

Syllable Integrity
Daniel L. Everett
University of Pittsburgh
dever@verb.linguist.pitt.edu

Abstract

This paper argues that syllables may be divided by feet, based on an analysis of Banawá, an Arawan language of the Brazilian Amazon. This result is important for phonological theory because it improves our understanding of the variable roles of syllables and moras in the phonological systems of individual languages and because it underscores the urgent need for more fieldwork on the prosody of the world's languages. Available data on the prosodic systems of the world's nearly 7000 languages lags so far behind the data available on other aspects of grammar that it seems premature to base predictions concerning prosodies on poorly informed notions of rarity or markedness in prosodic structures.

Syllable Integrity

1. Introduction¹

The purpose of this paper is to argue that syllables may be divided between separate prosodic feet, based on data from Banawá, an endangered language of the Brazilian Amazon. To state the problem in more technical jargon, the claim is that *Syllable Integrity* is not a universally inviolable constraint. Syllable Integrity refers to the widely accepted notion that syllables may not be divided into separate feet, i.e. that foot boundaries may not fall between syllable boundaries.² For example, Hayes (1995, 49) argues that syllable integrity is an inviolable constraint of Universal Grammar (UG) because feet are always built directly on syllables, never directly on subsyllabic units, e.g. moras (cf. also Blevins 1995). Therefore, there is no way for feet to be so constructed that they violate syllable integrity.³ In Optimality Theory (OT; Prince & Smolensky 1993), it is assumed that feet are so defined within the GEN(erate) function that their boundaries never fall within syllables. On the other hand, there are linguists who reject syllable integrity, most notably Halle & Vergnaud (1987, pp18ff). Halle & Vergnaud (HV) argue that feet may be built either on syllables or on moras. When they are built on moras then they may violate syllable integrity.

These positions seem to be at once too strong and too weak, although the HV position is closest to being correct. The prohibitions against violations of syllable integrity are too strong because, as I argue in this paper, syllable integrity can be violated (but does not usually seem to be) crosslinguistically. The HV position may be too weak because it offers no explanation as to why syllable integrity seems to be so rare in the languages studied to date.⁴ The research results reported here are important for phonological theory for a couple of reasons. First, they show that syllables and moras may play separate roles in a language. Prosodic structure thus appears less as a trivial result of the construction of larger units out of smaller units and more as a rich domain of interconnected but nonautomatic relationships between independently required elements. Second, they underscore the urgent need for more fieldwork on the prosody of the world's languages. Available data on the prosodic systems of the world's nearly 7000 languages lags so far behind the data available on other aspects of grammar that it seems premature to base predictions on what can or cannot be done in prosodic systems on ill-informed notions of rarity or markedness in prosodic structures. This is not to say that it is wrong to make predictions about the relationships we expect to find in future research, merely that frequency of occurrence and rarity ought not to count heavily in the formulation of those predictions.

This paper is relatively atheoretic. Although I use terms of OT, the primary contribution is to phonological knowledge, rather than theory per se, although the implication for theory is significant - because syllables may be divided by feet in some languages but not others, then Syllable Integrity must be interpreted as *violable* constraint of UG (or, alternatively, as a *parameter*). The discussion is organized as follows: First, I argue for the existence of CV and CVV syllables and against V syllables in Banawá, including new data recently collected in the field on hypocoristic forms. Next, I review the basic properties of Banawá stress, as originally analyzed in Buller, Buller, & Everett (1993, BBE). This section shows that foot boundaries fall within syllables, thus violating syllable integrity.

2. Syllable structure in Banawá

My task in this section is twofold. First, I must argue against V syllables in Banawá. This is because if it had V syllables, then we could say that feet are built on Vs, rather than moras and that, therefore, no violations of syllable integrity occur. Second, I must argue

for the existence of CVV syllables, since it is in this syllable type that we observe syllable integrity violations (incidentally, arguments for CV syllables are also given).

