Acoustic correlates of prominence in Nafsan

Rosey Billington[12], Janet Fletcher[12], Nick Thieberger[12], Ben Volchok[2]

[1]Centre of Excellence for the Dynamics of Language
[2]The University of Melbourne

rbil@unimelb.edu.au; janetf@unimelb.edu.au; thien@unimelb.edu.au; volchokb@unimelb.edu.au

Abstract

Though Oceanic languages are often described as preferring primary stress on penultimate syllables, many different patterns have been noted across and within language families, and may interact with segmental and phonotactic factors. This is exemplified across linguistically diverse Vanuatu. However, both impressionistic and instrumentally-based descriptions of prosodic patterns and their correlates are limited for languages of this region. This paper presents preliminary acoustic and durational results for Nafsan, an Oceanic language of Vanuatu, which suggest a preference for prominence at the right edge of words, with fundamental frequency as a primary correlate.

Index Terms: Oceanic, stress, duration, f0, intensity, F1, F2

1. Introduction

Nafsan (South Efate) is a Southern Oceanic language spoken by an estimated 6,000 people in three villages (Erakor, Eratap and Pango) on the island of Efate in Vanuatu. Though the phonology of Nafsan has been discussed in comparative and descriptive work [1] [2] [3], some challenges remain in understanding Nafsan segmental and prosodic patterns. Key questions are whether there is a vowel length distinction, why some vowels undergo deletion, what patterns of word- and phrase-level prominence are used, and how different parts of the sound system interact. This study is part of a wider project using instrumental phonetic approaches to address these questions, and presents selected results focusing on prominence patterns within words.

1.1. Segmental inventory of Nafsan

Nafsan has an inventory of 15 consonant phonemes, and has five contrastive vowel qualities, as is typical of Oceanic languages [3] [4]. Though the possibility of a length distinction has been mentioned in previous work [2], the status of length within the Nafsan vowel system was until recently unresolved. New phonetic data allowing for a targeted investigation of Nafsan vowels provides evidence for a monophthong inventory comprising /i, iː, e, eː, a, aː, o, oː, u, uː/; each of the five vowel qualities may occur either phonemically short or long, in various syllable types. Ongoing experimental work suggests that at least in CVC syllables, long vowels are close to twice as long as short vowels, and the duration difference between long and short vowels is approximately the same across all five vowel qualities [5].

1.2. Vowel deletion and phonotactic patterns in Nafsan

Nafsan phonotactic patterns are strikingly complex both compared to patterns for languages spoken further to the north in Vanuatu, and compared to the more typologically common preference for CV syllables among Oceanic languages [2] [4]. The language exhibits a range of heterorganic consonant clusters in syllable onsets, with various possible sonority profiles [3] [5]. There is some evidence, through comparisons with historical records and with cognates in closely-related languages, that these complex syllable onsets may have arisen through the deletion of selected medial vowels, but the status of and environments for vowel deletion have remained unclear [3]. Ongoing work indicates that vowel deletion is both a historical and productive process, and that at least for the productive process as used by contemporary speakers, short vowels are frequently deleted when they occur in the penultimate syllable of the word, though this appears to be mediated by lexical, grammatical, and speaker-specific factors [6]. These observations raise the question of whether vowel deletion in Nafsan, when it occurs, is pre-tonic, but given that prominence patterns in Nafsan remain under-described, this has not been clearly established.

1.3. Stress in Oceanic languages

It is often observed that many Oceanic languages display primary stress on penultimate syllables [4], but crosslinguistic examinations suggest that this tendency is not as widespread as previously thought [7]. Contemporary Oceanic languages exhibit a range of prominence patterns, including stress which is regularly penultimate, generally penultimate but final if the final syllable contains a coda and/or a long vowel or diphthong, final, initial, antepenultimate, lexically specified, or dependent on morphological factors [7]. All of these have been reported for at least some of the 130+ languages of Vanuatu, and varied prominence patterns are noted even across closely-related languages, such as Nafsan and neighbouring varieties [8, 9, 10]. For Nafsan, previously suggested patterns include initial and final stress, though the unclear status of vowel length has been a complicating factor [3, 11], and recent auditory impressionistic analyses suggest possible final prominence [5], but this has not yet been examined experimentally. Crosslinguistically, various acoustic and durational cues may correlate with lexical prominence [12], but for the languages of Vanuatu, impressions of stress correlates are only occasionally noted, and have not yet been supplemented by instrumental data. Perceived correlates include combinations of increased duration, pitch, and intensity/loudness for stressed vowels or syllables [13, 14, 10], but in some cases, for languages with a vowel length contrast such as Anejom̃, duration may not be a salient cue [15, 16].

