Alternatives to syllable-based accounts of consonantal phonotactics

Donca Steriade

UCLA

Abstract

Phonotactic statements characterize contextual restrictions on the occurrence of segments or feature values. This study argues that consonantal phonotactics are best understood as syllable-independent, string-based conditions reflecting positional differences in the perceptibility of contrasts. The analyses proposed here have better empirical coverage compared to syllable-based analyses that link a consonant's feature realization to its syllabic position. Syllable-based analyses require identification of word-medial syllabic divisions; the account proposed here does not and this may be a significant advantage. Word-medial syllable-edges are, under specific conditions, not uniformly identified by speakers; but comparable variability does not exist for phonotactic knowledge. The paper suggests that syllable-independent conditions define segmental phonotactics, and that word-edge phonotactics, in turn, are among the guidelines used by speakers to infer word-internal syllable divisions.

1. Introduction

The issue addressed here is the link between consonantal phonotactics and syllable structure conditions. Consider a paradigm case of the role attributed to syllables in current accounts of consonant distributions, based on Kahn (1976). Kahn suggests a syllable-based explanation of the fact that *tpt, dbd, tkt, dgd* strings - two coronal stops surrounding a non-coronal - are impossible in English. Kahn explains this gap by combining three hypotheses, two of which are specific to English, and a third with potential as a universal: (i) English onsets cannot contain stop-stop sequences; (ii) alveolar stops cannot precede non-coronals in English codas; and (iii) each segment/feature must belong to some syllable. This last hypothesis limits permissible strings of Cs to combinations of legal codas plus legal onsets. The system predicts that underlying strings like *VtptV* could not be parsed as *Vt.ptV* (in virtue of (i) above), or as *Vtp.tV* (cf. (ii)) or as *Vt.p.tV* (with unaffiliated *p*, cf. (iii)). In sum, the three hypotheses have the joint effect of predicting precisely the gap we sought to explain.

Proceedings of LP'98, edited by O. Fujimura, B.D.Joseph and B.Palek ISBN 80-246-0016-1

ISSN 0567-8269
The Karolinum Press:205-245

Syllable-based analyses have also been used to explain consonantal feature distributions. For example, the observation that German obstruents are voiceless word finally and before other obstruents has been attributed to the hypothesis that distinctive voicing is not possible in the syllable coda (Vennemann 1972; Hall 1992, Wiese 1996). For a full analysis parallel to Kahn's, we may have the following three hypotheses: (i) German syllables cannot begin with stop-obstruent sequences; (ii) German codas cannot contain distinctively voiced segments; and (iii) each segment/feature must belong to some syllable. Again, the analysis is simple: an underlying string *VgdV* cannot surface as (i) *V.gdV*, (ii) *Vg.dV* or (iii) *V.g.dV* (with unaffiliated *g*) due to each of the postulates listed above.

Syllable-based accounts have several attractive properties. First, to the extent that speakers have convergent intuitions of syllable constituency, it may be independently necessary to postulate a unit of analysis at this level. The syllabic unit will be endowed with specific constraints on its composition, and the segmental or featural hypotheses invoked above are possible instances of such constraints. More importantly, a model of syllable structure postulated to answer one phonotactic question may shed unexpected light on other, apparently unrelated, issues. One area where this is true is in the intersection of stress and syllabification. Take, for example, the correlation between phonotactics, stress attraction, and native syllabifications of stop-liquid (TR) clusters in Spanish and Cairene Arabic.[1] First, in the realm of phonotactics, words can begin with TR in Spanish, but not in Arabic, where all CC strings are disallowed word-initially.[2] Additionally, postconsonantal TR is possible in Spanish (e.g. *semblanza* 'sketch', *claustro* 'cloister') but not in Arabic. The stress facts require a similar distinction: any CC cluster, including TR, allows a preceding vowel to attract stress in Arabic: e.g. *tanábla* 'extremely lazy-pl'. In Spanish, on the other hand, many CC clusters have the same effect (e.g. *Salamánca, solémne*) but TR clusters do not (*fúnebre, idólatra*). Finally, native intuitions of syllabic division differ for the two languages: Spanish speakers consistently give V.TRV (e.g. *o.tros* 'others', *ha.bla* 'speaks') while Arabic speakers insist on VT.RV (e.g. *zak.ru* 'they studied', *qab.lu* 'they met'). It is possible to deduce the stress, phonotactic and division differences from a single factor: Spanish, but not Arabic, allows complex syllable onsets and such complex onsets include TR (but not RT, NT, etc.).

The suggestion emerging from examples of this sort is that the right hypothesis about syllables will generate structures from which both the segmental and the prosodic properties of a sound system can be predicted. What is attractive about theories of syllable structure is precisely their ability to simultaneously generate predictions in three distinct empirical domains: intuitions of string division, rhythmic phenomena like stress and constraints on permissible segment sequences. By the same token, the convergence of these phenomena on one syllable structure is the standard such approaches must meet. This means that when rhythmic patterns, phonotactics and speaker intuitions of constituency fail to converge on one set of syllabic structures, one is justified to look into alternatives. I outline next the ideas that structure the alternative discussed here.

Our starting observation is that, in certain core cases, the typology of phonotactic statements is illuminated by string-based, rather than by syllable-based conditions (Steriade 1997). Consider for instance the observation that distinctively voiced obstruents are typically followed by vowels or sonorants. The sequence obstruent-sonorant (TR) is frequently a possible onset cluster; and therefore the contextual limitations on obstruent voicing can be characterized either by focussing on the syllable, as in (1), or in linear terms, as in (2):

(1) a. A voiced obstruent is an onset. (Goldsmith 1990)
  b. A voiced obstruent is followed by a tautosyllabic sonorant. (Lombardi 1995).

(2) A voiced obstruent is followed by a sonorant.

The test case distinguishing (1) from (2) are strings of the form VD.RV, with distinctively voiced obstruents (D) in the coda, followed by heterosyllabic sonorants. It turns out that in at least some languages[3] (2) is the only accurate statement: distinctive voicing is preserved in the coda of VD.RV strings, but not in pre-obstruent or word-final codas. Therefore, in at least this case, a broader typological generalization - (2) - is obtained only if, in the choice between syllable-based and linear statements, we opt for the latter. In what follows I refer to this type of choice as Segmental Autonomy: the idea that phonotactic constraints are independent of the location of syllable boundaries.

We have seen that the prevalent context of voicing neutralization can be given a uniform statement in linear terms ("when no sonorant follows the obstruent") but not in syllabic terms.[4] Two comparable patterns - that of aspiration and place neutralization - are examined in this light below. To complete an argument for Segmental Autonomy we would also have to determine that there are no phonotactic laws that require a syllabic characterization. It is conceivable that some constraints are inherently linear - e.g. the condition on voicing in (2) - while others might be syllable-based. We should ask then: what would the typological data supporting a syllable-based phonotactic law look like? It would involve a phenomenon whose typology can be uniformly characterized syllabically but not linearly. A hypothetical case is the following: suppose that the occurrence of an anteriority s/ʃ contrast before consonants (e.g. in sp, ʃp, sm, ʃm, sl, ʃl etc. clusters) was limited to languages that permit syllabic parses such as V.spV, V.ʃmV, but was excluded from languages where the sibilant-C clusters are systematically heterosyllabic. This correlation between contrast distribution and syllable structure would require a syllable-based characterization, for instance "A distinctively [-anterior] sibilant occurs only in the onset."[5] In any case, I am unaware of any such correlations between neutralization and syllable parsing choices: indeed, Segmental Autonomy predicts that they will not occur.

There are broader generalizations which emerge when the positions segments occupy are classified not in syllabic terms (e.g. as onset vs. coda) or in linear terms (e.g. "before a vowel" vs. "after a consonant") but in terms of *relative perceptibility*, as positions where certain featural contrasts are more vs. less perceptible. This brings us to a second hypothesis, referred to here as Licensing by Cue: The likelihood that distinctive values of the feature F will occur in a given context is a function of the relative perceptibility of the F-contrast in that context. "Licensing" refers here to the fact that contrasts are said to be *licensed* in the contexts where the contrastive features are allowed to occur (Ito 1986, Goldsmith 1990.) Our hypothesis is that contexts where a contrast is allowed to occur differ from those where the contrast is prohibited in terms of the presence of more, or more infomative, perceptual cues.

We will observe in section 2 that the typological distribution of laryngeal and place contrasts supports Licensing by Cue, since positions of diminished perceptibility are systematically associated with greater incidence of neutralization. In section 3 I suggest that the same typology supports Segmental Autonomy, since the status of a contrast in a given position is not affected by shifting syllable boundaries within the string.

Recall now the correlation between stress, phonotactics and intuitions of syllable division, encountered earlier in the comparison between Spanish and Arabic. Under what assumptions are correlations of this sort consistent with the idea of Segmental Autonomy, which denies the relevance of syllables to the formulation of phonotactic conditions? To address this question, we need a third hypothesis, Word-Based Syllables, which can be summarized for now as follows: speakers rely on inference when they attempt to locate syllable boundaries in a multi-vowelled string, and one guideline in this process is that the segmental composition of word and syllable edges must be similar. Thus the possibility of parsing a string *abc* into *ab* and *c* is, in part, a function of the similarity between *b* and known word-finals and of the similarity between *c* and known word initials. One initial justification for this assumption is that syllables are domains whose edges appear to lack well-defined perceptual correlates: therefore any task that requires a phonologically-based partition of the string must rely on a form of inference that is substantially different from that involved in the segmental categorization of a string of speech sounds. I examine the forms taken by this syllabic inference process and its significance for the relation between segmental structure and syllabic organization.

In arguing against syllable-based accounts, one may spell out individual analyses using the *syllabic markedness* approach, in which context-free correspondence conditions (e.g. Ident [α voice]) are combined with syllable-sensitive markedness conditions (e.g. *[+voice]-in-Coda). But the arguments developed hold as well against the *syllabic faithfulness* approach (a version of positional faithfulness, cf. Jun 1995, Steriade 1995, Beckman 1997): this analytical technique combines syllable-sensitive correspondence conditions (e.g. Ident [α voice]-in-Onset) with context-free markedness conditions (e.g. *[+voice]). The choice between positional markedness and positional faithfulness constraints arises independently of one's views about the role of prosodic structure in segmental phonotactics. What is at issue here is the terms in which contexts are analyzed - segmental or prosodic - not the types of constraints referring to these contexts.

Sections 1 and 2 sketch the typological arguments for the hypotheses of Segmental Autonomy and Licensing by Cue. Section 3 outlines the evidence for Word-Based Syllables.

2. Licensing by Cue and Segmental Autonomy.

This section summarizes two surveys of consonant neutralization patterns: aspiration and place neutralization. In both cases certain implicational patterns emerge that correlate with the asymmetric distribution of basic perceptual correlates for each contrast. The overall generalization is that place and aspiration contrasts tend to be neutralized in contexts where they lack their primary perceptual correlates.

2.0 Left- and right-anchored contrasts

It is possible to classify segmental contrasts based on asymmetries in the distribution of their transitional cues. An example of a transitional cue is the Voice Onset Time (VOT) value, a primary correlate of voicing and aspiration. This cue is asymmetrically distributed: a vowel following an aspirated stop will be contextually devoiced by it and thus will provide information about the stop's laryngeal feature. A preceding vowel will not provide this information. For this reason, we say that contextual devoicing is an *asymmetrically distributed cue* to aspiration. Using this asymmetry as a classifying principle, we can distinguish right-anchored from left-anchored contrasts. In right-anchored contrasts, such as that between [t^h] and [t], the transitional cue to aspiration resides exclusively in the post-release interval. In left-anchored contrasts, on the other hand, the main transitional cue resides in the context preceding the onset of the consonant's closure. An example of a left-anchored contrast is that between the pre-aspirate [^{h}t] and [t]: in this case it is the context preceding [^{h}t] and [t] that will help differentiate their laryngeal categories, by manifesting the transitional devoicing effect associated with the aspirate.

