Pausing strategies and prosodic boundaries in Dalabon

Janet Fletcher, Nicholas Evans, and Belinda Ross
Department of Linguistics and Applied Linguistics
School of Languages
University of Melbourne
Australia
janetf@unimelb.edu.au

Abstract

Earlier impressionistic analyses of Dalabon indicate that the grammatical word is often realized as either an accentual phrase, or an intonational phrase, followed by a pause. Unusually, it can also be interrupted by a silent pause, with each section being realized as separate intonational phrases. Our results support these earlier impressions, although this use of the silent pause appears to be restricted to certain affix boundaries, and other phonological constraints relating to the following linguistic material.

1. Introduction

This paper examines prosodic structure, silent pause placement and silent pause duration in Dalabon, a polysynthetic Australian language of Arnhem Land. One interesting feature of Dalabon is that grammatical words can be interrupted by placing a silent pause after a pronominal prefix. 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. Evans et al. (submitted) suggested this is not a widespread feature of pausing in the language, but it makes it quite distinct from its nearest related language, Bininj Gun-wok. Figure 1 shows a typical example of this phenomenon. The pause shown here is not a disfluency (after Shriberg, 2001), i.e. a repair or filled pause, but a silent pause. The prefix "kenh" would normally be attached to the following verb complex in the utterance "his breath was really stinking".

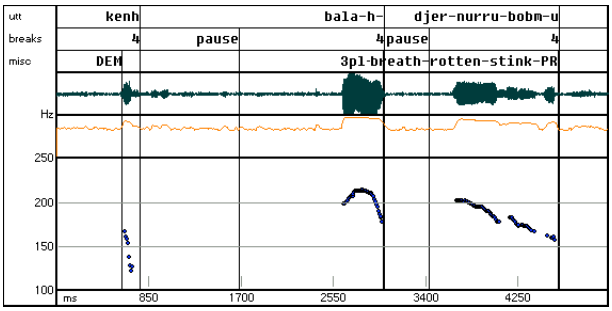

Figure 1: Speech waveform, fundamental frequency contour and rms amplitude trace for a stretch of speech illustrating an isolated prefix “kenh”, a demonstrative.

Before outlining our study, we present some general phonetic and phonological features of Dalabon in the following section.

2. General characteristics of Dalabon

Tables 1 and 2 show the consonant and vowel inventories of Dalabon, using both IPA symbols and the practical orthography that has been developed for this language

Table 2 Consonant phoneme inventory of Dalabon

		Place of Articulation					
		Peripheral		Apico-		Laminо-palatal	Glottal
		Bilabial	Velar	alveolar	retroflex		
Manner of Articulation	Short stop	p (b)	k (k)	t (d)	ʈ (rd)	c (dj)	ʔ (h)
	Long stop	pː (bb)	kː (kk)	tː (dd)	ʈː (rdd)	cː (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)	

Accepted after full paper review

Table 2. Vowel phoneme inventory of Dalabon.

	Front	Central	Back
Close	i	ʉ (û)	u
Mid	e		o
Open		a	

The segmental phonology of Dalabon is fairly typical of many Australian languages. There is a relatively rich set of place of articulation contrasts in the stop, and nasal series, no voicing contrast among the set of obstruents, and a long/short stop contrast. The language has 5 or 6 vowels depending on the phonological status of the close central vowel. There are only a handful of remaining speakers of the language, and the sixth vowel is merging with the close back rounded vowel in the speech of the remaining younger speakers. Phonotactically, all syllables have structure C_1V $(C_2)(C_3)(ʔ)$; 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 [ʉ] in some unstressed positions, and the conversion of sequences of identical stops into long stops within some compounds.

With respect to phrasal stress and intonation in Dalabon, our earlier studies suggest that phrasal stress or prominence (that is, the most intonationally prominent syllable in a phrase) tends to be located on the final, penultimate, or antepenultimate syllable of intonational phrases depending on the location of lexical stress (Fletcher & Evans, 2002). Due to its polysynthetic morphology, Dalabon also has extremely long words consisting of several morphemes. As a consequence, grammatical words are often realized as one, two or in certain rare cases, three intonational phrases. The general pitch pattern of these intonational phrases is like the classic "hat-pattern" described in traditional Dutch models of intonation (e.g. Cohen & t'Hart, 1967). In other words, intonational phrases tend to consist of one or two pitch peaks, with an initial rise to the first peak delimiting the left edge, and terminating with a final fall, whose start generally coincides with the final peak or "elbow" in the fundamental frequency contour for the phrase. Many phrases also end with a suspended mid-high tune, particularly in story-telling genre.

