higher. The typology of neutralization for F can then be modelled by simply referring to the *F constraint family projected by F's perceptibility scale. Whether this is in fact the right way to model both the typology of positional neutralization and its description in individual languages is the subject of a larger investigation, of which the present study is one part.

0.8. Excessive variability

Do neutralization patterns change with speed and style? They may well change in the sense that certain distinctions may be completely abandoned at faster speeds and in more casual registers. We could tell that this is so by observing that relevant gestures are not being performed at all at certain speech rates/registers. Before systematic work testing this has been carried out, it would be premature to exclude the possibility of rate-dependent neutralization[2].

But does a cue-based approach to neutralization predict an unrealistic amount of variability in the realization of phonological contrasts? For instance, what clearly does not happen is that when we slow down considerably - at an unnaturally slow rate or in unnaturally hyperarticulated speech - no new contrasts emerge[3]. No phonemic contrast between *s* and *z* will emerge in extra-slow and careful speech in English inter-obstruent positions (e.g. *ekstra* vs. *ekztra*) even though by slowing down we may provide two essential conditions for the detection of the *s/z* contrast, namely duration and lack of overlap.

[2]On the fact that certain contrasts emerge only in careful speech and are either imperceptible or perhaps not even attempted in hypo-articulated or fast speech, there is quite a bit of anecdotal evidence. For instance Shipley (1956: 236) notes that in rapid Maidu speech glottalized stops become so weakly glottalized that "the aspirated and glottalized series fall together to some extent". He goes on to note that the merging is incomplete, but that "only a practiced Maidu ear can clearly distinguish a glottalized from an unglottalized stop in an allegro utterance". Newman (1944: p) complains about Yokuts that most of the time the difference between glottalized and unglottalized consonants is imperceptible and that it emerges only in slow, careful speech. As already mentioned, these informal observations may reflect genuine rate-dependent neutralizations: but we cannot tell in the absence of articulatory data.

[3]Paul Smolensky first raised this worry with me (1993, p.c.).

This observation reflects a fact about language acquisition as well as a fact about the structure of adult grammars. There are standard speaking rates and styles and we learn the contrasts of our language at these rates and styles. We can obviously slow down beyond this standard, but since this is not the tempo at which we have been exposed to language, no new contrasts will emerge. Thus what must be built into any theory of phonology is the understanding that the contrasts of the language will be learned based on a limited range of rates and speech styles.

The conjecture about adult grammars that emerges from this discussion is that the effect of implementational factors on the system of contrasts is computed *relative to fixed standards of speech rate and degree of hyperarticulation* and then extended to other rates and styles through the effect of paradigmatic correspondence conditions (Burzio 1995, Flemming 1995, Kenstowicz 1995, Steriade 1995, 1996). I assume that the fixed standard corresponds to slow and careful speech, but this assumption is not essential for our purposes. What is essential is the existence of *some* standard, correspondence to which has the effect of limiting the degree of variation in realizing a contrast.

An illustration of this idea involves the Russian voicing distributions. The obstruents of this language occur as distinctively voiced or voiceless when, *in careful speech* , they are followed by vowels or sonorant-vowel sequences. Absence of such a right-hand context in the careful pronunciation results in voice neutralization: *goro[t]*, but *gorod-a* 'town'. In faster speech, vowels reduce and sometimes disappear completely, without however affecting the distribution of voicing: thus loss or compression of the medial stressless a in *sapag-á* 'boot' does not necessarily induce neutralization of the voiceless quality of [p]. I attribute this fact to the effect of constraints that require featural correspondence between the standard rate and style and all other rates and styles, including ultra-fast or hyper-careful speech[4].

0.9. Extensions

0.9.1 Direct reference to cues?

There are multiple reasons to view analyses like (8)-(9) as only first approximations. Revisions and extensions to other laryngeal features are discussed in later sections. This study begins by

[4]It is important to note that conditions inducing correspondence to a standard rate or style are required independently of one's views on the relation between phonology and phonetic implementation. The original observation establishing this is due to Mohanan (1986), who notes that some processes are immune to pause insertion between their participant segments: *divinity* undergoes Trisyllabic Shortening regardless of whether one inserts a hesitation pause between *divine* and *ity*. This shortening process, which can be viewed as foot optimization (Prince 1990), is immune to tempo- or pause-induced variation because, we argue, it is made invariable by reference to a standard pronunciation in which the [vini] substring is indeed a foot. In the realm of fast speech processes comparable correspondence effects are observed in English (Manuel et al. 1992), Korean (Jun and Beckman 1990) and French (Fougeron and Steriade 1997, Steriade 1996): the fast speech pronunciation of a word maintains selected articulatory properties, by correspondence to the careful speech variant.

using the schema in (9) as a preliminary means of demonstrating the empirical interest of cue-based analyses. I sketch now what will be lacking in such analyses so as to anticipate the direction of the revisions to come. First, the characterization of the link between perceptibility and neutralization given in (8)-(9) is very indirect: if it's lack of cues that causes neutralization, then one may wish to consider grammatical analyses where the cues themselves play an overt role, for instance by being referred to directly in constraints such as **[voice]/ if cue x is missing*. The possibility of direct reference to cues is discussed further in Part II.

0.9.2. Intersegmental timing

Second, the presence of cues to any feature F in some context depends frequently on the degree of overlap between segments carrying F and their neighbors. For instance, English pre-obstruent stops typically lack acoustically salient bursts because the canonical degree of overlap between adjacent consonantal gestures is quite extensive in this language (Browman and Goldstein 1992). This may explain the significant limitations on the composition of English obstruent clusters (Lamontagne 1993; Pierrehumbert 1994). In languages without significant interconsonantal overlap, contrasts cued by burst quality may be safer (Browman and Goldstein, 1992:176, and Lamontagne 1993). This conjecture can be verified by observing the lack of laryngeal neutralization in many Northwest American Indian languages, where adjacent consonants are impressionistically described as non-overlapped (Hoard 1978, Urbanczyk 1995, 1996). For instance, Lushootseed (Urbanczyk 1995, 1996), Twana (Drachman 1969) and Bella Coola (Nater 1984) maintain an ejection contrast in final, pre-obstruent and inter-obstruent position - in addition to all the more favorable contexts - presumably *because* all stops are audibly released in all positions in these languages. Since an audible release can be guaranteed only under certain timing conditions, it appears that the characterization of these Salish systems must refer explicitly to intersegmental timing patterns prevalent in the language. This then is another reason to view analyses like (9) as incomplete, since the connection of neutralization to gestural timing is not being explicitly modelled.

0.9.3. Intrasegmental timing

A further aspect in need of revision relates to the a different timing issue: the same pair of intrasegmental gestures, when differently timed, generate different cue packages. This can be observed by considering two ways of timing aspiration to a stop's oral constriction. The peak of glottal abduction may lead the onset of the oral closure, as the gestural score in (8.a) indicates, or else the abduction peak may align to the oral release, as shown in (8.b). Aspiration is cued, among other things, by its effect on the voice onset or offset of a neighboring sonorant: the diagrams in (8) show that when the timing relations change, the context carrying these transitional cues changes as well.

(10) a. Peak of laryngeal gesture timed to onset of oral constriction: e.g. ^{h}t

```
                    [-------glottal abduction-----]
                         [-----oral closure---release ----]
context cues for         ↑
laryngeal feature here
```

b. Peak of laryngeal gesture timed to release of oral constriction: e.g. t^{h}

```
                         [--------glottal abduction--------]
              [----oral closure ---------release-]
                                      context cues for     ↑
                                        laryngeal feature here
```

We will observe in Part II that neutralization sites for glottalization and aspiration are essentially those lacking contextual cues: pre-aspirated and pre-glottalized consonants neutralize, if at all, in the absence of a *preceding* vowel or sonorant, whereas post-aspirated consonants and ejectives typically neutralize in the absence of a *following* vowel or sonorant. This general observation supports the proposed connection between cues and neutralization. But, once again, in order to turn this into a prediction of the model, we will have to factor in some reference to timing relations.

0.9.4. Variable timing

A last observation related to timing is that the languages surveyed exercise three options when faced with a conflict between preferred timing patterns and unfavorable contexts. An underlying postaspirated stop - the preferred timing pattern for aspiration in obstruents - may happen to occur word-finally, a context where the critical VOT cue will normally be unvailable. The conflict in such a case will be between enforcing the generally prevailing timing pattern - preserving the timing in (10.b) - vs. generating some other transitional cue to aspiration. The three options in this respect are: (a) to keep the timing pattern of (10.b) and rely on impoverished cues to postaspiration (Maithili: Yadav 1984; Bengali dialects: Kenstowicz 1993); (b) to modify the timing to (10.a), and thus generate other contextual cues to aspiration (Icelandic: Thráinsson 1978; other cases discussed in Steriade 1993); and (c) to do neither but rather neutralize the aspiration contrast word finally (Greek, Sanskrit: below section 2.1). The existence of option (b) - contextually variable timing - supports the idea that the grammar is responsive to the range of cues being generated in different positions and with different timing options (see also Silverman 1997).

With these provisos, I set out to establish the first step in the argument: namely that syllable structure does not begin to describe, let alone explain, the patterns of laryngeal neutralization. This is the main object of Part I.

Part I: Against syllable-based accounts of neutralization

I document now the fact that syllable position provides neither a necessary nor an adequate characterization of the sites of laryngeal neutralization. The focus is first on showing that many classic patterns of devoicing and de-aspiration operating in the contexts _ # and _ [-son] cannot be given a syllabic analysis. Second, I show that even ambiguous instances of neutralization - which could be described as coda devoicing/deaspiration - receive a better treatment under the assumption that syllables are irrelevant here. Finally we will verify that the perceptually more impoverished contexts are always the first to induce neutralization.

1.1. 1 Lithuanian

Unlike better studied modern European languages, Lithuanian consonant clusters are heterosyllabic regardless of composition (e.g. *áuk.le*), yet the context of neutralization is identical to that observed in German or Russian: distinctive voicing is preserved before sonorants, lost elsewhere. My sources on Lithuanian are Senn 1966, Augustaitis 1964, the collective *Lietuvių kalbos gramatika* (vol.1: Fonetika ir morfologija) Vilnius 1965, and Dambriunas, Klimas and Schmalstieg 1966. Lithuanian voiced stops are fully voiced; voiceless stops are unaspirated (Senn 1966:67).

(11) a. **Lithuanian obstruent phoneme inventory**

p	t	č	k
b	d	ǰ	g
f	s	š	
v	z	ž	

Loss of distinctive voicing occurs before obstruents and word-finally[5].

(12) **Distribution of voicing in Lithuanian obstruents**

Acute, grave and circumflex accents indicate rime length and pitch accents: á = HL on V:, à = H, ã: = LH on V:.

b. Distinctive voicing preserved before sonorants

	Voiceless		Voiced	
i.	á**ukl**e 'governness'	nu**kn**iaũti 'klauen'	au**gl**ingas 'fruitful'	dre**gn**a 'feucht'
ii.	vikrùs 'geschickt'	cy**pl**ys '??'	**edr**us 'glutton'	žvìr**bl**is 'sparrow'
iii.	te**šm**uõ 'Euter'	a**sm**uõ 'Schneide'	zie**žm**uõ (place name)	ba**žn**ýčia 'church'

---

[5] This is described by Senn and Augustaitis (1964) as neutralizing (Senn 1966:66 "stimmhaft wird stimmlos", where *stimmlos* is the term describing the non-neutral voiceless series. The term for non-neutralizingly devoiced is *entsonorisiert*.)

iv.	sil**pn**as 'weak'	ry**tm**etỹs 'morning'	sko**bn**is 'table'	bā**dm**etys 'year of famine'
v.	**akm**uõ 'stone'	à**tm**inti 'to remember'	au**gm**uõ 'growth'	liū**dn**as 'sad'

c. Voicing neutralized word finally

daũ**g** [dau**k**] 'much'	kà**d** [ka**t**] 'that'

d. Voicing neutralized before obstruents

i.	**at-g**al̃ **[dg]** 'back'	mè**s-d**avau **[zd]**	dìr**b-ti [pt]** 'work-inf'	dè**g-ti [kt]** 'burn-inf'
ii.	míela**s** **d**raũgas **[zd]** 'dear friend'		daũ**g p**inigũ **[kp]** 'much money'	

Are there connections between syllable structure and the voicing neutralization? All Lithuanian grammars report that VCCV sequences are divided as VC.CV. A minor exception is the *Lietuvių kalbos gramatika* (LKG), which mentions variation in the assignment of s-stop clusters: *a-sta ~ as-ta.* The following quote from Senn is representative of the other sources: "Wenn zwischen zwei Vokalen oder Diphthongen zwei oder mehr Konsonanten sind [...] so wird nur der letzte zur folgenen Silbe gezogen; z.B.: *áuk-le, bars-tý-ti, ga-nyk-là, giñk-las, žvìrb-lis*".

(13) Syllable divisions: use of - as an indication of syllable boundary follows Lithuanian practice

a. Reported in LKG (1965: 124-126); not glossed
à**t-ne**-še, ìr**k-l**as, ge-rè**s-n**is, rak**š-t**is, be-du**g-n**e, ča**k-n**o-ti, čy**p-l**ys, ar**k-l**iõ-kas, ar**k-l**e-nà, am̃**b-r**y-ti, añ**t-r**as, dum**p-l**es, dum**b-l**as, gañ**g-r**in-ti, krem**b-l**ys, krem**z-l**e.

b. Reported in Dambriunas, Klimas and Schmalstieg (1966:18):

gañ**d-r**as 'stork', pir̃**š-t**as 'finger', res-pu**b-l**i-ka 'republic'

All obstruent + liquid clusters are heterosyllabic, indicating that at least some coda obstruents maintain distinctive voicing (cf. *dump-les, dumb-las* in (13.a)).

Substantially the same conclusion is reached by considering phonotactic restrictions on clusters. Suppose that the divisions reported in (13) are interpreted as showing *ambisyllabicity* of C_1 in certain

VC_1C_2V sequences, rather than coda status for C_1. One may then consider categorical phonotactic restrictions on possible initial clusters as an indication of what is a plausible Lithuánian onset. Many obstruent + liquid sequences are systematically disallowed initially as seen below.

(14) Initial cluster phonotactics (Augustaitis (1964))

Possible	Impossible
- sl, šl, žl, pl, bl, kl, gl	**no tl, dl**
- sr, pr, br, tr, dr, kr, gr, spr, str	
- sm, šm, žm, sn, šn, žn, kn, gn	**no tm, tn, dn, dm, km, gm, pm, bm, pn, bn**
- sv, šv, žv, kv, gv, tv, dv	**no pv, bv**[6]

Although disallowed initially, clusters like *tn, dn, dm, bn* do surface in the V_V context without neutralization: *skobnis, bãdmetis, liũdnas, àtneše.* One may infer from the restricted distribution of clusters like *dm*, that they occur only in contexts where they need not be tautosyllabic[7]. The fact that voicing is maintained in the coda *d* of *bãdmetis* shows then that there is no correlation between the sites of neutralization and either the reported syllable divisions or the divisions we may infer from cluster phonotactics.

