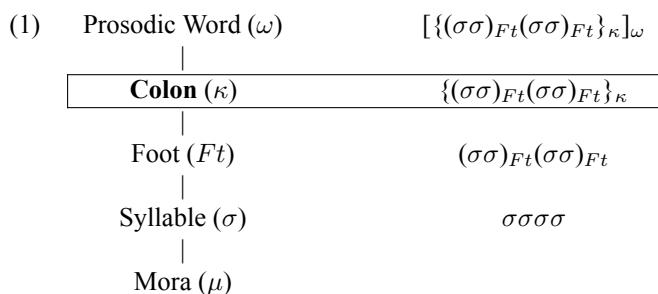


# The Colon as a Separate Prosodic Category: Tonal Evidence from Paicî (Oceanic, New Caledonia)

Florian Lionnet

## 1. Introduction

This paper presents new evidence supporting the inclusion of the colon ( $\kappa$ ) as a separate category in the Prosodic Hierarchy, on the basis of tonal data from Paicî, an Oceanic language of New Caledonia. The colon is a constituent intermediate between the foot and the word, made of two feet, as schematized in (1) (Stowell 1979; Halle & Clements 1983: 18-19; Hammond 1987; Hayes 1995: 119; Green 1997; a.o.).



Justification for the colon ( $\kappa$ ) rests mostly on the existence of tertiary stress. As clearly summarized by Green (1997: 102), “it is clear that in order to derive four levels of stress (primary, secondary, tertiary, unstressed), four levels of structure (prosodic word, colon, foot, syllable) are called for.” Colon-based analyses have so far been proposed for a dozen languages, listed in (2) below.

- (2) a. Passamaquoddy (Stowell 1979; Hayes 1995: 215-216; Green 1997: 104-109),  
 b. Tiberian Hebrew (Dresher 1981),  
 c. Garawa (Halle & Clements 1983: 20-21; Halle & Vergnaud 1987: 43; Hayes 1995: 202),  
 d. Hungarian (Hammond 1987; Hayes 1995: 330; Green 1997: 102-104),  
 e. Maithili (Hayes 1995: 149-162),  
 f. Eastern Ojibwa (Hayes 1995: 216-218; Green 1997: 109-112),  
 g. Asheninca (Hayes 1995: 288-296; Green 1997: 112-114),  
 h. the Neo-Štokavian dialect of Serbo-Croatian (Green 1997: 115-116),  
 i. three Goidelic varieties: Munster Irish, East Mayo Irish, and Manx (Green 1997: 120-133).  
 j. Iquito (Michael 2011; Topintzi 2016, 2017).

The analysis proposed by Hammond (1987) for Hungarian tertiary stress is given in (3) as an illustration, using Hammond’s notation, where the acute accent  $\langle \acute{\text{a}} \rangle$  indicates primary stress, the circumflex accent  $\langle \hat{\text{a}} \rangle$  secondary stress, and the grave accent  $\langle \grave{\text{a}} \rangle$  tertiary stress.<sup>1</sup>

\* Florian Lionnet, Princeton University, flionnet@princeton.edu. I would like to thank H el ene Nimbaye for kindly accepting to spend some of her precious time going over Jean-Claude Rivierre’s (1974, 1983) Paic i data with me: *Ol e!* I also thank the WCCFL 36 audience for helpful comments and feedback. All errors are my own.

<sup>1</sup> Note that Green (1997) uses a slightly different notation, where secondary stress is marked with a grave accent  $\langle \grave{\text{a}} \rangle$  and tertiary stress with a circumflex  $\langle \hat{\text{a}} \rangle$ .



## b. Consonants:

p	pw	t	c	k
b [ᵐb]	bw [ᵐbʷ]	d [ᵐd]	ʃ [ʃ]	g [ᵐg]
m	mw	n	ɲ	ŋ
	w	r		
		l		

As we will see, Paicî is a language that counts morae. Monomorphemic lexical items are rarely longer than five morae.<sup>4</sup>

## 2.1. Paicî tonal inventory

Paicî is a tonal language. The tone-bearing unit is the mora. Only vowels are moraic in Paicî, where coda consonants are not allowed. Rivierre (1974) posits three contrastive tone heights: high (H), mid (M), and low (L). I propose to reinterpret it as a two-tone system. There is, indeed, no distinction between Rivierre's underlying M and L. The tone system of Paicî can thus be described as involving only two underlying level tones: H (Rivierre's H) vs. L (Rivierre's M+L). Based on Rivierre's description, the L tone is indeed better analyzed as a derived downstepped L. The reanalysis I propose here, schematized in § 2.1, closely follows Rivierre's (1978: 430, 1993: 161) description and intuition.

