Péter Szigetvári

Why CVCV

One of the most important achievements of modern linguistics is the discovery of the use of emptiness. Its relevance can be likened to that of the concept of zero, without which it is hard to imagine the progress natural sciences have made in the last few centuries. One aim of the discussion that follows is to convince the reader that empty positions in the phonological skeleton are not merely a tricky device to ease the analysis, but rather a logical conclusion of various different lines of thought pursued by theorists of modern phonology.

Once one accepts the possibility of empty skeletal positions it is fairly obvious that they must be exploited. The evident use of the idea is that syllable structure can be radically simplified if we are not obliged to take surface adjacency in phonological strings to be an unrefutable symptom of adjacency in the underlying representation as well. In the second part of this paper I argue for the radically reduced syllable constituency proposed by Lowenstamm (1996) and advocated by his students. The skeleton can be likened to a binary system, each skeletal string contains strictly alternating consonantal and vocalic positions.

The paper is structured as follows. The first three sections argue for the existence of empty skeletal positions: section **1** shows that the widely accepted autosegmental framework of phonological representations calls for the recognition of melodiless points on the skeletal tier. In an excursus in section **2** on the competing models of representing melody I argue for the theoretical desirability of unary features and show why such a framework again necessitates the acceptance of empty segments, i.e., skeletal points without melodic content. Then, in section **3**, the theoretical status of empty skeletal points is discussed: I claim that it is only a question of scholarly tradition that many phonologists believe the null hypothesis to be a skeleton where only the pronounced segments are furnished with a slot. In the

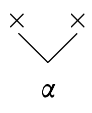
* People who have directly influenced this text are András Cser, Csaba Csides, Péter Dienes, Péter Siptár and Miklós Törkenczy. I was also supported by the Zsigmond Telegdi Scholarship. I thank them all.

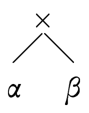
second part of the paper I survey different theories of syllable structure beginning with the standard argument for having syllables, some widespread ideas of parsing phonological strings into syllables (section **4**). After pinpointing some problems with these methods, I will introduce a framework that aims at solving them by positing empty positions in the skeleton, which by this time are hoped to be received members of the conceptual community (section **5**). Section **6** collects arguments brought up in favour of codas and aims at refuting the necessity of each. In section **7**, we are going to arrive at the most restrictive syllabic framework possible, CVCV phonology, and the last section (**8**) mentions some further arguments against syllabic constituency of any kind.

1 The skeleton–melody relationship

A not so recent advancement in phonological theory is the recognition of the necessity of separating the quantitative and the qualitative aspects of segments. In this line of research, first called prosodic, later reinvented as autosegmental phonology, the quantitative aspect is represented by a so-called skeletal tier, the qualitative by the melodic tier. The exact content of these two tiers is one of the most important issues of current research. The relationship between the elements of the two tiers is negotiated by association lines.

With the advent of the autosegmental model, it becomes necessary to explore the consequences of non-biunique relationships between the two. Having one batch of melody defining primes associated to two skeletal positions is the best-known and probably least controversial option, standardly employed to represent some acoustic property stretching across multiple timing slots (1a). The realization of this configuration ranges from long vowels ([aː]), through genuine geminate consonants ([tː]), to partially identical clusters, like adjacent monomorphemic homorganic consonants ([mb]). The complementary configuration — two distinct sets of melodic features linked to the same skeletal position — is less obviously plausible (1b).

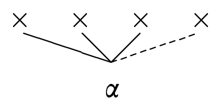
(1) a. 

b. 

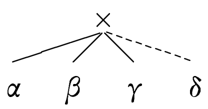
Unless one hypothesizes some kind of segment internal structuring of features, more specifically, a root node or something similar, it is hard to see how the two sets (α and β) could be kept separated. This kind of structure is occasionally proposed for the representation of affricates and pre- or post-nasalized stops. The idea is that incompatible features (or feature values) are linked to the same slot, noncontinuant and continuant in the first case and nonnasal and nasal in the second, yielding contour segments.[1] Alternative accounts are also available for these types of segment — affricates and stops, for example, may be distinguished by headedness, cf. Schafer 1995 and Szigetvári 1997 — , therefore accepting the idea of one-to-two association between a skeletal slot and melodic material is not an absolute necessity.[2]

Association of melodic material and skeletal slots includes not only one-to-two, but also one-to-three, one-to-four, etc., associations (2a). What is intriguing is that while such configurations cannot be excluded — vowel harmony and tone phenomena very often exemplify unbounded spreading of melodic material through longer skeletal strings — , three-long consonants ([t::]) or vowels ([a::]) are standardly explained away and supposed to be noncontrastive even if phonetically existent. To constrain representations theories may stipulate an upper limit of two on branching constituents (like nucleus and onset), this option, however, is not available for the type of skeletal organization I will argue for in this paper. (For a possible solution see Szigetvári (in prep.).) Similarly three-part contour segments (2b) are spurious, although Lass (1984 : 113) (citing Stephen Anderson) claims that medionasalized stops ([^{d}n^{d}]) occur in some languages, such segments are far too marginal to cater for in a first approach.[3] The difficulty with keeping contour segments from excessive arborescence increases one's desire to totally deny the existence of such structures.

(2) a.



b.



Two further options that deviate from the boring one-to-one relationship are available in an autosegmental model. One is melodic material

[1] The so-called light diphthongs, which are monomoraic, are also frequently conceived of as contour segments. In their case the features are usually compatible, still unmerged.

[2] Note that here and below I take α, β etc. to be segment-sized chunks of melody. Individual melodic primes, i.e., features, may obviously assume the configurations (1b) and (2b).

[3] Also, contrary to naive expectations, such contours can only occur with a very limited set of segments, definitely not three different types, for example.

without an associated skeletal slot. Such floating features are very useful in handling alternations where in what looks like the base form of a word there is nothing to indicate the presence of melodic material surfacing in some other, oblique form. An alternation of this type is especially intriguing when there exist other words with phonologically similar base forms which fail to manifest the same alternation. Such is the case in, for example, liaison phenomena, like the textbook RP[4] *grammar is* [græmə r ɪz] vs. *gamma is* [gæmə ɪz], where the base forms are [græmə] and [gæmə], respectively. The presence of the [r] in the first but not in the second case is neatly explained by assuming that [græmə] is lexically furnished with an [r] that lacks (or is unassociated to) a skeletal slot, while [gæmə] has no [r] of any kind, as shown in (3).

(3) a.
$$\begin{array}{ccccccccc} \times & \times & \times & \times & \times & & \times & \times & \times \\ | & | & | & | & | & & & | & | \\ \text{g} & \text{r} & \text{æ} & \text{m} & \text{ə} & \text{r} & & \text{ɪ} & \text{z} \end{array}$$

b.
$$\begin{array}{cccccccc} \times & \times & \times & \times & & \times & \times & \times \\ | & | & | & | & & & | & | \\ \text{g} & \text{æ} & \text{m} & \text{ə} & & & \text{ɪ} & \text{z} \end{array}$$

Such an account avoids the use of brute force deletion to explain the failure of the [r] to surface in case no vowel-initial string follows (e.g., *grammar book* [græmə bʊk]). It also presupposes that phonetic interpretation proceeds on the skeleton, realizing those and only those portions of the melody that are associated with the skeleton. If the mere presence of melodic material in the representation were enough for its being phonetically interpreted the option of unpronounced floating melody would not be viable.

The complementary situation is obviously a skeletal slot without any melodic content associated to it. This configuration comes handy again in dealing with liaison phenomena: for the floating liaison consonant to be interpreted it must be linked to a skeletal position. Since such consonants are typically pronounced only if a vowel-initial word (or suffix) follows, all need be hypothesized is that such words carry a skeletal slot at their beginning which is not associated to any melodic material lexically, like at the beginning of *is* [ɪz] in (3). The floating melody thus has a chance to associate and hence get interpreted. Though this account appears elegant at first sight,

[4] This dialect is sometimes claimed to be nonexistent outside prescriptively biased books on English pronunciation. It nevertheless illustrates the case in point.

there is some theoretical difficulty with it. If the phonetic interpreter takes consecutive skeletal positions as its input and realizes whatever melody is linked to each, one may wonder what should happen when it encounters a position to which no melody is associated. This problem brings us to query the nature of melodic primes, which is exactly what we are going to do in the next section.

2 The representation of melody

Western writing systems segment the continuous flow of the speech signal quite uniformly as regards segment size; disputes of what constitutes one or two segments are much less common — though not unheard of — than disputes about the domains of larger (syllable, foot) and smaller (subsegmental) units. A more robust segmentation, e.g., into chunks of the size of what is called the syllable, would ignore the fact that these units are recurrently made up of the same types of parts: syllable beginnings are by and large freely combinable with syllable endings (i.e., [ta ti tu] and [ta ma sa] are usually all possible in a given language). Therefore, a framework not analysing syllables into smaller segments would face a significant loss in economy. But it is also evident that the traditional segments labelled by the letters of the alphabet are not atomic: for one thing, phonological processes manipulate parts of these segments independently of the rest, for another, sets of segments recurrently pattern together in all sorts of different languages. Take voice assimilation, for example: in many languages two adjacent obstruents come to agree in voicing. In doing so one of them loses its own voicedness or voicelessness and assumes that of the other. For this reason, the analyst must posit voicedness and voicelessness as properties of segments distinct from the rest. Furthermore, obstruentness must be capturable, there must be something in common in all obstruents. Like voicing/voicelessness, place of articulation and many stricture properties also exhibit independent behaviour, leading phonologists to see segments as being made up of smaller subcomponents.

