The prosodic structure of Palu’e

MARK DONOHUE
National University of Singapore
mark@donohue.cc

Palu’e, an Austronesian language from Flores in central Indonesia, has prosodic patterns that can be explained simply in terms of existing constraints on foot parsing and foot size. The prosodic foot is more salient than the syllable in explaining the stress patterns found, due to conditions on the parsing of morphemes into separate feet.

1. Introduction

Palu’e (locally *Lu’a*) is an Austronesian language spoken on a small island of the same name off the north coast of Flores, in central Indonesia, three hours by boat west of Maumere. It is most closely related to Lio and Kéo, languages that are spoken on the mainland of Flores. The language has a segmentally simple phonology, but shows a complex prosodic structure that combines foot regularity with restrictions on parsing morphemes.

I shall present a prosodic phonology of Palu’e, arguing that there is a special status associated with the disyllabic foot (see Blust 1988 for a discussion of root structure in Austronesian languages). The extreme high frequency of disyllabic feet is associated with a highly-ranked constraint on foot binarity, applying in different ways to ensure that lexically monosyllabic roots are augmented, and that roots larger than two syllables in length are rare.

2. Basic segmental phonology and inventory of prosodic patterns

Palu’e has the following consonants: p t tʃ k ʔ d dʒ g ɓ̥ mb nd ŋg m n ŋ v s r l, and five vowels, i ɛ a ɔ and u (though the status of the prenasalised sequences as unit phonemes is not without some problems). The orthography used here represents them following their IPA symbols, with the following exceptions: *b* represents [ɓ̥], *c* represents [tʃ], *j* represents [dʒ]; the prenasalised stops are represent by the digraphs *mb*, *nd* and *ŋg*, and the mid vowels ɛ and ɔ are represented as *e* and *o*.[1]

Examining stress, we find that the patterns that need to be accounted for, in words of up to four syllables in length, are those shown in table 1. The categories that are relevant are: σ = syllable; σ́ = syllable with primary stress; σ̀ = syllable with secondary stress; σː = syllable with long vowel; σ̆ = syllable with a short schwa [ə̆] vowel as rime. In table 1 the line through the middle indicates that the patterns found below that line are all describable as being one of the patterns listed above the line with an extra unstressed syllable added to the right edge. They shall be discussed in section 6.

1 This transcription used here differs in some respects from the practical orthography in use on Palu’e. In the practical orthography *v* is represented by <w>, ŋ by the digraph <ng>, ŋg by <ngg>, and ʔ by <’>. Some speakers prefer, and on occasion use, <p> as a grapheme for ɓ̥.

Table 1. Phonetically attested stress patterns

	1-σ	2-σ	3-σ	4-σ
Penultimate		σ́ σ	σ σ́ σ	σ̀ σ σ́ σ
Penultimate, short unstressed σ		σ̆ σ́	σ̆ σ́ σ	σ̆ σ̀ σ́ σ
				σ̀ σ σ̆ σ́
				σ̆ σ̀ σ̆ σ̀
Long syllable			σ̀ː σ̆ σ́	
			σ̀ː σ́ σ	
Final long syllable stressed σ	σ́ː		σ̀ σ σ́ː	
		σ́ː σ	σ́ σ σ	σ σ́ σ σ
[matches a pattern from the			σ̆ σ́ σ	σ̆ σ́ σ σ
preceding column, with one extra				σ̀ː σ̆ σ́ σ
unstressed syllable]				σ̀ː σ́ σ σ
				σ̀ σ σ́ː σ

Some patterns might be predicted, but are not attested. For two and three-syllable words, these forms include *σ̀ː σ́ː and *σ̆ σ̀ σ́ː. Possible explanations for these, and other, gaps will be suggested in sections section 10.

Despite the great diversity of apparent stress types shown in table 1, the data can be plausibly reduced to a set of fairly simple regularities, as will be shown in the following sections.

3. Trochaicity

Palu'e words consist of binary feet of which the initial syllable is prominent. The final foot in the prosodic word displays primary stress, as can be seen in (1). The overwhelming majority of lexical roots are disyllabic, of the form (C)V(C)V, and so we shall start the exposition with these roots. All quadrisyllabic words for which adequate lexicographic investigation has been carried out are composed of two disyllabic roots, as in *obocava* below, which is composed of *obo* 'grass, creeper, runner (generic)' and *cava* 'pumpkin'. Trisyllabic words present a more complicated picture, as will be seen in section 7.

(1)	*loke*	'skin'	[ˈloke]	σ́ σ	* [loˈke]
	pana	'go'	[ˈpana]	σ́ σ	* [paˈna]
	kavala	'ribs'	[kaˈvala]	σ σ́ σ	* [ˌkavaˈla], etc.
	obocava	'pumpkin'	[ˌoɓoˈtʃava]	σ̀ σ σ́ σ	* [oˌɓotʃaˈva], etc.