2. Research aims

Given that recent work establishes that vowel length is distinctive in Nafsan, and that vowel deletion patterns and auditory impressions raise the possibility of word-final prominence, two key questions emerge. Are there phonetic differences in the realisation of vowels in final syllables, on the basis of duration, intensity, fundamental frequency, and formant frequencies, which suggest they are more prominent than preceding vowels? Relatedly, do long and short vowels show similar phonetic characteristics in the same word position or are they treated differently?

3. Method

3.1. Participants

The participants in this study were four adult speakers of Nafsan from Erakor village in Efate, Vanuatu: three men (GK, LE, MJ) and one woman (MK). All identify Nafsan as their first language, and the language they use at home. In addition, all speak Bislama, the English-lexified creole which is a lingua franca in Vanuatu, have knowledge of either English and a little French or French and a little English, and also have some knowledge of language varieties spoken in other parts of Vanuatu.

3.2. Materials and procedures

A set of two-syllable and three-syllable word forms was compiled as stimuli, drawing on existing databases and corpora for Nafsan. The words were selected to comprise only CV(C) structures (no complex onsets), and to have long and short vowels in different word positions, to allow investigation of whether these different phonotactic structures influence potential prominence patterns. The majority of the vowels were open /a, aː/. Results presented here pertain to two-syllable words only, with the following structures: CV.CVC, CV.CVVC, CVV.CVC, and CVV.CVVC. Examples include /rakat/ 'bite (DUAL)', /rapaːk/ 'delouse (DUAL)', /kaːkas/ 'be sweet', and /taːk͡paːr/ 'sin'. The duration values of vowels in initial and final syllables are not directly comparable in these structures, given that initial syllables are open and the final syllables are closed, but it is the characteristics of non-final CV syllables which are of particular interest, given that this is an environment in which vowel deletion occurs (though deletion is not attested for vowels in the word forms included here). In addition, though medial CVC syllables are possible in Nafsan, they are less frequent, and word-final CV syllables are much less common than final CVC syllables.

Each word was recorded three times in a medial frame, following a spoken prompt; the frame was *komam util ______ sernrak* 'We say ______ usually'. Recordings were made in a sheltered area during fieldwork in Erakor, and have been archived with other recordings relating to the wider project on Nafsan phonetic and phonological patterns [17]. Data were recorded at an archival sampling rate of 96kHz and 24-bit depth, using a Zoom H6 audio recorder and a Countryman H6 headset microphone with a hypercardioid polar pattern, and downsampled to 44.1kHz 16-bit for analysis. Tokens were balanced across speakers but not vowel length category and word structure (Table 1), given the limitations of available lexical data and the frequency with which different structures occur. The final dataset contained 1,114 vowel tokens drawn from 557 utterances produced (containing 46 different Nafsan words as the target word).

Table 1: *Number of tokens in dataset.*

word shape	initial /V/	final /V/	initial /Vː/	final /Vː/
CV.CVC	229	229	-	-
CV.CVVC	125	-	-	125
CVV.CVC	-	128	128	-
CVV.CVVC	-	-	75	75
total	**354**	**357**	**203**	**200**

