

PART I: WORD STRUCTURE AND WORD PROSODY.....	19
CHAPTER 2.....	19
ROOT, WORD AND COMPOUND STRESS	19
2.1 Simplex words.....	20
2.1.1 Analysis of simplex forms	22
2.1.2 Monosyllabic words.....	27
2.1.3 Analysis of monosyllabic words.....	29
2.1.4 Words with four or more syllables	34
2.2 Word compounds.....	36
2.2.1 Examples of WORD-compounds.....	36
2.2.2 Compounding morphology	38
2.2.3 Prosodic characteristics of WORD-compounds	40
2.2.4 Summary	47
2.3 ROOT-morphology and prosody	47
2.3.1 ROOT-level morphology	47
2.3.2 ROOT-compounds.....	50
2.3.3 Prosodic characteristics of ROOT-morphology.....	53
2.4 Phonological correlates of compounding: the distribution of apical distinctions in words.....	58
2.5 Conclusion.....	63

Part I: Word Structure and Word Prosody

Part I of the thesis provides a description and analysis of stress and morphological structure in Ngalakgan. The analysis is couched in Optimality Theory. The description and analysis provide the necessary background for the theoretical chapters which follow in Part II.

CHAPTER 2

Root, Word and Compound Stress

Compounds are prevalent in the Ngalakgan lexicon and in word production. There is a range of semantic relationships among the elements of compounds. This chapter provides an account of stress in compound words. I show that we require only two kinds of morpho-semantic relationships in order to account for stress patterns. These two kinds or 'levels' I call 'WORD-level' and 'ROOT-level'. Understanding the distinction between these levels is essential to the discussion of phenomena in other chapters: Chs 3 and 4.

I begin by outlining the basic prosodic patterns in simple words. I then show that for elements which are surface words, the metrical structure they have when independent is retained in compounds. Compounds of this kind are called WORD-compounds. WORD-compounds contrast with 'ROOT-compounds' which have the same metrical organisation as simple words. ROOT-compounds are lexicalised compounds, with largely idiosyncratic interpretations, whereas WORD-compounds are semantically compositional: the meaning of the compound is derived straightforwardly from the meanings of its constituent morphemes.

The distinction between WORD- and ROOT-level morphology in Ngalakgan is principled: it derives from observable differences in the meanings, distribution, and morpho-phonemics associated with each level. The prosodic differences between WORD- and ROOT-level compounds provide evidence that the distinction between WORD and ROOT-level is relevant to speakers, and is the basis for a regular system of correspondences between prosodic elements such as 'foot' and 'Prosodic Word' and elements of morphology such as 'affix' and 'stem'. The semantic and prosodic evidence of morphological levels in Ngalakgan provide mutually reinforcing arguments for the usefulness of the level distinction.

2.1 Simplex words

On the whole, Ngalakgan has a trochaic stress system. In simple, monomorphemic words, stress is initial. Most monomorphemic words are disyllabic or trisyllabic.

<u>Disyllables</u>				
(1)	a.	pere	[bére]	'brisket'
	b.	polo	[bóló]	'old person'
	c.	poɲi	[bóɲi]	'now'
	d.	kara	[gára]	'high up'
	e.	kaya	[gáya]	'same-sex sibling's child'
	f.	kiɬi	[gídɬi]	'young girl'
	g.	koyo	[góyo]	'freshwater crocodile'
	h.	cala	[jála]	'mouth'
	i.	cele	[jéle]	'urine'
	j.	ceɲe	[jéɲe]	'teat, breast, milk'
	k.	ciwi	[jíwi]	'liver'
	l.	colo	[jóló]	'inside meaty part'
	m.	coɲe	[jónɲe]	<i>Terminalia canescens</i>
	n.	maɲa	[máɲa]	'throat, neck'
	o.	maɭu	[máɭu]	'paperbark coolamon' (a vessel)
	p.	maɭa	[máɭa]	'centipede'
	q.	mira	[míra]	'head'
	r.	ɲaɭa	[ɲáɭa]	'uncircumcised boy'
	s.	ɲoro	[ɲóro]	'flower'
	t.	ɲuru	[ɲóru]	'catfish sp.'
	u.	ɬere	[ɬére]	'camp' (N)

v.	ɲamu	[námu] ¹	'woman's child'
w.	ɲaɬi	[náɬi]	'frog sp.'
x.	ɲimi	[ními]	'arse'
y.	ɬoro	[ɬóro]	'east'
z.	were	[wére]	'who?'; 'somebody'
aa.	wiɲi	[wíɲi]	'shortnecked freshwater tortoise sp.'
			(<i>Emydura</i> sp.)
bb.	wolo	[wólo]	'cooked, ripe'
cc.	yele	[yéle]	'animal hole'
dd.	kaɾe	[gáɾe]	'Eucalyptus sp.'
ee.	ɲili	[ɲíli]	'fog'

The examples in (1) represent all the disyllabic, open-syllabled roots which can occur as words.² All the trisyllabic open-syllabled examples are presented in (2). Again, stress is on the initial syllable. There is no stress on the final syllable.

	<u>Trisyllables</u>			
(2)	a.	picuɬu	[bíjuɬu]	'whirlwind'
	b.	puluki	[bóluɡi]	'bullock, cow'
	c.	kamala	[gámala]	'sky'
	d.	caworo	[jáworo]	'patrilineal clan'
	e.	ceraɬa	[jéraɬa]	'women's curing and related ceremonial singing'
	f.	cilara	[jílara]	'unidentified tree sp.'
	g.	malaɬi	[málaɬi]	'wedge-tailed eagle'
	h.	munana	[mónana]	'European'
	i.	muwaɬa	[múwaɬa]	'canoe'
	j.	ɲuliri	[ɲúliri]	'black duck sp.'
	k.	walama	[wálama]	'face'
	l.	waɬiya	[wáɬiya]	'multiparous woman'
	m.	welene	[wélene]	'young girl'; 'female'
	n.	yinoro	[yínoro]	'all around'
	o.	makara	[mágara]	'white clay'
	p.	wamara	[wámara]	'freshwater tortoise sp.'
	q.	ciwuru	[jíwuru]	'wild cassava'
	r.	palara	[bálara]	'string'
	s.	komo[o]	[gómo[o]	'white crane'

¹Utterance-initially, only apico-alveolars occur. Following a vowel, word-initial apicals are post-alveolar; hence the distinction is neutralised in this position. I assume that apicals are underlyingly post-alveolar, since constraints preventing postalveolars utterance-initially and post-consonantly are more easily motivated than ones preventing alveolars post-vocalically; see §2.4 for discussion.

²I discount frozen reduplicated forms such as /ɬaɬa/ 'sugarbag (native honey)'. There are a handful of roots about which I have insufficient knowledge to include here: e.g. /moɲo/ 'animal lair'.

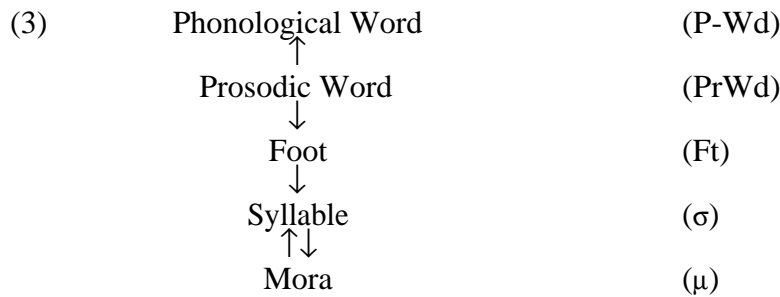
In the next section, I account for these patterns of prosodic structure.

2.1.1 Analysis of simplex forms

The prosodic form of simple words in Ngalakgan derives from constraint satisfaction and general principles of prosody. The general principles I assume are 'Prosodic Licensing' (Itô 1986), the 'Prosodic Hierarchy' and 'Strict Layering' (Selkirk 1980a, b), and 'Foot Binariness' (P&S 1993:47; cf. Prince 1980). I outline these briefly here, before moving on to an analysis of the metrical structure.

All morphological forms are subject to Prosodic Licensing. This is the proposal that all lexical words require some kind of prosodic structure in order to be pronounced (e.g. Selkirk 1980, Itô 1986). Prosodic structure is organised into a hierarchy of prosodic constituents: the Prosodic Hierarchy, as in (3). The form of the Prosodic Hierarchy in (3) is different from that proposed by Selkirk (1980), and assumed by M&P (1993a, 1993b). According to these authors, there is just one superordinate constituent, called 'Phonological Word' or 'Prosodic Word', respectively. I have assumed two superordinate constituents above the Foot. Arguments for this structure are provided in §2.2.4.

The Strict Layering hypothesis (e.g. Selkirk 1984) says that a prosodic constituent at level n must immediately dominate a prosodic constituent at level $n-1$. So any PrWd constituent must dominate some Foot constituent. I assume, following M&P (1993b:5), a weaker version of Strict Layering, such that the Prosodic Hierarchy 'expresses the domination relations among the prosodic constituents, but it does not express relations of immediate domination'. The particular dominance relations I assume are represented by arrows in (3).



This version of the Prosodic Hierarchy follows M&P (1993b) in all respects except the additional superordinate constituent P-Wd. The dominance relations expressed by the arrows are that any PrWd must be dominated by some P-Wd and must dominate some Foot, any Foot must dominate some syllable, and any mora must be dominated by some syllable and vice versa. Other relationships (which are not ruled out by the dominance relations in (3)) are also possible: P-Wd can dominate Feet and syllables directly, bypassing PrWd. Some version of the Prosodic Hierarchy, and of the Strict Layering principle, are held to be universal elements of GEN, they are not constraints which can be violated.³

Prosodic licensing in OT takes the form of a general constraint 'MCAT ≈ PRWD' (M&P 1993a:139; cf. P & S 1993:43). This constraint makes sure that words *qua* lexical entries (strings of segments) are realised as words *qua* prosodic constituents.

- (4) MCAT ≈ PRWD 'A member of the morphological category 'MCat' is a Prosodic Word' where MCat = Root, Stem, Lexical Word

³It is generally assumed that Strict Layering allows both recursion and domination of non-immediate constituents at certain points. The topmost constituent (whether PrWd or P-Wd) is assumed to allow embedded recursion, such that complex words which are PrWds can contain smaller 'words' (i.e. stems) which are also PrWds (Nespor and Vogel 1986, Rubach and Booij 1990). The syllable node must always dominate some mora, but not everything it dominates must in turn be dominated by a mora. The syllable node dominates onset directly in the representation assumed by e.g. Hayes (1989), M&P (1993ab). M&P (1993a:146) distinguish PrWd from what they call 'Word' level phonology in Axininca Campa. The latter is the domain of stress, just as P-Wd is in Ngalakgan.

For the moment, it is sufficient to regard the 'MCat' referred to in the constraint as a 'lexeme', that is a 'lexical word' in M&P's terminology. The question of the identity of MCat in this constraint as it pertains to Ngalakgan is addressed more carefully in following sections.

Metrical feet are binary, as encoded in the constraint FTBIN presented in (5). FTBIN rules out monomoraic, trisyllabic, and 'unbounded' feet (Halle and Vergnaud 1987).