There is no variation in the assignment of stress based on the different weight of different syllables; this reflects the fact that all syllables are open (with the exception of nouns marked by a third person genitive clitic – see 6 – which does not in any case affect stress), there are no lexically long vowels (though see section 4 below), and that all vowels are treated equally for the purposes of attracting stress. Sequences of two vowels in a root behave a separate syllables for the purposes of stress assignment; all possible vowel sequences are attested, except sequences of like vowels: there are no cases of **ii*, **ee*, **aa*, **oo* or **uu* in a lexical root, presumably due to the operation of a constraint forbidding long vowels (*$V_\alpha V_\alpha$).[2,3]

4. Monosyllables

The first complication to the description above involves lexically monosyllabic roots. Monosyllabic roots of the shape (C)V show universally long vowels, despite length not being a feature of the phonological system of the language. I assume that the length is present to satisfy a constraint on foot binarity, requiring that each foot exhibit two mora; this implies another constraint, f(oo)tBin(arity), which outranks the constraint against like vowels, as seen in (2). Examples of monosyllabic roots are shown in (3); as stated above, there are no roots of the form (C)V.

(2) ftBin » *$V_\alpha V_\alpha$

(3)	*pu*	'grandparent, grandchild'	['puː]	* ['pu]
	i	'flesh, meat'	['iː]	* ['i]
	bo	'cliff'	['ᵐboː]	* ['ᵐbo]
	te	'hear'	['teː]	* ['te]
	va	'forehead'	['vaː]	* ['va]

Other than these monosyllabic roots there are no long vowels in the language. Reasons for assuming that the constraint responsible for lengthening the vowels in monosyllables is a foot-size constraint, and not a constraint on minimal word size, will become apparent in section 7.

5. Final stress and epenthesis

Palu'e does not allow complex onsets or geminate consonants (nor sequences of the same vowel). The ban on clusters is relaxed only in the case of the complex prenasalised consonants, which must be syllabified over two syllables.[4] The lack of clusters is relevant to the analysis proposed for the following data concerning an apparent exception to regular penultimate stress, as well as the exceptional schwas in the language. Compare the stress patterns in the following forms to those in (1).

---

[2] It is possible that the constraint against adjacent like vowels in the same root, *$V_\alpha V_\alpha$, is satisfied by glottal stop insertion, and that many phonetic $V_\alpha ʔ V_\alpha$ sequences represent lexical $V_\alpha V_\alpha$ sequences. Examining roots of the form VʔV we find that approximately one third of all forms are of the shape $V_\alpha ʔ V_\alpha$, implying that there might be some truth to this hypothesis, from a diachronic perspective.

[3] There is at least one apparent exception in multimorphemic compounds, such as *tanaana* 'worm', < *tana* 'earth' and *ana* 'child'. It would be more accurate to say that Palu'e does not allow sequences of the same vowel within a morpheme, or across word boundaries.

[4] Unless they are present as the initial onset in a word, in which case they generally mark a word as being a loan. Even in this position there is at least one environment, monosyllabic roots, in which the prenasalised stops are not allowed in normal (non-onomatopoeic) lexical items. Thus *[mbiː], *[ndiː], *[ŋgiː] are not permitted lexical forms, while (for instance) [məˈᵐbu] 'hum' and [nə̈ˈdi] 'carry' are (see later this section).

(4)	*kru*	‘gecko’	[kə̆ˈru] ~ [kɨˈru]	σ σ́	* [ˈkəru], * [ˈkru]
	lba	‘wing’	[lə̆ˈɓa] ~ [lɐ̆ˈɓa]	σ σ́	* [ˈləɓa]
	vatikba	‘tomorrow’	[ˌvatikɐ̆ˈɓa]	σ̀ σ σ σ́	* [ˌvatiˈkəɓa], etc.

Here we can see that underlying consonant clusters are separated by an epenthetic vowel. (The exact quality of the vowel varies slightly depending on the height of the vowel in the following syllable; [ə] is always possible, but [ɨ] is also heard preceding a syllable with *i* or *u*, and [ɐ] is heard preceding a syllable with *a*.) These schwas appear not to be regular members of the Palu’e phonological system, for the following reasons:

- schwas never attract stress, regardless of their position in the word;
- schwas are never found in combination with other vowels, despite the otherwise unrestricted nature of V+V sequences across syllables (apart from the ban on geminates or long vowels – see (2));
- schwas are never the only vowel in a word, while other words can consist of a single quality in all vocalic positions;
- schwas are not found word finally or word initially; they are only ever heard between consonants;
- schwas are always found at the left edge of a foot, immediately preceding a stressed syllable.

These facts all indicate a non-lexical origin for the schwas as epenthetic vowels, and a general constraint against complex onsets (or possibly simply clusters; I note this as *CC for reasons of space) which works to introduce non-lexical vowels. These non-lexical vowels are not part of the input of the prosodic word, and so are not eligible to receive stress, through an input-output rule specifying that the stressed syllable must be a mora in the lexical root: *IO-σ́μ. This constraint must be ranked above the constraint that calls for trochaic stress assignment, as in (5), to allow for the presence of words with an initial epenthetic vowel that have final stress. These epenthetic vowels are sufficient to satisfy the bimoraicity requirement. (6) shows the interaction of constrains for *kru* ‘gecko’.

