PROSODIC MARKING OF FOCUS IN NAFSAN

Janet Fletcher[12], Rosey Billington[12], Nick Thieberger[12]
[1]Centre of Excellence for the Dynamics of Language, [2]The University of Melbourne
janetf@unimelb.edu.au; rbil@unimelb.edu.au; thien@unimelb.edu.au

ABSTRACT

Languages use a variety of means to realise informational structure categories like topicalisation and focus. The interaction between prosody and focus realisation strategies was examined in Nafsan, a Southern Oceanic language of Vanuatu, in a series of tasks that were designed to explore prosodic realisation of informational and contrastive focus on nouns that were subjects or objects in mini-dialogues where word-order was manipulated. All speakers produced utterance-initial or utterance-final focal elements with a major pitch movement associated with the focused noun (subject or object). Focused nouns were also realised with a wider pitch and often realised in their own prosodic phrase compared to the same item in non-focal contexts. There was also significant syllable lengthening at the right edge of in-focus words. In utterance-initial contexts, post-focal material in Nafsan was almost always produced in a relatively compressed pitch range and there was evidence of de-phrasing of non-focal nouns regardless of utterance position, suggesting prosodic phrasing patterns similar to other languages with edge-marking prominence.
Keywords: Nafsan, focus, accentual prominence, intonation

1. INTRODUCTION

In the last ten years or more, there has been an expansion of prosodic analyses of less well-studied languages (e.g. [13, 5, 4]). Compared to well-resourced European and Asian languages, only a handful of investigations have examined the interaction between prosody and information structure in Oceanic languages, with some notable exceptions (e.g. [3] for Samoan). It is generally accepted that tonal variation in languages like English is influenced by a combination of information structure and pragmatics. However, the phonological intonational devices that languages use to contrast narrow or contrastive versus broad focus are known to vary. These might include combinations of the following: manipulations of phrase-level pitch range (incorporating pitch level and pitch span after [8]), intonational and prosodic phrasing, and intonational prominence, including the use of different types of pitch accents for contrastive emphasis. Languages can also de-accent material (reducing the number of pitch accents in a phrase) and/or de-phrase non-focal material (reducing the number of intonational constituents) to promote a particular kind of discourse interpretation (after [5]).

Languages can use syntactic means to realise informational structure categories like topicalisation and contrastive focus. These devices include left dislocation of the constituent under focus, which has been noted in some Oceanic languages (e.g. in Vera'a, [12], and Nafsan, [14]). It has also been suggested that in such cases, intonation plays a lesser role in the realisation of semantic focus (compared to west-Germanic languages) with patterns of prosodic variation primarily the result of positional factors. In other words, if a language promotes left dislocation as a topicalisation or contrastive focus-marking strategy, the resulting prosodic patterns are because the item under focus is in initial position in a discourse segment. By contrast, others have suggested that so-called free word order languages also employ intonational devices (e.g. [17]), implying that there is a deliberate prosodic strategy to place a constituent in focus. Recent explorations of the complex interplay between prosody, pragmatics, and syntax in Samoan suggest prosodically-driven syntactic fronting is an important feature of information structure realisation in this language ([3]).

In this paper, we examined the intonational features of subject and object focus realisation in Nafsan, a Southern Oceanic language spoken by around 6,000 people on the island of Efate in Vanuatu. Nafsan has previously been analysed as a stress language although this is still under investigation (e.g. [1]). Many Oceanic languages have been impressionistically described as having penultimate main stress [9]. There has been relatively little quantitative investigation of lexical and post-lexical prosody in the languages of Vanuatu. Results of an initial analysis of accentual prominence in Nafsan suggest that final syllables rather than penultimate syllables exhibit higher F0 values in disyllables [1]. Earlier impressionistic analyses of intonation patterns suggest that the prosodic typology of Nafsan may be edge-marking, akin to Korean or French, rather than head-marking, like English or Dutch (see summary in [5]) but this remains to be fully investigated. The purpose of the current study was to extend this analysis to see what kinds of prosodic strategies, including intonational phrasing and prominence-lending tonal movements, are associated with the realisation of different focus constructions in Nafsan.

2. METHOD

2.1. Participants

The participants in this study were five male speakers, aged 18-48, from Erakor village in Vanuatu. All participants identify Nafsan as their first language. In addition, all speak Bislama, a lingua franca in Vanuatu, and have different levels of knowledge of English and French.