The representatives of sound properties are standardly referred to as features. Features usually come in one of two flavours: either as standalone units that indicate a property by being present in the representation — these are called unary[5] features —, or as ordered pairs consisting of a feature and its value. The latter, more complex type has two subtypes, binary features, which can have one of two values, usually marked '+' and '−' and

[5] Unary, as if they had one value, but these are in fact valueless.

scalar features, which can characterize the prominence of the given property with greater precision, typically marked by integers (e.g., 0VOICE, 1VOICE, 2VOICE, ... *n*VOICE).

2.1 Scalar versus binary/unary features

The theoretical difference between scalar features on the one hand and binary and unary features on the other is obvious: the former allows a theory to express much more subtle distinctions in the dimension of a certain phonological property than the latter two. There is, however, not very much need for such subtle distinctions: oppositions in natural language are overwhelmingly binary, that is, a property is either present in an expression or absent, there is hardly ever any need to make reference to more than two values of a feature.[6] This is not to say that a language may not have more than two degrees of, say, voicing on the surface, but as regards their phonological behaviour segments with a lesser or greater degree of voicing will always line up with one or the other pole, a given segment will always behave as either voiceless or voiced. Those cases which appear to call for a scalar feature analysis certainly do not immediately warrant the introduction of such objects. Because they are so powerful, one has to show first that scalar features are absolutely unavoidable, that they cannot be replaced by the simpler binary or unary features.

This manner of arguing for the rejection of scalar features may reasonably provoke an attack on the tacit assumption behind it: why should features be of the same type? Why could we not have some scalar, some binary and some unary features in our theory? Lass (1984 : 102f) asks these questions and calls the kind of stance I take here "the atomic fallacy." He puts it down primarily to Jakobson's and Chomsky & Halle's "cognitive" or "psycholinguistic" bias. While this may be true historically, I have the impression that mere theoretical elegance adequately justifies the desire that the primitives of a theory be uniform. Furthermore, the simpler the primitives, the more plausible they are. A unary feature is obviously simpler than a binary or a scalar feature in that it is one bit of information, while the others are two: the feature name and its value. Occam bears witness to categorial uniformity too: it is not only the number of categories but also the number of types of categories that must not be multiplied in vain.

[6] Natural language is unable to count, i.e., all it is sensitive to is whether something exists or not.

2.2 Binary versus unary features

Returning to a theoretical comparison of the types of phonological primes that have been suggested, we are left with unary and binary features as possible candidates. To illustrate the difference between them let us consider the way they express the voiced–voiceless contrast typically exhibited by obstruents. If we represent the segment that is the common denominator of both [t] and [d] as T,[7] i.e., an expression from which laryngeality has been abstracted, then in a binary framework [t] will be {T,−VOICE}, while [d] will be {T,+VOICE}. The typical unary framework has [t]={T} and [d]={T,VOICE}. The most conspicuous difference between the two approaches is that with binary features the number of units present in the representation of [t] and [d] is equal, while with unary features one of the contrasting segments ([t] in the case examined[8]) contains one unit less. Since usually there is no special theoretical relevance attributed to the feature values in binary frameworks, i.e., −VOICE and +VOICE are theoretically equivalent to +VOICELESS and −VOICELESS, respectively, this theory does not distinguish between the status of [t] and [d] in a given system. Consequently, in a feature theory containing only binary features privative oppositions are inexpressible as such, every opposition will appear as equipollent. Nevertheless, privative and equipollent oppositions may be modelled by a binary feature framework. In a privative opposition the value of only one feature is different in the two parties, if more than one feature values differ we face an equipollent opposition.

At this point, there appears to be no difference between the contrastive capacity of a unary and a binary feature: both types can distinguish two objects. With the introduction and rather general acceptance of underspecification theories (e.g., Kiparsky 1985, Archangeli 1988, etc.) the situation has changed radically. If a feature is allowed to have no value in addition to having its two explicit values, analysts will soon be using it with three distinct values: ∅VOICE, +VOICE and −VOICE. Various proposals have been put forward to manage this problem. Constraining underspecification to cases when the missing feature value is predictable or nondistinctive (or both) still makes the filling in of default feature values necessary and, in

[7] This symbol probably reminds the reader of the Trubetzkoyan archiphoneme. It is indeed similar in certain respects, but must not be equated with it. One may think of it as an underspecified segment, like the Firthian phonematic unit.

[8] It must be noted that [t]={T,+VOICELESS}, [d]={T} is also viable, though less often argued for.

addition, in some cases the fill-in rule and other phonological rules have to be extrinsically ordered.

The force driving underspecification is lexical minimality: "underlying representations must reduce to some minimum the phonological information used to distinguish lexical items" (Steriade 1995 : 114). Not assigning any value to certain features in the lexical representation is believed to lessen the burden of the memory. The same move, however, increases computation time and the direction the balance is tilted in is far from being obvious (Harris & Lindsey 1995 : 47f). Underspecification captures the insight that the two values of a feature are usually in an asymmetrical relationship. The problems with this approach stem from the obsession with full specification, the idea that "the output of the phonological component must contain fully (or at least maximally) specified feature matrices" (Steriade 1995 : 114). To take an example, in many languages sonorants only occur as voiced; this fact is encoded in the representation by not assigning any value to the VOICE feature. Underspecification theories nevertheless will require that sonorants get the '+' value for the VOICE feature during the derivation, even though this property is irrelevant for sonorants even at the surface. The voicing of sonorants is invisible for phonology throughout the derivation up to the point where the default rule fills in the value. If the filling-in process is allowed to occur before other phonological rules, i.e., rules may refer to the default value of the VOICE feature in sonorants, the theory faces the danger of serious overgeneration. If it must be the last step of the derivation, one wonders why have it at all.

2.3 Unary features

A decisive step, taking underspecification to its logical conclusion, is to dispense with fill-in rules altogether. Returning to the previous example, this would mean not specifying sonorants for voicing at all.[9] The significance of applying such features throughout the representation is the fact, already noted, that a privative phonological opposition is now represented by two segments one of which has one feature less than the other. Since unary features typically represent the marked value of the given opposition, the segment comprising n features will be more marked than the one comprising

[9] Three objections come to mind that could be raised against such a proposal: (i) how will the phonetic interpreter know that sonorants are pronounced with vocal fold vibration?; (ii) how come some sonorants seem to spread voicing?; (iii) how can some languages contrast voiced and voiceless sonorants? I address these questions in Szigetvári 1998 : 227f, for now let us simply assume that they can be dealt with.

$n-1$ features. There will be found segments that are even less marked, comprising $n-2$, $n-3$, ... features; eventually we arrive at a very unmarked segment that is represented by a single unary feature. It will at this point be justified to say that the given feature is not an abstract phonogical entity, but a concrete physical phenomenon, the sound it is used to represent.[10] A basic claim of the mainstream theories, that sounds are not atomic, has to be modified: some sounds are not atomic, but can be analysed as the combination of other less complex sounds. But the ultimate building blocks of sounds are themselves sounds, which are, in fact, atomic. This idea features very strongly in Government Phonology, especially in Harris 1996 and Harris & Lindsey 1995.

Single-feature segments are still marked: each feature is a mark. They form privative oppositions with two-feature segments on the one hand, and with the featureless segment on the other. The interpretation of the featureless segment must be the acoustically most unmarked segment, which, however, may not coincide with the segment most often occurring cross-linguistically. It very often occurs in language that the most common expressions do carry some mark to ease perception.

We can conclude that it follows from a model having exclusively unary features that empty segments, i.e., skeletal positions unassociated with any melodic material, ought to exist, furthermore they must be phonetically interpretable.[11]

[10] This is so unless the representation of a segment contains more than features and association lines linking them to a skeletal position. Such objects — root nodes, feature geometries of various flavours etc. — are often assumed in other theories, but disregarded here.

[11] Péter Dienes (p.c.) points out that this does not follow. Indeed, the matter depends on the meaning of skeletal slots, whether they represent a segment of the speech flow or a more abstract notion which may be interpreted as silence as well.

3 Empty skeletal positions and the null hypothesis

One way of classifying current phonological theories is by the criterion whether they allow skeletal positions to be empty or not. The stance one adopts in this issue is of substantial relevance to the whole of a given theory. There are several questions that the existence or nonexistence of empty skeletal positions bears upon. To mention but a few: the association of segments in phonological strings to syllabic constituents will be seen radically differently if empty positions may occur and cases of segments alternating with zero must also be analysed differently if we are reluctant to accept that a skeletal position may be empty: the destructive, non-monotonous device of resyllabification is very often called for if one wishes to have only positions with melodic content on the skeleton.

