

The syllable

Donca Steriade. MIT

1. Syllable: predictability and variation in parsing; universality of preferences

A syllable is a string of segments grouped around one obligatory vowel or vowel-like (*syllabic*) element. This segment is the syllable's *nucleus*; any preceding group of consonants within the syllable is the *onset*; and any following consonants form the *coda*. The word *algebra*, if divided as *al.ge.bra*, contains an initial syllable without onset but ending in a coda *l*, followed by two syllables with onsets but lacking codas.

How a string of segments is parsed into syllables can always be predicted from the segments' feature composition and from the location of morpheme boundaries. Thus, English speakers judge that in most inter-vocalic contexts, any string Stop-r forms a single onset (as in *al.ge.bra*, *se.cret*, *ni.strate*); but a word boundary separating the consonants (as in *night rate*) induces the judgment that the stop and *r* belong to separate syllables. Such parsing judgments are not always secure or invariant across speakers of one language (see below), but this does not alter the fact that lexical items do not minimally contrast in their syllabic division.

Languages differ in limited ways in their choices of syllable assignment. Thus a string like [–izm] is monosyllabic in Russian but disyllabic [izm] in English. A Stop-r string like *abra* is divided as *a.bra* in Romance languages, but as *ab.ra*, in Lithuanian. Some languages systematically require the coincidence of word and syllable boundaries (Polish: Booij and Rubach 1989) while, in others, syllabification largely ignores intervening boundaries (Cairene Arabic: Brodbeck 1976).

What is striking however is not so much the diversity of choices involved in syllable division as the uniformity of preference for certain syllable types, a cross-linguistic effect revealed by closer analysis. Thus a V.CV parse (where V = any vowel and C = any consonant) is always preferred to VC.V: for virtually all known languages, only

V.CV is possible morpheme-internally¹. The VC.V parse arises, if at all, from the preference to let morpheme and syllable boundaries coincide (McCarthy and Prince 1993); or from the effect of other preferences that occasionally override the favored V.CV parse. The preference for V.CV over VC.V stems from the fact that both *closed* syllables (syllables with coda) and onsetless syllables are disfavored cross-linguistically: thus the V.CV parse minimizes onsetless syllables and avoids codas altogether, in contrast to VC.V (Ito 1989, Prince & Smolensky 1993).

The preference for onsets and against codas is further suggested by implicational data of two sorts. First, in any given language, the morpheme types – roots, affixes - that permit CV syllables represent a superset of the morphemes that also allow onsetless or closed syllables (McCarthy & Prince 1995). Second, all languages permit CV syllables, but some disallow onsetless or closed syllables (no onsetless in Yokuts; no codas in Tahitian). The general principle underlying these points is that preferred structures occur more widely, and are subject to fewer distributional restrictions, than dispreferred ones.

This outlines the empirical arguments advanced to support the universality of the preference for CV. The fact that many languages allow, in addition to CV, dispreferred syllable types (closed CVC, onsetless V, onsetless and closed VC) has been attributed to the effect of conflicting constraints on linguistic structure (ref to OT entry), constraints whose variable ranking relative to the CV preference yields this fragment of syllable typology (Prince & Smolensky 1993).

2. Sonority effects: nuclear sonority and sonority sequencing

2.1. Nuclear sonority

The notion of *relative sonority* (Whitney 1865; Clements 1990 for further references) makes available a broader set of universal preferences in syllable parsing. Relative sonority refers to the scale of segment types in (1):

¹ See Kahn 1976 and Treiman and Danis 1988 for stress-related complications in English; and Breen and Pensalfini 1999 for a unique case of VC.V preference.

(1)	vowels	liquids	nasals	fricatives	stops
	most sonorous				least sonorous

The properties generating this scale are unknown (Ohala 1992) but it is possible that reference to the relative amount of acoustic energy present in the lowest formants of the oral tract, F1 and F2, may succeed in drawing some of the right distinctions.