(5) *IO-σ́μ » TROCH(AIC)

(6)

	kru	*CC	ft-Bin	*IO-σ́μ	TROCH
	krú	*!	*		
	krúː	*!			
	kə́ru			*!	
	kə́ruː		*!	*	
☞	kərú				*
	kərúː		*!		*

We can see that the new constraint, *CC, has two functions: firstly, it breaks up illicit consonant clusters. Secondly, by inserting epenthetic vowels where there are lexically none, it can produce forms that do not violate the foot binarity constraint, and so bleeds the application of the lengthening process seen earlier. Recognising these schwas as epenthetic allows us to account for the otherwise irregular final stress patterns found on words such as *kru*, *lba* and *vatikba*.

6. Antepenultimate stress and genitive enclitics

Exceptions to these patterns are found when clitics are added to words. Clitics form part of the same word, phonologically, as their host, but are not in the domain of stress assignment. This means that a word composed of a root and a clitic will show antepenultimate, rather than penultimate, stress.[5]

(7)	*loke-gu*	'my skin'	[ˈlokegu]	σ́ σ σ	* [loˈkegu]
	pana-ʔu	'already gone'	[ˈpanaʔu]	σ́ σ σ	* [paˈnaʔu]
	kavala-mo	'your ribs'	[kaˈvalamo]	σ σ́ σ σ	* [ˌkavaˈlamo]
	obocava-de	'our pumpkin'	[ˌoˀɓoˈtʃavade]	σ̀ σ σ́ σ σ	* [oˌˀɓotʃaˈvade]

Clearly the prosodic structure of these words contains extrametrical syllables; the genitive and aspectual clitics are not part of the material which is lexically specified for the word, and so are not parsed into the foot-structure that is built around the lexical items. At the same time they are unable to build their own metrical feet. Other evidence that the clitics are phonologically aberrant can easily be found:

- the most common vowel in the language, *a*, is not found in any clitics;
- the 1GEN and 12GEN clitics begin with *g* and *d*, consonants not otherwise found root-initially;[6]
- the 3GEN clitic *-n* contains no vowel, a pattern otherwise unattested for morphemes;
- the round vowels of the 1GEN and 2GEN clitics *-gu* and *-mo* can be resyllabified as [w] glides serving as onsets on a following V-initial word. Such vowel reduction is not found with the final vowels of lexical words.

In the case of the 2GEN clitic vowel resyllabification results in a nasal-final word, which is (other than in the case of a word inflected with the 3GEN clitic) otherwise unattested. That it is allowed is further testimony to the exceptional phonological status of the clitics. For and example of this, consider the common expression in (8).

(8) *Kau* *ɲara-mo* *ai?*
2SG name-2GEN who
'What's your name?'

This is commonly produced in fast speech as

(8)' [ˈɲaram ˈwai]

A possible reply to this might be

---

5 Words that are composed of two lexical morphemes, such as *obocava* 'pumpkin', which is made up of *obo* 'grass (generic)' + *cava* 'pumpkin', appear with both morphemes being parsed into stress-assigning feet. This will be discussed in section 8.

6 For an account of these voiced stops from an historical perspective, see Donohue (2004).

(9) *ŋara-gu* *Edo*
name-1GEN Edo
'My name is Edo.'

which surfaces as

(9)' ['ŋaraɡ̊ 'wedo] ~ ['ŋarak 'wedo]

in addition to the slow-speech ['ŋaragu 'edo]. As noted above, this sort of vowel reduction and resyllabification is not found with the final vowels of lexical words, as shown in (10) and (11).

(10) *rero* *ama-n*
friend father-3GEN
'friend's father'

(10)' ['rɛrɔ 'aman]
* ['rɛr 'waman], etc.

(11) *Kau* *mugu* *aba?*
2SG burn what
'What are you burning?'

(11)' ['kau 'mugu 'aᵐba]
* ['kau 'muk 'waᵐba], etc.

Another prominent pattern that indicates the extrametricality of the genitive clitics comes from examining monosyllabic roots, which have the structure CV or V. The addition of a genitive clitic does not satisfy the requirement for bimoraic feet, supporting the analysis of these morphemes as being extrametrical: since they are not part of the prosodic structure built by the lexical root, they are not in the domain of stress assignment. Since the clitics are not lexical roots, they are not capable of building their own prosodic feet and so cannot receive stress.[7]

(12) *pu-gu* 'my grandparent' ['puːgu] * ['pugu]
i-de 'our flesh, meat' ['iːde] * ['ide]
u-mo 'your wake (of a boat)' ['uːmo] * ['umo]
va-gu 'my forehead' ['vaːgu] * ['vagu]

I assume that the fact that the genitive clitic is not part of a prosodic foot means that its presence cannot satisfy the foot binarity constraint imposed on the lexical material, and so the non-lexical length of the vowel of the root is preserved.[8]

The data gathered from examining words with clitics shows us that it is possible to have syllables which are not parsed into feet themselves. This is an undesirable outcome, but is preferable to parsing non-lexical material into feet, a constraint that we have already seen in the non-parsing of epenthetic vowels in words with underlying consonant clusters. I shall assume a

---

7 The third person genitive clitic *-n*, not having a vowel, will also fail to satisfy moraicity: ['puːn] 'her/his/their grandparent', *['pun]; similarly ['iːn], *['in], ['uːn], *['un], and ['voːn], *['von], The fact that this non-vocalic clitic exists shows that there is no constraint against codas in Palu'e, simply a general lack of such roots. Further support for the lack of a NOCODA constraint can be found in the syllabification of prenasalised stops and the incipient 1SG subject proclitic *ak-*.