Taking the first case, let us assume the conventional syllable structure comprising an onset, a nucleus and a coda. In the standard textbook account all three constituents come with a practically unbounded branching potential, i.e., the onset in English may contain 0–3, the nucleus 1–2 and the coda 0–5 segments (e.g., Giegerich 1992 : 153, 167). Being empirically correct this analysis fares well for a description but is unusable when searching for an explanation; the number of branches for each constituent ranges within patently stipulative limits. One wonders why the onset may contain up to three segments, what inhibits it from having, say, four. The tacit assumptions behind this analysis are the axioms that syllable boundaries necessarily coincide with word boundaries[12] and that segments are fully integrated into the prosodic hierarchy, that is, each segment belongs to some syllabic constituent, each syllabic constituent belongs to some syllable and so on. The unfoundedness of the first axiom becomes apparent if we consider that on another level of the prosodic hierarchy, that of feet, boundaries do not necessarily coincide; words may begin with a degenerate foot and may end with a sole stressed syllable, which is not usually referred to as a degenerate foot, it still lacks a dependent second syllable. The second axiom, full integration of segments, has to be given up by theorists following this line as soon as it is realized that word edges tolerate a wider range of phonotactic

[12] E.g., Blevins 1995 : 209: "In all languages, syllable edges correspond with word/ utterance edges..."

freedom,[13] and to handle such phenomena the notion of extrasyllabicity has to be introduced.[14]

There is yet another reason why Giegerich's (or other analysts' similar) constraints are spurious theoretically: while the two consonantal constituents, the onset and the coda may be empty, i.e., they may contain zero segments, the same possibility is not available for the vocalic portion of the syllable, the nucleus. The excuse that may be brought up to explain this discrepancy is the head status of nuclei; as the head of the syllable they must not be empty. Again, if we move to other levels of the prosodic hierarchy the situation is different: both headless feet and headless segments[15] are possible.

As for segment–zero alternations, we have already seen a case where hypothesizing an empty skeletal position facilitates the analysis: liaison phenomena are neatly describable by positing an empty consonantal position before vowel-initial words. To take another instance, this time a vowel alternating with zero, consider the onset [m] of the unsyncopated [fæməli], which becomes a coda in the bisyllabic [fæmli]. A similar but converse situation often arises with morphological concatenation, e.g., the coda [k] of *take* becomes an onset in *taking.* Both of the latter cases involve resyllabification in theories that reject the possibility of having empty skeletal positions. Resyllabification, however, subverts the result of core syllabification, thereby representing a serious challenge to phonological parsing and violating the principle of monotonicity. One could argue that resyllabification is necessary because a word-final or preconsonantal consonant behaves differently from its prevocalic alternant. This, of course, is true, but one must also admit that resyllabification is simply a way of representing this fact, nothing that would offer any explanation. In such a framework we know a consonant is in coda position because it behaves like consonants in coda position usually do. Since *being* in coda position is not an empirical issue, codas have no theory-external status, we have no independent evidence for the codahood of a consonant apart from the fact that it behaves like other consonants that we believe to be in the coda.

[13] If syllables in English could in fact begin with three consonants and end in five, we would expect eight-consonant-long intervocalic sequences within words, but this also turns out to be a disappointed expectation.

[14] E.g., Goldsmith 1990 : 123: "*prosodic licensing,* which require[s] that all elements be a member of some syllable, or else be marked as contingently extrasyllabic."

[15] The head–nonhead distinction in segments is not universally accepted, but cf. Anderson & Ewen 1987, Kaye et al. 1985, Schafer 1995, among others.

What apparently justifies theories of the skeleton that reject the possibility of empty positions is the assumption that this is the null hypothesis. That is, empty skeletal positions ought not to be posited unless there is no other way to analyse phonological phenomena. While it is true that accepting skeletal positions that fail to be interpreted phonetically does bring some abstractness into a theory, it is controversial whether their rejection *is* the null hypothesis. The generative power of a theory having syllables of an unlimited size is just as excessive (if not more) as that of one having empty skeletal positions.

I will now argue against this, aiming to show that the prevalent view rejecting empty positions is somewhat accidental, relying on tradition. Let us imagine that modern phonological theory happened to be developed by people who spoke languages like Desano or Zulu, in which phonological domains are exclusively CVCV…CV on the surface. If they spoke about syllables at all, these phonologists would surely claim that syllables have the structure CV — or onset–nucleus, for us, Indo-European phonologists. After encountering more complex languages like Luo or Krenak, which allow consonants word-finally, or others like Japanese, which allow certain types of consonant clusters word-medially, it is far from obvious that our hypothetical phonologists would extend their syllable template by adding a further optional C position at the end. It is just as likely that they would hypothesize an empty V position between the two consonants or after the word-final one. In fact empirical evidence exists for the latter assumption: Harris & Gussmann (1998 : 141) claim that all syllabic writing systems assign such offending consonants to an independent syllable with an uninterpreted vowel (dummy syllables as Harris & Gussmann refers to them). Now one may argue that this is so for reasons of economy: one needs much less new symbols for the vowelless syllables — the number of all consonants in the worst case, but coda consonants are typically only a small subset of all consonants in a given language —, while, again in the worst case, the number of all CV syllables (maximally $C \times V$, where C is the number of consonants and V of vowels in the language) would be multiplied by the number of coda consonants if CVC symbols were introduced. However, one also has to admit that economical considerations do feature in scientific theories as noted by Occam quite a while ago, furthermore, once dummy syllable symbols are used literate people will unavoidably analyse their language as having empty vowels at certain points in phonological strings and phonologist will follow this tradition and, more importantly, they will take it to be the null hypothesis.

I hope to have shown that while the acceptance or otherwise of empty skeletal positions appears to be a matter of scholarly taste (analyses applying both approaches abound, after all), laying the burden of proof on theories with empty positions thinking that we have the null hypothesis on our side is not right after all. What the null hypothesis is in this issue is most probably a question of tradition.

4 Syllable structure

Many current theories of phonological representation assume one or more levels between feet and the skeleton in the prosodic hierarchy. These are occupied by so-called syllabic constituents which organize skeletal positions and other syllabic constituents into syllables. Syllabic constituents gain theoretical relevance when they prove to be indispensable in — or at least result in a substantial simplification of — the formulation of phonological generalizations.

Syllables, on the other hand, are not uncontroversial entities. The notion has been abandoned several times in the history of phonological theory, the best known case is probably that of the SPE (Chomsky & Halle 1968). From the 1970s mainstream phonology has gradually returned to applying this traditional concept, but interestingly in most cases it is not the syllable constituent itself that is necessary for the analyses, but its subconstituents, the onset, the nucleus and the coda.

4.1 Why have syllable structure?

It has been noticed — e.g., by Kahn (1976) — that certain consonantal processes favour the phonological environment depicted in (4).

(4) $$\_\left\{ \begin{array}{c} \mathrm{C} \\ \# \end{array} \right\}$$

If syllables have a theoretical status, the environment in (4) can simply be referred to as the end of the syllable, i.e., its coda. There are two problems with this formula: first, it is not true that all preconsonantal consonants exhibit coda-like behaviour, for example, we find glottalization in an English word like *A*[ˀt]*lantic* but aspiration in *a*[tʰ]*ractive,* although the [t] is preconsonantal in both cases. Thus it seems that syntagmatic relationships in the string of segments are not in themselves enough to properly capture phonological environments. Second, even if they were so, the formula in (4) makes use of an unnatural disjunction: there is nothing more common in the word boundary and consonants than in, say, the word boundary and vowels.

As we have seen, the two contexts, _C and _#, can be unified by assigning both types of consonants to a coda constituent. The relevant phonological rules can now be formulated by the structure in (5).

(5) coda
      |
      _

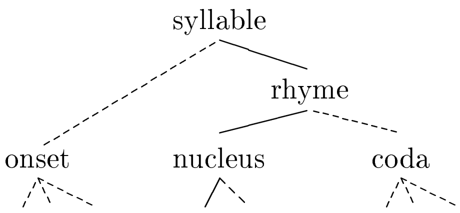
In the case of contrasts like *A*[ˀt]*lantic* vs. *a*[tʰ]*ractive* all there is to do is to assign one of the *t*'s to the coda and the other elsewhere — obviously to the following onset. In many cases such distinctions can be justified by independent evidence, in this one, for example, we can note that one of the clusters in question, *tl,* does not occur word-initially, the other, *tr,* does.

One cannot, however, be satisfied with this much. While a significant degree of descriptive adequacy is reached by the formulation in (5), explanatory adequacy is still wanting. For example, lenition, a phenomenon typically associated with the coda position, manifested as glottalization in the previous example, may be adequately captured by the generalization that coda consonants lenite, there is, nevertheless, no reason why it should be the coda position of all that triggers weakening. One promising initiative to an explanation is made by Itô (1986) and Goldsmith (1990), who claim that codas have a weaker prosodic license than other domains of the syllable, therefore coda consonants are more prone to lenition. There is still ground for insisting on the question why it is codas that have a weaker prosodic license. An answer couched in the Government Phonology framework is provided by Harris (1997), who posits a so-called licensing path in phonological domains ranging from the most prominent nucleus through least prominent ones to the onsets of these nuclei. The claim is that the further away a position is from the prime licensor, the more prone it is to lenition.