One use of the sonority scale involves the typology of syllable nuclei, about which a universal law can be formulated (Zec 1989):

- (2) **The nuclear sonority law:** If segment x is a possible nucleus in any language L, then any segment more sonorous than x is also a possible nucleus in L.

Thus English permits nasal nuclei (as in *chasm* [kæzm̩]); since liquids and vowels are more sonorous than nasals, they are correctly ruled in by (2) as possible nuclei (cf. syllabic [l] and [r] as in *settle* [setl], *acre* [ekr̩]). By comparison, Berber dialects that allow stop nuclei (Dell & Elmedlaoui 1985; e.g. [tftksṭt] ‘you-fem. sprained it’) are predicted to permit any other segment type, including fricatives, to occupy the nucleus position, since stops are at the bottom of the sonority scale: this prediction is also borne out.

The nuclear sonority law suggests that the sonority scale corresponds to a scale of preferences for syllabic nuclei: the more sonorous the segment the better the corresponding nucleus (Prince and Smolensky 1993:16). On this interpretation, the set of possible nuclei in a language can be defined by establishing a cut-off point on the scale, marking the boundary between tolerable and intolerable nuclei. The location of this boundary point is language specific, but the scale of preference itself – the distinction between better and worse nuclei – is universal. Independent evidence for the scalar preference interpretation comes from a generalization distinct from but related to (2) (cf. Dell & Elmedlaoui’s 1985 for the original proposal, as applied to Berber; and Prince and Smolensky 1993 for the cross-linguistic extension):

(3) If segment x has greater sonority than y, then a parse in which x is a nucleus is, ceteris paribus, preferred to a parse in which y is a nucleus.

An instantiation of (3) involves a Berber string such as /sawl-x/ realized as [sa.wlx] 'I spoke'. An alternative parse [saw.lx] is not excluded in principle, since Berber does allow syllabic fricatives in other contexts (e.g. txznt 'store'). To express the preference for [sa.wlx] over [saw.lx] one must invoke (3), which favors a parse with [l] over [x]. Although [x] is tolerated in this language, its lower sonority marks it as a worse nucleus than [l]: it is thus the availability of the better nucleus [l] that excludes the parse *[saw.lx].

2.2. Sonority Sequencing

The original use of the sonority scale was to identify preferred relations among adjacent segments belonging to the same syllable. This set of preferences (known as Sonority Sequencing: Selkirk 1982) can be summarized by defining first the notion of sonority peak:

(4) **Sonority Peak:** A segment a is a sonority peak iff there is no segment b, such that a and b are in the same syllable, a and b are adjacent, and a is less sonorous than b.

This definition identifies one sonority peak, [ɑ], in [harm]: this is because each segment other than [ɑ] is adjacent to one of higher sonority. In contrast, in the hypothetical syllable [hæmr], (4) defines two sonority peaks [æ] and [r]: [r] qualifies as a peak in this case because it is higher in sonority than the only segment adjacent to it, [m]. Using (4), an important preference law can be formulated:

(5) **Sonority Sequencing:** In any syllable, there is no more than one sonority peak.

This condition identifies a large class of dispreferred syllables: those with multiple peaks. Thus (5) distinguishes the disyllabic parse [hæ.mr] – two syllables, each with one peak - from the disfavored monosyllable [hæmr], with two peaks. The English judgment

that monosyllables like [hæmr] are impossible reflect awareness of (5). Note that when (5) is not at stake, the monosyllabic parse is preferred (as in *harm*; cf. *[ha.rm]) in virtue of the disfavored status of low sonority nuclei mentioned earlier.