2.2. Materials and procedures

A series of tasks were designed to explore prosodic realisation of informational and contrastive focus on subject or object nouns in mini-dialogues where word-order was manipulated. Two of these tasks formed the corpus for this paper. 30 target nouns ranging from one to four syllables and comprising various phonotactic structures were embedded in three different frames, forming the dialogue format for each task: an opening declarative statement to be produced by Participant A, for which INFORMATIONAL FOCUS was anticipated for the target noun (A1), a negated statement to be produced by Participant B, for which the target noun was expected to be NON-FOCAL (B1), and a final declarative statement, also produced by Participant B, for which CONTRASTIVE focus was anticipated for the (different) target noun (B2). In a), each noun functions as a SUBJECT and occurs in INITIAL position in both the A1 and B2 utterances, and in b), each noun functions as an OBJECT and occurs in FINAL position in the A1 and B2 utterances.

a) A1: SUBJ. INIT. ***menaal*** *imur nafnag*
barracuda wants food
B1: SUBJ. MED. *itiik,* ***menaal*** *ita mur nafnag mau*
no, **barracuda** doesn't want food
B2: SUBJ. INIT. ***rakum*** *imur nafnag*
crab wants food
b) A1: OBJ. FIN. *natam̃ool ipuetsok* ***menaal***
the person is holding **barracuda**
B1: OBJ. MED. *itiik, natam̃ool ita puetsok* ***menaal*** *mau*
no, the person isn't holding **barracuda**
B2: OBJ. FIN. *natam̃ool ipuetsok* ***rakum***
the person is holding **crab**

Participants were recorded in pairs in a sheltered area during fieldwork in Erakor, with stimuli presented in Nafsan orthography on slides, and arranged so that each noun would be produced by each speaker in each focus context (for each task). 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. The final dataset for these analyses contained 1012 utterances: 5 speakers x 30 nouns x 3 focus contexts x 2 tasks with different utterance conditions, plus occasional repetitions.

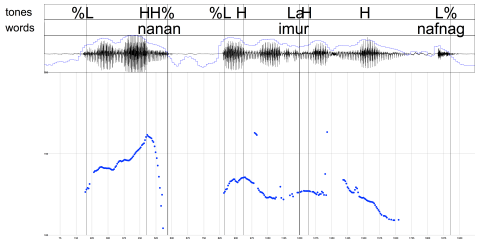
2.3. Data processing and analysis

Utterances were orthographically transcribed in Praat [2] and converted to SAMPA. Automatic segmentation of the speech signal was performed using the language-independent model of the Munich Automatic Segmentation System [6]. Segment boundaries were checked and corrected where necessary with reference to wideband spectrograms and corresponding waveforms. A hierarchical database was constructed using the EMU Speech Database Management System [16], including tiers for the utterance, prosodic words, syllables, phonemic segments, and intonational tone targets. If a syllable carried a major pitch movement at the right edge of the word, it was labelled S (strong). All other syllables in the token were labelled W (weak).

All target nouns were annotated using a preliminary Autosegmental-Metrical transcription system where peaks and elbows in the F0 contour are labelled as H (high) or L(low) [8]. In the SUBJECT INITIAL condition, low, rising, or high pitch movements across tokens were annotated as either %L, %LH or %HH, with the % symbol indicating an intonational phrase-initial L or H tone. In the OBJECT FINAL condition, the word initial tones were annotated as either %L, L, H, or %H and word final tones were annotated as HL% or HH% depending on whether the phrase-final tune was falling or rising. If there was no word-final pitch movement, only the word-initial tone was annotated.

An example of an intonationally-annotated utterance produced in the SUBJECT INITIAL condition and INFORMATIONAL FOCUS context is shown in Fig. 1. Tonal patterns were extracted for each target noun and F0 Hz values were extracted for tone targets using the *emuR* package in R [16, 11], and then converted to semitones using a formula with a base frequency of 50 Hz after [10]. Syllable duration was also extracted for S and W syllables (based on previous analyses of Nafsan that suggest final syllables are prosodically prominent [1]) and z-normalised to minimise inter-speaker variation. All extracted pitch values associated with H tone targets were included in a maximally-specified mixed effects model with the fixed factor FOCUS (A1: INFORMATIONAL FOCUS; B1: NON-FOCAL; B2: CONTRASTIVE FOCUS), with random slopes and intercepts for ITEM, SPEAKER using *lmertest* and *step* in R [7]. Syllable duration was also tested with the additional fixed factor PROMINENCE. A logistic regression analysis was also performed to see whether there was a significant effect of FOCUS on tonal pattern distribution associated with the target nouns across the utterance conditions: SUBJECT INITIAL, and OBJECT FINAL.

