Focus prosody in Fijian: A pilot study

Albert Lee[1], Candide Simard[2], Apolonia Tamata[2], Jiaying Sun[1], Tsz Ching Mut[1]

[1] The Education University of Hong Kong, [2]University of the South Pacific
albertlee@eduhk.hk, candide.simard@usp.ac.fj, apolonia.tamata@usp.ac.fj, jsun@eduhk.hk, tcmut@eduhk.hk

ABSTRACT

This paper presents an acoustic analysis of focus prosody in Fijian – a verb-first language. Through analyzing pitch (f_0), intensity, and duration data, we compared (i) VP (i.e. initial) vs. neutral focus, (ii) subject vs. corrective-subject focus, and (iii) object vs. corrective-object focus. We found significant on-focus mean f_0 raising for VP focus, but no post-focus compression. Significant on-focus *shortening* of syllable duration was also observed. Compared with subject focus, corrective-subject is marked by significant on-focus *increase* in intensity and *shortening* of syllable duration. No significant difference between object and corrective-object focus was observed in the on-focus domain. Our findings pave the way for future systematic investigations of Fijian prosody, a currently understudied area in phonetics.

Keywords: Fijian, focus

1. INTRODUCTION

Studying prosodic focus often requires comparing the acoustic realisation of the same utterance in different focus conditions, elicited in a controlled manner (see overview in [1]). However, singling out prosodic cues can be less straightforward for languages where focus marking involves compulsory syntactic means (e.g. fronting in Samoan [2]), such as verb-initial languages (see review in [3]). One such language is Fijian, which to date has not been systematically studied in a controlled production experiment.

Fijian is an Austronesian language spoken by about 400,000 as a first language [4] in Fiji. It has positional stress. The basic word order of Fijian is verb-object-subject (but note alternative accounts such as [4]). Currently, there is no known experimental study reporting on the acoustic correlates of Fijian focus.

For a (reasonably) nearby verb-initial language Samoan, it was found that the initial phonological phrase was always the most prominent [2]. Thus when the focus was fronted, it was maximally prominent. In VSO sentences, the verb and agent were in the initial phrase. Speakers raised the pitch on the object in object focus, and lowered it in agent focus; although they did not do this consistently. There was no prosodic marking of focus on the agent.

As a first study of Fijian focus prosody, here we focused on identifying prosodic cues to focus and refrained from analysing conditions that would necessitate using syntactic focus markers like fronting. Therefore, in this paper we compared the following focus conditions: (i) VP (i.e. initial) vs. neutral focus, (ii) subject vs. corrective-subject focus, and (iii) object vs. corrective-object focus. As VP is sentence initial, analyzing VP vs. neutral focus allows us to see in situ prosodic marking of narrow focus in Fijian. In addition, we are also interested in whether in a corrective context, the same narrow focus (i.e. subject and object focus) would be marked differently.

2. METHODS

2.1. Participants

The study included ten native Fijian speakers (5 female and 5 male). They were born and raised in Fiji and spoke fluent Fijian. They were university students at the time of testing. No one reported any speech or hearing impairments.

2.2. Materials

Five different declarative sentences were designed (Table 1) to elicit prosodic focus. S1 and S2 are typical verb-object-subject (VOS) sentences in Fijian and S3-S5 are verb-subject (VS) sentences. S4 is in the passive voice. For S1 and S2, each base sentence has six possible focus conditions, namely, neutral focus, VP (verb phrase) focus, object focus, subject focus, corrective-subject focus (i.e., the subject is corrected by the speaker and under focus), and corrective-object focus. S3, S4, and S5 (VS sentences) do not have (corrective) object focus. When defining the on-focus domain [5], only content words were included (except for VP focus condition). All focus conditions were elicited by a precursor question asked by the interviewer (Table 2).

S1	Era vaqitora tiko na lali na gone. 3pl play-prep ASP ART drum ART child 'The children are playing with the drum.'
S2	E volia tiko na dalo na yalewa. 3sg buy-TR ASP ART taro ART woman 'The woman is buying taro.'

S3	Eratau qaqalo tiko na gone. 3pl swim ASP ART child 'The kids are swimming.'
S4	E waluvu tiko na vale. 3sg flood ASP ART house 'The house is being flooded.'
S5	E sa vuka vakababa na bebe. 3sg ASP fly sideways ART butterfly 'The butterfly is flying.'

Table 1: Base sentences used in this study.

