Big words, small phrases: Mismatches between pause units and the polysynthetic word in Dalabon[1]

NICHOLAS EVANS, JANET FLETCHER, AND BELINDA ROSS

Abstract

This article uses instrumental data from natural speech to examine the phenomenon of pause placement within the verbal word in Dalabon, a polysynthetic Australian language of Arnhem Land. Though the phenomenon is incipient and in two sample texts occurs in only around 4% of verbs, there are clear possibilities for interrupting the grammatical word by pause after the pronominal prefix and some associated material at the left edge, though these within-word pauses are significantly shorter, on average, than those between words. Within-word pause placement is not random, but is restricted to certain affix boundaries; it requires that the paused-after material be at least dimoraic, and that the remaining material in the verbal word be at least disyllabic. Bininj Gun-wok, another polysynthetic language closely related to Dalabon, does not allow pauses to interrupt the verbal word, and the Dalabon development appears to be tied up with certain morphological innovations that have increased the proportion of closed syllables in the pronominal prefix zone of the verb. Though only incipient and not yet phonologized, pause placement in Dalabon verbs suggests a phonology-driven route by which polysynthetic languages may ultimately become less morphologically complex by fracturing into smaller units.

1. Introduction

The last few years have witnessed a revival of interest in the problems of defining the unit word, and in particular in how far the units delineated by grammatical and phonological criteria match up. Polysynthetic languages pose particularly acute challenges because of the considerable size of the grammatical words involved.

Some recent studies of polysynthetic languages in North America have argued that, even though there may be good grounds for postulating large

grammatical words, particularly in the case of verbs, the best candidates for phonological words are considerably smaller, so that a single grammatical word may comprise more than one phonological word. As Russell (1999: 220) argues for Cree and Dakota, ''the most likely candidates for being phonological words are a fair bit smaller than the 'words' that have traditionally been assumed in the linguistic literature.[2] Indeed the traditionally defined 'word', if it is a phonological constituent at all, seems to be at the level of the phonological phrase.'' Rice (1993) on Slave, and Dyck (1994) on Cayuga, make similar points. In other polysynthetic languages there appears to be a good match between phonological and grammatical words, e.g., Chukchee, where the phonological word can be shown by vowel harmony to be coextensive with the grammatical word (Dunn 2001), or in Bininj Gun-wok, a classic polysynthetic language, with a clear correlation between the unit word defined by both grammatical and phonological criteria (Bishop 2003; Evans 2003a).

Until now we have little relevant data on this problem for the polysynthetic languages of northern Australia (though see Baker 1999), and practically no data from anywhere that incorporates detailed phonetic data on pause. In this article we examine this issue for Dalabon[3], a polysynthetic language that, like Bininj Gun-wok, is a member of the Gunwinyguan family, a non-Pama-Nyungan group of languages spoken in Arnhem Land, Australia.[4] Unlike in Bininj Gun-wok, the phonological status of verbal words in Dalabon is problematic. Although Dalabon is basically polysynthetic, the situation is rendered more complex grammatically by the existence of coexisting alternatives at the left edge of the word, through the three alternative possibilities of representing object pronominals by prefix, proclitic or independent pronoun; the rather subtle factors governing the choice between these alternations are discussed in Section 4. Phonologically, difficulties arise because of the possibility of breaking up a single grammatical word into a number of phonological units, as defined by pause and the location of intonational accents; under certain conditions some of these detached units may then be re-attached to the preceding grammatical word. The conditions under which pausing within the verbal word is possible, and the phonetic details it manifests, constitute the main focus of this article.[5]

Diachronically, there are reasons to regard the possibility of breaking a single verbal word into a number of pause units as a Dalabon innovation. Comparison with other Gunwinyguan languages shows the situation in Bininj Gun-wok, rather than Dalabon, to be the norm and the more likely reconstructable state. Moreover, the emergence of subword pause units appears to be linked to a morphological innovation in Dalabon which has had important phonological consequences: the extension of a

codal glottal stop following pronominal prefixes to become the unmarked TAM value, signaling assertativity, rather than the marked type that it is in BGW, where it is confined to the much rarer ''immediate'' aspect (cf. [1a]–[1c]; orthographically, *h* indicates the glottal stop). A further development is the high frequency with which the sequential prefix -*yelûng*-'(and) then', often reduced to -*lng*- in Dalabon, fuses phonologically with the pronominal prefix that precedes it (2a). Again this creates a heavy syllable coda at the end of the foot containing the pronominal prefix, whereas in BGW the sequential prefix occurs in its full form -*weleng*-, leaving the foot containing the pronominal prefix to end in an open syllable (2b).

(1)   a. *Nga-bu-n.*
  BGW 1sg.A/3sg.O-hit-NP
      'I hit (pres) / will hit him/her.'
     b. *Nga-h-bu-n.*
  BGW 1sg.A/3sg.O-IMM-hit-NP
      'I am hitting him/her right now.'
     c. *Nga-h-bu-n.*
  DAL 1sgA/3sgO-As-hit-PR
      'I hit him/her.' (Present)
      [Phonemically, prefixal sequence is /ŋaʔ/.]
(2)   a. *Bûka-lng-h-bong*
  DAL 3SgA/3sg.hi.O-SEQ-As-hit+PP
      '(S)he hit him/her then.'
      [Phonemically, prefixal sequence is /pɨkalŋʔ/.]
     b. *Bi-weleng-bom.*
  BGW 3SgA/3sg.hi.O-SEQ-hit+PP
      '(S)he hit him/her then.'

These innovations mean that the typical pronominal prefix in Dalabon, unlike in BGW, ends in a closed syllable. Moreover, since every possible morpheme following the pronominal prefix is consonant-initial, and since syllable onsets in Dalabon have one and only one consonant, it is not possible to resyllabify the syllable-closing consonant into the next syllable. Because terminating in a closed syllable is almost always a precondition for a grammatical prefix being assigned to a distinct phonological word (see Section 5), the generalization of glottal-closed syllables appears to have laid the foundation for the uncoupling of grammatical from phonological units in Dalabon. Nonetheless, pausing within the verbal word is not common, and we shall show in Section 5 that even within the phonological conditions that allow pause, this option is restricted to a relatively small percentage of text tokens, suggesting that it is an incipient

phenomenon, at phonetic level, that has yet to become generalized. We briefly relate the significance of this phonetic possibility to our understanding of the diachronic phonology of polysynthetic languages in Section 6.

2. General characteristics of Dalabon

Tables 1 and 2 show the vowel and consonant inventories of Dalabon, using both IPA symbols and the practical orthography.

Phonotactically, all syllables have structure $C_1 V$ (C_2) (C_3) (h); C_3 must have lower sonority than C_2. Normally stops are voiced in onsets and voiceless in codas. Long stops are only found stem-internally. There are only a few morphophonemic rules: underlying forms of morphemes virtually always surface directly, except for the optional reduction of -*yelûng*- to -*lng*- mentioned above, the reduction of *e* and *i* to *û* in some unstressed positions, and the conversion of sequences of identical stops into long stops within some compounds.

Lexical morphemes (roots) must be dimoraic; this results in the non-phonemic vowel lengthening of CV roots, even when compounded, suffixed or followed by an enclitic: e.g., /bono/ 'river', made up of /bo-/

Table 1. *Phonemic vowel system*

	Front	Central	Back
High	i	(û)	u
Mid	e		o
Low		a	

Table 2. *Consonant phoneme inventory*

		Place of articulation					
		Peripheral bilabial	velar	Apico-alveolar	postalveolar	Lamino-palatal	Glottal
Manner of articulation	Short stop	p (b)	k (k)	t (d)	ʈ (rd)	c (dj)	ʔ (h)
	Long stop	p: (bb)	k: (kk)	t: (dd)	ʈ: (rdd)	c: (djdj)	
	Nasal	m (m)	ŋ (ng)	n (n)	ɳ (rn)	ɲ (nj)	
	Lateral			l (l)	ɭ (rl)		
	Rhotic			r (rr)	ɻ (r)		
	Semi-vowel	w (w)				j (y)	

'liquid' plus /-no/ 'his, her, its; part marker', is phonetically [bo:no],[6] showing that minimality constraints in terms of morae apply to the root rather than the phonological word.

Dalabon's polysynthetic characteristics can be illustrated with the following two examples from a mythological text. (3) illustrates the predominance of verbs in a typical passage, while (4) shows how verbal complexity is built up through a combination of reduplication (*murri-murridjka-*), noun incorporation (*borndok-murri-murridjka-*), argument affixation and prefixes marking cause (*buh-*), or sequence (*yelûng-*). (3) also illustrates the use of procliticized object pronouns (*bulu=*), while (4) illustrates the use of encliticized case markers (*=walûng* and *=kûn*). Note also that while there is limited suffixation for TAM, most of the morphology on verbs is prefixal.

(3)   ka-lng-yurdmi-nj   bulu=ka-h-yelûng-berrû-bawo-ng
       3sgS-SEQ-run-PP   3plO=3sgA-As-SEQ-many-leave-PP
       bala-buh-ngong-boyenj-ni-nj mahkih
       3plS-because-mob-big-be-PP because
       'He ran away then and left them all, because there were so many of
       them.'

(4)   Ka-h-yelûng-djed-djedm-inj   korlkkorlk-no
       3sgS-As-SEQ-ITER-make.new.one-PP   by.night-ADV
       yibungkarn-walûng=kûn   borndok-no=mûn
       himself-ABL=GEN   woomera-3sgPOSS=only
       bûla-buh-borndok-murri-murridjka-ng   mahkih.
       3plA(/3sgO)-because-woomera-ITER-break-PP   because
       'Then he made a new (woomera), (working) by himself through the
       night. because they had completely broken up his woomera.'

Turning to nominals, there is negligible prefixal morphology,[7] and limited suffixing, mainly confined to the possessor markers like *-ngan* 'my' and *-(r)no* 'his/her/its' and the feminine kin suffix *-djan*, as illustrated in (5a). Enclitics mark case relations and some other functions (e.g., 'only').[8] Possessive relationships are doubly marked, by possessor suffixes on the possessed noun and genitive enclitics on the possessor (5a), (5b). Such enclitics are positioned at the right edge of phrases.

(5)   a.   nah-ngan=kûn   wurlkûn-djan-rno
            mother-1sgPOSS=GEN   younger.sib-FEM-3sgPOSS
            'my mother's younger sister'
       b.   djikka-no   kûrdûkûrd-ngan=kûn
            breast-3sgPOSS   wife-1sgPOSS=GEN
            'my wife's breasts'

Certain case enclitics can also follow phrase-final verbs in adverbial sub-ordinate clauses;[9] (6) illustrates such a use with the locative clitic =*kah*; see also (7) below for an example with the genitive used in a purpose clause.

(6) "Nahda   nga-h-bo-niyan, kahnûn   na   kornbo
    this.way 1sg-As-go-FUT him   now   that.man
    nga-me-y=kah,   nge-y-na-rr-inj=kah"
    1sgA/3sgO-get-PP=LOC   1dis.duS-SUB-see-RR-PP=LOC
    ka-h-yin-inj.
    3sgS-As-say-PP
    '"I'll go this way, to where I picked that feller up, to where we two
    saw each other", he said.'

More rarely, possessor markers can be encliticized directly after verbs; this happens when they qualify an incorporated nominal. An example is =*ngan* 'my' in (7), which modifies incorporated -*mele*- 'swag':

(7) kardûko   ngayh-mele-monwo-yan=ngan,
    maybe   1sgA/3sgO:HORT-swag-prepare-FUT=my
    ngayh-yong-iyan=kûn
    1sgS:HORT-sleep-FUT=GEN
    'I better make up my swag well (since it's cold), so I can sleep.'

Morpheme boundaries coincide with syllable boundaries except in two parts of the grammatical word: (a) tense/aspect/mood suffixes, which may form rhymes or codas added to C(V) initials drawn from the lexical stem (b) the right edge of the pronominal prefix, where assertative, sequential and subordinating morphemes may be part of the rhyme, as exemplified in (2b). Significantly, both these directly precede normal or potential phonological word boundaries.

