An instrumental analysis of focus and juncture in Warlpiri

Andrew Butcher[†], **Jonathan Harrington** [‡]

† Centre for Human Communication Research, FMRI, Flinders University, Adelaide, Australia.

‡ Institute of Phonetics and Digital Speech Processing (IPDS), University of Kiel, Germany.

E-mail: andy.butcher@flinders.edu.au, jmh@ipds.uni-kiel.de

ABSTRACT

Two indigenous speakers of the Central Australian language Warlpiri produced dialogues that were designed to elicit two types of contrast: between phrase-initially focused and phrase-finally non-focused words; and between two phonemically equivalent lexical-items that either spanned a full word-boundary or were separated by a morpheme boundary in a compound. Focusing was marked by lengthening of the rhyme and supralaryngeal expansion, principally of the consonant of the rhyme, while the word/morpheme boundary distinction resulted in timing differences both preceding and following the boundary. The contrasts at these two levels were maintained independently of each other. We discuss these effects in terms of hierarchical prosodic models and hyperarticulation.

1. INTRODUCTION

In the last 20-30 years, there have been numerous impressionist studies of the sound structure of indigenous Australian languages but comparatively few instrumental investigations. This is particularly so as far as their prosody is concerned which, with recent exceptions [e.g., 1-3], has been almost exclusively analysed from auditory-based transcriptions. In the present study, we aim to apply acoustic measurements to controlled laboratory dialogues collected from two speakers of Warlpiri, a central Australian language spoken to the north and west of Alice Springs, with roughly 3000 speakers. One of our main aims is to disentangle the phonetic effects of focus (to the extent that this exists in Warlpiri) from lexical stress, given firstly that none of the impressionistic analyses of prominence patterns in Australian languages makes this distinction explicit (so it is difficult to know whether fieldworkers are hearing 'sentence-accent' or 'lexical-stress') and secondly because most of our knowledge of the articulatory and acoustic characteristics of the effects of focus and its relationship to sentence accent is restricted to Germanic languages (e.g., [4-6]). In the present study, we investigated the focused/unfocused distinction and in particular whether (as has been found for Germanic languages [5,6]) focus is accompanied by a durational and supralaryngeal expansion. A second aim was to investigate the extent to which boundaries at, and below, the word level are distinguished phonetically. Recent experimental investigations have shown that these types of junctural differences can be marked phonetically in English and in other languages [7,8,9]. A recent analysis of consonants in Korean also shows that speakers' articulatory patterns are influenced by morphological structure [10]. Here we extend these types of investigation to Warlpiri by comparing a sequence of words that were segmentally identical but differed in whether they were, or were not, compounded as a noun-phrase. Since it has been suggested that the second lexical item is metrically weaker in the compound than when it forms a word in its own right [11], we hypothesised that these low-level prosodic contrasts might be manifested as subtle timing differences and in particular a weakening of the initial segment of the second lexical-item when it was compounded. Finally, we seek answers to some basic questions about prosodic structure in Warlpiri such as whether focused words are marked by a pitch accent and whether Warlpiri (like many other languages) exhibits phrase-final lengthening.

In common with many Australian languages, which show a remarkable homogeneity in their segmental phonological structure, Warlpiri has 5 contrasting oral and nasal stop places of articulation, a rich set of sonorant consonants (including three contrasting laterals), no fricatives, and no voicing distinction. There is a 'triangular' vowel system including two mid-high vowels and an open central vowel as well as a vowel length distinction that is restricted to the initial syllable. Almost all words begin with a consonant, end in a vowel and are minimally CVCV. Because consonant clusters can only occur word-medially and because of the word-initial neutralisation of a retroflex/alveolar contrast, the phonetic and phonological structure of Warlpiri is more complex in medial than initial position. There is general agreement that 'stress' in Warlpiri is fixed on morpheme-initial syllables and that there may be 'secondary stresses' on every other subsequent odd-numbered syllable, with the exception that morpheme-final syllables are not stressed (e.g., /mánaŋkàra/, 'spinifex plain', [11]). Phrasal-stress does not seem to shift depending on the pragmatic interpretation: with the exception of the auxiliary, Warlpiri has largely free word order and prominence-marking or focusing is accomplished by putting a word in phrase -initial position [12].