A further argument can be based on the syllable alignment effects reported by all three Lithuanian grammars: in prefixation and compounding the prefix and stem boundaries are said to coincide with syllable boundaries (Senn 1966: 61, Dambriunas, Klimas and Schmalstieg (1966:18) and the LKG (1965:125-126)).

(15) ati-**tr**aukti (not *atit-raukti) 'drag towards'; **at-i**m̃ti (not *a-tim̃ti) 'to begin from'
akì-**pl**eša (not *akip-leša) 'freche Person', sikšno-**sp**arnis 'bat' (not *sikšno**s-p**arnis)

There are no cyclicity effects in devoicing: in compounds and prefixed words, the final consonant of the first member maintains distinctive voicing if followed by a sonorant in the second member, *despite the intervening syllable boundary.* The examples below come from LKG 1965:126.

(16) voicing preserved: sta**b-m**eldýste 'idolatry, heathenism' si̇̃**lk-m**edis 'silk-tree'

---

[6]The phonetic realization of Lithuanian v in different contexts is not documented in my sources. As in some of the Slavic languages, Lithuanian v is said to pattern as a sonorant, in that it allows voicing distinctions to be maintained. I do not know whether - or how - this behavior is related to its phonetic realization.

[7]Inferring syllable composition from phonotactic restrictions is not always a sound procedure. But if this sort of inference is rejected then the basis for Licensing by Prosody vanishes also.

voicing neutralized: smul**k-ž**emis [**gž**] 'GLOSS'

Consider now a form like *stab-meldŷste.* The discussion so far has established three distinct reasons to believe that the distinctively voiced **b** is an unambiguous coda: first, all comparable obstruent-sonorant clusters are intuited to be heterosyllabic by native grammarians. Second, *bm* is an impossible word initial cluster hence an implausible intervocalic onset. Third, the assignment of *b* in *stab-meldŷste* to the onset of the second syllable conflicts with an otherwise unviolated syllable-to-morpheme alignment condition.

The option of ambisyllabicity should be pursued now more carefully (cf. Merchant 1995; Calabrese 1996). Consider *skobnis* and assume that simply preserving [voice] values in a properly licensed in the onset is a sufficient reason to generate an ambisyllabic obstruent in such a string. We may represent ambisyllabicity graphically as the improper bracketting $[_{\sigma 1}\text{sko}[_{\sigma 2}\text{b}]_{\sigma 1}\text{nis}]_{\sigma 2}$. Let us assume a grammar where [voice] is subject to prosodic licensing, in virtue of a constraint akin to Lombardi's, e.g. ***voice / unless followed by [+son] in same syllable.** We abbreviate this constraint as **License [voice]** . Suppose now that the preservation of input voicing as well as **License [voice]** outrank any competing constraints on syllable well-formedness (e.g. ***Obstruent-Nasal Onset**). The ranking **(Preserve[voice], License [voice] >> *Obstruent-Nasal Onset)** will generate syllable divisions such as *sko.bnis* with onset *b* rather than ambisyllabic *b*. This output conflicts directly with the reported syllable divisions and therefore invalidates the analysis considered. Note further that no ranking of the constraints discussed - and more generally no ranking of generally justifiable constraints - can generate an ambisyllabic result. The reason is that an ambisyllabic candidate contains both a closed syllable (hence a *Coda violation) and a complex onset. Therefore a candidate like *[sko[b]nis]* - with ambisyllabic *[b]* will always be be inferior to *[sko.][bnis],* which satisfies at least ***Coda** and violates no additional constraint. This point is illustrated by a tableau that assumes the ranking mentioned above. But the argument is independent of the ranking assumed: candidate [V[C]CV] will violate both the constraints violated by the V.CCV parse and the *Coda constraint violated by VC.CV.

(17) No ambisyllabicity in **skobnis** (the ✣ symbol marks a wrong winner)

		License [voice] >>	*Obst-Nasal Onset,	*Coda
✣	[sko][bnis]		*	*
	[sko[b]nis]		*	*!*
	[skob][nis]	!*		**

Thus no general solution to the analysis of laryngeal neutralization can be obtained by appeal to ambisyllabicity because the necessary ambisyllabic parses cannot be enforced.

Having excluded the alternatives, I present the analysis of voicing neutralization in Lithuanian as an instance of (9.i), a ranking repeated below:

(18) *voice/ V_ [-son]
↓
*voice/ V_ #
↓
Preserve [voice]
↓
*voice/ V_ [+son]

The derivations of neutralized *g* in *dau[k]* and non-neutralized *b* in *sko[b].nis* are shown below. I begin by assuming that the output of neutralization is identical to an underlying voiceless stop, a position reconsidered in the next sections.

(19) a. skobnis from /skobnis/

	Preserve [voice]	>>	*[voice] /V_[+son]
√bn			*
pn	!*		

b. dau[k] from /daug/

	*[voice]/ V_#	>>	Preserve [voice]
g#	!*		
√k			*

The Lithuanian data indicates that there is no justification for characterizing the site of licensing or neutralization in terms of syllabic position. There are licensed onsets (*smagùs* 'cheerful' vs. *žmogùs* 'man') and neutralized onsets (*spalvà* 'color'[sp], *lìzdas* 'nest'); licensed codas (*aug.muõ* 'growth' *ak.muõ* 'stone') and neutralized codas (*daũ[k]* 'much'). Voicing in Lithuanian obstruents is neutralized in all and only the positions where the main contextual cues (VOT and other release-related cues) are missing.

1. 2. The representation of neutralized voicing

We may now look into some of the issues left open by the analysis. Consider first the nature of constraints like *[voice]/V_#. The analysis presented above relies on the assumption that this constraint bars voiced obstruents but not voiceless ones from the final position: for, if *[voice]/V_# is interpreted as applying to both [voice] values then neither candidate considered in (19.b) - *[dauk]* or *[daug]* - will satisfy the higher ranked constraint. I redo the relevant tableau to show the unwanted consequence of this interpretation: as above, ✣ marks a wrong winner.

(20) Attempting to derive [dauk] from /daug/ with the hierarchy in (18) and an extended interpretation of [voice] as [α voice]

	*[αvoice]/ V_#	>>	Preserve [αvoice]
☞ g	*		
k	*		!*

However, the interpretation adopted earlier - which views the *voice constraints as banning only voiced obstruents - was not satisfactory either. First, it is inconsistent with the basic idea that cues function as licensers of the voicing contrast. Most cues to voicing - or any other feature - involve an implicit comparison between two poles of some dimension. Thus, to evaluate the significance of a stop's burst amplitude for the stop's voicing category one must know both the range of values characterizing the bursts of voiced stops and the range for voiceless ones. It is the comparison between the two that yields information about the categorization of any given token. The same goes for VOT values: the same short lag VOT value of 20 ms cues voicelessness in French - by comparison to the even shorter VOT of the French voiced stops - but voicing in English - by comparison to the longer VOT values of the English voiceless stops (Keating 1984). Neutralization takes place word finally *because the relevant comparison between VOT values cannot be carried out* in that context: it is therefore arbitrary to select just one of the poles of the voicing dimension as being either the feature itself or the feature value banned in some position.

Of course, phonological grammars may turn out to be structured in ways that are arbitrary or unexpected from the standpoint of speech perception. In this instance, however, there is good reason to think otherwise: we know independently that the laryngeally neutralized obstruents involve different articulatory postures from distinctively voiceless ones (Hsu 1996). This then is a second reason to revise the analysis in (17-18), which fails to distinguish distinctive voiceless from neutralized stops. Hsu demonstrates, on the basis of Taiwanese data, that neutralized obstruents are, in Keating's (1990) terms, *targetless with respect to voicing*: they assume the laryngeal posture of a neighboring sound[8]. Devoicing in the V_# context is passive, an automatic consequence of equalization in transglottal pressure. In contrast, distinctive {p, t, k} achieve voicelessness actively through glottal abduction (cf. Dixit 1987). The difference between targetful voiceless {p, t, k} and neutralized targetless {P, T, K} has multiple consequences in the phonology of Taiwanese and other languages: in particular, the realization of

[8] *Targetless* (with respect to some gesture) means more than *unspecified* (with respect to the corresponding feature): non-specification can be phonetically interpreted in various ways, including through the assignment of a fixed articulatory target. This is the interpretation of the frequent statement "Laryngeally neutralized stops are phonologically unspecified and phonetically voiceless" (Mascaró 1987, Clements 1985, Lombardi 1995). In contrast a segment that is targetless for feature F is both unspecified and lacking in an invariant articulatory realization for the corresponding gesture.

neutralized stops is variable in context, precisely because they lack their own articulatory target, while distinctively voiceless ones are invariant in comparable positions. A case of this sort is discussed in section 7.

The facts reviewed suggest the following interpretation of neutralization: both distinctively voiced and voiceless stops possess specific auditory targets, to be implemented through specific articulatory routines. The positions of potential neutralization are those where the hearer is less likely to correctly evaluate the achievement of these auditory targets. For such positions the grammar may evaluate as optimal obstruent representations which place no perceptual burden on the hearer and require no articulatory effort from the speaker: there is no auditory target to achieve in the neutralized obstruent, no distinct auditory category to identify and therefore no specific, invariant set of articulatory gestures to perform. Under this interpretation, the neutralized obstruents are distinct from both the voiced and the voiceless ones and the *[voice] constraints exclude equally the distinctive voiceless and voiced series[9]. To show how this new interpretation of the *[voice] constraints operates, I revise the earlier tableaux in (18) below: [voice] is interpreted to mean [α voice] - both plus and minus - and {P, T, K} represent laryngeally neutralized stops whose phonological representations lack both the invariant auditory propeorties associated with [+voice] or [-voice] and the articulatory gestures used to implement these auditory targets.

(20) a. skobnis from /skobnis/

	Preserve [voice]	>>	*[voice] /V_[+son]
√bn			*
Pn	!*		
pn	!*		*

b. dau[k] from /daug/

	*[voice]/ V_ #	>>	Preserve [voice]
g#	!*		
√K			*
k	!*		*

One more comment on *[α voice] constraints. The *[α voice] condition can be interpreted in at least three ways. The first possibility is that *[αvoice] is penalizes the *articulatory* effort expended in implementing distinctive voicing values: cf. Kirchner (1996). If this line is pursued, then our representations must be expanded to distinguish explicitly the articulatory from perceptual correlates of

[9]The position taken by Hsu (1995) and here conflicts with the idea that voicing is universally a privative feature (Mester and Ito 1989, Lombardi 1991; but cf. Rubach 1996 for evidence to the contrary): the representations of voice-neutralized stops must differ, in at least some languages, from those of distinctively voiceless stops.

distinctive features: the articulatory interpretation of the *[α voice] constraints penalizes not the fact that the auditory correlates of voicing or voicelessness are being generated but the fact that specific articulatory gestures are required to do this. The neutralized stops do not violate the constraints because no specific gestures characterize them, insofar as voicing is concerned.

A second possibility is that *[αvoice] is used as a means to limit perceptual uncertainty: this accords with the fact that the ranking in (6) reflects a scale of information content. Distinctive voicing is least likely in the least informative contexts so we may also view these constraints as being listener-oriented, in the sense of Ohala 1981. What seems most likely, however, is that both the articulatory effort and the perceptual poverty are being referred to in these constraints, in the sense that what is penalized is an unfavorable ratio of effort expended to cues generated. For the same amount of articulatory effort spent in generating some degree of voicing across three contexts (_[+son], _#, _[-son]) the cues available to identify voicing are progressively worse or fewer and therefore the ratios of effort to cues differ in ways that mirror the perceptibility scale in (5). This ratio can be improved by spending less effort and falling short of the articulatory target or by categorically giving up on the distinction in the perceptually difficult context. The latter is the standard case of neutralization. The interpretation of constraints like *[αvoice]/X_Y along these lines seems both plausible and consistent with the documented existence of gradient reduction of oral constriction gestures in contexts of reduced perceptibility (Byrd 1994, Jun 1995). In order to formalize this type of solution, some specific quantification of effort as well as of the information content of cues will be necessary: these points cannot be addressed here. See Kirchner (1997) for part of the necessary solution. I will continue to employ in Part I statements like *[αvoice]/X_Y with the understanding that the rationale for such conditions is the fact that they prohibit progressively worse ratios of cues to effort.

1. 3. Word domain effects in voicing neutralization

A different issue that has arisen from the discussion of the Lithuanian data concerns the word-bound character of the voicing neutralization. Word final obstruents are reportedly neutralized regardless of whether a following word begins with an obstruent or a sonorant: e.g. *dau[k] akmens* 'many stones' from */daug/*. This phenomenon should not be tied to the aligned syllabification *daug.ak.mens.*, where the original /g/ is in the coda, for we have seen that other codas do maintain distinctive voicing in this language. My proposal is that in phrases like *dau[k] akmens* we're dealing with a cyclic effect. In procedural terms, */daug/* devoices on the word cycle, prior to the consideration of any licensing context offerred by the following word. The same suggestion can be modelled non-derivationally, with cyclic effects viewed as stemming from the action of constraints that limit paradigmatic alternations (Burzio 1995, 1997, Flemming 1995, Kenstowicz 1995, Steriade 1995, 1996, 1997). By having a constantly devoiced *k* in *dauk* and related forms, Lithuanian simplifies its phrasal paradigms and blocks the proliferation of allomorphs. This point is developed in section 5, where the phrasal realization of laryngeally neutralized stops is being considered.

2. Generalizing from Lithuanian

The next step in the argument is to show that Lithuanian is fully representative of the typology of voicing assimilation. I do this first by showing that identical patterns of devoicing are found in other languages: the aim here is to establish that any observable connection between being a coda and being laryngeally neutralized represents an accidental by-product of facts unrelated to syllable structure. Second, I show that the analysis proposed for Lithuanian extends straightforwardly to a related but distinct style of voicing neutralization: the case where voicing distinctions are maintained word-finally but eliminated before obstruents. Syllable-based analyses of this pattern are also shown to be untenable. Finally, the languages considered in this section permit an extension of the analysis to other laryngeal contrasts cued by VOT.

To determine what counts as relevant evidence in comparing cue-based to syllable-based licensing, the reader should note that many languages - such as Korean (Kim-Renaud1974) - disallow heterosyllabic obstruent-sonorant sequences. However this effect is analyzed (cf. Vennemann 1988), its consequence is that many languages will lack the very sequences whose behavior is differently predicted by the two analyses considered. The Licensing by Cue model presented so far predicts that strings like *Vp.rV, Vb.rV* may maintain distinctive voicing, whereas Licensing by Prosody, in all of its versions, predicts that the *p/b* distinction will be neutralized. If such sequences are either lacking (as in Korean), or must be tautosyllabic (as in modern Romance), then the data will not distinguish the proposals we compare. For this reason, languages like Korean will not be mentioned here.