- | | | |
|-----|-----------------------|---|
| (5) | Foot Binarity (FTBIN) | Feet are binary at some level of analysis (mora μ , syllable σ) |
|-----|-----------------------|---|

From the Prosodic Hierarchy, Strict Layering, and FTBIN, it then follows that PrWds are minimally bimoraic. And since lexemes are required to be PrWds, by $MCAT \approx PRWD$, then every lexeme must also be bimoraic at the surface (M&P 1993a, P&S 1993).

$MCAT \approx PRWD$ does not say anything about the form of feet, or their association with particular edges of words or morphemes (foot 'directionality'). The form of feet is decided by constraints given the general term 'Foot Form' (P&S 1993:53, 63).⁴

- | | | |
|-----|------------------------------|------------------------|
| (6) | Foot Form: Trochaic (FTFORM) | 'Feet are left headed' |
|-----|------------------------------|------------------------|

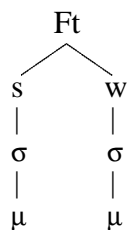
This constraint requires that any feet created by metrical processes have the form of *trochees*: left-headed feet. Given the constraint in (5), the only allowable foot in Ngalakgan is that shown in (7) (where 's' 'w' label 'strong' or 'head' and 'weak' or 'tail' branches, respectively, of the foot).

⁴In P&S (1993), FtForm is a cover term for distinct constraints determining headedness (left/right) called 'RhType=T' (trochaic rhythm) and 'RhType=I' (Iambic Rhythm).



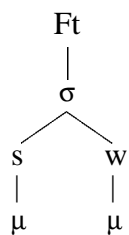
Trochees are *bimoraic* in Ngalakgan. The moras can be associated to each of two syllable peaks, giving disyllabic feet as in (8).

(8) Disyllabic foot



Otherwise both moras can be associated to the same syllable: where one mora is associated to a peak and the other to a coda, or both moras are associated to the peak, giving monosyllabic, 'heavy' feet as in (9).

(9) Monosyllabic heavy foot



The definition of syllable weight in Ngalakgan is a complex issue, and is discussed in Ch 5. For the purposes of this chapter and the next, the only heavy syllables considered will be those where the vowel is long.

The fact that stress is initial in all of the disyllabic and trisyllabic words above is determined by the following constraint:

- (10) ALIGNL(PRWD, FT): 'Align the left edge of every Prosodic Word with the left edge of some foot.' (Assign a violation mark for every syllable separating a PrWd left edge from a foot left edge.)

ALIGNL encodes the observation that PrWds are always left-headed in Ngalakgan, modulo the effect of heavy syllables (Ch 5).

All of these constraints are active in deriving a trisyllabic form like [jéraɖa], as shown in (11). Foot boundaries are shown with parentheses, PrWd boundaries with square brackets.

(11)

/ceraɖa/ 'women's ceremony'	FTBIN	FTFORM	MCAT ≈ PRWD	ALIGNL (PRWD, FT)
a. [(céra)ɖa]				
b. [ce(ráɖa)]				*!
c. [ce(raɖá)]		*!		*
d. [(céraɖa)]	*!			
e. [(céra)(ɖà)]	*!			
f. ceraɖa			*!	

I have only shown a few of the possible parses here. None of these constraints conflict. The constraints are therefore unranked - represented by the dotted boundary.

FTBIN and FTFORM rule out any parse which includes feet which are iambic (right-headed), as in (11c), or which are less or more than binary (11d, e).

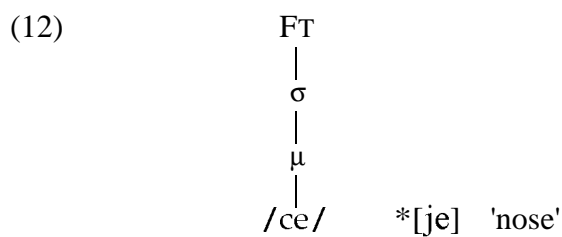
MCAT ≈ PRWD rules out the 'Null Parse' in (11f), where a morphological form is not supplied with any prosodic realisation at all. ALIGNL(PRWD, FT) prefers left-headed words, and so picks (11a) over the misaligned (11b).

All simple polysyllabic words are derived straightforwardly with the constraints presented in this section. Now let us consider some more problematic cases.

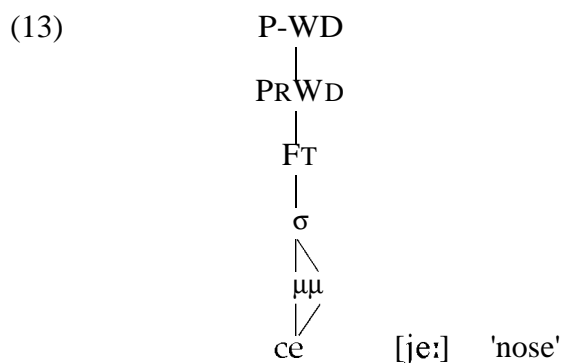
2.1.2 Monosyllabic words

All monosyllabic roots in Ngalakgan are required to correspond to a bimoraic minimum, as proposed above. A monosyllabic word consisting of an underlyingly short vowel, such as /ce/, the Ngalakgan word for 'nose', cannot satisfy the word minimality requirement: this word is not bimoraic. These words present a problem then: they cannot be parsed Faithfully without violating $\text{MCAT} \approx \text{PRWD}$.⁵

A short vowel, according to Moraic Theory, is a vowel which is associated to one mora. The word in (12) violates FTBIN because a minimum of two moras is required to form a foot.



Since the word in (12) must be parsed by PRWD, and every PRWD must dominate a foot, and every foot must be minimally bimoraic, the outcome for Ngalakgan (and many other languages) is that the smallest possible word (the *minimal word*) is bimoraic. This is shown in (13).



⁵ The reasons for adopting an underlying short vowel are discussed more fully in §2.1.3. I postpone considerations of 'Richness of the Base' (P&S 1993, §6) until then.

In (13), the word satisfies the FTBIN requirement because it is realised with a bimoraic, that is long, vowel. (14) presents examples of underlying monosyllabic, open-syllabled roots, pronounced with long vowels in surface forms.⁶

(14)	a.	/ce/	[je:]	*[je]	'nose'
	b.	/po/	[bo:]	*[bo]	'river'
	c.	/ke/	[ge:]	*[ge]	'man's child'

CV words cannot be realised with a short vowel at the surface.⁷

In monosyllabic closed roots the vowel is short. This is true regardless of whether the coda is a sonorant (15), or a stop (16).⁸

(15)	a.	/lel/	[lɛl]	*[le:l]	'aquatic plant sp.'
	b.	/wom/	[wɔm]	*[wɔ:m]	'black plum' (<i>Vitex glabrata</i>)
	c.	/pim/	[bɪm]	*[bɪ:m]	'white ochre'
	d.	/pin/	[bɪn]	*[bɪ:n]	'(1) rock; (2) money'
	e.	/pun/	[bʊn]	*[bʊ:n]	'knee'
	f.	/ker/	[gɛr]	*[gɛ:r]	'kurrajong' (<i>Brachychiton paradoxum</i>)
	g.	/kon/	[gɔn]	*[gɔ:n]	'kangaroo (generic)'
(16)	a.	/pot/	[bɔt]	*[bɔ:t]	'fly'
	b.	/kot/	[gɔt]	*[gɔ:t]	'paperbark'
	c.	/rok/	[ɾɔk]	*[ɾɔ:k]	'pandanus'
	d.	/pok/	[bɔk]	*[bɔ:k]	'small creek'
	e.	/pak/	[bak]	*[ba:k]	'pond algae'
	f.	/mic/	[mic]	*[mi:c]	'louse'
	g.	/muc/	[muc]	*[mu:c]	'Rainbow Serpent' (a mythological figure).

⁶These examples are the only CV roots which can occur unaffixed. Aside from the three nouns, there are nine inflected verb forms which are CV words. These do not occur without a prefix. Affixed CV roots are examined in Ch 3.

⁷McKay (1975:26) makes a similar observation for Rembarnga. He notes that no contrasting pairs of short vowel/long vowel are available, and that 'all words consistently pronounced with noticeably lengthened vowels are monosyllables'. He provides the following examples, of which he notes just one is a closed syllable: [ɛ:n] 'ground honeycomb'; [pi:] 'man', [mɔ:] 'knee', [mɛ:] 'vegetable food', [ɛ:] '[class of animals]', [bɛ:] 'north'. He notes further 'these words regularly have long vowels, even when inflected with suffixes'. Changes in the vowel systems of Ngalakgan and Rembarnga have led to the situation where cognates which are monosyllabic or monophthongal in one language may be disyllabic or diphthongal in the other, and vice versa. For the monophthongal Rembarnga items cited above, cf. Ngkn cognates/may/ 'vegetable food', /ɹay/ 'meat; class of meat animal', /pay/ 'north', all of which are realised with a diphthong nucleus [ai] ~ [ay]. And for a correspondence monophthong:disyllable compare Ngkn /po/ Rmba /puwa/ 'river'; Ngkn /po/ Rmba /puwa/ 'hit' (PP); Ngkn /kon/ Rmba /kuwɛn/ 'kangaroo (generic)'; Ngkn /ɲe/ Rmba /ɲiya/ 'burn' (PR).

⁸Glottal stop codas in CVC words are a separate case. These are examined in Ch 5.

The list of monosyllabic roots with sonorant codas in (15) is not exhaustive, but the examples in (16) represent all the monosyllabic nominal roots with obstruent codas in my data.⁹

The next section presents an OT analysis of the monosyllabic words.

2.1.3 Analysis of monosyllabic words

In this section I argue that the long vowels in CV words are not present in underlying forms, but are realised in satisfaction of MCAT \approx PRWD and FTBIN: a condition called 'word minimality' in M&P (1986).

Is there any evidence for the opposing view - that is, that the long-vowel form in (13) [jé:] is the underlying form? If long vowels were contrastive in some environment in Ngalakgan, then we would be forced to postulate an underlying vowel contrast in the language. But the distribution of long and short vowels is entirely predictable, and therefore I assume that there is no underlying contrast, but rather, the surface realisation of vowels is an outcome of constraint satisfaction.

Long vowels in Ngalakgan only occur in monosyllabic open roots, and in no other environment. Vowels in every other position are short. Long vowels and short vowels are in complementary distribution, therefore, they are allophones of the same underlying phoneme (Trubetzkoy 1969:49). Vowel length is non-distinctive in Ngalakgan.

The long allophone of vowels is the marked allophone. It occurs in an environment which must be specified: in monosyllabic, open roots. The short allophone occurs in the elsewhere case, every root which is not monosyllabic, or is monosyllabic but closed. The environment which conditions the distribution of the long allophone is a proper subset of that which describes the distribution of the short allophone, so I assume the latter is unmarked with respect to the former. Therefore, we can postulate the constraint ranking shown in (17) (cf. Harvey and Borowsky 1999:91).

⁹There are one or two additional words of this form in Merlan (1983) whose realisation I am unsure of: e.g. /cet/ 'stone oven'.