8 The preservation of the length of the monosyllables can also be accounted for by the assumption, following Borowsky and Harvey (1997), that the phonetic identity of the root remains constant through a Vowel Weight-Identity constraint.

parsing constraint that requires syllables to be assigned to feet, which is outranked by an input-output constraint that restricts metrical structures to syllables that are part of the lexical input of a word.

(13) *IO-σμ » PARSE(σ,ft)

The importance of the relative ranking of parsing constraints in the metrical structure of Palu'e becomes apparent when we examine trisyllabic words. These words also show that we are dealing with foot binarity, and not a minimal word constraint.

7. Trisyllabic patterns

The trisyllabic root that we have seen so far, *kavala* 'ribs', conforms nicely to the patterns that we have described. As it stands, our constraints would predict that *kavala* should be parsed into two separate feet, one of which would, by foot-binarity, show a lengthened vowel: *[ˌkaːˈvala] or *[ˌkavaˈlaː]. These are not acceptable pronunciations of the word, but are patterns which are attested with other trisyllables. The words in (14) show some of the other patterns that we need to account for; *puloga* and *erombu* show the prosodic patterns that are not allowed for *kavala*.

(14)	*kavala*	'ribs'	[kaˈvala]	σ σ́ σ	* [ˌkavaˈla], etc.
	puloga	'whale'	[ˌpuːˈloga]	σ̀ː σ́ σ	* [puˈloga]
	erombu	'bee'	[ˌeromˈbuː]	σ̀ σ σ́ː	* [eˈrombu]

How can we explain these different prosodic patterns? The answer lies in the morphological structure of the different types of words. While *kavala* does appear to be genuinely monomorphemic, *puloga* is made up of two roots, *pu* 'grandparent/child' and *loga* 'spout hole (of a whale)'. Similarly *erombu* is composed of *ero* 'bee' and *mbu* 'buzzing noise made by insect or small bird'.[9] The stress patterns can be accounted for by a constraint that requires each lexical morpheme to support its own foot structure, and be modelled with a single metrical foot: Parse(M,ft).[10] Since it is, however, possible for monosyllabic and trisyllabic roots to occur, we must rank this constraint below ftBin to ensure that lengthening applies where applicable, and that ternary feet are not generated. Within each foot, the conditions on cluster separation, foot binarity and trochaic stress assignment apply. The application of this constraint is shown in (16) - (18). With a monomorphemic root the general trochaic stress principle is the one that decides the correct parsing of the word, with various unlikely parsing being ruled out by the other constraints.

(15) ftBin » PARSE(M,ft)

Sample tableaux illustrating the application of these constraints to correctly predict the forms found in (14) are shown below in (16) - (18). In (16) the trisyllabic form *kavala* 'ribs' appears with trochaic parsing, and stress on the penultimate syllable and an unparsed initial syllable. Note that we require the morpheme-parsing constraint to outrank the syllable-parsing constraint.

---

9 See footnote 4.

10 This constraint, combined with the ftBin constraint, explains the overwhelming predominance of disyllabic roots in the language.

(16)

	kavala	ft-Bin	PARSE(M,ft)	PARSE(σ,ft)
☞	ka(vála)			*
	(ka)(vála)	*!	*	
	(kà)(và)(lá)	***!	**	
	(kà.va)(lá)	*!	*	
	(ka.vá.la)	*!		
	ka.va.la		*!	***

When we examine a bimorphemic root such as *puloga*, a very different interaction can be seen. Here, because there are two morphemes each of which requires their own foot, the first syllable must be parsed into a foot of its own; vowel lengthening means that the ftBin constraint is not violated (lengthening of monosyllabic feet was not shown in (16) when dealing with *kavala*, though it would not have affected the outcome).

(17)

	puloga	ft-Bin	PARSE(M,ft)	PARSE(σ,ft)
	pu(lóga)		*!	*
	(pù)(lóga)	*!		
☞	(pùː)(lóga)			
	(pù)(lò)(gá)	***!	*	
	(pùː)(lòː)(gáː)		*!	
	(pù.lo)(gá)	*!	**	
	(pù.lo)(gáː)		**!	
	(pu.ló.ga)	*!	*	
	pu.lo.ga		**!	***

Similarly with words such as *erombu*, in which the first two syllables form a morpheme compounded with the final syllable, the optimal candidate differs from that selected for *kavala*. Note that the optimal candidate, [ˌerom'buː], is in fact in violation of the low-ranking TROCH constraint due to the stress placement on the final syllable (it might be argued that this does place the stress on the *pen*ultimate mora, but since there is otherwise no evidence for morae functioning as a linguistically relevant unit in Palu’e, and since it does not affect any outcomes, I shall not address this issue), but still emerges as the most optimal candidate.