2.1. Greek and Sanskrit

The distribution of laryngeal features in the older Indo-European languages has been taken to reflect syllable-based licensing conditions (Steriade 1982). Greek possesses the distinction between voiced, voiceless and voiceless aspirated stops but implements it only before sonorants. Sanskrit contrasts voiced, voiceless, voiced aspirated (murmured) and voiceless aspirated stops: here too the laryngeal contrasts surface only before sonorants. Examples of contrast and neutralization in both languages appear below. Stops are not allowed word-finally in Greek. I choose [-anterior] stops to exemplify the pattern because these segments occur with relatively fewer distributional restrictions.

(22) a. Greek laryngeal contrasts and neutralization in velar stops

	Voiceless	Voiced	Aspirated
a. Pre-sonorant:	deik-nu:-mi "I show'	zeug-nu:-mi 'I yoke'	a**k**$^{\mathbf{h}}$-nu:-mai "I am troubled'
b. Pre-vocalic:	t^{h}o:ra**k**-os 'thorax-GENsg"	laryn**g**-os 'larynx-GENsg'	tri**k**$^{\mathbf{h}}$-os 'hair-GENsg'
c. Pre-stop:	dei**k**-teos 'to be shown'	zeu**k**-teos 'to be yoked'	he**k**-teos (cf. e**k**$^{\mathbf{h}}$-o: 'I have') 'to be had'
d. Pre-s:	t^{h}ora**k**-si 'thorax-DATpl'	laryn**k**-si 'larynx-DATpl'	t^{h}ri**k**-si 'hair-DAT-pl'

(22) b. Sanskrit laryngeal contrasts and neutralization in [-anterior] stops

	Voiceless unaspirated	Voiceless aspirated	Voiced unaspirated	Voiced aspirated
a. Pre-sonorant	va**c**-mi 'I speak'	ca**k**$^{\mathbf{h}}$ya:-u 'has seen'	ti**g**-ma 'sharp-pointed'	da**g**$^{\mathbf{h}}$-nu:-yat 'reaching'
b. Pre-vocalic	u-va:**c**-a 'has spoken'	**k**$^{\mathbf{h}}$an-ati 'digs'	ni-ne**j**-a 'has washed'	da**g**$^{\mathbf{h}}$-at 'has reached'

	Voiceless unaspirated	Voiceless aspirated	Voiced unaspirated	Voiced aspirated
c. Word-final	va:**k** 'voice'	no examples	var**k** (cf. var**j**-a) 'twisting'	d^{h}a**k** 'has reached'

	Voiceless unaspirated	Voiceless aspirated	Voiced unaspirated	Voiced aspirated
d. Pre-obstruent	u**k**-ta 'spoken'	no examples	ni**k**-ta 'washed"	d^{h}a**k**-tam you 2 reached'

These patterns of neutralization must be analyzed in the same terms as the Lithuanian facts. The observation establishing this will be that for both Sanskrit and Greek syllable divisions in obstruent-sonorant clusters were variable, depending on the dialect, the period, the literary style and the juncture separating the consonants. In contrast, and this is fundamental, there was no variation in the pattern of laryngeal neutralization: in styles or dialects where VC.CV divisions were the norm for all clusters, laryngeal neutralization did not take place before heterosyllabic sonorants. Therefore the syllabic assignment of clusters and the licensing of laryngeal features are independent of each other.

2.1.1. Sanskrit

According to Mishra (1972:200ff) - a compendium of the opinions on syllable divisions held by Sanskrit grammarians - "the most general rule is that the first member of the consonant group [...] consisting of either two or more than two consonants, belongs to the preceding vowel. Thus the word "pitre" will be divided as pit+re and not pi+tre. " The only exceptions to this statement are: (a) the opinions of certain grammarians that RigVedic Sanskrit had the option of syllabifying stop-liquid sequences as onsets and (b) the view expressed by commentators on the YajurVeda that stop-glide and stop-fricative sequences are tautosyllabic. Even if taken at face value, these exception statements leave us with possible heterosyllabic assignments for all sequences of stop-consonantal sonorant in the RigVeda and obligatory VC.CV assignments for all stop-consonantal sonorant sequences in the later language. This means that the aspirated voiced stop in *dag^h-nu:-yat* 'has reached' is, as a rule, a coda segment: what licenses its laryngeal features is something else than syllabic position[10]. I also note that, as in Lithuanian, most stop--nasal sequences present intervocalically are impossible initially, and thus implausible as onsets: they do however maintain distinctive voicing, as in *agni-* 'fire', *stab^h-na:-* 'establish-present'.

Furthermore, the patterns of vowel lengthening in the intensive prefix suggest that in fact all clusters were heterosyllabic during all periods, an opinion shared by Mishra (1972) and supported by the metrical evidence. Thus the vowel-final reduplicating prefixes of the intensive are short before any consonant cluster and long before a single root-initial C (Whitney 1889:365). This rule, strictly obeyed in the earlier stages of Sanskrit, indicates that the syllable division was VC.CV in all cases. Below I highlight the pre-stem rimes containing the vowel with variable length.

(23) Pre-stem vowel length in the intensive

(i) Long vowel	(ii) Short vowel
gan-**i:**-gam- 'go'	gan-**i-g.**m-atam 'go'
mar-**i:**-mṛj- 'wipe'	kan-**i-k.**rand 'cry out'
b^{h}ar-**i:**-b^{h}ṛ- 'bear'	b^{h}ar-**i-b^h.**ra-ti 'bear'
tav-**i:**-tuat- 'be strong'	dav-**i-d.**yut- 'shine'

[10]For a discussion of post-Vedic syllabification see Vaux 1992, who argues that a shift in boundaries took place in later Sanskrit. Vaux's arguments also support the point made in the text: despite the extensive variation in cluster assignment across styles and periods, the patterns of laryngeal neutralization do not change. They are therefore independent of syllable position.

The requirement is clearly that pre-stem syllables be heavy in the intensive[11]. This condition is satisfied in the (ii) column by heterosyllabification: *ga.nig.ma.tam, da.vid.yut,* b^h*a.rib*^{h}*.ra.ti* and by lengthening in the (i) column: *ga.ni:.gam,* b^h*a.ri:.b*^{h}*ṛ-.*

Note now that the coda stops *g, d, bh* in (23.ii) do not lose their laryngeal features: heterosyllabification - as in b^h*a.rib*^{h}*.ra.ti* - does not induce laryngeal neutralization. Note further that this argument is independent of the method of syllabification in the rest of Sanskrit: even if closed syllables were avoided in other contexts, the last syllable of the intensive would still have to be analyzed as heavy, and therefore the syllabic divisions *ga.nig.ma.tam, da.vid.yut,* b^h*a.rib*^{h}*.ra.ti* would still suffice to establish the fact that laryngeal licensing is independent of syllabic position.

2.1.2. Greek

The facts of Greek syllabic division have been discussed in Hermann (1924), Steriade (1982) and most recently by Devine and Stephens (1994:32-42 and passim). Certain Attic poets syllabify consistently the tautomorphemic stop-sonorant clusters as onsets. But this cannot be the basis for a syllable-licensing account of the laryngeal neutralization facts shown earlier in (20). First, the laryngeal neutralization pattern is pan-Hellenic, spanning dialects, literary styles and periods, whereas the syllabic divisions are highly variable. Homer tends to assign all word-internal intervocalic clusters to separate syllables, as shown by the weight of syllables reflected in the meter. Since forms like *dak.ru, ag.rion* and *ak*^{h}*.nu:mai* contain coda velars, laryngeal licensing for this variety of Greek cannot be syllable based.

Further, the attested voiced stop-nasal clusters (limited to *dn, dm* in most dialects) are scanned heterosyllabically in all literary styles (Koster 1952:34), without however losing distinctive voicing on *d.* In Attic both *dm* and *dn* are absent word-initially, a fact consistent with the assumption that they cannot be onsets. The voiced-stop-l clusters (*gl, bl*) are variably heterosyllabic in dialects like Attic, where other obstruent liquid sequences form complex onsets (Steriade 1982): but this does not cause variable or even occasional loss of distinctive voicing.

Further arguments for a syllable-independent statement of laryngeal neutralization emerge from a consideration of allomorphy in the thematic comparative and superlative forms of adjectives (Devine and Stephens 1994:40, 104): the allomorphs are *-oteros* (comparative), *-otatos* (superlative) after a

[11] This condition is also evidenced by the alternative form of the intensive, where no -i- intervenes between the reduplicated syllable and the stem: *can-krand-* (alternative to *kani-krand*), *ja:-gam* (alternative to *gani:gam-*) etc.

heavy syllable and *-o:teros*, *-o:tatos* after a light one. Below I highlight the syllable immediately before the comparative suffix:

(24) Comparative allomorphy in Greek

(i) After heavy syllable	(ii) After light syllable
de:.l-oteros 'clearer'	k^{h}a.**le**.p-o:teros 'more difficult'
sem.n-oteros 'more venerable'	**ne**.-o:teros 'younger'
pis.t-oteros 'more faithful'	**so**.p^{h}-o:teros 'wiser'
mak.r-oteros 'longer'	**p^{h}i**.l-o:teros 'dearer'

The argument here is exactly parallel to that based on the Sanskrit intensives. The allomorphy facts require syllabic divisions that place some of the non-neutralized obstruents in the coda (cf. *mak.roteros* vs. *ag.roteros)*. Therefore what is constant for all laryngeally licensed obstruents is not the syllabic position but rather the presence of a following sonorant. The constraint hierarchy generating laryngeal neutralization in Greek and Sanskrit is thus identical to that argued for in Lithuanian[12]. The fact that these languages have occasional differences in syllable assignment only reinforces our argument for uncoupling neutralization from the syllable.

2.1.3. Aspiration neutralized

Note that Greek and Sanskrit aspiration is neutralized in the same contexts as voicing: word-finally and before obstruents. This is due to the fact that the aspirated stops of both languages are *post-*

---

[12]Steriade (1982) has argued that the reduplication patterns - which treat root-initial stop-sonorant clusters on a par with single consonants - establish the onset status of all such clusters. Thus *lu:-o:* reduplicates as *le-lu:-ka, graph-o:* reduplicates as *ge-graph-a, blapt-o:* as *be-blaph-a* but *ktiz-o:* reduplicates as *e-ktis-ma* and *stell-o:* as *e-stal-ka*. The point is that the clearly heterosyllabic *kt* and *st* clusters pattern differently from the arguably tautosyllabic *gr, bl*. Had *gr* and *bl* been heterosyllabic in all contexts, the facts of reduplication would remain unexplained: we wouldn't be able to predict the difference between *ge-grapha* and *e-ktisma*. This may well be true, but this argument does not establish that stop-sonorant clusters are onsets in all contexts. Root-initially, the cluster assignment was probably subject to additional constraints, which reflect the preference for alignment between root and syllable boundaries. Indeed, Devine and Stephens (1994) provide considerable metrical evidence for the tendency to avoid misaligned syllabifications in Greek poetry and prose, at all levels of the prosodic hierarchy. Clearly, the preference for aligned root and syllable edges was overridden by the dispreference for marked *kt, st, mn* onsets: hence *es.tal.ka* rather than *e.stal.ka*. This then explains the difference in the patterning of *st-, kt-, mn-* initials vs. *bl, gr*, etc. The alignment constraints were irrelevant in other positions: therefore *agr-oteros, pukn-oteros* must have been syllabified, as argued above, *ag.ro.te.ros, puk.no.te.ros* in the same language where *gegrapha* was syllabified - frequently or invariably - as *ge.gra.pha*. The argument against syllable-based licensing of laryngeal features formulated earlier stands.

aspirated: meaning that the chief effect of aspiration is to prolong the voicing lag. Therefore VOT values must have played a major role in differentiating all laryngeal categories in Greek and Sanskrit and contexts where these values cannot be observed and compared are likely neutralization sites. This explains why the contexts of neutralization for voicing and aspiration are identical in the data observed so far. Had the aspirated stops been pre-aspirated, the context where *t*: ^{h}t contrasts are lost may well have been different from the context where *t: d* contrasts are suspended, as we see in Part II.

The range of available cues to postaspiration has been studied less than those signalling voicing (cf. however Schiefer 1992). It is nonetheless quite plausible that the cues - and therefore the contexts where these cues can be observed - might differ. Schiefer (1992) does not mention either F0, F1 or durational differences that would reliably separate p from p^h, or b from b^h. It seems likely that all the cues to such contrasts are contextual (VOT and burst) and occur exclusively the vicinity of the release. Therefore, if this range of cues is characteristic of all postaspiration contrasts, we may predict that the neutralization of such distinctions will *always* occur in unreleased word-final and pre-obstruent stops. This prediction is confirmed by languages like Khasi (Henderson 1976), in which distinctive voicing is maintained initially before all consonants (cf. *bt, pd, pt,* bt^h, *bs, bn, bl, pl, pn* initial clusters) whereas distinctive aspiration is maintained only before sonorants (t^hl, k^hm, t^hm, p^hn, but * t^hp, *p^hd, etc.) Khasi also preserves a marginal voicing contrast finally, but not the aspiration contrast. The point here is that voicing and post-aspiration, although frequently parallel in licensing behaviors, are nonetheless different in their overall typology. They are different because their cue distribution is not exactly the same: voicing contrasts can be and are maintained in the absence of the onset cues, whereas postaspiration contrasts cannot be maintained - because they cannot be perceived - under the same circumstances. This strengthens the argument for cues as contrast licensers, since the cues are specific to the feature and timing relation involved.

2.2. Voicing neutralization in Russian and Polish

The evidence against a syllable-based analysis of laryngeal licensing in Russian has been gathered by Pilch (1967:1564) and Darden (1991). This pattern of voiced/voiceless neutralization is in fact identical to that of Greek, Sanskrit and Lithuanian, modulo independent differences in phonotactics and laryngeal inventories. The facts of Russian have been analyzed by Jakobson (1962 (=1956):503), Halle (1959), Hayes (1984), Kiparsky (1985), Mascaró (1987) and Lombardi (1991). The core of the paradigm is the fact that the voiced/voiceless distinction in obstruents is neutralized in word-final position and before other obstruents. It is not neutralized before vowels or sonorants belonging to the same word or the same prefix-stem unit (cf. Jakobson 1962 for details). The neutralized obstruents surface as voiceless unless followed by voiced obstruents, in which case they are voiced.

Since this paradigm is otherwise identical to those observed in earlier sections, I will concentrate below only on the Russian evidence for laryngeal licensing of *coda* obstruents. Pilch (1967:1564) reports that non -neutralized voiced obstruents may occur as codas if followed within the same word or close-knit phrase by neutral elements with respect to voicing, i.e. by sonorants.