2. METHOD

Acoustic and kinematic recordings were made at the Speech Hearing and Language Research Centre (SHLRC), Macquarie University in July 2000 from four indigenous

ISBN 1-876346-48-5

speakers of Warlpiri: 3 adult female (BP, KR, TR) and one adult male (RR). We report here on data from the 2 subjects (BP, TR) only: the untimely death of KR soon after the recordings were made rendered the immediate analysis of her data culturally inappropriate. The segmentation and labelling of the data from RR were in many ways more complicated than those of the other subjects since he used a greater number of reductions than the other two subjects and, at the time of writing, the labelling and analysis of his data could not be completed. Since we will only report on the acoustic variables in this study, the description of the kinematic techniques is not considered in any further detail here.

Two sets of materials, which were constructed with the extensive help of Mary Laughren, David Nash and Bess Price (one of the subjects), varied the focus with word order as well as the presence or absence of morpheme boundaries in some well-known minimal pair phrases. We report here on the first set of materials that included the minimal pair /kuju#puŋu/ (a verb-phrase meaning 'killed the animal' and with a word boundary between the lexical-items) and /kuju+puŋu/ ('game-killer' a compound noun-phrase in which the items are separated by a morpheme-boundary). The materials were constructed in such a way that each of these word-pairs was phrase-initial in a focused context and phrase-final in a non-focused context. The focus was constructed by means of a question and answer dialogue. The phrase-initial focused contexts that we analysed are shown in bold below, the phrase-final unfocused contexts are underlined. The Warlpiri sentences are given in a broad phonemic transcription.

Dialogue 1. Q: /ɲija ŋanta puŋu napanaŋkaɭu/
(What is Napanangka reported to have killed?)
A. /**kuju#**puŋu ŋanta napanaŋkaɭu/
(Napanangka is reported to have killed **the animal**).
Dialogue 2. Q. /ŋanaŋku ŋanta kuju#puŋu/
(Who is reported to have killed the animal?)
A. /napanaŋkaɭu ŋanta kuju#puŋu/
(Napanangka is reported to have killed the animal).
Dialogue 3. Q. /ɲija pija ŋanta ka ɲina napanaŋka /
(What is Napanangka reported to be?)
A. /**kuju+puŋu** ŋanta ka ɲina napanaŋka /
(Napanangka is reported to be a good **game-killer**).
Dialogue 4. Q: /ŋana ŋanta ka ɲina kuju+puŋu/
(Who is reported to be a good game-killer?)
A: / napanaŋka ŋanta ka ɲina kuju+puŋu/
(Napanangka is reported to be a good game-killer).

Each of these four dialogues formed a block with four other dialogues from the other set of materials that is not analysed in this paper (i.e., 8 dialogues per block). The blocks were repeated 10 times and the dialogues within each block were randomised separately. In each case, the question and answer parts of the dialogue were produced by different subjects. The subjects were asked to produce the answers as a single prosodic phrase and they repeated the answer if they introduced a phrase boundary.

The synchronised acoustic and kinematic data were digitised at sampling frequencies of 20 kHz and 0.5 kHz respectively. The fundamental and formant frequencies of the answer part of each dialogue were segmented and labelled into sub-phonemic units using the EMU system for speech database analysis. As far as the present analysis is concerned, the main segments and events of interest that were extracted are as follows. Firstly, the /j/-target of /kuju/ which was defined to occur at an F2-maximum and which was clearly identifiable in all cases. Secondly, the voiced section of /kuju/ (henceforth /uju/) which extended almost immediately following the release of /k/ (stops are normally unaspirated in Warlpiri) to the onset of the following /p/ closure of /puŋu/. Thirdly, the acoustic closure of /p/ in /puŋu/. Fourthly, /ŋu/ of /puŋu/ which was segmented (sometimes with difficulty) from the preceding /u/ and the following /ŋ/ (in the focused dialogues) based on both an auditory analysis and on sometimes abrupt changes in the formant tracks. Phrase-final /u/ of /ŋu/ was often almost completely elided (especially for TR) resulting in a realisation of this word as [puŋ] or [puŋ:].