- (17) *[V:] >> *[V]
 'Long vowels are marked with respect to neutral (non-long) vowels.'¹⁰

This constraint ranking, like markedness constraints in general (P&S 1993:3, 187-8; M&P 1995b), is held to be universal on the grounds that:

- (18) a. There are more languages lacking in (contrastive) long vowels, than there are possessing them.
 b. Long vowels are often more restricted in their distribution than short vowels.¹¹

Thus, the restricted and predictable distribution of long vowels, and their complementarity with short vowels, leads to the conclusion that the UR of the root for 'nose' is /ce/, and for 'man's child' is /ke/, with short vowels.

A surface realisation with a long vowel in Ngalakgan violates constraints on the relation between the *input* and the *output* in OT. The relevant constraint here is IO-IDENT[mora], regulating the length of segments in the output with respect to the input (Urbanczyk 1995; McCarthy and Prince 1995b:265, McCarthy 1995):

- (19) IO-IDENT[mora]: 'Vowels in the output must have the same moraic affiliation as their correspondents in the input.'

IO-IDENT[mora] checks that segments which are underlyingly monomoraic are monomoraic in surface form, and likewise that underlyingly bimoraic segments are bimoraic at the surface, on the assumption that length is encoded in underlying moraic affiliation (e.g. Hayes 1989).¹²

¹⁰I avoid using 'short' here to characterise the non-long member of this opposition, since vowels can be distinctively short in opposition to a neutral vowel also (e.g. Ossetic: Abaev 1964).

¹¹We might also add the following observation: 'Every language which has long vowels also has short vowels, but not vice versa'. However, there is an element of circularity in this claim which I am not willing to delve into.

¹²For this chapter and the following one I will assume the only moraic elements allowed - outside of monosyllabic words - are vowels; the issue of moraic consonants is addressed in Ch 5.

The distribution of long vowels in monosyllabic words can be derived in the form of an OT tableau, as in (20). I have given both the phonemic and phonetic representations for the candidate forms. The phonemic form is surrounded by parentheses, representing a foot. The fact that the surface form - the attested output - is [bo:] and not [bo] constitutes a ranking argument for the domination of FTBIN over both IDENT and *[V:]. Even though candidate (a) violates both *[V:] and IDENT, it is considered better than candidate (b), which surfaces with a short vowel, because candidate (b) violates FTBIN. Candidate (c) does not parse the word as a foot, satisfying FTBIN (spuriously) but fatally violating MCAT \approx PRWD.

(20)

/po:/ 'river'	MCAT \approx PRWD	FTBIN	IDENT[μ]	*[V:]	*[V]
☞ a. (po _{μμ}) ~ [bo:]			*	*	
b. (po _μ) ~ [bo]		*!			*
c. po _μ ~ [bo]	*!				*

Since FTBIN is decisive in the evaluation of candidates in (20), IDENT[mora] can do no work. The same hierarchy will derive the correct output even if we assume the input has a long vowel, as shown in (21). Regardless of whether the UR has a long vowel or a short vowel, the outcome must have a long vowel, because of the dominance of MCAT \approx PRWD and FTBIN in the language.

(21)

/po:/ 'river'	MCAT \approx PRWD	FTBIN	IDENT[μ]	*[V:]	*[V]
☞ a. (po _{μμ}) ~ [bo:]				*	
b. (po _μ) ~ [bo]		*!	*		*

Are there any principles which can decide which is more likely to be the underlying form? Considerations of 'Lexicon Optimisation' will lead learners to postulate an underlying form with a *long* vowel. Lexicon Optimisation is defined by P&S (1993:192) as follows.

- (22) Lexicon Optimisation: Suppose that several different inputs I_1, I_2, \dots, I_n when parsed by a grammar G lead to corresponding outputs O_1, O_2, \dots, O_n , all of which are realised as the same phonetic form Φ - these inputs are all *phonetically equivalent* with respect to G . Now one of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labelled O_k . Then the learner should choose, as the underlying form for Φ , the input I_k .

The idea behind the principle is that, of two or more possible sources of a surface form, learners should postulate the one which incurs the minimal violation with respect to the other potential inputs, an idea they attribute to Stampe (P&S 1993:50, §9, citing Stampe 1973[1979]; cf. also Hale 1973). Unbridled Lexicon Optimisation predicts that underlying forms are the same as surface forms in any language, as far as possible.


Comparing the two tableaux (20) and (21), it can be shown that postulating an underlying form with a long vowel, as in (21), results in an evaluation which is minimally disharmonic. In (21), the attested form violates only the markedness constraint $*[V:]$. In (20), the attested form violates both $*[V:]$ and IDENT[mora]. Therefore, of the two potential inputs, that with the long vowel, in (21) induces the least significant violations, compared to (20). According to Lexicon Optimisation, the learner should postulate an underlying form with a long vowel.

Balancing Lexicon Optimisation is something P&S dub 'Minimal Redundancy': 'to the maximal extent possible, information should be excluded from the lexicon which is predictable from grammatical constraints' (1993:195; cf. Kiparsky 1982). This principle will lead learners to postulate a *short* vowel underlyingly, since the long vowel realisation is entirely predictable, based on the constraints constituting the grammar. I

suggest that Minimal Redundancy outweighs Lexicon Optimisation in this case. Long vowels are only found in environments where their form is entirely determined by constraints which are needed anyway: FTBIN and MCAT \approx PRWD. Therefore, I conclude that the underlying form has a short vowel. This results in a grammar which is more parsimonious.

In CVC words there are no long vowels at the surface. I interpret this to mean the coda is moraic. CVC words satisfy FTBIN on that assumption. Candidate (23c) is optimal because it constitutes a bimoraic foot, unlike (b), and it does not violate *[V:], unlike (a). Since long vowels are unattested in CVC roots, I assume the input has a short vowel. In that case, candidate (a) also violates IDENT[μ].

(23)

/le _μ l _μ /	MCAT \approx PRWD	FTBIN	IDENT[μ]	*[V:]	*[V]
'aquatic plant sp.'					
a. (le _μ l _μ) ~ [lɛ:l]			*!	*	
b. (le _μ l) ~ [lɛl]		*!	*		*
 c. (le _μ l _μ) ~ [lɛl]					*

I have assumed an *input* in (23) in which the coda is moraic. Candidate (c) is optimal regardless of whether the coda is underlyingly moraic or not - so this assumption does not affect the outcome. The assumption follows again from Lexicon Optimisation (P&S 1993:§9). Given that codas *can* be moraic, there is no reason to restrict their presence in the input: OT is unable to achieve this because of the 'Richness of the Base' hypothesis (P&S 1993, §6).¹³

¹³What P&S call Richness of the Base is the assumption that there are no principled limits on the amount of structure that can be posited in the input. All possible inputs for any given output are subject to evaluation according to the Lexicon Optimisation proposal. Given the obvious complexities in this notion I will not go into it here.

2.1.4 Words with four or more syllables

This section describes the stress patterns in simple words of more than three syllables. I call these 'bipedal' words for short, because they must be parsed by more than one foot. Each quadrisyllabic word with open syllables divides neatly into two disyllabic feet, each of which is left-headed: trochaic. These forms therefore provide the best evidence for trochaic rhythm in words.

There are 15 quadrisyllabic roots with all-open syllables for which I have stress data. Initial primary stress is observed in 8 of the 15 forms; these are shown in (24).¹⁴

- | | | | | |
|------|----|-------------|------------|----------------------------|
| (24) | a. | /paɪamunu/ | [báɪamùnu] | 'sand goanna' |
| | b. | /paɪaŋari/ | [báɪaŋàri] | 'Mitchell's water goanna' |
| | c. | /paɪikuluʔ/ | [bádɪgùlɥ] | ' <i>E. ferruginea</i> ' |
| | d. | /purukulu/ | [búruɡùlu] | 'snake sp.' |
| | e. | /kaŋamuru/ | [gáŋamùru] | '"long-nose" wild bee' |
| | f. | /kaŋawaraʔ/ | [gáŋawàrɔ] | 'tree-dwelling goanna sp.' |
| | g. | /karatata/ | [gáradàda] | 'chest brace' |
| | h. | /ŋoɔomoro/ | [ŋóɔomòro] | 'nail-tailed wallaby' |

I have not heard these forms pronounced with alternative stress patterns.


7 of the 15 roots vary in pronunciation between initial and penultimate stress; these are shown in (25).

- | | | | | |
|------|----|------------|-------------------------|-------------------------------|
| (25) | a. | /kurucaɬu/ | [gòrujáɬu] ~ [górujàɬu] | 'olive python' |
| | b. | /camaɭara/ | [jàmaɭára] ~ [jámaɭàra] | 'the Pleiades' |
| | c. | /ciriɬiti/ | [jìriɬídi] ~ [jírɪɬìdi] | 'fly catcher' |
| | d. | /ŋalaɭara/ | [ŋàlaɬára] ~ [ŋálaɬàra] | 'archer fish' |
| | e. | /tarapiya/ | [dàrabíya] ~ [dárabíya] | 'black cockatoo' |
| | f. | /wariɬila/ | [wàriɬíla] ~ [wáriɬíla] | 'hooked boomerang' |
| | g. | /kuɲɪpara/ | [gòɲɪpára] ~ [góɲɪpàra] | <i>Terminalia grandiflora</i> |

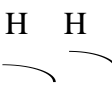
I have given all the forms for which I have reasonably clear intuitions on stress here. The generalisation then is, any form which takes penultimate primary stress also has a variant with initial primary stress.

¹⁴Syllables which are open except for a glottal stop count as open syllables; see Ch 5 for discussion.

The tonal characteristics of each type differ. Forms with an initial primary stress have a single salient pitch movement:

- (26)  bá.ɬamùnu
/paɬamunu/ 'sand goanna'

Forms with a penultimate primary stress always have a salient pitch movement on the initial syllable also:

- (27)  gùrujádu
/kurucaɬu/ 'olive python'

There are just a few simple words longer than four syllables, none of them entirely composed of open syllables. (28a) presents the only pentasyllabic form with mostly open syllables, and (28b) the only sextisyllabic form. Both forms have initial primary stress.

- (28) a. /woroŋomoloʔ/ [wóroŋomɔ̃lɔ] 'grass sp.'
b. /kaɬaŋkaŋaŋini/ [gá.ɬaŋgàŋaŋĩni] 'large macropod sp.'

Monomorphemic roots in Ngalakgan almost always have some stress on the first syllable. In disyllabic and trisyllabic forms, initial stress is also primary. In quadrisyllabic words, the initial stress may be primary or secondary, but in either case it is associated with a pitch movement. In longer words, stress is initial also. A formal account of these patterns is given in §2.2.4. First, I turn to a description of word

compounds. I will show that bipedal roots and compounds have a similar prosodic structure.¹⁵

2.2 Word compounds

Section 2.1 has shown that monomorphemic words always have stress on their initial syllable. When monomorphemic words are compounded - in a 'WORD-level compound' - each WORD retains the stress accent on its initial syllable. Prosodic structure therefore reflects the morphological structure of WORD-compounds. A similar relationship between stress and the structure of compounds is found in many other languages, for example English (e.g. Selkirk 1984), and Malayalam (Mohan 1986).