(25) a. Licensed voiced obstruents in medial Russian codas:

zəmʲorz.lə	'frozen'
skorb.nij	'full of grief'
əd.no	'one-neut.'

b. Licensed voiced obstruents in final Russian codas:

tʲigr	'tiger'	(contrast mokr 'damp')
ʒiznʲ	'life'	(contrast pesnʲ 'song')

Darden (1991) makes some complementary observations. First, he notes that not all laryngeally neutralized obstruents occupy the coda position: *[g]zʲemlʲe* (from */k zʲemlʲe/* to-earth') does not begin with a coda. Nor can the non-neutralized obstruents seen in (26.b) be onsets: *n* is not syllabic in either *ʒiznʲ* or *pesnʲ* and therefore the preceding fricative cannot be an onset. These observations render untenable a simple onset licensing analysis of Russian voicing along the lines proposed by Rubach (1990) for German. Pilch's data (26.a) also argue against Lombardi's (1991, 1995) version of laryngeal licensing, under which a tautosyllabic sonorant must follow any distinctive laryngeal value: *d* in *əd.no* is not tautosyllabic with *n*. Further, according to Darden, Russian displays the same stem-to-syllable alignment effects noted earlier for Lithuanian: e.g. *pəd-jexatʲ* 'to exit' has heterosyllabic *d.j* with non-neutralized *d*. (cf. contrasting *ət-jexatʲ* 'to enter'). Darden's evidence for this is based on the observation that regressive palatality assimilation always obtains within the same syllable, but fails to cross certain prefix-stem boundaries, such as the one in *pəd-jexatʲ*. Palatality assimilation applies only when the affix is monoconsonantal: e.g. *s-jes-tj* [sʲjesʲtʲ] 'to eat with'. I infer from this that prefix-stem boundaries align to syllable boundaries in Russian, unless the resulting syllable are grossly ill-formed (**[s.jes.tʲ]*. An analysis in terms of constraint ranking can be proposed, where the constraints of stem-to-syllable alignment are outranked only by avoidance of the worst syllable types (syllabic obstruents). The upshot is, in any case, that aligned candidates are optimal in forms like *pəd-jexatʲ* and that this results in syllable divisions that block tautosyllabic palatality assimilation (**pədʲ-jexatʲ*.) Most relevant here is that the syllabic division enforced by alignment is irrelevant to laryngeal licensing: voicing in the coda *d* is linearly licensed by the cues inherent in the following heterosyllabic *jV* sequence.

The paradigm of Polish voicing has been studied by Bethin (1984), Rubach and Booij (1990), Gussmann (1992), Lombardi (1991) and Rubach (1996). Rubach and Booij provide important evidence for stem-to-syllable alignment, even in cases where the resulting syllable structure is quite unusual: for instance the aligned syllabification of *o-mdlec'* 'faint' *[o.mdlec']* is an acceptable structure in Polish[13] along with [om.dlec']. Lesser violations of syllabic well-formedness result in an even greater likelihood of aligned boundaries: the subjects of Rubach and Booij's investigation were unanymous in preferring *bez.alkoholowy* 'alcohol free' to misaligned (but syllabically improved) *be.zalkoholowy.* Thus underlying **z** is viewed as a coda in this form, but preserves distinctive voicing[14]. Given this data, it is unnecessary to discuss the possible stem-internal syllabifications of the notoriously complex clusters attested in Polish: whether *Piotrka* is syllabified as *Pio.trka* (cf. Gussmann 1992, Lombardi 1991, 1995) or otherwise, we know, based on Rubach and Booij's alignment effects, that syllabic divisions do not condition laryngeal licensing in Polish either[15]. This point is in agreement with the major conclusion reached by Rubach (1996).

Both Russian and Polish display complex patterns of voice licensing and assimilation when the obstruent is followed by a non-prevocalic sonorant, as in Russian *bobr* 'beaver'. These cases have to be considered because they bear on the idea that a sonorant carries the VOT cue of a preceding obstruent and in that sense, licenses its voicing value. At first sight, one might expect that a sonorant following the obstruent will always insure the intact realization of the obstruent's voicing value. Since this is not always the case in these languages, additional analysis is called for. The paradigm for the two languages is assembled in (26): abbreviations used are O = obstruent and R = sonorant. The sources of this data are Jakobson (1956), Hayes (1984) Kiparsky (1985) and, for Polish, Rubach and Booij (1990), Gussmann (1992) and Rubach (1996).

(26)	Russian	Polish
OR#	*optional devoicing,* *not neutralizing* (bobr, bobr̥, bopr̥ 'beaver')	*obligatory neutralization* *when the sonorant is non-syllabic* (spa[sm], spa[sm̥] /spazm/ 'spasm')
$O_1R\#O_2(R)V$	*no assimilation* my[slJ̇ ʒ]e 'thought, though'	*obligatory assimilation to* O_2 wia[dr z]achodni 'westerly wind'

[13]The judgments reported by Rubach and Booij have been verified by Alicja Gorecka.

[14]The final *z* of *bez* in *bez alkoholowy* is not voiced by assimilation to the following voiced vowel: although voicing assimilation initiated by sonorants is attested in Krakow Polish, the dialect described by Rubach and Booij (1990) and Rubach (1996) is that spoken in Warsaw, where only obstruents induce assimilation. Therefore the *z* in *bez.alkoholowy* is a coda that maintains its underlying voicing value. The same point is made by Gussmann (1992:33, fn. 4) a propos of *bez nadziei* 'without hope', another case where the distinctive voicing is maintained in a coda obstruent.

[15]Gussmann's suggestion that voicing is licensed in syllabified obstruents but not in stray ones is also inconsistent with the alignment data adduced by Rubach and Booij.

		li[dr v]ody 'liter of water
O1#RO2 (R)V	*obligatory assimilation to O2* o[d mzd]y 'from the bribe' i[sm̥ts]enska] 'from Mcensk'	*no assimilation* wido[k mg]ły 'sight of mist'
O1 RO2 (R)V	*obligatory assimilation to O2* [g mzd]y 'with the bribe'	*obligatory assimilation to O2* kon[tr̥f]ors 'contrefort' [drg]nąc' 'shudder' [str̥f]onic' 'squander'

Reformatskij (1971, cited in Hayes 1984) notes that the transparent sonorants of Russian must be non-syllabic: syllabic *m̩* in a sequence like [zm̩ts] will block voicing assimilation. For Polish, most recent sources mention non-syllabic sonorants in the positions listed in (27) but A.Gorecka (p.c.) points out that her variety of Polish is characterized by facts very reminiscent of Reformatskij's generalizations: word final sonorants in forms like *spasm* are syllabic and block devoicing: *spa[zm̩]* contrasting in voice with *pa[sm̩]* 'stripes-Gen pl'. The syllabic sonorants of Gorecka's dialect also block assimilation: *litr wody* 'a liter of water' is realized as *[litr̩ vodi]*, with syllabic *r* and unassimilated *t*, in contrast with the standard pronunciation *[lidr vodi]*, with non-syllabic *r* and assimilated *d*. Rubach (1974) also mentions the option *spa[zm]* and the possibility of syllabic sonorants, although his subjects differ on this point. The phonetic realization of the Russian sonorants in (26) has been only partially documented, by Hayes (1984), who shows that they are variably voiceless in forms like *rta* and *bobr*. An impressionistic report to the same effect is made by Gussmann, Rubach and Booij about Polish forms such as *Pio[tr̥]*, *spa[sm̩]* and *[str̥f]onic'*. Gussmann (1992:33) however reports that the initial sonorant of Polish *rtęc'* 'mercury' is fully voiced.

Before proposing an analysis of voicing in ORO sequences, let us note the major descriptive generalizations shared by Russian and Polish. First, distinctive voicing is maintained - regardless of syllable divisions - in OV and ORV strings, in both languages. Voicing is neutralized in ORO and OO strings in both languages. Syllabic sonorants - when allowed in either language - act exactly like vowels in blocking neutralization and assimilation of a preceding obstruent. The differences between Russian and Polish involve then only the status of OR# and #RO strings: final non-syllabic R allows voicing to be maintained in Russian OR#, but not in Polish; and initial non-syllabic R blocks assimilation in Polish but not Russian O#RO strings. The key observation for our analysis will be that the OV, ORV and OR̩ strings (where R̩ is a syllabic sonorant) allow optimal manifestation of the onset cues to voicing, esp. VOT: in these cases, the sonorant string following the obstruent is longer than in the case of ORO clusters, with non-syllabic R. In complementary fashion, we note that in ORO strings - with non-syllabic, word-internal R - the string potentially manifesting onset cues is shortest and doubly overlapped. Therefore, if the duration of the cues to voicing plays a role in the perceptibility of the distinction, then we expect that O_1's voicing in O_1RO_2 will be least perceptible. This corresponds to the fact that in O_1RO_2 strings both languages neutralize voicing in O_1. The corelation between overall duration of the sonorant string and neutralization is outlined below:

(27)			
	Obstruent followed by longest sonorant string		Obstruent followed by shortest sonorant string
	OV, ORV, OṚ, OṚ O	OR#	ORO
	Voicing contrast preserved	Variation	Voicing contrast neutralized

Further support for this correlation between the duration of the sonorant and its laryngeal-licensing abilities comes from Klamath (Blevins 1993, Barker 1964: 23, 26-27) where an O_1RO_2 sequence contains what Barker perceives to be a syllabic sonorant preceded by a laryngeally licensed O_1: contrast *makḷga* 'camps at' with *ntʃik'ḷga* 'drips down" and *wdogḷgi* 'comes to beat someone' (Barker 1964:23). This syllabic sonorant is however ignored for the purpose of metrical counting. As Price (1980) shows, duration is one of the correlates of syllabicity: sonorants perceived as syllabic differ from their non-syllabic counterparts primarily in being longer. Thus we may infer that the Klamath inter-obstruent sonorants are perceived by Barker as syllabic on acount of their duration alone. The fact that these longer sonorants help preserve the distinctive laryngeal qualities of the preceding obstruent - unlike their non-syllabic Polish counterparts - should be attributed to the effect of duration on the perceptibility of VOT distinctions.

In word final OR# sequences, where R is overlapped on only one side, Polish and Russian diverge: Russian counts this final R as a still valid licensing context - since devoicing in forms like *bobr* is gradient and non-neutralizing - whereas Polish does not: final OR# is neutralized in Polish, when the sonorant is non-syllabic. This difference in the treatment of final OR# sequences with non-syllabic R may relate to differences in the degree of temporal reduction of the final sonorant but it more likely reflects a different categorization of otherwise identical phonetic data. Assuming the latter, I expand the perceptibility scale proposed in (7) to distinguish three classes of contexts in which a sonorant follows an obstruent: (a) a context in which the obstruent is followed by a long sonorous stretch (either V, RV or syllabic R); (b) the context in which the obstruent is followed by a word-final, non-syllabic R; and (c) the context where the obstruent is followed by the shortest R, non-syllabic and doubly overlapped (ORO). The distinction between contexts (b) and (c) stems from the lesser extent of overlap in (b) and the likelihood of final lengthening, which may result in a longer R word finally. In both respects then, the _R# context emerges as more favorable for the perception of voicing that the _RO context.

(28) Scale of perceptibility in obstruent voicing according to context (revised)

Notation: [long son] = long sonorous stretch (V, RV, or syllabic R)
[son] = shorter sonorous stretch (R#)
[short son] = shortest sonorous stretch (_RO)

V_[long son] ▶V_[son,] ▶ V_[short son] ▶ V_# ▶ V_[-son] ▶ {[-son] __[-son], [-son] __#, #__[-son] }

Corresponding to this expanded scale, we have an expanded set of *[α voice] constraints.

(29) Revised *[α voice] constraints corresponding to (28)

(iii) *αvoice/ V_ #
↓
(iv) *αvoice/ V_ [short son]
↓
(v) *αvoice/ V_ [son]
↓
(v) *αvoice/ V_ [long son]

Both the scale in (28) and the corresponding constraints should be read by interpreting [short son], and [long son] to refer to the durational category of the overall string of sonorants that follows the obstruent. It does not matter whether an obstruent is followed by one or more sonorants provided that the overall sequence is such as to allow a reliable identification of VOT distinctions.

Given these additions, we may describe the variation between licensed Russian *bobr* and neutralized Polish *bopr* as a simple ranking effect: final r in /bobr/ for both languages corresponds to the sonorous string of intermediate duration ([son]) identified earlier.

(30)

Polish	Russian
*αvoice/ V_#	*αvoice/ V_#
↓	↓
*αvoice/ V_ [short son]	*αvoice/ V_ [short son]
↓	↓
*αvoice/ V_ [son]	**Preserve [voice]**
↓	↓
Preserve voice	*αvoice/ V_ [son]
↓	↓
*αvoice/ V_ [long son]	*αvoice/ V_ [long son]

Since Polish, but not Russian, neutralize O_1 in $O_1R\#$ sequences, it follows that Polish, but not Russian will assimilate O1 in $O_1R\#O_2$, as seen in (27). Following earlier work (esp. Kiparsky 1985; Mascaró 1987; Cho 1990), I assume that the target of assimilation is identified as the neutralized segment: therefore a non-neutralized obstruent (Russian *b* in *bobr*, for instance) will not be a target. This is then sufficient to derive the Polish-Russian differences on this point.

The last case to consider is the behavior of Polish and Russian $O_1\#RO_2$ sequences. Here again my remarks are speculative, in the absence of a systematic phonetic investigation. We should observe at the outset that in any context $O_1\#R$ - regardless of what follows R - O_1 is neutralized in both languages, unless it belongs to a prefix or proclitic: an example of neutralization in this context is a Russian phrase such as *gorod Moskva [gorət məskva]* 'the city of Moscow', with devoiced /d/. This is a cyclic effect, as in Lithuanian: both Polish and Russian prefer to generalize to all contexts the devoiced quality that is justified phrase-finally. The only question is how the neutralized O_1 will be realized: fully voiceless, as expected phrase-finally, or partially assimilated to some other consonant. I claim that the occurrence of assimilation in this case depends on three conflicting factors: the preference for paradigm uniformity or morphemic invariance, whose effect would be to generalize phrase final allomorphs like [gorət] to all contexts (Steriade 1996); the avoidance of increasingly faster articulatory adjustments, which militates against abutting voiceless-voiced sequences; and the need to preserve lexical contrasts. The last factor requires that, in an obstruent whose glottal state is distinctive, it must be initiated early so that the obstruent will be fully voiced or fully voiceless throughout its duration: therefore in $\#RO_2$ (with voiceless O_2) R is partially or fully devoiced (as shown by Hayes 1984 for *rta*). Similarly, in $O_1\#O_2$ sequences, the neutralized O_1 will reflect the transition to the glottal state of O_2 and thus will be categorized as belonging to the same laryngeal class as O_2. This is the basis of voicing assimilation in such cases. However, in strings like $O_1\#RO_2$, the glottal state of O_2 need not be initiated during O_1: a slow enough transition between glottal states is possible even if O_1 preserves its phrase-final voiceless quality. Whether the transition between O_1 and O_2 is slow enough depends on how long R may be: if we take seriously Gussmann's (1992) report that R in forms like Polish *rtęć* 'mercury' is fully voiced and compare it with the partial devoicing found in Russian *rta* 'mouth-Gen. sg', then this suggests a durational difference between the two languages which correlates as predicted with the facts of assimilation. The shorter R of Russian in O_1#RO_2 strings is an insufficient buffer between opposing glottal states and thus induces assimilation, whereas the longer R of Polish permits O_1 to maintain the devoiced quality that characterizes the phrase-final allomorph. The fact that Reformatskij (1971) explicitly rules out assimilation in Russian O_1#RO_2 when R is syllabic further supports this line of analysis: as suggested earlier, the syllabic R is the longest R and thus will necessarily allow O_1 to preserve its paradigmatically dictated voiceless state[16].

