



The IP as the Domain of Syllabification

Bruce Birch

Department of Linguistics and Applied Linguistics
University of Melbourne, Australia
b.birch@linguistics.unimelb.edu.au

Abstract

Evidence suggests that in Iwajja, a non-Pama-Nyungan language of northern Australia, the Intonation Phrase (IP) is an integrated entity, one which is typical of a *cursus* language (Pulgram, 1970) like French, in which words in connected speech give up some of the properties they exhibit in isolation. A salient feature of the prosody of this language is syllabification across content word boundaries. The paper examines the way in which this phenomenon relates to two influential versions of the prosodic hierarchy, both of which suggest that syllabification is subsumed, one way or another, within the boundaries of word-sized units. On the basis of acoustic analysis, the paper suggests that the IP, rather than the Phonological or Prosodic Word may form the domain for syllabification in Iwajja.

1. Introduction

1.1. Syllabification and prosodic hierarchies

The aim of this paper is to explore the idea that a constituent other than the phonological or prosodic word forms the domain for syllabification in Iwajja. This notion is not a generally accepted one. Whilst models of the prosodic structure of speech differ both in terms of the number of levels they allow and the way in which those levels are defined, they tend to agree that syllables are formed within the boundaries of word-sized constituents. The position of these constituents in prosodic hierarchies may vary enormously, however.

Table 1: A comparison of two hierarchies.

N + V (1986)	OT (Kager 1999)
Utterance	
Intonation Phrase	
Phonological Phrase	
Clitic Group	
Phonological Word	Prosodic Word
Foot	Foot
Syllable	Syllable
	Mora

1.2. Nespor and Vogel

Nespor and Vogel (1986), for example, present a seven-level hierarchy (see Table 1, Column 1). The authors suggest that a hierarchy of prosodic constituents is motivated by the fact that not all phonological processes can be adequately explained by reference to morpho-syntactic categories. Each

of the seven levels, therefore, is justified on the basis that it forms a domain for phonological processes. For example, the aspiration of /t/ in *satire*, but not in *satyr*, is explained as a syllable juncture rule operating on the domain of the *foot*. (p.91). Or, to take a higher-level constituent, the *phonological phrase* is required as the domain for Radoppiamento Sintattico in Italian. The domain of application for this rule "cannot be identified with any syntactic constituent since syntactic constituents of the same type may behave differently with respect to the rule." (p.165).

The failure of the boundaries of syllables and feet to consistently align with proposed phonological word boundaries is a problem for all models of prosodic structure which posit a word level constituent. Regarding the relationship between feet and phonological words, Nespor and Vogel state that

"as required by the Strict Layer Hypothesis . . . , all the feet of a given string must be grouped into phonological words, and no other category may be so grouped. Each foot is thus exhaustively included in a ω ; that is, it is never the case that the syllables of a single foot belong to two different phonological words." (Nespor and Vogel, 1986, p.109)

The authors deal with phenomena such as liaison in French by positing two levels of syllabification – one at the phonological word level, and a second *resyllabification* at a higher level. (See below for further discussion.)

1.3. Optimality Theory

Another model which is concerned with the mapping between morphosyntax and prosody is that used predominantly in Optimality Theory. This model originally proposed a three-level hierarchy (McCarthy and Prince, 1993), although its more recent manifestations (Kager, 1999) include the mora as a fourth level below the syllable (see Table 1, Column 2). The motivations for such a comparatively minimal structure seem to be theory-internal. McCarthy and Prince (1993) state that their "results will be most secure" if they "adhere to a minimally elaborated theory, so that they [the results] emerge from the judicious application of Alignment rather than some dubious cleverness in the assignment of the constituents themselves." Whilst this model allows that alignment of syllable boundaries with the edges of Prosodic Words is dependent on language-specific rankings of universal constraints, it does not allow that the relevant boundary for syllable alignment be a phrase-level constituent, for the simple reason that the theory does not include such constituents in its inventory.