2.1. Phonotactics

The first source of evidence for syllable structure in Banawá comes from phonotactics. In Banawá there are no sequences of more than two adjacent vowels. So, for example, words like the hypothetical examples in (1) and (2) are not observed:

- (1) a. *paeaba
 b. pae a ba
 \|/ | \|/
 σ ? σ
- (2) a. *fuauni, etc.
 b. fua u ni
 \|/ | \|/
 σ ? σ

This fact is predicted without further comment if the language lacks V syllables, as shown in the hypothetical parses in (1b) and (2b). However, if it does have V syllables, then we must also add a stipulation to the effect that V syllables can only follow CV syllables. It might be argued that such a stipulation follows from an independent principle, e.g. an intersyllabic sonority restriction. However, since sonority is generally understood as a syllable-internal principle, this analysis would not be entirely satisfactory. No such complication arises under the ‘no V-syllable’ hypothesis. But the case against V-syllables does not rest exclusively on phonotactics. Let’s consider a second set of problems for the V-syllable hypothesis.

2.2 Alignment

A problem for the hypothesis that V-syllables are part of Banawá prosodic structure comes, paradoxically at first glance, from the behavior of word-initial vowels.⁵ Consider words like those in (3) (cf. also Ladefoged, Ladefoged, & Everett 1997 for phonetic analysis of Banawá consonants, vowels, and stress):

- (3) a. *enémedè* ‘child’
 b. *énéki* ‘middle’
 c. *abébirì* ‘gnat’
 d. *abárikò* ‘moon’
 e. *ukúmu* ‘parasite’
 f. *uwáribìsei* ‘one more (masc. adj.)’

One must ask at least the following questions about the examples in (3). First, why are word-initial vowels never stressed (except in bimoraic words, as pointed out in BBE and 3.1. below)? Second, why is no more than one vowel permitted in word-initial position?⁶ Third, how are Vs allowed in word-initial position at all if no V syllables are allowed in the language? The answer to the final question is simple and leads to an answer to the first: initial Vs are found in the input. Since there are no V syllables, these word-initial Vs are anomalies. They may result from some yet unknown historical development (or some other reason), but they are anomalies for the synchronic phonology nonetheless. The phonology must somehow accommodate itself to them. BBE argue that this accommodation takes the form of extrametricality (which targets a peripheral constituent and places it outside of metrical structure at the relevant level(s)), i.e. that word-initial vowels are invisible to the

metrical structure and hence avoid a violation of syllable structure. In Everett 1996, it is argued that these vowels are outside the prosodic word, because prosodic words align with syllables. Thus word-initial vowels are not stressed because they are not well-formed syllables and are therefore omitted from the prosodic word. It remains to say why there are no word-initial two vowel sequences. In BBE, this results from extrametricality. Two vowels could not be extrametrical because they do not form a constituent and extrametricality targets constituents. Therefore, the noninitial member of the two (or more) word-initial vowels would not be extrametrical. But if it is not extrametrical it would have to be assigned to a syllable. Without an onset, however, it would produce a violation of syllable structure, rendering the word ungrammatical. In an OT approach, the absence of words beginning with vowel sequences is due to an *alignment* constraint, namely, that the lexical word should align with the prosodic word (cf. Everett 1996 for details).⁷ The point is clear in either approach, however: word-initial vowels are unstressed because they are not parsed by feet. This failure to be parsed by feet follows straightforwardly under the assumption that they are not well-formed syllables. Other accounts of this fact are imaginable, but less plausible than the assumption that V-syllables do not exist in Banawá.⁸

2.3. Hypocoristics

Polysyllabic names in Banawá also have bisyllabic, hypocoristic forms. However, a name which is less than three syllables in length has no special hypocoristic form. As the data below show, this phenomenon cannot be stated in terms of moras or V-syllables but only in terms of CV and CVV syllables. Consider the names in (93)-(96). The a. forms give the full form, the b. forms the hypocoristic forms:

	Long Form	vs.	Hypocoristic Form
(7)	a. <i>Sàbatáo</i>		b. <i>Bátað</i> ⁹ (long form from Portuguese <i>Sabatão</i>)
(8)	a. <i>Hóbetò</i>		b. <i>Béto</i> (long form from Portuguese <i>Roberto</i>)
(9)	a. <i>Atídekè</i>		b. <i>Déka</i> ¹⁰
(10)	a. <i>Sólimào</i>		b. <i>Rímað</i> (long form from Portuguese <i>Solimão</i>)
(11)	a. <i>Téresina</i>		b. <i>Sína</i>
(12)	a. <i>Tóefi</i>		b. N/A (cf. <i>*Efi</i>)