4.2 Problems with the standard view

In (6) I give a diagram that shows my interpretation of the syllable tree most widespread in the literature (e.g., Lass 1984 : 252, Durand 1990 : 204, Giegerich 1992 : 138, Carr 1993 : 196, Kenstowicz 1994 : 253, Roca 1994 : 141, Blevins 1995 : 213).

(6)



The solid lines in (6) represent obligatory associations, the dashed lines are optional, i.e., one nuclear segment is obligatory for any syllable, all the others — another nuclear segment and practically any number of onset and coda segments — may or may not be added to complete a syllable.

Given this syllable template syllabifying strings is still not a trivial issue: the length of both onsets and codas is rather flexible. Nuclei can be found applying the SONORITY SEQUENCING PRINCIPLE, one possible wording of which is quoted in (7).

(7) The Sonority Sequencing Principle (SSP)

Within a syllable sonority rises from the onset towards the nucleus and falls from the nucleus towards the coda.

That is, the sonority peaks of a certain string, away from which sonority falls in both directions, can be identified with the syllabic nuclei. Even if nuclei are spotted easily, the consonantal interlude streching between two sonority peaks must be properly distributed among the coda and the onset. To be able to do this in a principled way the ONSET MAXIMIZATION PRINCIPLE[16] is formulated to the effect of (8).

(8) The Onset Maximization Principle (OMP)

If a consonant can be assigned both to a coda and the following onset, assign it to the onset.

Equipped with this principle, consonantal interludes can be unambiguously divided: in a $VC_4C_3C_2C_1V$ string C_1 always goes with the second vowel, then one has to test whether C_2C_1 is a valid onset, if yes it goes with the second vowel, else the syllable boundary is between C_2 and C_1, and so on.

[16] An alternative, negative name of the principle could be the "coda minimalization principle." Both names convey the superiority of onsets over codas. In Optimality Theory the same idea is manifest in the ONSET and NOCODA constraints.

One difficulty comes with deciding whether a given consonant cluster is a valid onset or not. The assumption that the set of word-initial clusters is coextensive with that of valid onsets — and likewise that of word-final clusters with that of valid codas — is often accepted (cf. footnote 12) but rarely if ever supported by any evidence. In fact, what can be supported by empirical evidence is the falsity of this hypothesis, as, for example, the CLOSED SYLLABLE ADJUSTMENT rule of French shows: according to this rule [e] and [ə] surface as [ɛ] in closed syllables, and althought sC clusters do occur word-initially, they also close a syllable (Lowenstamm 1981 : 598f). If sC clusters are heterosyllabic within a word, then it cannot be concluded that the set of well-formed onsets is that of word-initial clusters. On the other hand, in most — perhaps all — languages single consonants that can turn up before a vowel may also turn up word-initially.[17] On the other hand, it is not true that in all languages single consonants that can turn up before a consonant may also turn up word-finally — this is most evident in the case of languages that have word-internal codas, but lack word-final consonants, like Italian. Also word-final consonants can very often not stand before a consonant word-medially — the distribution of English [ð] and the affricates could exemplify this situation.[18] Therefore, we may conclude that the only inference that can be drawn is the following: whatever is an onset may turn up at the beginning of a word. To schematize:

(9) The relationship of consonant(s) at word and syllable margins

NAIVE VIEW
word-initial consonant(s) $\Leftrightarrow$ syllable-initial consonant(s)
word-final consonant(s) $\Leftrightarrow$ syllable-final consonant(s)

EVIDENCED VIEW
word-initial consonant(s) $\overset{\nRightarrow}{\Leftarrow}$ syllable-initial consonant(s)
word-final consonant(s) $\overset{\nRightarrow}{\nLeftarrow}$ syllable-final consonant(s)

Another method that may be of use in determining the end of the coda and the beginning of the onset, i.e., the syllable boundary, is provided by

[17] Counterexamples include [ɾ] and [ŋ] in English, as Péter Siptár (p.c.) points out. To explain them away, the first is a variant of [t] or [d], thus its status is not obvious, the second is always followed by a word boundary when prevocalic.

[18] The only counterexamples are *rhythmic* and *logarithmic* for [ð], and some other syncope created clusters like in *natural* [ˈnæʧrəl] for the affricates.

the SONORITY DISPERSION PRINCIPLE proposed by Clements (1990), quoted in (10).

(10) The Sonority Dispersion Principle (SDP)

a. The preferred initial demisyllable maximizes the dispersion in sonority.
b. The preferred final demisyllable minimizes the dispersion in sonority.

An initial demisyllable is the first half of the syllable up to and including the vowel — with certain language specific differences in the case of long vowels and diphthongs —, a final demisyllable is the second half from and including the vowel; i.e., the onset with the (first half of the) nucleus and the (second half of the) nucleus with the coda, respectively. Sonority dispersion is maximized if the individual members of the demisyllable are evenly distributed on the sonority scale: in an initial demisyllable the first member being the least sonorous (an obstruent[19]), the last the most sonorous (a vowel) and if there is a further member between them then that should be a liquid. In the final demisyllable, sonority dispersion is minimized, that is, the best case is not to have a coda at all, or at least have very sonorous segments in it. The OMP is a derivate of the SDP: it is not only preferable not to have a coda, but also to have an onset and thereby a large — or at least some — sonority distance in the onset–nucleus sequence.

In the case of a string like *atla* both the SDP and the OMP prefers the syllabification *a.tla,* yet in many languages, including English or French, *at.la* is the accepted division, since *tl* is not encountered word-initially and — as already noted — the [t] behaves differently before [l] and [r]. The third logical possibility, *atl.a,* is the worst, it even violates the SSP, introduced in (7). What we end up with are both principles, the SDP and the OMP, only partially satisfied. One way out of this situation is to abandon the apparently self-evident hypothesis that superficial adjacency is evidence of adjacency at all levels. Syntacticians have long noticed this fact,[20] for phonologists it still is not always obvious. Accepting the — let's call it — adjacency hypothesis makes it seem trivial to determine syllable structure simply by looking at

[19] Clements assumes a five-step sonority scale: obstruents < nasals < liquids < glides < vowels. He claims that the algorithm he gives for measuring sonority dispersion also works for more refined scales, but argues that such scales lose cross-linguistic generalizations and become too language specific.

[20] For example, current syntax posits an empty category in the string *the man I want ∅ to go* but not in *I want to go* in order to explain, among other things, the impossibility of *wanna*-contraction in the first.

the string of segments constituting the word. The price to pay is that we have to content ourselves with dispreferred syllable structures and contacts, on the one hand, and the unbelievable complexity and number that syllable types will exhibit, on the other. If we are not willing to pay this price, we have to allow some degree of abstraction — although it is controversial whether this is indeed an abstraction after all, as shown in section **3** —, dispensing with the view that adjacent segments are necessarily adjacent underlyingly. In this way, syllable structure can be radically simplified.

5 Empty nuclei in the skeleton

In this section I am going to introduce a train of thought that allows skeletal positions to remain empty, abandoning the adjacency hypothesis. Government Phonology (GP), especially Kaye et al. (1990) and Kaye (1990), is one, but not the only theory that uses empty positions.[21]

One motivation for Kaye et al. (1990) to assume empty nuclei bears close resemblance to the impasse situation encountered above, the syllabification of *atla.* The claim is that any two consonants that are indeed adjacent are in a governing relationship with each other, i.e., one of them governs the other. The governing potential of specific consonants is determined by their melodic content:[22] some consonants are typically governors, others typically governees. As a result, if a consonant cluster xy is established as a coda x followed by an onset y — in which then y governs x —, the opposite, yx, will definitely not be the same type of cluster, coda–onset in this case, since that would require the previous governing relationship to be swapped, the governor y to now be governed by the governee x. This is deemed impossible, because codas are always governed by the following onset.

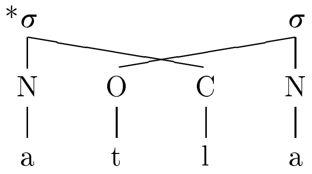
Translated to our case, if *alta* is syllabified *al.ta* — and there is good reason to do that: having a small sonority distance in the nucleus–coda sequence and a great one in the onset–nucleus sequence, it perfectly matches the requirements of the SDP —, *atla* cannot be analysed as a coda–onset

[21] It is interesting to note that hypothesizing empty consonantal positions is more common (e.g., Selkirk & Vergnaud 1973, Clements & Keyser 1983), than empty vocalic positions: the silence of the former is more obvious than that of the latter.

[22] In some versions of the theory governing potential is a function of the charm value of the given segment, but then charm is dependent on melodic content.