Summarizing then, I have suggested that two different phenomena are reflected in the paradigm in (26). One is the ability of a shortened and overlapped, non-syllabic R to offer useable transitional cues to a preceding O: Polish and Russian agree in categorizing the shortest instance of R (in ORO contexts) as an insufficient licenser. They differ in the categorization of final R: Russian, but not Polish, counts this R as a possible licenser of distinctive voicing. This is reflected in the ranking differences seen in (30). The other phenomenon is also related to the duration of R but involves not its licensing ability but rather its function as a buffer between the opposite glottal states of neutralized O_1 and distinctive O_2. When R is too short, O_1 will assimilate, as if it was adjacent to O_2; when R is sufficiently long, O_1 will maintain a voiceless quality.

To implement this analysis it was necessary to expand the perceptibility scale and the related set of *α voice constraints. These additions do not represent unwanted epicycles: the scale of perceptibility involves a ranking of contexts according to the quality of the information they offer for the identification of voice categories. We have started out with a small set of contexts (in (7)) but any realistic consideration of a larger set of contexts will result in refinments of the sort just introduced. What is constant throughout such revisions is the idea we started out with: that the likelihood of neutralization is determined by the quality of information provided by the context.

---

[16] The analysis presented here agrees with Rubach's (1996) important proposal that the distribution of Polish voicing is independent of the syllable. However the two analyses diverge on the mechanisms of assimilation in ORO, OR#O and O#RO strings. Rubach suggests that medial R in ORO and final R in OR are unsyllabified and, for this reason, lack a voicing value: therefore the voicing values of the surrounding obstruents in ORO, OR#O strings are in fact adjacent and this triggers assimilation. I do not pursue this line of analysis: nothing establishes that the non-syllabic sonorants of OR# or ORO strings are in fact unsyllabified (as against non-syllabic) and there is little reason to believe that unsyllabified segments will be ignored by redundancy rules, assuming that the latter exist at all (cf. Steriade 1995).

2.3. German syllabification and devoicing

The case of German devoicing is discussed because several syllable-based analyses have been offerred for this phenomenon (Vennemann 1982, Rubach 1990, Hall 1992): I will suggest that they are not supported by the evidence. A second reason to look at German is that neutralization here is clearly incomplete, with durational differences continuing to maintain a lexical contrast between {p, t, k} and {b, d, g} in stem-final position: this is relevant in the present context because the property being neutralized is the VOT, which loses its cues in the _# context. The property being maintained is the durational aspect of voicing, which continues to offer detectable differences in the _# position. Therefore, the German data strengthens the correlation proposed here between cue distribution and licensing sites. Finally, German devoicing is cyclic not only at the phrasal level, as in the languages discussed so far, but also at the stem level: devoicing affects stem final obstruents before all productive consonant-initial suffixes. The incomplete character of German neutralization and its cyclic aspects are discussed elsewhere (section 5; Steriade 1996). This section will consider only the relationship between syllable divisions and neutralization sites.

2.3.1. The facts

The facts of German have been assembled most recently by Rubach (1990) and Hall (1992). As Rubach points out (1990:83), there is no directly observable correlation between syllable positions and licensing sites, since voiced codas are found in forms like *ebnen* 'to even out' [e:b.nən] or *Ordnung* 'order' [ɔrd.nuŋ]. Rubach himself draws a rather different conclusion from this fact, but the bare fact is significant when compared to the identical situation observed in Lithuanian, Greek, Sanskrit, Polish and Russian. We have seen that the syllable is irrelevant to voice licensing in these languages: it would therefore be surprising if the very similar German facts called for a different analysis. In linear terms, the paradigm of German devoicing is - modulo the cyclic effects - identical to that of the languages discussed earlier: the stops are partially neutralized word-finally and before obstruents. We will see that there is no reason to assume that German differs fundamentally from languages analyzed so far in its characterization of [voice] licensing contexts.

The contexts of potential devoicing we have to consider for German are shown below. There are several cases of particular interest. First, we must consider OR clusters that cannot occur initially or occur only marginally there (stop-nasal, alveolar-l; class (b) below): these are relevant because they are *a priori* implausible onsets and have been reported as heterosyllabic by at least some of my sources. A second relevant class are O#R sequences (where # is a stem or prefix boundary; class (c) below); these illustrate the cyclic effects in devoicing. Finally, we will consider the OR#V sequences, where R alternates between syllabic and non-syllabic (e.g. *neblig* 'foggy' *[ne:bliç]*, *Nebel* 'fog' *[ne:bl̩]* ; class (d) below). The table in (32) provides a synoptic view of devoicing and licensing in various contexts.

Since judgments on syllabification in most classes - esp. (b) and (c) - are debated, the issue of syllabic division is separately discussed below.

(32) Contexts for German [voice] licensing and neutralization
(data from Bithell 1952, Duden 1962, 1966, Cassell 1978, Vennemann 1982, Rubach 1990, Hall 1992):

	Underlying voiced	Underlying voiceless
a. ORV (OR allowed in #_):	I[gl]u 'igloo' ~[kl]	Eklat 'altercation'
distinctive voicing maintained;	Bi[bl]iothek 'library' ~ [pl]	Diplom 'diploma'
optional devoicing reported by some.		
b. ORV(OR disallowed/ marginal in #_)	Ma[gm]a ~ [km]	Akme
distinctive voicing maintained;	A[dl]er 'eagle' ~ [tl]	Atlas 'atlas; satin'
optional devoicing reported by some.	Mö[dl]ing ~ [tl]	Rütli (place names)
	Si[gn]al 'signal' ~[kn]	Akne 'acne'
	La[dn]er ~[tn]	Eitner (proper names)
	Teu[bn]er ~[pn] (proper name)	
c. OR#V	Han[dl]-ung 'action' ~ [tl]	
distinctive voicing maintained;	Ra[dl]-er 'bicyclist' ~ [tl]	Schüttl-er 'shaker'
optional devoicing reported by some	ne[bl]-ich 'foggy' ~ [pl]	Kuppl-ung 'coupling'
	Schul[dn]-er 'debtor' ~[tn]	
	Re[dn]-er 'speaker' ~ [tn]	
	re[gn]-en 'rain' ~ [kn]	Trockn-er 'dryer'
	e[bn]-en 'flatten' ~ [pn]	
	wi[dm]-en 'dedicate' ~ [tm]	Atm-ung 'breathing'
d. O#RV	Bil[t]-nis 'picture'	
voicing neutralized	bil[t]-lich 'pictorial, graphic'	rät-lich 'advisable'
	Erlau[p]-nis 'permission'	
e. O# #	Lan[t] 'country'	bun[t]'colorful'
voicing neutralized	ga[p] 'gave'	kna[p] 'tight'
f. O_1O_2	geha[p]t 'had (ppl)'	geden[k]t 'thought (ppl)'
voicing neutralized	gesa[k]t 'said (ppl)'	gepa[k]t 'packed up'(ppl.)
	sa[k]te 'said (3sg.)'	

Devoicing in final and pre-obstruent position is predicted by all accounts considered here, and therefore offers no diagnostic value. Neutralization in the O#RV context is a cyclic effect:; this point is elaborated in section 5. It is therefore occurrence of devoicing in other VORV contexts that we must concentrate on.

Devoicing is reported in the Duden Grammatik (1966: 57) to apply optionally in the colloquial language (Umgangssprache) in all contexts: for forms like *sigmatisch, Ebne, Händler, regnet,* pronunciations like *[zi:kmatiʃ], [e:pne], [hɛntlər], [rɛknət]* are said to coexist with standard *[zi:gmatiʃ], [e:bne], [hɛndlər], [re:gnət]. g41*
Devoicing is crucially not limited to certain clusters: it applies generally, including in simplex onsets, as indicated by data from the same Duden passage: colloquial *Greis* is transcribed as devoiced *[g̊rais], Bärchen* as *[b̥ɛrçən] leide* as *[laɪd̥ə].* Therefore this variety of colloquial devoicing does not distinguish theories of voicing neutralization, being essentially context-free[17]. In the Hochsprache, on the other hand, devoicing takes place only in the contexts (d), (e) and (f) above: before obstruents, end of word and at the end of cyclic domains[18].

2.3.2. Correlations between neutralization and the syllable

We may now consider the evidence for correlating the contexts of devoicing in the Hochsprache with the location of syllable boundaries. On this point, there appear to exist two traditions. One is implicit in the syllabifications given by Cassell's dictionary and by Rubach (1990). According to Cassell's transcriptions, at least the medial stop-nasal, stop-j clusters are heterosyllabic: *Magnet, Signal, Magnesia, Adjunkt, Dogma* are transcribed as *[mag'nɛt], [zɪg'nal], [mag'nezia], [ad'jʊŋkt], [dɔgma]* (the latter with laxing in closed syllable indicating *Dog.ma*) whereas unmarked OR clusters are transcribed as tautosyllabic: *Diplom [di.'plo:m].* Rubach (1990:83) reports subject responses to the question of syllabifying OR#V sequences that coincide with Cassell's: *Handlung [hand.lʊŋ],*

---

17 Duden's report of optional devoicing in the colloquial language appears in abbreviated form in Hall (1992), where devoicing is said to occur optionally, in fast speech, in OR sequences (I[gl]u ~ I[kl]u, Du[bl]in ~ Du[pl]in). Hall , who advocates a syllable-based theory of voicing neutralization, analyzes this devoicing as the effect of an optional boundary shift (I.glu -> Ig.lu -> Ik.lu). Note however that fast speech devoicing occurs in cases where no boundary shift is possible: e.g. *Bärchen* realized as devoiced [pɛrçən].

18 We must distinguish neutralization - i.e. wholesale loss of voicing contrast - from the devoicing of individual tokens of voiced obstruents: thus *Admiral* is reported as containing devoiced [t] by the Duden Aussprache Worterbuch (1962) but appears with voiced [d] in Cassell's German Dictionary. This is an individual lexical matter, since the t/d contrast is clearly maintained before m according to both dictionaries: *Bodmen* 'floor', *Bodmerei* 'bottomry' are transcribed with voiced [d] in both of these sources, as well as Bithell (1952), Jessen (1996). Similarly, *Abner* is sometimes reported as devoiced [apner] but the b/p contrast is maintained before nasals through lexical items such as *Ebner, Grabner* (Jessen 1996). The incidence of devoicing in individual lexical items and its causes has not been studied.

Ordnung [ɔrd.nʊŋ]. ebnen [e:b.nən] Although Rubach does not consider monomorphemic strings like *Dogma*, the syllabifications he reports suggest, like Cassell's, that speakers divide the OR clusters in the same way as English and (most) Romance speakers: unmarked OR is tautosyllabic, marked OR (stop-nasal, alveolar-l) is heterosyllabic. On this view of German syllable structure, voice neutralization is unrelated to the coda position, since voicing is licensed in many codas, exactly as in the languages considered earlier.

A different, more influential view of German syllabification goes back at least to the Duden Grammatik, where the claim is explicitly made that voiced obstruents cannot stand at end of the syllable, "haben keine Silbengrenze unmittelbar nach sich" (1966:49). This view inspires the Duden to dictate syllabifications such as *Ma.gma, Re.dner, He.dschra [he.dʒra], Pilsner [pɪl.znər], fa.sre [fa.zrə]* and, presumably, *Bo.dmen, A.dler.* We can identify this tradition as the major source for Vennemann's (1972, 1982) and Hall's (1992) theories of voice neutralization in German. It should be noted however that the Duden is not endorsing a general policy of maximizing onset clusters, since for *tm, tl, km* clusters, the assignment reported is heterosyllabic: we are told to divide *Hyp.nose* not *Hy.pnose* (1966:410), *At.mo.sphä.re* (1965:561) not *A.tmo.sphä.re, Ath.let, At.lan.tik ([at'lɛt], [at'lantik]* both in 1962:136), not *A.tlan.tik, Ak.me ([ak'me:],* 1962:106) not *A.kme, Drechs.ler* (1965: 559) not *Drech.sler, Tech.netium [tʰɛç.'netsium]* (1962:729) not *Te.chnetium*[19]. Bithell (1952:375) reports that some speakers follow some but not all aspects of the Duden system: *Adler* 'eagle' is indeed syllabified *A.dler* by these speakers but *Redner* 'speaker' is syllabified *Red.ner.* The difference, according to Bithell, involves awareness of a connection to *Rede* 'speech': the syllable boundary in *Red.ner* is felt to coincide with the major morphological division. Bithell's data then also suggests that syllable boundaries are orthogonal to devoicing since the speakers that intuit *Red.ner* nonetheless fail to devoice the [d].

The truly puzzling question that has emerged from this discussion is the basis for the intuition codified by the Duden that syllable boundaries should always precede the voiced stop: why *Ma.gma* vs. *Ak.me*, why *Bo.dmen* vs. *At.mos.phä.re* , *Pil.[zn]er* vs. *Te[ç.n]etium* . Given what we know about the typology of onset clusters in other languages (Blevins 1985, Clements 1990, Ito 1986, Steriade 1982) one would expect the exact opposite assignments. Thus *dm, dn* are heterosyllabic in Attic Greek

---

[19]The unmarked OR clusters are reported as tautosyllabic at least in: *Di.plom, Zy.klus, Sa.krament* (1965:560), although for native words the division sanctioned by the Duden is such that only the last consonant in a cluster opens the second syllable: *zweifenst.rig,* (1965:559). The heterosyllabic division of *pn, km, tm, tl* is not borrowed from the Greek sources of these words and must be counted as reflecting an aspect of German grammar: the Greek tradition on how to divide these clusters varies with dialect and period, as pointed out earlier.

while *tm, tn* are possibly tautosyllabic. In Icelandic no voiced stop can form a complex onset with a following liquid: *Vg.rV but V.krV.* Why is German different?