Linear asymmetries in the distribution of transitional cues represent a substantial factor in the typology of neutralization. The basic observation establishing this is that right-anchored and left-anchored contrasts display mirror-image patterns of neutralization. For right anchored contrasts - e.g. [t^h] vs. [t] - it is the nature of the following context that is the critical factor in neutralization. Simplifying somewhat, the preservation of a [t] vs. [t^h] contrast depends on whether a vowel or sonorant *follows* the stop. Only in such a context is a post-aspiration contrast typically preserved. For left-anchored contrasts - e.g. [^{h}t] vs. [t] - it is the nature of the preceding context that plays this role: to determine whether the [^{h}t] vs. [t] contrast is likely to neutralize, the critical question is whether a vowel or sonorant *precedes* the stop. The generalization that encompasses both contrast types is that the optimal context for the manifestation of a contrast is that in which its major transitional cue is present. This is one observation that supports the idea of Licensing by Cue.

2.1 Laryngeal neutralization

Consider first the typology of neutralization in aspiration-based contrasts. (Neutralization of voicing and glottalization contrasts follow a similar pattern.) The most common variety of aspiration contrast is that between post-aspirated stops (T^h) and plain voiceless ones (T). A less common distinction is that between plain voiceless stops and pre-aspirated ones (^{h}T). The main cue to pre-aspiration, the voice offset time (VoffT[6]), precedes the onset of oral closure. The gestural difference between pre- and post-aspirated stops can be diagrammed as below: the significant aspect in these diagrams is the timing relation between the **offsets** of oral and glottal gestures (in T^h) and the timing between the **onsets** of oral and glottal gestures (in ^{h}T).

(3) Post-aspirated stops: $\mathbf{T^h}$, D^h

Peak of laryngeal gesture timed to release of oral constriction: e.g. t^h
[----------glottal abduction----------------]
[------oral closure -------------release-]

(4) Pre-aspirated stops: ^{h}T

Peak of laryngeal gesture timed to onset of oral constriction: e.g. ^{h}t
[------------glottal abduction-----------------]
[----oral closure---release--------------]

If one compares patterns of neutralization for languages with pre- vs. post-aspiration a striking observation emerges: the T^h vs. T contrast is typically neutralized in the absence of a following sonorant; whereas the ^{h}T vs. T contrast is lost in the absence of a preceding sonorant. This difference between pre- and post-aspiration can be illustrated through a comparison of Sanskrit (a T^h language) and Gaelic (a ^{h}T language).

Sanskrit (Whitney 1889) allows voiced and voiceless post-aspirates (D^h, T^h) before sonorants and neutralizes aspiration contrasts finally and before obstruents, i.e. when no sonorant follows. What segments, if any, precede the stop has no effect on the realization of aspiration in this language.

(4) Sanskrit post-aspiration (Whitney 1889)

- Contrast before sonorants:

	D^h	D
g^h vs. g before V	dag^{h}at 'has reached'	b^{h}a:ga 'share'
g^h vs.g before sonorant	dag^{h}nuyat 'reaching'	agnih 'fire'

- No contrast before obstruents: d^h**ak**tam 'you two reached'
- No contrast word finally: d^h**ak** 'has reached'

• Context that must follow T^h, D^h: sonorants
• Context preceding T^h, D^h: any segment, or no segment (#_)

Scots Gaelic (Leurbost dialect, Island of Lewis: Oftedal 1956) illustrates an aspiration contrast between voiceless ^{h}T and T. The pre-aspirates are systematically missing word initially and after voiceless sounds, for instance when they are stops in the onset of a syllable preceded by an obstruent coda (as in *paskat*, where *k* cannot contrast as to aspiration). This is explicitly recognized by Oftedal (1956: 43): "stops that follow an ordinarily voiceless consonant are always unaspirated." Aspiration is not absent word-initially but it is realized there as post-aspiration. Distinctive pre-aspirates may occur in contexts where

Sanskrit aspirates are forbidden, for instance word-finally. The segments, if any, following the pre-aspirated stop have no consequences for the realization of aspiration:

(5) Leurbost Gaelic distinctive pre-aspiration (Oftedal 1956)

- Contrast after sonorants:

	^hT	T
^{h}t vs. **t after V**	k^{h}a^{h}t 'cat'	paskat 'basket'
^{h}t vs. **t between V**	i^{h}tə 'feather'	pi:təɣ 'to bite'
^{h}k vs. **k after sonorant**	paŋ^{h}kə 'bank'	laŋkətər 'seaweed'

- Neutralization after obstruents: cf.neutralized [k] in *paskat*
- Pre-aspiration disallowed word-initially: *^{h}ka-

  - Context following ^hT: any segment, or no segment (_#).
  - Context that must precede ^hT: a sonorant.

The contexts of neutralization for Sanskrit T^h (word-final and before obstruent) are essentially the mirror image of the contexts of neutralization for Leurbost Gaelic ^hT (word-initial and post-obstruent).

Both pre- and post-aspiration contrasts show a certain amount of cross linguistic variation but the variation is lawful and maintains this mirror image relation. That is, post-aspirates prefer to be followed by a sonorant (preferably a vowel) while pre-aspirates prefer to follow a sonorant (more likely a vowel or a stressed vowel). Item (6) summarizes the results of the survey that has brought these preferences to light. There are significantly fewer systems attested with pre-aspiration contrasts and my count of ^hT systems is conservative - thus excluding languages like Icelandic (Kingston 1990) where pre-aspirates might be analyzed as clusters. Despite this, the trend is clear: any restriction imposed on the distribution of the contrast involves contexts that follow T^h and contexts that precede ^hT.

(6) An abbreviated survey of post-aspiration (**right-anchored**) contrasts

- Class A: distinctive T^h occurs only **before** a vowel.
  [E.g. Takelma (Sapir 1912), Sre (Manley 1972]
- Class B: distinctive T^h occurs only **before** a sonorant
  [E.g. Greek, Klamath (Barker 1967), Khasi (Nagaraja 1985)]
- Class C: distinctive T^h occurs only **before** sonorants and word-finally.
  [E.g. Gujarati (Cardona 1965), Parsi Gujarati (Gajendragadkar 1974)]
- Class D: distinctive T^h occurs in all positions where stops may occur.
  [E.g. Hindi (careful speech only: M.Ohala 1987), Yokuts (Newman 1944).]

The post-aspiration systems surveyed in (6) share two properties: (a) constraints on the occurrence of distinctive T^h characterize its right-hand context and (b) a scale of preference for the following context emerges:

(7) $T^hV > T^h[+\text{sonorant}, -\text{syll}] > T^h\# > T^h[-\text{sonorant}]$.

The preference scale in (7) reflects the observation that presence of distinctive T^h before an obstruent - the least preferred context for the realization of post-aspiration - implies, in any given system, that of word final T^h (written $T^h\#$ above). In turn, the existence of word final distinctive T^h implies that of T^h before a non-syllabic sonorant, and T^h before a non-syllabic sonorant implies the possibility of T^h before a vowel.

Pre-aspiration systems display the opposite orientation: constraints on the occurrence of distinctive ^hT relate to the composition of the *preceding* sound.

(8) An abbreviated survey of pre-aspiration (**left-anchored**) contrasts

- Class A: distinctive ^hT only **after** a (stressed) vowel.
  [e.g. Toreva Hopi (Whorf 1956), Bernera Gaelic (Oftedal 1956)]
- Class A or B: distinctive ^hT only **after** a vowel (no sonorant-stop clusters attested).
  [e.g. Fox (Jones 1910)]
- Class B: distinctive ^hT only **after** a sonorant (including a vowel).
  [e.g. Leurbost Gaelic (Oftedal 1956), Papago (Saxton 1963, Fitzgerald 1996), Tarascan (Foster 1969), Lappish (Engstrom 1987)]
- Class C: distinctive ^hT only **after** sonorants and word-initially.
  [Huautla Mazateco (Pike and Pike 1947)]

Despite the difference in left-right orientation between pre- and post-aspiration systems, both types of contrasts are preferably realized in the vicinity of a vowel or sonorant. The implicational scale observed for post-aspiration is found, in mirror image, in the case of pre-aspiration as well: a language permitting ^hT initially, also allows it after a vowel or sonorant; and a language permitting ^hT after a consonantal sonorant also allows ^hT after a vowel, but not necessarily the other way around.

We can characterize different types of aspiration systems while abstracting away from the directionality issue, as below:

(9) Contexts selectively licensing aspiration contrasts:
class A: An aspirated stop must be adjacent to vowel.
(V is right-adjacent to T^h; left-adjacent to ^hT)
class B: An aspirated stop must be adjacent to a sonorant (incl. a vowel).
(the sonorant is right-adjacent to T^h; left-adjacent to ^hT)
class C: An aspirated stop must be adjacent to a vowel or a sonorant as in class B, **or** adjacent to a word boundary (#).
(# is right-adjacent to T^h, # is left-adjacent to ^hT)

An account of the data outlined so far must answer three questions: why is it preferable for there to be *a vowel or sonorant* adjacent to the aspirated stop; why must this vowel or sonorant surface *on the same side as the aspiration*; and what is the analysis of the boundary effects in Class C systems?

Take the first two questions. Why do aspirates need an adjacent vowel or sonorant? All aspiration contrasts are cued primarily by a period of transitional voicelessness, which is best perceived when it audibly modifies a neighboring sound. The best sound to be modified in this way is a sonorant, because sonorants are normally voiced, and thus can be contextually devoiced. The hearer can rely on this contextual devoicing effect to identify the presence of aspiration. Second, sonorants provide good aspiration contexts because, typically, they are not themselves contrastively voiced: they may be substantially devoiced by neighboring T^h or ^hT without sacrificing any laryngeal distinctions of their own. Vowels, especially stressed vowels, appear to be the best contexts for the aspirated segment: in languages like Toreva Hopi and Bernera Gaelic, distinctive ^hT occurs only after a stressed vowel. This is perhaps due to the fact that vowels, particularly stressed vowels, are sufficiently long to preserve a portion of themselves unaffected by aspiration. A *partially* devoiced neighboring sound may be the most reliable cue to aspiration.

What about the location of the vowel or sonorant relative to the stop? If the sonorant is to provide a contextual cue for the stop's aspiration then it must be located on the side of aspiration, and that means after T^h, and before ^hT. In effect then, we have observed that pre- and post-aspiration contrasts neutralize in identical circumstances: when aspiration cannot manifest itself contextually, as devoicing of a near-by sonorant. It follows then that, despite appearances to the contrary, what differentiates the pattern of ^hT-neutralization from that of T^h-neutralization is just the timing of oral and glottal articulations. Modulo this difference, we can say that aspiration contrasts - whether right- or left-anchored - are typically neutralized in contexts where aspiration cannot be perceived through its devoicing effects on context. This is then a first example of the role of perceptibility in the analysis of phonotactic effects. The broader hypothesis of Licensing by Cue simply generalizes this observation to other features.

In Class C aspiration systems, the contrast is permitted in positions where contextual aspiration cues (i.e. VOT and VoffT) are available, and, additionally, in word-final post-vocalic position (for T^h) and word-initially in pre-vocalic position (for ^hT). There are in principle two ways to analyze Class C aspiration. One line of analysis would be based on the idea that a richer range of contrasts is attested, *for all features*, at the edges of prosodic domains. If this holds for all features, in all combinations, and for both domain edges, then the simplest analysis is one that invokes directly a positional priviledge associated with domain edges. This can be modelled either as a positional faithfulness ranking schema or as positional markedness, but for simplicity we will consider only the former option: preservation of any feature occurring at the edge of any domain D is more highly ranked than preservation of the same feature when not occurring at the edge. Using MAX (αF) as the preservation constraint[7], we write:

(10) For any feature F, any domain D:
MAX (αF)/ $[_D$__ , MAX (αF)/ __$]_D$ >> MAX (αF).

