ACOUSTIC CORRELATES OF LEXICAL STRESS IN WUBUY

Brett Baker[1], Rikke Bundgaard-Nielsen[2], Sarah Babinski[3], Janet Fletcher[1]

University of Melbourne[1], MARCS Institute Western Sydney University[2], Yale University[3]
bjbaker@unimelb.edu.au, rikkelou@gmail.com, sarah.babinski@yale.edu, j.fletcher@unimelb.edu.au

ABSTRACT

We examined the acoustic correlates of lexical stress in the non-Pama-Nyungan language Wubuy (Northern Territory, Australia). We tested two hypotheses about stress: that stress is determined by (1) a combination of syllable position in prosodic word and quantity sensitivity, or (2) by position alone. To test these hypotheses, we elicited trisyllabic noun roots differing in position of heavy syllables in frame-final environments from 3 speakers. We found that both position and predicted stress based on prior phonological descriptions could account for many correlates (segment and syllable duration, f0, intensity, vowel formants) although overall syllable position appeared to account for more of the variance.

Keywords: Stress, weight, Australian languages

1. INTRODUCTION

Stress in Australian Indigenous languages has been the topic of descriptive (impressionistic) and instrumental examination, but has proven somewhat elusive to characterise in terms of standard acoustic correlates. The primary cue to main stress appears to be f0 excursions, not necessarily located on the metrically prominent syllable, with conflicting evidence for other parameters (e.g. see **Error! Reference source not found.** for an overview). Among the correlates which have been suggested are segment duration e.g. Bardi: [10], Kayardild: [11]; syllable duration e.g. Bininj Gun-wok: [3]; vowel quality e.g. Dalabon: [6]; and vowel intensity e.g. Dalabon: [6]. Current analyses of metrical structure have been based on largely impressionistic descriptions of fieldworkers. Most Australian languages are reported to have main stress initially in words or roots [5], [1], although a number of northern Australian languages are reported to prefer penultimate main stress. Long vowels are reported to attract stress in a number of Australian languages, such as Nhanda, Banyjima, and Ngiyampaa [2]. Some Australian languages have been described as having longer consonants following stressed syllables ('post-tonic lengthening': see e.g. **Error! Reference source not found.**). This dataset did not contain a sufficient material to test this hypothesis, although a small sample showed differences in this direction (not reported here). Finally, a small number of Australian languages have been reported to have quantity-sensitive stress determined by closed syllables, as opposed to long vowels, notably Ngalakgan [1], but with suggestions that the pattern may be found more generally in Arnhem Land.

Here, we test the phonological description of stress in noun roots of the non-Pama-Nyungan Australian language Wubuy (a.k.a. Nunggubuyu: [8]). Wubuy is a northern Australian language, and like many of those, appears to have a conflict between initial and penultimate stress in trisyllabic words [1]. Wubuy has previously been analysed in a metrical framework as having quantity-sensitive stress, where heavy syllables are defined as those with long vowels [9]. Impressionistically, however, syllables where a vowel is followed by a heterorganic cluster also appear to attract stress, as in Ngalakgan [1], and it may be unique in having this combination of heavy syllables. Note that [9] treats homorganic nasal-stop clusters as complex segments in Wubuy, and hence, these do not count as closed syllables for the stress algorithm (and compare [1] on Ngalakgan). However, and impressionistically again, the first author hears initial stress on syllables where a vowel is followed by a homorganic cluster, even if the following syllable is heavy, and they may thus be heavy in this position, exceptionally. Under an *a priori* metrical analysis like that of Ngalakgan [1], we predict that trisyllabic roots will be assigned strictly bimoraic, trochaic feet from the left edge, where heavy syllables are defined as those with long vowels or those where the nucleus is followed by a heterorganic cluster, or, in initial syllables, where the nucleus is followed by any cluster. Thus, roots of the form /l̪aŋata/ 'jungle' should be stressed as [ˈl̪aŋata] i.e. (l̪aŋa)ta, where parentheses indicate bimoraic foot boundaries; while a form such as /cuɭurpi/ 'salmon' should be stressed as [cuˈɭurpi], i.e. cu(ɭur)pi, because of the medial closed syllable. Furthermore, a form such as /ciminʈi/ 'harpoon spike' should also receive initial stress, as [ˈciminʈi], because the homorganic cluster does not contribute weight to the preceding syllable (see [1] for discussion implications for syllable theory), and hence the metrical structure is (cimi)ɳʈi. Finally, /ɭuŋkurma/ 'northeast wind' should be stressed as [ˌɭuŋˈkurma] where the initial closed syllable is exceptionally marked as heavy, hence the metrical structure is (ɭuŋ)(kur)ma.