Rivierre (1974)	Reanalysis
H	H
M	L
L	↓L

**Table 1:** Paicî tone inventory

All of Rivierre's M-toned words will henceforth accordingly be transcribed as L-toned. Most lexical items in Paicî are isotonal, i.e. either all H or all L. The following examples illustrate the H vs. L tonal contrast with a few minimal pairs taken from Rivierre (1983).

- (5)
- |    |       |                     |       |                       |
|----|-------|---------------------|-------|-----------------------|
| 1μ | ú     | 'bad spirit'        | ù     | 'breath'              |
| 2μ | kóó   | 'humidity, cold'    | kòò   | 'tree sp.'            |
|    | pʌdí  | 'to hit, to thrash' | pʌdì  | 'to share, to divide' |
| 3μ | pwáái | 'to fill, to load'  | pwàài | 'tree sp.'            |
|    | údłrî | 'to catch on fire'  | ùdłrî | 'to disjoin'          |

## 2.2. Downstep

H tones are stable, i.e. they are not affected by any tonal processes, and in general never change: a H-toned mora always surfaces as H. L tones, on the other hand, are targeted by two distinct tonal processes: downstep and juncture H-tone spreading (see § 3.2). Downstep is illustrated in (6): while L-toned words of one to three morae are realized with a level L tone throughout (6a), words of four morae and above undergo a register drop after the second mora (6b).

- (6) a. 1 ~ 3μ words:
- |        |       |            |        |              |
|--------|-------|------------|--------|--------------|
| ̀̀̀:   | ù     | 'breath'   | pwl̃   | 'turtle'     |
| ̀̀̀̀:  | nêê   | 'name'     | cłm̀̀  | 'to plant'   |
| ̀̀̀̀̀: | pwàài | 'tree sp.' | ùdłr̀̀ | 'to disjoin' |

<sup>4</sup> There is both language-internal and language-external evidence for the importance of moraic count in Paicî: every line in the traditional epic poems called *ténô* [ténɔ̃] consists (in theory) of eight morae (cf. Bensa & Rivierre 1976, 1994).

- b. 4 $\mu$ + words:  
 $\dot{\mu}\dot{\mu}^4\dot{\mu}\dot{\mu}$ : àù+kòð ‘cagou’ pɔ̀ʒàʔjì ‘molar teeth’  
 $\dot{\mu}\dot{\mu}^4\dot{\mu}\dot{\mu}\dot{\mu}$ : èàʔàràbwà ‘crab sp.’ pwèrèʔtòðtì ‘wind’

The downstep analysis is justified by the fact that this lower register is then maintained throughout the remainder of the utterance, as shown in (7) (from Rivierre 1974).

- (7) a. /àbòrò kírí/ [a<sup>3</sup>bo<sup>3</sup>ro<sup>3</sup> ki<sup>5</sup>ri<sup>5</sup>]  
 person small  
 ‘small/short person’
- b. /àbòrò m̀̀ìǹ̀/ [a<sup>3</sup>bo<sup>3</sup>ro<sup>3</sup> m̀̀i<sup>3</sup>ǹ̀i<sup>3</sup>]  
 person big  
 ‘big/tall person’
- c. /pàìèè kírí/ [pa<sup>3</sup>i<sup>3</sup>e<sup>1</sup>e<sup>1</sup> ki<sup>3</sup>ri<sup>3</sup>] i.e. [pàìʔèè kírí|]  
 serpentine small \*[pa<sup>3</sup>i<sup>3</sup>e<sup>1</sup>e<sup>1</sup> ki<sup>5</sup>ri<sup>5</sup>]  
 ‘small (piece of) serpentine’
- d. /pàìèè m̀̀ìǹ̀/ [pa<sup>3</sup>i<sup>3</sup>e<sup>1</sup>e<sup>1</sup> m̀̀i<sup>1</sup>ǹ̀i<sup>1</sup>] i.e. [pàìʔèè m̀̀ìǹ̀]  
 serpentine big \*[pa<sup>3</sup>i<sup>3</sup>e<sup>1</sup>e<sup>1</sup> m̀̀i<sup>3</sup>ǹ̀i<sup>3</sup>]  
 ‘big (piece of) serpentine’
- e. /pàìèè bẁ̀h̀̀c̀̀r̀̀/ [pa<sup>3</sup>i<sup>3</sup>e<sup>2</sup>e<sup>2</sup> bẁ̀i<sup>2</sup>i<sup>2</sup>ci<sup>1</sup>ri<sup>1</sup>] i.e. [pàìʔèè bẁ̀h̀̀c̀̀r̀̀]  
 serpentine green \*[pa<sup>3</sup>i<sup>3</sup>e<sup>1</sup>e<sup>1</sup> bẁ̀i<sup>3</sup>i<sup>3</sup>ci<sup>1</sup>ri<sup>1</sup>]  
 ‘green serpentine’

### 3. Analysis

The two main questions posed by the data above are 1) how to explain the downstep after the second mora in 4 $\mu$ + words; and 2) how to explain the absence of downstep after the second mora in 1 ~ 3 $\mu$  words?