Whilst both hierarchies in the above table include a word-sized constituent, the OT model positions it at the top, whereas Nespor and Vogel place four constituents above it

(although the latter do not go so far as to provide representations of actual utterances, showing their constituency from the lowest level through to the highest, and demonstrating how the pieces fit together).

The fact that the OT literature positions the Prosodic Word at the top of its hierarchy means that, whilst much interesting analysis is presented in the literature regarding the interaction between roots and affixes and between content words and function words, we are rarely, if ever, shown how the constituents posited on the basis of such interaction are fitted together to form phrases and utterances in a discourse. In other words, there is a failure on the part of the OT literature to examine *context*, and therefore to build context-sensitive variation in the prosodic structure of words into its analyses. The following presentation of data from Iwajja, showing the formation of syllables across word boundaries, suggests that the inclusion of context is essential if an accurate picture of prosodic structure is to be drawn.

2. Syllabification in Iwajja

2.1. The morphological word in isolation

The following examples are taken from two related field recordings made on Croker Island between 1996 and 2000. A partially transcribed recording of an Iwajja speaker explaining aspects of rituals surrounding childbirth, was played back to a second speaker, who was asked to translate parts of the text. The second speaker's explication was also recorded. As part of this process, the second speaker was asked to hyperarticulate some of the more difficult words for transcription purposes. This resulted in the emergence of syllables which were absent from tokens of the same word in the first recording. The word used in the examples occurs several times spontaneously in the original text. It consists of a verbal root and two affixes, i.e., a prefix encoding the verb's arguments, and the 'repetitive' suffix (see below).

For the purposes of the analysis, the following minimal prosodic structure is assumed. The basic set of Iwajjan syllable structures is {[V(C)], [CV], and [CVC(C)]}, with a fourth type - [CVV(C)] - occurring in all forms of speech other than hyperarticulation. Feet are either bimoraic or disyllabic trochees. The only prosodic constituent above the foot for which there is clear evidence at this stage is the IP. The boundaries of the IP are marked by both tonal and durational features and contain a nucleus marked by the alignment of a pitch accent with the head of a foot.

Table 2: The target word.

aŋku -	wijbani -	ki
3plA>3sgO -	pile-up -	REP
arrange		
"They arrange it."		

In hyperarticulated tokens of this word, all syllable boundaries are clearly represented in waveform, intensity and f0 trace. A slight release is often audible after the coda of the syllable [wij]. The fact that such a release tends to occur only at IP boundaries suggests that the speaker separated the word into two IP's. Note that the word-internal IP boundary does not coincide with a morpheme boundary.

1. (aŋkuwij)_{IP} (baniki)_{IP}

When the speaker produced a rendition at the level of articulation typical of casual speech, such a release was no longer in evidence, suggesting that the entire word was isomorphic with a single IP. In addition, the glide onset to the syllable [wij] was deleted. All acoustic analyses (waveform, intensity trace, f0 trace and spectrogram) suggest a transition to a single nucleus resulting in the formation of a bimoraic closed syllable, i.e., [ku] + [wij] → [kuij].

2. (aŋkuijbaniki)_{IP}

It is important to note here that variation in the prosodic realization of a morphological word in isolation is not simply a result of varying rates of speech. It must also be seen as reflecting differing constituency structures, e.g. the realization of the word as two IP's versus a single IP, or the merging of two syllables into one. This highlights the fact that 'connected speech processes' (in this case glide deletion) are not solely processes occurring *between* words, but also *within* them. I mention this in order to press the point that an approach which positions citation forms *outside* of the category 'connected speech' is not valid if processes diagnostic of connected speech are observable within isolated tokens.

2.2. The morphological word in IP-medial position

Examples 3 and 4 (below) show the target word embedded in longer IP's and subjected to further phonological processes, namely vowel coalescence and *i*-deletion. In both tokens of the target word, the vowel-initial first syllable has gained an onset from the final syllable of the preceding word, and the monosyllabic suffix has lost its nucleus. This has resulted in the onset of the final syllable docking with the onsetless initial syllable of the following word. In Example 3, the following word is a content word, whilst in Example 4 it is a function word. The same process is seen to apply in both cases.