The crucial observation in these examples is that hypocoristic forms in Banawá are *bisyllabic*, not *bimoraic*. Thus, the name *Sabatao* corresponds to *Batao*, not **Tao*. Although *Tao* is bimoraic and thus satisfies the word minimality constraint discussed in section 3.1 below, it is not a possible hypocoristic form because it is monosyllabic (CVV). Again, however, there is nothing ill-formed about this form per se. For example, (97) is an actual Banawá name:

(13) *Nao* ‘woman’s name’

The name *Nao* is bimoraic, not bisyllabic. But **Tao* is not a possible hypocoristic form, in spite of the fact that nonhypocoristic names can be monosyllabic (e.g. (13)), because hypocoristic forms are based on syllabic rather than moraic structure. The forms in (7)-(12) are also interesting because they illustrate that a CVV sequence is monosyllabic rather than bisyllabic. If there were V syllables, we would once again be at a loss to account for the failure of *Sabatao* to shorten to *Tao*, rather than grammatical *Batao* (syllabified in this hypothetical case as *ta.o*).

The failure of the name *Toefi* in (12) to have a hypocoristic form is especially interesting in this regard. This failure is not due to the absence of words with the same shape, since words with an initial vowel do occur:

(14) *Ife* ‘man’s name’

The problem with **Efi* in (12b) is that it is not an acceptable hypocoristic form because it is not bisyllabic (although the form itself is grammatical even in common nouns, e.g. *efe* ‘leaf’). Of course, this explanation crucially assumes that CVV (but not V) is a possible syllable in Banawá. If V were a syllable, we would have no explanation for the absence of the form in (12b). On the other hand, one might argue that hypocoristic forms must be consonant initial, thus ruling out **Efi* by an alternative route. However, this suggestion is not what is wanted here, because it is superfluous. We need the bisyllabic constraint to rule out consonant-initial forms like **Tao*. This very same constraint will also eliminate the V-initial forms like **Efi*. Therefore, nothing is gained by adding the hypothesized consonant-initial constraint. The stress patterns in (7)-(12) also show that the presence or absence of a hypocoristic form is not due to stress placement, since the words do not differ in any significant way in this regard. (cf. LLE for further argumentation and discussion of Banawá’s otherwise unremarkable stress pattern.) Banawá thus offers strong evidence that the constraint in (15) must be stated in terms of CV and CVV syllables:

(15) HYPCORISTIC CONSTRAINT: A hypocoristic form is bisyllabic.

We can only account for hypocoristic forms in Banawá if we assume that the language has CVV (and CV) syllables, but not V syllables. Alternative accounts which avoid this implication are not able to account for all the hypocoristic data. For example, one could not simply write a rule such as ‘remove the first syllable of the base to form the hypocoristic’, because this would fail to account for (9), (11), and (12). Therefore, hypocoristic forms

support both (i) the necessity of the syllable in analyzing Banawá prosody and (ii) the CV/CVV analysis proposed above, as opposed to the alternative V/CV analysis.

2.4. Conclusion

Let's sum up the evidence against V-syllables in Banawá:

(16) Against V-syllables:

- a. There are no vowel sequences longer than two vowels¹¹.
- b. Word-initial vowels are not stressed (because they fall outside the prosodic structure of the word, i.e. the prosodic structure above the mora).
- c. Hypocoristics cannot be explained if the language has V syllables.

We have thus established that Banawá has CV and CVV syllables, but no V syllables. It remains to show that CVV syllables can be split by feet. To do this, we must understand the role of the mora in Banawá prosody.

3. Moras and Syllable Integrity

The mora is a crucial component of Banawá phonology. Its significance is seen especially in a constraint on minimal word size and in foot construction. We begin with minimal word size.

3.1. Word Minimality

As argued in BBE (p290ff), Banawá prohibits words of less than two moras. This is restated here as the constraint in (17):

(17) Word Minimality (W_DMIN): Words are minimally bimoraic (W_D ≥ μμ).