There is also a class of 'ROOT-compounds' which we will see in §2.2.3 have just one stress on their initial syllable, like simple words. There are no prosodic indicators of internal morphological structure in ROOT-compounds. This accords with their semantic characteristics: they are non-compositional.

WORD-compounds also have a distinctive superordinate prosodic structure, where metrically strong positions are associated to salient pitch movements, like the roots in previous sections. Based on the prosodic characteristics of WORD-compounds, in §2.2.4 I propose three distinct levels of metrical and prosodic constituency: Foot, Prosodic Word, and Phonological Word, each of which has its own headedness, phonetic, and organisational characteristics.

2.2.1 Examples of WORD-compounds

Compounding is a productive process in Ngalakgan, applying to all major lexical categories: noun+adjective, noun+verb, adverb+verb. There are numerous examples in

¹⁵ The lack of a pitch accent on a secondary stress following any main stress appears to be a general characteristic of intonational phrasing; see Truckenbrodt (1999). It is probable that further investigation of the nature of intonational phrasing in Ngalakgan will reveal the alternant placement of primary stress within bipedal roots to be determined at the Intonational Phrase level, which may in turn be equivalent to the Phonological Word described here. This would make complex words in Ngalakgan prosodically equivalent to some kinds of phrases in a language such as English, an equivalence which is reflected in the semantics.

the data, of which the following is a sample. The examples in (29-34) are composed of two elements, each of which can also occur alone as an independent word.¹⁶

- (29) Disyllable+disyllable
- | | | |
|----|--|--|
| a. | <i>gurndu-gakgeng</i>
country-far | [gùŋdugákkɛŋ]
'long way' |
| b. | <i>langga-ganyah</i>
lagoon-small | [làŋgagáŋa?]
'little billabong' |
| c. | <i>gurndu-yotjjong</i>
country-not.sacred | [gùŋduyóccɔŋ]
'clear country' |
| d. | <i>mala-borno</i>
group-another | [màlabónɔ]
'different person/group' |
| e. | <i>binyi-ngolkgo</i>
water-big | [bìŋiŋólkkɔ]
'big water' |
- (30) Disyllable+ trisyllable
- | | | |
|----|--|-------------------------------------|
| a. | <i>gurndu-japburru</i>
country-sacred | [gùŋdujápɸpuru]
'sacred country' |
| b. | <i>jolkgo-borrama</i>
ground-good | [jòlkkobórama]
'good ground' |
- (31) Disyllable+monosyllable
- | | | |
|----|--|--------------------------------------|
| a. | <i>gurndu-mah</i>
country-good | [gùŋdumáa?]
'good country/place' |
| b. | <i>jolkgo-derh</i>
ground-hard | [jòlkkodéɻ?]
'hard ground' |
| c. | [<i>gu-gu</i>]- <i>binyi-wurrk</i> ¹⁷
IRR-IV-water-swallow | [bìŋiwúrk]
'he swallows water' |
| d. | [<i>gu-gu</i>]- <i>binyi-ngalh</i>
IRR-IV-water-rise | [bìŋiŋál?]
'the water is rising' |
| e. | [<i>gu-gu</i>]- <i>mili-ngalh</i>
IRR-IV-water-rise | [mìliŋál?]
'the water is rising' |
| f. | [<i>gu</i>]- <i>binyi-jow</i>
IRR-water-flood | [bìŋijów]
'the water is flooding' |
| g. | <i>weya-mah</i>
shade-good | [wèiyamáa?]
'good shade' |

¹⁶There are some exceptions to this statement, which I return to below.

¹⁷ Prefixation with the Irrealis prefix *gu-* is obligatory for verbal predicates in the present tense. As shown in §3.3.2, prefixes such as these are consistently without stress in Ngalakgan, they do not affect the prosodic structure of the rest of the word.

- (32) Monosyllable+disyllable
- | | | |
|----|--|--|
| a. | <i>bolh-genggeng</i>
track-long | [bəl̩ʔgɛŋgɛŋ]
'long footprint' |
| b. | <i>bolh-jaruk</i>
track-short | [bəl̩ʔjɑ̌ɾʊk]
'short footprint' |
| c. | <i>rum-borno</i>
law-different | [ɾʊmbóno]
'(in a) different way/manner' |
| d. | <i>weh-balkginy</i>
water-salty/dangerous | [wɛʔbálkkɪŋ]
'saltwater; beer' |
| e. | [gu-]bo-wana
IRR-river-will.follow | [bò:wána]
'will follow the river' |
- (33) Monosyllable+monosyllable
- | | | |
|----|--------------------------------|-----------------------------|
| a. | <i>weh-mah</i>
water-good | [wɛʔmáaʔ]
'freshwater' |
| b. | <i>jeny-yarrh</i>
fish-lots | [jèɪnyáʔ]
'lots of fish' |
- (34) Trisyllable+monosyllable
- | | | |
|----|--|--------------------------------------|
| a. | <i>jardugal-yarrh</i>
kangaroo-lots | [jɑ̌dʊgalyáʔ]
'lots of kangaroos' |
|----|--|--------------------------------------|
- Trisyllable+disyllable
- | | | |
|----|--|-------------------------------------|
| b. | <i>mirpbarra-ganyah</i>
child-small | [mɪɾpparagánaʔ]
'a small child' |
| c. | <i>nganjurla-bordewk</i>
eye-bad | [ŋànju[abodɛwk]
'bad eyes' |
| d. | <i>ngu-bu-mirpbarra-naniny</i>
1mS-3a-child-saw | [mɪɾpparanánɪŋ]
'I saw children' |

2.2.2 Compounding morphology

The constituents of WORD-compounds retain the morphological and semantic characteristics that they have in other environments. Each element in a WORD-compound is typically a freely-occurring word.

- (35)
- ```

 WORD-compound
 / \
 WORD WORD

```

WORD-compounding is a morphological operation on WORDs, a 'WORD-level' operation.

The morphological relationship between each element in a WORD-compound can be characterised as in (36):

- (36) WORD-level characteristics
- a. Freedom of appearance
  - b. Lack of morphophonemic selectional requirements
  - c. Compositional contribution to meaning

These are characteristic of other WORD-level processes in Ngalakgan also - affixation and reduplication - which are examined in Chs 3 and 5, respectively.

'Freedom of appearance' (36a) encodes the fact that each element of a WORD-compound is a word which can appear in many other compounds. There are no idiosyncratic morphophonemic requirements which obtain between the elements of WORD-compounds (36b); the combination of words to form WORD-compounds is in principle unrestricted. Lastly, each element in a WORD-compound retains the meaning that it has as a word (36c), and words in compounds are interpreted in a standard fashion, with the right-most element constituting the head.

I define 'head' as that element in a complex form which determines the morphological behaviour of the form as a whole. In a noun+adjective compound, as in (37), the adjective *japburru* 'not sacred' is the head. The word as a whole has the function of a predicate in this clause: only predicates take tense-inflected negatives such as *-hmolk*. Since the word is a predicate, the adjective must be the head: only adjectives can readily function as predicates or as referring expressions (nominals) in Ngalakgan.

- (37) *gu-gurndu-japburru-hmolk*      *gurnmarnh*  
 IV-country-sacred-PNEG      maybe  
 'Maybe it's not sacred.'

[3/9/97:2A]

Given the meanings of two elements, and the right-headed form of compounds, the interpretation of WORD-compounds is straightforward.<sup>18</sup>

I define '*stem*' as any form of a root or word in a morphologically complex environment. In Ngalakgan there are four types of environments of this kind: affixation, compounding, reduplication, cliticisation. All of the elements in the WORD-compounds in (29-34) are WORD-level stems, though not all of them are freely occurring words.

For instance, in (31c-f), the two morphemes meaning 'water', *binyi-* and *mili-*, are both bound, WORD-level stems. They are not 'words' because they cannot occur alone, but must be compounded with some other (WORD-level) element. In all other respects, they have the characteristics of WORD-level morphology: freedom of appearance, lack of morphophonemic restrictions, and compositional contribution to meaning.

### 2.2.3 Prosodic characteristics of WORD-compounds

The prosodic structure of WORD-compounds provides evidence about the nature of prosodic constituents in Ngalakgan. I show that there are distinct Prosodic Word (PrWd) and Phonological Word (P-Wd) constituents; PrWds are left-headed and P-Wds right-headed.

WORD-compounds have a distinctive prosodic structure. Each of the words of the compound is stressed as a word. In the compound as a whole, the first element takes secondary stress, while the second takes primary stress, as shown in (38).

---

<sup>18</sup>Although Merlan (1983:57) found only adjectival and human category (including kinship) nouns as predicates, other kinds of nouns can occasionally function as predicates; for example generics (i.e. not species names), as in (87).

- (87)    a.    *yanipbi*            *gu-baramulk*,            *gu-mu-ngu+n*,  
              whatsit            IV-vine.sp            IRR-III-eat+PR
- b.    *bo-wi*            *nu-gohje*            *gu-mu-jung-gun*  
              river-LAT        I-DEM            IRR-III-shade-REL  
              'It eats *baramulk* (wild cucumber), at the river where that one is shaded.'    [10/9/97]

The phrase *gu-mu-jung-gun* 'where it is shaded' can only be a predicate in this example because predicates alone take the IRRealis prefix *gu-*. Class IV *baramulk* shows 'disagreement' in this example with the class III prefix on the verb. This phenomenon is addressed separately (Baker 1997c).

- (38) *gurndu-japburru* [gùŋɖuǰáppuru]  
country-sacred 'sacred country'

Unlike the bipedal roots discussed in §2.1.4, in WORD-compounds the second stress accent is always the primary one; these words do not show variation in the relationship between secondary and primary stress.

We can see from the examples in (39) the stress that each of these words has in isolation. Comparing (38) and (39), we can see that whatever stress pattern a word has in isolation is retained in the stress organisation of a compound in which it appears.

- (39) a. *gurndu* [gùŋɖu] 'country'  
b. *japburru* [jáppuru] 'sacred'

Each stem in a compound is therefore a distinct metrical domain. This is because each stem is subject to  $\text{MCAT} \approx \text{PRWD}$  and  $\text{ALIGNL}(\text{PrWd}, \text{Foot})$ . Therefore, each stem must minimally constitute a bimoraic foot, and each stem is associated with an initial stress.

Since not all stems are freely occurring words, I revise  $\text{MCAT} \approx \text{PRWD}$  accordingly. In Ngalakgan, this constraint has a more particular instantiation, where  $\text{MCat}$  is 'Morphological Word' (MWd). A 'Morphological Word' is a WORD-level stem, as defined previously.

The constraint relevant then to Ngalakgan is given in (40):

- (40)  $\text{MWD} \approx \text{PRWD}$ : 'Every Morphological Word corresponds to a Prosodic Word'.

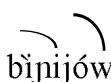
In a compound, each stress is associated with a distinctive, salient pitch movement, as shown in (41).<sup>19</sup>

---

<sup>19</sup>The IRR prefix /ku-/ is obligatory in Present tense, the representation in (41) abstracts away from such extraneous elements.