In this study, we test two competing hypotheses. One hypothesis is that stress is assigned initially unless there is a heavy syllable in medial position, as in [1]. The other hypothesis is that stress is simply associated with penultimate position, as in [9]. The acoustic correlates targeted were those identified as being the most indicative cross-linguistically: duration, fundamental frequency (f0), and intensity, in that order, according to a recent survey [7] (See *Section 3.1*). In addition to these acoustic parameters, we include measures of F1 and F2, under the hypothesis that stressed vowels are more acoustically extreme (i.e. hyperarticulated) than unstressed vowels, as in the related language Dalabon [6] (see *Section 3.2*).

2. METHOD

Three literate female native speakers (aged between around 55-65) produced five repetitions of 30 target words in a frame-final position (*nun-jamayn* ___?: 'Did you say ___?), where the stimuli were presented on a laptop screen in the Wubuy orthography. However, the stimulus was visible only for the first production, after which the screen was lowered (in an attempt to avoid reading intonation).

All target items were trisyllabic independent noun roots of Wubuy, containing a variety of open and closed syllables, resulting in a total of 869 analysable syllables. Recordings were hand-segmented into vowel and consonant segments using visible landmarks and the following measures were extracted using *praat* [4]: vowel (VDur) and syllable duration (SyllDur), pitch maximum (PMax) and minimum (PMin), pitch range (PR), and normalised intensity max/segment (IntMax). We classified each syllable as either 'stressed' or 'un-stressed' on the basis of the phonological analysis described in §1.

We aimed to test two competing hypotheses: that stress is assigned on the basis of metrical principles, including quantity sensitivity, or that stress is assigned purely on the basis of position in word (regardless of weight), where the favoured metrical position in Wubuy is penultimate, following [9]. In *Section 3.1*, below, we present the results from an analysis of differences in pitch, intensity and duration in Syll 1 and Syll 2 position, and according to assigned stress, in Wubuy. *Section 3.2* presents the results of the acoustic analyses of differences in vowel F1 and F2, by the three Wubuy vowels /a i u/.

3. RESULTS

3.1. Pitch, intensity and duration in syllables

The dataset provided a total of 465 unstressed syllables (156 in Syllable 1; 309 Syllable 2), and a total of 404 stressed syllables (278 Syllable 1; 126 Syllable 2) (see *Table 1*), where 'stressed' is computed according to the *a priori* metrical analysis involving quantity-sensitivity. We excluded long vowels from the analysis because there were insufficient numbers to validly test the effect of long vowels on acoustic correlates.

We conducted a 2x6 Multivariate Analysis of Variance with the independent variables 'syllable position' (Syllable 1 vs Syllable 2) and 'stress' (Stressed versus Unstressed) and the dependent variables pitch maximum (PMax), pitch minimum (PMin), pitch range (PR), intensity max in the target vowel (IntMax), vowel duration (VDur) and syllable duration (SyllDur). There was a significant effect of 'syllable position' (*DoF* 1, 865 in all cases) for PMax ($F = 27.985$, $p < .001$); PR ($F = 37.779$, $p < .001$); IntMax ($F = 11.004$, $p = .001$); SegDur ($F = 250.297$, $p < .001$); and SyllDur ($F = 653.574$, $p < .001$), with the two latter contributing 22% and 43% of the residual variance, respectively. Similarly, there was a significant effect of 'stress' for PMin ($F = 7.776$, p .005); IntMax ($F = 6.114$, $p = .014$); SegDur ($F = 15.194$, $p < .001$); and SyllDur ($F = 88.772$, $p < .001$) though the effect of 'stress' was much smaller: < 2% of the residual variance for all, except SyllDur for which 'stress' accounted for 9%. There was an interaction between 'syllable position' and 'stress' only for SyllDur ($F = 33.376$, $p < .001$).