Under this analysis, Class C systems represent grammars in which one of the edge faithfulness constraints, either **MAX** (αF)/$[_D$__ or **MAX** (αF)/__$]_D$, outranks a conflicting ***αF/ in K** condition, where K designates a class of contexts that includes $[_D$__ or __$]_D$ respectively. For instance, if K represents the class of contexts where T^h is *not* followed by a sonorant (e.g. the word final and pre-obstruent contexts), then this set of contexts includes the __$]_{Word}$ context.[8] Under a ranking such as **MAX (α aspirate)/** __$]_{Word}$ **>> *[α aspirate]/ not followed by sonorant >> MAX (α aspirate)** the aspiration will be preserved word finally but not before an obstruent. This will yield a description of the Class C post-aspirate languages like Gujarati.

However, the ranking schema in (10) should be rejected because it is capable of generating aspiration systems in which all initial and all final T^h or ^hT are permitted but none of the word medial ones are. Consider the ranking **MAX** (α aspirate)/ $[_{Word}$__, **MAX (α aspirate)/**__$]_{Word}$ **>> {*α aspirate/in** K^1, ***α aspirate /in** K^2, **... *α aspirate/in** K^n **} >> MAX (α aspirate)**. In this case, the entire set of ***α aspirate/in K** constraints, regardless of what K is, is outranked by the edge-faithfulness conditions and outranks in turn the general faithfulness ondition **MAX (α aspirate).** Therefore word-initial and final aspirates are rotected regardless of their relative perceptibility, while all word- medial as- pirates, are prohibited - again, regardless of *their* contextual perceptibility. No uch systems are in fact attested. This interestingly wrong prediction is gener- ated by any analysis that allows a set of edge-based constraints - either edge- based faithfulness conditions or edge-sensitive markedness conditions - to freely interact with an independent set of perceptibility-based constraints. The problem is resolved when we understand that the edge effects reflect additional distinctions of perceptibility and must therefore be accounted for by the same family of conditions that regulates contrast distribution in terms of contextual perceptibility.

The other option for analyzing Class C systems involves an additional as- sumption regarding the perceptibility of aspiration. So far we have distin- guished three classes of contexts: those in which the contextual devoicing that cues aspiration is manifested on a vowel; those where it is manifested on a son- orant; and the context deprived of contextual cues, where no neighboring seg- ment can be perceptibly devoiced by aspiration. Suppose however that this last class of contexts actually subdivides further in terms of relative perceptibility. In word final position a final T^h can be distinguished from T via its burst dura- tion and amplitude. In word medial position, on the other hand, a pre-obstruent T^h will lack - under certain conditions of inter-consonantal overlap - an audi- ble burst altogether, thus obliterating this residual cue to aspiration. We will as- sume that a comparable overlap effect explains the difference between word-initial pre-vocalic ^hT (attested in Huatla Mazateco) and post-obstruent ^hT, which remains unattested. Under this interpretation, Class C systems toler- ate aspiration in a position of intermediate perceptibility: the aspirated stop is identified through a puff of air (preceding closure, in the case of ^hT; following release in the case of T^h) but not by contextual devoicing of a neighboring sound. The word-edge position is superior to the pre-obstruent position (for T^h) and to the post-obstruent position (for ^hT) because the consonant on the edge escapes the masking effects of consonantal overlap.

We can now sketch a synchronic analysis of aspiration neutralization which incorporates the main lines of the perceptibility scenario outlined so far and which holds the promise of generating the range of systems observed typolo- gically. The key elements in the analysis are statements which model speakers' awareness of relative perceptibility: for instance the fact that the pre-aspiration contrast is more reliably identified in post-sonorant than in post-obstruent po- sition. Knowledge of this sort is encoded in the form of scales which compare the perceptibility of a feature contrast in a range of contexts.

(11) Contextual perceptibility scales:
([+F]/[-F] refers to the contrast between these values;
➤= *more perceptible than*)
[+F]/[-F] in context K1 ➤[+F]/[-F] in K2 ...➤ [+F]/[-F] in K3...

e.g.

T/T^h before vowel ➤T/T^h before sonorant ➤T/T^h word finally ➤T/T^h before obstruent

$T/^hT$ after vowel ➤ $T/^hT$ after sonorant ➤$T/^hT$ word initially ➤$T/^hT$ after obstruent

Such perceptibility scales are assumed to project constraints whose general form is: *Do not realize the contrast between +F and -F in a context of dimin- ished perceptibility.* This means that for any F-perceptibility scale of the form in (11) there is a corresponding *F-constraint family in which the worst context for the perception of the feature F, say context K, gives rise to the highest ranked *F constraint, *F/_K, the next worst context K' generates the second highest *F constraint, *F/ __K' and so forth. There are a variety of constraints of this sort, corresponding to the various contexts where the feature F occurs but, critically, their relative ranking mirrors the relative position of the contexts on the perceptibility scale.

(12) Perceptibility conditions for F aligned to F's contextual perceptibility scale
[+F]/[-F] in context K1 ➤ [+F]/[-F] in K2 ..➤ [+F]/[-F] in K3...
*F in K3 >>*F in K2 >> *F in K1...

Thus the perceptibility scales in (11) will project the constraint families below:

(13)(a) $*T/T^h$ before obstruent >> $*T/T^h$ word finally >> $*T/T^h$ before sonorant

(b) $*T/^hT$ after obstruent >> $*T/^hT$ word initially >> $*T/^hT$ after sonorant

These constraints can interact with faithfulness conmditions such as MAX [F] and this interaction can yield a description of the attested systems and a characterization of the typology. We observe this below for the T^h-systems: note that the difference between Class A, B, C and D systems is reflected in the relative ranking of the faithfulness condition, MAX [aspiration], and the mem- bers of the constraint family in (13.a). What is invariant - and what explains the implicational relations observed between T^h-systems - is the internal ranking of the constraint set in (13.a) which mirrors the perceptibility differences ob- served across contexts.

(14)

Class A:

*T/T^h before obstruent
↓
*T/T^h word finally
↓
*T/T^h before sonorant
↓
MAX [aspiration]
↓
*T/T^h before vowel

Class B:

*T/T^h before obstruent
↓
*T/T^h word finally
↓
MAX [aspiration]
↓
*T/T^h before sonorant
↓
*T/T^h before vowel

Class C:

*T/T^h before obstruent
↓
MAX [aspiration]
↓
*T/T^h word finally
↓
*T/T^h before sonorant
↓
*T/T^h before vowel

Class D:

MAX [aspiration]
↓
*T/T^h before obstruent
↓
*T/T^h word finally
↓
*T/T^h before sonorant
↓
*T/T^h before vowel

The same ranking variation between MAX [aspiration] and the constraint family in (13.b) characterizes the implicational relations between pre-aspiration systems.

Finally, we assume statements of intergestural timing, inspired by Browman and Goldstein's ideas (1992), which indicate whether the language implements its aspiration contrast as pre- or post-aspiration.

(15) Oral-to-glottal timing statements
(a) T^h : The peak of glottal opening aligns to the oral release.
(b) ^hT: The peak of glottal opening aligns to the onset of oral closure.

These are also violable conditions: for instance (15.b) is violated in Leurbost Gaelic when the aspirated stop is realized as a post-aspirate in initial position (cf. 5). This violation is triggered by the different perceptibility of T^h and ^hT in a word initial, prevocalic context: ^hT would be considerably harder to identify there than T^h. Thus the timing modification insures that the aspiration contrast is perceptibly realized across contexts. The link we propose between timing and neutralization is indirect: different timing relations between oral and glottal features create a different distribution of cues to aspiration, and the cue distribution in turn, in the form of the perceptibility constraints, is directly responsible for neutralization.

The weight of the explanation proposed for aspiration neutralization is on two ideas: first, that the neutralization of any contrast is more likely in contexts of diminished perceptibility. Second, that knowledge of differential perceptibility enters the grammar in the form of constraints like (13), which are projected from perceptibility scales.

2.2 Aspiration and syllable structure

Earlier, I mentioned the possible convergence of different types of evidence as the best reason to explore syllabic accounts of segmental structure. I examine now the relation between syllable structure and the facts of aspiration neutralization. The question is how successful are syllable-based accounts of the data considered so far.

Consider first the languages with post-aspiration, surveyed in (6). The neutralization in some of the T^h-languages of class B has been analyzed - for instance by Lombardi (1995) - on the assumption that stop-sonorant ($T(^h)R$) sequences are tautosyllabic. Indeed, in certain prominent Class B T^h-languages, such as Greek and Sanskrit, the stop-liquid or stop-nasal clusters occur word initially. The actual proposals regarding the position of neutralization differ somewhat, but all share the assumption that distinctive aspiration may only be realized in obstruents that are followed within the same syllable by a sonorant. If this assumption is made, then the grammatical statements responsible for the neutralization of aspiration are syllable-based, as below:

(15) Syllabic analyses of Class B T^h-languages

- Distinctive T^h must be licensed by the onset.

or

- Distinctive T^h must be licensed by tautosyllabic sonorants.

The evidence supporting tautosyllabification of T^hR clusters in the Class B T^h-languages is limited to the observation that initial $T(^h)R$-clusters exist, coupled with the assumption of onset maximization, i.e. that clusters attested initially are parsed tautosyllabically in all contexts. However if we consider evidence that bears more directly on the issue of word medial syllable division, the picture changes considerably.

The meter in Sanskrit and in at least some dialects of Greek indicates that all stop-sonorant sequences were in fact divided into separate syllables. Thus for a sequence like Sanskrit VT^hRV (cf. (4)) the metrical evidence shows that the first syllable is heavy: this argues for parses like $VT^h.RV$, with a distinctively aspirated stop in the coda (Mishra 1972). Similarly, Homeric Greek $VT(^h)RV$ strings are metrically scanned with an initial heavy, as if divided $VT(^h).RV$. Weight-dependent allomorphy effects - cases in which one allomorph occurs next to a heavy stem syllable and the other allomorph next to light syllables - also support the view that TR clusters were heterosyllabic: for such allomorphic purposes, $VT(^h)RV$ sequences contain an initial heavy, hence $VT(^h)$. RV (Devine and Stevens 1994: 104 on Greek; Whitney 1889: §1002 on Sanskrit).

For some T^h-languages, available grammars record the syllabic intuitions of native speakers. In many of these cases, TR clusters are reported to be heterosyllabic: e.g. Gujarati (Cardona 1965) and Parsi Gujarati (Gajendragadkar 1974), two Class C languages with reportedly heterosyllabic medial $T(^h)R$. The mere presence of a following sonorant is all that is required for the licensing of aspiration: the syllabic relation between this sonorant and the preceding stop is irrelevant.

Certain class A systems provide a different argument for the claim that T/T^h neutralization is unrelated to syllable structure. In these languages[9] complex onsets of the form TR do occur, but pre-consonantal stops cannot be aspirated, regardless of syllabification, hence T^hR is impossible. Thus the sonorant in TR onsets fails to license aspiration on the stop. The syllabic position of the aspirated stop is therefore neither necessary as a condition of neutralization, nor sufficient: it is not sufficient in these cases because the correct characterization of aspiration licensing contexts must specifically mention a neighboring vowel rather than the onset position.

Consider now ^hT languages, like Leurbost Gaelic. For these, the connection between aspiration neutralization and syllable structure is impossible to make in any form. Not all Leurbost coda stops lose aspiration: none do if they follow a vowel or sonorant, as [^{h}**t**] does in [k^ha^ht]. Nor is it possible to claim that it is the onset stops which neutralize aspiration, since intervocalic stops always maintain aspiration and initial stops maintain it as well, albeit with changed timing.

Our observations can be summarized as follows. For a significant subset of T^h-systems the predictions of the syllable-based analyses diverge from those proposed in (14). The divergence concerns the status of VT^hRV strings: syllable-based analyses predict that T^h is allowed in such strings in all and only the languages admitting a $V.T^hRV$ parse. This prediction is repeatedly contradicted by the evidence. Moreover, none of the ^hT-systems are analyzable in syllabic terms. The perceptibility-based analysis offerred for T^h-systems extends straightforwardly to ^hT as well: the contexts where pre-aspiration is neutralized are parallel, in terms of availability of contextual cues, to those where post-aspiration is lost.