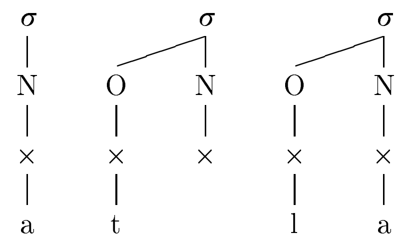
cluster too, i.e., **at.la*.[23] If we are also unable to squeeze both consonants into the onset (**a.tla*) or the coda (**atl.a*), there is no possible syllabification in a model that accepts the adjacency hypothesis. It would be desirable to say that the *t* of *atla* is an onset and the *l* a coda, since — as the SDP suggests — *t* is an ideal onset consonant, and *l* is okay for a coda. This would unfortunately lead to a violation of the constraint banning crossing lines as shown in (11), where σ denotes the syllable node, O, N and C should be obvious.

(11)



Allowing melodically empty skeletal positions into our theory offers a solution to this problem: we are now able to say that the two consonants are not adjacent underlyingly, there is an empty vocalic[24] position ($\emptyset$) between them. Thus we can have both consonants in separate onsets (*a.t∅.la*), in an onset and a coda (*a.t∅l.a,* this is a possible manifestation of the idea in (11)) or in separate codas (*at.∅l.a*), though the second option is a bit strange, the last one rather perverse and neither is favoured by the SDP. The two-onset representation is the most plausible, (12) shows this option syllabified with an empty skeletal position. The skeletal tier is now included since once we have empty positions on it the alphabetic symbols abbreviating melody cannot simultaneously represent skeletal positions anymore.

(12)



[23] Note that Clements's (1990) theory would allow this option, albeit as a highly marked and unpreferable syllable contact. By doing so, Clements is paving the way towards Optimality Theory, where "anything goes," constraints are more or less preferably violable.

[24] Of course, one might hypothesize an empty consonantal position between the two consonants but that would not bring him any closer to a viable analysis: hosting the extra C position is yet another pain in the neck.

It is an interesting question to ask how the SDP would react to the syllabification *a.t∅.la.* The sonority of an unpronounced segment is undefined, therefore the sonority rise in the syllable *t∅* is indeterminable. Nonetheless, the absence of codas is of merit in the eyes of the SDP; onset maximization is fully performed.

There seems to be a difficulty with this solution. As we have seen in section **2**, melodically empty skeletal positions do have a phonetic interpretation: the most unmarked vocalic segment ([ə, ɯ, ɨ] or something similar) if dominated by a nuclear position, or the most unmarked consonantal segment (the identity of which is debatable and indeed debated in the literature) if dominated by a nonnuclear position, i.e., the onset or the coda. This means that the phonetic interpretation of the representation in (12) should be [atəla] or [atɯla], a pronunciation that would cause no debate in phonologist circles as regards its syllabification. If we are to maintain the results of section **2** and posit *unpronounced* empty positions simultaneously we have to claim that some melodically empty skeletal positions are pronounced, others are not. The theory must provide some means to predict the pronunciation or nonpronunciation of a skeletal position in each case. GP's solution is the formulation of the phonological EMPTY CATEGORY PRINCIPLE, which is given in (13). This formulation is in its essence identical to that of Kaye (1995 : 295).

(13) The Empty Category Principle (ECP)

A melodically empty skeletal position remains unpronounced if

i. properly governed,
ii. domain-final (parametric) or
iii. enclosed within an onset-to-onset governing domain.

Let us first examine the first clause (13i) of the ECP. To do this we need a definition of PROPER GOVERNMENT. This is given in (14), again cited almost verbatim[25] from the same locus.

[25] I have modified clause b of Kaye's definition. Originally it run "α is not itself licensed." What causes the problem here is the two meanings of the term licensing (cf. Brockhaus 1999 : 208, fn. 22): a position may be licensed to be part of the representation, but also to remain unpronounced. Kaye here has the latter meaning in mind.

(14) Proper Government (PG)

A nuclear position α properly governs a nuclear position β iff
a. α is adjacent to β on its projection,
b. α is not unpronounced,
c. no governing domain separates α from β.

Clause (14a) simply means that the governing and the governed vowels cannot be separated by a third nuclear position, clause (14c) means that no consonant cluster may intervene — recall, there is a governing domain created by any truly adjacent consonants in GP, while clusters of nonadjacent consonants contain an empty nuclear position, hence fall under clause (14a). The middle, empty nucleus of *at∅la* satisfies all the requirements for proper government, and therefore may remain uninterpreted.[26] We may conclude that the phonetic interpretation of a melodically empty skeletal position depends on the syntagmatic relations it has with other positions in its string.

Another location where GP posits empty nuclear positions is after word-final consonants (cf. Kaye 1990). The suggestion is based on the observation that word-final consonants often exhibit different behaviour from word-medial codas: they usually do not count for syllable heaviness, they often host segments that word-medial codas cannot (e.g., affricates in English), they often do not cause closed syllable shortening, etc. These phenomena — captured by the notion of extraprosodicity in another part of the mainstream literature — are neatly, though as we are going to see not unproblematically, explained by denying the codahood of word-final consonants. While it is true that these consonants also show properties of codas (e.g., they are more ready to undergo lenition than onsets proper) thus it is not uncontroversial that they should be onsets, the extraprosodicity account, which distinguishes them from both codas and onsets, misses the generalization that word-final consonant clusters in, for example, English are with a few exceptions the same as intervocalic coda–onset clusters. If the second of a two-member word-final consonant cluster is made extraprosodic, the grammar duplicates the phonotactic statements on VC.CV clusters by having to make them again on VC⟨C⟩#[27] clusters (Harris 1994 : 74f).

[26] The careful reader may now ask whether it is the word-initial or the word-final *a* of *at∅la* that properly governs the empty nucleus. Being a manifestation of interconstituent government, it is standardly assumed to propagate from right to left, therefore it is the last *a* that does the job. One notable exception is Rowicka (1998), who argues that PG is left-to-right.

[27] As is conventional, I use angled brackets to denote extraprosodic segments.

The possibility of word-final empty nuclei is language specific: it is a parameter of UNIVERSAL GRAMMAR that is not available in the default case (not having word-final consonants is the unmarked case), but set on in some languages, which thus may have a consonant at the end of the word (cf. (13ii)).[28] This property, Kaye (1990 : 323f) claims, is distinct from having codas, the two parameters produce a cross-classification: languages may have codas only word-medially, like Italian, or only word-finally, like Luo, in addition to having them in both locations or neither. The existence of four different groups of languages with regard to non-prevocalic consonants is further evidence for denying the coda status of word-final consonants.

To conclude the discussion of empty positions in GP, we may say that by positing empty nuclear positions in the skeleton the theory reduces the cases where consonants are syllabified into the coda position. This tendency is in line with the generally accepted view that onsets are to be preferred over codas in syllabification. One salient feature of GP is its affinity to turn generalizations that other theories look at as universal preference statements into unviolable constraints. This property distinguishes the approach quite radically from Optimality Theory, where any constraint is violable. In the case discussed above, the fact that an obstruent–liquid cluster is a dispreferred coda–onset cluster is tightened to the claim that it *is never* a coda–onset cluster. If one dares take this thought to its conclusion, the next question to ask is if codas exist at all, after all the optimal final demisyllable is one without a coda. We are going to proceed in this direction.

6 Does the coda exist?

What we have to examine is the arguments supporting the existence of the coda position. As it was already noted there is a sharp asymmetry between the two margins of the syllable, the onset and the coda. The most unmarked syllable type, available in all languages, is CV, i.e., one that contains an onset but no coda. Furthermore, while in the unmarked case the onset is

[28] The plausibility of the ECP is reduced by the fact that it includes clauses that are of so different types. I do not discuss this further here, but see Szigetvári (in prep.).

obligatory — contrary to what Blevins (1995 : 218–220) claims[29] —, it is the marked case to have a coda.

One of the reasons why codas are posited in the first place is the assumption that syllable boundaries and word boundaries coincide. If consonants are found at the right margin of words then they obviously occupy the right margin of a syllable. But, as we have seen, there is also phonological evidence which indicates that word-final consonants are not uncontroversially codas.

Codas also have explanational value in the formalization of stress rules. In languages with unfixed stress, rules are often sensitive to syllable weight. The standard case is that syllables with only a short vowel count as light (therefore usually unstressable), while syllables more fleshy than that — either closed by a consonant or containing a long vowel — are heavy (and attract stress). Positing a constituent, the rhyme, dominating the nucleus and the coda facilitates the definition of heaviness: syllables with rhymes containing one segment are light, those with multisegmental rhymes are heavy. Unfortunately, neither the branching of the rhyme, nor that of the nucleus may be held to be responsible for heaviness, all we can say is that one of the two must branch. Another problematic aspect of this approach to syllable weight is the fact that onsets (apart from very few and therefore suspect cases) do not contribute to it. One either stipulates that only the size of the rhyme is relevant or offers some theory that assigns weight, standardly referred to as mora, to the appropriate segments. However, even the latter option does no more than formalizing the observation that coda consonants

[29] Blevins says "the unmarked case is that onsets are not obligatory." Interestingly, of the four arguments she brings up to support the claims about unmarkedness in syllable constituency two explicitly argue *for* obligatory onsets being the default case: "(3) All languages have CV syllables" — while, apparenly, only *some* have V syllables; to me this argues for onsetless syllables being more marked — and "(4) ... there are a variety of phonological processes which take marked syllable types to unmarked types..., but there are few if any rules which consistently result in [marked syllable types]." The avoidance of hiatus is a widespread phenomenon, which aims at getting rid of onsetless syllables, while losing onsets is typical only of intervocalic/posttonic, not of other types of onset. Blevins also says: "In second language acquisition, speakers have little difficulty in shifting from a 'yes' value to a 'no' value fro a given parameter, but do show difficulty in switching from a 'no' value to a 'yes' value" (*ibid.*). Yet itt is hard to see any difficulty in getting from a language having both CV- and V-type syllables to one which only has the former, which allegedly is the marked type. Furthermore, the "yes" and "no" values crucially depend on the formulation of the parameter: "obligatory onset" for Blevins, but it might as well be "nonobligatory onset".

do, while onset consonants do not influence the weight of a syllable, without explaining why this and not the opposite should be the case. The alternative below (in section **7.1**) fares better in both respects: it explains why both closed and long-vowelled syllables are heavy and why onsets do not count.