In general, parses that generate single peak syllables are systematically preferred to alternatives, with the proviso that languages differ on whether they will tolerate lower sonority nuclei for the sake of satisfying (5). To illustrate first the general preference for satisfying (5), consider a hypothetical string like [karta], which, in virtue of (5), must contain at least two syllables, corresponding to the two [a]-peaks. Moreover, (5) excludes the disyllabic parse [ka.rta], whose second syllable contains two peaks. Indeed, most known languages reject parses of this form. To illustrate now the circumstances under which two-peak syllables like [rta] are permissible, consider Polish, a language in which only vowels can be nuclei. The Polish lexicon contains words like [rta], which begin with a consonantal, non-syllabic, sonority peak: this can be attributed to the fact that avoidance of nuclear consonants - i.e. avoidance of *[r.ta] - takes priority in this language over (5). Since Polish sanctions violations of (5) in this circumstance, one might expect that this very fact will open up two possible parses for strings like [karta]: [kar.ta] and [ka.rta], the latter involving a two-peak second syllable, identical to the initially attested [rta]. The evidence (Bethin 1992) shows however that the parse [ka.rta] is strongly disfavored. The suggestion emerging from these observations is that the only source of cross-linguistic variation regarding (5) involves instances in which sonority sequencing can only be achieved at the cost of creating low sonority nuclei: as in other cases, the variation arises here from the existence of multiple solutions to conflicting preferences. In strings like [karta], unlike in initial [rta], the conflict between nuclear sonority preferences and (5) does not arise: hence no parsing variation is expected here and indeed almost none is found².

² Cf. Mohanan (1989) for reported parses like *ka.rta* in Malayalam and Steriade (1999) for discussion. Note that the preference for open syllables is never satisfied at the expense of (5): even in Malayalam, where divisions like *ka.rta* exist, this mode of parsing relates to orthographic and word-edge properties of the language, not to coda avoidance per se.

Sonority Sequencing is also central to our understanding of a related aspect of syllable typology, the laws pertaining to possible initial and final consonant clusters (Greenberg 1978). Although these generalizations involve word edges, they reflect preferences for syllable edges as well. Greenberg's laws take the form "the presence of an initial/final C_iC_j sequence implies the presence, in the same language, of an initial/final C_nC_m sequence." and they reveal a structural preference for C_nC_m over the corresponding C_iC_j . Two clustering laws relevant to sonority sequencing can be cited. First, languages that possess initial liquid-obstruent clusters (e.g. *rta*) also possess obstruent-liquid clusters (e.g. *tra*), but not conversely. Thus Spanish (Harris 1983) permits complex onsets but forbids multiple peaks per syllable: this allows *tr-* initials and prohibits *rt*. Polish (Bethin 1992) tolerates *rt*-initials along with more frequent *tr-*. What does not occur is a language all of whose complex onsets violate (5). The mirror image of the *rt-implies-tr* law is documented for final sequences: here *tr* finals imply *rt*. In this respect too, languages appear to give priority to structures that satisfy certain structural conditions – e.g. (5): they enrich their lexical stock by letting in additional structures only if the unmarked, well-formed syllables are being utilized as well.

2.3. Other sonority profile conditions

There exist other, less securely established, preferences for the aligning sonority profiles with syllables.

Within the onset, a preference may exist for abrupt rises in sonority (Selkirk 1984, Clements 1990). This would explain the fact that gradually rising onsets like *pn*, *mr* are systematically dispreferred to more abruptly rising onsets, such as *pr*, as suggested by some of Greenberg's (1978) clustering laws. See however Ohala (1992) for an alternative explanation of this data which has somewhat greater empirical coverage: the suggestion is that the difference between *pn*, *mr*, on the one hand, and favored *pr* on the other, lies in the magnitude of the acoustic modulation between adjacent consonants.

Across the boundary between syllables, sonority rises – as in *ab.ra*, *am.ra* – whether abrupt or otherwise, are dispreferred according to Vennemann (1988).