Figure 1: Example of a tonally annotated SUBJECT INITIAL utterance in the CONTRASTIVE focus context for the sentence ***nanan*** *imur nafnag* "the **goat** wants food".



3. RESULTS

3.1. Tonal distribution

Table 1 summarises the distribution of tonal patterns across INFORMATIONAL FOCUS, NON-FOCAL, and CONTRASTIVE subjects and objects in the two utterance conditions – SUBJECT INITIAL, OBJECT FINAL - examined in this paper. In the SUBJECT INITIAL condition, there was a significant effect of focus condition on the distribution of tone patterns (χ^2=35 p<0.0001). CONTRASTIVE and INFORMATIONAL FOCUS tokens were realised primarily with a left edge low tone (%L) and final high tone (H) regardless of word-length. In other words, the rising movement was delayed across the full word in cases of longer tokens with the H realised on the final syllable. However, the LH movement was fully realised across monosyllabic tokens. A number of tokens were also realised with a %HH pattern. In most cases, the right-edge tone was also phrase-final, with clear pitch range re-set on following material. NON-FOCAL words were occasionally realised with the same kind of rising tonal movement but in many cases there was no final H tone at the right edge and no evidence of an intonational boundary.

In the OBJECT FINAL condition, there was also a significant effect of FOCUS on distribution of tonal patterns (χ^2=12.8 p<0.01). Many INFORMATIONAL FOCUS and CONTRASTIVE focused tokens were realised with either a %LHL%, LHL% or HL% pattern, with the initial L tone realised on the initial syllable and final falling tune (HL%) realised on the final syllable of longer words. Monosyllabic tokens were usually realised with a HL% pattern, with an optional left edge L tone. A number of object tokens were also realised in intonational phrase-initial position due to the presence of a pause and a full intonational phrase boundary before the token. However, NON-FOCAL words were often realised with a single left edge H or L tone with the utterance-final negative particle *mau* attracting the intonational phrase-final HL% pitch movement. In other words, there was no evidence of a final H tone at the right edge of NON-FOCAL tokens, the site of putative accentual prominence in Nafsan.

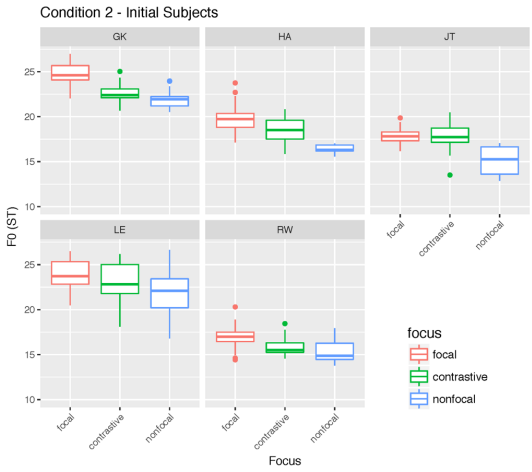
Table 1. Distribution of tonal patterns in 3 focus contexts a 2 utterance conditions: SUBJECT INITIAL, OBJECT FINAL.

	Context	Tonal Pattern	Number
SUBJECT INITIAL	INF. FOCUS	%HH	20
		%LH	148
	NON-FOCAL	%LH	85
		%L	48
		%H	25
	CONTRASTIVE	%HH	22
		%LH	137
OBJECT FINAL	INF. FOCUS	(%L)HH%	22
		(%L)HL%	120
	NON-FOCAL	L	82
		LH	24
		H	9
	CONTRASTIVE	(%L)HL%	91
		(%L)HH%	36

3.2. F0 scaling & duration: initial subjects/final objects

Fig. 2 plots the pitch level of H tones (semitones) associated with the final syllable of each token in SUBJECT INITIAL position for the three focal contexts, INFORMATIONAL and CONTRASTIVE FOCUS, and NON-FOCAL, for the five speakers. For SUBJECT INITIAL tokens, there was a significant effect of FOCUS on pitch level, with CONTRASTIVE and INFORMATIONAL FOCUS H tones scaled consistently higher than H tones associated with the right or left edge of NON-FOCAL tokens (F=17.208, p<0.001). The pitch span of the LH movement in INFORMATIONAL FOCUS and CONTRASTIVE focus tokens was also somewhat greater in magnitude compared to NON-FOCAL tokens.