The minimal word phenomenon, that constrains the size of lexical words in a number of languages as diverse as English, Hungarian, Beijing Mandarin, Khalkha Mongolian and Turkish (for the last three cf. Denwood 1998), also depends on a plausible formulation of heavy syllables. The observation is that in these languages a lexical word cannot be a single light syllable, it must either be a heavy syllable or two light syllables. In monosyllables the necessary weight is provided either by the length of the vowel or a final, allegedly coda, consonant.

Another reason for assuming codas is the widely observed phenomenon of closed syllable shortening. If the size of the rhyme is limited to two segments, the largest syllable types available are an open syllable with a long vowel (-VV) or a closed syllable (-VC) whose vowel must be short since three segments are too many in a rhyme.[30] Consequently, if an open syllable containing a long vowel is made closed by some phonological process its vowel must shorten so that the domain does not exceed its limits. If the one involving codas were the only plausible explanation for closed syllable shortening, codas would be safe. We are going to see below (in section **7.2**) that this is not so.

One of the standard arguments for constituenthood in the subsyllabic domain is the existence of phonotactic constraints. For instance, the almost universal absence of homorganic consonants in branching onsets may be seen as evidence that such consonants do form a constituent. Similarly, in nuclei the types of attested vowel clusters, i.e., diphthongs and long vowels, are restricted to a small subset of all the possibilities. As opposed to this, very few qualitative[31] phonotactic constraints apply to VC clusters, that

[30] This is stated as the Binary Theorem by Kaye (1990 : 306) and Kaye et al. (1990 : 199) and deduced from different principles in different ways by Kaye (et al.) on the one hand and Charette (1989 : 161ff) and Lowenstamm & Kaye 1985–1986 on the other. Harris (1994 : 68ff) in his analysis of English is forced to abandon the restriction.

[31] The quantitative constraint, referred to as closed syllable shortening, was treated above.

is, within the rhyme.[32] Where we do encounter phonotactic constraints between consonants is in intervocalic and word-final clusters. Intervocalic clusters of the type [nt], [mp] are rather unanimously analysed as heterosyllabic, coda–onset clusters. Yet, it is not usual to consider these clusters as members of the same syllabic constituent. Therefore, we may conclude that the existence of some phonotactic constraint between two segments does not necessarily imply that they share their host constituent.

Recall that different syllabifications were suggested for *al.ta* and *a.t∅.la,* as shown here. If we accept that some intervocalic consonant clusters are coda–onset clusters, while others are onset–onset clusters containing an empty nucleus between them, our theory becomes indeterminate. Nothing excludes the syllabification *a.l∅.ta*: there will be no way of knowing whether a cluster that satisfies the criteria for coda–onset clusters is to be analysed as such or as an onset–onset cluster that accidentally happens to contain consonants which would also make a coda–onset cluster.[33]

To summarize: the theoretical status of the coda is strongly challenged. It is an outcast in markedness universals: onsets may even be obligatory but are never impossible in languages, codas are never obligatory and may even be impossible. Though positing a coda position seems to help in distinguishing heavy and light syllables, there are serious problems with the formulation. While closed syllable shortening can be explained by reference to the impossibility of a coda following a branching nucleus, there are alternative explanations to be discussed below. Finally, the possibility of analysing some clusters both as coda–onset and as onset–onset clusters loosens the theoretical tightness of the framework.

[32] The most often cited such constraint bans noncoronal consonants after [aʊ] in English (Anderson 1986). Hungarian does not allow nonround vowels ([i ɛ]) before word-final labial nasal-stop clusters (Törkenczy 1994 : 338), and only [ø] is possible before a word-final lateral liquid–palatal stop cluster (viz., [lɟ]). Nobody has very much to say about such constraints, they do not seem to be systematic and may even be attributed to lexical/historical accidents.

[33] Of course, phenomena like closed syllable shortening or heaviness for stress assignment may tilt the balance in this or that direction, but only in case C.C and C∅.C are treated differently in the analysis of these phenomena.

7 Without codas

Making a constraint out of the preference of the Sonority Dispersion Principle, one may claim that all syllables have an onset and none have a coda (cf. Lowenstamm 1996).[34] Setting aside for the time being the possibility of having more than one consonant in a single onset constituent, this means that whenever we find a consonant that is not followed by a vowel it must be followed by an empty nucleus — to make it, at least theoretically, an onset.

It is important to bear in mind that the question whether something is in coda position or not is not an empirical one; this property does not in itself have any physical correlate. The rationale of positing a coda position is to unify the contexts that pattern together in certain phonological phenomena. If these contexts may be unified by other means there is no strong argument for keeping codas in the theoretical vocabulary, unless one needs it for descriptive purposes, as a dated but useful term, similarly to the way a syntactician would refer to S(entence)s even after showing that they are I(nflection)P(hrase)s or C(omplementizer)P(hrase)s. This is the sense the word coda will be used hereafter. Actually, if codas do not have a theoretical status then it does not make much sense to talk about onsets either, even if — what is kept in benign ignorance — they are imagined to be potentially branching. What are left of syllabic constituents is a consonantal and a vocalic constituent.

Having stripped syllabic constituency so brutally, one might as well take the last move and claim that neither the consonantal, nor the vocalic constituent ever branches, that is, the skeleton contains a strict alternation of consonantal and vocalic *positions*; this is exactly what Lowenstamm (1996) does. Arguments for this final step do not readily offer themselves, some motivations will, nevertheless, be pointed out in the next section. Even without explicitly arguing against branching nuclei and branching onsets, formal simplicity is a criterion that opts for nonbranching constituents. Recall (from section **3**), if current phonological theory were initiated by Zulu phonologists a CVCV skeleton would surely have been the starting point.

[34] Note that GP theorists regularly argue that the coda *consituent* is nonexistent in their theory. There still are codas in GP, since rhymes may branch, what the right branch dominates is the coda as opposed to the other two consonantal positions that are in the onset (which may also be branching), i.e., the term coda is a shorthand for the "postnuclear rhymal complement". My aim above, however, was to show that as regards their skeletal status all consonantal positions are equal, the only difference is whether a consonantal position is followed by an interpreted vocalic position or not.

In this section we are going to see the way the CVCV framework handles some coda-related phenomena discussed in section **6**.

7.1 Heavy versus light syllables

In a theory comprising only CV pairs to represent syllable structure, a light syllable will be made up of one such pair, while a heavy syllable will contain two of them as shown in (15), where the Greek letters stand for any, potentially identical, melodic material (if identical, the two symbols are merged in (15b)):

(15) a. LIGHT SYLLABLE b. HEAVY SYLLABLE TYPE I c. HEAVY SYLLABLE TYPE II

$$\begin{array}{cc} C & V \\ | & | \\ \alpha & \beta \end{array} \qquad\qquad \begin{array}{cccc} C & V & C & V \\ | & | & & | \\ \alpha & \beta & & \gamma \end{array} \qquad\qquad \begin{array}{cccc} C & V & C & V \\ | & | & | & \\ \alpha & \beta & \gamma & \end{array}$$

The advantages of the representations in (15) are the following: (i) the formulation of what constitutes a heavy syllable is much less clumsy than if we were using the coda, all that has to be distinguished is one vs. two CV pairs, as opposed to statements like "*either* the nucleus *or* the rhyme is branching." (ii) We get an explanation of why onsets do not contribute to syllable weight: paradoxically rhymes do not contribute either, the question itself loses its significance. All we need for a heavy syllable is two pronounced CV pairs, that is two CV pairs both containing some melodic material.[35] The onset of such a syllable is the C of the first pair but whether it is filled or not is immaterial, since its V will be filled, that is why it is taken to be a syllable in the traditional approach. In a sense then a CV slice of the skeleton is the equivalent of the mora in frameworks that measure syllable weight by that means, but unlike moraic frameworks we get a nonstipulative account for the lack of onset weight. The CVCV approach, however, still owes an explanation for why word-final consonants often fail to contribute to syllable weight.