I suggest that the intuitions of syllabification reported in the Duden reflect the use of phonetic vowel length as the unique cue to syllable divisions: vowels are generally shorter in closed syllables and this is a readily useable indication as to how clusters are divided (Maddieson 1985). Since vowels are also significantly longer before voiced obstruents, the extra length of the vowel in sequences like *Ma[:gm]a, Bo[:dm]en* may be misanalyzed as cueing syllable boundaries rather than simply voicing. It is likely that for many speakers - including Bithell's, Cassell's and Rubach's - vowel length is just one of several conflicting indications of syllable assignment and may be overridden. Morphemic composition appears to be the overriding factor in Bithell's reported divisions such as *Red.ner.* For Cassell's and Rubach's speakers, considerations of onset markedness, as well as morphemic divisions appear to have a higher priority in deciding how to divide the clusters. On the other hand, for the Duden speakers, it appears that vowel length is either the only or the decisive cue to syllabification: if the vowel is tense in V:CCV, then the following cluster cannot close its syllable. Thus, by assuming that different speakers give different weight to various diagnostics of syllabification, we can understand the nature of the variation reported in the literature without rejecting any of the judgments. The very fact that the syllable divisions are murky and variable across individuals, whereas voicing neutralization in the Hochsprache is remarkably invariant, supports the idea that no interesting connections exist between the two[20].

To flesh this out, we suggest that the variability in syllable assignment should be modelled by the use of variably ranked constraints on syllable structure, such as *[tense V] in closed syllable, *obstruent-nasal onset, Align (root, R, syllable, R) (cf. McCarthy and Prince 1993)[21]. The category [tense V] abbreviates the durational category to which vowels in pre-voiced contexts belong. An undominated condition requires tense vowels of this category to occur before voiced obstruents. The difference between the Duden division (e.g. *Re.dner, Ma.gma*), the Bithell division (*Red.ner, Ma.gma*) and the Cassell-Rubach division (*Red.ner, Mag.ma*) emerges in this case as a simple matter of ranking:

[20] This point was verified for the present study when three speakers of German with identical devoicing patterns were asked to pronounce and then syllabify the critical sequences VORV (*Dogma, ebnen, Adler, Redner*). All three speakers had voiced obstruents in their pronunciation of these words but all three reported uncertainty and variation with respect to the location of the syllable boundaries: two speakers (of Alemannic and Bavarian respectively) reported syllabification judgments identical to the Duden's except that ebnen and Redner were felt to be more likely heterosyllabic ['e:b.nən], ['re:d.nər], as reported in Rubach's study. A third speaker (of the Hesse dialect) reported heterosyllabic assignments for all clusters. All three speakers agreed that the marked OR clusters with a voiceless first member (e.g. pn in Hypnose, tm in atmospherisch) are heterosyllabic, as indicated in the Duden.

[21] I use the Align constraint here as shorthand for a set of intraparadigmatic correspondence conditions whose indirect effect is to generate the perception of aligned syllable and morpheme boundaries.

(33) Variable cluster divisions in German as the effect of variable ranking of cues to syllable structure :

a. The Duden division (*R[e:].dner, M[a:].gma*):

*[tense V] in closed syllable >> *obstruent-nasal onset, Align (root, R, syllable, R)

b. The Bithell division (*R[e:]d.ner, M[a:].gma)*:

Align (root, R, syllable, R) >> *[tense V] in closed syllable >> *obstruent-nasal onset

c. The Cassell-Rubach division (*R[e:]d.ner, M[a:]g.ma*):

*obstruent-nasal onset >> *[tense V] in closed syllable

I should emphasize, in closing this part of the discussion, that not even Duden's syllabification can be taken to support the idea of [voice] being licensed by the onset: what we have seen is that the Duden assigns the boundaries on the basis of the surface voicing of the obstruent. Therefore it is the voicing - or its consequences for the tense quality of the vowel - that induces the perception of syllable boundaries, rather than the syllable position dictating voicing possibilities. Only a circular theory of Licensing by Prosody may take comfort in the Duden data[22]. Since the surface distribution of voicing in German is either independent of the syllable (for Bithell's, Cassell's and Rubach's speakers) or actually dictates syllabification (for the Duden), we must characterize in syllable-independent terms where the voicing contrast is possible and where it's not. The same constraints that were justified for Lithuanian, Russian, Greek and Sanskrit are operative here: these are the cue-based constraints in (6).

3. A second voicing neutralization pattern: before obstruents only

The voicing contrast is frequently preserved word-finally, but not before obstruents. The analysis of this pattern will be based on three observations all of which justify the global statement that distinctive voicing is less perceptible and harder to implement before obstruents than in final, post-vocalic position. The first observation is that word-final obstruents have an advantage in the identification of the offset transitional correlates of voicing relative to word medial obstruents, since they are preceded by longer vowels: durational differences between final vowels, as well as F1 and F0 differences are probably easier to evaluate. Second, word final stops are more likely to be audibly

[22] The circularity is apparent in some characterizations of German syllable divisions: "The syllabification [...] is based on the non-application of Devoicing to the medial obstruent." (Hall 1992:89). Similarly, when discussing forms like *Ordnung*, whose voicing fluctuates ([ɔrdnuŋ] ~ [ɔrtnuŋ]) Hall attributes the fluctuation to differences in syllabification and states (1992:92-93): "Phonetic evidence for the two possible syllabifications [*Ord.nung and Or.dnung*] is that [...] voiced obstruents can surface as either voiced or voiceless." In other words, the devoicing is caused by the syllable structure, and the phonetic evidence for the syllable structure is the devoicing itself. Merchant (1995) presents a variant of this analysis under which the obstruent in clusters like *dn* is ambisyllabic.

released than stops in word internal stop-obstruent clusters: the quality of the burst itself may cue voicing or voicelessness (Lisker 1986; though see Hillenbrand et alii 1992). Third, the perception of voicing in word-medial obstruent clusters (e.g. *agta, akda*) may be determined by the laryngeal category of the second, prevocalic obstruent, as documented by Slis (1986): *agta* may be perceived as *akta* and *akda* as *agda*. We can speculate that this is because the second obstruent benefits from the more informative release and VOT cues, whereas the first lacks them[23]. In contrast, the perception of voicing in a word-final stop preceded by a vowel (e.g. *ag*) will not be affected by the voicing of any other segment. In addition to this perceptual advantage of word final obstruents, there is an articulatory one: in a word-internal obstruent cluster whose members disagree in voicing (e.g. *agta, akda*) a fast transition is required between opposite states of the glottis (Hsu 1996). This is not the case when the obstruent stands alone at the end of the word. All these points justify the ranking of the two contexts V_# and V_[-son] on the perceptibility scale in (6). As anticipated earlier, there are no cases in which voicing is licensed before obstruents but not word-finally: this unattested pattern of devoicing is one that our ranking schema cannot in fact generate. There are numerous neutralization patterns of the sort anticipated in (7.ii), where voicing is neutralized before obstruents but not word-finally after vowels. They are discussed in this section.

Although this second voicing neutralization type has been documented (Mascaró 1987, Cho 1990, Lombardi 1991, 1995) the cases discussed in the recent literature involve only languages in which the site of neutralization could be roughly identified with the medial coda. This is the case with most contemporary Romance languages: French, for instance, has regular voicing neutralization before obstruents (e.g. *absent [apsã];* Dell 1995) but maintains the contrast word finally (*laide [lɛd]* 'ugly-fem' vs. *Lette [lɛt]* 'Latvian'; *laisse [lɛs]* 'let' vs. *lèse [lɛz]* 'injures'). Thus, non-final codas are mostly neutralized in French and only phonotactically marginal forms like *dogme [dɔg.mə]* 'dogma' show that the licensed voiced stops can also be medial codas.

The languages considered in this section neutralize voicing in the same contexts as French but are markedly different in syllable structure: I present here Hungarian and Kolami data that support a purely linear, syllable-independent analysis of this type of voicing neutralization. Both languages lack complex onsets, both initially and medially. Both languages preserve voicing before heterosyllabic sonorants as well as word-finally.

The Hungarian data (from Vago 1974, Njeki 1988) below shows neutralization before obstruents as well as *h*, a voiceless sonorant. This is exactly what we might expect if the absence of

[23]This reasoning is modeled on Ohala's (1990) results involving the primacy of release place cues over V-C transitions.

distinctive VOT values is a relevant factor in defining the context of voicing neutralization: a non-modal sonorant like *h* or or *n̥* or *w̰* will delay the onset of modal voicing and thus act like an obstruent in removing the VOT cue. No neutralization is observed before consonantal sonorants despite the fact that all CC clusters are heterosyllabic:

(34) Hungarian: a. Neutralization before obstruents

abcug	[apt^{s}ug]	'resign!'
habcsók	[hapčo:k] '	meringue'
lakzi	[lagzi]	'wedding'

b. Neutralization before h:

hívhat	[hi:fhat]	'he may call'
alábbhagy	[ala:p:haɟ]	'diminishes'

c. No neutralization word-finally:

rab	[rɔb]	'prisoner'
kalap	[kɔlɔp]	'hat'

d. No neutralization before heterosyllabic R:

vedmeg	[ved.meg]	'buy it !'
halottnak	[halot.nak]	'death-elative'
átmegy	[a:t.meɟ]	'to cross
továbmegy	[tova:b.meɟ]	'to go forward'

The Kolami pattern is identical (Emmeneau 1955):

(35) Kolami a. Neutralization before obstruents:

va:ŋk-tan 'I poured' va:ŋg-dun 'I was pouring'

b. No neutralization word finally :

novvod '90' kudug 'thigh', ga:z 'bangle', sa:yeb 'sahib'

c. No neutralization before (heterosyllabic) sonorant:

voiced	voiceless
saye:b.na 'of the sahib'	teḍep.ne 'of cloth'

bag.li.ak 'white heron'	paṭ.lak 'sharp edge'
sib.le 'man of exogamous division'	kop.li 'mouth'
saye:b.ral 'European woman'	tok.re 'shell'
sa:z.re 'good'	tis.re 'third'
pand.ri 'bathhouse'	kebut.ri 'pidgeon'
ad.ne 'of her'	put.niak 'brother's son'
	paṭ.lak.net 'of headman'

Voicing neutralization can be modelled in both of these languages as anticipated in (9.ii) above: the critical ranking is shown below.

(36) $*\alpha$ voice/ __[-son]

↓

Preserve α voice

↓

$*\alpha$ voice/ __#

↓

$*\alpha$ voice/ __[+son]

(For Hungarian, the analysis in (36) is incomplete as the role of h is ignored. I defer the discussion of modal and non-modal sonorants as laryngeal licensing contexts to Part II.) In closing this section, we should consider two alternatives to the analysis proposed here. One is the possibility of accounting for the behavior of final obstruents by manipulating the boundaries of the prosodic word, in line with suggestions by Inkelas (1987): if the word-final consonant falls outside the boundaries of the prosodic word, then one might think that the context _# does not characterize its position. The diagram in (37) illustrates this mode of analysis (cf. also Lombardi 1991, 1995).

(37) Final C extraprosodicity:

```
morpho-syntactic word
/ / | \ \\ \
CVCVCV Cf
\ \| / //
prosodic word
```

C_f in (37) is not at the right edge of the prosodic word and therefore no constraint mentioning that site will be applicable to C_f. This then appears to provide an alternative account for the difference in voice licensing between Russian-German and Hungarian-Kolami: C_f may be said to be effectively final in Russian but not in Kolami, if the representation in (37) applies to the latter. However, this suggestion is unworkable for practically all languages of the Kolami class: most of these impose minimality conditions on the size of the prosodic word, conditions that can be met only if C_f is counted in. Hungarian and Kolami lack C_0V content words, a fact normally interpreted as a symptom of undominated minimal size requirements. If C_f is excluded from the prosodic word in items like [rɔb] - to describe its voicing - then we cannot explain the absence of equally sized words like *[rɔ]. Further, if the option to ignore C_f is introduced, then one predicts patterns of neutralization in which final obstruent clusters (i.e. VO_1O_2#) require O_1 to be neutralized but not necessarily O_2: this means hypothetical contrasts like *makt* vs. *makd.* Such cases are unattested (Greenberg 1968; cf. Cho 1990 for discussion). Finally, C_f is not systematically extra-prosodic in either Hungarian or Kolami: it triggers epenthesis in both languages when the word ends in an impermissible cluster. This damages even more the prospect of describing the voicing facts by declaring C_f selectively extraprosodic.

The second alternative is that pursued by Cho (1990), who distinguishes syllable-based from cluster-based laryngeal neutralizations (cf. also Rubach 1996). For Cho, German exemplifies the syllable-based pattern: but we have seen in section 2 that the facts do not support this option. Hungarian can be analyzed, in Cho's terms, as cluster devoicing. What remains unexplained, under this theory, is why the segmental environments of voice neutralization are so remarkably similar in languages with cluster-based and syllable-based neutralization: German-Russian-Polish-Greek-Sanskrit-Lithuanian and Hungarian-Kolami-French differ only, as far as voicing is concerned, in the status of word-final obstruents. It is unjustified to invoke fundamentally different neutralization mechanisms for these two patterns. Our analysis accounts for this minimal difference in terms of a minimal ranking change, (9.i) vs. (9.ii).

Thus the only viable analysis seems the one proposed here: voicing in French, Hungarian and Kolami is licensed word finally, not before obstruents, because the V_# position is superior to the V_O context in the range of cues it offers for the detection of contrastive voicing. Equally important is the fact that this analysis is the only one that explains the data by reference to independently observable facts: the different cue-to-effort ratios required for implementing contrasts in different contexts. In contrast, alternative analyses either fail to provide the principles of a general typology of voicing neutralization or else rely on representational properties like (37), the only evidence for which is the very data they are meant to derive.

4. Further patterns of voicing neutralization

The cases considered so far indicate that many apparent instances of coda devoicing should be reanalyzed as final or pre-obstruent neutralization. One argument against syllable-based statements in such cases was that codas can possess distinctive voicing, when they are followed by a sonorant, as in Polish *bez.alkoholowy,* Sanskrit *tig.ma.* or Hungarian *ved.meg.*

This observation raises however some further questions. First, are there cases of voicing neutralization before all consonants, regardless of obstruency: are there languages where inputs like *tig.ma* are realized as *tiK.ma* with a laryngeally neutralized stop in pre-sonorant position? Can such patterns be analyzed without reference to the syllable? Should they be so analyzed? Second, do we also find instances of voicing neutralization in the onset? We predict that such cases may occur if certain onset positions present diminished cues to voicing relative to the contexts where voicing is maintained.

This section takes up these questions with a view to expand our inventory of voice neutralization cases beyond the two positions (_ #, _ [-son]) studied so far.

4.1. Neutralization in the absence of following vowel

Voicing is reportedly neutralized in a number of languages in final and all pre-consonantal positions. Wantoat (Davis 1969), a New Guinea language, contrasts two stop series - {p, t, k} vs. {b, d, g} - prevocalically but allows only one undifferentiated class in other contexts. The consonant clusters contain at most two members, the first of which is always transcribed as voiceless, regardless of what follows: *jak.ŋa* 'leaves', *u.jap.ma* 'my younger sibling', *pa.kap.zon* 'you all bring them', *put.da* 'let us two break them'. The clusters *pn* and *tn* are also attested with voice-neutralized obstruents. A similar pattern is attested in West Tarangan (Nivens 1992), a language of Indonesia: the voiced-voiceless contrast in stops is lost word finally and before all consonants: thus intermediate *ke-b-laba* 'plank' is realized as *[kep.la.ba]* (p. 220) while *mata-b-sebar* 'eye discharge' surfaces as *[ma.tap.se.bar]*. As in Wantoat, final stops are neutralized to voiceless: *pit* 'night', *guk* 'suck', *tɔp* 'short. Dialects of Quechua are also reported to possess this sort of neutralization.