3.   The grammatical word

Grammatical words are by and large easy to identify in Dalabon, with the exception of certain problems surrounding clitics which we treat in Section 4. The first two primary tests for grammatical wordhood identified in Dixon and Aikhenvald (2002: 19) can be applied straightforwardly:

  (a)   the elements within a grammatical word always occur together,
        i.e., although the free word order of Dalabon means there is
        great variation in where grammatical words go with respect to

one another, the constituent elements of these words always cohere, and

(b) the constituent elements of the grammatical word occur in a fixed order. For each of the major word classes (verbs, nouns and adjectives) a morphological template can be given, with strictly identified slots in a fixed order. A simplified version of the verbal template is given in Figure 1.

The claim of fixed order is subject to two caveats.

Firstly, the position of the sequential morpheme shows some variability. In its full form *-yelûng-* it follows the assertative prefix *-h-* (see [9],

Position	Slot
=	Object proclitic
−12	Subject(/Object)
(−11)	Subordinate
(−10)	Sequential
(−9)	Status
(−8)	Cause
(−7)*	Various adverbial
(−6)	Benefactive/Instrumental
(−5)	Various adverbial
(−4)	Incorporated generic nominal
(−3)	Incorporated body part nominal
(−2)	Numerospatial
(−1)	Comitative
(E−4)	(Embedded verb) Incorporated generic nominal
(E−0)	Stem [open]
(E+3)	Incorporated verb form
0	Stem [open]
(+1)	Reflexive/reciprocal
+2	Tense/aspect/mood
=	Case enclitic

Figure 1. *Morphological template for the verb in Dalabon*

Optional slots are shown in brackets. Potentially fused segments are shaded together. To the left are object proclitics and to the right are case enclitics. At most one verb may be embedded in another, provided that the main verb belongs to a small host class including *bon* 'go' and *wo* 'give, cause'. The subject pronominal prefix is shown here as a single unit, but may itself be morphologically complex, e.g., *burra-* '3duA', decomposable into 3nsg *b*, transitive subject *u*, and dual *rr(a)-*. Stems may likewise be morphologically complex (e.g., *murridjka-* 'break (tr.)' splits into prepound *murridj-* 'break' plus *ka-* 'carry'), and so may be incorporated body parts, which may be compounds of two body part roots. Other slots are monomorphemic.

[17], [24] for examples), and when the 'just' adverbial prefix -*dja*- is pres-
ent it follows this (9). However in its reduced form -*lng*- it normally
moves left of the assertative-marking glottal stop, though the fact that
glottal closure is an independent gesture means that there is some pho-
netic variation in the timing of these two segments, ranging from a stop
following the *lng* sequence, to concurrent partial glottal constriction in
the most reduced form of realization of the *h*.[10] In our data we have gen-
erally transcribed the assertative morphemes as occurring after the short-
ened sequential morpheme. In the written text considered in Section 5.2
there are also examples of sequential -*lng*- being placed after benefactive
-*marnû*-, merging with it to give the sequence -*malng*-.

Secondly, there are a number of alternate constructions whose main
difference lies in whether an element occurs internally to the verb, exter-
nally, or both. The commonest is incorporation: alongside verbs in which
absolutive nominal roots are incorporated (such as *bûla-buh-borndok-*
murri-murridjkang 'because they had completely broken up [his] woo-
mera' in [4]), are virtually synonymous alternatives in which the nominal
root is not incorporated at all. It then appears externally, in which case it
may bear possessive suffixation such as third person singular possessive
-*no*, e.g., *borndok-no bûla-buh-murri-murridjkang*. It is also possible for
the nominal root to be doubled, occurring both internally and externally,
which is actually the case in (4).

The second type of alternate construction is prepound-extraction. Some
verb stems are morphologically complex, e.g., *dukka* 'tie', comprising the
verb root *ka* 'carry' plus prepound *duk*. The *duk* in *dukka* occurs nowhere
else, and there are many other examples like this (e.g., *wo* in *wona*
'hear' < *na* 'see'). However, a subset of prepounds are of an ideophonic
character and may occur either as part of a verb stem (e.g., *dadjka*- 'cut'
or *walkka*- 'hide') or independently (e.g., *dadj!* '[someone] cut [some-
thing]'). Now just in the case of ideophonic prepounds, there are alterna-
tive constructions: a single integrated verb may be used (e.g., *kalngwalk-*
karrinj 'then [s]he hid him/herself'), or the ideophonic prepound may be
extracted and placed immediately before the verbal word:

(8) Yulyul   walk   ka-lng-ka-rr-inj
     run.crouching.under.cover   hide   3sgS-SEQ-take-RR-PP
     'He ran along crouching and hid himself.'

A more complex variant occurs when what is extracted is not just the
ideophonic prepound, but also a preceding incorporated nominal. (9)
gives an example where such a sequence is given first in extracted form,
then repeated inside the verbal word:

(9) Djirrh   ngûrh   bûka-h-yelûng-ngurl-dulubo-ng.
     Right   life.force 3sgA/3sg.hiO-As-SEQ-heart-spear-PP
     ngurl-wirb   bûka-h-ngurl-wirbme
     heart-rip   3sgA/3sg.hiO-As-heart-ripPRES
     bûka-h-dja-lng-kom-dengkohm-inj.
     3sgA/3sg.hiO-As-just-SEQ-neck-knock-PP
     'And he speared him then right where his life-force was, right in the
     heart. Heart-rip! He ripped his heart out and knocked him in the
     back of the neck.'

Clearly these possibilities do not create alternative orderings *within the word* — rather they offer two alternative constructions, in one of which the relevant material occurs within the verbal root, and in the other of which it doesn't.

The criterion of "conventionalized coherence," the third criterion in Dixon and Aikhenvald (2002: 20), indicates "that the speakers of a language think of a word as having its own coherence and meaning. That is, they may talk about a word (but are unlikely to talk about a morpheme)." Now it is generally true that it is the entire grammatical word that Dalabon speakers take as the unit for metalinguistic discussions. Maggie Tukumba's definitions of verbs for the Dalabon dictionary, for example, are always given using fully inflected verbs. In this sense the "conventionalized coherence" criterion holds straightforwardly. But there is a stronger sense of "conventionalized coherence" that does not necessarily hold for Dalabon. It has been asserted for some polysynthetic languages, such as Mohawk, that speakers remember exactly, in gestaltlike form, which words they have heard before, without being aware of their component parts (Mithun 1998: 178). However, Dalabon speakers (even when nonliterate) are typically at ease with partial segmentations of complex words. Some of these segmentations involve the breaking off of the pronominal prefix complex, and ideophones (see below), but others interrupt single phonological words — for example speakers will isolate incorporated nominal roots for discussion.[11]

There are two subparts of the grammatical word, on the other hand, which will each never be broken up. These are (a) the sequence taking in the pronominal prefix plus sequential, causal, assertative and subordinate markers on the one hand (the "pronominal prefix zone"), and the sequence taking in the comitative applicative, root, reflexive/reciprocal and TAM suffixes on the other (the "root zone").

This suggests that what we are calling the 'verbal word' is made up of a number of "coherence zones", which are chunked together independently before being assembled. Note also that there are no dependencies, either

phonological or grammatical, between these two parts of the polysynthetic word, so logically it makes sense to process them separately — the only exception is the fact that whether to choose the transitive or intransitive prefix set depends on the valence determined by the stem and comitative applicative in the root zone.

Two other criteria sometimes used to identify grammatical words deserve brief comment.

Firstly, completeness of utterance. It is clear that, at least for the major word classes of verb, noun and adjective, grammatical words can form complete utterances. Verbs regularly form complete predications, supplied as they are with argument information through the system of pronominal prefixes and proclitics. Nouns may form complete utterances in context, for example in answer to the question ''who is it?'' one may reply *nah-ngan* [mother-my] '(it's) my mother'. Adjectives may be used as predicates, in which case they take the subject prefix set: examples are *ka-h-weh-no* [3sg-As-bad-ADJ] 'it's bad' and *dji-moyh-no* [2sg.APPR-sick-ADJ] 'you might get sick'. Subject prefixes are also sometimes used with nominals, if they refer to social category terms, e.g., *barra-h-duwa* [3du-As-*duwa* (a patrimoiety name)] 'the two of them are of Duwa patrimoiety'.

Secondly, positioning of inflectional material at the boundaries of grammatical words. This is clearly the case in Dalabon. Within the verb, the complex morphological sequence with its many derivational slots is bracketed by inflectional slots at each edge — the pronominal prefixes at the left edge, and the TAM suffixes at the right edge, these being the only obligatory affixes in the sense that every verb must have an exponent of these slots. Within the noun, possessive suffixes, which are obligatory if the possessive relationship is being expressed, are located at the right edge, except for enclitics.

4.   Clitics

There are a number of clitic types in Dalabon, and we only deal with some of them in this article. Essentially, clitics in Dalabon are positioned with respect to a phrasal or clausal constituent. For example, case-like relational enclitics may be positioned at the end of a NP ([5a], [5b]), or after a clause whose relation to another clause they are signaling, which typically ends in a verb ([6], [7]). Possessor markers, though normally suffixes, may be positioned after verbs that have incorporated the possessed noun (if no other exponent of the possessed NP is available), as in (7). In each case, clitics are phonologically integrated with the word they end up positioned next to; if disyllabic (e.g., ablative =*walûng* or 3plO *bulu*=) they

normally take secondary stress on their first syllable, but they cannot take primary stress. But when a string of one or more enclitics reaches or exceeds two syllables, this sequence may form its own phonological word. An example is the sequence =*walûng*=*kûn* [=ABL=GEN], which can be separated off into a separate word, e.g., *ngey wálûng*=*kûn* 'me, by myself'. Since such divisions into several phonological words can affect single grammatical words as well — see Section 5 — this does not count as a special property of clitics. Enclitics also participate in various optional phonological processes found within the phonological word, such as optional palatalization of *k* to *y* after *i* or *y* (e.g., *dubmi*=*kah* 'today=LOC', [dupmijaʔ]), and palatal glides spreading leftwards through glottal stops (e.g., *kah-nguh-yu* '3sg-shit-lies' > [kaʔŋujʔju] 'there is shit there').

We now turn to focus on the most interesting case, that of object proclitics to the verb. Pronominal object encoding is complex in Dalabon. For some subject/object combinations there are simply special portmanteau prefix forms, e.g., *djila-* '3pl.A > 2sgO; 1pl.A > 2sgO'; *ka-* '3sgA > 1sgO; 2sgA > 1sgO', which pattern grammatically exactly like regular prefixes. The second of these (*ka-*) also occurs as in intransitive prefix form, with the meaning '3sg subject', but the first (*djila-*) only occurs in transitive combinations.

For most subject/object combinations, however, only the subject is represented by an unambiguous prefix, and objects are represented by forms that allow alternative realizations as (a) free pronouns positioned to the left (not necessarily immediately) of the verbal word (b) proclitics positioned immediately to the left of the verbal word (c) a reduced form of the pronominal clitic, which may drop its final vowel and merge with the subject prefix.

Previous materials on Dalabon have differed in how they represent object pronominals. Capell (1962: 117) wrote them as separate words, whereas Alpher (1982: 117) and Merlan (1993) wrote them as part of the same word as the rest of the verb. Their representations are reproduced here as the top lines of (10)–(12) respectively, followed by renditions in the current practical orthography. Note that Capell wrote the glottal stop as "ʔ", and Merlan as an apostrophe, Merlan writes the palatal nasal as *yn* and long stops with single voiceless symbols, and Alpher uses voiceless stop symbols syllable-finally and voiced symbols syllable-initially.