3.3. Data processing and analysis

Utterances were transcribed orthographically in Praat [18], and orthographic transcriptions were converted to phonemic transcriptions in the Speech Assessment Methods Phonetic Alphabet (SAMPA). Using the TextGrid files and associated WAV files, automatic segmentation of the speech signal was performed via the web interface of the Munich Automatic Segmentation System (WebMAUS) [19], using the language-independent model for segment identification. Segment boundaries in the output TextGrid files were checked and manually corrected where necessary with reference to wideband spectrograms and corresponding waveforms in Praat. A hierarchical database was constructed using the EMU Speech Database Management System [20], including tiers for the phonemic segments, syllables, and words. The acoustic and durational characteristics of vowel tokens produced in the target words were queried and analysed using the `emuR` package in R [21, 22]. Measures of interest in this study are vowel duration (in ms), intensity at vowel midpoints relative to the midpoint of the (typically light) coda lateral in the preceding word (root mean square amplitude, in dB), fundamental frequency at vowel midpoints (in Hz), and first and second formant frequencies at vowel midpoints (in Hz). The data were tested with linear mixed-effects models using the `lme4` package [23] with random slopes and intercepts for speaker and word. Model validity was checked using a likelihood ratio test, and differences are reported based on Tukey's Honest Significant Difference post-hoc tests.

4. Results

4.1. Duration

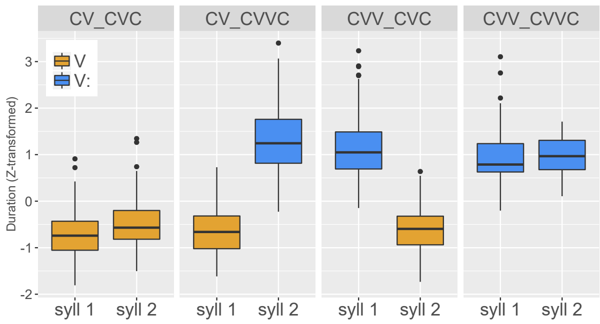

Figure 1: *Duration (Lobanov-normalised) of short and long vowels in initial and final syllables of disyllabic Nafsan words (by word structure).*

Duration values for vowels, shown in Figure 1, clearly illustrate the vowel length distinction across the different word structures in this data. The likelihood ratio test shows that there is a significant effect of context ($\chi^2(7)$= 1304, p<0.001), but post-hoc tests reveal that almost all significant differences are between vowels whose phonemic length differs. Final long vowels are an estimated 72 $\pm$ 2 ms longer than initial short vowels in CV.CVVC words (p<0.001), and initial long vowels are an estimated 61 $\pm$ 2 ms longer than final short vowels in CVV.CVC words (p<0.001). There are no significant differences between long vowels in any of their four contexts. For short vowels, there is a small but significant difference of 8 $\pm$ 2 ms between initial and final short vowels in CV.CVC words (p<0.001), but there are no other notable differences. Though larger differences may be observed for initial and final vowels of the same quantity in data with open rather than closed final syllables, the distributions shown here suggest that any duration differences correlating with word position are likely much smaller than those correlating with phonemic length. Vowel quality was not found to be a significant factor affecting duration when statistical models were compared (recalling that most tokens, 91%, were /a, aː/). Short vowels were on average 61ms in initial syllables and 66ms in final syllables, and long vowels 126ms in initial syllables and 134ms in final syllables.

4.2. Intensity

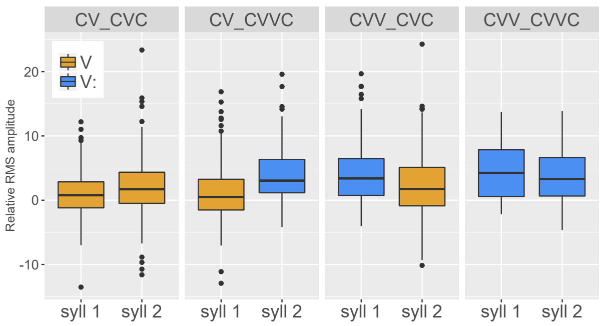

Figure 2: *Root mean square amplitude at midpoints of short and long vowels (relative to midpoints of sonorant coda in preceding word) in initial and final syllables of disyllabic Nafsan words (by word structure).*