(41)

  
 bĩnjíów  
 /piɲi-cow/  
 water-flood    'water is flooding'

I call these salient pitch movements 'pitch accents'. The distribution of metrical heads and pitch accents is not necessarily the same. In the word in (42) below, there are three metrical feet but only two MWds, and there are correspondingly only two pitch accents. Only those metrical prominences which are associated with pitch accents are marked with accent diacritics (secondary ` and primary ´) in examples.

|      |           |     |     |          |   |      |
|------|-----------|-----|-----|----------|---|------|
|      | ( .       |     | x   | )        | 3 |      |
|      | ( x       | . ) | ( x | . )      | 2 | P-Wd |
|      | ( x       | . ) | ( x | . )      | 1 | Foot |
| (42) | k         | à   | r   | a        | k | k    |
|      | a         | r   | a   | -        | c | á    |
|      | ŋ         | +   | a   | n        |   |      |
|      | together- |     |     | stand+PR |   |      |

The distinction between metrical structure, pitch accents, and secondary vs primary accent in a word suggests three levels of prosodic structure. The first level ('Foot') corresponds to 'rhythmic prominence': the prominence given to syllables because of their organisation into rhythmic (metrical) feet. At the second level ('Prosodic Word'), prominent syllables have tonal characteristics. At the third level ('Phonological Word'), the only prominent syllable is the one which has the greatest amplitude: that is, has primary stress.<sup>20</sup>

The three prosodic constituents - Foot, PrWd, and P-Wd - are subject to distinct constraints. The Foot is subject to constraints FTBIN and FTFORM as discussed previously. The PrWd and P-Wd constituents are subject to distinct *Alignment* constraints.

---

<sup>20</sup>Acoustic study of Ngalakgan is at a preliminary stage, and I present this as my impressionistic view.

While the P-Wd is characterised by right-headedness, Prosodic Words are left-headed: they are both regulated by Alignment constraints referring to prosodic 'peaks' - which are local metrical maxima (McCarthy and Prince 1986:9):

(43) ALIGNL(PK, PRWD): 'Align the peak left in a PrWd'.

(44) ALIGNR(PK, P-WD): 'Align the peak right in a P-Wd'.

These two constraints define two prosodic domains within which peaks are computed. The P-Wd is defined over the whole word, is right-headed, and determines the position of primary stress. Being the superordinate constituent, the P-Wd may be recursively embedded.

The other domain, the PrWd, is smaller than or equal to the P-Wd. It is left-headed, and determines secondary *as well as* primary stresses. PrWds may not be recursively embedded - as we would expect, being the domain of contrastive tonal effects - though there may be more than one PrWd per P-Wd. PrWds must form adjacent, non-embedded domains.

I assume that only the heads of feet can be associated with pitch accents. I encode this as a constraint 'Peak-to-Head' (cf. Prince's 1983 'Continuous Column Constraint').

(45) PK-TO-HEAD: 'Peaks (at level  $n+1$ ) are associated to metrical heads (at level  $n$ )'.

We have seen that the location of feet, and thereby heads, is determined by morphological structure. Hence, given  $MWD \approx PRWD$ , and  $ALIGNL(PRWD, FT)$ , we derive MWds with the form shown in (46). The *leftmost* head of each MWd is projected as the peak at the PrWd level, by  $ALIGNL(PK, PRWD)$ .

$$(46) \quad \begin{array}{cc} \begin{array}{c} (x \quad . \quad ) \\ (x \quad .)(x \quad .) \end{array} & \begin{array}{c} (x \quad . \quad ) \\ (x \quad .)(x \quad .) \end{array} \\ [MWd(\mu \quad \mu)(\mu \quad \mu) \dots] - [MWd(\mu \quad \mu)(\mu \quad \mu) \dots] \end{array}$$

By ALIGNR(PK, P-WD), the *rightmost* of these strong positions will realise primary stress, and the other position will realise secondary stress:

$$(47) \quad \begin{array}{ccc} \begin{array}{c} ( \quad . \quad ) \\ (x \quad . \quad ) \\ (x \quad .)(x \quad .) \end{array} & \begin{array}{c} x \quad ) \\ (x \quad . \quad ) \\ (x \quad .)(x \quad .) \end{array} & \begin{array}{c} P-Wd \\ PrWd \\ Foot \end{array} \\ [MWd(\mu \quad \mu)(\mu \quad \mu) \dots] - [MWd(\mu \quad \mu)(\mu \quad \mu) \dots] \end{array}$$

With Pk-to-Head undominated, I assume that ALIGNL(PK, PRWD) and ALIGNR(PK, P-WD) are only violated when a peak is projected from any foot which is non-initial in MWd. For instance, if the second foot head of (47) were to realise the local PrWd peak this would violate ALIGNL. Any other realisation of the peak would violate Pk-to-Head.

The effects of the various constraints can be seen in consideration of a simple example such as that in (48). None of the constraints shown here are in conflict. The word has two metrically strong positions, each associated with the initial syllable of a MWd, by  $MWD \approx PRWD$  and ALIGNL(PRWD, Ft). The rightmost of these must be projected to primary stress, by ALIGNR(PK, P-WD). Candidate (a) violates this constraint thereby. I assume candidate (d) violates  $MWD \approx PRWD$  because it is not true that every MWd - there are two - corresponds to its own PrWd domain - there is just one. While this is a representational difference, it allows the distribution of PrWds and P-Wds to be described uniformly.

I assume certain possibilities are ruled out by universals of metrical organisation, governed by an undominated constraint I have simply called 'WORDFORM'. A form such as (48c), where each prosodic word receives a primary stress is ruled out by what Liberman and Prince (1977:262) call 'Culminativity', the fact that 'each word or phrase normally has a single strongest syllable bearing the main stress' (Hayes 1995:24).

(48)

| /kúnʦu-cappuru/<br>country-sacred | WD<br>FORM | MWD ≈<br>PRWD | ALIGNL<br>(PRWD,<br>FT) | ALIGNL<br>(PK,<br>PRWD) | ALIGNR<br>(PK, P-WD) |
|-----------------------------------|------------|---------------|-------------------------|-------------------------|----------------------|
| a. [(kúnʦu) I-[(càppu)ru] I]      |            |               |                         |                         | *!                   |
| ☞ b. [(kùnʦu) I-[(cáppu)ru] I]    |            |               |                         |                         |                      |
| c. [(kúnʦu) I-[(cáppu)ru] I]      | *!         |               |                         |                         |                      |
| d. [(kùnʦu)-(cáppu)ru] I]         |            | *!            |                         |                         |                      |
| e. [(kúnʦu) I-[(cap(púru)] I]     |            |               | *!                      |                         |                      |

The characteristics of compounds then are the characteristics of words, and in particular bipedal words with penultimate stress. Stress is hierarchically organised and culminative: resulting in a single peak. Each stressed syllable corresponds to the initial syllable of a WORD.

The more complex example of (42) above is analysed in (49). A form, as in (49c), where each foot in a PrWd realises a distinctive stress violates ALIGNL(PK, PRWD), since at least one peak is not leftmost. But it is probably to be ruled out by Culminativity as before in any case, since within each domain, stresses should be hierarchically organised.

(49)

| /karakkara-caṇa+n/<br>together-stand+PR | ALIGNL<br>(PRWD, FT) | ALIGNL<br>(PK, PRWD) | ALIGNR<br>(PK, P-WD) |
|-----------------------------------------|----------------------|----------------------|----------------------|
| a. [(kárak)(kara) I-[(càṇa)n] I]        |                      |                      | *!                   |
| ☞ b. [(kàrak)(kara) I-[(cáṇa)n] I]      |                      |                      |                      |
| c. [(kàrak)(kàra) I-[(cáṇa)n] I]        |                      | *!                   |                      |

In a few cases, the preference for no more than two pitch accents overrides  $MWD \approx PRWD$ , as in (50) below. This example has three MWds, /piʈic/ 'nearly', /ɲariɲ/ 'hand' and /peɲ/ 'bit', but only two salient pitch movements occur.<sup>21</sup>

- (50) a. [nʊnbiʈicɲariɲbéɲ]  
 /nʊn-piʈic-ɲariɲ-pe+ɲ/  
 2mO-nearly-hand-bite+PP  
 '[It: dog] nearly bit your hand.' [31/5/96:1B]

Compounds with more than two stems are rare, and so the maximum number of pitch accents in a word is normally two. The only attested examples from my data with more than two pitch accents are given in (51):

- (51) a. [gugàmalagàrakkaraǰáɲan]  
 /ku-kamala-karakkara-caɲ+an/  
 IRR-cloud/sky-together-stand+PR  
 'daylight breaks out' [28/6/96:2B]
- b. [juɲàmʊlugèɲɛɲána]  
 /cu-ɲamulu-kɛɲɛ-ɲa+na/  
 2mS-really-body-see+FUT  
*yu labta faindim prabli*  
 'You'll have to find that animal properly.' [2/9/97:1B]
- c. [ɲubugèɲɛbòɲnanɪkkóro]  
 /ɲu-pu-kɛɲɛ-pur?+ɲa+ni+kkoro/  
 1mS-3a-body-know+[see]+POT+PRNEG  
 'I don't know them' [3/9/97]

Comparing (50) and the examples in (51) above, it appears that two syllables can intervene between successive pitch accents. But an interval of three or more syllables is too long a stretch without some kind of metrical prominence, whether or not this is reflected in the intonation contour.

The existence of forms like (50), though very rare, suggests that the Alignment constraints on peaks dominate  $MWD \approx PRWD$ , so that only the leftmost and rightmost stress accents are projected to pitch. Ideally, there is just one leftmost head and one

<sup>21</sup>Stress on the syllable [ʈic] of prefix /piʈic-/ here is due to syllable weight effects on stress examined in Chapter 5.

rightmost head. A series of three PrWds, as in /kàmala-kàrakkara-cáŋ+an/, disrupts this symmetry. Such forms are avoided: most complex words in Ngalakgan contain no more than two pitch accents.

#### **2.2.4 Summary**

This section has shown that WORD-level stems - MWds - retain their prosodic characteristics when compounded. These complex words in turn have distinctive, and consistent, prosodic characteristics. While PrWds form left-headed and adjacent tone domains, P-Wds are right-headed and culminative metrical domains.

In the next sections I examine the phonological, semantic and prosodic characteristics of ROOT-level morphology, which are quite different from that of the WORD-level.

### **2.3 ROOT-morphology and prosody**

In the following sections I show that the prosodic and semantic characteristics of words in Ngalakgan can be captured by the generalisation in (52).

(52) Stems which are ROOT-complex are prosodically equivalent to simple WORDs.

ROOT-level morphology is invisible to prosodic constraints.

In what follows I firstly discuss the distributional and semantic characteristics of ROOT-morphology in the lexicon. I then show that for two kinds of ROOT-morphology - affixation and compounding - the statement in (52) is true. I do this by describing the prosodic characteristics of ROOT-complex words, and comparing these to the prosody of words which are complex at the WORD-level.