(10) bulu   bila-ʔ-na-n
      bulu   bila-h-na-n
      3plO   3plA-As-see-PRES
      'They see them.'
      (Capell 1962: 117)

(11) bat-go   burnu-nga-h-meyi
      bad-ko   bunu nga-h-meyi
      stone-DYAD3duO-1sgA-As-getPP
      'I got two stones.'
      (Alpher 1982: 29)
(12) bulu-bula-'-ngapuniyn
      bulubula-h-ngabbuniny
      3plO-3plA-As-givePP
      'They gave it to them.'
      (Merlan 1993)

These representational differences reflect selections from real variation in how pronominal objects behave. In examples we have transcribed that are equivalent to those represented by (10) to (12), the object pronoun takes its own secondary stress but is rhythmically integrated with the following verb, whose prefix also takes its own primary stress, e.g., *bùlu-búlah-nan* for (10). We analyze such cases as proclitics. However, in addition to this possibility there are both less and more integrated alternatives.

   Object pronominals, though they must lie left of the verb, can be separated from it by an intervening word such as *wanjh* 'already' in (13), an emphatic subject pronoun like *ngey* 'I' in (14), or *mak* 'not', also in (14); in such cases they may take their own primary stress. (Note that the 3du form varies between *bunu*, *burnu* and *bulno*; the last form only occurs when it is separated from the verb, but the others can occur separated or procliticized).

(13) rolu bunu   wanjh   ka-h-yelûng-banj
      dog 3duO   already   3sgA-As-next-bitePP
      'The dog has already bitten them.'
(14) bulno   ngey nga-h-nanhna-n,   yala-h-ni
      3duO   1sg 1sgA-As-look.after-PR plS-As-sitPR
      wadda=kah,   nidjarra marrmo=njelng,   mak
      country=LOC   this.way clan=1plPOSS   not
      kini=kah   marrmo-kinikin=kah, bulno   mak
      different=LOC   clan-different=LOC   3duO   not
      nga-munkuyung
      1sgA-sendPR
      'I look after them two, even though we're sitting on our own land here, this is our clan land here, it doesn't belong to another clan, yet I don't send them away.'

On the other hand, the object marker can drop its final vowel and coalesce with the prefix into a single unit, with the second element losing

its stress (e.g., *bùlu dá-* > *búlda-* [2sgA > 3plO], *bùlu ká-* > *búlka-* [3sgA > 3plO]), *bùlu búla-* > *búlbula-* '3plA > 3plO'. An example is (15); the three dots indicate a pause.

(15)   bul-ka-lng-h-marnû-...   bad-modûk-wo-ng
        3plO-3sgA-SEQ-As-BEN- rock-closed-make-PP
        'Then he sealed them up inside the rock cave.'

To summarize, object pronouns present several options:

   (a)   for some combinations, they are obligatorily part of a portman-
         teau prefix
   (b)   for other combinations, they allow three constructional options:
         (i)   free pronoun positioned to the left of the verb ([13]–[14])
         (ii) pronoun procliticized directly to the left of the verb ([10]–
               [12])
         (iii) reduced pronoun fully integrated into the verb (15)

Comparison of prefixal forms with the other Gunwinyguan languages suggests that (a) is the original situation, but that various neutralizations in the Dalabon prefixal paradigm have created a need, in the case of a range of subject/object combinations, for "patches" which have been filled by positioning free pronouns to the left of the verb. These are in the process of grammaticalizing back into prefixes. But the process is not advanced and several alternative constructions are currently available.[12]

5.   Problems in delineating the phonological word

Dalabon is not a language with prosodically scoped morphophonemic processes that allow us to define a phonologically word clearly, such as vowel harmony, assimilatory or dissimilatory alternations — morphemes are simply assembled together without modification. Nor are there clear phonotactic differences between phonological word edges and syllable edges, which are nearly identical. The only exceptions are that the trilled *r* (orthographically *rr*) and the long stops can only occur word-internally, but unfortunately since they are also debarred from occurring morpheme-initially, and the initials of the problematic units we will be dealing with always align with morpheme breaks, this cannot be used as a test where we want it. As mentioned in Section 2, there is a bimoraic minimum on roots, leading to nonphonemic lengthening of monosyllabic CV roots, but this applies to them qua morphemes, not words.

   In this respect Dalabon contrasts with some other polysynthetic languages, such as Chukchi (Dunn 2001), or Cree (Russell 1999), in which morphophonemic processes allow the analyst to clearly delimit

phonological words — in Chukchi, coextensive with the grammatical word, and in Cree fitting inside it. Stress, though assigned over the whole grammatical word, takes the form of a number of stressed foot-initial syllables, whose relative prominence is determined by higher level intonational factors rather than positioning with respect to the word boundary, so one cannot use location of ''primary stress'' as a way of locating phonological word boundaries. There are a few fast-speech assimilations, such as palatalization, palatal glide spreading, and nasal devoicing, but since these only occur within domains that are defined by pause and rhythm, they are not primary means of investigating phonological boundaries.

Hall (1999) mentions three types of evidence bearing on the domain of phonological words — phonological rules, phonotactic generalizations, and minimality constraints (e.g., number of syllables or morae in a word). As the remarks in the previous paragraph indicate, none of these types of evidence help us delineate a phonological word in Dalabon, since there are no significant morphophonemic rules, phonotactic generalizations apply at the syllable and morpheme levels rather than the word, and the minimality constraint applies to the unit ''root'' rather than to the phonological word.[13]

The status of pause in determining phonological word boundaries is controversial, and neither Hall (1999) nor Dixon and Aikhenvald (2002) treat it as a valid diagnostic.[14] Nonetheless, it is analytically easy to identify and measure (Section 5), and we can expect that it is within pause-delimited units that the stress and fast-speech assimilations which diachronically engender phonological rules take place. In what follows we use ''...'' to denote a pause in our examples, supplemented by Break Index values of 2, 3, or 4 in the examples for which we supply transcriptions in the TOBI [Tone and Break Index] framework (see Section 5.4). Most examples are based on analysis of naturalistic data, though several are based on dictations made to the first author by the late Peter Mandeberru (who was not literate).

In Section 5.1 we give an initial characterization of the conditions that allow a single grammatical word to be broken up by pause(s) into two or more phonological words, or that allow prefixes of one or two syllables to be grouped with the preceding grammatical word. We expand upon this evidence in subsequent sections, bringing in orthographic evidence in Section 5.2, noting correlations with intonational patterns in Section 5.3, and then making a more thorough quantitative study of a sample of detached prefixes from two subcorpora in Section 5.4. Since intraword pausing is a possibility rather than a requirement, and it is common for even quite long verbs to be realized as a single phonological unit, in Section 5.5 we give figures on the actual frequency of intraword pausing in two texts.

5.1. *Intraword pauses: an initial characterization*

The two relevant phonological requirements that must be met before pause can occur are that:

(a) this must not split a foot, though since feet are constructed over morphemes in virtually all cases,[15] this could also be phrased as a morphological requirement. Because feet have a bimoraic minimum, this effectively means that the prefix must be closed, if monosyllabic, and otherwise satisfy a disyllabic minimum

(b) the remainder of the word must have at least two syllables.

Two initial examples from dictated speech[16] are (16), illustrating pause after the pronominal prefix and then again after an incorporated nominal, and (17), illustrating pause after the pronominal prefix and then after an incorporated adverbial. Note that in each case the paused-after phonological unit is closed (by *h*, *k* or *rrk*); somewhat unusually all are monosyllables in these examples.

(16) ka-h-...rak-...m-iyan
      3sgA/3sg.lo.O-As-wood-get-FUT
      'He ... will get ... firewood.'

(17) dje-h-... djarrk-...ning-iyan
      12disS-As-together-sit-FUT
      'We (disharmonic) will sit together.'

In spontaneous speech it is common to find pausing within grammatical words, regulated by the phonological and morphological principles just given. There are three common reasons for this happening:

(a) self-correction, e.g., after using the wrong prefix. The self-repair interjection *kenh!* 'Oops, I mean' is often added right after the pause, as in (18):

(18) dubmi   wungurr-no   bûla-lng-h-... kenh!
      Now   shadow-3sgPOSS   3pl-SEQ-As-...oops
      'Now his shadow they ... I mean ...' [goes on to correct to 'he']

(b) suspense or continuation, typically accompanied by a high edge tone just before the pause and always including the sequential marker *-lng-*; this high edge tone is regularly associated with suspense, as in (19) and (20), or continuation of a request following a slight pause signaling deferral to the hearer, as in (21).[17] Following the pause, the prefix may be repeated, as in (19) and (20), or the speaker may simply proceed to the rest of the verb (21).

(19)　kenbo　bûka-lng-h-ka-ka-ng　　　　　　　　　bala-lng-h-...
　　　later　3sgA/3sg.hi.O-SEQ-As-ITER-take-PP　3plS-SEQ-As-
　　　bala-lng-h-njong-buddi-nj
　　　3plS-SEQ-As-many-be-PP
　　　[A *mimih* spirit tricks a hunter into going with him, stating that he
　　　is on his own and therefore doesn't present any danger, to a cave
　　　where the *mimih*'s allies are lying in wait. Here the pause precedes
　　　the moment at which the *mimih*'s full forces are revealed.]
　　　'Later he took him (to his place) and then they ... there were
　　　many (mimihs) there.'

(20)　[This section describes the hunter's revenge, and the pause
　　　precedes his administration of the coup de grace, that will kill the
　　　mimih by knocking the back of his neck.]
　　　bûka-h-yelûng-...　　　　　bûka-lng-h-...dengkohm-inj
　　　3sgA/3sg.hi.O-As-SEQ-　3sgA/3sg.hi.O-SEQ-As-knock-PP
　　　'Then he ... he knocked him out.'

(21)　[The speaker is suggesting to his companion that they go out from
　　　the windbreak, now it is dry; because the two are in a
　　　'disharmonic' kin relation it is appropriate to speak in a somewhat
　　　tentative style in making suggestions.]
　　　dje-h-lng-...djarrk-bulhm-iyan
　　　12disS-As-SEQ-together-come.out-FUT
　　　'Then we (disharmonic) will ... come out together.'

　　(c)　the possibility of extracting ideophonelike prepounds was men-
　　　　tioned in Section 3. A further possibility, sometimes used for dra-
　　　　matic effect, is to iterate them in place, i.e., within their regular
　　　　position inside the grammatical word. In such cases each iteration
　　　　has its own intonation contour (a high level contour, in this case
　　　　— see [25] for pitch trace), and all but the last may be followed
　　　　by pause:

(22)　ka-h-dja-berrûh-...　　　berrûh-...　　berrûh-m-inj
　　　3sgS-As-just-come.out-. .come.out-...come.out-VZR-PP
　　　'Then he edged out. .out...out.'
　　　[Context: the protagonist is waiting for his *mimih* enemy to come
　　　far enough out of a rock crevice to spear him, but the *mimih* is
　　　only inching out bit by bit.]

In the examples given so far, the paused-after prefix has formed its own
phonological unit. However, it is also possible to merge it with the pre-
ceding phonological unit, in the sense that there is no pause between
them. This is shown in our transcription by a '_' between the joined

grammatical words, as in (23) and (24); for typographical reasons we sometimes extend this underscore so that we can continue to align the morpheme glosses with the start of grammatical words. Where a paused-after prefix attaches to just one grammatical word, that word must be minimally dimoraic, i.e., either disyllabic or a heavy monosyllable. This produces a mismatch between the grammatical word boundaries, shown by '|w', and the phonological word boundaries, shown by '|ω'.

(23) Grammatical |w |w |w
      boundary:
      Phonological |ω |ω |ω
      boundary:
                    Dorrng-no-duninj__ka-lng-h-. . .dorrng-bulhm-inj.
                    body-3POSS-real   3sgS-SEQ-As-body-appear-PP
                    'Then his actual body (i.e., not his shadow)
                    appeared'
(24) Gr. |w |w |w |w |w
      boundary:
      Ph. |ω |ω |ω |ω |ω
      Boundary:
                    Woy!   Djulu-kah djarra_ye-h-. . .djarrk-ni-ngiyan.
                    Come! Fire-LOC here   1inc-As-together-sit-FUT
                    'Come over here and we'll sit together by the fire.'