(18)

	erombu	ft-Bin	PARSE(M,ft)	PARSE(σ,ft)	TROCH
	e(rómbu)		**!	*	
	(è)(rómbu)	*!	**		
	(èː)(rómbu)		**!		
	(è)(ròm)(bú)	***!	*		*
	(èː)(ròːm)(búː)		*!		*
	(è.rom)(bú)	*!			*
☞	(è.rom)(búː)				*
	(e.róm.bu)	*!	*		
	e.rom.bu		**!	***	*

The other trisyllabic patterns involve words with lexical CC sequences, such as the following:

(19)	*tuŋ̩go*	‘praying mantis’	[ˌtuːtə̆ŋˈgo]	σ̀ː σ σ́	* [ˈtutə̆ŋgo], etc.
	nbeto	‘hide’	[nə̆ˈᵐbeto]	σ̆ σ́ σ	* [nəˈᵐbeto]

Accounting for these patterns simply requires us to employ the *CC constraint, and to posit a bimorphemic structure for *tuŋ̩go*, *tu* ‘knee’ + *ŋ̩go* ‘bend/hook a limb’. (15) does not show the parsing constraints, nor the various possible non-optimal candidates with which they would interact (such as *[(ˌtùːt)(ŋ̩góː)]), to save space.

(20)

	tuŋ̩go	*CC	ft-Bin	*IO-σ́μ	TROCH
	(tú.ŋ̩go)	*!			
	(tù)(ŋ̩gó)	*!	**		*
	(tùː)(ŋ̩gó)	*!	*		*
	(tù)(ŋ̩góː)	*!	*		*
	(tùː)(ŋ̩góː)	*!			*
	(tù)(təŋgó)		*!		*
☞	(tùː)(təŋgó)				*
	(tù)(təŋgóː)		**!		*
	(tùː)(təŋgóː)		**!		*
	(tù)(tə́ŋgo)		*!	*	
	(tùː)(tə́ŋgo)			*!	
	(tù)(tə́ŋgoː)			*!	
	(tùː)(tə́ŋgoː)		*!	*	
	tu.ŋ̩go	*!			

Note that these analyses presented above crucially rely on there being a difference in the morphological make-up of trisyllabic words: this then ‘activates’ the ft-Bin constraint, critical in selecting the correct candidates. Despite this, a small minority are monomorphemic. The tableaux for *nbeto* ‘hide’, which is monomorphemic but does include an illicit initial consonant cluster, is shown in (21).

(21)

	nɓeto	*CC	ft-Bin	PARSE(M,ft)	PARSE(σ,ft)	*IO-σ́μ	TROCH
	(nɓé.to)	*!					
	(nɓè)(tó)	*!	**	*			*
	(nɓèː)(tó)	*!	*	*			*
	(nɓè)(tóː)	*!	*	*			*
	(nɓèː)(tóː)	*!		*			*
☞	(nǎɓé.to)						
	(nàɓe)(tó)			*!		*	*
	(nǎɓè)(tó)		*!	*			*
	(nǎɓèː)(tó)		*!	*			*
	(nǎɓè)(tóː)			*!			*
	(nǎɓèː)(tóː)		*!	*			*
	(nǎɓé)(to)			*!			
	(nǎɓéː)(to)		*!	*			
	(nǎɓé)(toː)			*!			
	(nǎɓéː)(toː)		*!	*			
	nɓe.to	*!		*	**		

Supporting evidence for the analysis of the differences in foot structure of words greater than one syllable in length being based on morphemic structure can be found by examining the positions in which epenthetic vowels may occur. These non-lexical vowels are only found at the left edge of a metrical foot. The different attested structures for monomorphemic metrical feet are shown in (22).

	Root	example		
(22)	(V)	*i*	'flesh'	['iː]
	(CV)	*tu*	'knee'	['tuː]
	(VV)	*iu*	'shark'	['i(j)u]
	(VCV)	*iku*	'elbow'	['iku]
	(CVV)	*vae*	'water'	['vae]
	(CVCV)	*ŋgeʔu*	'baby'	['ŋgeʔu]
	(CCV)	*ʔra*	'sun'	[ʔə̆'ra]
	(CCVV)	*kcoa*	'cockroach'[11]	[kə̆'tʃo(w)a]
	(CCVCV)	*btiko*	'firefly'	[ᵐɓə̆'tiko]
	*(VCCV)	V.CCV		* ['VCəCV]
	*(CVCCV)	CV.CCV		* ['CVCəCV]

From this data we can see that we would be led to assign *tuŋgo* 'praying mantis' to two separate morphemes, even if we did not know any etymologies for *tu* and *tŋgo*. Assuming a single-foot CVCCV skeleton would not only violate the observed patterns in the language, but would not account for the lengthening of the first syllable to [tuː] (recall that *ŋg* is a single unit prenasalised stop phoneme, and so not subject to epenthesis; the prenasalised stops only

[11] This is a loanword from Indonesian, but conforms completely to the Palu'e phonological system, as do other loan words. For instance, Indonesian *dosa* 'sins' has been borrowed as *ndosa*, enforcing the ban on root-initial *d*.

contrast with sequences of nasal + stop in onomatapoeic words, however, making this analysis less certain than could be desired).