#### 3.1. Foot-based account

To answer the first question, the downstep pattern strongly suggests that the two initial morae of 4 $\mu$ + words form a bimoraic foot ( $\dot{\mu}\dot{\mu}$ ). A foot-based analysis thus seems to be in order. The presence of one word-initial L-toned foot is, however, not sufficient to account for the downstep pattern, since it would predict downstep after the second mora of di- and trisyllabic words, contrary to the facts presented in section 2, as shown in (8) below.

- (8)
- |           | Prediction   | Facts                                    |          |   |
|-----------|--|--|----------|---|
| 1 $\mu$ : | $\dot{\mu}$  | $\dot{\mu}$                              | pẁ̀     | ✓ |
| 2 $\mu$ : | ( $\dot{\mu}\dot{\mu}$ ) <sup>4</sup>                          | $\dot{\mu}\dot{\mu}$                     | c̀̀m̀̀   | * |
| 3 $\mu$ : | ( $\dot{\mu}\dot{\mu}$ ) <sup>4</sup> $\dot{\mu}$              | $\dot{\mu}\dot{\mu}\dot{\mu}$            | ùd̀̀r̀̀  | * |
| 4 $\mu$ : | ( $\dot{\mu}\dot{\mu}$ ) <sup>4</sup> ( $\dot{\mu}\dot{\mu}$ ) | $\dot{\mu}\dot{\mu}^4\dot{\mu}\dot{\mu}$ | pɔ̀ʒàʔjì | ✓ |

The downstep seems to occur between two L-toned feet, a tonal dissimilation akin to an Obligatory Contour Principle effect. This analysis accounts for the data we have seen so far:

- (9)
- |           | Prediction   | Facts                                    |          |   |
|-----------|--|--|----------|---|
| 1 $\mu$ : | $\dot{\mu}$  | $\dot{\mu}$                              | pẁ̀     | ✓ |
| 2 $\mu$ : | ( $\dot{\mu}\dot{\mu}$ )                                       | $\dot{\mu}\dot{\mu}$                     | c̀̀m̀̀   | ✓ |
| 3 $\mu$ : | ( $\dot{\mu}\dot{\mu}$ ) $\dot{\mu}$                           | $\dot{\mu}\dot{\mu}\dot{\mu}$            | ùd̀̀r̀̀  | ✓ |
| 4 $\mu$ : | ( $\dot{\mu}\dot{\mu}$ ) <sup>4</sup> ( $\dot{\mu}\dot{\mu}$ ) | $\dot{\mu}\dot{\mu}^4\dot{\mu}\dot{\mu}$ | pɔ̀ʒàʔjì | ✓ |

As seen, downstep can only occur if there are two adjacent L-toned bimoraic feet, i.e. if there are at least four morae. The four-mora threshold is thus explained with a simple foot-based analysis. We will see in the next section that this analysis fails to correctly predict the conditions of application of another tonal process at work in Paicî: juncture H-tone spread.

### 3.2. H-tone spread: further restrictions on foot-parsing

In certain head+complement sequences, the dependency relation between the head and complement is marked by a juncture H-tone, realized on the initial mora of the complement. The head may be a subset of determiners (followed by their nominal complement), a subset of preverbal TAM markers, all nominal and verbal derivational prefixes (e.g. causative /pà-/ , agent /à-/), the head noun of a (head-initial) genitive construction, or a transitive verb followed by its incorporated object<sup>5</sup> (Rivierre 1974). The juncture H tone is only attested after monosyllabic L-toned heads, followed by a L-toned complement, as illustrated in (10), with the MIDDLE derivational prefix /pì-/.

- (10) /rλ wâdò/ → [rλ wâdò] ‘They drink.’  
 they drink  
 /rλ pì- +H+ wâdò/ → [rλ pì-wâdò] ‘They are getting drunk.’  
 they MID- drink

Interestingly, the discrepancy between  $1 \sim 3\mu$  and  $4\mu+$  words highlighted by the downstep data in section 2.2 is confirmed by the realization of the juncture H-tone. The juncture H-tone is realized on the initial mora of the complement if it is less than  $4\mu$  long, on the first two morae, i.e. on the entire initial foot, if it is a  $4\mu+$  word, as shown in (11b).