Summarizing so far, it is possible to interrupt the grammatical word by pauses in the prefix series, provided that the resultant prefix has a bi-moraic minimum, and this is done for a variety of communicative goals[18] including dictation, self-correction, suspense and dramatic itera-tion of ideophonic prepounds. The resultant units, whose domain is smaller than the grammatical word, may either form free-standing pause units of their own, or, be joined on to the preceding phonological word.[19]

5.2.   *Intraword pauses: orthographic evidence*

Here it is worth commenting briefly on how verbal words are represented in Dalabon orthography. This has only been developed in the last couple of decades, in connection with basic teaching of the language in schools, and the few passages of written Dalabon have been produced by speakers for whom Dalabon is not their dominant language (typically it is an L2 or L3 after Kriol and Aboriginal English). Moreover, the conventions for writing word-boundaries have not stabilized and existing texts typically exhibit considerable variation in the degree to which word spaces and hy-phens are used to break up single grammatical words; we also cannot rule

Table 3. *Examples of various orthographic representations of phonological units in the written text "The brolga and the quai" by David Jentian Nangan:golod. (In the orthography he uses, v = û, and both g and k represent k.)*

Strategy	Example (as written)	Translation	Morphemic division	Interlinear gloss
Spaces	*narrah durrun*	'you two argue'	*narrah-du-rr-un*	2duS-swear.at-RR-PR
Hyphen or underscore	*buga_marnu_yininj*	'he said to him'	*bûka-marnû-yin-inj*	3sgA/3sg.hiO-BEN-say-PP
Hyphen/space	*ngarra-dja bon*	'let's just go'	*ngarra-dja-bo-n*	12plS-just-go-PR
Joined/underscore	*gahmalng_yininj*	'then he said'	*ka-h-ma-lng-yin-inj*	3A/3O-R-BEN-SEQ-say-PP
Joined/space	*ngarralung bon*	'let's go then'	*ngarra-lung-bo-n*	12plS-SEQ-go-PR

out some influence from English orthography which is the dominant language in which the author of the texts is used to writing.

Nonetheless, it is striking how often single grammatical words are broken up by one or more word spaces. To consider the representation of verbs in a single such text (Nangan:golod 1975), of 30 verbs none are written without any indication of separation of components; 16 are written with word spaces between prefixes and verb stem; 6 separate the prefix from the verb stem by a dash ("-" or "_"), five mix spaces and hyphens (with the space always preceding the verb stem, and hyphens used between different prefixes), one joins the prefixes en bloc then joins it to the verb by an underscore, and two write the prefixes en bloc then leave a space before the verb stem. Table 3 gives one example of each type.

While it would be naive to take this evidence too seriously, given the lack of an established orthographic tradition, these figures nonetheless suggest that Dalabon speakers feel that verbal grammatical words are made up of several phonological units, and are employing spaces to indicate this. The contrast with Kunwinjku is striking — each of the words given in Table 3 would be written en bloc in Kunwinjku (respectively *ngurridurren, kabimarneyimeng, karridjalre, kabimarneyimeng,* and *karri-welengre,* with identical morpheme structure to the Dalabon forms). Admittedly, Kunwinjku has a more established orthographic tradition, its orthography being learned in school by L1 speakers who have produced some substantial written materials, so we cannot discount the possibility that the difference simply reflects the difficulties faced in writing long words by speakers first learning to write polysynthetic languages.[20] Nonetheless, an equally plausible hypothesis is that the use of spaces and hyphens by Dalabon speakers shows an accurate insight into the phonological structure of their language, and that the difference from Kunwinjku

reflects the greater possibilities in Dalabon for splitting up comparable polysynthetic verbs into several prosodic units.[21]

5.3.   *Intraword pauses and intonational movements*

We now return to phonetic as opposed to orthographic arguments. So far the possibility of splitting up grammatical words into more than one phonological domain has been presented as simply a matter of pause units. However, it also a matter of intonation units. The pause-delineated units mentioned above also display pitch movements typical of certain types of independent intonational phrases. Although we lack the space to present this fully here, we comment briefly on two typical patterns, bringing in sound spectrographs and intonational transcriptions in the TOBI framework, which analyses intonational melodies into fixed points and edges — both "tones", between which pitch movements are projected as transitions. See Ladd (1996) on the overall framework, and Fletcher and Evans (2000), Bishop (2003) and Bishop and Fletcher (2005) for an example of its application to Bininj Gun-wok.

The two commonest intonational patterns associated with pauses after material inside the verbal word are:

(a)   pattern we transcribe as H* !H%, a high contour with a slight drop at the right edge (indicated by the downstepped high symbol "!H" preceding the edge symbol "%"), and characteristically associated with suspense. This is exemplified in (25), derived from (21).

(25)   speech waveform, RMS amplitude trace and F_0 contour illustrating (a)
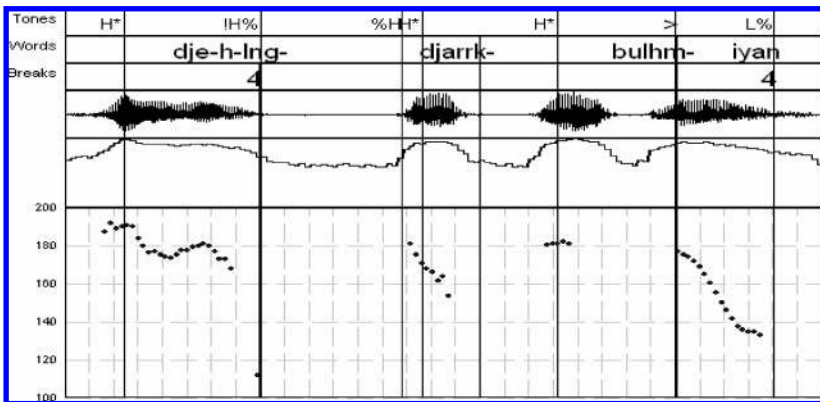
(b)   a pattern we transcribe as H*H%, a high contour without significant final drop at the right edge (which stays high, indicated by the edge symbol "H%"). This contour is characteristically associated with continuation of activity. An example is (26), derived from (22) above, which used an iterated ideophonic prepound to describe the continued step-by-step emergence of a targeted victim from a rock crevice.

(26)   Speech waveform, RMS amplitude trace and F_0 contour illustrating (b)
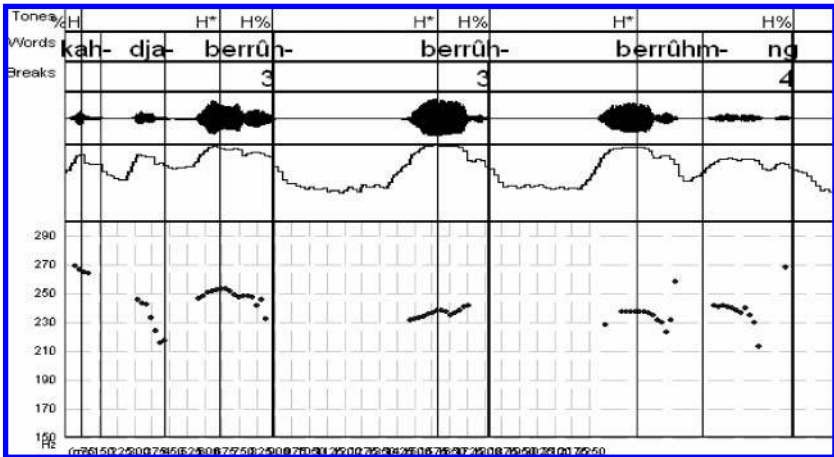
The intonational phrases on intraverbal prosodic units that we have been able to transcribe so far only represent a subset of the set of melody types, in particular those melody types occupying nonfinal positions in phonological utterances. The lowest edge tones they exhibit are the 'downstepped high', transcribed by !H% (which is the second highest in a series of four taking in H, !H, L and final L); this is consistent with their nonfinal nature, since the latter two edge tones are limited to clause and paragraph boundaries. However, outside the phenomenon of prosodically detached prefixes being considered here, H% and !H% edge tones are found in a number of other continuation and suspense type constructions associated with multi-word and multiphrase constructions, so there is certainly no intrinsic link between these intonational melodies and detached prefixes (or parts of words). Rather, we are dealing with the association of intonational melodies, normally associated with phrasal and clausal units, with smaller intraword units in particular discourse contexts.

5.4. *Intraword pauses: quantitative data*

We now pass to a more thorough quantitative study, summarizing the findings in Ross (2003), which investigated the nonisomorphism of the grammatical and phonological word of Dalabon through pause location and duration, and intonational contours. This section deals with examples that were preselected from the corpus by listening for instances of intraword pause; we hold off on giving actual frequency-of-occurrence data until the next section.

The initial data from which the pause tokens were selected consists of a corpus of thirty minutes of spontaneous speech from two different Dalabon speakers, which was digitized at 22.05 kHz using ESPS/Waves+ on a Sun Ultra workstation and labeled using the EMU Speech Database System. The digitized segments were labeled on four tiers; the utterance tier, the word tier, the break tier and the gloss tier. Statistics were extracted using the statistical package R version 1.6.2.

On the break tier breaks were labeled using modified ToBI (Tone and Break Index) transcription conventions (e.g., Beckman and Ayers Elam 1997; Fletcher and Evans 2002). Break indices were given on purely phonetic grounds, without reference to the morphosyntax of Dalabon. These are outlined below.

Juncture was determined on auditory grounds. The break index 1 was used to indicate a word boundary with a minimal degree of perceived juncture between a pair of words. The break index 3 was used to indicate a greater perceived juncture than a break index 1 but less juncture than a break index 4. The cues indicating a break index 3 juncture may consist of one or more of the following: a brief juncture of no more than 200 ms, lengthening or prominence of the final syllable, and pitch reset of the following constituent. For the purpose of this study 200 ms was considered the minimum length of a pause, as the phonemic inventory of Dalabon includes long voiceless stops (Fletcher and Evans 2002: 124), which might otherwise incorrectly be construed as a silent pause. We follow the usual definition of silent pause as any silent interval in a stretch of speech that cannot be attributed to a stop closure or any other phonetic segment. For speaker MT, long stops range in duration from 130 ms to 203 ms, with a mean duration of 163 ms.

A break index 4 is given to mark the highest degree of perceived juncture. The phonetic cue indicating a break 4 constituent is a pause of more than 200 ms. For purposes of convenience, a constituent marked by a break index 1 will be referred to as a word (a grammatical word that does not coincide with a tonally marked prosodic boundary), a constituent marked by a break index 3 will be referred to as an intonational

phrase, while a constituent marked by a break index 4 will be referred to as an utterance. We do not make use of a break index 2. This study is largely focusing on the role of pause in relation to word boundaries so it was felt that we did not need BI 2. This is in keeping with current applications of ToBI modeling (e.g., Jun 2005) which question the status of this break index. Moreover, most applications of ToBI to languages other than English use only BI 1, 3, and 4.

5.4.1.   *Distribution of break indices.*   The total distribution of break indices, determined by the procedures above, and without regard to grammatical environment, is shown in (27).

(27)   Distribution of break indices

Distribution of break indices
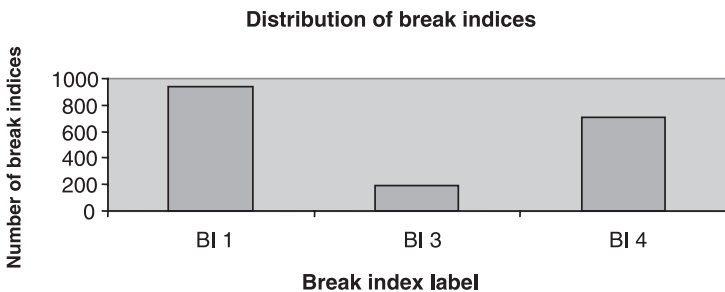
In the data there are 939 occurrences of the break index 1, 189 occurrences of the break index 3, and 698 occurrences of the break index 4. On average an intonational phrase or utterance contains approximately 1.9 grammatical words. Example (28) illustrates the distribution of sequences of break 1 indices per break 3 or 4 index.