The intonation contours in the answers of our corpus matched quite closely Bishop & Fletcher's [2] description of the declarative tune in the Northern Australian language Bininj Gun-wok: a flat hat-pattern and a fall to a low boundary tone. As in [2], the declarative tune's hat pattern sometimes (but less frequently) had a final rise at the right edge of the phrase for subject BP, typically over the extent of [puŋu] when it occurred as the last item in dialogues 2 and 4. The only clearly identifiable pitch-accent occurred within the phrase-initial /kuju/ in dialogues 1 and 3.

Ten (repetitions) x 2 (focused/unfocused) x 2 (word/morpheme boundary) = 40 tokens of /kuju(#/+)puŋu/ were available for analysis for BP. For TR, the same number of dialogues was recorded but due to a technical error, only the first 6 blocks were digitised. For TR, there were therefore 24 tokens available for analysis.

The statistics were derived from an ANOVA with various acoustic parameters as dependent variables and with three independent variables: Speaker (BP or TR), Boundary (/kuju#puŋu/ or /kuju+puŋu/) and Focus (focused or unfocused dialogue). Unless otherwise stated, there was no significant interaction between the independent variables.

3. RESULTS

3.1 /kuju/ As suggested above, focused and unfocused words were differentiated by the presence vs. absence of a pitch-accent that results in a clear F0-peak in the former as shown in Fig. 1. For BP, the timing of the F0-peak also differentiated /kuju#puŋu/ from /kuju+puŋu/ in the focused context: as Fig. 1 shows, the F0-peak preceded the /j/ target by 50 ms whereas in /kuju+puŋu/ it was usually at the /j/ target. TR did not make such a distinction.

ISBN 1-876346-48-5

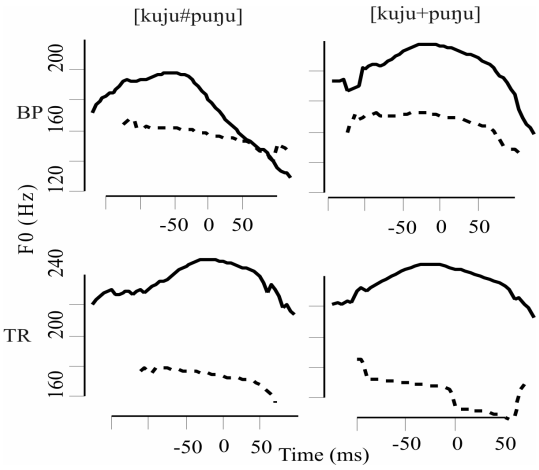

Fig. 1. F0-contours averaged after time-alignment at the /j/-target in focused (solid) and unfocused (dashed) dialogues for speakers BP (row 1) and TR (row 2) in the word-boundary (column 1) and morpheme-boundary (column 2) contexts.

The acoustic duration of /uju/ was significantly greater in the focused than in the unfocused dialogue ($F=21.1$, $p < 0.0001$). A subsequent analysis showed that this expansion of /uju/ was predominantly due to a significantly greater lengthening of the interval from the acoustic onset of /u/ to the /j/-target ($F = 13.6$, $p < 0.001$), i.e. an expansion of the rhyme /uj/. In addition, the duration was significantly greater for the word-boundary compared with the morpheme-boundary context over this rhyme interval ($F = 7.3$, $p < 0.01$). These results, which are summarised in the bar plot in Fig. 2, indicate that the /uj/ duration is greater in /kuju#puŋu/ than in /kuju+puŋu/ regardless of the focusing: the main effect of focusing is to magnify an existing distinction between /kuju#puŋu/ and /kuju+puŋu/ for both speakers, as Fig. 2 shows.

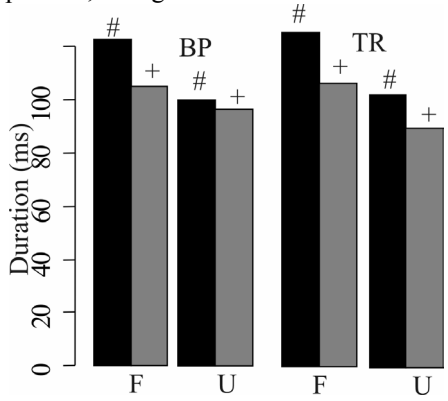

Fig. 2. Average duration from the acoustic onset of /u/ to the F2-target of /j/ in /uju/ in F (focused) and U (unfocused) dialogues and in /kuju#puŋu/ (#, black) and /kuju+puŋu/ (+, grey) contexts for BP (left) and TR (right).