Table 1: Pitch, intensity and duration differences according to the factors of syllable position (1 = initial, 2 = penultimate) and stress.

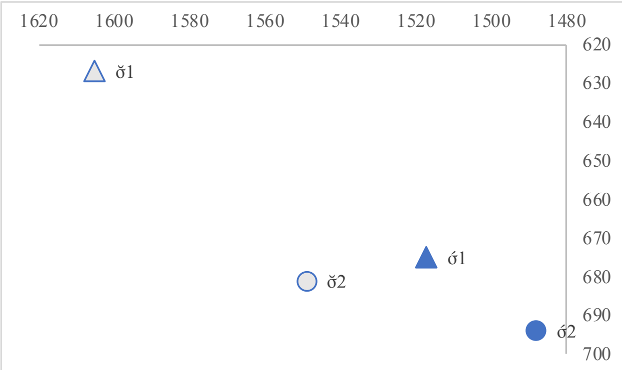
	Syll #	Unstr. M(SD)	Str. M(SD)
PMax	1	178 (15)	180 (16)
	2	185 (19)	187 (16)
PMin	1	165 (16)	168 (14)
	2	166 (21)	170 (16)
PR	1	14 (9)	13 (7)
	2	19 (14)	17 (8)
IntMax	1	81 (3)	82 (3)
	2	82 (3)	82 (4)
VDur	1	84 (17)	N/A
	2	117 (29)	124 (41)
SyllDur	1	168 (37)	180 (45)
	2	235 (53)	287 (47)

3.2. F1 and F2 in syllables

The dataset provided a total of 590 tokens of /a/ (109 unstressed and 233 stressed tokens in Syllable 1; 188 unstressed and 60 stressed tokens in Syllable 2). It also provided 94 /i/ tokens (15 unstressed and 15 stressed tokens in Syllable 1; and 30 unstressed and 34 stressed tokens in Syllable 2). Finally, the dataset provided 185 /u/ tokens (32 unstressed and 30 stressed in Syllable 1; 91 unstressed and 32 stressed in Syllable 2). Mean F1 and F2 values (including standard deviations in parentheses) are presented in *Table 2*.

We conducted a series of three Multivariate Analyses of Variance, treating each vowel separately (*DoF* 1, 585 for /a/; *DoF* 1, 90 for /i/; and *DoF* 1, 181 for /u/).In the case of /a/, the MANOVA indicated that there was a significant effect of ‘syllable position’ for both F1 (F = 19.427, p < .001), and F2 (F = 5.433, p = .02), though the effect was very small for both (3% for F1; 1% for F2). There was also an effect of ‘stress’ for both F1 (F = 13.373, p < .001) and F2 (F = 16.504, p < .001), though, again, the effect was small (2% for F1; 3% for F2, respectively); see *Figure 1*.

Figure 1: Formant distribution of stressed (darker shades) and unstressed (paler shades) /a/ in syllable 1 (triangles) and syllable 2 (circles).



In the case of /i/, the MANOVA indicated that there was a significant effect of ‘syllable position’ for both F1 (F = 4.529, p = .036) and F2 (F = 49.785, p < .001), the first of which accounts for 5% and the latter for 36% of the residual variance. Closer inspection of the *Mean* unstressed F2 values (see *Table 2*) provide some explanation for this finding: the F2 value of unstressed /i/ in Syllable 2 is notably centralised in our dataset. This is likely due to /i/ being followed by retroflex segments in all tokens in this position, resulting in a lower F2 value for the target due to anticipatory co-articulation. There was also a significant effect of ‘stress’ for F2 (F = 30.462, p < .001) (here, accounting for 25% of the variance). There was a significant interaction for both F1 (F = 4.315, p = .041) and F2 (F = 5.809, p = .018); see *Figure 2*.