Earlier prosodic studies of Dalabon , as well as the current one, are located within the autosegmental-metrical (A-M) intonational framework (e.g. Pierrehumbert, 1980; Ladd, 1996) among others. A version of ToBI (Tones and Break Indices) has been devised for the language, whereby levels of prosodic constituency are also annotated, along with intonational targets (after Beckman & Ayers-Elam, 1994). Earlier work by Ross (2003) and Bishop and Fletcher (2004) suggests that minimally three levels of constituency need to be acknowledged in Dalabon: a break index value of 1 suggests there is minimal juncture between adjacent words; a break index value of 3 suggests there is tonal juncture of some kind, i.e. a falling or rising intonation contour the end of a word, followed by a pitch restart on a following word. In this study, we will describe this constituent as an accentual phrase. Finally, a break index value of 4 indicates a full intonational phrase. The latter is almost always followed by a silent pause. The aims of this study were to see whether intra-word pause duration is different from inter-word pause duration, and to see what implications this might have for modeling prosodic constituency in Dalabon.

3. Method and materials

The corpus consisted of one spontaneous narrative produced by a female speaker of Dalabon (speaker MT). Approximately 30 minutes of connected speech were analysed for the speaker. An additional speaker (JC) was recorded on another occasion, but the majority of results reported here are based on analyses of speaker MT's speech. The corpus was recorded by the second author in Northern Arnhem land during a linguistic fieldtrip. The field tapes were transcribed in practical orthography (see Tables 1 & 2), and glossed and translated by the second author, also. The recordings were subsequently digitized at 22Khz using ESPS/Waves+, running on a SUN work station in the Phonetics Laboratory, University of Melbourne. The signal files were then annotated using EMU labeler (Cassidy & Harrington, 2001). Silent pauses were identified and annotated using the acoustic waveform and spectrogram as a guide. Any silent gap in the waveform of 200 ms or more was labeled as a pause. A conservative measure was chosen because Dalabon contrasts short and long stops. The main acoustic cue to a long stop is a long silent closure interval.

Break indices were also annotated. As noted above, break index 4 marks the highest degree of perceived juncture between prosodic constituents. The phonetic cue indicating a break 4 constituent is a phrase-final intonational movement (generally a fall or a sustained mid or high level tune), lengthening of a final syllable, and generally a silent pause. Break indices 1 and 3 were also annotated, representing a 'word', and accentual phrase, respectively. The 'word' level break inex corresponded generally to a morpho-syntactic or 'grammatical' word. The distribution of pauses, median pause duration, and distribution of prosodic constituents as indicated by break indices, were calculated using the R and Emu interface package. Instances of grammatical words that were interrupted by a pause were also noted. The median duration of these intra-word pauses was also calculated, and compared with the duration of inter-word pauses.

Accepted after full paper review

4. Results

4.1. Distribution of silent pauses

Figure 2 plots the durational distribution of pauses for speaker MT. Altogether, 521 silent pauses were measured. As would be expected, the distribution is skewed to the left. The median duration of all silent pauses for this speaker was 941ms. The pauses ranged in duration from 207ms to 6160 ms.