#### **2.3.1 ROOT-level morphology**

Elements in a ROOT-level relationship have the following semantic and distributional characteristics (and compare the preceding list in 36):

- (53) ROOT-level characteristics
- Restricted appearance
  - Morpho-phonemic selectional requirements
  - Non-compositional contribution to meaning

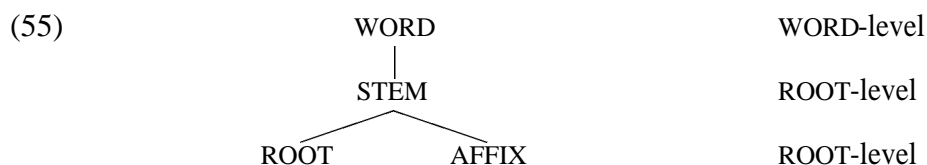
Bound roots cannot occur on the surface in a root form - they are obligatorily subject to ROOT-level morphology. This is the content of characteristic (53a), and bound roots contrast in this respect with *free* roots, which may be surface words. The morphological roots examined in the preceding sections were on the whole 'free roots': able to appear at the surface as WORDs, or freely combined with other WORD-level stems.

For example, corresponding to the surface word in (54a), there is no freestanding word such as [bu:] 'hit'.

- (54) a. /pu/+/niɲ/            [buniɲ] '(he/she/it) was hitting (s.t/s.o)'  
              'hit (PC)'
- b. /pu/                    \*[bu:]

Morphemes in a ROOT-level relationship are separated by a '+' boundary symbol, those in a WORD-level relationship are separated by a '-' boundary symbol.

ROOT-affixation, ROOT-compounding, and ROOT-reduplication all create ROOT-level stems out of elements that cannot be words by themselves. Some of these stems are subject to further derivation, in order to create WORDs. Others can surface as WORDs without further derivation. This claim is represented schematically in (55) (and compare the preceding diagram in (35)):



The most important set of forms in the class of bound roots are the 32 inflecting verb roots - such as *bu+* 'hit' and *ma+* 'get'. This class is called the *finite* verb class. All other predicates in the language inflect by means of a construction which is morphologically a compound, using one of these roots as an inflected auxiliary.

The affixes to finite verb roots are paradigmatically highly irregular, idiosyncratic and unproductive, thus satisfying characteristics (53a) and (b). It is not the case that all of a verb's inflectional forms can be predicted given any of the others.

Consider the following representative sample for instance:

|      |            |           |               |            |         |
|------|------------|-----------|---------------|------------|---------|
| (56) | √pu+       | √ma+      | √ca(ŋ)+       | √karpe+    | √co+    |
|      | 'hit'      | 'get'     | 'stand (itv)' | 'crawl'    | 'chop'  |
| PP   | poʔpo      | meʔme     | ca+n          | karrpe+n   | ce      |
| PR   | pu+n       | maʔma     | caŋa+n        | karrpe     | co      |
| EVIT | "          | ma        | "             | "          | "       |
| PC   | pun+iŋ     | maŋ+iŋ    | caŋan+iŋ      | karrpen+iŋ | coŋ+ iŋ |
| FUT  | pun+a      | maŋ+a     | caŋan+a       | karrpen+a  | coŋ+a   |
| POT  | pun+i      | maŋ+i     | caŋan+i       | karrpen+i  | coŋ+i   |
| RR   | pu-yci-    | maŋi-cci- |               |            |         |
|      | ~ pu-ycci- |           |               |            |         |

Some inflections have a predictable realisation in *most* verbs. For almost all finite verbs, the Future and Potential inflections are equal to the Past Continuous, with -a and -i substituted for -iŋ.<sup>22</sup> We can therefore identify a Non-Past (NP) stem, equivalent to the inflected form minus these three inflections. For the verb *ma+* 'to get', for instance, the NP stem is *maŋ*.

The relationship between these inflections (PC, FUT, POT) and the older ones - Present and Past Punctual - is much less regular. A number of verbs show an ablaut relation between the PR and PP, as in *co* 'chop' PR, *ce* PP. Other verbs have an overt PP

<sup>22</sup>There are four exceptional verbs which do not display this regular relationship. The verb *caŋe* 'twirl firedrill', for example, has PC *caŋi+n* but FUT, POT in *caŋiŋ+a*, *caŋiŋ+i* respectively. The Future and Potential inflectional forms in Ngalakgan are of comparatively recent origin. They serve to distinguish Rembarrnga and Ngalakgan from their nearest genetic relatives - Mayali, Dalabon and Ngandi - which do not have a regular correspondence like this between PC, FUT and POT (or their equivalents).



suffix, e.g. ca+n 'stood' (PP), cf. caŋ+an 'be standing' (PR). Some verbs have both an overt suffix and ablaut changes in the stem: ŋow+iŋ 'ate' (PP), cf. ŋu+n 'eat' (PR).

It is possible to group the conjugations into inflection classes of paradigms showing some broad correspondences. But the inflectional forms of finite verb roots, with the possible exception of the Future, Potential, and Reflexive-Reciprocal forms, are irregular enough to warrant that each verb paradigm be separately listed in the lexicon in full. Since there are just 32 finite verbs, this is not such a great hardship for the learner. Hence, lexical entries for finite verbs will have the following form:

|      |                          |                        |                           |                           |
|------|--------------------------|------------------------|---------------------------|---------------------------|
| (57) | <u>Phonological form</u> | <u>Lexical content</u> | <u>Argument structure</u> | <u>Syntactic features</u> |
|      | /pun/                    | 'hit'                  | <x <y>>                   | V, (PR)                   |
|      | /punij/                  | 'hit'                  | <x <y>>                   | V, (PC)                   |
|      | /ma/                     | 'get'                  | <x <y>>                   | V, (EVIT)                 |
|      | /maŋij/                  | 'get'                  | <x <y>>                   | V, (PC)                   |

The inflected form of finite verb roots are amenable to the 'full-listing' approach (cf. Jackendoff 1975; Aronoff 1976), since the rules and subcategorisations required to generate inflected forms from separately listed roots and affixes would not appreciably increase the parsimony of the grammar.<sup>23</sup>

The last characteristic of ROOT-level morphology proposed in (53) - non-compositional contribution to interpretation - is illustrated by ROOT-level compounding forms.

### 2.3.2 ROOT-compounds

There are around 150 verbs in the language which are formally ROOT-level compounds. ROOT-compounds have idiosyncratic meanings, and often consist of at

---

<sup>23</sup>An approach such as Jackendoff's (1975) would minimise the drawbacks of full-listing. 'Lexical redundancy' relationships would lessen the overall cost of listing by eliminating redundancies such as those between the PC, Fut, Pot and RR forms, which are on the whole in a regular relationship.

least one root which is not found elsewhere in the morphology. Some examples are given in (58-59).

- (58) a. /koʔ+ŋa+nɪp/      [gɔ̌:ŋaniɲ]      'had' (PC)  
       b. /ŋa+nɪp/            [naniɲ]            'saw' (PC)  
       c. \*/koʔ/
- (59) a. /wi+ŋa+n/            [wɪŋan]            'forgets' (PR)  
       b. /ŋa+n/              [nan]                'sees' (PR)  
       c. \*/wi/

I will refer to the initial element in a verb compound (whether ROOT- or WORD-level) as the *coverb*, in contrast to the *finite* verb, which is the final element. It is the finite verb which hosts the inflection of the verb compound.

In the ROOT-compound /koʔ+ŋa+/ 'to have', we recognise the root /ŋa+/ 'to see'. The ROOT-compound /koʔ+ŋa+/ inflects exactly like the finite verb /ŋa+/. This is the case with all ROOT-compounds - the inflection of the whole is that of the final finite verb root. However, the meaning of the ROOT-compound does not include, or is not derived from, the meaning of at least the second member /ŋa+/ 'to see' in any obvious way. For this reason, in glossing root compounds, I enclose the meaning of the finite verb in square brackets: /wi+ŋa+n/ 'forget+[see+]PR'.

The same remarks apply to (59) and other root compounds, although occasionally the compound does bear some (unpredictable) relationship with the meaning of the verb root (Merlan 1983:111); some examples are presented in (60).

- (60) a. /purʔ+ŋa/            'to know'  
       b. /pop+ŋa+/        'to smell'  
       c. /tɪc+ŋa+/        'to look hard at'

The first element of such compounds typically has no status whatsoever outside the compound. There is no element /koʔ/ (or /wi/ or /purʔ/ or /tɪc/) which occurs otherwise with any component of the meaning of the ROOT-compounds in (58)-(60).

Although again, there are occasional exceptions. The element /pop/ for instance can also be a verb stem in its own right: /pop-mi+/ 'to be smelly'.<sup>24</sup>

ROOT-compounds therefore contrast with WORD-compounds, as in (61), in which each element occurs in many other complex environments with the same interpretation:

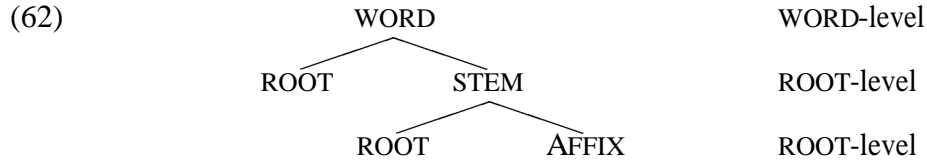
- |      |    |                                                    |                                   |
|------|----|----------------------------------------------------|-----------------------------------|
| (61) | a. | /ney/                                              | 'name (N)'                        |
|      | b. | /ɲur-pu-ney-wu+na/<br>12aS-3a-name-give+FUT        | 'we will give them names'         |
|      | c. | /ɲun-ney-wi+ɲa+n/<br>1mO-name-forget+[see+]PR      | 'he's forgotten my name'          |
|      | d. | /pur-kun̥tu-ney+pu+n/<br>3aS-country-name+[hit+]PR | 'they name the country [toponym]' |

The free root *ngey* 'name' can occur alone as a noun, as in (a). It can be incorporated, to form a compound verb as in (b) and (c), where its contribution to the meaning of the word remains constant. Example (d) shows that it also occurs in ROOT-compounds. In this case, the meaning of the whole is not recoverable from the meaning of the parts: *ngey+bu+* has a fixed, idiosyncratic interpretation like other ROOT-compounds.<sup>25</sup>

<sup>24</sup> The finite root /-mi+/ (no meaning) is the verbaliser for the open ('thematic') class of verb roots. This class is productive in Ngalakgan; for instance, it derives all attested loaned verbs from English via Kriol, e.g. /ɲu-laytimap-mi+ɲ/ (1mS-set.alight-AUX+PP) 'I lit it up', /ku-caɭim?-ma-kan/ (IRR-start(tr)-FUT-PURP) 'so he can start it'. The Future and Potential tense/moods in this class are derived by geminating the final supra-oral segment of the stem. See §5.5.2 for more examples.

<sup>25</sup> Around half of the compounds with pu+ have a first element which occurs elsewhere as a word. Typically the pu+ compound 'verbalises' the meaning of the noun, as here. In the related GN languages Warray and Jawoyn, pu+ is a productive 'Factitive' auxiliary, with a meaning 'cause X to come into being', where X is a noun. It can for instance take loanwords from English in Warray (Harvey p.c., and MSb): [ɛyp+pu+ 'make a [cassette] tape [recording]'. Many pu+ compounds in Ngalakgan also have this kind of interpretation: cuɲ+pu+ 'build a bough shade' (ku-cuɲ 'a bough shade'), cele+pu+ 'to urinate' (ku-cele 'urine'). But this seems to be no longer productive in Ngalakgan. The analogous loanword example to Warray occurs in Ngalakgan as [ɛyp-man̥ɲɲ-mi+ 'make a tape', where man̥ɲɲ 'make, build' is a WORD-level verb stem of the thematic class.