Finally, within the syllable, it is not only the case that a single sonority peak is favored, but also that this peak will most likely be the nucleus: syllables like [jm̩] – with a sonority drop from onset to nucleus - are only rarely attested. This is a direct consequence of the preference for high sonority nuclei: to understand what makes [jm̩] possible, we have to ask what excludes the alternate parse [im], which assigns nuclear status to the vocoid, a more sonorous segment. Indeed the few attested instances of syllables like [jm̩] find their explanation in the fact that [j] is the only available onset (Dell and Elmedlaoui 1985).

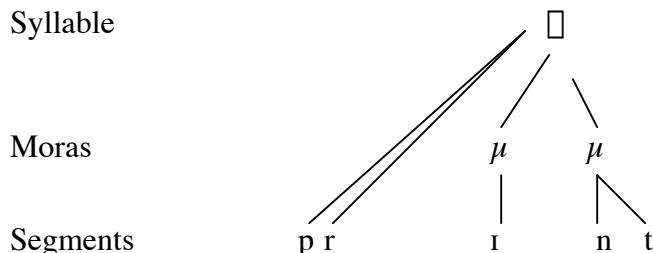
3. Syllable weight and subsyllabic constituents

Linguistic rhythm phenomena - stress and meter – frequently invoke the distinction between *light* syllables (generally those ending in a short V; but see below) and *heavy* syllables (all others). Metrically strong positions, in particular stress, tend to be associated with heavy syllables. There are two facets of this association (Prince 1992). On the one hand, if a syllable is heavy, it tends to attract stress. On the other hand, if a syllable acquires stress - for reasons possibly unrelated to its weight - then it tends to surface as heavy: syllables that are inherently light (e.g. contain reduced vowels that can't be lengthened) are systematically avoided as stress recipients.

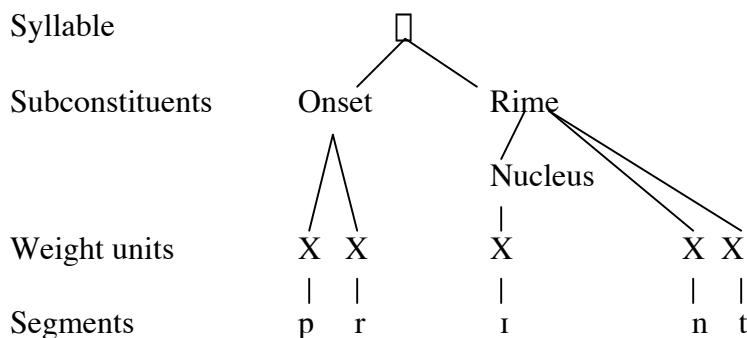
Until recently, weight distinctions were taken to be strictly binary (heavy vs. light) rather than scalar. But the number of documented languages in which more than two degrees of weight need to be recognized is steadily growing (cf. Hayes 1995). A non-binary weight scale includes, from heaviest to lightest, the list of syllables: C₀VVC, C₀VCC, C₀VV, C₀VC, C₀V (where VV represents either a long vowel or a diphthong and C₀ is a string of C's, possibly empty). Most languages collapse this extended weight scale into a binary or ternary scale, without however reversing its rankings: thus C₀VC never counts as heavier than C₀VV, and may count as lighter than C₀VV. Another recent finding (Kenstowicz 1994, Gordon 2001) is that many languages let factors other than segment count or segment length determine the weight of the syllable: these include the

height of vowels (lower vowels make for heavier syllables) and the distinction between coda obstruents (less likely to contribute weight) and coda sonorants (more likely).