- (11) a.  $1\mu$  /pì- +H + cò/ [pì-cò] ‘move forward’  
 $2\mu$  /pì- +H + wâdò/ [pì-wâdò] ‘get drunk’ \* [pì-(wâdò)]  
 $3\mu$  /pì- +H + tλmλrî/ [pì-tλmλrî] ‘give birth’ \* [pì-(tλmλ)rî]  
 b.  $4\mu$  /pì- +H + nλjâirî/ [pì-(nλjâ)irî] ‘to curse’ \* [pì-nλjâirî]  
 /pì- +H + tɔ̀wârî/ [pì-(tɔ̀)wârî] ‘reimburse’ \* [pì-tɔ̀wârî]

As seen above, when the juncture H-tone is realized on a mora belonging to a bimoraic foot, it spreads to the entire foot: /Cv̂+(LL)/→\*Cv̂+(HL)→[Cv̂+(HH)]. The realization of the juncture H-tone can thus be used as a diagnosis for foot parsing.

The data in (11a) clearly show that words that are made up of less than four morae are not parsed into feet. Foot parsing thus appears to be dependent on the possibility to parse at least two adjacent bimoraic feet, which the foot-based analysis proposed in (9) above fails to account for, since it predicts that the initial two morae in  $2\mu$  and  $3\mu$  L-toned words should be parsed into a binary foot, as shown in (12).

- (12)
- |              | Prediction | Facts |         |   |
|--------------|------------|-------|---------|---|
| H + $1\mu$ : | ́          | ́     | có      | ✓ |
| H + $2\mu$ : | (́́)       | ́́    | wâdò    | * |
| H + $3\mu$ : | (́́)̀      | ́́̀   | tλmλrî  | * |
| H + $4\mu$ : | (́́)(̀̀)   | ́́̀̀  | tɔ̀wârî | ✓ |

We thus need to model a type of foot parsing that is conditional on the presence of at least two feet, and the challenge is to model it in a restrictive way, in order to avoid predicting unlikely patterns.

### 3.3. Colon-based analysis

This puzzle is easily solved by analyzing foot parsing in Paicî as conditioned, not directly by the presence of a following foot, but by whether it is licensed by, i.e. part of a (strictly binary) colon.

<sup>5</sup> In this case the transitive verb is always made to fit a /Cv̂/ template, whatever its original shape: /cárí/ ‘cut’ [gò cárí ì úpwárá] ‘I cut the tree’, [gò cà + úpwárá] ‘I cut trees, I am a tree-cutter.’

This straightforwardly explains the  $4\mu$  minimum condition for foot parsing: if both feet and cola are strictly binary and degenerate feet and cola are not allowed, four morae are necessary to parse a colon, as schematized in (13).



This analysis has the advantage of being very restrictive: parsing can only be conditional upon the presence of at most one adjacent foot. This avoids the unwanted prediction of unlikely patterns mentioned above, such as foot parsing dependent on the presence of more than two adjacent feet in the word, or gapped dependent parsing, i.e. parsing dependent on the presence of a non-adjacent foot in the word (e.g. parse the two initial morae into a bimoraic foot only if the last two are also parsed into a foot, irrespective of the number of morae in between).

### 3.4. OT implementation

In the Optimality Theoretic (Prince and Smolensky 1993/2004) analysis proposed below, the conditional parsing is accounted for by the constraint  $*Ft \not\subset \kappa$ , penalizing any foot that is not licensed by a colon, outranking  $PARSE-\mu$ , requiring that every mora be parsed into a foot (inconsequentially violated by  $2\mu$  and  $3\mu$  words). High-ranked BIN penalizes non-binary feet and cola, enforcing strict binarity (in particular ruling out degenerate feet/cola). Downstep is accounted for by the OCP constraint  $*L_{Ft}L_{Ft}$  (two adjacent feet cannot both be L), outranking  $IDENT-register$ . Finally  $IDENT-register$  is outranked by  $IDENT-tone$ , since the repair for  $*L_{Ft}L_{Ft}$  is not  $L \rightarrow H$  dissimilation.