2.3 Place features

The Licensing by Cue analysis can be applied to the neutralization of place distinctions. There are two types of place features, whose patterns of neutralization parallel the T^h and ^hT systems discussed earlier. The better studied place contrasts are right-anchored: their main cues lie in the post-release interval (burst - in the case of stops - and CV transitions).[10] These features include those insuring the distinction between labials, coronals and velars; the anteriority contrast between laminals (for languages contrasting dental [t̪] and [tʃ]) and the apical-laminal contrast ([t̪] and [tʃ] vs.[t] and [ʈ]). Left-anchored place features, although less common, exist as well: they are similar to pre-aspiration in that their primary cues lie in the interval preceding the onset of closure. Retroflexion is one such feature: retroflexes display significantly lowered F3 and F4 but these effects are manifested mainly in the V-C transitions.[10] The retroflex-to-V transitions are similar or identical to those of apico-alveolars (Stevens and Blumstein 1975). There is a clear articulatory explanation for this fact: the tongue tip slides forward during the retroflex closure so that the constriction site reached just prior to release is hard to distinguish from that of an apicoalveolar.[11] The forward shift in constriction site means that retroflexes produced after a consonant (esp. after a stop) or in word-initial position will not cue their distinctive point of articulation by audibly lowering the F3 and F4 of a preceding sound. They might not be impossible to identify, but their perceptibility in such contexts is decreased relative to that of postvocalic retroflexes.

The patterns of neutralization in place features mirror those of T^h and ^hT: the right-anchored place contrasts (e.g. {b vs. d vs. g}, the inter-laminal contrast {d̪ vs. dʒ} and the apical vs. laminal contrast {ʈ d, ɖ vs. d̪, d̪ } typically neutralize in the absence of a following vowel or sonorant while the left-anchored retroflexion contrast (ʈ e.g. ɖ vs. d) neutralizes in the absence of a preceding vowel or sonorant. Indeed, right-anchored place contrasts appear to fall into the by-now familiar four classes of place-neutralization patterns, sketched below in (16); while left-anchored place contrasts give rise to the mirror image pattern of neutralization, surveyed in (17). Note the preference for a *following* vowel or sonorant in the case of major place contrasts (16); and the opposite preference for a *preceding* vowel in the case of retroflexion.

(16) Left/Right asymmetries in the location of place neutralization:
Right-anchored place contrasts[12]

Class A: contrast permitted only before V
(e.g. Japanese)

Class B: contrast permitted only before V and approximants
(e.g. Late Latin)

Class C: contrast permitted only before V and approximants and in V _# (e.g. Diola Fogny)

Class D: contrast permitted in all or most contexts where obstruents occur. (e.g. English)

(20) Left/Right asymmetries in the location of place neutralization:
Left-anchored place contrasts[13]

Class A: contrast permitted only after V
(e.g. Gooniyandi: McGregor 1990)

Class B: contrast permitted only after V and approximants (not documented)

Class C: contrast permitted only after V and in #__
(e.g. Djinang: Waters 1980)

Class D: contrast permitted everywhere
(e.g. Hindi: M.Ohala 1987)

This classification of place neutralization types reinforces a point made earlier: contexts where a place contrast is neutralized are contexts where cues specific to that contrast are significantly diminished when compared to the cues available in positions where the contrast is maintained. To the extent that the distribution of contextual cues differs between two contrasts - as it differs for the alveolar / retroflex contrast compared to the dental / palato-alveolar contrast - to that extent the positions of typical neutralization will also differ. The weakness of the syllabic analysis of phonotactic conditions is the idea that a single syllable position (e.g. the onset) or a small set of positions (the syllable onset, the word onset, the stressed syllable) can be used as all-purpose licensers for a large set of diverse features. Cues are contrast-specific: it turns out that licensing patterns are contrast-specific as well, and that they match closely the cue distribution.

3. Phonotactics and syllable division

3.1 Learning syllables

Up to this point we have assumed that it is possible for the linguist to define - and for the native speaker to discover - principles of syllabic organization in a way that does not rely on knowledge of phonotactics. Starting from this assumption, we have shown that there is no systematic correlation between the syllable and the positions where two classes of consonantal features are licensed. Ohala (1990:329) anticipates this finding as follows: "syllables are logically subsequent, not antecedent, to constructing the optimal segment stream itself." A given string is marked or unmarked, admissible or not in a given language, regardless of how it is divided into subconstituents. A segment string deemed acceptable will be prosodically parsed, if produced. But the decision to assign to it a specific prosodic structure has nothing to do with accepting the string in the first place.

This position leaves unanswered a key question: if there is even a partial match between phonotactic facts and syllabic intuitions, what is its source? Are syllable-to-phonotactics correlations like those observed in the Spanish-Arabic comparison accidental? This question is addressed in a largely speculative way in this section. I suggest that speakers' intuitions of syllabic division are derived through an inference process in which a large part is played by knowledge of word phonotactics: in particular, by knowledge of the range of permissible initials and finals. This is the hypothesis of Word-Based Syllables mentioned earlier. On this view, syllabic intuitions correlate in part with word-phonotactics because the latter determine the former, not the other way around.

At the core of the Word-Based Syllables hypothesis lies the assumption that the boundaries of prosodic domains must be perceptibly expressed in order to be learned uniformly by members of a speech community. Such a requirement does not affect morpho-syntactic constituents, because these can be learned through paradigmatic reasoning. Thus, to discover that *bomber* [bɑmər] is bimorphemic, the learner must possess the paradigmatic knowledge that *bomber* is related to *bomb*. This suffices to locate the internal constituent boundary, without any need for a perceptible boundary mark. The process of learning syllable division is fundamentally different, because it cannot be assisted by paradigmatic reasoning of this sort. It must either be based on direct perceptual correlates of syllable structure or else it must proceed indirectly, by exploiting the assumption that syllable edges bear a structural similarity to the edges of better known constituents, such as words.

I assume that the second option is closer to reality. This is not to deny that one can occasionally pinpoint articulatory correlates differentiating syllable-initial from syllable-final segments (cf. Boucher 1988, Browman and Goldstein 1988, Krakow 1989, Turk 1995) although much of this research conflates, of necessity, word and syllable positions. But, more to the point, it has not been shown that the articulatory differences documented in such studies translate into perceptual correlates that are reliably exploited by listeners. Further, there is reason to believe that such syllabic correlates as might exist are limited to certain segments and possess limited cue-value. Experimental work suggesting this appears in Portele (1995).

The same conclusion can be reached by considering the results of psycholinguistic investigations of syllable division. Speakers of English are divided on whether words like *lemon* should be parsed as *lem.on, le[m]on* (with ambisyllabic [m]) or *le.mon*. But this fact emerges from studies (such as Treiman and Danis' 1988 and Derwing's 1992) which probe opinions on this matter, not through observation of speakers' production. Those volunteering *lem.on* were not observed to differ in their production from those who prefer *le.mon*. I suggest that the variation in responses reflects different ways of reasoning about possible parses of *the same auditory stimulus*. This explains why no perceptible differences were associated with the syllable production of different classes of subjects. Similarly, there are no reported isoglosses involving syllable division, no dialectal areas differentiated, say, by the distribution of ambisyllabic C's or by the division of TR clusters. This could be because, in normal circumstances, one cannot read the syllable boundaries in the auditory stimulus. A phonological isogloss involves a perceptible difference between two alternative pronunciations of the same set of forms: if syllable constituency yields negligible cues, then we understand why there are no parsing isoglosses. Finally, the universal non-existence of lexical contrasts of syllable division - contrasts such as [a.ska] vs. [as.ka] - may be explained in the same way. The diachronic survival of a contrast depends on its meeting some minimal discriminability standards: I am suggesting that whatever articulatory differences may result from alternate syllabic parses, the perceptual consequences of such differences fall short with respect to discriminability.

If we grant then that language learners are unable to find clear indications of syllabic constituency in the auditory stimulus, it follows that they must rely on inference in discovering what is the proper partition of a string of segments. And, since syllable divisions are, to an extent at least, language specific, inferences about syllables cannot be based entirely on innnate preferences. What else are they based on? How do speakers acquire the preference for one syllabic cut as against another? By studying the phonotactics: by using readily accessible evidence on the structure of word-edges to shed light on an abstract structural question. The Word-Based Syllables hypothesis is that speakers infer the location of syllable boundaries by assuming that syllables are one-vowelled words, hence that they bear certain structural similarities to words; and that the central dimension of similarity involves the segmental composition of the edges. On this view, speakers will opt - *ceteris paribus* - for parses that maximize the similarity of word edges (whose structure is known) to syllable edges (which are to be discovered).

So far, our Word-Based Syllable hypothesis is indistinguishable from a principle that has been invoked in earlier work: cf. Treiman and Danis's Legality Principle and comparable ideas found in Pulgram 1970, Kahn 1976 and earlier writers. The Legality Principle states that syllable edges must represent possible word edges. A critical difference between this principle and our hypothesis is that, in our view, the similarity between word and syllable edges is a heuristic guideline - one among others - in a process of inference. For Kahn, Treiman and also for researchers who reject the Legality Principle, syllable boundaries are not inferred but somehow directly perceived. Another difference is that we assume that, in cases where the word-edge guideline conflicts with others, the speaker has few means to determine how the conflict is resolved by other members of his speech community, precisely because the syllabic divisions lack clear perceptual correlates. Therefore the Word-Based Syllable hypothesis predicts that, in circumstances where multiple guidelines to division conflict, the conflict will be reflected in increased response variability to tasks like hyphenation, which rely directly on the placement of syllable boundaries. We assume that variability is a reflection of the speakers' uncertainty as to how to partition the string. Conversely, strings that elicit consensus in syllable division tasks are predicted to coincide with cases where all guidelines to division converge on one answer.

In what follows I review several phenomena supportive of this view. I document first the prediction regarding variability in syllable division in cases where (a) certain parsing preferences conflict with phonotactic guidelines to syllable division or (b) the word-edge phonotactic guidelines fail to rule out all but one parse. I consider then a case where multiple phonotactic guidelines conflict with each other; and finally the effect of word phonotactics on syllable division in languages with unusual word edge properties.

3.2 Variability in parsing and ambisyllabicity

Recall that our guiding assumption in reasoning about these matters, is that at normal speech rates there are no reliable perceptual correlates of syllable divisions: the learner who acquires strings like [CɛCəC] cannot determine whether other speakers mean to convey [Cɛ.CəC] or [CɛC.əC] and is thus free to impose his own parse. A bias in favor of one parse, revealed in an experimental setting, must then be due either to a preference for some structure (e.g. open syllables) or for some dimension of similarity between syllable and word edges (e.g. no final [ɛ]). I concentrate on two predictions of this syllable-learning scenario, which distinguish it from the prevalent assumption that syllable divisions are directly accessible to the learner.

Our scenario predicts response variation and uncertainty in cases where the assumption of word-to-syllable correspondence conflicts with other guidelines to syllabic division. For the purposes of this discussion, I assume, in line with much work in phonology, that there exist universal preferences for syllable parsing. (This assumption may need revisiting in the long run, but the arguments presented here do not require this.) One such universal preference is that for the parse V.CV as against VC.V: the preferred parse minimizes the number of closed syllables and maximizes the number of syllables with onsets. The question we address now is how this parsing preference interacts with the edge-based phonotactic guidelines to syllable division. The case of interest is that in which the preference for V.CV parses conflicts with what the speaker knows about possible word edges. For instance, in the case of words like *lemon* [lɛmən], the conflict takes the following form: on the one hand, [ɛ] is not a possible word-final vowel, hence [lɛ.mən] fails the word-based phonotactic test of possible finals; on the other hand, the alternative [lɛm.ən] violates the preference for V.CV. For concreteness, we adopt the principles in (21) as preliminary versions of edge-based phonotactic guidelines to syllable division:

(21) Word-to-syllable identity conditions

W-S(I): For any I, a syllable-initial segment, there is a word such that its initial segment is identical to I.