Intensity patterns for vowels, based on the difference in root mean square (RMS) amplitude between the vowel midpoint and the midpoint of the sonorant coda in the preceding word, are shown in Figure 2. The effect of context is significant ($\chi^2(7)$= 88, p<0.001). However, post-hoc tests show that there are no significant differences between vowels of the same phonemic length in CV.CVC words (p=0.14) and CVV.CVVC words (p=0.97), nor between any vowels of the same phonemic length in other comparisons. For CV.CVVC words, there is a significantly larger increase in intensity for the final long vowels than the initial short vowels, by an estimated 2.9 ± 0.4 dB (p<0.001). For CVV.CVC words there are also significant differences but with higher values for the initial long vowel rather than the final vowel, of an estimated 1.7 ± 0.4 dB (p<0.001). As for duration, then, differences on the basis of RMS amplitude correlate more with vowel length than word position.

4.3. Fundamental frequency

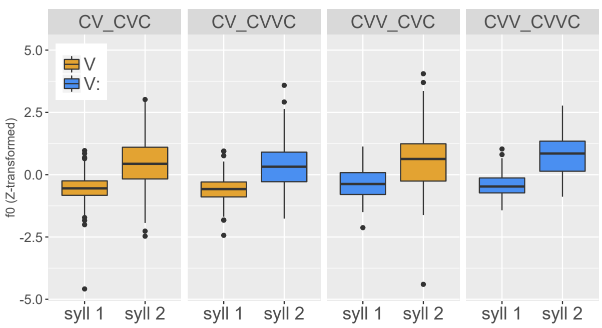

Figure 3: *Fundamental frequency (Lobanov-normalised) at midpoints of short and long vowels in initial and final syllables of disyllabic Nafsan words (by word structure).*

For fundamental frequency (f0), measured at the vowel midpoint, a subset of the data was used; 74 vowel tokens which returned zero values, due to breathiness or devoicing of the vowel, were excluded. Of these, 70% were from the same speaker, whose speech rate was noticeably faster than that of the other participants. 92% were short vowels; 45% of these were short vowels in initial syllables, and 49% were short vowels in final syllables. Results for the remaining 1,040 tokens are shown in Figure 3, and a consistent pattern can be seen across the four word structures. The effect of context is significant ($\chi^2(7)$= 361, p<0.001), and post-hoc tests confirm that there are significantly higher f0 values in final compared to initial syllables in each case. The differences are also of a similar magnitude; vowels in final syllables are an estimated 15 ± 1 Hz higher in CV.CVC words (p<0.001), 14 ± 2 Hz higher in CV.CVVC words (p<0.001), 14 ± 1 Hz higher in CVV.CVC words (p<0.001) and 19 ± 2 Hz higher in CVV.CVVC words (p<0.001). There are no significant f0 differences between vowels of the same phonemic length occurring in the same word position.

4.4. First and second formant frequency

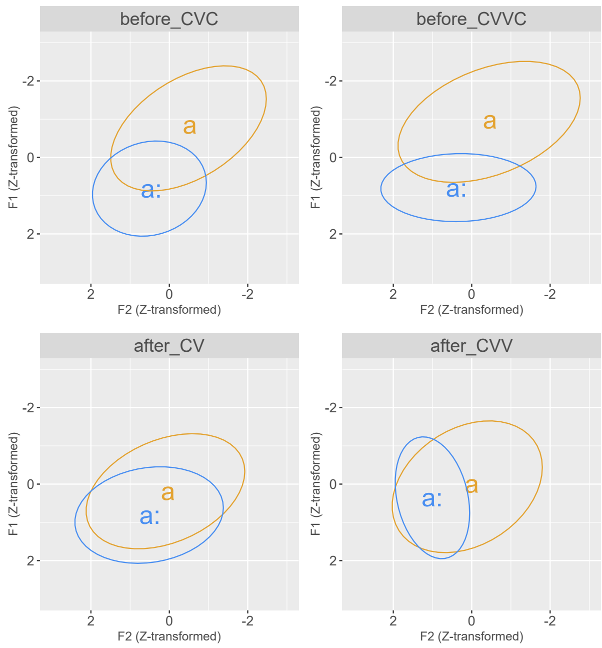

Figure 4: *First and second formant frequency (Lobanov-normalised) at midpoints of short and long vowels in initial and final syllables of disyllabic Nafsan words (by word position).*