Finally, in the case of /u/, the MANOVA indicated that there was also an effect of ‘syllable position’ for F1 (F = 9.485, p = .002) and F2 (F = 61.71, p < .001), with ‘syllable position’ accounting for 25% of the residual variance observed. There was also an effect of ‘stress’ for F2 (F = 4.136, p = .043), accounting for 2% of the residual variance. There was a significant interaction for F2 (F = 5.765, p = .017); see *Figure 3*.

Table 2: Formant differences in syllable 1 and 2.

Formant	Syll #	Stress	*M (SD)*	*N*
F1 /a/	**1**	**Ustr.**	627 (96)	109
		Str.	675 (94)	233
	2	**Ustr.**	681 (76)	188
		Str.	694 (83)	60
F2 /a/	**1**	**Ustr.**	1605 (225)	109
		Str.	1517 (175)	233
	2	**Ustr.**	1549 (180)	188
		Str.	1488 (242)	60
F1 /i/	**1**	**Ustr.**	389 (17)	15
		Str.	356 (18)	15
	2	**Ustr.**	390 (68)	30
		Str.	395 (21)	34
F2 /i/	**1**	**Ustr.**	2143 (227)	15
		Str.	2338 (100)	15
	2	**Ustr.**	1549 (290)	30
		Str.	2046 (345)	34
F1 /u/	**1**	**Ustr.**	401 (40)	32
		Str.	403 (27)	30
	2	**Ustr.**	422 (57)	91
		Str.	429 (35)	32
F2 /u/	**1**	**Ustr.**	1265 (269)	32
		Str.	1255 (215)	30
	2	**Ustr.**	968 (134)	91
		Str.	1096 (124)	32

Figure 2: Formant distribution of stressed (darker shades) and unstressed (paler shades) /i/ in syllable 1 (triangles) and syllable 2 (circles).

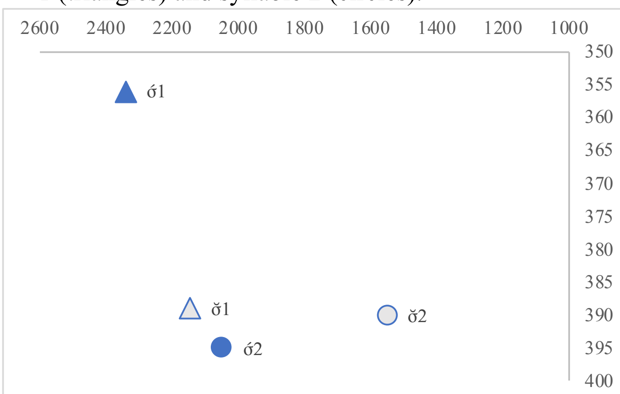
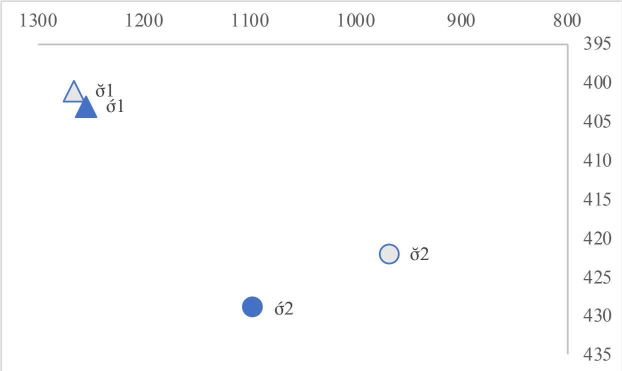


Figure 3: Formant distribution of stressed (darker shades) and unstressed (paler shades) /u/ in syllable 1 (triangles) and syllable 2 (circles).