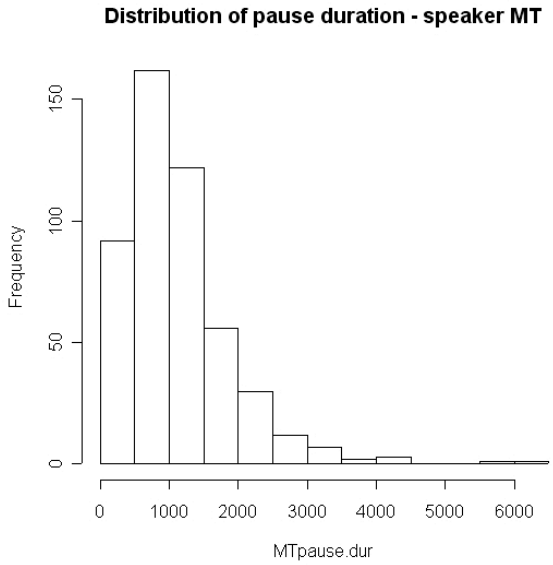

Figure 2. Distribution of silent pauses for Speaker 1 (MT)

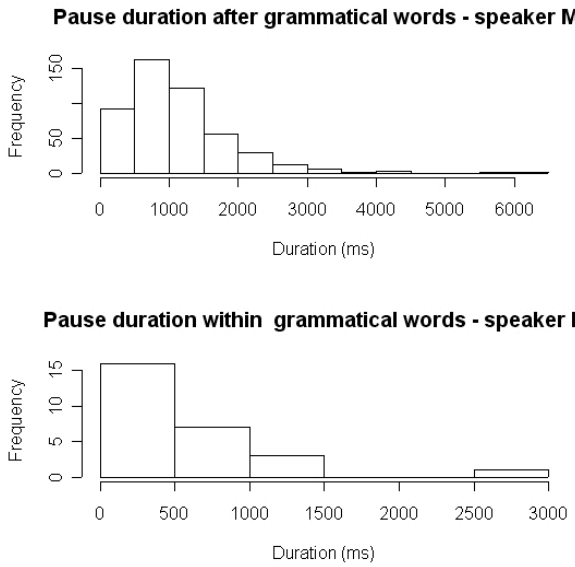

Figure 3. Comparison of pauses that occurred word and intonational phrase finally, and pauses the occurred within a grammatical word.

Figure 3 plots the distribution of pauses between full grammatical words, i.e. standard pauses, and silent pauses that interrupt grammatical words. As predicted in the introduction, there were far fewer pauses within grammatical words (27 versus 488 instances) than standard pauses that occurred after complete grammatical words. The difference in median duration between the two types of pause was highly significant ($t=5.05$; $p<0.0001$), with standard pauses being substantially longer than intra-word pauses (981ms versus 482 ms).

4.2 Distribution of break indices

The total distribution of break indices is shown in Figure 4.

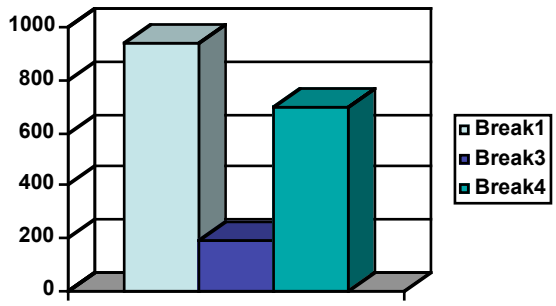

Figure 4. Distribution of break indices

In the corpus there were 939 occurrences of break index 1 constituents, 189 occurrences of break index 3 constituents, and 698 occurrences of break index 4 consstituents. An average intonational phrase in this corpus consisted of approximately 1.9 grammatical words. Figure 5 illustrates the distribution of sequences of words (break 1 indices) per intonational phrase (break 4 indices).

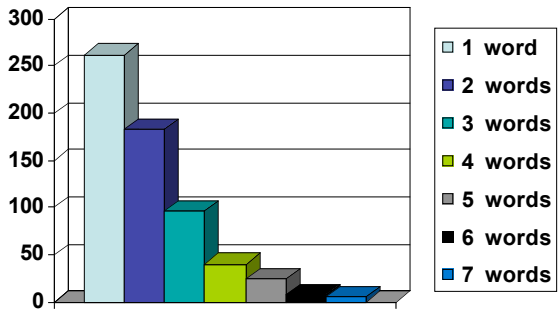

Figure 5. Distribution of intonational phrases based on number of words.