There are three main arguments for (17). The first is negative. As BBE observe, Banawá lacks monomoraic words (*=unattested):

(18) **bi*, **ka*, **mu*, etc.

According to BBE, positive evidence is most clearly seen in two facts: (i) the restriction of geminate or long vowels to bimoraic words, (19), and (ii) the fact that the prohibition against stressing word-initial vowels is systematically violated in bimoraic words, (20):

(19) *fáa* 'water'; *dáa* 'the'; *bíi* 'fan'; *búu* 'beat'; *háa* 'also'. (cf. **faaha*; **baadi*; etc.)¹²

(20) *ába* 'fish'; *áwa* 'wood'; *údei* 'I spear'¹³ (cf. *udéibúna* 'I fish-spear' vs. **údeibuna*).

The analysis of (19) proposed by BBE is that long vowels are required to satisfy W_DMIN, which outranks a constraint against adjacent vowels of identical sonority (cf. the starred hypothetical forms in (19)). The analysis of (20) proposed in BBE is that the word-initial vowels cannot remain outside the prosodic word if that would result in a violation of W_DMIN (i.e. if this would leave the word with less than two moras). A third kind of evidence (closely related to the evidence in (19)) for word-minimality is seen in verb types. Verb roots to which inflectional morphology can attach directly can be monomoraic (since they will never surface in this way, given that at least some inflection is obligatory in Banawá). However, verbs which take an auxiliary must be at least bimoraic, since inflection must appear on the auxiliary, not the verb directly. This means that the verb root will correspond to an independent phonological word and so will need to satisfy W_DMIN.

Moreover, long vowels are only found in these latter verbs, i.e. those which take an auxiliary verb.¹⁴

(21) *soo* ‘lay’; *saa* ‘let go of’; *sii* ‘dip out’; etc.

3.2. Stress Placement and Syllable Integrity

3.2.1. Stress Placement

As originally described in BBE, Banawá stresses every other mora from left-to-right within the word, skipping word-initial vowels in words with more than two syllables. The words in (23) illustrate the basic stress rule (see also LLE):

- (23)
- a. *súirì* ‘penis’
 - b. *wárabù* ‘ear’
 - c. *wánakàri* ‘spider’
 - d. *yífuyà* ‘fire’
 - e. *uwárià* ‘one’
 - f. *úwi* ‘cry’
 - g. *uwía* ‘go out (as a fire)’
 - h. *káwarise* *-i* ‘rafter (masc.)’
rafter *-masc.*
 - i. *dábikàri* ‘bee’

3.2.2. Syllable Integrity

The stress pattern described above leads to violations of syllable integrity, as seen in (24)-(27) (see BBE and Everett (1996) for fuller argumentation for that stress pattern; where . = syllable boundary and () = foot boundary; relevant portions are marked in bold):

(24) a. sáyiè 'sound out'

b. (.sa.yi)(e.)
 \ / \ \ /
 σ σ

(25) a. kèrewéduàma 'turn end over end'

b. (.ke.re.)(.we.du)(a.ma)
 \ / \ / \ / \ / / \ /
 σ σ σ σ σ

(26) a. tìkadámuè 'you forget'

b. (.ti.ka.)(.da.mu)(e.)
 \ / \ / \ / \ / /
 σ σ σ σ

(27) a. díhià 'to catch'

b. (.di.hi)(a.)
 \ / \ / /
 σ σ

There are no morpheme boundaries at the relevant syllable boundaries, so we cannot escape the implications of these examples by saying that the 'violated' syllables are heteromorphemic. The conclusion such violations of syllable integrity force on us is that syllable integrity is not an inviolable universal. e.g. built into OT's GENERATE function or deriving from the fact that feet are obligatorily built on syllables (as in Itô & Mester's 1992 'strict layering' hypothesis; cf. also Hayes 1995), but rather that syllable integrity is a violable constraint, outranked by PARSE μ (which requires all moras to appear in feet) in Banawá. Consider the OT tableau in (28), where the optimal (F) parse is the one in which all moras are footed, even when this violates Syllable Integrity.