If these duration changes are associated with supralaryngeal expansion, then a greater change in F2 from a low value for the /u/ vowels to a high value for the /j/ target could be expected. For both speakers, the F2-change in /uj/ and in /ju/, as measured from the F2-difference between the acoustic vowel onset/offset and the /j/-target was significantly greater in the focused than the non-focused dialogues (/uj/: $F=13.3$, $p < 0.001$; /ju/: $F=8.7$, $p < 0.01$), with no significant effect in either /uj/ or /ju/ for Boundary. This suggests that focusing is indeed accompanied by greater supralaryngeal/spectral change, whereas the prosodically lower-level contrast between a word and morpheme boundary is differentiated by rhythmic/durational differences of the kind discussed earlier.

We investigated whether the F2-change was primarily due to F2-raising of /j/ or to F2-lowering at vowel onset or offset in /uju/. The results of ANOVAs with the two independent variables Speaker and Focus (i.e. collapsing across Boundary) showed that the F2-peak for /j/ was significantly higher in the focused dialogue ($F=21.2$, $p < 0.0001$) but that there were no significant differences in F2 at the acoustic onset nor at the acoustic offset of /uju/. Taken together, these results imply that the greater F2-change in /uj/ and /ju/ in the focused dialogue is attributable to a higher F2 at the /j/ target, rather than to F2-lowering at the onset or offset of /u/.

The plots in Fig. 3 of F2-trajectories of /uju/ which were first linearly time-normalised and then averaged separately in the focused and the non-focused dialogues, are consistent with this interpretation. For both speakers the trajectories show a rapid change in the vowels towards a very clear /j/-target, which is higher for BP and slightly higher for TR in the focused dialogue (the mean F2 values in the focused and non-focused dialogues are 2102 Hz and 1973 Hz respectively for BP; and 1812 Hz and 1746 Hz respectively for TR). There is no distinct target in the /u/ vowels and, with the exception of F2-offset for BP and F2-onset for TR, little evidence to suggest that F2 is lower (and hence phonetically more peripheral).

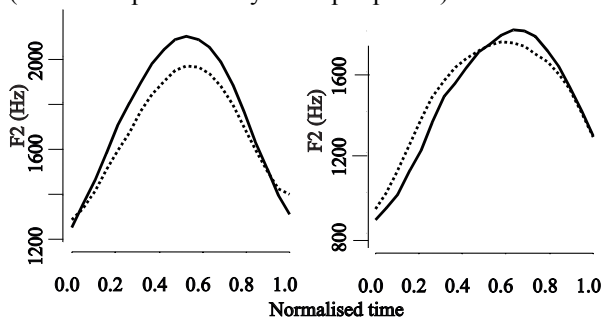

Fig. 3. F2-trajectories in /uju/ for BP (left) and TR (right) which were derived by linear time-normalisation and then averaging in the focused (solid) and non-focused (dotted) dialogues, collapsed across the word/morpheme boundary distinction.

3.2 /puŋu/ This word was analysed for phrase-final lengthening effects and in order to assess whether /p/-closure duration was affected by the presence of a preceding word as opposed to morpheme boundary. The duration of /ŋu/ in the non-focused dialogue, in which it was also phrase-final was significantly greater than in the focused dialogue ($F=38.7$, $p < 0.00001$) with no effect for Boundary.

ISBN 1-876346-48-5

The duration of the /p/-closure was significantly greater in /kuju#puŋu/ than in /kuju+puŋu/ ($F=11.8$, $p < 0.01$). This result, together with the bar plots in Figs. 2 and 4, shows that the boundary effect is preserved for both speakers across the changes in focus.

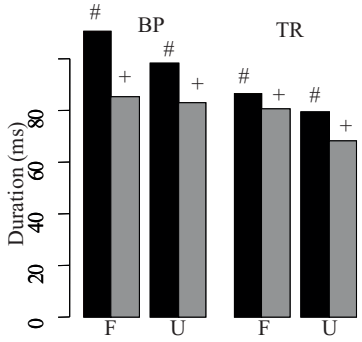

Fig. 4. Average duration of the /p/ closure in F (focused) and U (unfocused) dialogues and in /kuju#puŋu/ (#, black) and /kuju+puŋu/ (+, grey) contexts for BP (left) and TR (right).