Note also that in languages like Latin or English, where stress rules typically take the form "if the (pen)ult is heavy stress it, if it is light stress

[35] In a subset of the languages distinguishing heavy and light syllables only (C)VV, but not (C)VC counts as heavy. In such languages it is apparently the pronunciation of the V part of the CV unit that is taken into account.

the (ante)penult," there is room for a simplified formulation: e.g., stress the third last CV pair,[36] boxed in the Latin words illustrating the rule in (16):[37]

(16) a. *domínica* 'lord adj.fem.'

C V $\boxed{\text{C V}}$ C V C V
d o m i n i k a

b. *arḗna* 'sand'

C V $\boxed{\text{C V}}$ C V C V
a r e n a

c. *agénda* 'things to do'

C V $\boxed{\text{C V}}$ C V C V
a g e n d a

It is rather complicated to capture the minimal word constraint, which limits the size of content words to two moras at least, in the traditional GP framework. Since word-final consonants are claimed to be onsets followed by an empty nuclear position, one has to say that either the nucleus of the only syllable of the minimal word must branch or the word must contain two onset–rhyme sequences.[38] The CVCV formulation is trivial: the minimal word contains two CV pairs[39] (perhaps in order to be stressable).

7.2 Closed syllable shortening

To show what the CVCV approach can do with the phenomenon of closed syllable shortening, I repeat and comment on what Lowenstamm (1996 : 12–13) has to say on the issue.

Of the two hypothetical forms, *[kaːtpi] and [kaːtupi], the first is ungrammatical because it contains a long vowel in a closed syllable. Using only CV pairs in the representation, a closed syllable takes the form of at

[36] If, as I argued above, natural language is unable to count, "third last" can only be a descriptive formulation of the rule.

[37] The situation is not as neat as depicted here. Difficulties arise in the following cases: the third last CV pair may contain an empty V position, stress in this case appearing on the fourth (*fórmula* 'rule'), word-final long vowels count as if short (*fáciō* 'make') and word-final consonants do not count (*ácidus* 'sour').

[38] An alternative, slightly less disjunctive but no more plausible formulation is the following: a minimal word must contain two slots dominated by a nuclear node.

[39] Again, we are forced to do counting. The formulation is descriptively adequate and better than standard GP's, but theoretically problematic.

least two CV pairs with the latter containing an empty V position. The representation of a long vowel is also two CV pairs of which the second is totally devoid of melodic material, it is the vocalic melody of the first pair that is interpreted in the V position of the second pair, as shown in (17b):

(17) a. katpi, *kaːtpi

C	V	C	$V_{\$}$	C	$V_{£}$	C	V
k	a			t		p	i

b. kaːtupi

C	V	C	$V_{\$}$	C	$V_{£}$	C	V
k	a			t	u	p	i

For the vocalic melody to be interpreted in the $V_{\$}$ position that position must be licensed. It is licensed if it is properly governed (cf. (14)). However, an empty position which is itself licensed is unable to properly govern the preceding position. While in (17b) the relevant position, $V_{£}$ is unlicensed and pronounced, hence can properly govern and thereby license the $V_{\$}$ position, in which the vocalic melody of the first V can be interpreted, this does not hold in (17a).

The rhyme-maximizing theory of closed syllable shortening faces a serious challenge: if seen as a dynamic process it violates the principle of structure preservation. A coda position not present lexically may not be created by some phonological event. If, however, we cannot create a coda position in a rhyme containing a branching nucleus (i.e., after the long vowel), closed syllable shortening cannot be motivated. This problem is overcome by the CVCV approach.

The analysis of closed syllable shortening in (17), however, is not fully satisfactory. On the one hand, it does not incorporate the observation that close syllable shortening does not happen before any consonant cluster. I am not thinking of branching onsets here, for which Lowenstamm has an account, but of clusters created by syncope, which in the standard GP analysis were onset–onset clusters — as opposed to coda–onset clusters. According to the present analysis these should also shorten: onset–onset and coda–onset clusters are not distinguished here. No shortening is attested in English before syncope sites, cf. e.g., *favourite* [ˈfeɪv~~ə~~rɪt]. The account is thus descriptively inadequate. The theory-internal problem with explaining closed syllable shortening as presented here is that proper government results in an unwarranted dual behaviour: it silences some vocalic positions and makes the pronunciation of others possible.

7.3 Compensatory lengthening

Compensatory lengthening is another phenomenon that appears to call for coda positions in representations. After the total lenition of a consonant in a weak prosodic position the loss is made up for by the propagation of either the preceding vocalic or the following consonantal material, for example, the reconstructed Greek form *[esmi] is realized in Classical Attic as [eːmi] 'I am', while Aeolic has [emːi]. The latter event, where the place of a consonant is taken up by another consonant, is rather easy to handle for both theories. Vowel lengthening on the other hand happens again in violation of structure preservation in the coda approach: what used to be a consonantal position, coda, is lost and a vocalic, nuclear position appears instead. The model offered by the CVCV approach does not face such problems: the vacation of the C position by the loss of [s] either opens the way for the following C position to occupy it (18c) or removes the obstacle that has prevented the preceding V from taking it (18b). Which of the two strategies is applied can be predicted on a language—here dialect—specific basis: it looks very much like a parameter.[40]

(18) a. Reconstr. *esmi

C V C V C V
e s m i

b. Attic eːmi

C V C V C V
e m i

c. Aeolic emːi

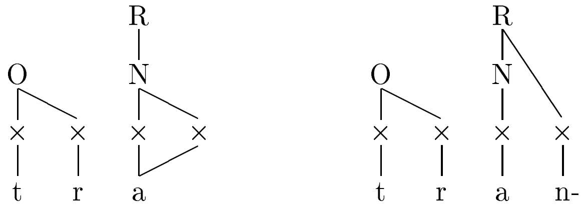
C V C V C V
e m i

[40] Lowenstamm & Kaye (1985–1986 : 114) argue that gemination, (18c), is the default case, while vowel lengthening, (18b), only occurs if the system lacks the relevant geminate consonant. This may be related to the fact that coda–onset clusters appear to be less marked than long vowels.

8 Against constituency

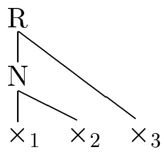
All three syllabic constituents, the onset, the rhyme and the nucleus, are imagined to be potentially branching by GP theorists and most mainstream researchers alike (for the latter, even the coda is potentially branching). In the former, more restrictive, framework a maximal syllable has one of the structures depicted in (19).

(19)



In a GP-like framework the nonexistence of codas amounts to the claim that the rhyme constituent does not branch, and if it does not branch it is not a syllabic constituent — it shares the fate the coda has suffered earlier. It is in fact a felicitous development of the theory to have got rid of the rhyme constituent, which is a nuisance in more than one respects. For one thing, the rhyme is the only syllabic constituent that does not dominate exclusively skeletal slots but also another syllabic constituent, the nucleus. This fact has led to uncertainty about whether and why a branching nucleus may occur in a branching rhyme. In one view (that of, e.g., Kaye et al. 1990) it cannot, because in such a constituent — shown in (20) for those with a visual disposition — no head can be assigned; this is the BINARITY THEOREM already mentioned in section **6**. The two constraints that head and dependent must be adjacent and that their relationship is unidirectional destroys the hopes of all three possible candidates: the first is not adjacent to the third, the second would have one dependend on the left, one on the right, the third is not adjacent to the first.

(20)



However, when forced to accept the structure in (20), as Harris (1994: 68f, 76f, 82f) is in order to cater for words like *dainty, easter, b*[ɑː]*sket,* *saint, post, wild*[41] etc., one may seek refuge in the idea that the head of the rhyme is not on the skeleton, but it is the nuclear node itself. It is not unreasonable to look for the head of a constituent among its daughters, after all. If the rhyme should no more exist, the dilemma also perishes.[42]

If syllable heaviness is not (merely) a function of the number of skeletal positions in the rhyme, representing long vowels and diphthongs by branching nuclei becomes much less obvious. The wish to keep syllables together as onset–rhyme sequences is also in vain if codas are let loose. The "phonetic unity" of long vowels — whatever that should mean — is not a strong argument: a long vowel is just as much a unit as a long consonant, the latter is, nevertheless, a coda–onset cluster, thus not one constituent, in most frameworks. (Not to mention the fact that without codas long consonants hopelessly become C∅C clusters.)

The claim that "all feet are minimally binary and that the word in many languages must consist minimally of a foot" (made by McCarthy & Prince (1986) and quoted by Harris (1997)) suggests that just as [tata] and [tat] (the latter obviously *tat∅*) are binary feet — hence qualify for minimal words in the languages concerned —, [taː] must also somehow make a binary foot. The number of vocalic positions involved in the string is undisputedly two, but the immediate constituens of foot nodes are usually either syllable nodes or, in their absence, nuclei. Only by analysing the [aː] as two nuclei, i.e., N∅N, do we obtain a binary foot, thus satisfying the minimal limit on word size. Note that the same argument was already brought up in section **7.1** cast in a slightly different form.