The first line in each example shows the syllable structure for hyperarticulated tokens of each component word. The second line shows the instantiated syllabification and foot structure for each IP. As there is a tendency toward isochrony at the level of the foot in Iwajja, feet are diagnosed on the basis of durational measurements, as well as segmental structure. Hence, whereas in Example 3 the first two syllables of the target word form separate feet, in Example 4 they form a single foot. It is worth noting that the release associated with the final segment of *atɔŋ* in Example 4, evident in both the waveform and the intensity trace, is diagnostic of an IP boundary. Whilst it does not occur consistently, it is *only* found following IP-final tokens, i.e., no IP-medial occurrences have been found.

3. [ji . jiiŋ . ka] [aŋ . ku.wij . ba.ni.ki] [a.ɖa . muŋ.kun]

((jiiŋ) (kaŋ) (kuij) (ba.ni) (ka.ɖa) (muŋ.kun))_{IP}

Table 3: Gloss and translation of Example 3.

jiiŋka	aŋku -	wijbani -	ki	aɖamuŋkun
ground-oven	3A>3O -	pile-up -	REP	antbed
arrange				
"They arrange the termite-mounds for the ground-oven."				

4. [bu. wa. ni] [aŋ. ku. wij. ba. ni. ki] [aʊaʃ]

(bə(waan) (yaŋ.kuij) (ba.ni) (kaaʃ))_{IP}

Table 4: Gloss and translation of Example 4.

buwani	aŋku -	wijbani -	ki	aʊaʃ
<i>3pl-sit</i>	<i>3A>3O-</i>	<i>pile-up -</i>	<i>REP</i>	<i>inside</i>
		<i>arrange</i>		
"They arrange them inside."				

3. Discussion

3.1. Syllabification and footing in the IP domain

As Fudge (1999) has suggested, "a complication in stating the relation between feet and phonological words is that many words may be 'footed' in several different ways . . ." (p.281). One possible solution to this type of conflict is to shift the boundaries of the phonological word to align with syllable and foot boundaries. This has been suggested for Italian by Peperkamp (1997). Thus for a phrase combining the two morphological words *bar* and *aperto*, the syllabification *ba.ra.per.to* is extended to the phonological word level, yielding the two 'restructured' constituents (ba)_ω and (raperto)_ω. As Hall (1999) points out, this analysis proves problematic in two ways, however. Firstly, according (ba) the status of a phonological word violates the bimoraic minimum which Peperkamp requires for phonological words in Italian, and secondly, the non-alignment of phonological word boundaries with morpheme boundaries violates the generalization that phonological words "apply to morphemes and not to arbitrary sequences of sounds." (Hall 1999, p.15).

Fudge's statement (above), which obviously owes far more to Abercrombie (1964) than to the proponents of Optimality Theory and the Strict Layer Hypothesis, brings into focus the role of context in the prosodic realization of morphological words and opens for discussion a second possibility that, in some languages at least, the IP, or at least some phrase-level unit, rather than the phonological or prosodic word, may be the relevant domain for syllabification and foot construction.

Under such a view, the alignment of syllables and feet with the boundaries of morphological words uttered in isolation is interpreted as alignment with the boundaries of the intonation unit with which the word is isomorphic (see Example 2). In the case of Example 1 (above), where a single word spans more than one IP, alignment of syllables and feet with word-internal IP boundaries is predicted. The non-alignment of syllables and feet with IP-internal morphological word boundaries (as in Examples 3 and 4) is also predicted, and therefore potentially unproblematic. The boundaries of morphological words are not required to be marked prosodically and may be obscured by the foot structure, which in turn is facilitated by processes such as deletion and coalescence.