(14)

	/àùkðð/	$*Ft \not\subset \kappa$	BIN	ID-tone	$*L_{Ft}L_{Ft}$	$PARSE-\mu$	ID-reg
a.	àùkðð					*!***	
b.	(àù)kðð	*!				**	
c.	(àù)(kðð)	*!*					
d.	{(àù)(kðð)}				*!		
e.	{(àù)(kɔɔ)}			*!			
f.	{(àù) <sup>+</sup> (kðð)}						*

(15)

	/ùdλrî/	$*Ft \not\subset \kappa$	BIN	ID-tone	$*L_{Ft}L_{Ft}$	$PARSE-\mu$	ID-reg
a.	ùdλrî					***	
b.	(ùdλ)rî	*!				*	
c.	{(ùdλ)(rî)}		*!		*		
d.	{(ùdλ)(rî)}		*!	*			
e.	{(ùdλ) <sup>+</sup> (rî)}		*!				*

## 4. Evaluating the colon analysis and alternatives

### 4.1. Potential weaknesses of the colon analysis

The first potential weakness of the colon analysis has to do with the fact that monomorphemic words in Paicî seem to be at most 5 morae long<sup>6</sup>, which seems to be too short to allow the colon to historically arise as a helpful category. Indeed, one can easily see how the colon could arise historically in languages

<sup>6</sup> In a rapid survey of Rivierre's (1983) dictionary, I was not able to find a word longer than  $5\mu$ . If such words exist, they are very rare, and most likely not much longer than  $5\mu$ .

with very long words (as in most cases of tertiary stress mentioned in the introduction). But prosodic groupings of feet in a language that has short to-mid-length words seems unlikely.

There is evidence, however, that the prosodic word in Paicî extends beyond the lexical word, and includes all the following tonal enclitics (the term is borrowed from Rivierre 1974). Several types of tonal enclitics exist in Paicî, distinguished by their tonal behavior. The most frequent type consists in toneless functional words (mostly adverbials and prepositions), which are systematically realized with the same tone as the closest preceding toned word (the ‘tonal nucleus’, after Rivierre’s *centre tonal*).

- (16) a. tópwó = boo = nǎǎ = wee/  
 [tópwó = bóó = nǎǎ = wéé]<sub>ω</sub>  
 put down at there
- b. /gèè = mǎ = nǎǎ = nǎ = boo/  
 [{}(gèè) = {}(mǎ = nǎ)}ǎ = nǎ] = bǎò]<sub>ω</sub>  
 cross to.here at here down

Evidence that tonal enclitics form a phonological constituent with their tonal nucleus comes from the fact this constituent is the domain of application of the downstep process described in section 2. This is seen in (16b), where the colon which licenses foot parsing (and eventually the application of downstep) contains morae from both the tonal center and following enclitics. Example (16b) further shows that there can only be one downstep per prosodic word in Paicî. Indeed, (16b) is not realized with two downsteps as in (17), despite the fact that it contains enough morae to parse two cola. Only the leftmost downstep is ever realized.

- (17) \*[]{{(gèè) = {}(mǎ = nǎ)}{}(ǎ = nǎ) = {}(bǎò)}]<sub>ω</sub>

Note that the domain of application of downstep is indeed the prosodic word –defined as a tonal nucleus and all following tonal enclitics–, and not simply any sufficiently long sequence of surface L tones –which would make downstep a late rule insensitive to prosodic structure. This is shown in examples (7a, e) above, repeated in (18a, b) respectively.

- (18) a. /ǎbǎòò mǎìnǎ/ [ǎbǎòò]<sub>ω</sub> [mǎìnǎ]<sub>ω</sub> \*[]{{(ǎbǎò)}(ò mǎ)}ìnǎ]  
 person big  
 ‘big/tall person’
- b. /pǎìèè bwǎìcǎì/ [{}(pǎì)(èè)]<sub>ω</sub> [{}(bwǎì)(cǎì)]<sub>ω</sub> \*[]{{(pǎì)(èè)} bwǎìcǎì]  
 serpentine green  
 ‘green serpentine’

As seen, no downstep occurs in (18a), because [ǎbǎòò] and [mǎìnǎ] are two independent prosodic words, none of which reaches the 4μ threshold for the application of downstep. Similarly in (18b), [pǎìèè] and [bwǎìcǎì] are two independent 4μ-long prosodic words, each with their own downstep; if the sequence of eight L-toned morae they constitute were the domain of application of downstep, only the first downstep would be realized (see (16b) and (17) above).

The prosodic word can thus be quite long in Paicî. However, this does not really eliminate the weakness mentioned above. Indeed, we have seen that there is ever only one downstep per word. This means that the colon is only ever relevant for the first four morae of a prosodic word, and there is no evidence for foot or colon parsing past the fourth mora. The role of the colon is thus rather limited, and does not appear to have any relation with the potentially great length of the prosodic word.

The second potential weakness of this analysis is that the colon is needed only as a licenser: the downstep process is not directly conditioned or triggered by the colon, but by the presence of two adjacent L-toned feet. In general, the colon is neither a trigger nor a target of any phonological process. This analysis thus uneconomically adds a category in the prosodic hierarchy for a very limited gain.