A subset of the data containing only open vowels /a, aː/, which as noted comprised the majority of the data (1,012 tokens), was used to examine any differences in first formant frequency (F1) and second formant frequency (F2). Results for F1 and F2 measured at the vowel midpoint are shown in Figure 4. Likelihood ratio tests show that the effect of context is significant for both F1 ($\chi^2(7)$= 587, p<0.001) and F2 ($\chi^2(7)$= 217, p<0.001). The results of post-hoc tests include significant differences between short and long vowels in initial syllables; preceding CVC syllables, short /a/ has F1 values an estimated 189 ± 12 Hz lower than long /aː/ (p<0.001), and F2 values an estimated 144 ± 31 Hz lower (p<0.001). Preceding CVVC syllables, short /a/ similarly has F1 values an estimated 196 ± 21 Hz lower than long /aː/ (p<0.001), but F2 differences are not significant (p=0.16). In final syllables, the formant characteristics of /a/ and /aː/ differ less. Following CV syllables, F1 values for /a/ are an estimated 76 ± 13 Hz lower than for /aː/ (p<0.001); this is smaller than the quite substantial F1 difference in initial syllables, and there are no significant F2 differences between /a/ and /aː/ in this context (p=0.6). Following CVV syllables, there are no significant differences between short and long vowels on the basis of F1 (p=1.00) or F2 (p=0.5).

5. Discussion and conclusions

These findings provide compelling evidence that disyllabic Nafsan words are more prominent at the right edge. The consistent pattern of higher f0 values in final compared to initial syllables, regardless of the phonemic length of the vowel, suggests that f0 likely plays an important role in prominence marking. Results for F1 and F2 at midpoints of vowels in different contexts offer supporting evidence for right-edge prominence; in final syllables, the spectral differences between /a/ and /aː/ are minimal, but in initial syllables, F1 values are substantially lower for short /a/, and F2 values are somewhat lower, indicating centralisation and some degree of retraction for these short vowels. This is of particular interest given impressions that where productive vowel deletion occurs, it is generally of penultimate short vowels in CV syllables [2, 3, 6]. Penultimate short vowels in this dataset are not deleted, but still show some reduction.

Duration and intensity are likely only minor cues to prominence; the only difference between vowels of the same phonemic length was for the duration of initial and final short vowels in CV.CVC words, and the size of the estimated difference was very small in this case. Instead, duration and intensity appear to be robust correlates of vowel length and bolster evidence for the status of quantity distinctions in Nafsan [5]. Vowels identified as being phonemically long show reliably and substantially higher duration values than short vowels in different contexts, as well as higher intensity values as indicated by RMS amplitude at vowel midpoints. While there may be larger duration differences correlating with word position in data where both the initial and final syllables are either open or closed, duration differences correlating with vowel length are likely to be preserved regardless of syllable prominence. Similar observations have been made based on exploratory duration measures for vowels in the closely-related language Lelepa [10].

The results presented here accord with other impressions of final prominence in Nafsan [11, 5], and in pointing towards the importance of f0 as a major correlate, also suggest the possibility that prominence marking in Nafsan is more like that of languages such as Japanese which have been described as having non-stress accent [24]. Given that other preliminary work on Nafsan indicates that high f0 targets at the right edge of a word may demarcate the right edge of an accentual phrase, it will also be worth considering whether the language has a prosodic system like that of Korean or French, with high tone targets at right edges relating to constituents that are not necessarily lexical. This is the subject of ongoing research. The present results for vowels produced in an utterance-medial frame controlling for word length and syllable structure will provide a useful point of reference in ongoing work, which includes comparisons of the phonetic characteristics of Nafsan words of different lengths and syllable structures in initial, medial and final contexts.