(28)   Distribution of break index sequences

Break 1 per break 3/4 constituent
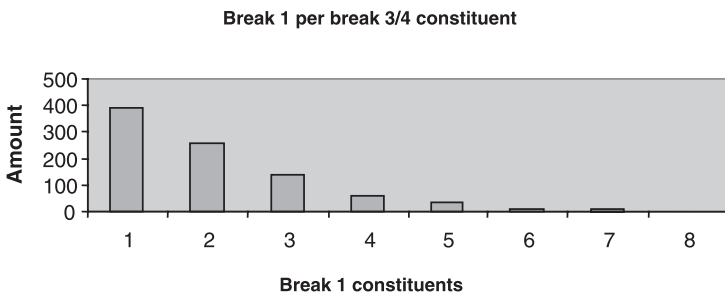
Example (28) reveals that an intonational phrase, as defined by a break index 3 or 4, contains minimally one word (with 386 occurrences in the

corpus) and maximally eight words (1 occurrence in the corpus). Thus a typical intonational phrase in Dalabon consists of very few words. These results support the findings of the closely related dialects of Bininj Gun-wok; Manyallaluk, Mayali, Kune and Kunwinjku which have a ratio of between 1.5 and 2.1 words per intonational phrase (Bishop and Fletcher 2005).

5.4.2. *Pause types sorted by grammatical environment.*   On the word tier of the transcription, silent pauses were divided into three categories and labeled accordingly. These were: standard pauses, disfluency pauses, and pauses within a word. Standard pauses are pauses that occur between grammatical units. Disfluency pauses occur in connection with repetition, deletion or substitution of units, such as pronominal prefixes, which should otherwise be attached to a following verbal unit. Pauses within a word are a clear example of an interruption of a grammatical word. As such, these are the clearest example of a nonisomorphism between a grammatical word and a phonological word. Disfluency pauses differ from this type of pause, as the unit preceding a pause is either repeated or changed to a different unit. The latter two categories of pause have in common that the preceding constituent must attach to another constitu-ent to form a complete grammatical word.

   A sample t-test on the durations of disfluency pauses revealed no signif-icant difference between the pause durations of the speakers ($p > 0.05$). For this reason, the following pause duration results have been drawn from the combined results of speakers A and B.

   (29) shows mean duration (ms) of disfluency pauses, pauses within words, and standard pauses respectively.

(29)   Pause duration in Dalabon
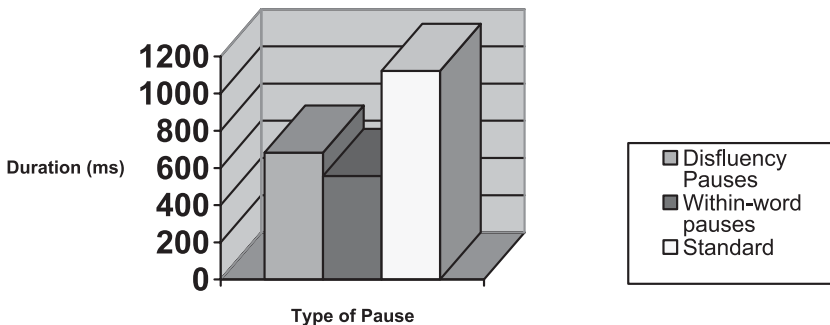
Of the pause types, there were 68 disfluency pauses, 29 instances of within-word pauses, and 492 "standard" pauses, i.e., those pauses that

occurred at the edge of a full prosodic constituent that coincided with a fully inflected grammatical word. As one would expect, the latter type of silent pause was the most frequent in the corpus. Example (29) shows that the standard pause type is of a substantially longer duration (mean = 1122.0 ms) than pauses within words (mean = 556.5 ms) and disfluency pauses (mean = 683.6 ms). A one-way ANOVA revealed a highly significant difference ($p < 0.0001$) between the three pause types. Post-hoc t-tests revealed the highly significant difference in duration lay between pauses within words and standard pauses ($p < 0.0001$), as well as disfluency pauses and standard pauses ($p < 0.0001$), but that there was no difference in duration between disfluency pauses and pauses within words ($p > 0.05$).

Where a grammatical word is interrupted by pause, the result is a sequence of units that display pitch movements typical of an intonational phrase with a peak followed by a trough. Where a pronominal prefix detaches from the grammatical word, one of two things may occur:

a) the pronominal prefix may form its own intonational phrase as defined by at least one peak, a final falling pitch movement at the rightmost edge and the presence of a notable pause to either side of the prefix, or

b) a pronominal prefix may attach to a preceding unit to form an intonational phrase with that unit. The latter possibility is the most radical nonisomorphism of a grammatical and phonological word. In these circumstances the phonological word spans one grammatical word and a reattached "prefix". In some rare cases a grammatical word may be interrupted by two pauses resulting in the grammatical word spanning three intonational phrases.

As mentioned above, the phonological requirements that must be met in order for a prefix to detach from the verbal word are that the prefix have a bimoraic minimum (either monosyllabic and closed, or disyllabic), and that the remainder of the grammatical word have at least two syllables. Generally detached prefixes include the assertative marker, the subordinate marker or the sequential markers. For the most part detached units end in closed syllables, apart from where the sequential marker *dja*, the subordinate marker *ye* or the benefactive marker *marnu* attach to the prefix. The units which typically attach to a pronominal prefix to form a closed unit are the assertative marker, the sequential marker or a combination of both.

Certain morphemes of the verbal word, such as the assertative marker or subordinate marker, always attach to the detached unit containing the

pronominal prefix. Likewise the verb root is never separated from the TAM markers by a pause. This suggests that certain sequences of morphemes must form a coherent unit. These units may therefore be regarded as coherence zones as they may not be separated by a pause. Other morphemes, however, such as the benefactive marker, show some variation as to which unit they attach to suggesting that these morphemes do not form coherence zones together with the pronominal prefix as they may be separated by a pause.

The following examples illustrate the cases where a detached prefix does not attach to a preceding unit, which we will call "isolated prefixes". The constraint operating in these examples limits isolated units to be minimally bimoraic — that is either disyllabic, or monosyllabic and super heavy.

The verbal word in (30) was uttered hesitantly as the verb has unpleasant connotations in Dalabon and is considered insulting. Here the detached pronominal prefix *bala-h* forms its own intonational phrase with a single peak followed by a final falling pitch contour.

(30) kanh [2076 ms]   bala-h [672 ms]   djer-nurru-bobm-u
     DEM   3pl-R   breath-rotten-stink-PR
     'They have bad breath'

(31) Speech waveform, RMS amplitude and F_0 contour of an isolated prefix (Tape6a14m34s)
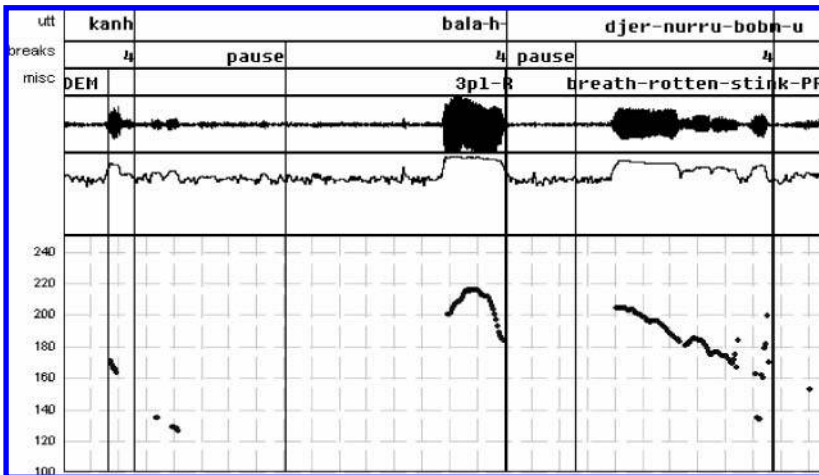
(32) illustrates a detached prefix whose final syllable is not closed, but which satisfies the disyllabicity requirement.

(32)   nunh__mahkih_____da-ye [596ms]
       DEM maybe   2/3-SUB
       birrwoyin-ngandung-iyan_____da-ye   [2963ms]
       br.in.law-call.someone-IRR   2/3-SUB
       wawurd-ngandung-iyan
       big.brother-call.someone-IRR
       'Maybe you would call him "brother in law", you would call him
       "big brother"'

(33)   Speech waveform, RMS amplitude and F_0 contour of an isolated
       prefix "da-ye" (Tape5b4m31s)
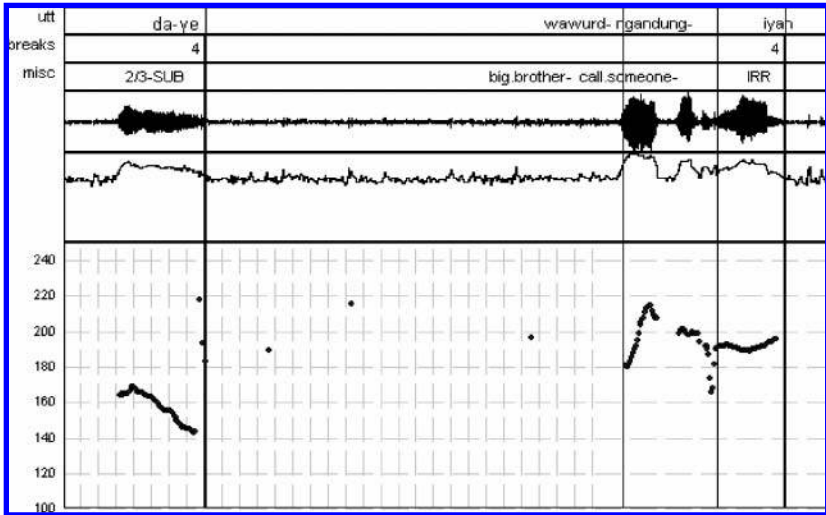
Example (34) illustrates the sole occurrence in the data where a monosyl-
labic prefix detaches from the verbal word to stand isolated. The prefix is
super heavy and thus conforms to the bimoraic minimum of phonological
words.

(34)   ka-h-lng [222ms] maryami-nj
       3-R-SEQ   get. lost-PI
       'Then he got lost'

Where a unit reattaches to the previous intonational phrase without an
intervening pause, this preceding phrase may consist of minimally one
and maximally four words. Table 4 shows the total set of pause-detached
units, both those that stand alone and those that reattach to preceding
material. It is arranged according to the number of words in the preced-
ing material, ranging from 0 (i.e., the prefix stands independently) to 4

Table 4. *Full list of the grammatical words preceding a detached unit*

+4	+3	+2	+1	Detached unit	Rest of word
				bala-h	djer-nurru-bobmu
				buka-h	burlhkeywo-yan
				buka-h-dja-m	men-were-minj
				da-ye	wawurd-ngandung-iyan
				ka-h-dja	marlaworr-karnkarn-no
				ka-h-dja-lng	kurangarrk-burlhmu
				ka-h-lng	maryhahm-inj
				nga-h-kom	dum-kiyan
				ngarra-h	yirrbi-yirrbih-kun
				ngurra-h	djorm-iyan
				yawoyh-dorrng	marra-ngyi
				-yirri-kun-	bon-inj
			kardu	bala-h	yirri-kun
			kanh	bula-h-lng	nguni-njyi
			bulnu	bula-h	warlkoni-njyi
			warhdu	ka-h	marnu-yenjdju-ng
			dorrng-no-duninj	ka-lng	dorrng-burlhm-inj
			-dum-kiyan	nga-h	kom-muhm-iyan
			melbe	nga-h	nahbu-ng
			makmak	yila-yawoyh	djorhk-iyan
			buka-h-lng-kom-dengkohm-inj	buka-h-lng	kom-deyhm-inj
		nunh	mahkih	da-ye	birrwoyin-ngandung-iyan
		wadda	*bulu*	ka-ye-marnu	wudjm-iyan
		kerrino	kanh	ka-h	kerri-derrk-
		bod	*ngorr*	ka-h	kirdikirrbu-n
		nunh	kardu	ngarra-h	yawoyh-dorrng
	kabo-kah	nunh	kanh	ka-h-lng-kolk	burkwo-n
	wurdang	nunh	kanunh	buka-marnu-bad	murduk-won-inj
nunh	yala-yi-n	yala-yirrbbih-kun	nunh	yala-h	nakka-ng

(i.e., attached as the last element of an intonation unit containing four grammatical words).