W-S(F): For any F, a syllable-final segment, there is a word such that its final segment is identical to F.

It is clear that W-S(F) and the preference for V.CV parses conflict in dividing forms like [lɛmən]. What is less obvious is that any attempt to resolve this conflict faces two sources of uncertainty. One is discovering the ranking of conflicting constraints: since, *ex hypothesi*, other speakers' solutions to this problem cannot be identified in the auditory output, the learner does not know whether to accept the disfavored parse VC.V ([lɛm.ən]), which satisfies W-S(F), or the preferred parse V.CV ([lɛ.mən]) which violates W-S(F).

The other source of uncertainty is the fact that any form of correspondence, including that between word and syllable edges, may be sought at different levels of detail. For instance if some C's, but not others, are permitted word initially, the syllable learner can interpret this fact in multiple ways, depending on the degree of strictness with which the requirement of word-to-syllable similarity is interpreted: for instance, is a syllable initial [p] sufficiently similar to an (obligatorily aspirated) word initial [pʰ] to satisfy W-S(I)? Is a syllable final [ɛ] sufficiently similar to an attested word final [e] to pass W-S(F)? We lack a theory of segmental similarity to provide specific answers to these questions. The right direction seems to assume that the greater the similarity between a given syllable edge and an attested word edge, the greater the confidence of the speaker in the relevant parse. Thus we should not assume that the mere existence of W-S(F) leads speakers to confidently reject the parse [lɛ.mən] (cf. Derwing 1992, for evidence showing that they do not): the learner has no information about the linguistic norms that determine the strictness of syllable-to-word correspondence or the ranking between such conditions and conflicting constraints. The phonotactic fact that [ɛ] cannot occur word finally only has the effect of reducing the confidence in [lɛ.mən] relative to otherwise comparable parses such as [di.mən] (i.e. *demon*.) which do not run up against W-S (F). One manifestation of non-confidence may be response variability, in cases where several parses exist with compensating advantages and disadvantages.

The prediction of greater variation in parsing [lɛmən] than [dimən] is confirmed by the results of Treiman and Danis (1988) and Derwing (1992): Derwing's subjects split into two nearly equal groups in parsing forms like *lemon* (with 51% opting for [lɛm.ən], 37% opting for [lɛ.mən] and the 12% residue opting for ambisyllabic [m]). In contrast, there was a solid 82% consensus for parses like [di.mən].[14] Similar patterns were observed with intervocalic obstruents and laterals: for forms like *melon*, *seven* (with lax V_1) there was significantly more parsing variation than for forms like *select*, *depend* (where V_1 was either schwa or tense, hence a possible final). Derwing did give his subjects the option of ambisyllabic assignment, but this was disfavored, except when orthographic geminates occur. Thus appeals to ambisyllabicity do not represent a viable alternative to our view that syllable structure represents the outcome of a process of phonotactically-based inference. Schiller et al's (1997) results for Dutch (where lax vowels are likewise impossible word finally) are similar to the English pattern: CəCVC and CeCVC strings were divided predominantly as Cə.CVC (97.2%) and Ce.CVC (97.9%) whereas CɛCVC gave rise to a much narrower 61% (CɛC.VC) to 36% (Cɛ.CVC) split in responses.

A second prediction of our approach involves cases in which the phonotactic guidelines fail to exclude enough parses. Here too we predict greater variation and uncertainty compared to cases where the phonotactics converge on a unique parse. This point can be illustrated through a comparison of Polish and English: Polish allows a richer array of initial clusters, including sequences of steeply decreasing sonority like [rt], [mʃ], and [wʒ]. Not all conceivable combinations of consonants occur initially in Polish, but there are no broadly definable phonotactic gaps among the initial CC sequences, as there are in English. Intuitions of division for CC clusters, when intervocalic, vary significantly among Polish speakers. Dictionary writers (Jodłowski and Taszycki 1958; Szymczak 1975) allow the V.CC(C(C))V parse for any choice of C-cluster, including cases like *Tu.rcja, ka.rtʃma, a.bstrakt*. In contrast, the speakers who participated in Dubiel's (1994) study differentiate clusters according to their sonority profile: parses like [ka.rta], [ma.wʒa], [za.mʃovɨ], [poko.rn] were volunteered in 12% to 30% of the responses from this group, as against an average 40.6% V.CCV parses in stop-sonorant clusters like [rɪ.tmɨ], [ka.dra].

The first significant fact here is that Polish speakers do not unanymously reject the V.CC(C)V parse for strings of decreasing sonority: [ka.rta] is a marginally viable option in Polish, not an unthinkable one, as in English (cf. Fallows 1981). Speakers of both languages may prefer to avoid onsets of decreasing sonority, but the matter is clear-cut only in English, where word-phonotactics strongly confirm the dispreference. A second point is that the parsing variation in Polish is tied to whether a cluster or cluster type is found initially: for [ns], [ɲtʃ], [lʃ], [ltʃ], which are absent initially, the V.CCV parses in Dubiel's study fall under 6%. There is no general structural difference between clusters like [rt] and [ltʃ]: the difference between the two sets - which correlates with the rates of tautosyllabic parsing in Dubiel's study - involves only word-initial attestation of the cluster type.[15]

Polish speakers, judging from Dubiel 1994 and from the range of hyphenations sanctioned by Szymczak 1975, reject unconditionally onsetless parses of the form [VC.V]. There is no variation here. In this respect too they differ from English speakers, for whom $V_1C.V_2$, with lax V_1, is one of two roughly equally favored parses (V.CV, VC.V: Derwing 1992). The variation involved in parsing English VCV is largely limited to strings with a lax penult, like [lɛmən]. This difference between English and Polish is also phonotactically based: all Polish vowels are permitted word finally, but not all English vowels are. Therefore the V.CV parse is not penalized by any consideration in Polish.

To repeat: the variation in responses stems from the fact that the language cannot impose an enforceable norm on syllabification, since parsing differences lack perceptible consequences. Therefore when parsing preferences conflict, speakers are free to resolve the conflict individually, or to not resolve it all. There is no variation only when both phonotactic and syllabic preferences happen to converge on one parse.

3.4 Conflicts between phonotactic guidelines

We have seen that syllabic well-formedness and phonotactic guidelines to syllabification are in potential conflict. The phonotactic guidelines may also conflict with each other: if a string of segments abc cannot be tautosyllabic in its entirety and if a is impossible word finally and c is impossible word initially, the division into syllables is expected to vary because each speaker must decide on his own how to resolve the conflict. The logic of this situation is outlined below.

(23) Parsing uncertainty due to phonotactic conflict:

	Parse	Phonotactic penalizing it
Dividing abc into syllables	*ab-c:	c is impossible word initially.
	*a-bc	a is impossible word finally.

This case is encountered in English sequences of the form: V_1sTV_2, where V_1 is a lax vowel, T is a voiceless stop and V_2 is stressed: for instance *mystérious, vestígial, despótic*. In such strings, the $V_1.sTV_2$ parse is difficult because V_1 is impossible finally. The $V_1s.TV_2$ parse is also problematic: T is unaspirated in the original - where it follows [s] - but unaspirated initial T is impossible, particularly when it precedes a stressed vowel. The parse $V_1sT.V_2$ is difficult in English not only because it involves an onsetless syllable but also on phonotactic grounds: word-initial stressed vowels are prefixed by [ʔ] or glottalized (Pierrehumbert and Talkin 1990). Thus none of the conceivable divisions are phonotactically sanctioned.

(24) Parsing uncertainty due to phonotactic conflict: the case of English V_1sTV_2 (V_1 is lax; T is a voiceless stop; V_2 is stressed)
Dividing [dɛspɔ́rɪk]

Parse	Phonotactic penalizing it
*[de.spɔ́...]	[ɛ] impossible word finally
*[de.spɔ́...]	[p] impossible initially before ' V;[p^h] expected
*[de.spɔ́...]	[ɔ́] is impossible initially; [ʔɔ́] expected

Contrast *despótic* with minimally different *asbéstos* [æzbɛ́stəs] and *moslémic* [mɑslɛmɪk]. What has changed here is the status of C_2: voiced stops like [b] and sonorants like [l] can occur initially, in contrast to voiceless unaspirated stops. Therefore nothing precludes the parses [æz.bɛ́s...] and [mɑs.lɛ́m...]: we predict no variation or uncertainty for such forms, since all other parses continue to be penalized.

The predicted difference in parsing *sl* and *sp* is confirmed in Pritchett (1984), a study of expletive insertion. Pritchett observed that words like *despotic* tend to split the [s] under infixation (e.g. *des-fucking-spotic*) and that syllable division in such cases is uncertain. His study - in which 20 Harvard Law School students were asked to apply expletive infixation to a selection of 100 target words - aimed to verify the correlation between uncertain syllable boundaries and split [s]'s. Pritchett classifies his data in terms of whether the syllable parse is unambiguous to him. But the subjects' own ideas on syllabification can only be inferred from their performance on expletive insertion. We have therefore retabulated Pritchett's data looking for a correlation between split [s]'s ([s]'s that both precede and follow the expletive) and certain segmental characteristics of the surrounding string which lead, along the lines explained earlier, to parsing indeterminacy. I interpret segment splitting as the strategy followed when speakers are uncertain how to parse the string without opting for a ranking among conflicting word-edge phonotactics. Indeterminacy in parsing is expected in V_1sTV_2 where V_1 is lax and V_2 is stressed, but not when V_1 is schwa or tense (e.g. *destrúctable*), or in V_1CsTV_2 (e.g. *Tolstoy*), since the parse $V_1C.sTV_2$ raises no phonotactic difficulty. In V_1sTV_2 strings where V_1 is lax and V_2 is stressless e.g. *mustard*), one expects uncertainty as well, but there are not enough examples with this structure in Pritchett's materials to draw any conclusions. The retabulated results appear below:

(25) Data from Pritchett (1984) reclassified

	Vs-**F**-tV	Vs-**F**-stV	Other (mainly V-**F**-sTV)
Type 1 e.g. [dɛspárɪk] V_1sTV_2, V_1 lax, V_2 stressed	12.8%	80.5%	6.7%
Type 2 e.g. [maslémɪk] V_1sRV_2, with V_1 lax,	86.3%	0%	13.7%
Type 3 e.g. [dəstrʌktabl̩] V_1sTV_2, V_1 schwa	11.6%	37.3	51.1%
Type 4. e.g. [tɔ́lstoj] $V_1Cs T\acute{V}_2$	26.1	16.8%	58.1%

The 80.5% split [s] responses for Type 1 ([dɛspárɪk]) confirm the expectation that V_1sTV_2 strings (with lax V_1 and stressed V_2) will pose a parsing dilemma: no partition of [dɛspárɪk] can match each syllable-final with a word-final and each syllable-initial with a word-initial. No other word category in Pritchett's data comes close to the ratio of split [s]'s of Type 1. Our proposal predicts this imbalance: in every other string type there is some syllabic division that satisfies the word-to-syllable edge matching requirement.[16]

The second aspect of the data that can be explained is the difference between Type 2 and Types 3-4 with respect to non-split infixation patterns (V(C)s-**F**-CV vs. V(C)-**F**-sCV). Types 3-4 favor positioning the infix before sC (e.g. *de-F-structable, Tol-F-stoj*) while Type 2 prefers infixing in the middle of the sC cluster (*mos-F-lemic*). This follows from the fact that Type 2 strings contain a lax V_1: the pattern *mos-F-lemic* is the only parse that shields this lax V_1 from the syllable final position. In contrast, the segment immediately preceding [s] in Types 3 and 4 is a vowel ([ə]) or consonant, both of which can occur word finally. Therefore the infixation pattern V_1(C)-**F**-sTV_2 is possible. Infixation before sT in such cases is not only possible but also preferred, as the data indicates. This too can be explained: a large number of the Type 3 and 4 strings contain a stressed V_2, which would require an aspirated T if T was word-initial. To preserve the unaspirated stop of the original string, the sT cluster must remain intact in such forms.