Figure 5 shows that an intonational phrase for this speaker of Dalabon, can contain minimally one word (with 262 occurrences in the corpus), and maximally seven words (5 instances). Longer intonational phrases generally consisted of one or more accentual phrases. It could be argued that intonational phrases that contain

Accepted after full paper review

prefixes isolated from the interrupted grammatical word, constitute prosodic constituents of "less" than one word. This would not be surprising, given the highly polysynthetic nature of the language, compared to English, for example. These results are also comparable to earlier findings for Bininj Gun-wok, which has a mean ratio of between 1.5 and 2.1 words per intonational phrase (Bishop & Fletcher 2004).

5. Discussion

This study focused on the distribution and durational aspects of silent pauses in Dalabon spontaneous discourse. Two major kinds of silent pauses were observed: so-called standard pauses, generally following a grammatical word and coinciding with an intonational phrase boundary, and pauses within a grammatical word. The distribution of pause durations was strongly skewed to the left and reflected a pattern noted by other researchers for spontaneous discourse (e.g. Campione & Veronis, 2002). Only 27 pauses within a word were observed in our corpus, with the remainder of silent pauses occurring after a full grammatical word. Furthermore the durational characteristics of these two types of pauses were not dissimilar from the patterns described by Campione and Veronis, with pauses grouping into two categories, medium pauses (around 500ms) and long pauses (around 1000ms or more). Pauses within words tended to be half as long as pauses between words. As predicted in the introduction, the former provided a clear example of an interruption of a grammatical word.

These intra-word pause results also have implications for prosodic structure in Dalabon. The detached prefixes are the clearest example of mismatch between a grammatical word and a phonological word. In other words, the grammatical word is effectively realized as two phonological words, accentual phrases or intonational phrases, following the principles of the strict layering hypothesis (Nespor and Vogel, 1986). More specifically, where a grammatical word is interrupted by pause, the result is a sequence of units that display pitch movements typical of an intonational phrase (Break index 4) with a pitch peak followed by phrase-final pitch fall. 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 (i.e.a Break index 4 constituent) 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. 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. This appears to be a typologically unique feature of Dalabon, relative to other Australian languages examined so far. It remains to be seen whether other heavily polysynthetic languages exhibit the same phenomenon. It is also not at all clear what the discourse function of such pauses is, if we can assume that speakers use pausing strategies as one means of structuring their spontaneous discourse. This will be the subject of a follow-up study. Recall that we also did not examine disfluencies in this paper. There were, however, a number of disfluencies in this corpus, i.e. repetition, deletion or substitution of units, such as pronominal prefixes, which should otherwise be attached to a following verbal unit. These will be the subject of a wider study. Disfluency pauses differ from intra-word pauses, as the unit preceding a pause is either repeated or changed to a different unit.

With respect to the overall patterning of intonational units in this language, our findings concur with earlier findings for the closely related language Bininj Gun-wok, in that the largest number of intonational phrases in our corpus consisted of one grammatical word. This suggests that theories of prosodic constituency need to accommodate a range of different language types before any unified universal theory can be assumed.

6. References

Beckman, M. & Ayers-Elam, G.(1994). Guidelines for ToBI labelling http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/singer_tobi.html

Bishop, J. & Fletcher, J. (2004) Intonation in six dialects of Bininj Gun-wok. In Sun-Ah Jun (ed.) *Prosodic typology: a unified approach*. Oxford: Oxford University Press. 331-361

Cohen, A.& t'Hart, J. (1967) On the anatomy of intonation. *Lingua* 19, 177-192.

Campione, E. & Veronis, J. (2002). A large-scale multilingual study of silent pause duration. *Proceedings of Eurospeech 2002*, 199-202

Cassidy, S. and Harrington, J. (2001). Multi-level annotation in the Emu speech database management system. *Speech Communication*, 33, 61-77.

Evans, N., Fletcher, J. & Ross, B. (submitted) Big words, small phrases: mismatches between pause units and the polysynthetic word in Dalabon.

Fletcher, J.& Evans, N. (2002). An acoustic intonational study of intonational prominence in two Australian languages. *JIPA* 32, 123-140.

Ladd, D.R (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Unpublished MIT PhD thesis.

Ross, B. (2003). The phonological / grammatical mismatch in the Dalabon word Unpublished Honours Thesis, Department of Linguistics & Applied Linguistics, University of Melbourne.

Accepted after full paper review