Of the 29 examples, 12 stand isolated and 17 attach to a previous pho-nological unit consisting of up to four grammatical words; of these 17, nine attach to a single preceding word, five attach to a sequence of two preceding words (though two of these, shown in italics, include object pronouns), two attach to a phrase containing three words, and one to a group of four preceding words. Had we made the analytic decision to treat cases where the preceding material includes the object pronoun

(recall our discussion of the complex situation with object pronouns in Section 2), the figures would have shifted to 11 out of 17 "reattached units" being added to preceding single-word units.

Where a detached unit reattaches to just one grammatical word, that word must be minimally disyllabic. The sole exception is one case where there is reattachment to the monosyllable *kanh* 'that'; note that this (a) is heavy, with a complex two-segment coda, and (b) is reduced, in any case, from the fuller disyllabic form *kanunh*. The commonest elements hosting reattached prefixes are the demonstratives *nunh*, *kanh*, *kanunh* and the adjective *kardu*.

Occasionally the verbal word is broken into more than two parts by pause. An example is (35), where the detached unit reattaches to the previous grammatical word *kardu*, and there is a further pause between the incorporated adverbial *yirri-kun*[22] and the inflected verb stem *boninj*.

(35)   kardu__bala-h [213ms] yirri-kun [467.8ms]   bon-inj
        maybe 3pl-R   spread.out-INCORP.ADV   travel-PI
        'Maybe they travel spread out by themselves'

(36)   Speech waveform, RMS amplitude trace and F0 contour of a reattached unit showing within word pauses (pauseww) (Tape6b4m37s)
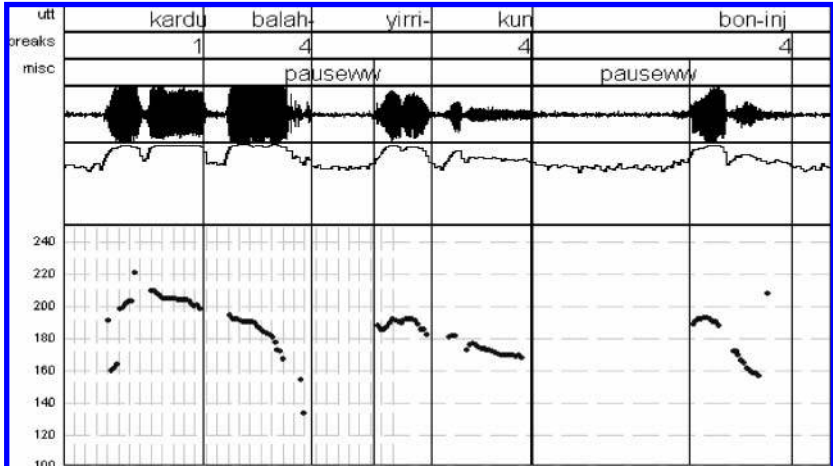
We will now give a few more examples of the various types illustrated in Table 4. (37) is an example of the detached unit forming an intonational phrase with the preceding grammatical word — the demonstrative *kanh*, mentioned above already as the only example of a detached unit reattaching to a single grammatical word that is monosyllabic.

(37) kanh__bula-h-lng [524ms] nguni-njyi_____wurrhwurrungu
      DEM 3-R-SEQ   eat-PCust   old.person
      'Then the old people eat it'
(38) Speech waveform, RMS amplitude trace, and F0 contour of a re-
      attached unit — example (39) (Tape1b27m59s)
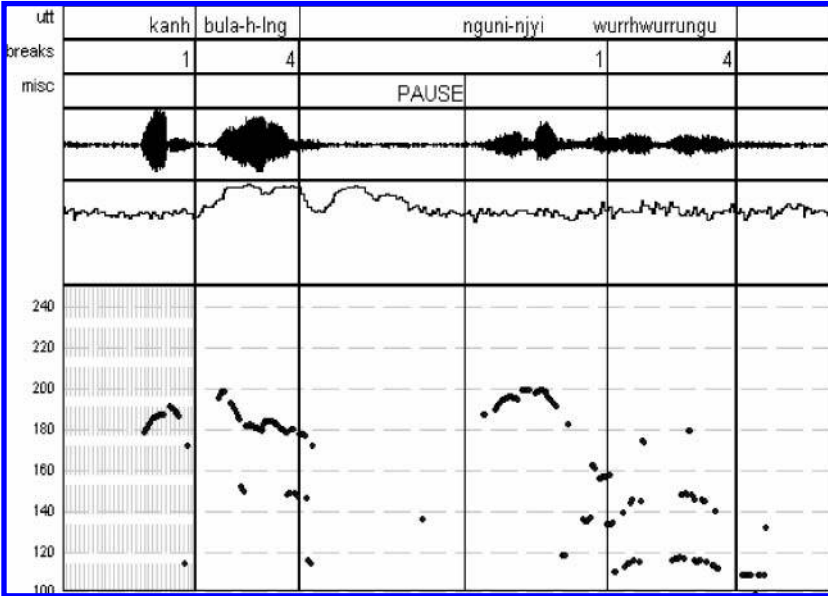
The opposite extreme is shown in (39), where the pronominal prefix
reattaches to an intonational phrase consisting of four grammatical words
excluding the prefix.

(39) nunh__yala-yi-n____yala-yirrbbih-kun__nunh__yala-h
      DEM 1pl-say-PR 1pl-inland-from   DEM 1pl-R
      [266ms]nakka-ng
      come.from-PP
      'We say we are from the inland, that is where we came from'

Pauses within grammatical words are not limited exclusively to prefixes,
but may also separate other components of the verbal word, such as in-
corporated nominals and adverbial prefixes. Typically, an incorporated
nominal groups with the unit containing the prefix, though this is a not a
requirement.
     In (40) the detached unit containing the prefix also contains the incor-
porated nominal *kolk*.

(40)   ka-h-lng-kolk   [1111 ms] burkwo-n
      3-R-SEQ-water   dry.out-PR
      'The water dries out'

In (41) the adverbial prefix *yawoyh* attaches to the pronominal prefix and forms an intonational phrase with the preceding unit.

(41)   makmak_____yila-yawoyh   [315 ms]   djorlhk-iyan
      not.at.all   1pl-again   change-FUT
      'We will not change again'

Where a pause within a word occurs, incorporated nominals and adverbials tend to attach to the unit containing the prefix, as illustrated in examples (40) and (41). However, they need not do so as illustrated in (17) and (25) where the incorporated nominals attach to the unit containing the verbal root.

The degree of freedom available can be seen by comparing the very similar examples in (42) and (43), both involving the same incorporated body-part root *kom* 'neck'. In (42) the second pause separates the sequence, comprising prefix + assertative marker + incorporated nominal *kom*, from the verb root. In (43), by contrast, the pause separates the pronominal prefix + assertative marker from the incorporated nominal *kom* + verb root. Thus an incorporated nominal may behave in different ways by either attaching to the prefix or attaching to the verbal root.

(42)   nga-h [380.1ms] nga-h-kom   [538 ms] dum-kiyan
      1/3-R   1/3-R-throat   make.hole-FUT
      'I, I will make a hole in the throat'
      (Tape1a25m34.40824s)
(43)   nga-h   [440 ms] kom-buhm-iyan
      1/3-R   throat-blow-FUT
      'I will blow into the throat'
      (Tape1a25m36.98887s)

The following two examples illustrate comparable variation with regard to whether the benefactive marker *marnu* groups with the following root, or the preceding prefix.

(44)   warhdu____ka-h   [215 ms]   marnu-yenjdjung
      devil   3-R   BEN-talk.together-PR
      'The devil, he will talk with you'
      (Tape5b20m38s)
(45)   wadda____bulu__ka-ye-marnu   [809 ms] wudjm-iyan
      country   3plO   3-SUB-BEN   be.gone-FUT
      'Our country will be gone'
      (Tape5a21m31s)

In (44) the benefactive marker *marnu* attaches to the unit containing the verbal root, whereas in (45) *marnu* attaches to the unit containing the pronominal prefix + subordinate marker *ye* resulting in an open syllable.

The above-discussed examples, then, illustrate an inconsistency in the location of the pause within a word regarding which units the benefactive marker and incorporated nominal may group with prosodically. This free variation is not observed with the sequential markers *dja* and *-lng* and the assertative marker *h* suggesting that these units must form a coherent unit together with the prefix, and that this unit may not be interrupted by pause.

Reattached units create a nonisomorphism between a grammatical word and a phonological word, as the grammatical word spans two intonational phrases. An even more extreme case of a nonisomorphism between a grammatical word and a phonological word is presented in example (46) as the verbal word is separated by two pauses and therefore belongs to three intonational phrases. Here, the detached prefix forms an intonational unit with the preceding unit; the incorporated adjectives form their own separate intonational unit as do the verb root + TAM marker (see [35] above for a further example).

(46) nunh__kardu   ngarra-h [482 ms] yawoyh-dorrng [509 ms]
     DEM maybe   1pl-R   again-alive
     marra-ngyi
     join.together-Pcust
     'Maybe we would come back to life again'

5.5. *Actual textual frequency of prefix detachment*

Since the material above was based on selected occurrences of the phenomenon picked out specifically because they show pause-detached prefixes, it cannot provide us with information about the frequency of the phenomenon. For this reason a further investigation into the frequency of pauses within grammatical words was conducted. This investigation was based on two different Dalabon texts: Korlomomo (recorded from Maggie Tukumba)[23] and Mimih and Naworneng, recorded from the late Jack Chadum. The figures extracted from these texts reveal the actual textual frequency of the phenomenon in question.

Table 5 displays several figures: the total number of verbs in the two texts; the number of verbs structurally eligible for prefix detachment according to the phonological requirements of prefix detachment stated above; the number of actual detached prefixes; and the percentage of detached prefixes both overall and per eligible verbal word.

Table 5. *Frequency of detached prefixes in the Mimih and Naworneng and Korlomomo texts (% rounded off to nearest integer)*

	Total number of verbal words	Structurally eligible verbal words	Detached prefixes	% of detached prefixes per verbal word	% of detached prefixes per *eligible* verbal word
Korlomomo text	54	40	2	4	5
Mimih and Naworneng text	132	104	6	5	6

The ''Korlomomo'' text, told by Maggie Tukumba, contains 54 verbal words in total. Of these, only 40 verbal words are structurally eligible for prefix detachment. Of these structurally eligible verbs only 2 are actual instances of prefix detachment.

The ''Mimih and Naworneng'' text, told by Jack Chadum, contains 132 verbal words in total. Of these, 104 verbal words are structurally eligible for prefix detachment. Of these structurally eligible verbs only 6 are actually instances of prefix detachment.

These figures reveal that the frequency of prefix detachment in spontaneous speech is somewhat uncommon, hovering around 5% of eligible verbal words in both texts and only taking in around 4% of verbal words overall. This means that the whole phenomenon discussed in this article is rather marginal from a statistical point of view, and obviously does not represent the statistically normal situation, which is for phonological and grammatical words to coincide. This does not diminish its importance, however: it is likely that most linguistic innovations start as low-frequency occurrences, and spread from there, so that this snapshot of Dalabon phonetic structure allows us an early glimpse of a phenomenon that has yet to be generalized to become the commonest realization.

6. Diachronic implications: prefixal reattachment in the bud

In the preceding section we characterize the phenomenon of post-prefix pause as an incipient development, still low-frequency and yet to be phonologized. This statement, of course, presupposes the diachronic scenario mentioned in Section 1, where we give reasons for regarding this as a Dalabon innovation within the Gunwinyguan family. We must stress, however, that at present no comparable study has been carried out for other Gunwinyguan languages, although the extensive empirical data in Bishop (2003) for Bininj Gun-wok, the language most closely related to Dalabon, shows no evidence of an equivalent phenomenon there.