A further point of interest that can be gleaned from Pritchett's data is that the tautosyllabic parse for Type 2 *sR* clusters is dispreferred in relation to the frequency of the voiced pronunciation of the fricative: the fricative in *islamic* or *moslemic* can be [s] or [z], but that of *Tasmania, Bismarck* is exclusively voiced [z]. Correspondingly, forms like *Tasmania* give rise to 100% Vs-F-RV infixing patterns, whereas *moslemic*, shows a weaker (80%) preference for Vs-F-RV infixation. Word-edge phonotactics have an effect here because zC sequences are impossible word initially: thus a parse such as [tʰæ-**F**-zmenjə] is doubly impossible, first because of the final [æ] and second because of the initial [zm]. In contrast only the final [ɑ] rules against [mɑ-**F**-slɛmɪk]. The hypothesis of phonotactically based syllabification explains in this way both the major trends in Pritchett's data and some of the smaller details. Note in this connection that there is no general principle related to syllable well-formedness that explains why zC clusters are excluded from initial position. Sonority rises in [zm] about as steeply as it does in [sm] and voicing, by itself, does not disqualify obstruents from occurring in complex onsets. Thus the general absence of [z] from initial clusters cannot be deduced from independent syllabic laws: it is just a fact about English edge phonotactics. But this fact does affect syllabic parsing, as we have seen.

Like the [m] of *lemon*, the [s] of [sT] strings like *despotic* is reported by some as ambisyllabic (cf. Kahn's 1976 algorithm, which means to derive this). The perspective on syllables as phonotactically-based inferences leads us to equate the intuition of ambisyllabicity with sheer uncertainty regarding the right parse. The [s] of *despotic* is not in both syllables at once: rather the syllable learner cannot assign it to either. What all accounts must explain is that [m] is felt to be ambisyllabic in *lemon* but not in *demon*, that [s] is ambisyllabic in *despotic* but less so in *Tolstoy* or *destructive* and not at all in *Islamic*. The idea that syllabification is uncertain when the phonotactic guidelines conflict predicts these facts.

3.5 Parsing universals and exceptions

Two generalizations on syllabic division are widely assumed to hold universally. First, when a parse exists which minimizes the number of onsetless syllables, that parse is necessarily adopted: hence $VC_0C.V$ is impossible, because $VC_0.CV$ is available. The second generalization concerns the sonority profile of consonantal interludes: an intervocalic cluster of decreasing sonority is generally not assigned *in toto* to the onset. Hence *a.rta, a.lma, a.msa* are impossible or disfavored too. Two isolated languages, Malayalam and Arrernte, are known to challenge both of these generalizations. Speakers of colloquial Malayalam (a Dravidian language) are reported to systematically opt for the V.RTV parse (Mohanan 1986, T.Mohanan 1989). Speakers of Arrernte (Australian Aboriginal) parse VC_0CV as $VC_0C.V$ (Breen n.d; Henderson 1998).

Analyzed in Optimality Theoretic terms, the problems raised by Malayalam and Arrernte seem quite different. Consider Malayalam [ka.mpi]. If this parse is penalized by constraints on the complexity and sonority profile of the onset, then the relevant comment is that such constraints conflict with No Coda. The ranking No Coda >> Sonority Sequencing, *Complex will yield exactly [ka.mpi] and all patterns reported by the Mohanans. The question then is not why Malayalam deviates from the alleged universal on sonority decrease in onsets but why no other language does. The problem raised by Arrernte however cannot be overcome by ranking: a chief feature of Prince and Smolensky's (1993) analysis of syllabic typology is that it guarantees that the $VC_0C.V$ parse will lose to $VC_0.CV$, because $VC_0.CV$ encurs a proper subset of the violations of $VC_0C.V$. Then Malayalam is straightforward but Arrernte should not exist.

If syllabic parsing is an inference process based largely on word-edge properties, we can explain both exceptional patterns in similar ways. The key observation is that both languages have unusual word-edge properties: Malayalam words seldom end in C's, and then only in a very limited range of C's; Malayalam words frequently begin with C-clusters, some of which decrease in sonority. Details are given below. Arrernte word edges have opposite properties: at a modest level of abstraction, Arrernte words never begin with a C, and always end in C. The assumption that syllable boundaries are identified by matching the segmental quality of word boundaries predicts $V.C_0CV$ for Malayalam and $VC_0C.V$ for Arrernte. The syllabic parsing strategies encountered in these languages are unusual simply because each one accidentally presents an unusual conjunction of edge phonotactics. But given the phonotactics, the parses are straightforwardly predicted. In what follows I outline briefly the evidence supporting these suggestions.

3.5.1 Malayalam

Medial C-clusters in Malayalam include sequences of increasing and decreasing sonority as well as sonority dipping clusters.

(26) Malayalam medial clusters (Asher and Kumari 1997)

(a) increasing or level sonority:	tr, tn̪, ʃw, kk, sm, nm, ps, kt, ktr.
(b) decreasing sonority:	ʂʈ, ɾg, pn, ɽʂ, lpp, rkk,
(c) dipping sonority:	ndr, rgj, rdw, str, nsp, lsj

Initial clusters in careful speech include the sequences in (26.a) and s-stop clusters. The range of initial clusters is broader in casual speech (K.P.Mohanan 1986 and p.c.): initial [i] is lost regardless of the C sequence following it and this may create initials like [ɲtʃ] in [ɲtʃi] (from careful speech [iɲtʃi] 'ginger') or [ndra] (from careful speech [indra]. The word final consonants of colloquial Malayalam are limited to the nasals [m] and [n]. There is productive epenthesis of [ə] after all other C-finals, including non-anterior nasals:

(27) Final ə-epenthesis in colloquial Malayalam (Mohanan 1986, T.Mohanan 1989)

word medial	word final	gloss
[wirak-]	[wirakə]	'firewood'
[kun̪iʃ-]	[kun̪iʃə]	'Holy Cross'
[itaɭ -]	[itaɭə]	'petal'
[aaɳ-]	[aaɳə]	'male'
[wayaɾ]	[wayaɾə]	'stomach'

We now consider the evidence for open syllabification in Malayalam. Mohanan (1986) reports that literate and illiterate speakers of colloquial Malayalam assign the entire consonantal interlude to the second syllable. This is confirmed by a syllable-based form of speech disguise in which [pa] is inserted before every maximal C*V string of the word.

(28) Malayalam pa-language (Mohanan 1986, T.Mohanan 1989)

Base word	Pa-variant	Gloss
band^hanam	→ paba-pand^ha-panam	'imprisonment'
daɾ̆anam	→ pada-paɾ̆a-panam	'vision'
ammu	→ paʔa-pammu	(name)
tʃaṉḏran	→ patʃa-paṉḏran	(name)

Malayalam-Hindi bilinguals play the pa-game differently depending on whether it is used with Malayalam words or Hindi (a language allowing unrestricted C-finals): in Hindi, the pa-string is inserted before the lowest sonority point in a medial C-cluster:

(29) Hindi pa-language (T.Mohanan 1989):

Base word	Pa-variant	Gloss
band^hanam	→ paban-pad^ha-panam	'imprisonment'
daɾ̆anam	→ padaɾ-pa̋a-panam	'vision'
ammu	→ paʔam-pamu	(name)
tʃaṉḏran	→ patʃaṉ-paḏran	(name)

Forms of speech disguise such as this are typically taught using minimal instruction and minimal examplification. It is critical then that some abstraction such as *the syllable* be mutually understandable between game teachers and game learners, or inferrable from initial game examples, since the specific behavior of individual clusters is not being drilled. Therefore, to produce *padaɾ-pa̋a-panam*, when speaking Hindi, but *pada-paɾ̆apanam*, when speaking Malayalam, the learner must have reached the conclusion that syllables define the locus of *pa*-insertion and that they are systematically different between the two languages.[17] This must have been the reasoning behind Mohanan's argument for Malayalam open syllables. But on what basis does the learner *know* that the Malayalam parse is [da.ɾ̆a.nam] and that it differs from the Hindi parse? He knows only two critical facts: that most C's, including [ɾ], are impossible word finally and that initial clusters of decreasing sonority are possible in the colloquial variety of Malayalam he is first exposed to. He also knows that, in these respects, Hindi is different.

The Malayalam *Praasam* rhyme is a versification form based on the identity of the second maximal C*V string in each verse (K.P.Mohanan 1986). The second syllable of Malayalam words is typically stressed. The Praasam rhyme is then similar to the accentually based rhyming schemes of modern Indo-European languages. Pairs such as *kuṯi* and *maṯi*, *suṉayani* and *iiṉarttaki*, *kampi* and *tumpi* rhyme in Praasam. The identical strings are underlined: note that they correspond to the second syllable in each pair of words, as identified by Mohanan's syllabification rules. Conversely, pairs like *kampi* and *tuppi* do not rhyme: this indicates that the shared string [pi] does not correspond to a prosodic domain. On the plausible assumption that Praasam identity is computed on the same unit as that identified in the pa-language game, this data strengthens Mohanan's argument for a system of open syllabification.

The simple existence of parses such as *ii.na.ɾtta.ki* in Malayalam does not exclude the analysis anticipated earlier, in which No Coda outranks Sonority Sequencing. The objection to this analysis is that it fails to link the unique parses of Malayalam to its unique word edge properties. By linking the two we explain how bilinguals come to know that Hindi and Malayalam must be parsed differently. Accounts based on re-ranking Sonority Sequencing and No Coda do not explain why languages with the phonotactics of Hindi never adopt Malayalam's ranking. More generally, such accounts also leave us in the dark regarding the process by which speakers discover the evidence for a given ranking: if syllabic parses are imperceptible and phonotactic evidence is irrelevant, the learner will not know whether to construct a grammar that syllabifies *ii.na.ɾtta.ki* or *ii.naɾt.ta.ki.*

Our account leaves one property of Malayalam syllabification unexplained: given the word-final occurrence of [m] and [n], strings like *kampi* should be parsable as either *ka.mpi* or *kam.pi*. It is unclear at this point what considerations rule out the second parse or to what extent this judgment is constant across speakers.

3.5.2 Arrernte and Oykangand

Word-initial consonants have been lost in a number of Australian languages. At least two of these, Oykangand-Olgol and Arrernte, have also lost - or neutralized to [ə] - their final vowels. With final [ə] analyzed as a predictable release of the final C (Henderson 1998), this historical development yields words of the form $VCC_0VCC_0.VCC_0$. There are no systematic restrictions on the range of C's admissible finally or on the V's admissible initially. Then, if speakers rely on edge-phonotactics as their chief guideline to syllabic parsing, each maximal VCC_0 sequence represents a syllable. The boundaries of the syllable will emerge in syllabic forms of reduplication, in speech disguise forms involving syllable reversal and in metalinguistic tasks. Proposals with this character appear in Sommer's (1970) Oykangand studies and in Breen's (n.d) and Henderson's (1998) work on Arrernte dialects. I summarize here some of the arguments for onsetless syllabification in Arrernte based on evidence presented by Breen and Henderson, and largely confirmed in Henderson's recent informant work.