8. Pre-stem bound morphology

While there is no regular system of agreement on verbs in Palu'e, there is an emergent 1SG nominative proclitic, *ak*-. This morpheme can (but need not) attach to any verb to indicate a 1SG subject, and its use precludes the appearance of the free pronoun *aku* in the same clause Examples are shown in (23).

(23)	*te*	'hear'	[ˈteː]	σ́ː
	ak-te	'I hear.'	[ˈakte]	σ́ σ
	pana	'go'	[ˈpana]	σ́ σ
	ak-pana	'I go.'	[akˈpana]	σ σ́ σ
	alu	'pestle'	[ˈalu]	σ́ σ
	ak-alu	'I pound with a pestle.'	[aˈkalu]	σ σ́ σ
	tetaʼi	'defecate'	[teˈtaʔi]	σ σ́ σ
	ak-tetaʼi	'I defecate.'	[ˌakteˈtaʔi]	σ̀ σ σ́ σ

The first point we can note is that, being proclitics rather than enclitics, the penultimate stress assignment is not disrupted by presence or absence of the clitic. The location of the stressed syllable is constant whether the agreement clitic is used or not (though, of course, a perfective enclitic may be added to an inflected verb, which would disrupt the stress: *ak-te-ʼu* [ˈakteʔu] 'I have heard', for instance). Secondly, the proclitic is parsed into a metrical foot, unlike the enclitics. This can be seen in the secondary stress it acquires in a word like *ak-tetaʼi* above, or the primary stress in *ak-te*. This means that the proclitic is not affected by the constraint against non-lexical material being parsed into a foot, while the enclitic material is. This is independently justified, based on the distributional behaviour of the proclitic: as mentioned above, it is not obligatory on verbs, and may not cooccur in the same clause as an independent pronoun. Examine the following partial paradigm for a clause using the verb *pana* 'go', showing its appearance with a singular subject of different persons. The clause is grammatical with either a free pronoun, or a proclitic serving to mark the subject, and is ungrammatical with either no pronominal marking, or both types of marking simultaneously.

1SG subject

(24) a. *Aku* *pana.*
1SG go
'I went.'

b. *Ak-pana.*
1SG.NOM-go

c. * *aku* *ak-pana*
1SG 1SG.NOM-go

d. * *pana*
go

2SG subject

(25) a. *Kau* *pana.*
2SG go
'You went.'

b. * *pana*
go

3SG subject

(26) a. *Ia* *pana*. b. * *pana*
3SG go go
'He/She/It went.'

Clearly whatever lexical material is included in a free pronoun is also present in the 1SG proclitic; it is completely justified to treat the *ak*- clitic as having the same foot-forming privileges as other lexical material, since it shares the same amount of lexical specification (specifically, is has the same level of pronominal information as the free pronoun *aku*).

Similar behaviour is found with the proclitic versions of oblique case markers. The alternations in (27) - (28) show that when *noʔo* is used as a clitic it forms a single metrical word with the lexical root, and is assigned stress as if part of a single stress domain.

(27)	*noʔo*	'with, be'	[ˈnoʔo]	σ́ σ
	noʔo ia	'with her/him'	[ˈnoʔo ˈia]	σ́ σ σ́ σ
	no(ʔ)ja	'with her/him'	[ˈno(ʔ)ja]	σ́ σ
(28)	*noʔo kau*	'with you'	[ˈnoʔo ˈkau]	σ́ σ σ́ σ
	no(ʔ)kau	'with you'	[no(ʔ)ˈkau]	σ σ́ σ

One other pre-stem formative is found in Palu'e, the non-productive prefix *te*-, found only in the word *tetaʔi* 'defecate' (see above), clearly derived from *taʔi* 'faeces'. The word *tetaʔi* behaves as a monomorphemic trisyllabic root, showing the same prosodic patterns as are found for words like *kandera* 'chair', *kavala* 'ribs', *kululu* 'mouse (sp.)', and *taŋala* 'mirror'.