Oceanic languages are under-represented in prosodic research, and as overviews of proposed prominence patterns show, there is much that remains to be understood [7]. The languages of Vanuatu are especially well-suited to investigations of the different ways that segmental and prosodic phenomena interact; they show enormous diversity in their sound systems, and appear to have responded in different ways to various sound changes which can be traced back to Proto-Oceanic. These results show that a more detailed understanding of language-internal prosodic patterns may offer insights into the phonetic mechanisms underpinning historical changes, while also contributing to a better understanding of the typological profile of these languages.

6. Acknowledgements

Sincere thanks to all the Nafsan speakers who have participated in and facilitated this and earlier work, in particular Gray Kaltap̃au, Lionel Emil, Michael Joseph and Marinette Kalpram for their involvement in this study, and Carol Lingkary and Yvanna Ataurua for their assistance during fieldwork in Erakor. This research was conducted with support from the ARC Centre of Excellence for the Dynamics of Language (Project ID: CE140100041).

7. References

[1] Clark, R. "The Efate dialects", Te Reo, 28:3–35, 1985.

[2] Lynch, J. "South Efate phonological history", Oceanic Linguistics, 39(2):320–338, 2000.

[3] Thieberger, N. A grammar of South Efate: An Oceanic language of Vanuatu. University of Hawaii Press, 2006.

[4] Lynch, J., Ross, M., and Crowley, T. The Oceanic languages. Curzon Press, 2002.

[5] Billington, R., Fletcher, J., and Thieberger, N. "Nafsan", submitted.

[6] Billington, R., Thieberger, N., and Fletcher, J. "Phonetic evidence for phonotactic change in Nafsan", submitted.

[7] Lynch, J. "Reconstructing Proto-Oceanic stress", Oceanic Linguistics, 39(1):53–82, 2000.

[8] Schütz, A. Nguna grammar. University of Hawaii Press, 1969.

[9] Sperlich, W. B. Namakir: A description of a Central Vanuatu language. PhD thesis, University of Auckland, Auckland, New Zealand, 1991.

[10] Lacrampe, S. Lelepa: Topics in the grammar of a Vanuatu language. PhD thesis, Australian National University, Canberra, Australia, 2014.

[11] Capell, A. The Nguna-Efate dialects, 1930-1980. http://catalog.paradisec.org.au/repository/AC2/VNEFAT11.

[12] Gordon, M. K. and Roettger, T. "Acoustic correlates of word stress: A cross-linguistic survey", Linguistics Vanguard, 3(1):1–11, 2017.

[13] Guy, J. B. M. A grammar of the northern dialect of Sakao. Number 33 in Series B. Pacific Linguistics, 1974.

[14] Jauncey, D. G. Tamambo, the language of west Malo, Vanuatu. Pacific Linguistics, 2011.

[15] Lynch, J. A grammar of Anejom̃. Pacific Linguistics, 2000.

[16] Lunden, A., Campbell, J., Hutchens, M., and Kalivoda, N. "Vowel-length contrasts and phonetic cues to stress: An investigation of their relation", Phonology, 34:565–580, 2017.

[17] Billington, R. Rosey Billington Nafsan materials (BR1), Digital collection managed by PARADISEC, 2017.

[18] Boersma, P. and Weenink, D. Praat, 2018. Version 6.0.40.

[19] Kisler, T., Reichel, U., and Schiel, F. "Multilingual processing of speech via web services", Computer Speech & Language, 45:326–347, 2017.

[20] Winkelmann, R., Harrington, J., and Jänsch, K. "EMU-SDMS: Advanced speech database management and analysis in R", Computer Speech & Language, 45:392–410, 2017.

[21] R Core Team. R: A language and environment for statistical computing, 2018. Version 3.4.2.

[22] Winkelmann, R., Jänsch, K., Cassidy, S., and Harrington, J. emuR: Main package of the EMU Speech Database Management System, 2018. R package, Version 1.0.0.

[23] Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. "Fitting linear mixed-effects models using lme4", Journal of Statistical Software, 67(1):1–48, 2015.

[24] Beckman, M. Stress and non-stress accent. Foris Publications, 1986.