Theories of syllable-internal structure have been proposed which relate to the representation of weight. Their goal is, in general, to discover constituent structures from which certain facts about weight computation follow, in particular the fact that onsets do not affect the weight of the syllable: CV is as light as CCCV, while CCCV is lighter than VC. For instance, moraic representations (Hyman 1985) assign a maximum of two and a minimum of one weight elements (or *moras*) to every syllable: segments contributing weight (V's; selected coda C's) associate to these moras, while onsets, because they are weightless, attach directly to the syllable node. The resulting structure is illustrated below by the syllable *print*:



An alternative representation of weight (Halle and Vergnaud 1978, McCarthy 1979, Levin (Blevins) 1985) holds that every segment possesses one weight unit (here an X), and long segments possess two. Weight is read off a representation of syllables in which nuclei and codas - the potentially weight bearing elements - form one constituent, the rime: the syllable is heavy iff its Rime or, for some languages, its Nucleus, dominate more than one X.



This body of work is further surveyed by Hayes (1995). However, the status of the empirical generalizations on weight is now shifting: it is no longer clear that weight distinctions are binary, as the two representational theories reviewed above assume. Nor is it clear that the weight representations designed for the analysis of stress are equally useful for other phenomena: tonal association, compensatory lengthening (Hayes 1989) or minimality conditions (Hayes 1995). An alternative view (Gordon 2001) is that different phonetic properties of the syllable are relevant to different phonological phenomena: on this view stress will access representations derived from durational properties of a syllable (since duration is a correlate of stress), or from measures that combine duration and amplitude, while tone will access representations derived from properties involving strictly the realization of F0 contours.

4. Knowledge of syllable structure and its sources

The syllable is perhaps the only construct of phonological analysis about whose existence untrained speakers appear to possess conscious knowledge (Liberman et al. 1974). Yet no aspect of the acoustic signal has been found to uniquely identify syllable boundaries (Laver 1994:113ff; but cf. Maddieson 1985); and attempts to pinpoint the acoustic/perceptual basis for judgments of syllable division suggest that speakers can tell apart, say, V.CV from VC.V, primarily by relying on the phonetic properties identifying *word* boundaries (glottal stops, silent intervals, aspiration: Gårding 1967), or on lexical knowledge, not syllable edges per se. Despite this, speakers are not only able to count syllables but also reach substantial agreement on some aspects of syllable division (Treiman 1988).

Several unresolved issues arise here. When subjects agree on some parses, we do not know what the consensus is based on: for instance, what suggests to English learners that *algebra* is *al.ge.bra* and not *al.geb.ra*? And, when subjects fail to agree on other aspects of syllable division, the roots of disagreement remain equally unclear. Thus Derwing's (1992) data on VCV division reflects agreement on the division of some strings (e.g. 86% of responses favor *de.mand*, 82% favor *mo.ment*) but uncertainty on others (51% responses for *lem.on*, 37% for *le.mon* and 12% for *lem.mon*, with ambisyllabic *m*). The selective areas of consensus, as well as the lack of clear perceptual correlates of syllable division, suggests that speakers learn their syllables through an inference process in which a number of possibly conflicting factors play a role: knowledge of possible word-initial and final sequences (as these are necessarily syllable-initials and finals as well), and certain universal preferences outlined earlier. Lack of consensus may arise in cases where these factors yield contradictory guidelines to division (Steriade 1999): thus the *lem.on* vs. *le.mon* variation may reflect the conflict between the preference for V.CV vs. the knowledge that the [ɛ] of *lemon* is not a possible word final segment.

5. Syllabic explanations and alternatives

These unresolved issues have suggested to some that syllables are dispensable units of analysis (cf. Dressler and Dziubalska-Kolaczyk 1992). This view is yet to be reconciled with evidence that speakers group segments into syllables in production (Krakow 1999) and perhaps in perception (Mehler et al. 1981), as well as with basic facts of phonological analysis. For instance, the optimal distance between stresses or between stresses and the edges of stress domains (Prince 1983) is typically measurable in syllables, not in milliseconds, segments or sonority peaks. And certain common distributional restrictions – e.g. Berber schwa occurs only before CC or C# - find a uniform statement in syllabic terms: Berber schwa occurs in only closed syllables.