9. Quadrisyllabic words

The analysis of quadrisyllabic words follows from the treatment of trisyllables, with the exception that there are not any native monomorphemic quadrisyllabic roots in the language. This means that almost all quadrisyllables must be parsed into two or more feet. Overwhelmingly it is two feet, reflecting the preference for disyllabic roots in general, and compounding as a productive process in the language. Occasionally a quadrisyllabic root is found which cannot be analysed as being a compound made up of two independent lexical roots. In these cases the majority involve some kind of reduplication (sometimes with modification). Examples of simple reduplication are shown in (29). Examples of modified reduplication are found in (30), with modification of a consonant, one or both vowels, or a complex (but well-attested) *Coe* reduplicant pattern. (31) shows what might prove to be genuine quadrisyllabic roots, though *cakacora* is conceivably a (heavily) modified reduplication, after the pattern of *valuvoe* and *ŋgiliŋgoe*, and a ruthless etymologist could find *lava* 'long' and *luja* 'part of a loom that tightens threads', but they are unlikely to contribute to the meaning of 'sea tern'.

Clear reduplication

(29)	*teitei*	'later'	[ˌteiˈtei]	σ̀ σ σ́ σ

	anaana	'red ant'	[ˌanaˈana]	σ̀ σ σ́ σ
	kajakaja	'sandals, thongs'	[ˌkadʒaˈkadʒa]	σ̀ σ σ́ σ
	sokosoko	'hat'	[ˌsokoˈsoko]	σ̀ σ σ́ σ
	baːobaːo	'slow(ly)'	[ˌˀɓaʔoˈˀɓaʔo]	σ̀ σ σ́ σ
	basabasa	'dolphin'	[ˌˀɓasaˈˀɓasa]	σ̀ σ σ́ σ
	Modified reduplication			
(30)	*kererere*[12]	'lizard'	[ˌkereˈrere]	σ̀ σ σ́ σ
	tgolgo	'tickle'	[təˌgoləˈgo]	σ σ̀ σ σ́
	kiokau	'noisy'	[ˌkioˈkau]	σ̀ σ σ́ σ
	ŋguruŋgero	'crab (sp.)'	[ˌŋguruŋˈgero]	σ̀ σ σ́ σ
	kidokada	'whirlpooly'	[ˌkidoˈkada]	σ̀ σ σ́ σ
	valuvoe	'twirl'	[ˌvaluˈvoe]	σ̀ σ σ́ σ
	ŋgiliŋgoe	'whirl'	[ˌŋgiliŋˈgoe]	σ̀ σ σ́ σ
	No obvious reduplication source			
(31)	*cakacora*	'centipede (sp.)'	[ˌtʃakaˈtʃora]	σ̀ σ σ́ σ
	mbabambou	'centipede'	[ˌmbaˀɓamˈbou]	σ̀ σ σ́ σ
	karapau	'buffalo'	[ˌkaraˈpau]	σ̀ σ σ́ σ
	lavaluja	'sea tern'	[ˌlavaˈludʒa]	σ̀ σ σ́ σ
	bolovua	'wasp'	[ˌˀɓoloˈvua]	σ̀ σ σ́ σ
	boimbsi	'cicada'	[ˌˀɓoimbəˈsi]	σ̀ σ σ σ́
	salavdo	'fish (sp.)'	[ˌsalavəˈdo]	σ̀ σ σ σ́
	vatikba	'tomorrow'	[ˌvatikəˀɓa]	σ̀ σ σ σ́
	lbomaŋga	'sweets'	[ləˌˀɓoˈmaŋga]	σ σ̀ σ́ σ

With the not obviously bimorphemic words in (25) stress is clearly assigned to two different feet, just as with the more obviously identifiable forms in (23) and (24). Either there is a constraint operating that blocks the appearance of two unfooted adjacent syllables (thus disallowing form such as *[σ σ (σ́ σ)]), or else the words have been folk-etymologised, reanalysed into two bound morphemes, With 'buffalo', for instance, we might imagine a splitting into *kara-* (also attested in the word for spider, *karamboro*) and *-pau* (possible folk-etymologised by analogy with *bau* 'mango'; no other *p.a.u* sequences are found in Palu'e). An argument against this folk-etymologising is the fact that *lavaluja* 'sea tern' is not considered to be related to *lava* 'long' and *luja* 'part of loom', despite the phonological plausibility.

In sum, quadrisyllabic roots do not require any new morphemes, but provide further evidence that ftBin is ranked higher than PARSE(M,ft).

10. Summary of constraints and phonological implications

We have seen evidence that the following constraints are relevant for a description of stress in Palu'e:

---

12 This word is sometimes heard as [ˌkerːeˈre].

Contraints bearing on stress assignment in Palu'e

(32)

TROCH	feet are parsed into trochaic units.
$*V_\alpha V_\alpha$	sequences of like vowels are disallowed.
ftBin	prosodic words are parsed into binary feet.
*CC	'clusters are not allowed' (with complications and conditions).
*IO-σ́μ	a stressed mora in the output must correspond to a phonological unit that was present in the input.
PARSE(σ,ft)	lexical syllables in the input must be parsed into prosodic feet.
PARSE(M,ft)	each individual morpheme in the input must be parsed into a separate prosodic foot; a single morpheme may not be parsed into multiple feet.

Furthermore, we have evidence for the rankings of these constraints as shown in (33) (this has been distilled from (2), (5), (13), (15), (18) and (21) in the preceding sections).