The morphological structure of ROOT-compounds is schematised in (62):



ROOT-compounding creates WORDs. The WORD constituent in (62) is predicted to behave like other, morphologically-simple WORD-level stems. That this is the case is shown in the following section.

### 2.3.3 Prosodic characteristics of ROOT-morphology

Words consisting of bound root+affix have the prosodic characteristics of simple words: disyllabic, trisyllabic and quadrisyllabic forms have initial stress (in quadrisyllabics initial stress is secondary, primary stress is penultimate). This is shown in the examples in (63-65), and compare the simplex words in §2.1.

(63)      Disyllables

- |    |                              |                                     |
|----|------------------------------|-------------------------------------|
| a. | <i>bu+ni</i><br>hit+POT      | [búni]<br>'he/she wants to hit it'  |
| b. | <i>ga+nginy</i><br>take+PP   | [gáŋɪŋ]<br>'he/she took it'         |
| c. | <i>jo+nginy</i><br>chop+PC   | [jɔŋɪŋ]<br>'he/she was chopping it' |
| d. | <i>garrbe+ny</i><br>crawl+PP | [gárbeŋ]<br>'he/she crawled'        |

(64)      Trisyllables

- |    |                                         |                                                 |
|----|-----------------------------------------|-------------------------------------------------|
| a. | <i>rabo+niny</i><br>go+PC               | [ɾábonɪŋ]<br>'he/she was going'                 |
| b. | <i>jurruwe+ny</i><br>rush+PP            | [júruweŋ]<br>'he/she was rushing'               |
| c. | <i>yini+ngi</i><br>'say/do (like that)' | [yíniŋi]<br>'he/she should say/do it like that' |
| d. | <i>ma+ngi+tjji+ny</i><br>get+POT+RR+PP  | [máŋɪccɪŋ]<br>'he/she got him/herself'          |

- (65) Quadrisyllables
- |    |                       |                          |
|----|-----------------------|--------------------------|
| a. | <i>jurruwen+iny</i>   | [jùruwénɪŋ]              |
|    | rush+PP               | 'he/she was rushing'     |
| b. | <i>ma+ngi+tjji+ni</i> | [màŋiccíni]              |
|    | get+POT+RR+PC         | 'he/she got him/herself' |

Affixed forms which are ROOT-complex contrast with words affixed at the WORD-level, as the contrast between (66) and (67) shows. Nouns take only WORD-level affixes. ROOT-level affixes are restricted to the class of inflecting verb roots.

- |      | <u>Verb</u>                                           | <u>Noun</u>                                             |
|------|-------------------------------------------------------|---------------------------------------------------------|
| (66) | a. [jáŋana]<br>/ca+ŋana/<br>stand+FUT<br>'will stand' | b. [gè:ŋíni]<br>/ke=ŋini/<br>son=1mDAT<br>'my son'      |
| (67) | a. [búni]<br>/pu+ni/<br>hit+POT<br>'might hit'        | b. [bó:wi]<br>/po-wi/<br>river-LAT<br>'along the river' |

Examples (66a), (67a) are complex at the ROOT-level but not the WORD-level, so they are realised with a single, initial stress, and the initial syllable is short. Words like this contrast with forms such as (66b), (67b), which are complex at the WORD-level. Both contain a MWd which is monosyllabic and open, followed in (66b) by a WORD-level enclitic, and in (67b) by a WORD-level suffix. The MWd must constitute a PrWd, and so it must be realised with a long vowel. The prosody of affixes is discussed in detail in Ch 3. The difference between the two kinds of words is shown in (68).

- (68)
- |                                  |                                     |
|----------------------------------|-------------------------------------|
| WORD                             | WORD                                |
| ROOT      AFFIX<br>ca+      ŋana | WORD      ENCLITIC<br>ke      =ŋini |

The vowels of the monosyllabic open roots /ca+/ and /pu+/ of the (a) examples are not long at the surface, and the following affix is not stressed. The word as

a whole takes a single stress. The roots of ROOT-affixed words, then, are not MWds, and are not themselves required to constitute PrWds.

The same pattern is observed in ROOT-compounds. Disyllabic and trisyllabic compounds have a single stress on the initial syllable of the compound. (69) and (70) exemplify the ROOT-compound stress pattern.

(69) Disyllables

- |    |                                   |                       |
|----|-----------------------------------|-----------------------|
| a. | <i>goh+na+n</i><br>have+[see+]PR  | [gɔ́ːnən]<br>'has'    |
| b. | <i>wotj+ma</i><br>steal+[get]PR   | [wɔ́cma]<br>'steals'  |
| c. | <i>wernh+mi+ny</i><br>look+AUX+PP | [wɛ́ŋmɪɲ]<br>'looked' |
| d. | <i>lak+bo</i><br>split+[hit]PP    | [lákbɔ]<br>'splits'   |

(70) Trisyllables

- |    |                                     |                                       |
|----|-------------------------------------|---------------------------------------|
| a. | <i>goh+nan+iny</i><br>have+[see]+PC | [gɔ́ːnaniɲ]<br>'had'                  |
| b. | <i>lak+bu+na</i><br>split+[hit+]FUT | [lákbuna]<br>'will split'             |
| c. | <i>binje+pbo</i><br>quarter+[hit]PP | [bíɲjeppo]<br>'quartered [a carcass]' |
| d. | <i>yinih+gan</i><br>say+[take]PR    | [yínɪgan]<br>'tells someone to'       |
| e. | <i>gangah+wo</i><br>ask+[give]PP    | [gáŋgawo]<br>'asked'                  |

In the examples in (71) which follow, the contrast between WORD and ROOT-level prosody is made clear. Disyllabic forms which are complex at the ROOT-level, such as (69a), contrast with the double-headed prosody of forms which are complex at the WORD-level, as in (71b). ROOT-compounds therefore contrast with WORD-compounds, which take a stress on each WORD-level constituent.

- (71) a. [gɔ́ːɲan]  
/koʔ+ɲa+n/  
have+[see+]PR
- b. [wèɛmáa]  
/weʔ-maʔ/  
water-good

Similarly, trisyllabic ROOT-compounds (72a), (73a) contrast with trisyllabic WORD-compounds (72b), (73b). Trisyllabic ROOT-compounds have a single initial stress, while trisyllabic WORD-compounds have two stresses: one on each MWd.

- (72) a. [gɔ́ːɲanɪɲ]  
/koʔ+ɲan+ɪɲ/  
have+[see+]PC
- b. [ɲòò<sup>2</sup>máɲɪɲ]  
/ɲoʔ-maɲ+ɪɲ/  
guts-get+PC
- (73) a. [míɲjɪppun]  
/mɪɲci+ppu+n/  
remember+[hit+]PR
- b. [mèlɛɲán]  
/meleʔ-ɲa+n/  
lest-see+PR

Note that trisyllabic ROOT-compounds behave in the same way regardless of their constituency at the ROOT-level. In WORD-compounds, by contrast, the location of stress depends on the internal constituency. A form, as in (72b), with a final disyllabic MWd has penultimate stress, whereas a form with a final monosyllabic MWd has final stress, as in (73b).

These facts follow from the constraints already discussed. An affixed or compounded ROOT consists of a single MWd. By  $MWD \approx PRWD$  then, a ROOT-compound is only required to correspond to a single PrWd and a single foot. This is shown in tableaux (74) and (75) below. Tableau (74) presents an analysis of a WORD-compound. In (74a), each MWd is separately footed; separate footing is a pre-requisite to separate PrWd-hood, by the Prosodic Hierarchy hypothesis. To not do so, as in candidate (b), violates  $MWD \approx PRWD$ . This is the case even though candidate (a) here violates the constraint against a short vowel:long vowel correspondence -

IDENT[mora].<sup>26</sup>

<sup>26</sup>The possible candidate where glottal stop bears the mora is ruled out by other constraints examined separately in §5.5.3.

(74)

| /ŋoʔ-maŋ+iŋ/<br>guts-get+PC          | FTBIN | ALIGNL<br>(PRWD,<br>FT) | MWD ≈<br>PRWD | IDENT[μ] |
|--------------------------------------|-------|-------------------------|---------------|----------|
| ☞ a. [(ŋo <sub>μ</sub> ʔ)-l-(maŋiŋ)] |       |                         |               | *        |
| b. [(ŋo <sub>μ</sub> ʔ-ma)ŋiŋ]       |       |                         | *!            |          |
| c. (ŋo <sub>μ</sub> ʔ)-l-(maŋiŋ)]    |       |                         | *!            | *        |

Tableau (75) presents the analogous ROOT-compound form. Candidate (b) is the best option in (75) because the PrWd, and hence the foot, is left-aligned to the only morpheme which is 'visible' to prosody: the ROOT-compound /koʔ+ŋan+iŋ/ meaning 'have (PC)'. This constituent is footed only once, because to build a heavy foot on the initial syllable /koʔ/ violates the constraint against long vowels. Violation of this constraint is licensed in (74), because of higher-ranked MWD ≈ PRWD. In (75a), it is an unnecessary violation of well-formedness and (b) is the optimal candidate.

(75)

| /koʔ+ŋan+iŋ/<br>have+[see]+PC   | FTBIN | ALIGNL<br>(PRWD,<br>FT) | MWD ≈<br>PRWD | IDENT[μ] |
|---------------------------------|-------|-------------------------|---------------|----------|
| a. [(ko <sub>μ</sub> ʔ)(ŋaniŋ)] |       |                         |               | *!       |
| ☞ b. [(koʔŋa)niŋ]               |       |                         |               |          |
| c. [koʔ(ŋaniŋ)]                 |       | *!                      |               |          |

The preceding sections have shown that there is a clear prosodic contrast between ROOT-compounds and WORD-compounds. ROOT-compounds have the prosodic characteristics of simple words, whereas WORD-compounds retain the prosodic characteristics of each of their constituent words. The prosodic contrast reflects the morpho-semantic contrast: ROOT-compounds are lexicalised, WORD-compounds are productive. ROOT-compounds have idiosyncratic meanings, and their constituents are



often not found in other morphological contexts. WORD-compounds have compositional meanings, and are formed from constituents which are commonly attested in a range of morphological environments.

I now turn to the characteristics of the apical contrast in words. The distribution of this contrast correlates with internal as well as external root and stem boundaries, and with the ROOT-level/WORD-level distinction in affixes. It is therefore an independent test of the morphological constituency proposed in preceding sections, though ROOT- and WORD-level *stems* (as opposed to affixes) are not distinguished by the distribution of apical distinctions. It will also be important for establishing the constituency of affixes in Ch 3, and the effect of glottal stops on neighbouring consonants, in Ch 5.