(28) Tableau (* = head of foot; . = syllable boundary)

	PARSE _μ >>	SYLLINT
* * *		*
F(.ke.re.)(.we.du)(a.ma.)		
* * * *	**!	
(.ke.re.)(.we.) .dua. (ma)		

4. Conclusion

This paper has shown that syllables can be divided between different prosodic feet. A new type of evidence collected in recent fieldwork, expressed in terms of the HYPOCORISITIC CONSTRAINT (which requires that short forms of proper names be bisyllabic), in addition to the atonicity of word-initial vowels and phonotactics, forces this conclusion. This leads to a reconsideration of the nature of the relationship between different prosodic levels. It also suggests that our knowledge of prosody in the world's languages is not sufficient to build a theory around what is 'rare' or 'unattested'. Arguments for phonological theory must come from other considerations. Nevertheless, Hayes (1995, 49ff) makes a convincing case that Syllable Integrity is respected in various languages. The conclusion forced upon us, then, is that Syllable Integrity is a constraint of UG, ranked highly in some languages (those which do not violate it) and lower in other languages.¹⁵

References

- Blevins, Juliette. 1995. 'The syllable in phonological theory', In: John Goldsmith (ed.), *The handbook of phonological theory*, Blackwell, Oxford, pp206-244.
- Buller, Barbara, Ernest Buller and Daniel L. Everett. 1993. 'Stress placement, syllable structure, and minimality in Banawá', *IJAL* 59: 280-293.
- Everett, Daniel L. 1990. 'Extraprosodicity and minimality in Kamã and Banawá', *Pitt Working Papers in Linguistics* 1, 17-41.
- Everett, Daniel L. 1995a. 'Word minimality, word binarity, and foot minimality in Suruwahá', *Linguistics Society of America Annual Meeting*, New Orleans.
- Everett, Daniel L. 1995b. 'Sistemas prosódicas da família Arawá', In: Leo Wetzels (ed.), *Estudos fonológicos de línguas indígenas brasileiras*, Editora da Universidade Federal do Rio de Janeiro.
- Everett, Daniel L. 1996. 'Prosodic Levels and Constraints in Banawá and Suruwahá', *Rutgers Optimality Archive* (<http://ruccs.rutgers.edu/roa.html>), paper ID = ROA-121-0496.
- Halle Morris and Jean-Roger Vergnaud. 1987. *An essay on stress*, MIT Press, Cambridge, MA.
- Hayes, Bruce. 1995. *Metrical stress theory: principles and case studies*, University of Chicago Press, Chicago.
- Itô, Junko and R. Armin Mester. 1992. *Weak layering and word binarity*, Linguistics Research Center, Cowell College, UCSC, Santa Cruz, CA.
- Ladefoged, Peter, Jenny Ladefoged, and Daniel Everett. 1996. 'The sounds of Banawá: an endangered language', submitted to *Phonetica*.
- McCarthy, John and Alan Prince. 1995. *Faithfulness and reduplicative identity*, ms., Optimality Archive, Rutgers University.
- Prince, Alan & Paul Smolensky. 1993. *Optimality theory: constraint interaction in generative grammar*, Technical Reports of the Rutgers University Center for Cognitive Science.
- Zec, Draga. 1988. *Sonority constraints on prosodic structures*, unpublished Ph.D. dissertation, Stanford University.

Notes

¹ I would like to thank Peter Ladefoged, Ernie Buller, Annie Buller, John McCarthy, Bruce Hayes, Colleen Fitzgerald, and the stimulating group who attended the Fourth Annual Phonology Conference at the University of Manchester for helpful comments. They are not to be blamed for any silliness or lack of clarity of my own which has survived in this paper. This research was funded by NSF Grants SBR93-10221 and SBR96-31322. I owe a special thanks to Paul Chapin for suggestions in writing the proposals for these grants, without which the present results would be nonexistent. Previous studies of Banawá include Everett (1990, 1995a, 1995b, and 1996) and Buller, Buller, and Everett (1993). Banawá is spoken by approximately seventy people near the Kitiá/Banawá river in the state of Amazonas, Brazil. It is a member of the Arawan language family, which also includes Suruwahá, Paumari, Kulina, Deni, Yarawara, and Yamamadi, cf. Everett (1995b) for an overview of the prosodic systems of all the known Arawan languages.