3.2. The notion of resyllabification

In identifying the phonological word as the domain in which syllabification occurs, Nespor and Vogel (1986) suggest that there are some languages (like English and Dutch) where syllabification tends to occur within

phonological word boundaries and others (like Spanish and French) where syllabification across phonological word boundaries is commonplace. They suggest however, that although it is tempting to propose some unit larger than the word as the default domain of syllabification in languages which permit syllabification across word boundaries, such a proposal would be inappropriate since, even in such languages, there exist phenomena which can only be adequately explained by ordered rules which refer to the phonological word.

The one example they cite in support of this claim involves the rule in non-meridional French whereby [e] and [ə] may not occur in closed syllables. In this environment, both vowels are realized as [ɛ]. For instance, *premier* contrasts with *première*. The authors assert that in contexts where the final segment of *première* forms the onset of a following onsetless syllable via enchaînement (e.g. *la première année* which syllabifies as *la.pre.myè.ra.née*) the quality of the vowel remains as if it were still in a closed syllable. It is necessary therefore, they suggest, to refer to two stages of syllabification – one at the word level, which happens first, and a second, a resyllabification,¹ at some higher, yet to be clearly defined level, which can only occur on the basis of the first.

This observation, which seems to lack experimental support, does perhaps suggest two levels of syllabification in French, and provides a clear example of the type of serial derivation which presents major challenges for Optimality Theory.² The fact that such a phenomenon has been observed for French does not entail, however, that all languages which exhibit frequent syllabification across word boundaries must also have two stages of syllabification.

3.3. Psycholinguistic evidence

Syllabification and foot-construction at the phrase-level finds support from psycholinguistic research, where it has been suggested that phrases should be understood "not as the sequencing of any sort of entity, but rather as the unfolding of a temporally extended, integrated, hierarchically organized rhythmic act." (Martin, 1983). "Studies have proposed the existence of what has been termed a "post-lexical buffer into which the phonological content of the words in the utterance is transferred." (Pierrehumbert, 1993). As Pierrehumbert points out, Shattuck-Hufnagel (1979), for example, uses speech error data such as spoonerisms to argue for the existence of such a buffer, which suggests a prosodic blueprint exists for phrases and utterances before they are spoken, and Sternberg et al (1980) suggest that the units of the buffer are metrical feet. Pierrehumbert (1993) points out that there is also strong linguistic evidence for such a buffer "since the phrasal prosody must be constructed somewhere." (Pierrehumbert, (1993) p.272). She continues,

"Since the phrasal prosody constructed over word tokens does not permanently impact the lexical

¹ It should be noted here that, given that the authors allow for the possibility that syllables are formed 'prior' to words and adjusted according to "word formation processes", a resyllabification at phrase-level implies the possibility of a *three-stage process* - 'pre-word', word, and phrase.

² See McCarthy (1999) for an OT analysis using the notion of sympathy.

representations of the word types, it is clear that this construction must take place outside of the lexicon." (ibid.)

Whilst the author is clearly referring here to intonational features, both her argument and her evidence are readily adaptable to the notion of 'post-lexical' feet and syllables, i.e. those which ignore word boundaries.

Conclusion

As mentioned earlier, authors who take as their starting point the McCarthy and Prince hierarchy, and along with it the Optimality Theory framework, do not tend to confront issues such as syllabification and footing across content word boundaries. This is because they tend to take 'decontextualized' examples of words and examine their internal structure rather than look at what happens when these words combine with each other to form phrases. For such authors, the notion of *connected speech* is frequently taken to refer to a categorically distinct form of production, in which *connected speech processes* affect to a greater or lesser extent the prosodic realization of words which exhibit their default prosodic patterning in citation form. In the view adopted here, the citation form, i.e. the uttering of a morphological word in isolation, rather than being characterized as a neutral or default form, is viewed as no less contextually bound than forms of the word embedded in larger utterances.