#### **2.4 Phonological correlates of compounding: the distribution of apical distinctions in words**

Apical consonants in Ngalakgan contrast for alveolar and postalveolar place of articulation. This contrast is neutralised in a range of positions. The contrast can only be distinguished after vowels in Ngalakgan, as in many languages (Hamilton 1996:48).<sup>27</sup> Root-internally, the contrast distinguishes minimal and near-minimal pairs:

- |      |    |          |         |                           |
|------|----|----------|---------|---------------------------|
| (76) | a. | /paŋarʔ/ | [baŋar] | 'marble tree'             |
|      | b. | /wanar/  | [wanar] | 'unidentified mammal sp.' |

In other positions, apicals either neutralise to postalveolar, or assimilate completely in place of articulation to a preceding apical. There are no words like those in (78).<sup>28</sup>

---

<sup>27</sup>Apicals can also be postalveolar after a vowel+glottal stop sequence both root-internally and heteromorphemically. The reason for this is discussed in §5.5.2. Butcher (1995), based on palatographic evidence from Warlpiri, Ngaanyatjarra, Murrinh-Patha, Guugu Yimidhirr, and Nyangumarta, has claimed that neutralised apicals are gesturally distinct from both contrastive apico-alveolars and apico-postalveolars. Rather, he claims, neutralised apicals are in fact articulated at a distinct place midway between the alveolar and post-alveolar regions, with an apical, rather than sublaminal, gesture. I have no data on this issue in Ngalakgan. But I note that Butcher only discusses the articulation of word-initial apicals in isolation; he does not discuss the effect of preceding vowels on the articulation of neutralised apicals.

<sup>28</sup>In roots, only homorganic nasal+stop clusters of apicals are permitted. This is despite the fact that heterorganic nasal+stop clusters (of coronal+peripheral place of articulation) are relatively common.

- (77) a. /kaŋ<sup>h</sup>ta/ [gaŋda] 'lower leg'  
b. /cakaŋ<sup>h</sup>ta/ [jaɡanda] 'female plains kangaroo'
- (78) a. \*/kaŋ<sup>h</sup>ta/ [gaŋta]  
b. \*/cakaŋ<sup>h</sup>ta/ [jaɡanda]

The contrast is neutralised word-initially. After a pause, apicals are alveolar, as in examples (79a-c). Following a vowel, apical-initial stems and words are postalveolar-initial, as in the prefixed examples in (79d-f). Square brackets enclose the phonetic realisation in each case.

- |      |                                              |                                                          |
|------|----------------------------------------------|----------------------------------------------------------|
| (79) | a. [n]aŋkuru<br>'saltwater crocodile'        | d. ŋu-ku-[ŋ]aŋkuru<br>I-FAM-crocodile<br>'the crocodile' |
|      | b. [n]occo<br>'grass' (generic term)         | e. ku-[ŋ]occo<br>IV-grass<br>'(some) grass'              |
|      | c. [n]e+p<br>burn+PP<br>'[someone] burnt it' | f. ŋu-ku-[ŋ]e+p<br>1mS-IV-burn+PP<br>'I burnt it'        |

I represent the initial-apical in words as underlyingly postalveolar (following Merlan 1983:9). A neutralisation of postalveolar to alveolar is easier to motivate on perceptual grounds. Since the primary cue for retroflexion is 'r-colouring' on the preceding vowel (Hamilton 1996:48), in the absence of a preceding vowel, postalveolars are difficult to distinguish from alveolars. Neutralisation of the apical contrast in word-initial position is phonetically-grounded therefore. There is no such motivation for the opposite position: that underlying apico-alveolars are regularly realised as postalveolar following vowels.

The same distribution is observed in complex words, both at the ROOT- and WORD-levels. An apical-initial stem is postalveolar when it follows a vowel or

There are no morpheme-internal homorganic liquid+stop clusters in the language.

vowel+glottal stop, (80a-c), (81a) but alveolar when it follows a (non-postalveolar) consonant (80d), (81b).

- (80) a. cu-ŋamulu-kerŋe-[ŋ]a+na  
[juŋàmulugèrŋeŋána]  
2mS-really-body-see+FUT  
'you'll have to really look for an animal'
- d. ŋu-pu-ŋic+[ŋ]a+na-ppira?  
[ŋubuŋicnanappíra]  
1mS-3a-stare.at+[see+]FUT-DU  
'I'll stare at those two'
- b. ŋu-ku-wi+[ŋ]a+n  
[ŋuguwíŋan]  
1mS-IV-forget+[see+]PR  
'I forget it [name: IV]'
- c. pur-koʔ+[ŋ]a+n  
[burgóŋan]  
3aS-have+[see+]PR  
'they have some'
- (81) a. ku-mili-[ŋ]e+ŋ  
[gumĩliŋéŋ]  
IV-water-burn+PP  
'he boiled the water'
- b. yir-puyppuy+[ŋ]e+ŋ<sup>29</sup>  
[yirbùippuínéŋ]  
1aS-burn.hair-burn+PP  
'we singed the hair'

Note that the alveolar/postalveolar distinction does not correlate with WORD-level/ROOT-level distinctions in the case of stems, since both WORD-level (80a) and ROOT-level compounds (80b-c) behave in the same way.

Morpheme-initial apicals assimilate to the POA of preceding apicals; (82a-b) show assimilation to preceding alveolars, and (82c-d) to preceding post-alveolars.<sup>30</sup>

- Prefixed word
- (82) a. yir-[ŋ]e+ŋa  
12aS-burn+FUT  
'we'll cook it'
- c. ŋurun-[ŋ]e  
12aO-burn.PR  
'it burns us'
- WORD-compound
- b. yir-ɬalʔ-[ŋ]e+ŋa  
1aS-hair-burn+FUT  
'we have to burn the hair now'
- d. ŋu-ku-piŋ-[ŋ]a+na  
1mS-IV-rock-see+FUT  
'I can see the hill'

<sup>29</sup>I have confirmed this form on a spectrogram. The spectrogram shows a brief period of voicelessness between the diphthong [ui] and the following apical [n]. This indicates either that the palatal glide is a fricative in this position or that the nasal is prestopped. The diphthong shows no trace of depression in the third formant, which is the usual indicator of retroflexion, and I can hear none.

<sup>30</sup>There are no examples in my data of apical-initial stems following an apico-postalveolar-final stem (as against postalveolar-final prefixes, which are common). I regard this as a lacuna in the data, rather than an important fact. The example in (82d) is an invented one which is consistent with the facts otherwise.

In stems, then, apical distinctions are neutralised in initial position, with the unmarked realisation being postalveolar. The WORD-level and ROOT-level distinction is irrelevant in this case, stems at both levels behave in the same fashion.

The distribution of apical distinctions in words *does* distinguish ROOT-suffixes from WORD-suffixes, and WORD-suffixes from clitics and modifying prefixes. ROOT-suffixes have the only morpheme-initial apicals which do not show alternation between alveolar and postalveolar: they are always apico-alveolar (Merlan 1983:9).

- (83)
- |    |                      |                          |                         |
|----|----------------------|--------------------------|-------------------------|
| a. | pu+nɪp<br>hit+PC     | [bunɪp]<br>'was hitting' | *[bunɪp]                |
| b. | caŋa+na<br>stand+FUT | [jaŋana]<br>'will stand' | *[jaŋaŋa] <sup>31</sup> |

Apico-postalveolar realisations of verb ROOT-level suffixes are unattested.<sup>32</sup>

Several clitics and modifying prefixes are apical-initial. They behave in the same way as roots and stems: initial apicals are neutralised in favour of apico-postalveolar. In (84a), a prefix /ŋamulu-/ 'really' is apico-postalveolar initial following vowels, but apico-alveolar following a consonant (84b).

- (84)
- |    |                                                                                                                                                                    |              |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| a. | cu-[ŋ]amulu-kerŋe-ŋa+na<br>2mS-really-body-see+FUT<br><i>yu labta faindim prabli</i><br>'You'll have to really find something.'                                    | [2/9/97:1B]  |
| b. | macci      poŋewk-me+n    [ŋ]amulu    ku-ɲapo+n-ci?<br>indeed      bad-be+PR      really      IRR-go+PR-NEG<br>'He's bad [i.e. sick], he really can't get around.' | [31/5/96:1B] |

The examples in (85) show the same range of behaviour for an enclitic /ŋowi/ (3nfemDAT) 'his, her, its'. The initial apical is postalveolar following a vowel, as in (85a),

<sup>31</sup>The morpheme division in inflected roots is open to some dispute.

<sup>32</sup>There are roughly 32 verb roots, with 5 distinctive tense suffix categories (for most verbs) giving a total of c. 192 inflected forms. Of the 32 roots, 17 have a paradigm in which the Present, Future, Potential, and Past Continuous inflections are built on a suffix beginning, arguably, in an apico-alveolar nasal +n. Therefore, of 192 forms, 85 - or just under half - involve an affix which is apico-alveolar.

and alveolar following a non-postalveolar consonant, as in (85b). Enclitics behave like roots and stems with respect to apical neutralisation.

- (85) a. ku-ce=[**ŋ**]owi  
IV-nose=3nfemDAT  
'his nose'
- b. cu-palak=[**n**]owi  
II-WiMo=3nfemDAT  
'his mother-in-law'

While many ROOT-suffixes are apical-initial, there are no apical-initial WORD-suffixes at all (Merlan 1983:9). Therefore we find a three-way contrast in the distribution of apical distinctions in bound morphemes: ROOT-affixes are always apico-alveolar, clitics and modifying prefixes are apical-neutral, while WORD-suffixes do not allow apical-initial forms. The distribution of apicals in the various categories of bound morphemes is presented in table (86). The number on the left in each cell represents the number of tokens in each morphological category ('ROOT-suffix' etc) which have an initial segment of the type given in each column, over the number of entire tokens in the category.

(86)

|                               | Apico-alveolar | Apico-postalveolar<br>(neutralised) | Non-apical |
|-------------------------------|----------------|-------------------------------------|------------|
| ROOT-suffix                   | 85/192         | 0/192                               | 107/192    |
| WORD-suffix                   | 0/9            | 0/9                                 | 9/9        |
| WORD-prefix<br>(Inflectional) | 0/30           | 5/30                                | 25/30      |
| Enclitic                      | 0/16           | 2/16                                | 14/16      |
| Modifying prefix              | 0/17           | 1/17                                | 16/17      |

The neutralisation of apical distinctions at morpheme boundaries provides additional phonological evidence of internal morphological constituency in complex words, and of the WORD-level/ROOT-level distinction in affixes.<sup>33</sup>

<sup>33</sup> Heath (1976) uses differences in the distribution of apical contrasts in initial positions of morphemes as a heuristic test of morphological constituency in Ngandi and Nunggubuyu.

## **2.5 Conclusion**

This chapter has shown that the prosodic characteristics of morphologically-simple surface words are also true of MWds: the WORD-level constituents of compounds and affixed forms. By contrast, morphological constituents which are not MWds do not have a special relationship to prosodic structure: forms which are complex only at the ROOT-level are treated in the same way as simple words.

In the next chapter, I consider the prosodic characteristics of WORD-level affixes and clitics. These provide evidence that a relationship between all WORD-level morphemes and prosodic structure is general to the language.