In these languages, it appears that voicing is indeed limited to onsets: a heterosyllabic sonorant does not help maintain the contrast, since sequences like *bl, bn* are impossible. Before concluding that onset-licensing remains a necessary ingredient in the analysis of laryngeal neutralization, we must note that laryngeal contrasts are sometimes neutralized before consonantal sonorants even in languages where the relevant clusters form complex onsets. While Tarangan and Wantoat simply lack these, languages like Pacoh (Mon-Khmer: Watson 1964) and Sre (Mon-Khmer: Manley 1972) have complex onsets of the form stop-liquid but limit laryngeal contrasts to the immediately pre-vocalic position, as seen below. All consonants, including the voiced stops, occur as simple onsets in Pacoh but only the plain voiceless stops occur finally and pre-consonantally.

(38) a. Pacoh consonants:

p	t	tʃ	k	ʔ
	s			
b	d	dʒ	g	
m	n	ɲ	ŋ	
w	l	r	y	

b. Pacoh clusters:

(i) onsets:

pʰ	tʰ		kʰ
pr	tr	tʃr	kr
pl			kl

(ii) $C_1.C_2(C_3)$: {l, r, m, n, ɲ, ŋ} + { any onset)

It is also possible to view Pacoh as possessing unit aspirated stops, rather than stop-h clusters. Either way, the conclusion that *obstruent laryngeal contrasts are allowed only in prevocalic position* is clearly warranted: only this explains the systematic absence of onset clusters like *br, bl, pʰr, pʰl.* What is significant here is that laryngeal neutralization occurs both in codas (only plain voiceless stops are allowed word finally) and in complex onsets (when a liquid follows the stop). Coda devoicing can be invoked, but is neither sufficient nor necessary to characterize the distribution of voicing and aspiration. The Sre facts discussed by Manley (1972) are comparable: a three-way laryngeal contrast among the stops (plain, voiced-implosive, aspirated) is allowed only prevocalically. Complex onsets include laryngeally neutralized stops in C_1 position. All other pre-consonantal or final stops are also neutralized. Mikir (Grüssner 1978) is said to contrast voiceless unaspirated, voiced and voiceless aspirated stops before vowels, but only plain and aspirated stops before liquids. The partially neutralized Mikir stop-liquid sequences are tautosyllabic. All laryngeal contrasts are lost word finally and before obstruents. Sherpa (Tibeto-Burman: Hale 1970) exhibits the same properties, even though complex onsets seem limited to word-medial position[24].

The Tarangan and Wantoat facts now appear in a very different new light: what looked at first like coda devoicing in these languages turns out to be an instance of a more general phenomenon of laryngeal neutralization caused by the absence of a following vowel. We conclude that this process is just as unrelated to syllable positions as the devoicing facts studied in earlier sections.

---

24 The survey of laryngeal neutralization on which this study is based includes languages where voicing but not aspiration contrasts are neutralized pre-consonantally - including before tautosyllabic sonorants - as well as languages where aspiration is neutralized in the same positions but not voicing. I do not understand the basis for this variation and would conjecture that it may involve differences in transcriptional practices rather than actual production. Chepang (a Tibeto-Burman language; Caughley 1970) contrasts before vowels voiceless, voiced, aspirated and voiced aspirated stops, but limits the contrast to voiced-voiceless before liquids and neutralizes to a unique voiceless series in all other contexts. As in Mikir, the stop-liquid clusters of Chepang form complex onsets, indicating that this partial collapse of laryngeal categories may target the onset.

What remains unclear are the factors differentiating the pre-vocalic position - where distinctive voicing and post-aspiration are always permitted - from the pre-liquid, pre-nasal position, where some languages neutralize voicing/aspiration. One possibly relevant observation is made by Docherty (1992: 44ff) who notes significant increases in VOT between a given stop class before non-syllabic sonorants (OR) as compared to the same class in pre-vocalic position (OV). While in English this increase in VOT caused by the following sonorant is non-neutralizing, it is conceivable that in other systems this phenomenon renders the VOT values harder to interpret as voicing cues in the OR context, precisely because the voicing lag is different in the OV and OR strings. This is pure speculation. What is important is that, whatever the source of the difficulty in maintaining a voicing/aspiration contrast in OR sequences, it has nothing to do with syllables.

We may ask then: what would count as genuine evidence for syllable final devoicing? The simple answer is: any system that allows us to compare voicing maintenance in onset OR sequences with voicing neutralization in heterosyllabic O.R[25]. Thus the hypothetical language in (45) - inspired by but critically distinct from French - distinguishes voiced obstruents in the OR sequences functioning as onsets, but neutralizes voicing in every other obstruent-C sequence, including in heterosyllabic OR. The reader will find the comparison between fictitious and real French quite instructive since the facts of the real language are unintelligible under prosodically based analyses of voice neutralization[26]:

(39) a. Fictitious language similar to French: voicing neutralized in all codas — b. Real French: voicing neutralized before O

	a.		b.	
onset obstruents	a.bri vs. a.pri		a.bri vs. 'shelter'	a.pri 'learned'
	a.vrə vs. a.frə		a.vrə vs. 'haven'	a.frə 'terrors'
	e.klə vs. e.glə		sjɛ.klə vs. 'century'	ɛ.glə 'eagle'

[25]Cases like German *Liebling* (with heterosyllabic *b.l* and a laryngeally neutralized stop [li:plɪŋ]) require a cyclic analysis, rather than coda devoicing: this point is discussed in section 5.

[26] Catalan is cited by Cho (1990:150) as possessing a pattern of voice neutralization similar to the fictitious language in (39). Cho analyzes Catalan on the assumption that this is a syllable-based neutralization. However, the data cited is very limited and involves segmentable prefixes like *sub* (*supmari* 'submarine') whose devoicing can also be attributed to a cyclic effect. Another language cited by Cho as possessing syllable-final devoicing - Dutch - turns out to pattern like Lithuanian: obstruents are neutralized word-finally and before obstruents, but not before heterosyllabic sonorants, in forms like *Ariadne* (vs. *Etna*) or *Abner* (vs. *hypnose*). This observation is due to Harkema (1997).

coda obstruents	doK.mə	dog.mə vs. 'dogma'	ak.me 'peak'
	maT.len	mad.lɛn vs. (name)	ʃat.lɛn castle-lady
	aK.sa	ak.sã 'accent'	

We may conclude this section by reiterating the essential point: coda neutralization is insufficient, unnecessary and inadequate as a general account of the voicing typology. It is an insufficient account for languages like Pacoh, where onset stops are also neutralized preconsonantally; it is unnecessary for languages like Wantoat, whose devoicing patterns can be characterized in linear terms identical to those needed for Pacoh: all obstruents neutralize unless prevocalic. Finally, coda devoicing is drastically inadequate for all other languages studied so far, since the codas in O.R sequences do not neutralize in any of them. This class of cases is by far the most abundantly documented.

4.2. Onset devoicing

The existence of onset neutralization in Pacoh requires us to turn now to one of the key predictions of the model presented here. We have focussed so far on two classes of contexts where the voicing contrast necessarily lacks VOT and other release cues: these are the final and pre-consonantal position. In both of these contexts a stop is most likely to be syllabified as a coda. But certain unambiguous onset positions can also be identified in which the perception of voicing is made difficult by the absence of - or difficulty in implementing - other voicing cues. We predict for such cases as well a greater likelihood of neutralization relative to the contexts in which all voicing cues are equally available: the fact that these are onset positions is immaterial.

Thus utterance initial as well as post-obstruent onsets will lack all cues normally residing in the preceding vowel or sonorant (V_1 duration, F_0, F_1 values at the onset of closure) and may lack the closure duration cue. Therefore the #_V context is less likely, ceteris paribus, to maintain voicing contrasts than the V_V context. Since the utterance-initial properties of words are frequently extended to the utterance-medial realizations, we have here the potential source of word-initial voicing neutralization. Moreover, vocal cord vibration during the obstruent's closure will be more difficult to insure in utterance-initial (and, by paradigmatic extension, word-initial) position, because of insufficient subglottal pressure (Flege 1982, Flege and Brown 1982, Westbury and Keating 1985). This is a second reason why we might expect word-initial stops to be targetted for voicing neutralization. After a voiceless obstruent, voicing during closure will also be difficult to implement: in a cluster such as *asda*, if *s* is truly voiceless, vocal cord vibration can be obtained for *d* only through precisely timed articulatory maneuvers that create a sudden transglottal pressure drop. As Westbury and Keating (1985:162) point out, "expenditures of the latter sort are articulatorily costly." If such efforts are not undertaken, then all post-voiceless stops will lack voicing during closure. The investigations of Flege (1982), Docherty (1992) for English and Jessen (1995) for German indicate that in languages where the voicing contrast is maintained in initial and post-obstruent stops, it is maintained mostly without the benefit of the closure voicing cue.

With this in mind, let us consider the fate of the voicing contrast in the word-initial and post-obstruent context. We limit the discussion to two contexts where voicing is in principle identifiable on the basis of VOT distinctions and other release-bound cues: #_(R)V and O_(R)V. Consider now the comparison between the #_(R)V context and the V_(R)V context. We have seen that the #_(R)V context benefits from fewer cues to voicing relative to the V_(R)V context; and also that the closure voicing cue, which is potentially available in both contexts, is articulatorily harder to obtain in #_(R)V than in V_(R)V. Therefore a cue-based model of neutralization predicts the existence of systems where distinctive voicing is maintained intervocalically but lost initially. The conflicting preference for deploying feature contrasts word-initially (Bosch 1992; Casali 1995a, b; Hsu 1995; MacEachern 1995; McCarthy and Prince 1995; Steriade 1993, 1995) softens somewhat the strength of this prediction and explains why initial neutralization is considerably less frequent than one might otherwise expect. But the prediction stands.

That initial voice neutralization does exist has already been pointed out by Keating et al. (1983) and Westbury and Keating (1985). Here are two cases not included in their survey. Lac Simon (Kaye 1979, 1981, Iverson 1983), an Algonquian language, has an intervocalic voice contrast among obstruents, and voice-neutralized obstruents initially. The initial neutralization affects loan words: [pa:na:n] is Lac Simon for *banana*[27]. In addition, Lac Simon neutralizes voicing word finally and in all obstruent clusters: attested obstruent sequences are *sk, sp, ʃt, ʃk, ʃp*. The other case is found in Totontepec Mixe (Crawford 1963), a Mixtecan language of Oaxaca. The Totontepec obstruents contrast for [voice] intervocalically but not initially. The clearest reflex of this limitation involves the fricatives: *ş* and *ʐ̧* contrast in *wʌʐ̧oy* 'embroidered' vs. *nʌşoya* 'shirt' but only ş occurs initially, finally, as well as before or after obstruents: *şu:şpa* 'musician', *tʌdʌpuş* 'he already cut it', *mnʌhkşup* 'you're going'. The same holds for all other obstruents, with the only difference that voiced stops are lenited intervocalically, hence VtV contrasts with VðV rather than VdV.

[27]The neutralized initial stops of Lac Simon are realized as voiced when preceded by a vowel across certain boundaries. We discuss this effect in section 5.

Note that this pattern of neutralization is also impossible to characterize in syllabic terms: what neutralizes in Lac Simon or Totontepec are all the codas plus a subset of the onsets. The initial neutralization can be straightforwardly analyzed by observing the difference between V_V and #_V contexts on the perceptibility scale, with #_V inferior to V_V. This is mirrored by the ranking of constraints: *α *voice/ #_V* >> *α *voice/ V_V*. What distinguishes Lac Simon and Totontepec from the languages studied so far is the fact that *Preserve voice* ranks in this language below *α*voice/ #_V* but above *α *voice/ V_V*. Both contexts mentioned are relatively favorable to the identification of voicing, but this class of languages apparently maintains voicing only under the most favorable circumstances, namely only when all internal and transitional cues are easily implemented.

A different instance of onset neutralization is progressive devoicing, a process attested in Basque (Hualde 1991, Artiagoitia 1993): after a voiceless sound, all obstruents surface as voiceless. The directionality of the process can be determined by observing the [+voice] value of the second obstruent after vowels or after **r**, contexts where distinctive voicing is maintained in Basque. Progressive devoicing applies both word-internally and across the boundaries of certain function words, but does not normally affect the initial of content words. The voicing contrast in Basque is limited to the contexts #_V, V_V and R_V: no voicing distinctions occur finally or in obstruent clusters. Progressive devoicing is illustrated below:

(40)	a.	-garren	(ordinal morpheme)	(Artiagoitia 1993: 267)
		amar-garren	'tenth'	
		/boçt-garren/ [boçkarren]	"fifth'	
	b.	da	"is"	
		laguna da	"it's a friend"	
		es ta [esta]	"is not"	

The same progressive devoicing process results in incomplete and/or variable neutralization in Dutch (Slis 1986) and for certain German speakers (Jessen 1995). The German voiced stops following a voiceless sound (as in the *[kd]* of *Ma[kd]eburg*) are realized without closure voicing but maintain a shorter VOT sufficient to distinguish them from the underlyingly voiceless stops. The German data confirm that the primary cause of post-voiceless neutralization is the difficulty in implementing closure voicing, possibly combined with the absence of offset cues such as V_1 duration. What happens in Basque is then simply the categorical version of the same process.

The analysis of the Basque pattern appears in (41). The constraints in the left column are projected by an expanded version of the voice perceptibility scale in (7), one which includes the difference between implementing voicing in the optimal V_V context vs. the #_V or [-voice]_ V contexts. The difference between V_V and #_V or [-voice] _V involves the possibility of generating and sustaining the closure voicing cue. This difference accounts for the ranking between (41.v) and (41.vi). The right hand column in (41) contains two **Preserve voice** constraints. The second, ranked at the top, is the more specific condition requiring preservation of voicing in word- initial position: **Preserve voice in #_** . The tendency to preserve *all* attributes of the word's left edge is well attested: this is formalized here as the ranking **Preserve voice in #_ >> Preserve voice** (cf. Casali 1995, Beckman 1995, MacEachern 1995, McCarthy and Prince1995). I assume here that the Basque function words subject to progressive devoicing are part of the prosodic domain defined by the preceding content word (Selkirk 1995): therefore /da/ in /es-da/ is in effect word-medial and the constraint **Preserve voice in #_** does not protect /d/'s voicing in this context.