Two linked problems in the historical linguistics of Australian languages concern:

(a) the need to give clear diachronic accounts of the development of 'auxiliaries', comprising reduced pronouns often attached to a 'catalyst' or base, sometimes tense- or mood-specific, which are widespread in Australia (see map on p. 340 of Dixon 2002), found in around seventy of the 200 languages lacking pronominal affixation to the verb; the placement of these is rather variable across languages, including (i) placement in second position within the clause or intonational unit, as in Warlpiri (ii) free placement, as in Karrwa or Warumungu (iii) placement before the verb, as in Kugu Nganhcara.

(b) the need to give well-supported pathways of change between languages with verbal affixation and those with reduced pronouns attached to auxiliaries.

The huge number, diversity, and geographical noncontiguity of languages and families involved means that many separate transition events must have occurred, and it is likely that several distinct mechanisms have been involved in different cases.

The mechanism that is most commonly invoked involves the development of clitics to verb affixes, via (a) the reduction of free pronouns to clitics in various positions, e.g., in Wackernagel's position, or as proclitics to the verb (b) their further development to verbal prefixes (or, more rarely, to verbal suffixes). A clear example of this direction of change is the development of pronominal prefixes in Yanyuwa, partly motivated by areal convergence to neighboring prefixing languages (see Dixon 2002: 380–381). With some caveats regarding languages like Wambaya that have reduced prefixed verbs to the status of second-position auxiliaries, with further development from auxiliary to suffix in Jingulu, Dixon (2002) takes the view that there is a one-way relation between pronominal prefixation and bound pronouns in auxiliaries.[24]

However, could it also be the case that in at least some cases, auxiliaries have arisen through prefixal detachment?[25] The Dalabon data suggest that it might be possible, through a series of steps that would include:

(a) a systematization of silent pause placement within the word, as illustrated in this article

(b) a generalization of this tendency so that it becomes the normative treatment of grammatical prefixes, i.e., the phonological word would now be systematically noncongruent with the grammatical word. Connected with this, within-word pauses would approach between-word pauses in length

(c)   a reanalysis of the grammatical word so that it realigns with the new phonological word in treating the (erstwhile) pronominal prefixes as separate grammatical entities. This type is exemplified by Kugu Nganhcara, for example (Smith and Johnson 2000), where the auxiliary is always placed in immediate preverbal position

(d)   as appropriate, the development of language-specific placement rules for the detached prefixal material, e.g., in Wackernagel's position, Warlpiri style, or placed freely within the clause, as in Garrwa or Warumungu.

The Dalabon data considered in this article only furnishes evidence for the first step, but the third and fourth steps involve possibilities that are also exemplified by other Australian languages — Kugu Nganhcara for (c), Warlpiri and Garrwa among many others for (d).[26] However, it at least suggests that development in the opposite direction to what is normally assumed — i.e., from bound to free material — is conceivable. A fuller evaluation of this scenario will require us to bring in closely analyzed diachronic material from a range of other languages, since the variation in any given language state is unlikely to span more than two of the above four steps. Also relevant here will be diachronic accounts of other domain mismatches between grammatical and phonological units. Some examples are

–   the well-known phenomenon of Kwakw'ala demonstratives which, though grammatically relevant to the following element, are phonologically attached to the preceding one (Anderson 1992: 199),
–   phonological / grammatical domain mismatches in the Bantu language Kukuya (Paulian 1974; Hyman 1987), where grammatical prefixes constitute a phonological domain with the preceding word for the purposes of (a) linking lexical tone melodies and (b) the conditioning of domain-internal processes such as m- and l-gemination, and casual speech deletion of [b] and [m],
–   so-called Aoyagi prefixes in Japanese (Poser 1990), which, though grammatically part of a larger word, are followed by a minor phrase boundary for the purposes of assigning pitch accent.

The developmental trajectory of head-marking and polysynthetic languages is often seen as a one-way street, with ever-more complex morphology developing typological "sinks" — structures that can only be disassembled by catastrophic processes like pidginization (see Lee [1987] for an example involving another polysynthetic Australian language, Tiwi). However, even Nichols (1986: 87), whose principle of "Headward Migration" posits a one-way path of development in the movement of

affixal morphology from dependent to head, allows that "[r]eversal of the headward-migration principle can only occur because of boundary-shifts." The pause phenomena examined in this article may well represent the first stage of such a boundary shift in Dalabon.

7.   Conclusion

All polysynthetic languages are characterized by complex grammatical words capable of expressing a clause worth of information, but there is significant crosslinguistic variation in whether these generously stuffed grammatical words are coterminous with phonological words (as in Chukchi or Bininj Gun-wok), or break up into more than one phonological unit (as in Cree).[27] Dalabon clearly belongs to the second type — apart from some wrinkles with clitics the grammatical word is easy to define, but it may readily (though not necessarily) break up into a number of smaller prosodic units, delineated by pause, and in certain cases it is even possible to merge the first of these smaller units with the preceding word. As a result, the relation between grammatical and phonological units is far less straightforward in Dalabon than in the closely related Bininj Gun-wok, even though our study of text tokens revealed it to be incipient only, being both restricted in its phonological conditions, and relatively rare even when these restricted conditions are met.

In terms of diachronic trajectory this development appears to be the incipient first step on a course which, had the natural process of development of Dalabon not been arrested through a disrupted process of intergenerational transmission, could be expected to lead to the breaking up of its large verbal words into a number of smaller phonological units, and the attachment of closed monosyllabic prefixes, as enclitics to preceding words.

Appendix A. Abbreviations

A	Transitive subject
ABL	Ablative
ADV	Adverbalizer
APPR	Apprehensive
As	Assertative
dis.du	disharmonic dual (members of set in odd-numbered generations)

FUT	Future
GEN	Genitive
hi	higher animate (object, only with 3sg subject)
HORT	hortative
INCORP ADV	incorporated adverb
ITER	Iterative
lo	lower animate (object, only with 3sg subject)
LOC	Locative
NP	Nonpast
O	Object
POSS	Possessed noun
PP	Past perfective
PRT	Part
RR	reflexive/reciprocal
S	Intransitive subject
sg	singular
SUB	Subordinate

Appendix B. DAL_ToBI (Dalabon Tone and Break Indices) tonal labels

H*	H(igh) intonational pitch accent
!H*	Lowered High intonational pitch accent
%H	High left boundary tone
H%	High right boundary tone
!H%	Mid-level boundary tone
L%	Low right boundary tone

Notes

2. Russell believes (e-mail to NE, 13/5/04) that, for Cree, "The looseness and interruptability of the 'prefixes' also seems very old and is shared by (at least) the surviving

central Algonquian languages. Ojibwe is almost as loose as Cree, Fox even looser ... These morphemes seem to have always been syntactic particles, distributed according to syntactic principles. What is significant isn't the handful of combinations that have gained an idiomatic status, but the fact that they have resisted wholesale grammaticalization as prefixes in so many daughter languages for so long.''

3. The language is also known as Dangbon, Ngalkbon, and Buwan. Each of these names is favored by a different language bordering Dalabon: Dangbon by Kune and Kunwinjku, Ngalkbon by Jawoyn, and Buwan by Rembarrnga. Dalabon is the commonest autonym though the other terms are used by Dalabon speakers as well when in regular contact with the aforementioned languages. The terms Dalabon, Dangbon and Ngalkbon are all formed by compounding the root for 'mouth' in the respective language with the Dalabon word *-bon* 'go (present)': Dalabon *dalû-no*, Kunwinjku *kun-dang*, and Jawoyn *ngan-ngalk* all mean 'mouth' (*kun-* and *ngan-* are noun class prefixes). The Rembarrnga term is simply a regular reflex of the root *bon*: in Rembarrnga *o* descends regularly as *uwa* in certain environments (Harvey 2003a). The terms are interesting for the purposes of this article because they show that speakers of these languages are sufficiently aware of lexical stems like *-bon* to make them symbols of linguistic difference, even though they never occur as free forms.

4. Fewer than a dozen speakers of Dalabon remain. Although there has been some work on the language (Capell 1962; Alpher 1982; Sandefur and Gentian 1977; Merlan 1993; Evans et al. 2001; Fletcher and Evans 2002; Evans and Merlan 2003; Evans et al. 2004, Evans forthcoming) there is still no comprehensive grammar of the language and many aspects of the phonology remain unanalyzed. The current article must therefore be regarded as provisional in its treatment.

5. Though there are also interesting phenomena involving pause in nominal words, in this article we focus exclusively on verbal words.

6. Phonetic lengthening under rather similar conditions has been reported for other Gunwinyguan languages: see Harvey and Borowski (1999) on Warray and Baker (1999) on Ngalakgan.

7. The only significant prefixes are a directional prefix *berre-* on compass terms, originating as a compound meaning 'chest' (i.e., etymologically 'chest-east' for 'eastwards'), and a few masculine vs. feminine pairs of prefixes to human terms, which are the last residues of an original prefixal noun-class system still found in BGW.

8. It is interesting that in each major word class the predominant clitic patterns lie on the same side as the nonhead morphemes: verbs, with their rich prefixal system and right-headed structure (compounded nominal and verbal roots lie to the left of the main verbal root), have proclitics, while nouns, which form left-headed compounds (and predominantly suffixing), have enclitics.

9. Subordinate clauses have special prefix forms, some of which alter the final vowel and others of which are identified by the absence of the ''assertative'' marking glottal stop. Because the latter are homophonous with the basic form of the prefix they are not overtly glossed as subordinate here.

10. See Harvey (1991) on the phonology of glottal stops in Arnhem Land: Merlan (1983: 6) and Baker (1999: 15) point out that in the related language Ngalakgan complete glottal closure may be dispensed with leaving only creaky voice associated with glottal constriction as the phonetic signal.

11. A *Linguistics* reviewer ''wonders here whether speakers who isolate incorporated nominals might not be influenced by working with linguists.'' To the extent that such questions are answerable, we do not believe this to be the reason, since some of our Dalabon informants employed segmentations of this type quite early on in the process

of language investigation, at a time when they were still quite unfamiliar with the various expectations we had about elicitation of material.

12. A *Linguistics* reviewer points out that "in Bininj Gun-wok at least, free pronouns (as the first author of this article has argued) presently have a distinct discourse role from bound pronominals, being used in the introduction and focusing of referents" and raises the question of whether "there a possibility at least that this distinct discourse role might help stabilize the wordhood of the free pronouns, producing resistance to these processes of (re-)reduction?" The issue here concerns the relevant timing of the phonological process of cliticization and ultimate reduction to prefixal status on the one hand, and of the semantic process of loss of referentiality and givenness on the other. At this stage of research we do not have sufficient material on the referentiality of these clitic pronouns to say how far they have lost their focusing and referentializing function, if at all.

13. For those phonologists unhappy about stating prosodic constraints over morphologically-defined units as opposed to those defined in pure phonological terms, it would of course be possible to translate this constraint by making all roots a possible domain over which feet can be constructed (whereas affixes do not inherently constitute feet), and then stating the bimoraic minimum as a constraint on feet.

14. Dixon and Aikhenvald (2002: 24) make the commonsense observation that "the longer the words of a language are, the more likely there are to be pauses in the middle of them," without trying to relate this observation to grammatical or phonological criteria.

15. The only exceptions are where the morpheme does not left-align with a syllable onset, as well as a couple of TAM suffixes which must always be footed with the verb stem.

16. These "dictations" were made to the first author by Peter Mandeberru. Though he was not literate in any language, he naturally segmented the words as shown, in his efforts to slow down the stream of speech to a point where it could be transcribed in real time.

17. See Streeck (1996) on an interesting parallel in Ilokano — the use of drawn-out final vowel sounds of certain grammatical units to maintain turns. Although here they are not actual prefixes but prepositions, articles and linkers, nonetheless they resemble the prefixes discussed here in projecting forward to a following grammatical head.