Figure 2. Tone scaling – SUBJECT INITIAL tokens



For three of the five speakers (GK, RW, and HA), H tones were actually scaled slightly higher in INFORMATIONAL FOCUS compared to CONTRASTIVE focused tokens, suggesting that the first mention of a subject noun in the dialogue attracted a higher tone target. For the remaining speakers, post-hoc tests revealed minimal differences between

INFORMATIONAL FOCUS and CONTRASTIVE focused tokens, but for all speakers, there was a significant difference between INFORMATIONAL FOCUS and NON-FOCAL H tones (t=7.441 p<0.0001) and CONTRASTIVE and NON-FOCAL H tones (t=4.702 p <0.0001). However, as shown in Fig. 1, there was clearly also a high level of pitch compression of post-focal material in the CONTRASTIVE focus condition which was also observed in INFORMATIONAL FOCUS contexts.

In the OBJECT FINAL condition shown in Fig. 3, H tone targets associated with the INFORMATIONAL FOCUS context were scaled higher than NON-FOCAL tokens (t=8.57, p<0.0001) and CONTRASTIVE focus tokens (t=11.054, p<0.0001). Tokens in the CONTRASTIVE context were actually scaled *lower* than NON-FOCAL tokens for one speaker (HA), and there was no significant difference between CONTRASTIVE and NON-FOCAL contexts for a second speaker (RW). However, many tokens in the CONTRASTIVE context were realised in separate intonational constituents from preceding material, which was not the case for NON-FOCAL tokens.

Figure 3. Tone scaling – OBJECT FINAL tokens

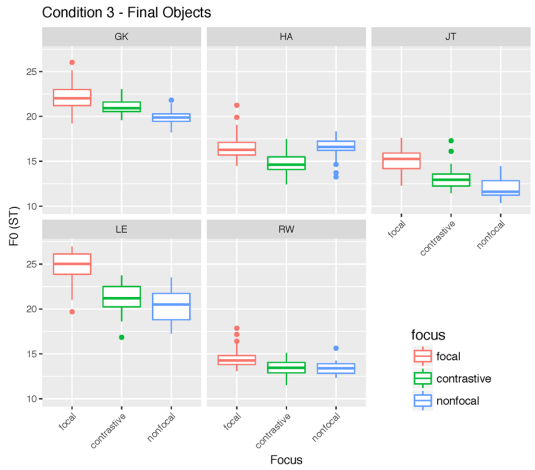
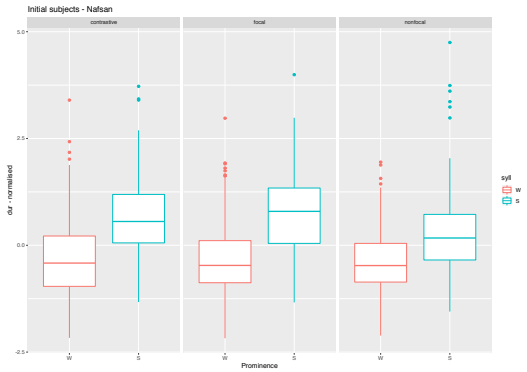


Figure 4. Duration patterns – SUBJECT INITIAL tokens



There were significant main effects of FOCUS (F=5.18, p<0.05) and PROMINENCE (S or W) (F=71.89, p<0.0001) on syllable duration in SUBJECT INITIAL and OBJECT FINAL tokens (see Fig. 4 for SUBJECT INITIAL tokens). Word-final (S) syllables that were associated with H tone targets in SUBJECT INITIAL INFORMATIONAL FOCUS and CONTRASTIVE focus contexts tended to be significantly longer than non-final (W) syllables in the same word (t= 8.86, p<0.0001;t=7.66, p<0.0001). They were also longer compared to final syllables in NON-FOCAL tokens (t=5.84, p<0.0001). In the OBJECT FINAL context, final (S) syllables were significantly longer in CONTRASTIVE tokens compared to the final (S) syllable in NON-FOCAL tokens (t=5.38 p<0.001). Non-final (W) syllables were shorter than final syllables in INFORMATIONAL FOCUS and CONTRASTIVE focus contexts (t=17.94, p<0.0001, t=18.29), p<0.0001, with smaller lengthening effects observed in NON-FOCAL words (t=8.49; p<0.0001). Final syllables in focussed words were also likely to be subject to pre-boundary lengthening given that these tokens were in utterance final position by contrast with NON-FOCAL tokens.