Kaye (1985 : 290f) and Lowenstamm & Kaye (1985–1986 : 99f) claim that there is an implicational relationship between branching rhymes and branching onsets. The observation, called the RHYME-DOMINANT PRINCIPLE, is that languages having branching onsets invariably have branching rhymes (i.e., closed syllables), while the opposite is not true, languages with branching rhymes may or may not have branching onsets. To put it in other

[41] Though Harris does allow type (20) superheavy rhymes (1994 : 69, 83), he also has to strictly limit their occurrence to ones with coronal and very few other consonant clusters.

[42] The problem of superheavy rhymes unfortunately does not disappear with this move.

words, branching onsets are more marked than closed syllables.[43] Whether this argues for the abandonment of the hypothesis that onsets, or rather, the consonantal constituent, may branch is not fully obvious. The question basically boils down to the markedness of branching constituents and that of empty skeletal positions.[44] Theoretical uniformity requires either the retention of constituency throughout the whole range of syllabic constituents or their total abandonment, which means positing a C∅C structure to branching onsets as well.

One last consideration that is relevant for the total rejection of syllabic constituency is that if the skeleton contains strictly alternating C and V positions—no adjacent Cs and no adjacent Vs—then it is trivial to parse a phonological string, provided the listener can distinguish consonants and vowels: whenever he encounters two instances of the same category an empty position of the opposite type must be inserted between them, while two different categories will be adjacent.[45] This advantage is not available in a system where at some points one may assume two adjacent Cs or Vs, at another they will be separated by an empty category. Consequently, allowing empty skeletal positions into phonological representations concludes to the hypothesis that the phonological skeleton must be made up of strictly alternating Cs and Vs.

9 Conclusion and invitation

I have tried to convince the reader of the advantages of accepting empty skeletal positions and hypothesizing phonological skeletons that are made up of strictly alternating consonantal and vocalic positions. Both ideas can be shown to follow from rather general assumptions shared by larger or smaller subsets of the phonological society. On the other hand, I could not help mentioning the difficulties a theory using strict CV skeletons faces. Some of these problems appear to be solvable. The interested reader is invited to consult

[43] Lowenstamm & Kaye (1985–1986 : 111) also claim that long vowels are more marked than closed syllables, that is, there exist no languages with long vowels and/or heavy diphthongs and only open syllables. If one accepts the proposal suggested here, this is a further argument for the V∅V representation of long vowels.

[44] There is a third possibility, branching onsets could be considered to be contour segments (cf. Rennison 1998). This idea includes large scale reshuffling of segmental representations, space limitations inhibit further discussion here.

[45] This is only true if two adjacent empty positions are not allowed, two instances of the opposing categories may or may not be adjacent (C∅∅V or V∅∅C).

Dienes & Szigetvári 1999 or the present author's PhD thesis — the first two chapters of which have been converted into the present paper —, both downloadable from HTTP://BUDLING.NYTUD.HU/~SZIGETVA/PAPERS.HTML.

REFERENCES

Anderson, John M. 1986. Suprasegmental dependencies. In: Jacques Durand (ed.) Dependency and Non-Linear Phonology. London: Croom Helm. 55–133.

Anderson, John M. and Colin J. Ewen. 1987. Principles of Dependency Phonology. Cambridge: Cambridge University Press.

Archangeli, Diana. 1988. Aspects of Underspecification Theory. Phonology 5 : 183–205.

Blevins, Juliette. 1995. The syllable in phonological theory. In: Goldsmith 1995 : 206–244.

Brockhaus, Wiebke. 1999. The syllable in German: exloring an alternative. In: Harry van der Hulst and Nancy A. Ritter (eds.) The Syllable: Views and Facts. Berlin: Mouton–de Gruyter. 169–218.

Carr, Philip. 1993. Phonology. Houndmills & London: Macmillan.

Charette, Monik. 1989. The minimality condition in phonology. Journal of Linguistics 25 : 159–187.

Chomsky, Noam and Morris Halle. 1968. The Sound Pattern of English. New York: Harper & Row.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In: John Kingston and Mary E. Beckman (eds.) Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech. Cambridge: Cambridge University Press. 283–333.

Cyran, Eugeniusz (ed.) 1998. Structure and Interpretation: Studies in Phonology. PASE Studies & Monographs 4. Lublin: Wydawnictwo Folium.

Denwood, Ann. 1998. A template for Turkish. SOAS Working Papers in Linguistics & Phonetics 8 : 177–190.

Dienes, Péter and Péter Szigetvári. 1999. Repartitioning the skeleton: VC Phonology. Ms., Eötvös Loránd University.

Durand, Jacques. 1990. Generative and Non-Linear Phonology. Harlow: Longman.

Durand, Jacques and Francis Katamba (eds.). 1995. Frontiers of Phonology: Atoms, Structures, Derivations. Harlow: Longman.

Durand, Jacques and Bernard Laks (eds.). 1996. Current Trends in Phonology: Models and Methods. European Studies Research Institute, University of Salford Publications.

Giegerich, Heinz J. 1992. English Phonology: An Introduction. Cambridge: Cambridge University Press.

Goldsmith, John A. 1990. Autosegmental and Metrical Phonology. Oxford & Cambridge, Mass.: Blackwell.

Goldsmith, John A. (ed.) 1995. The Handbook of Phonological Theory. Cambridge, Mass. & Oxford: Blackwell.

Harris, John. 1994. English Sound Structure. Oxford & Cambridge, Mass.: Blackwell.

Harris, John. 1996. Phonological output is redundancy-free and fully interpretable. Durand & Laks 1996 : 305–332. Also appeared in UCL Working Papers in Linguistics 8.

Harris, John. 1997. Licensing Inheritance: an integrated theory of neutralisation. Phonology 14 : 315–370.

Harris, John and Edmund Gussmann. 1998. Final codas: why the west was wrong. In: Cyran 1998 : 139–162.

Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In: Durand & Katamba 1995 : 34–79.

Itô, Junko. 1986. Syllabic Theory in Prosodic Phonology. PhD dissertation, University of Massachusetts. Published 1988, New York: Garland.

Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. PhD dissertation, Massachusetts Institute of Technology.

Kaye, Jonathan. 1985. On the syllable structure of certain West African languages. In: Didier Goyvaerts (ed). African Linguistics: Essays in Memory of M. W. K. Semikenke. Amsterdam: John Benjamins. 285–308.

Kaye, Jonathan. 1990. 'Coda' Licensing. Phonology 7 : 301–330.

Kaye, Jonathan. 1995. Derivations and interfaces. In: Durand & Katamba 1995 : 289–332.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1985. The internal structure of phonological elements: a theory of charm and government. Phonology Yearbook 2 : 305–328.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. Phonology 7 : 193–232.

Kenstowicz, Michael. 1994. Phonology in Generative Grammar. Cambridge, Mass. & Oxford: Blackwell.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. Phonology Yearbook 2 : 85–138.

Lass, Roger. 1984. Phonology: An Introduction to Basic Concepts. Cambridge: Cambridge University Press.

Lowenstamm, Jean. 1981. On the maximal cluster approach to syllable structure. Linguistic Inquiry 12 : 575–604.

Lowenstamm, Jean. 1996. CV as the only syllable type. Ms., Université Paris 7. Appeared in Durand & Laks 1996 : 419–442.

Lowenstamm, Jean and Jonathan Kaye. 1985–1986.[46] Compensatory lengthening in Tiberian Hebrew. In: Leo Wetzels and Engin Sezer (eds.) Studies in Compensatory Lengthening. Dordrecht: Foris. 97–132.

McCarthy, John J. and Alan Prince. 1986. Prosodic morphology. Ms., University of Massachusetts, Amherst and Brandeis University.

Rennison, John R. 1998. Branching onsets are contour segments. Paper presented at the 31st Poznań Linguistic Meeting, Government Phonology Workshop, 2 May 1998.

Roca, Iggy. 1994. Generative Phonology. London & New York: Routledge.

Rowicka, Grażyna J. 1998. The audibility and visibility of Mohawk ghosts. In: Cyran 1998 : 247–260.

Schafer, Robin. 1995. Headedness in the representation of affricates. The Linguistic Review 12 : 61–87.

Selkirk, Elisabeth O. and Jean-Roger Vergnaud. 1973. How abstract is French phonology? Foundations of Language 10 : 249–254.

Steriade, Donca. 1995. Underspecification and markedness. In: Goldsmith 1995 : 114–174.

Szigetvári, Péter. 1997. On affricates. In: Ágnes B. Farkas (ed.). Proceedings of the First Symposium of Doctoral Students in Linguistics. Budapest: Theoretical Linguistics Programme, Eötvös Loránd University (ELTE). 94–105.

Szigetvári, Péter. 1998. Voice assimilation in Hungarian: the hitches. The Even Yearbook 3 : 223–237.

Szigetvári, Péter. in prep. VC Phonology: a theory of lenition and phonotactics. PhD dissertation, Eötvös Loránd University.

Törkenczy, Miklós. 1994. A szótag. In: Kiefer Ferenc (ed.) Strukturális magyar nyelvtan 2: Fonológia. Budapest: Akadémiai Kiadó. 273–392.

[46] The book containing this paper has 1986 on its title page and 1985 on its copyright page. Some refer it by one, others by the other date.