Despite these two weaknesses, the colon analysis accounts for the Paicî facts in a more elegant way than the more conservative or economic alternatives considered in the next section.

#### 4.2. Alternative 1: extrametricality

An alternative analysis using only the traditional categories and mechanisms of metrical phonology is not impossible, but it is much less satisfying. For instance, if one analyzes the last two morae of every word in Paicî as extrametrical, the  $4\mu+$  minimum for foot parsing is explained:  $4\mu+$  words are the shortest words containing at least two parsable morae. Downstep can then be analyzed as occurring after the initial bimoraic foot (similarly to the first hypothesis entertained in section 3.1). This is illustrated in (19), where extrametrical constituents appear in angle brackets  $\langle \dots \rangle$ .

(19)	Prediction	Facts		
$1\mu$ :	$\langle \dot{\mu} \rangle$	$\dot{\mu}$	$pw\lambda$	✓
$2\mu$ :	$\langle \dot{\mu} \rangle \langle \dot{\mu} \rangle$	$\dot{\mu}\dot{\mu}$	$c\lambda m\dot{\imath}$	✓
$3\mu$ :	$\dot{\mu} \langle \dot{\mu} \rangle \langle \dot{\mu} \rangle$	$\dot{\mu}\dot{\mu}\dot{\mu}$	$\dot{\imath}d\lambda r\dot{\imath}$	✓
$4\mu$ :	$(\dot{\mu}\dot{\mu})^+ \langle \dot{\mu} \rangle \langle \dot{\mu} \rangle$	$\dot{\mu}\dot{\mu}^+ \dot{\mu}\dot{\mu}$	$p\lambda j\dot{\alpha}^+ j\dot{\imath}$	✓

However, this account violates the key restriction of extrametricality theory that only one single prosodic constituent at an edge is allowed to be extrametrical. Loosening this restriction drastically weakens the theory, by allowing it to predict systems where a sequence of any number of a given prosodic constituent (mora, syllable, etc.) may be extrametrical.

One possible solution to this problem is to say that it is not the last two morae, but the bimoraic foot they constitute that is extrametrical, as in (20).

(20)	Prediction	Facts		
$1\mu$ :	$\langle \dot{\mu} \rangle$	$\dot{\mu}$	$pw\lambda$	✓
$2\mu$ :	$\langle \dot{\mu}\dot{\mu} \rangle$	$\dot{\mu}\dot{\mu}$	$c\lambda m\dot{\imath}$	✓
$3\mu$ :	$\dot{\mu} \langle \dot{\mu}\dot{\mu} \rangle$	$\dot{\mu}\dot{\mu}\dot{\mu}$	$\dot{\imath}d\lambda r\dot{\imath}$	✓
$4\mu$ :	$(\dot{\mu}\dot{\mu})^+ \langle \dot{\mu}\dot{\mu} \rangle$	$\dot{\mu}\dot{\mu}^+ \dot{\mu}\dot{\mu}$	$p\lambda j\dot{\alpha}^+ j\dot{\imath}$	✓

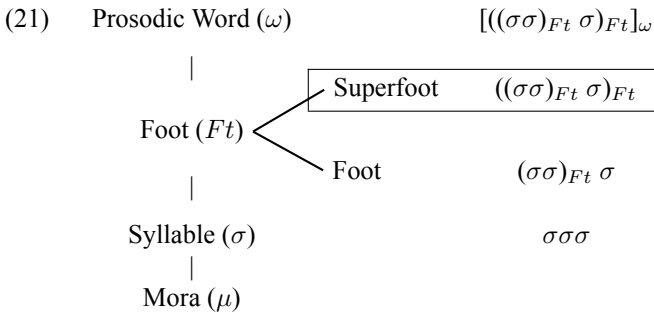
This seems to account for the Paicî data. However, one reaches a contradiction: resorting to extrametricality aims precisely at preventing the last two morae from being parsed into a foot. The extrametricality of the last foot simply means that it is not visible to parsing at the next level up the Prosodic Hierarchy, i.e. that it will not be parsed into the Prosodic Word. Extrametricality here has nothing to do with the application of downstep, which, if analyzed as foot-triggered, will occur whether the foot in question is visible to prosodic-word parsing or not.

A way out of this contradiction is to consider that the final foot is not extrametrical, but simply stipulated not to count for the application of automatic downstep. This implies the existence of two different kinds of feet in Paicî: a word-final foot, and a word-initial one, only the latter triggering downstep. As can be seen, this analysis is more stipulative, unnecessarily complicated, and explanatorily weak.