(33) *CC » ftBin » PARSE(M,ft) » PARSE(σ,ft) » *IO-σ́μ » $*V_\alpha V_\alpha$ » TROCH

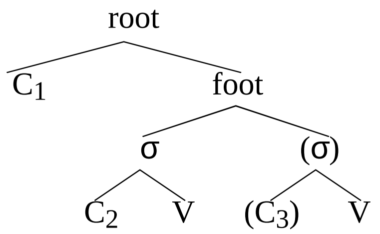
We should note that the constraint given as *CC is in fact probably decomposable into several related constraints. As mentioned in section 5, the main function of this constraint in this exposition about stress is to favour the insertion of epenthetic vowels to break up what would otherwise be complex onsets. There are two reasons to believe that the constraint *CC is not global:

- although they do in general pattern as single unit phonemes with the same distribution as other stops, the prenasalised stops /mb, nd, ŋg/ have some peculiarities:
    - prenasalised stops do not occur as the C in CVː words;
    - prenasalised stops occur only vanishingly rarely in Cə̆CV words.
- the 1SG subject proclitic *ak-*, when combined with a consonant-initial verb, allows a sequence of the coda [k] and the onset of the following syllable, as illustrated in (23).

The gaps in the attested word patterns mentioned in section 2 can now be addressed. A discussion of these relies on a model of the bimoraic foot, which based on the evidence we have seen can be modelled as in (34).

The phonotactic structure of lexical roots in Palu'e

(34)



The vowel positions are not constrained as to combinations, the only requirement beingthat two like vowels may not be adjacent in a single morpheme (= foot). The consonants that may occur in the different positions depend partly on which consonants fill the other positions in a root, and the combinations are shown in (35).

Consonants and root-position restrictions

(35) If the root consists of … Then the possiblities for … exclude:

If the root consists of …	C_1	C_2	C_3
C_2V	–	(*d*, *j*, *g*)	
$(C_2)V(C_3)V$	–	(*d*, *j*, *g*)	(*p*, *c*)
C_1C_2V $((C_3)V)$	(*d*, *j*, *g*)	(*p*, *c*, ɻ)	(*p*, *c*)

In section 2 we noted the absence of the disyllabic pattern *σ̀ː σ́ː, and to this we could also add variants such as *σ σ́, *σ́ σ̆̈, and *σ̆̈ σ̆̈. The first two of these forms would be ruled out as violating of the TROCH constraint, and the last by violations of PARSE(σ,ft). For *σ́ σ̈ we note that, given the language facts described above, we would expect this word shape only if the segmental template was of the form (C)VCC. While CCV is possible, (C)VCC is not, involving as it does at least one violation of *CODA, an inviolable constraint for lexical material. The best possible condition for (C)VCC to arise in would involve a (C)VC root affixes with the third person genitive clitic *-n*, but there are no such roots.

A word with a *σ̀ː σ́ː pattern could only arise as a result of two monosyllabic roots being compounded together; two words, such as /pu/ and /tu/, would be required to be parsed in separate feet, and would each qualify for lenthening to satisfy foot binarity: [ˌpuːˈtuː]. I sugget that the highly ranked constraint calling for foot binarity could well have been responsible forthe historical reanalysis of any such compounds as monomorphemic roots. If this had indeed happened then we would have a completely normal disyllabic root with the expected stress pattern σ́ σ.

Among the trisyllabic words the pattern *σ̈ σ̀ σ́ː is unattested (along with other less likely candidates). This would dictate a CV structure of (CCV)(CV), which should be possible in a word with two lexical roots, each requiring a prosodic foot. Its absence in the data must be taken as a gap, and not a genuine pattern, since (CV)(CCV) is attested (in *tu.tɻ̥go*), and complex predicates of the right (CCV)(CV) shape, such as *m.bu ɻ̥gu* ‘hum’, are found. It would not be a great step for the two roots of *m.bu ɻ̥gu* to be reanalysed as a single word or a single root, perhaps with the same structure as *nbeto* ‘hide’ (CCVCV).

The *σ̈ σ̀ σ́ː pattern would arise from a compound of two morphemes such as /ptu/ and /ku/. *CC would require a vowel between the *p* and the *t*, incidentally satisfying foot binarity, and the same foot binarity constraint would call for vowel lengthening of the second morpheme: [pə̆̈ˌtuˈkuː].

In sum, the various, disparate, patterns seen in table 1 can all be shown to be related in a trochaic foot pattern, in a language with no contrastive vowel length. The unusual requirement that lexical roots must be individually parsed into feet, combined with a rather ‘normal’ foot binarity constraint, interact with some less surprising phonological conditions to create the range of different syllable types that are attested.

References

BLUST, ROBERT. 1988. *Austronesian Root Theory: An Essay on the Limits of Morphology*. Studies in Language Companion Series Vol 19. Amsterdam: John Benjamins

BOROWSKY, TONI, and MARK HARVEY. 1997. Vowel-length in Warray and weight-identity. *Phonology* 14: 161-175.

DONOHUE, MARK. 2004. *Syntactic category identity and the diffusion of a sound change*. MS, Department of English, Language, and Literature, National University of Singapore.