However it is also possible that the explanatory role of the syllable has been occasionally overstated. Thus processes like final devoicing or the place neutralization,

which tend to apply before a consonant or at the end of the word, have been analyzed since Hooper (1972) and Vennemann (1972), as syllabically based: final devoicing, for instance, is said to apply to coda consonants. In this case, however, a closer look at the typology of voice and place neutralization suggests that it is not the coda position per se that causes laryngeal or place neutralization but rather the contextual loss of perceptual cues to voice or certain place distinctions (Steriade 1999). It remains to be seen what class of segmental processes are best understood as syllabically conditioned.

References

Bethin, Christina 1992 *Polish Syllables: The Role of Prosody in Phonology and Morphology*, Columbus, Ohio: Slavica Publishers, Inc.

Blevins, Juliette (2001) “The independent nature of phonotactic constraints: An alternative to syllable-based accounts”. In C. Féry and R. van de Vijver (eds.), *The syllable in Optimality Theory*. Cambridge University Press.

Booij, Gert and Jerzy Rubach 1989 “Edge of constituent effects in Polish”, *Natural Language and Linguistic Theory*

Breen, Gavan and Rob Pensalfini 1999 “Arrernte: A Language with No Syllable Onsets,” *Linguistic Inquiry*, 30, 1 – 25

Broselow, Ellen 1976 *The Phonology of Egyptian Arabic*, UMassachusetts, Amherst, PhD. thesis

Clements, G.N. 1990 “The role of the sonority cycle in core syllabification,” in John Kingston and Mary Beckman (eds.) *Papers from the 1st Conference in Laboratory Phonology: Between the Grammar and the Physics of Speech*, Cambridge University Press. 283-333

Dell, François & Mohammed Elmedlaoui 1985 “Syllabic Consonants and Syllabification in Imdlawn Tachlhiyt Berber, *Journal of African Linguistics*, 7, 105-130.

Derwing, Bruce 1992 “A pause-break task for eliciting syllable boundary judgments from literate and illiterate speakers: preliminary results from five diverse languages,” *Language and Speech*, 35, 219-235

Dressler, Wolfgang and Katarzyna Dziubalska-Kolaczyk 1992 “Evidence of syllabic writing against the phonological syllable,” in W.Dressler, M.Prinzhorn and J.R.Rennison (eds.) *Phonologica 1992: Proceedings of the 7th International Phonology Meeting*. Torino: Rosenberg and Sellier.

Gårding, Eva 1967 *Internal juncture in Swedish*, Travaux de l’Institut de Phonétique de Lund: Gleerup.

Gordon, Matthew (to appear) 2001 “A phonetically-driven account of syllable weight,” *Language*, vol. 77.

Greenberg, Joseph 1978 “Some Generalizations Concerning Initial and Final Consonant Clusters,” in Joseph Greenberg (ed.) *Universals of Human Language*, vol.. 2: *Phonology*., 243-280.

Halle, Morris and Jean-Roger Vergnaud 1978 “Metrical Structures in Phonology,” unpublished ms. MIT.

Harris, James 1983 *Syllable structure and stress in Spanish : a nonlinear analysis* Cambridge, Mass. : MIT Press, 1983.

Hayes, Bruce 1995 *Metrical stress theory : principles and case studies*, Chicago : University of Chicago Press.

Hooper, Joan 1972 “The Syllable in Phonological Theory,” *Language*, 48, 525-540.

Hyman, Larry M. 1985 *A theory of phonological weight*, Publications in language sciences Dordrecht, Holland: Foris

Ito, Junko 1989 “A prosodic theory of epenthesis,” *Natural Language and Linguistic Theory*, 7, 217-259

Kahn, Daniel 1976 *Syllable-based generalizations in English phonology*, MIT Ph.D. thesis

Kenstowicz, Michael 1994 “Sonority-driven stress”, *Rutgers Optimality Archive* 33-1094.