Consider Arrernte reduplication patterns, which involve suffixation of a disyllable; prefixation of a monosyllable copied from the left edge; and prefixation of a right edge monosyllable:

(32) Reduplication (Eastern Arrernte, Henderson 1998)

• Suffixing:	əṉtakəṉ	→	əṉtakəṉ - akəṉ
	arətjal	→	arətjal-ətjal
• Prefixing:	amər	→	am-amər
	ulpəɳ	→	ulp-ulpəɳ
• Opposite prefixing:	arip	→	ip-arip
	əlap	→	ap- əlap

Note first that all three patterns of reduplication are consistent with the assumption that members of reduplicated syllables occupy identical syllabic positions relative to their correspondents in the surface base, assuming onsetless parsing. In other words, reduplication in Arrernte can be viewed as copying the syllables defined in an onsetless ...VCC_0.VC... parse: this assumption explains why *ulpaɳ*, parsed as [ulp.aŋ] reduplicates as [ulp-ulpaŋ] instead of [ul-ulpaŋ]. However the clearest argument supporting this view comes from comparing suffixing forms like *əṉtakəṉ - akəṉ* with phonotactically acceptable but incorrect fuller copies such as **aṉtakəṉ takəṉ*. Note that both forms satisfy all conditions on possible initial and medial sequences in Arrernte. Moreover, the incorrect **aṉtakəṉ takəṉ* should be, *ceteris paribus*, favored because it represents a fuller copy of the base string (cf. McCarthy and Prince 1995). What is then that favors *aṉtakəṉ-akəṉ?* The answer is, again, that reduplicated segments maintain their original syllabic positions: [t] is a coda (*əṉtak.əṉ*) hence its reduplicated correspondent cannot be an onset (cf. Steriade 1988 for discussion of comparable conditions and their effects in reduplication).

A further argument for the onsetless parse comes from Breen's (n.d.) discussion of a syllable transposition speech disguise similar to French Verlan (Lefkowitz 1991): the initial VCC_0 moves to the end of the word:

(33) Arrernte Rabbit Talk (Breen n.d.; Henderson 1998)

Rabbit Talk	Normal
ampaŋkəm	*aŋkəmamp*
əjaṉ	*aṉəj*

Most types of reversal that involve strings longer than one segment can be analyzed as changing precedence relations among syllables. If this is true of Rabbit Talk, then the moving string VCC_0 is a syllable: then *aŋkəmamp* must be *parsed aŋk.əm.amp.*[18] This type of analysis is supported by the reduplication ev- idence discussed above and is consistent with the Breen's and Henderson's report that Arrernte speakers prefer onsetless parses of the form *u'ṉatət (ə)*→ *u'n (ə) ṉatət (ə)*, where (ə) represents the automatic word-final release. Insofar as these speakers could partition words at all, they appear to use as their main guide in this process the word-edge phonotactics. This is exactly what the evidence from reduplication and speech disguise indicate.

The analyses presented in the last two sections are significant for obvious empirical reasons: no alternative proposals exist that assign both Malayalam and Arrernte a clear place in the typology of syllabic parsing. The data acquires further significance when we consider how the conflict between syllable well-formedness and phonotactic guidelines to parsing is resolved in these languages. In the Malayalam and Arrernte data, the requirement of identity between word and syllable edges always determines the outcome, even when all relevant constraints on the structure of syllables weigh against this solution. I illustrate below the Arrernte competition between syllable structure constraints (Onset, No Coda, *Complex) and word-to-syllable identity conditions (W-S(I) from (21).

(35)

	W-S(I)	No Coda	Onset	Complex
☞ *aŋk. m.amp*		***	***	**
aŋ.k .mamp	*! (k), *(m)	**	*	*

The tableau in (35) illustrates the fact that W-S(I) is critical in deriving the parses required for the analysis of reduplication and Rabbit Talk, while constraints on syllable structure proper, like No Coda and Onset, remain inactive.

There is a further question prompted by this discussion, which, for lack of space and a satisfactory answer, I will only briefly identify. In the discussion of ambisyllabicity (section 3.2) I attributed the variation in parsing forms like *lemon* to the conflict between the preference for V.CV parses (a preference due to principles governing the structure of syllables) and word-to-syllable correspondence (21). The suggestion was that variability is the necessary outcome of such conflicts, because the learner cannot tell how other members of his speech community resolve parsing conflicts and hence cannot align his solution to theirs. We should expect then similar variability in the case of the Malayalam and Arrernte processes dependent on syllable division: but this is not what has been reported.

There are two possible differences between the English ambisyllabicity data and the Malayalam and Arrente facts. One difference is that we are comparing subject responses to an experimental task (in English) with conventionalized linguistic behavior (in the Malayalam Praasam rhyme; the Arrente Rabbit talk and reduplication). It is possible that a speaker may remain fundamentally undecided between two syllabic parses, while at the same time learning from overt evidence that only one of these parses provides the representations required for some specific, conventionalized linguistic activity, such as rhyming or speech disguise.[19] The other difference concerns the generality of the phonotactic principles available to English as against Arrernte learners. Three broad generalizations emerge from a study of Arrernte words: no words begin with consonants, no words end in vowels and the range of clusters permissible finally is identical to those appearing medially. These generalizations are based on the most basic phonological contrast, that between vowels and consonants, and they unambiguously point to the $VCC_0.V$ parse. By contrast, consider two of the phonotactic principles governing English word edges and relevant to syllable division: "No lax vowels distinct from schwa are permitted finally"; "No unaspirated voiceless stops are permitted initially." These principles refer to contrasts like [ɛ] vs. [e], [t^h] vs. [t] which are considerably less perceptually robust than the contrast between V and C, and which, moreover, lack the ability to evenly partition the segment inventory, since the tense/lax distinction remains undefined for consonants. I will use the term *generality* to refer both to the general applicability of the contrast and to the perceptual robustness of the contrast targetted by a given phonotactic law: a phonotactic law is more or less general in relation to the robustness and general applicability of the contrast it affects. One can speculate then that it is the relative degree of generality of a phonotactic principle that determines whether it will prevail in case of conflict with other factors: thus, the edge-phonotactics of Arrernte may have a more decisive effect on syllable division than those of English, precisely because they are more general.

Conclusions

The first half of this study has suggested a line of argument from which one may eventually conclude that the well-formedness of a segment sequence can and should be characterized in terms of relative perceptibility and not in syllable-sensitive terms. Further work leading to one or the other of these two conclusions appears in Blevins 1993, 1999, Kotchetov 1999, Lamontagne 1993, Rubach 1996, Steriade 1995, 1997.

The second half of the paper has sought to address further questions that arise from the rejection of syllable-based phonotactics. The question explicitly addressed so far was this: why do we occasionally encounter correlations between intuited syllable division and phonotactics? In answering this, I have suggested that the phonotactics of word edges play a basic role in guiding speakers' inferences about where the syllable boundaries might lie. The correlation arises then, when it does, not because the phonotactics are syllable-sensitive but because the process of inferring syllable boundaries is phonotactically guided. The two examples discussed at the outset, Spanish and Arabic, lend themselves to straighforward analysis along these lines: Arabic lacks all word initial clusters whereas Spanish allows TR initials. At the other end of the word, Arabic permits a broad range of C's to occur finally, whereas native Spanish words end in sonorants or [s] only, not in stops. Syllabic intuitions can be deduced entirely from these word-edge differences: Spanish favors V.TRV over VT.RV parses because (a) word final stops are missing in the native lexicon and (b) TR initials are possible. Arabic, on the other hand, rejects V.TRV in favor of VT.RV because (a) TR initials (and all CC initials) are impossible and (b) VT finals are not ruled out.

Syllable structure, whether directly perceived or inferred, is an undeniable aspect of phonological representations. This study suggests however that it does not condition segment realization. To gain some partial understanding of this fact, consider again the phenomenon of German final devoicing. The voicing/aspiration contrast of German has well-understood perceptual correlates. A child acquiring German can learn to categorize phones as [+voice] and [-voice] and has sufficient information to do so in substantially the same way as other speakers of his dialect. Based on this information, he can identify contexts of neutralization, where stops surface as exclusively voiceless: word finally and before obstruents. Because his understanding of the defining characteristics of voicing is secure, he learns the distributional properties of the voicing contrast in much the same way as other speakers of German. The learner's syllable knowledge, on the other hand, is necessarily the result of indirect and, frequently, insecure inference. Why then would the German learner, who can clearly identify segmentally all the contexts where voicing does and does not contrast, attempt to recast the well-understood distribution of voicing in terms of an uncertain and partial understanding of syllable boundaries? Why would he not attempt the exact opposite: locate the word-medial syllable boundaries by relying on the clear knowledge that voicing does not contrast word-finally.

Notes

1. Spanish data is from Harris 1983. The Cairene Arabic (Arabic, from hereon) pattern is discussed by Broselow 1976.

2. CC clusters are tolerated phrase-initially in Cairene Arabic. See Broselow 1976 for analysis of these exceptions to the general pattern.

3. Lithuanian, Polish and Russian and others reviewed in Steriade 1997.

4. Details of this argument are spelled out in Steriade 1997.

5.Nothing is said in this hypothetical case about the s/ʃ contrast among the word-final sibilants. That is because there are enough assumptions in the arsenal of syllable-based descriptions to make either possibility consistent with the proposed condition.

6.Term adopted from Engstrand 1987.

7.Cf. McCarthy and Prince 1995 for a discussion of the concepts of faithfulness and correspondence.

8. The analysis of word final phrase-medial cases is provided in Steriade 1997.

9. Burushasky (Berger 1974), Sre (Manley 1972), Brokskat (Ramaswami 1982), Jeh (Gradin 1966), Chepang (Caughley 1972).

10. The primacy of CV over VC transitions in the perception of place categories is established in work by Fujimura, Macchi and Streeter 1976 and Ohala 1990.

11.The forward slide of the tongue tip in the articulation of retroflexes has been documented by Butcher (1993) and Henderson (1998).

12.Cf. Ohala 1990, Mohanan 1993, Jun 1995.

13. Surveys in Hamilton 1995 and Steriade 1995.

14. The categories *lemon* and *demon* in Derwing's study correspond to trochaic disyllables in which the first vowel is short and lax (as in [lɛmən]) vs. tense (as in [dimən]).

15. Szpyra (n.d) addresses this point by noting that minimally different strings - e.g. VrdV and VrbV - occasionally have the same percentages of tautosyllabic parsing despite the fact that only one of the clusters (rd, in this case) is attested initially. Szpyra's conclusion is that initial attestation has no effect on the speakers' acceptance of tautosyllabification. The question though is how specific is the word-to-syllable similarity sought by the syllable learner: the relevant generalizations concerning rC clusters in Polish need not refer to individual segments but to broader classes, such as rT or rO (T = stop, O = obstruent). Thus the fact that rt and rd occur initially, may well be interpreted as partial evidence that all rO are possible word initials, hence that all rO are possible syllable initials. The learner may rationally assume that the non-attestation of initial rb is accidental, since b and d are otherwise behaving similarly and the global attestation of all rC initials is sparse.

16. Two other strings, Type 3 ([dəstrʌktəb]]) and Type 4 ([tɔlstoj]), have substantial ratios of split [s] but I believe that both figures are artificially high. Pritchett did not track the individual pronunciation of his subjects and thus we do not know whether those who volunteered Type 3 forms like *mines-fucking-strone* pronounce [mìnəstrón] or [mìnɛstrón]: the [ɛ]-form predicts s-splitting, the [ə]-form does not. Of the 16 items in Type 3, seven suffer from this vocalic ambiguity. The Type 4 figure (16.8%) is largely the contribution of two rare items, *capstan* and tungsten which inexplicably gave rise to 60% and 40% split-[s] infixations. Once again, we do not know how the subjects pronounced these words. A re-make of the experiment in which individual pronunciations are tracked will plausibly yield an even clearer confirmation of the phonotactic hypothesis. A further question that needs clarification involves splitting the segment as a form of conflict resolution: it is unclear why this strategy was not favored for overt partition tasks (like Derwing's) or for VCV strings.

17. T.Mohanan (1989) has studied the behavior of a bilingual Konkini-Kannada child (where Konkini has the relevant phonotactics of Hindi and Kannada those of Malayalam) during her first exposure to the pa-game. The child's behavior was as outlined in the text: she systematically varied the location of epenthesis depending on which language she was speaking, placing *pa* before the maximal C*V string in Kannada but before the lowest sonority point in each Konkini interlude.

18. In this case, if we ignore the reduplication evidence for $VCC_0.V$ parses, there are alternative interpretations of the data. Thus, if a VCVC string was parsed V.CVC then syllable reversal will yield CVC.V: but word-initial C's are impossible in Arrernte, and this may affect the outcome of reversal. It is thus possible to sketch an analysis of syllable reversal that denies Breen's basic claim about Arrernte syllabification and relies on the idea that a final syllable moved to the front of the word must satisfy word phonotactics, including the ban on initial C's.