#### 4.3. Alternative 2: recursive foot parsing

Another more economic alternative to the colon analysis would be to describe the prosodic constituent responsible for downstep not as a separate category, but as a type of foot: a “superfoot”, also known as “layered foot,” derived through recursive foot parsing, as illustrated in (21) (Selkirk 1980; Prince 1980; McCarthy 1982; Shih 1986; Martínez-Paricio 2012, 2013; Martínez-Paricio & Kager 2015, 2016, a.o.).





A slight disadvantage of such a recursively derived category is that it violates the Strict Layer Hypothesis (Selkirk 1984), which says that a prosodic category of one level is exhaustively parsed into constituents of the next-lower level, and that all constituents at the lower level are of the same type. The advantage, on the other hand, is that it offers a more economic theory, with only three well-motivated categories below the prosodic word.

To account for the Paicî downstep pattern, one would need a superfoot derived through double recursion, as schematized in (22)

$$(22) \quad \dot{\mu}\dot{\mu}\dot{\mu}\dot{\mu} \rightarrow (\dot{\mu}\dot{\mu})\dot{\mu}\dot{\mu} \rightarrow ((\dot{\mu}\dot{\mu})\dot{\mu})\dot{\mu} \rightarrow (((\dot{\mu}\dot{\mu})\dot{\mu})\dot{\mu}) \Rightarrow (((\dot{\mu}\dot{\mu})^4)\dot{\mu})$$

It is unclear, however, what the descriptive generalization should be here. Downstep would have to be described as occurring between a foot and the next mora within the same superfoot, only if that mora is itself followed by another mora parsed within that same superfoot. This description sounds stipulative and falls short of offering a convincing and explanatory account of downstep in Paicî. Furthermore, there is no evidence for the intermediate iteration (i.e. the first instance of recursion). Finally, this analysis supposes unrestricted recursive parsing (or recursion arbitrarily capped at two iterations), which runs the risk of overpredicting (e.g. a system where stress is assigned to the head of a foot on the condition that it be derived through five iterations of recursive parsing, which no one would expect to be attested or possible).

## 5. Conclusion

In conclusion, I have proposed to account for the Paicî downstep and juncture H-spread by appealing to the colon. I have shown that the colon analysis has weaknesses. In particular, it uneconomically adds a new category to the Prosodic Hierarchy for relatively little gain, since the workload of the colon, used only to license foot-parsing in Paicî, is pretty light.

However, the colon analysis has crucial advantages that makes it preferable to all the alternatives presented in section 4. It is both descriptively and explanatorily more adequate. It is much more restrictive and does not overpredict: foot parsing can only ever be conditional on the presence of at most one adjacent foot. The Paicî tone system, together with the prosodic system of Iquito, thus constitute the most solid arguments to date in favor of the inclusion of the colon in the Prosodic Hierarchy.

## References

- Bensa, Alban & Jean-Claude Rivierre (1976). De quelques genres littéraires en paicî (Nouvelle-Calédonie). *Journal de la Société des Océanistes* 50:2, 31–66.
- Bensa, Alban & Jean-Claude Rivierre (1994). *Les Filles du rocher Até: contes et récits paicî*. Geuthner-ADCK, Paris-Nouméa.
- Dresher, B. Elan (1981). Metrical structure and secondary stress in Tiberian Hebrew. Chapin, C. (ed.), *Brown University Working Papers in Linguistics*, Rhode Island, RI: Brown University., vol. 4, 24–37.
- Elenbaas, Nine & René Kager (1999). Ternary rhythm and the lapse constraint. *Phonology* 16, 273–329.
- Gordon, Matthew & Ian Maddieson (2004). The phonetics of Paicî vowels. *Oceanic Linguistics* 43:2, 296–310.
- Green, Antony D. (1997). *The Prosodic Structure of Irish, Scots Gaelic and Manx*. Ph.D. dissertation, Cornell University.