(41) Voice Perceptibility Conditions — Faithfulness Conditions

(i) *αvoice / [-son] __[-son], [-son] __#, #__[-son]
↓
(iii) *αvoice/ V_ [-son]
↓
(iv) *αvoice/ V_ # ↘ Preserve #[αvoice]
↓ ↙ ↓
(v) *αvoice / #_ V, [-voice] __ V
↓ ↘ ↓
↓ Preserve [αvoice]
↓ ↙
(vi) *αvoice/ V_ V

The hierarchy in (41) will derive the following aspects of the Basque voice contrast: (a) neutralization before obstruents and word-finally is induced by the ranking *αvoice/ V_ [-son] >> *αvoice/ V_ # >> Preserve [voice]; (b) preservation of voicing in the #_ V context is due to the ranking Preserve #[αvoice] >> *αvoice / #_ V, [-voice] __ V; (d) neutralization of voicing in post-voiceless contexts derives from *αvoice / #_ V, [-voice] __ V >> Preserve [αvoice]. I postulate a single constraint *αvoice / #_ V, [-voice] __ V - whose source is the difficulty of implementing closure voicing in these contexts - and derive the difference between word-initial and post-voiceless positions by letting this constraint interact with the principles of faithfulness (Preserve [α voice]) and positional faithfulness (Preserve #[αvoice]). I have not offerred here any evidence for the ranking between the constraints (iii) and (v); the facts of Basque could have been generated under alternative rankings. What may justify this aspect of the hierarchy is the observation that initial and post-voiceless neutralization is known to happen only in languages like Totontepec or Lac Simon, where final and pre-obstruent stops are also neutralized. And similarly, post-voiceless neutralization happens in Basque, where final and pre-obstruent stops are neutralized as well. A more general way to formulate the conjecture that underlies the ranking of *αvoice constraints in (41) is that voicing is less perceptible in contexts lacking onset cues (burst-plus-transitions) than it is in contexts lacking closure voicing and the offset cues.

5. Cyclicity, uniformity and related effects in voice neutralization

Most voice neutralization processes are word-bound: the obstruent is neutralized by reference to word boundary domains, regardless of potential cues to voicing that might lie outside of its word domain. The typical case, that of Lithuanian phrases like *dauk akmens* has been mentioned earlier: the underlying */g/* of */daug/* is word-final and neutralized here despite the fact that a vowel follows that would have permitted recovery of the voicing correlates[28]. The converse case occurs as well: Lac Simon obstruents are neutralized word-initially as well, even when preceded by vowels within the phrase.

The suggested analysis for these facts will appeal to the notions of morpheme invariance or paradigm uniformity that have recently resurfaced in an OT context (cf. Kiparsky 1970, 1968; and Benua 1995, Flemming 1995, Kenstowicz 1995, Steriade 1995, 1996). I suggest that in the languages where obstruents neutralize at word edges, regardless of the larger syntactic context, their realization is determined by the interplay between phonetic implementational factors and constraints promoting morpheme invariance or minimization of allomorphy. More specifically, the suggestion is that the voice-neutralized word edge takes on invariably *the form that would be phonetically natural in the citation form.* Consider again the case of Lithuanian *dauk*. In the citation form of this word, the underlyingly voiced /g/ occurs utterance finally, where onset cues to voicing cannot be detected. We assume it is neutralized utterance finally for this reason, in virtue of the ranking in (9.i), as shown earlier. All utterance final consonants show a general tendency to devoice, due to the loss of subglottal pressure (Westbury and Keating 1985). The neutralized utterance-final stops - which lack a voicing specification - will therefore be subject to passive devoicing and perceived as voiceless. Constraints promoting paradigm uniformity are then responsible for generalizing the properties of the citation form to utterance-medial positions. In the case of */daug/*, the devoiced *[k]* is extended to those allomorphs of /daug/ that stand in a productive paradigmatic relation to the citation form. For our purposes, simplifying somewhat, we'll assume that this type of paradigmatic extension takes place only between word-forms. Therefore citation form properties will not be extended to allomorphs created through affixation (e.g. *daug-a*) , but they will be extended to utterance-medial, word-final instances of /daug/. This accounts for phrases such

---

[28]It should be noted that voicing distinctions *can* be licensed by the larger phrasal context. In Czech, for instance, distinctively voiced stops occur before a sonorant, even across word boundaries, as in /naro:d roste' 'the nation grows', which may be realized as [naro:d roste] (Kucera (1961:59).

as *dau[k] akmens.* The specifics of this proposal are justified elsewhere (Steriade 1996). One relevant constraint is (42).

(42) **Paradigm Uniformity (right edge)** - abbreviated **PU edge**
Assume that the string Σ represents the last demisyllable in the citation (utterance-final) form of morpheme μ; and that the string Σ' represents the correspondent of Σ in a word-final, utterance medial position; then Σ and Σ' must be identical in feature composition.

To obtain invariably neutralized word-final stops, as in *dauk akmens,* regardless of utterance position, we must let **PU edge** outrank **Preserve [voice]** in the constraint hierarchy of (7.i). The effect of this ranking is illustrated below. We evaluate simultaneously the citation form of */daug/* and the realization of */daug/* phrase internally in forms like */daug akmens/.* Each mini-paradigm is a distinct candidate. We will correspondingly change our assumptions about the meaning of the context __# : henceforth, we assume that _# means final in the utterance, rather than followed by the end of the word. End-of-utterance properties may become invariant end-of-word properties through the effect of uniformity conditions like (42).

(43) PU edge , *α voice/V_ # >> Preserve [voice]

		PU edge,	*α voice/V # >>	Preserve [voice]
	daug daug akmens		!*	
√	dauk dauk akmens			* *
	dauk daug akmens	!*		*

We must consider now the difference between Lithuanian, where neutralized stops are invariably realized as voiceless, and languages like Sanskrit in which the word-final stops are neutralized, but their realization varies with the syntactic context. I discuss this case below, because it continues our argument that references to the syllable are unnecessary in laryngeal neutralization, and because it sheds some light on an assumption made earlier that laryngeally neutralized consonants differ from distinctively voiceless unaspirated stops.

Consider a Sanskrit root like */d^hag^h/* 'reach to' realized as *[d^hak]* in pre-pausal context. This word possesses a voice-neutralized final stop in all phrasal contexts where it occurs. This fact is to be accounted for in the terms suggested above: the citation realization of the stop is extended to all other syntactic contexts. However, unlike Lithuanian *k* in *dauK*, the neutralized Sanskrit *K* is contextually variable: all word-final stops, regardless of derivational origin, are realized as voiced before voiced segments, and as voiceless before voiceless ones. As the data below indicates, this contextual voicing process affects only the neutralized stops, not the distinctively voiceless ones.

(43) Laryngeally neutral stops in sandhi (data from Allen 1955)

	Underlying	Phrase final	Before sonorant	Before obstruent
a.	/triṣṭubʰ/	triṣṭu[p]	triṣtu[b] nu:nam	triṣṭu[p]-tu
b.	/arwa:c/	arwa:[k]	arva:[g] ra:dhah	
c.	/gamat/	gama[t]	gama[d] wa:jebhih	
d.	/ṣaṣ/	ṣa[ṭ]	ṣa[ḍ] aṣi:tayah	
e.	/tat/	ta[t]	ta[d] asti	

	Word-internal, non-neutralized stops no voicing assimilation	Word-final neutralized stops voicing assimilation
f.	stabʰnoti, apnuvanti	triṣtu[b] nu:nam

Let us first understand the variability in the realization of neutralized stops. The puzzle here is this: we claim that the word-final stop is neutralized utterance-finally in a form like *triṣṭu[p]*, and that the neutralized stop is generalized through **PU edge** to other positions within the utterance. But **PU Edge** requires featural identity between utterance final *[p]* and its utterance-medial correspondents: therefore **PU Edge** should inhibit voicing in *triṣtu[b] nu:nam*, contrary to fact. A general constraint inducing voice assimilation and outranking **PU Edge** cannot be invoked here because, as the data in (43.f) shows, only neutralized stops assimilate.

The solution will invoke the distinction between auditory and articulatory features (Flemming 1995). I suggest that paradigmatic conditions like (42) may require uniformity either with respect to auditory properties or with respect to articulatory properties. In languages like Lithuanian, the **PU Edge** condition requires *auditory* identity between the devoiced utterance-final stops and their utterance-medial corespondents. This means that all word-final stops must *sound* voiceless, even though this result will have to be obtained through different articulatory means in different positions of the utterance. In languages like Sanskrit, I suggest that the paradigmatic uniformity effect involves articulatory posture rather than auditory results: all word-final stops, regardless of utterance position, must possess articulatory representations that lack a glottal target. In the gestural terms of Browman and Goldstein (1992) word-final stops must be identical to utterance final stops in possessing an articulatory scores whose glottal tier is empty of all action. The result will be, as suggested by Hsu (1995) for Taiwanese, that word-final stops will be realized with laryngeal states interpolated from neighboring segments. When surrounded by voiced segments, the neutralized stops will be realized with an adducted glottis: the pre-sonorant voicing shown in (43) is due to this interpolation effect.

The same proposal may account for the voice alternations observed by Kaye and Iverson in Lac Simon, the Algonquian language in which word-initial obstruent are voice-neutralized. The word-initial stops are normally realized as voiceless, but they voice when prefixed with a vowel. This is the mirror image of the Sanskrit alternations in (43): the two languages share the fact that only positionally neutralized stops are subject to contextual voicing.

The last remark on the cyclicity effects in voicing neutralization involves the German paradigm mentioned section 2.3. The German datum in need of explanation is the difference between vowel initial and consonantal sonorant-initial suffixes. The latter induce final devoicing, the former do not. Compare the forms in (44):

(44)	word-final	before +V	before -r, -l or -n suffixes
	Bil[t]	bild-en	Bil[t]-nis, bil[t]-lich
	ga[p]	geb-en, Ergeb-ung	Erge[p]-nis

We assume that the voiceless stops in *Bil[t]nis* and *Erbe[p]nis* are PU effects: the stops will have voiceless realizations when utterance final and voicelessness is generalized to other allomorphs. But then it appears that we should also expect **bil[t]en, *ge[p]en,* by the same argument. A possible solution will take the following form. Paradigmatic uniformity is more stringently enforced between allomorphs that are already very similar to each other, in phonological or morphosyntactic properties. This point is illustrated abundantly in the literature on analogical extension (cf. Hock 1986 for review). Conversely, allomorphs that are necessarily different in one respect, accentually or segmentally, are under less pressure to be strictly identical in their other phonological properties (Steriade 1997). Note now that the stem final stops in (44) are syllable-final in word-final and pre-consonantal position (*ga[p], Erge[p].nis*) but syllable-initial before V-initial suffixes (*Er.ge.[b]ung).* It is then possible that the paradigm uniformity condition responsible for the distribution of voicing in German stem-final stops, requires featural identity only between stops that occupy the same position in the syllable: this will require then that the coda labial stops be featurally identical, but will allow onsets (as in *Er.ge.bung*) and codas (as in *ga[p]*) to differ in voicing.

Note that this analysis invokes the syllable only to require that corresponding segments with identical syllabic positions also be featurally identical. Syllable assignment is not invoked as the explanatory factor in the distribution of voicing.

6. Summary to this point and transition to Part II

The discussion so far has established several points. First, voicing neutralization is not driven by syllabification in any context. We have not encountered a single genuine instance of devoicing that can be said to be caused by the coda assignment of the obstruent. We have observed languages with different syllable structures but identical linear contexts for voicing neutralization (Lithuanian vs. Sanskrit; French vs. Kolami). We have also observed that fluctuations in the syllabic assignment of OR clusters (V.ORV ~ VO.RV) are not accompanied by fluctuations in the application of voice neutralization. Distinctive voicing is frequently maintained in codas (e.g. Lithuanian *skob.nis*) or lost in onsets (e.g. Lac Simon *pana:n*), thus disconfirming in multiple ways the predictions of Licensing by Prosody. Further, we have observed an implicational relation between contexts of neutralization (e.g. if devoiced in _# then devoiced in __[-son] but not vice versa) which relates clearly to cue distribution rather than prosodic assignment.

These findings are all consistent with the hypothesis of Licensing by Cue. As noted earlier, however, there are several ways in which contrast perceptibility, the factor identified here as the key to neutralization, may be reflected in the grammar. So far, we have provided only descriptions based on statements like ***[α voice]/ #_ [-son],** which identify indirectly, in standard featural vocabulary, a context prone to neutralization. There is no mention of the missing cues in the description #_ [-son]. The alternative yet to be explored is that the grammatical statements refer directly to the quality of the information provided by the context: for instance, a conceivable substitute for ***[α voice]/ #_ [-son]** is ***[α voice]/ in contexts where voicing lacks transitional cues** . This is a plausible interpretation in the context of our claim that ***[α voice]** conditions are projected from perceptibility scales: if the latter are truly observations about relative perceptibility, then it is more likely that they refer directly to cues, rather than indirectly, through mention of the features that facilitate their perception. We consider in Part II the advantages of this second type of description, as we extend the study of laryngeal neutralization to contrast types not considered yet.

One argument developed in Part II will be based on the observation that the contexts of neutralization depend on the distribution of transitional cues in the surrounding context, which in turn depends on the timing of laryngeal gestures relative to oral constrictions (cf. Kingston 1985). The timing facts differ from language to language: aspirated stops are postaspirated in Sanskrit but pre-aspirated in Tarascan. Timing differences engender drastic differences in cue distribution, which in turn entail differences in neutralization contexts: indeed Tarascan and Sanskrit neutralize their aspiration contrast in very different contexts. We will suggest that direct reference to the quality of cues can provide unified descriptions for languages that differ in laryngeal timing, descriptions that allow a better recognition of the ways in which grammars do and do not differ from each other. Thus, we will suggest that Tarascan and Sanskrit differ from each other in the choice of laryngeal timing patterns, but not in the range of constraints relevant to the description of neutralization: those constraints may be the same, but their interaction with timing conditions yields different surface results.

A different argument for cue-based descriptions will have to do with the adequacy of features like [sonorant] for the description of laryngeal neutralization contexts. A statement like ***[α voice]/ _ [-son]** is useful to the extent that [-sonorant] encapsulates the correct class of neutralizing segments. However, we have seen that the segments inducing voice neutralization to their left include [h], a sonorant under most assumptions (Chomsky and Halle 1968:305), as well as strings in which the potential duration of modal voicing following the obstruent is too brief to serve as a useable indication of VOT category. The last point has been established in the analysis of Russian and Polish voicing: word medial O_1RO_2 sequences - with shorter R - neutralize O_1 whereas word final OR# - with longer, less overlapped R - maintains distinctive voicing in Russian. Similarly, non-syllabic R in the O_1RO_2 context fails to license voicing in O_1 whereas syllabic R does: the difference between the class of sonorant strings allowing distinctive voicing to their left and those which do not is a function of the duration of the string on which transitional cues like VOT are manifested. The global conclusion then is that we cannot adequately characterize the context that licenses voicing in languages like Hungarian or Polish as _[+sonorant], since a short, overlapped or inherently voiceless sonorant will not be a licenser. The invariant licensing factor is the ability of the context to express transitional onset cues such as VOT, F0 and F1 values. This argument is extended in Part II.