19. The speaker may also learn, again from overt evidence, that one parse yields the right representations for one process, while the other is the required parse for some other process. Thus Attic versification requires the V.TRV parse whereas Attic rhythmic allomorphy requires the VT.RV parse (Devine and Stevens 1994).

20. I refer here to the fact that - as for other languages - the syllable divisions elicited from speakers of German are variable and insecure: for instance the subjects reported by Rubach (1990) disagreed with the judgments provided by Vennemann 1972 and others; and native speakers like Heike (1992) indicate uncertainty about the division of medial clusters. My own attempts to elicit from native speakers syllable division judgments in TN clusters (as in *Dogma*, *Magma*) bear out this basic fact.

References

Asher, R.E. and T.C Kumari 1997. *Malayalam Descriptive Gramar*, Routledge

Barker, M.A.R. 1967. *Klamath Grammar*, UCPL 32, Berkeley and Los Angeles.

Beckman, J. 1997. *Positional Faithfulness*. PhD.Dissertation. The University of Massachusetts, Amherst.

Berger, H. 1974. *Das Yasin-Burushaski (Werchikwar):Grammatik, Texte, Worterbuch*. Otto Harrassowitz: Wiesbaden.

Blevins, J. 1993. "Klamath Laryngeal Phonology," *International Journal of American Linguistics*. 59:237-279.

Blevins, J. 1999. "The independent nature of phonotactic constraints: an alternative to syllable-based approaches," to appear in C.Féry and R. van de Weijer (eds.) *The Syllable in Optimality Theoy,* Cambridge University Press.

Boucher, V. 1988. "A parameter of syllabification for VstopV and relative-timing invariance," *Journal of Phonetics*, 1988, 16, 299-326.

Breen, Gavan n.d. "Arrernte Phonology and Antisyllables," ms.

Broselow, E. 1976. *The Phonology of Egyptian Arabic*, PhD.Dissertation. The University of Massachusetts, Amherst.

Browman, C. and L. Goldstein 1988. "Some notes on sylable structure in articulatory phonology," *Phonetica* 45, 140-155.

Browman, Cathe and Louis Goldstein 1992. "Articulatory Phonology: An Overview," *Phonetica*, 49, 155-180.

Butcher, A. 1993. *The Phonetics of Australian Languages*, ms. Flinders University, Adelaide.

Cardona, G. 1965. A Gujarati Reference Grammar. University of Pennsylvania Press, Philadelphia.

Derwing, B. 1992. "A 'pause-break' task for eliciting syllable-boundary judgments from literate and illiterate speakers: Preliminary results for five diverse languages, " *Language and Speech* 35, 219-235.

Devine, A. and L. Stephens 1994. *The Prosody of Greek Speech*. Oxford: Oxford University Press.

Dubiel, S. 1994. *Syllables and syllabification: a case study of Polish.* MA thesis, the Lublin Institute of English Philology.

Engstrand, O. 1987. "Preaspiration and the Voicing Contrast in Lule Sami," *Phonetica* 44, 103-116.

Fallows, D. 1981. "Experimental evidence for English syllabification and syllable structure," Journal of Linguistics 17, 309-317.

Fitzgerald, C. 1996. "Aspiration in Tohono O'odham" paper presented at the LSA Annual Meeting, San Diego.

Foster, M. L. 1969. *The Tarascan Language*, University of California Publications in Linguistics, 56, Berkeley and Los Angeles.

Fujimura, O., M. Macchi and T. Streeter 1978. "Perception of stop consonants with conflicting transitional cues," *Language and Speech* 21, 337-343.

Gajendragadkar, S.N. 1974. *Parsi-Gujarati: a descriptive analysis.* University of Bombay Press. Bombay.

Goldsmith, J. 1990. *Autosegmental and Metrical Phonology*, Blackwell.

Gradin, D. 1966. *Jeh language lessons.* Summer Institue of Linguistics. Manila.

Hall, T. 1992. *Syllable structure and syllable related processes in German*. Niemeyer, Tübingen.

Hamilton, P. 1995. *Constraints and markedness in the phonotactics of Australian Aboriginal languages*, Ph.D. Dissertation, University of Toronto.

Harris, J. 1983. *Spanish Phonology*, MIT Press.

Heike, G. 1992. "Zur Phonetik der Silbe," in P.Eisenberg, H. Vater and K.-H. Ramers (eds.) *Silbenphonologie des Deutschen Sprache*, Narr. Tübingen.

Henderson, J. 1998. *Topics in Eastern and Central Arrernte Grammar*. Ph.D.Dissertation, The University of Western Australia .

Ito, J. 1986. *Syllable Theory in Prosodic Phonology*, Ph.D. Dissertation, University of Massachusetts, Amherst.

Jodłowski, S. and W.Taszycki 1958. *Slownik ortograficzny i prawidla pisowni polskiej.* Zakl. Narodowy im. Ossolinskich. Wroclaw.

Jones, W. 1910. *Algonquian (Fox)* in F. Boas (ed) *Handbook of American Indian Languages.*

Jun J. 1995. *A Constraint Based Analysis of Place Assimilation Typology*, UCLA Ph.D.Dissertation.

Kahn, D. 1976. *Syllable-based generalizations in English phonology*, Ph.D.Dissertation, MIT. [distributed by Garland Press].

Kingston, J. 1990. "Articulatory Binding," in J.Kingston and M. Beckman (eds.) *Papers in Laboratory Phonology I*, Cambridge University Press, Cambridge, 406-434.

Kotchetov, A. 1999. "What makes a coda contrastive? Plain and palatalized coronal stops in Russian," to appear in E.Hume and K.Johnson (eds.) *The Role of Speech Perception Phenomena in Phonology*, Academic Press.

Krakow, R. 1989. *The articulatory organization of syllables: a kinematic analysis of labial and velic gestures.* Ph.D.Dissertation, Yale University.

Lamontagne, G. 1993. *Syllabification and consonant cluster conditions* UMassachusetts, Amherst Ph.D. Dissertation.

Lefkowitz, N. 1991. *Talking backwards, looking forwards. The French language game Verlan.* Narr, Tübingen.

Lombardi, L. 1995. "Laryngeal neutralization and syllable well-formedness," Natural Language and Linguistic Theory, 13: 39-74.

McGregor W. 1990. *A Functional Grammar of Gooniyandi*. John Benjamins'. Amsterdam; Philadelphia.

Manley, T. 1972. *Outline of Sre Structure*. University of Hawaii Press. Honolulu.

McCarthy, J. and A. Prince 1995. "Faithfulness and Reduplicative Correspondence," *University of Massachusetts Occasional Paper in Linguistics*, vol. 5, University of Massachusetts, Amherst.

Mishra, V. 1972. *A critical study of Sanskrit Phonetics,* Varanasi: The Chomba Sanskrit Series Service.

Mohanan, K.P. 1986. *The theory of lexical phonology*. Dordrecht, Reidel.

Mohanan, T. 1989. "Syllable structure in Malayalam," *Linguistic Inquiry*, 20, 589-625

Nagaraja, K.S. 1985. *Khasi, A Descriptive Analysis* , Deccan College, Poone.

Newman, S. 1944. *Yokuts Language of California*, Viking Fund Publications in Anthropology, New York.

Oftedal, M. 1956. *A linguistic Survey of the Gaelic dialects of Scotland*, Norsk tidsskrift for sprogvidenskap. Supplement. series.

Ohala, J. 1990a. "Alternatives to the Sonority Hierarchy", in M. Ziolkowski, M.Noske and K. Deaton (eds.) *Papers from the 26th Regional Meeting of the Chicago Linguistics Society, Vol. 2: The parasession on the syllable in phonetics and phonology*. Chicago Linguistic Society: Chicago.

Ohala, John 1990b. "The phonetics and phonology of aspects of assimilation," in J.Kingston and M.Beckman (eds) *Papers in Laboratory Phonology 1*, (pp. 258-275) Cambridge University Press.

Ohala, M. 1987. *Aspects of Hindi Phonology*. Motilal Banarsidass: Delhi.

Pierrehumbert, Janet and David Talkin 1992. "Lenition of h and glottal stop," in G.Docherty and R.Ladd (eds.) *Papers in Laboratory Phonology 2*, (pp. 90-116) Cambridge University Press.

Pike, Kenneth and Eunice Pike 1947. "Immediate Constituents of Mazateco Syllables," *International Journal of American Linguistics*, 13. 78-91.

Portele, T. 1995. "The Influence of the syllable boundary on consonant-consonant realizations," Proceedings of the International Congress of Phonetic Sciences, Stockholm, vol. 2, 594-597.

Prince, A. and P.Smolensky 1993. *Optimality Theory: Constraint Interaction in Generative Grammar.* Ms. Rutgers University and University of Colorado, Boulder.

Pulgram, E. 1970. *Syllable, word, nexus, cursus*. The Hague and Paris. Mouton.

Pritchett, B. 1984. *The Mys-fuckin-sterious Behavior of s-stop clusters in English Expletive Infixation.* ms.. Harvard University Linguistics Department.

Ramaswami, N. 1982. *Brokskat Grammar.* Central Instute of Indian Languages. Mysore.

Rubach, J. 1990. "Final devoicing and cyclic syllabification in German, " *Linguistic Inquiry*, 21: 79-94.

Rubach, J. 1996. "Nonsyllabic Analysis of Voice Assimilation in Polish," *Linguistic Inquiry*, 27, 69-110.

Sapir, E. 1912. *Takelma language of southwestern Oregon*, reprinted in V.Golla (ed.) *The Collected Works of Edward Sapir*, vol. 8. Mouton de Gruyter, Berlin: New York.

Saxton, D. 1963. "Papago Phonemes", *International Journal of American Linguistics*, 29, 29-35.

Schiller, Niels, A. Meyer and W. Levelt 1997. "The syllabic structure of spoken words: evidence from the syllabification of intervocalic consonants," *Language and Speech*, 40, 103-140.

Sommer, B. 1970. "An Australian language without CV syllables," *International Journal of American Linguistics*, 36.1.57-58.

Steriade, D. 1988. "Reduplication and syllable transfer in Sanskrit and elsewhere," Phonology, 5.

Steriade, D. 1995. *Positional Neutralization*, UCLA ms.

Steriade, Donca 1997. "Phonetics in Phonology: the Case of Laryngeal Neutralization", to appear in M.Gordon (ed.) UCLA Working Papers in Phonology, vol. 2.

Stevens, K. and S. Blumstein 1975. "Quantal aspects of consonant production and perception: a study of retroflex consonants," *Journal of Phonetics*, 3, 215-234.

Szpyra, J. (n.d.) "Sonority-based syllabification in Polish" ms. Lublin University.

Szymczak. M. 1975. *Slownik ortograficzny jezyka polskiego: wraz i zasadami pisowni i interpunkcji*. Panstwowe Wydawn. Naukowe. Warszawa.

Treiman, R. and C. Danis 1988. "Syllabification of intervocalic consonants," *Journal of Memory and Language*, 27, 87-104.

Turk, A. 1995. "Articulatory phonetic clues to syllable affiliation: gestural characteristics of bilabial stops," in P.Keating (ed.) *Papers in Laboratory Phonology 3*, Cambridge University Press.

Vennemann, T. 1972. "Zur Silbenstruktur der deutchen Standardsprache" in T.Vennemann (ed.) *Silben, Segmente, Akzente*. Niemeyer, Tübingen.

Waters, B. 1979. *A distinctive features approach to Djinang phonology and verb morphology*. Summer Institute of Linguistics, Australian Aboriginal Branch. Darwin.

Whitney, William Dwight 1889. *Sanskrit Grammar*, Harvard University Press.

Wiese, R. 1996. *The Phonology of German*. Oxford, Clarendon Press.

Whorf, B.I 1956. *The Hopi Language* University of Chicago Library, Chicago.