4. DISCUSSION

Our study was designed to test two competing hypotheses about lexical stress in Wubuy. The first, following [1], would predict initial stress in trisyllabic roots except those containing a medial closed syllable, which should behave as metrically heavy. The second, following [9], predicts consistent penultimate stress in trisyllabic roots. On balance, the second hypothesis receives more support from the acoustic measures tested with this dataset, although the picture is not fully consistent. On the one hand, the location of stressed syllables in a word according to hypothesis 1 appears to have effects on F0, intensity, segment and syllable duration. On the other hand, raw syllable position (which does not take into account potential syllable weight), also has effects on these correlates, and in most cases the effects are greater. Syllable position (initial vs penultimate) and stress also interact, and this interaction is perhaps most prominent in its effects on syllable duration, where we see (Table 1) that a second syllable which is also stressed (according to hypothesis 1) is much longer than either an unstressed second syllable (by an average of 52 ms) or a stressed initial syllable (by an average of 107 ms). The magnitude of difference between an unstressed initial syllable and a stressed second syllable is also greater than that between Syllable 1 and Syllable 2 overall, which is 176 ms in Syllable 1 vs. 250 ms in Syllable 2, a difference of around 75 ms. In order to test whether there is a consistent effect on syllable duration, however, we would need a much more balanced dataset, taking into account syllable structure and segments.

The effects on vowel quality, reported in *Section 3.2*, are also mixed. As with the results for pitch, intensity and duration, we find that predicted stress according to hypothesis 1 and raw syllable position both contribute to the variation, with the added (and unexamined) factor of coarticulation from neighbouring consonants contributing to the difficulty in assigning this variation to one or the other variable. For the low vowel /a/ (Figure 1), it appears that stress results in more open and back vowels, relative to their unstressed counterparts. The results for the high front vowel /i/ are difficult to interpret: stressed vowels in syllable 1 are significantly more peripheral, but stressed vowels in Syllable 2 are much like unstressed vowels in Syllable 1. Coarticulation effects are presumably at work here, but these results are also reminiscent of similar findings for other Australian languages including Pitjantjatjara [12]. In the case of the high back vowel /u/, it is again difficult to see any clear picture, although syllable position appears to play a much greater role than stress. Stressed and unstressed /u/ in Syllable 1 are in much the same position, while /u/ in syllable 2 appears to be lower and more back, relative to the qualities in syllable 1.

Finally, some impressionistic comments on pitch. The F0 contours in this dataset have the appearance of a phrasal boundary tone consistently realized as either HL% or HH% (depending on speaker and utterance), aligned to the final two syllables, largely irrespective of the location of the predicted stressed syllable. It is possible that this reflects a task effect, such that the frame-final position has produced pitch contours that largely reflect phrasal, rather than lexical, accent. Further data collection using a range of intonational frames should help to examine this hypothesis.

In conclusion, we think there are still a number of issues remaining to be explored with respect to stress in Wubuy. In particular, the existence and correlates of quantity-sensitive stress remain elusive. While syllable position can account for much of the variance that we observe in the data, however, there are some interactions (such as syllable duration) that produce greater effects in combination with stress than we would expect if stress were a purely illusory phenomenon.

5. REFERENCES

[1] Baker, B. 2008. *Word structure in Ngalakgan*. Stanford: CSLI.

[2] Baker, B. 2014. Word structure in Australian languages. In: Koch, H., Nordlinger, R. (eds), *The Languages and Linguistics of Australia.* Berlin: Mouton de Gruyter, 139-213.

[3] Bishop, J. 2002. Aspects of intonation and prosody in Bininj Gun-wok. University of Melbourne, PhD thesis.

[4] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer. Retrieved from http:\\www.praat.org.

[5] Dixon, R. 2002. Australian languages. Cambridge: Cambridge University Press.

[6] Fletcher, J., Evans, N. 2002. An acoustic phonetic analysis of intonational prominence in two Australian languages. *JIPA* 32, 123–140.

[7] Gordon, M., Roettger, T. 2017. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard* 3, 1–11.
[8] Heath, J. 1984. Functional grammar of Nunggubuyu. Canberra: AIAS.
[9] Hore, M. 1981. Syllable length and stress in Nunggubuyu. In B. Waters (ed.) Papers in Australian phonologies. Darwin: SIL.
[10] Katsika, A. 2008. Acoustic correlates of primary stress in Bardi. Unpublished paper. Yale University.
[11] Round, E. 2012. Durational correlates of prosodic structure in Kayardild vowels. *Proc. 14th SST*, Sydney, 221–224.
[12] Tabain, M., Fletcher, J., Butcher, A. 2014. Lexical stress in Pitjantjatjara. *JPhon* 42, 52–66.