As Pierrehumbert (in press) has recently written in support of the approach taken by Bruce (1973) in his study of Swedish intonation,

"Much early work on prosody and intonation (such as Fry 1958) takes citation forms of words as basic. Insofar as the intonation of continuous speech was treated at all, it was in terms of concatenation and reduction of word patterns which could have been found in isolation. Bruce, in contrast, adopted the working hypothesis that the "basic" word patterns were abstract patterns whose character would be revealed by examining the full range of variation found when words are produced in different contexts.

... The citation form is then reconstructed as the form produced in a specific prosodic context – when the word is both phrase-final and bears the main stress of the phrase. The importance of this point cannot be overemphasized. In effect there is no such thing as an intonation pattern without a prosodic context. The nuclear position and the phrase-final position are both particular contexts, and as such leave their traces in the intonation pattern."

Pierrehumbert, J. *Tonal Elements and their alignment* (in press).

What Bruce has shown for Swedish applies equally to Iwajja, for example. When a word is uttered in isolation, it carries with it all the prosodic features which go to make up a well-formed IP. Its boundaries are marked by tones and by durational features which are demarcators of IP boundaries and it contains a head or nucleus which is the head or nucleus of an IP, but *not* clearly the head or nucleus of some word-level prosodic constituent. And isolated tokens of words tell

us virtually nothing as to the domain in which syllabification and footing occur. In order to capture the tendency for the boundaries of morphological words in Iwajja not to be clearly demarcated phrase-internally, it would seem necessary therefore, that the following approach be taken. Rather than attempt to identify a single prosodic structure for morphological words in Iwajja, we identify a set of *context-sensitive variants* and articulate, where possible, which elements in the context are the factors relevant to the observed prosodic outcomes.

References

- [1] Abercrombie, D. 1964. Syllable quantity and enclitics in English. In Abercrombie, D.; Fry, D.; MacCarthy, P.; Scott, N.; Trim, J. eds. *In Honour of Daniel Jones*. London: Longmans.
- [2] Fudge, Eric. 1999. Words and Feet. *Journal of Linguistics*; 1999, 35, 2, July, 273-296.
- [3] Hall, Alan. T.; Kleinheinz, Ursula. eds. 1999. *Studies on the Phonological Word*. Amsterdam: J. Benjamins.
- [4] Kager, Rene. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- [5] Martin, James.G. 1983. Aspects of rhythmic structure in speech perception. *Rhythm in Psychological, Linguistic and Musical Processes*. Illinois: Charles C. Thomas.
- [6] McCarthy, John, J.; Prince, Alan. 1993. Generalized Alignment. *Yearbook of Morphology*; 1993, 79-153.
- [7] Nespor, M.; Vogel, I. 1986. *Prosodic Phonology*. Dordrecht, Holland: Foris
- [8] Pierrehumbert, Janet. 1993. Alignment and Prosodic Heads. *Proceedings, Eastern States Conference on Linguistics-(ESCOL)*; 1993, 10, 268-286.
- [9] Pierrehumbert, Janet. (In press). Tonal Elements and their Alignment.
- [10] Peperkamp, Sharon. 1997. Output-to-Output Identity in Word-Level Phonology. *Linguistics-in-the-Netherlands*; 1997, 14, 159-170.
- [11] Pulgram, E. 1970. *Syllable, word, nexus, cursus*. The Hague: Mouton.
- [12] Shattuck-Hufnagel, Stefanie.; Klatt, Dennis. H. 1979. The Limited Use of Distinctive Features and Markedness in Speech Production: Evidence from Speech Error Data. *Journal of Verbal Learning and Verbal Behavior*; 1979, 18, 1, Feb, 41-55.
- [13] Sternberg, S.; Wright, C.E.; Knoll, R.L.; Monsell, S. (1980) Motor programs in rapid speech: Additional evidence. In R.A. Cole (ed) *Perception and production of fluent speech*. Hillsdale NJ: Lawrence Erlbaum Associates.

Additional material, including sound files and acoustic data for the examples in this paper, are available at:

<http://www.linguistics.unimelb.edu.au/contact/studentsites/birch/index.html>