- Halle, Morris & George N. Clements (1983). *Problem Book in Phonology: A Workbook for Introductory Courses in Linguistics and in Modern Phonology*. MIT Press, Cambridge, MA.
- Halle, Morris & Jean-Roger Vergnaud (1987). *An Essay on Stress*. MIT Press.
- Hammond, Michael (1987). Hungarian cola. *Phonology Yearbook* 4, 267–269.
- Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press, Chicago.
- Hyde, Brett (2002). A restrictive theory of metrical stress. *Phonology* 19, 313–359.
- Leben, William R. (1997). Tonal feet and the adaptation of english borrowings into hausa. *Studies in African Linguistics* 25, 139–154.
- Leben, William R. (2001). Tonal feet. Gut, Ulrike & Dafydd Gibbon (eds.), *Proceedings, Typology of African Prosodic Systems (Bielefeld Occasional Papers in Typology 1)*, 27–40.
- Leben, William R. (2003). Tonal feet as tonal domains. Mugane, J. (ed.), *Linguistic typology and representations of African languages*, Africa World Press, Trenton, NJ, vol. 4 of *Trends in African Linguistics*, 24–37.
- LeSourd, Philip S. (1993). *Accent and Syllable Structure in Passamaquoddy*. Garland, New York.
- Martínez-Paricio, Violeta (2012). Superfeet as recursion. *Proceedings of the 30th West Coast Conference on Formal Linguistics (WCCFL 30)*, 259–269.
- Martínez-Paricio, Violeta (2013). *An Exploration of Minimal and Maximal Feet*. Ph.D. dissertation, University of Tromsø.
- Martínez-Paricio, Violeta & René Kager (2015). The binary-to-ternary rhythmic continuum in stress typology: layered feet and non-intervention constraints. *Phonology* 32, 459–504.
- Martínez-Paricio, Violeta & René Kager (2016). Metrically conditioned pitch and layered feet in Chugach Alutiiq. *Loquens* 3:2.
- McCarthy, John J. (1982). *Prosodic Organization in Morphology*.
- Michael, Lev (2011). The interaction of tone and stress in the prosodic system of iquito (zaparoan, peru). *Amerindia* 35, 53–74.
- Pearce, Mary (2006). The interaction between metrical structure and tone in kera 23:2, 259–286.
- Pearce, Mary (2013). *The interaction of tone with voicing and foot structure*. CSLI Publications, Stanford, CA.
- Prince, Alan (1980). A metrical theory for Estonian quantity 11, 511–562.
- Prince, Alan & Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Available as ROA-537 on the Rutgers Optimality Archive, <http://roa.rutgers.edu>.
- Prince, Alan & Paul Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell.
- Rivierre, Jean-Claude (1974). Tons et segments du discours en paicî (Nouvelle-Calédonie). *Bulletin de la Société de Linguistique de Paris* 69:1, 325–340.
- Rivierre, Jean-Claude (1978). Accents, tons et inversion tonale en Nouvelle-Calédonie. *Bulletin de la Société de Linguistique de Paris* 73:1, 415–43.
- Rivierre, Jean-Claude (1983). *Dictionnaire paicî-français (Nouvelle-Calédonie)*. Société d'Études Linguistiques et Anthropologiques de France, Paris.
- Rivierre, Jean-Claude (1993). Tonogenesis in New Caledonia. *Oceanic Linguistics Special Publications* 24 Tonality in Austronesian Languages, 155–173.
- Selkirk, Elisabeth (1980). Prosodic domains in phonology: Sanskrit revisited. Aronoff, Mark, M.-L. Kean, Mark Aronoff & M.-L. Kean (eds.), *Juncture*, Anma Libri, 107–129.
- Selkirk, Elisabeth (1984). *Phonology and Syntax: The Relation between Sound and Structure*. MIT Press, Cambridge, MA.
- Shih, Chi-Lin (1986). *The Prosodic Domain of Tone Sandhi in Chinese*. Ph.D. dissertation, University of California, San Diego.
- Siptár, Péter & Miklós Törkenczy (2000). *The Phonology of Hungarian*. Oxford University Press, Oxford.
- Stowell, Timothy A. (1979). Stress systems of the world, unite! *MIT Working Papers in Linguistics* 1, 51–76.
- Topintzi, Nina (2016). Iquito: The prosodic colon and challenges to ot stress accounts. Heinz, Jeff, Harry van der Hulst & Rob Goedemans (eds.), *Dimensions of Phonological Stress*, Cambridge University Press, 123–167.
- Topintzi, Nina (2017). The prosodic colon in stress, tone and prosodic templates: evidence from iquito and elsewhere. *Selected papers from the 22nd International Symposium on Theoretical and Applied Linguistics (ISTAL 22)*, 466–479.

# Proceedings of the 36th West Coast Conference on Formal Linguistics

edited by Richard Stockwell, Maura O’Leary,  
Zhongshi Xu, and Z.L. Zhou

Cascadilla Proceedings Project Somerville, MA 2019

## Copyright information

Proceedings of the 36th West Coast Conference on Formal Linguistics  
© 2019 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-474-4 hardback

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the printed edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Lionnet, Florian. 2019. The Colon as a Separate Prosodic Category: Tonal Evidence from Paicî (Oceanic, New Caledonia). In *Proceedings of the 36th West Coast Conference on Formal Linguistics*, ed. Richard Stockwell et al., 250-259. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #3469.