² For instance, the feet in (i) violate syllable integrity, but the feet in (ii) do not (where . = syllable boundary and () = foot):

- (i) (.μ.μ)(μ.μ)(μ.)
 (ii) (.μ.μμ.)(μμ.)

³ Hayes claims, for instance, that ‘Following earlier work ..., I adopt the view that in stress languages, the stress-bearing unit is the syllable.’ (Hayes 1995, 49)
 If this view is correct, then, as Hayes argues, we expect that languages will not allow syllables to be divided between feet. This is so because stress is determined by foot structure. If only syllables can bear stress, then the immediate constituents of feet can only be syllables, never moras.

⁴ Of course, HV might turn out to be correct, i.e. it could turn out to be the case that syllable integrity is commonly violated, as we see more data.

⁵ ‘Paradoxically, because these would usually strike the fieldworker as evidence for V-syllables, under the average fieldworker’s assumption that segments do not appear outside of syllables.

⁶ So, for example, words like those in (i) are not found:

- (i) *aiba, *eupi, *aeisu, etc.

⁷ One might ask why unsyllabified vowels are not simply deleted. The answer within OT is that some CORRESPONDENCE (cf. McCarthy & Prince 1995) constraint must be ranked so as to prevent this deletion. The relevant constraints are:

- (i) MAX IO: Every segment of the input has a correspondent in the output.
 (ii) DEP IO: Every segment of the output has a correspondent in the input.

Consider the Tableau in (iv) and the treatment of word-initial, unsyllabified vowels (where < > = deleted):

(iii) *ufábonè* ‘I will drink.’

(iv) Tableau

	MAX IO >>	PARSE _μ >>	DEP IO
<u>[(fabo)(ne)]	*!		
F u[(fabo)(ne)]		*	

Therefore, any deletion rule will fatally violate MAX IO, effectively barring deletion for word-initial vowels.

⁸ For example, one could simply say that the prosodic word prefers to align with a consonant. If this were true, there would be no need to appeal to the lack of V-syllables to account for the lack of stress on word-initial Vs. However, while it is quite natural to state that a prosodic unit, the prosodic word, must align with another prosodic unit, e.g. a syllable, it is difficult to see what the intuitive basis might be for saying that prosodic words like to align with consonants. The role of consonants in the alignment process is surely derived from their role in the syllable, i.e. as onset.

⁹ For the sake of illustration here, I have shown the form of the hypocoristic that would be found in connected text, rather than the vocative form found in isolation. The isolated vocative forms are stressed by iambic feet (and stress goes on the most sonorant vowel of the syllable in the vocative forms):

(i) Vocative forms: *Batáao*; *Betó*; *Deká*; *Rimáao*; *Siná*.

¹⁰ I do not know why there is an [a] in the shortened form in this word.

¹¹ Also relevant is the fact that all adjacent vowels have different sonority values - except in bimoraic words; cf. (19) and discussion preceding it in the next section. These facts are predicted if adjacent vowels are tautosyllabic within CVV syllables, subject to syllable-internal sonority sequencing restrictions; cf. Zec 1988 and Everett 1996.

¹² There are exceptions to this generalization. Some of these exceptions (about a dozen words in all) are interjections, e.g. *eee*, an agreement particle. Others are the result of consonant deletion in morphophonemic combinations. But one or two appear to have no synchronic explanation at all. The most obvious example is *baama* 'catfish'. This derives historically from *bahama* (and in fact still has this form in most other Arawan languages). That is, the intervocalic /h/ has been dropped. But since the word is morphologically simple, so far as I can tell, it should have been shortened to *bama*. Indeed some speakers alternate between *baama* and *bama* in their pronunciation. However, at this point I have no satisfying account of this word.

¹³ The masculine suffix, {-i} - in words like *udei* - 'I spear' is never stressed when word-final.

¹⁴ Interestingly, as shown in Everett (1996) these facts also mean that although feet in Banawá are normally bimoraic, we cannot derive WDMIN from minimum foot size, because Banawá allows monomoraic, degenerate feet, yet prohibits monomoraic words.

¹⁵ Alternatively, one could propose the inherently less interesting idea that it is present in some languages but absent in others.