Krakow, Rena 1999 “Physiological organization of syllables: a review,” *Journal of Phonetics*, 27, 23-54

Laver, John 1994 *Principles of phonetics*, Cambridge textbooks in linguistics, Cambridge; New York, NY : Cambridge University Press, 1994

Levin (Blevins) 1985 *A metrical theory of syllabicity*, MIT Ph.D. thesis.

Liberman, I.Y., D. Shankweiler, F.W. Fisher, B. Carter 1974 “Explicit syllable and phoneme segmentation in the young child”, *Journal of Experimental Child Psychology*, 18, 201-212

Maddieson, Ian 1985 “Phonetic cues to syllabification,” in Victoria A. Fromkin (ed.) *Phonetic linguistics : essays in honor of Peter Ladefoged*, Orlando : Academic Press, 1985.[pages]

McCarthy, John J. 1979 “On stress and syllabification,” *Linguistic Inquiry*, 8, 249-336

McCarthy John J and Alan Prince 1993 “Generalized Alignment” in G.E.Booij and J.van de Marle (eds.) *Yearbook of Morphology 1993*, Dordrecht: Kluwer 79-153

McCarthy John J and Alan Prince 1995 “The emergence of the unmarked: optimality in prosodic morphology” in M.Gonzàlez (ed.) *Proceedings of the North East Linguistic Society* 24, 333-379.

Mehler, J., J. Y Dommergues, U. Frauenfelder, and J Segui. 1981 The syllable's role in speech segmentation, *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305.

Mohanan, Tara 1989 “Syllable structure in Malayalam,” *Linguistic Inquiry*, 20, 589-625

Ohala, John 1992 “Alternatives to the Sonority Hierarchy for Explaining Segmental Sequential Constraints,” in M.Ziolkowsky, M.Noske and K.Deaton (eds.) *Papers from the Chicago Linguistic Society* 26:2, 319-339

Prince, Alan 1983 “Relating to the Grid,” *Linguistic Inquiry* 14, 19-100

Prince, Alan 1992 “Quantitative consequences of rhythmic organization,” in M.Ziolkowsky, M.Noske and K.Deaton (eds.) *Papers from the Chicago Linguistic Society* 26:2, 355-398

Prince, Alan & Paul Smolensky 1993 *Optimality Theory: Constraint Interaction in Generative Grammar*, Technical Report #2 of the Rutgers University Center for Cognitive Science, Rutgers University.

Selkirk, Elisabeth 1982 “The syllable,” in Harry van der Hulst and Norval Smith (eds.) *The Structure of Phonological Representations*, 2, 337-384.

Selkirk, Elisabeth 1984 “On the Major Class Features and Syllable Theory,” in Mark Aronoff and Richard T. Oehrle (eds.) *Language sound structure : studies in phonology / presented to Morris Halle by his teacher and students* Cambridge, Mass: MIT Press, 107-113.

Steriade, Donca 1999 “Alternatives to syllable-based accounts of consonantal phonotactics,” in O.Fujimura, B.Joseph and B.Palek (eds.) *Proceedings of LP 1998*, Prague: The Karolinum Press. 205-245.

Treiman and Danis 1988 “Syllabification of Intervocalic Consonants,” *Journal of Memory and Language*, 27, 87-104

Vennemann, Theo 1972 “On the theory of syllabic phonology,” *Linguistische Berichte*, 18, 1-18.

Vennemann, Theo 1988 Preference laws for syllable structure and the explanation of sound change : with special reference to German, Germanic, Italian, and Latin; Berlin ; New York : Mouton de Gruyter.

Whitney, William Dwight 1865 “On the relation of Vowels and Consonants” reprinted in Michael Silverstein (ed.) 1971 *Whitney on Language: Selected Writings of William Dwight Whitney*, Cambridge, Mass.: MIT Press.

Zec, Draga 1989 *Sonority Constraints on Syllable Structure*, Stanford University PhD. thesis.