DISCUSSION

In phrase-initial position and in a focused context, the initial rhyme of the focused word is lengthened relative to the same word in an unfocused and phrase-medial context. The phrase-initial rhyme is also associated with what seems to be a pitch-accent and we have found some evidence that its consonant is hyperarticulated. In a subsequent experiment, we will need to compare the effects of narrow with broad focus in order to assess whether these characteristics are simply attributable to positional contrast (phrase-initial vs. phrase-medial), or whether the segment lengthening and associated hyperarticulation are induced by narrow focusing. Our analysis has also shown a phrase-final lengthening effect. Additionally, we found that the lower-level contrast that arises from a sequence of two words that are either do or do not form a compound is realised by subtle timing differences that persist independently of their phrasal position (focusing). This is in general consistent with a model in which contrasts at different levels of the prosodic hierarchy leave their own independent signature on the production of speech [5,9,13]. The finding of a longer domain-initial /p/-closure when it is word-initial compared with morpheme-initial is also compatible with Keating et al's model of domain-initial strengthening [8,9], on the assumption that the /p-/ in the compound is prosodically more deeply embedded (and hence weaker according to [8,9]) than when the word is not compounded.

The finding that the *consonant* of the rhyme is hyperarticulated in Warlpiri could suggest a language-specific effect: in accented/focused words in English, it is usually the *vowel* that undergoes lengthening and hyperarticulation [5,6]. These language differences may be reconciled in terms of their different phoneme inventories. In English, there are usually around 20 contrasting vowel phonemes, whereas in Warlpiri there are three plus a restricted length contrast. In Warlpiri, there are also more place of articulation constrasts in stops, nasals and liquids. The hyperarticulation of the vowel in English, but of the rhyme's consonant in Warlpiri would therefore both have the same effect of enhancing the greatest number of contrasts in their own phonemic system. From another perspective, there would be little to be gained in terms of clarity of the linguistic code by enhancing the distinctions between vowels in Warlpiri since, in constrast with English, they carry far fewer meaning contrasts.

REFERENCES

[1] J. Fletcher and N. Evans, "Intonational downtrends in Mayali," *Australian J. of Linguistics*, vol. 20, pp 23-38, 2000.

[2] J. Bishop and J. Fletcher, "Intonation in five dialects of Bininj Gun-wok," In *Prosody and Typology*, S-A. Jun (Ed.), Oxford:OUP, in press.

[3] H. King, *The Intonation of Djirbal*, Lincom Europa, in press.

[4] A. Sluijter and. V. van Heuven, "Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch," *Phonetica*, Vol. 52, pp. 71-89, 1995.

[5] J. Harrington et al., "Manner and place conflicts in the articulation of accent in Australian English," In *Laboratory Phonology V*, J. Broe (Ed.). pp. 40-55, 2000.

[6] K. de Jong, "The supraglottal articulation of prominence in English: linguistic stress as localized hyperarticulation," *J.Acoust.Soc.Am*, Vol. 97, pp.491-504, 1995.

[7] A. Turk and S. Shattuck-Hufnagel, "Word-boundary-related duration patterns in English.," *Journal of Phonetics*, Vol. 28, pp. 397-440, 2000

[8] C. Fougeron, C. and P. Keating, "Articulatory strengthening at edges of prosodic domains," *J.Acoust.Soc.Am*, Vol. 101, pp. 3728-3740, 1997.

[9] P. Keating et al., "Domain-initial articulatory strengthening in four languages," In *Laboratory Phonology VI*, J. Local (Ed.), in press.

[10] T. Cho, "Effects of morpheme boundaries on intergestural timing: evidence from Korean," *Phonetica*, Vol. 58, pp. 129-162, 2001.

[11] D. Nash, *Topics in Warlpiri Grammar*, New York: Garland Publishing, 1986.

[12] M. Laughren et al., *A Learner's Guide to Warlpiri*, Alice Springs: IAD Press, 1996.

[13] M.E. Beckman et al. "Prosodic structure and tempo in a sonority model of articulatory dynamics," In Laboratory Phonology II, G.Docherty and D.R. Ladd (Eds.), Cambridge: CUP, pp. 3-16, 1992.