18. It is beyond the scope of this article to examine the functions of pause more systematically, since our main concern is with the phonological and grammatical constraints on pause placement and we are only concerned with pauses within word boundaries. For a detailed study of the functions of pauses (this time at word and indeed paragraph boundaries) in the Kunwinjku dialect of Bininj Gun-wok, see Carroll (1995).

19. This is by no means a unique situation crosslinguistically — the situation with Nootka case-markers, which signal the clausal role of the following phrase but group phonologically with the preceding word, is perhaps the best known case, while the grouping of prefixes with the preceding word in terms of prosodic domain in Kukuya is another one (Hyman 1987). In Australia it has particular significance since it suggests a possible route by which Warlpiri-type second-position auxiliaries could arise from prefixes diachronically through prosodic regrouping of prefixes with their preceding words, a point we return to in Section 6.

20. At a Dalabon orthography workshop held at Blue Waters outstation, N.T. in October 2000 and coordinated by the first author and Francesca Merlan the issue of how to write such words was discussed, and the Dalabon speakers present there proposed the use of hyphens rather than spaces or writing the morphemes en bloc.

21. Cf. Russell's (1999: 206) observations on Cree speakers' use of spaces between the person prefix, each preverb and the verb stem when using their syllabic orthography, a

practice he argues aligns with various phonological arguments that the units so defined constitute separate phonological words. A further interesting parallel with Dalabon is that suffixes to the verb stem are phonologically integrated.

22. Dalabon permits the incorporation of a range of adverbials into the verbs, including nominal expressions denoting locations. In such cases the suffix *-kûn* must be suffixed to the incorporated element; on external nominals this same suffix marks genitive and purposive relations. In some cases the same suffix is also found on unincorporated versions, as with *molk-kûn* 'without the relevant person knowing: unannounced, without permission, unsuspected, sneakingly': cf. the unincorporated form in *kardû molk-kûn rakalk wuku-marnû-burlhmu* [maybe UNBEKNOWN killer 3A/1O:APPR-BEN-come.outPR] 'maybe a killer will attack me without anyone knowing' and the incorporated form in *kardû kah-molk-kûn-kolh-yu* [maybe 3sgAss-unbeknown-water-liePR] 'maybe there is water there that no-one knows about'. In other cases, such as the incorporation of nouns bearing a locative relation to the predicate, the *-kûn* is only found in the incorporated version: cf. *balah-djarrk-ni bad-dun-no-kah* [3plAss-together-sit rock-hole-3POSS-LOC] 'they are camping together in the rock cave', with the external nominal *bad-dun-no-kah* [rock-cave-PART-LOC] 'in the rock cave', and its incorporated equivalent *balah-djarrk-dun-kûn-ni* [3plAss-together-hole-GEN-sitPR]. Since the stripping away of affixal material is a normal characteristic of incorporation, such constructions are highly unusual crosslinguistically anyway; whatever their cross-linguistic status, it is clear that the addition of *-kûn* to the incorporated root makes it a larger prosodic unit more inclined to independent prosodic status.

23. For an analysis of this text, including sound files, see: http://socialstudies.cartagena.es/ images/PDF/no0/evanssasse_searching.pdf.

24. A separate issue which we lack the space to go into here concerns the question of antiquity. It is often assumed that all bound morphology originates as free material, at some point in a language's history, so that pronominal prefixes in all Australian languages ultimately derive from free, via cliticized, pronouns. However, whereas Dixon (1980, 2002) sees this as a recent and convergent process among the non-Pama-Nyungan languages, there is good evidence, based on the formal resemblances between intransitive prefixes in particular, that at least part of the prefixal systems of most non-Pama-Nyungan languages can be derived from a common ancestor, and is therefore of far greater antiquity. See Evans (2003b) and Harvey (2003b) for details.

25. Capell (1972) may be expected to have taken this view, through his choice of the term "affix-transferring languages" to describe languages like Warlpiri with a second-position auxiliary hosting pronominal clitics for subject (and often object), a reading reinforced by his wording "the transference of person markers (subject and object) from the verb, *where they logically belong* [italics ours] to the head-word of the utterance" (Capell 1972: 5). However, the diachronic scenario he puts forward in this article is essentially the same as Dixon's, and, as the italicized words in the above quote show, his use of the term seems merely to have reflected a sort of "prescriptive typology" that sees verbal agreement as somehow more "logical" than the positioning of agreement elements on an auxiliary.

26. This is not to say this is necessarily the correct historical explanation for each of the languages mentioned, merely that three out of the four states in the scenario are exemplified by attested Australian languages.

27. There are also less clear cases, where it remains controversial how many phonological words are involved. Cayuga, for example, has traditionally been treated as a language where grammatical and phonological words coincide, but Dyck (1994) argues that a single grammatical word is a phonological phrase made up of several phonological words.

References

Alpher, Barry J. (1982). Dalabon dual-subject prefixes, kinship categories and generation skewing. In *Languages of Kinship in Aboriginal Australia*, Oceania Linguistic Monographs 24, Jeffrey Heath, Francesca Merlan, and Alan Rumsey (eds.), 19–30. Sydney: University of Sydney.

Anderson, Stephen R. (1992). *A-Morphous Morphology.* Cambridge: Cambridge University Press.

Baker, Brett (1999). Word structure in Ngalakgan. Unpublished doctoral dissertation, University of Sydney.

Beckman, M. E. and Ayers Elam, G. (1997). Guidelines for ToBI labelling. Version 3, March 1997. Unpublished ms., Ohio State University.

Bishop, Judith (2003). Aspects of intonation and prosody in Bininj Gun-wok: an auto-segmental-metrical analysis. Unpublished doctoral dissertation, University of Melbourne.

—and Fletcher, Janet (2005). Intonation in six dialects of Bininj Gun-wok. In *Prosodic Typology and Transcription: A Unified Approach*, Sun Ah-Jun (ed.), 331–361. Oxford: Oxford University Press.

Capell, Arthur (1962). *Some Linguistic Types in Australia: Handbook of Australian Languages, Part 2*. Oceania Linguistic Monographs 7. Sydney: University of Sydney.

—(1972). The affix-transferring languages of Australia. *Linguistics* 87, 5–36.

Carroll, Peter (1995). The old people told us: verbal art in Western Arnhem Land. Unpublished doctoral dissertation, University of Queensland.

Dixon, R. M. W. (1980). *The Languages of Australia.* Cambridge: Cambridge University Press.

—(2002). *Australian Languages: Their Nature and Development.* Cambridge: Cambridge University Press.

—and Aikhenvald, A. Y. (eds.) (2002). *Word: A Cross-Linguistic Typology*. Cambridge: Cambridge University Press.

Dunn, Michael (2001). The word in Chukchi. Unpublished manuscript.

Dyck, Carrie (1994). On the definition of Word in a polysynthetic language. In *Proceedings of the 1994 Annual Meeting of the Canadian Linguistics Association*, 131–153. Toronto: Toronto Working Papers in Linguistics.

Evans, Nicholas (2003a). *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune.* Canberra: Pacific Linguistics.

—(ed.) (2003b). *The Non-Pama-Nyungan Languages of Northern Australia: Comparative Studies of the Continent's Most Linguistically Complex Region*. Canberra: Pacific Linguistics.

—(2006). Who said polysynthetic languages avoid subordination? Multiple subordination strategies in Dalabon. *Australian Journal of Linguistics* 26(1), 31–58.

—; Brown, Dunstan; and Corbett, Greville (2001). Dalabon pronominal prefixes and the typology of syncretism: a network morphology analysis. *Yearbook of Morphology 2000*, 187–231.

—and Merlan, Francesca (2003). Dalabon conjugations. In *The Non-Pama-Nyungan Languages of Northern Australia: Comparative Studies of the Continent's Most Linguistically Complex Region*, Nicholas Evans (ed.), 269–283. Canberra: Pacific Linguistics.

—; Merlan, Francesca; and Tukumba, Maggie (2004). *A First Dictionary of Dalabon*. Maningrida: Bawinanga Aboriginal Corporation.

Fletcher, Janet and Evans, Nicholas (2000). Intonational downtrends in Mayali. *Australian Journal of Linguistics* 20, 23–31.

—and Evans, Nicholas (2002). An acoustic phonetic analysis of intonational prominence in two Australian languages. *Journal of the International Phonetics Association* 32, 123–140.

Hall, T. Alan (1999). The phonological word: a review. In *Studies on the Phonological Word*, T. Alan Hall and Ursula Kleinhenz (eds.), 1–22. Amsterdam: John Benjamins.

Harvey, Mark (1991). Glottal stop, underspecification and syllable structures among the Top End languages. *Australian Journal of Linguistics* 11(1), 67–105.

—(2003a). An initial reconstruction of proto-Gunwinyguan phonology. In *The Non-Pama-Nyungan Languages of Northern Australia: Comparative Studies of the Continent's Most Linguistically Complex Region*, Nicholas Evans (ed.), 185–201. Canberra: Pacific Linguistics.

—(2003b). Reconstruction of pronominals among the non-Pama-Nyungan languages. In *The Non-Pama-Nyungan Languages of Northern Australia: Comparative Studies of the Continent's Most Linguistically Complex Region*, Nicholas Evans (ed.), 475–511.

—and Borowsky, Toni (1999). The minimum word in Warray. *Australian Journal of Linguistics* 19(1), 77–88.

Hyman, Larry (1987). Prosodic domains in Kukuya. *Natural Language and Linguistic Theory* 5, 311–333.

Ladd, D. Robert (1996). *Intonational Phonology.* Cambridge: Cambridge University Press.

Lee, Jennifer (1987). *Tiwi Today: A Study of Language Change in a Contact Situation.* Canberra: Pacific Linguistics.

Merlan, Francesca (1983). *Ngalakan Grammar, Texts and Vocabulary.* Canberra: Pacific Linguistics.

—(1993). Ngalkbon pronominal prefixes, from information collected January 1993 from Daisy Borduk and Lily Bennett. Unpublished manuscript.

Millar, J. B.; Harrington, J. M.; and Vonwiller, J. P. (1997). Spoken language resources for Australian speech technology. *Journal of Electrical and Engineering Australia* 17, 13–23.

Mithun, Marianne (1998). The significance of diversity in language endangerment and preservation. In *Endangered Languages*, Lenore A. Grenoble and Lindsay J. Whaley (eds.), 163–191. Cambridge: Cambridge University Press.

Nangan:gold, David Jentian (1975). *Ngal gordow djirrirti. The brolga and the quail.* Bamyili, N.T.: Bamyili School.

Nichols, Johanna (1986). Head-marking and dependent-marking grammar. *Language* 62(1), 56–119.

Paulian, Christiane (1974). *Le Kukuya: langue teke du Congo.* Paris: SELAF.

Poser, William J. (1990). Word-internal phrase boundary in Japanese. In *The Phonology-Syntax Connection*, Sharon Inkelas and Draga Zec (eds.), 279–288. Stanford: CSLI.

Rice, Keren (1993). The structure of the Slave verb. In *Studies in Lexical Phonology*, Sharon Hargus and Elle Kaisse (eds.), 145–172. New York: Academic Press.

Ross, Belinda (2003). The phonological/grammatical mismatch in the Dalabon word: a phonetic study. Unpublished honors thesis, University of Melbourne.

Russell, Kevin (1999). The "word" in two polysynthetic languages. In *Studies on the Phonological Word*, T. Alan Hall and Ursula Kleinhenz (eds.), 203–221. Amsterdam: John Benjamins.

Sandefur, John and Nangan:gold, David Jentian (1977). A tentative description of the phonemes of the Ngalkbun language (including a small word list). In *Five Papers in Australian Phonologies*, Joyce Hudson (ed.), 57–96. Darwin: SIL-AAB.

Smith, I. and Johnson, S. (2000). Kugu Nganhcara. In *Handbook of Australian Languages, Vol. 5*, R. M. W. Dixon and Barry J. Blake (eds.), 357–489. Oxford: Oxford University Press.

Streeck, Jürgen (1996). A little Ilokano grammar as it appears in interaction. *Journal of Pragmatics* 26, 189–213.