4. DISCUSSION

In this study, the tonal and duration patterns of SUBJECT INITIAL and OBJECT FINAL nouns in different focus contexts were examined in Nafsan. CONTRASTIVE and INFORMATIONAL FOCUS tokens are realised with significantly higher F0 targets than NON-FOCAL tokens, although there is a high level of gradience. The consistent alignment of the right edge H tone with the final syllable of words in focus confirms findings from a previous study of accentual prominence [1] in Nafsan, where a higher F0 value was observed in word-final syllables. In NON-FOCAL contexts, a number of SUBJECT INITIAL tokens did not have a right edge tone, suggesting a prosodic pattern akin to prosodic de-phrasing with the final H tone realised on the following verb.

In the OBJECT FINAL condition, INFORMATIONAL FOCUS tokens were realised with a strong HL falling tonal pattern on the final syllable with the H tone scaled consistently higher than tokens in CONTRASTIVE FOCUS tokens. In NON-FOCAL objects, there was a clear loss of a final H tone also suggesting a pattern of de-phrasing in this context, with the final negative particle attracting the phrase-final prominence. Durational patterns of subject and object tokens also confirmed consistent lengthening of final versus non-final syllables in words that were in INFORMATIONAL and CONTRASTIVE FOCUS with smaller effects evident in NON-FOCAL contexts. Results for SUBJECT INITIAL tokens, at least, suggest that phrasing as well as pitch span are important strategies used in focus marking in Nafsan. These results also confirm our earlier impressions that Nafsan has a phrasal edge-marking prosodic typology [5], and that accentual prominence may be phrasal, as for Vanuatu languages such as Dakaaka [15] and a range of Austronesian languages [e.g. 4].

5. ACKNOWLEDGEMENTS

Sincere thanks to all the Nafsan speakers who have participated in and facilitated this and earlier work, and thanks to Ben Volchok for assisting with data processing. This research was conducted with support from the ARC Centre of Excellence for the Dynamics of Language (Project ID: CE140100041).

6. REFERENCES

[1] Billington, R., Fletcher, J., Thieberger, N., Volchok, B. 2018. Acoustic correlates of prominence in Nafsan. In J. Epps, J. Wolfe, J. Smith, & C. Jones (eds.). *Proceedings of the 17th Australasian International Speech Science and Technology Conference*. Sydney: Australasian Speech Science and Technology Association, 137–140.

[2] Boersma, P., Weenink, D. 2018. Praat (Version 6.0.40). Retrieved from http://www.praat.org/

[3] Calhoun, S. 2015. The interaction of prosody and syntax in Samoan focus marking. *Lingua* 165, 205-229.

[4] Himmelmann, N., Kaufman, D. (in press). Prosodic Systems: Austronesian. In Gussenhoven, C., Chen, A. (eds). *The Oxford Handbook of Prosody.* Oxford: Oxford University Press.

[5] Jun, S-A. 2014. Prosodic typology by prominence type, word prosody, and macro-rhythm. In S-A Jun (ed.). *Prosodic Typology II*. Oxford: Oxford University Press, 520-539.

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

[7] Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13), 1–26.

[8] Ladd, D.R. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press

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

[10] Nolan, F. 2003. Intonational equivalence: An experimental evaluation of pitch scales. In M. J. Solé, D. Recasens, and J. Romero (eds.). *Proceedings of the 15th International Congress of Phonetic Sciences.* Barcelona: Causal Productions, 771-774.

[11] R Core Team. 2018. R: A language and environment for statistical computing (Version 3.4.2). Retrieved from http://www.R-project.org

[12] Schnell, S. 2018. Whence subject-verb agreement? Investigating the role of topicality, accessibility, and frequency in Vera'a texts. *Linguistics*, 735-780.

[13] Steindel-Burdin, R, Phillips-Bourass, S., Turnbull, R., Yasavul, M., Clopper, C. G., Tonhauser, J. 2015. Variation in the prosody of focus in head- and head/edge-prominence languages. *Lingua* 165, 254-275.

[14] Thieberger, N. 2006. *A Grammar of South Efate: An Oceanic Language of Vanuatu*. Honolulu: University of Hawaii Press.

[15] von Prince, K. 2015. *A grammar of Daakaka*. Berlin: De Gruyter Mouton.

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

[17] Zimmerman, M., Onea, E. 2011. Focus marking and focus interpretation. *Lingua* 121, 1651-1670.