Focus	**Precursor questions**	**Target sentences**
Neutral	Na cava e yaco tiko? ART what 3sg happen ASP 'What is happening?'	Era vaqitora tiko na lali na gone. 3pl play-prep ASP ART drum ART child 'The children are playing with the drum.'
VP	Na cava era cakava tiko na gone? ART what 3pl do ASP ART children 'What are the children doing?'	Era vaqitora tiko na lali na gone. 3pl play-prep ASP ART drum ART child 'The children are playing with the drum.'
Subject	O cei e vaqitora tiko na lali? ART who 3sg play-prep ASP ART drum 'Who is playing with the drum?'	O ira na gone, Era vaqitora tiko na lali. ART 3pl ART child 3pl play-prep ASP ART drum 'The children are playing with the drum.'
Object	Na cava era vaqitora tiko na gone? ART what 3pl play-prep ASP ART children 'What are the children playing with? '	Era vaqitora tiko na lali na gone. 3pl play-prep ASP ART drum ART child 'The children are playing with the drum.'
Corrective Subject	Eta vaqitora tiko beka na dramu? 1pl play-prep ASP perhaps ART drum 'Are we playing with the drum?'	Sega, o ira na gone era vaqitora tiko na lali. Not ART 3pl ART child 3pl play-prep ASP ART drum 'No, the children are playing with the drum.'
Corrective Object	Era qito rakavi tiko? ART play-TR rugby ASP 'Are the children playing rugby?'	Sega, era vaqitora tiko na lali na gone. Not 3pl play-prep ASP ART drum ART child 'No, the children are playing with the drum.'

Table 2: Examples of precursor questions and target sentences (the focused elements are underlined).

Speakers were asked to say each sentence three times. Altogether, we recorded 720 utterances ((2 sentences * 6 focus condition + 3 sentences * 4 focus condition) * 3 repetitions * 10 speakers). However, 168 (23.3%) utterances had to be discarded due to missing or redundant syllables, or unexpected word order. Consequently, a total of 552 sentences were retained for analysis. The seemingly high discard rate is considered to be due to the unscripted nature of the task (see §2.3).

2.3. Recording procedure

Recording took place in a room at the University of the South Pacific. We produced 21 cards, each with an image and a sentence describing it. The interviewer showed the participant the card and asked them to answer questions using the card's information. The presentation order of the cards was random. For the corrective focus condition, participants were required to correct the inaccurate information provided by the interviewer and give the complete and correct sentences. Before recording commenced, there were two tests in which the speakers were asked to describe the images, so that we could ensure that they understood the experimental procedure. The total recording time for each participant was about one hour.

2.4. Annotation and measurements

The raw sound data were first chunked into individual utterances, and then labelled by syllable in PRAAT [6]. As responses were unscripted, many deviated from the target list (see §2.2). Thus when labelling the recordings, we retained extra syllables that were semantically consistent with the expected answer. In cases where finding the syllable boundary was impossible (e.g. due to fast speech), it was placed in the dead centre of the two syllables concerned. Pauses, repeated syllables, and mispronounced syllables were not labelled.

After annotation, we checked and rectified vocal pulse markings manually for accurate f_0 tracking and obtained the time-normalised f_0, intensity, and duration with ProsodyPro [7].

3. RESULTS

3.1 F0

We fitted linear mixed effects models to pre-focus, on-focus, and post-focus mean f_0 (Hz) data using *lmerTest()* [8]. Model construction followed a bottom-up approach. Post-hoc comparisons were done using *emmeans()* [9]. For on-focus mean f_0, the best fitting model contained the fixed factor of focus condition (VP, neutral, subjective, Corrective-subject, object, Corrective-object), as well as by-subject and by-sentence random intercepts. Including random slopes prevented model convergence.

Intensity (§3.2) and duration (§3.3) data were analysed using the same approach.

The main effect of focus on on-focus f_0 mean was significant, $X^2(6) = 214.46$, $p < .001$. Post-hoc test shows that **VP focus** had significantly higher on-focus mean f_0 than corresponding neutral focus ($p < .0001$). No other significant contrast was found. However, SS ANOVA [10] plots for individual sentences (e.g. Figure 1) reveal that this difference was non-significant for 3 of 5 sentences.

The main effect of focus on post-focus mean f_0 was significant too, $X^2(6) = 159.71$, $p < .001$. The Post-hoc test shows that **Corrective-Subject focus** had significantly lower post-focus mean f_0 than corresponding **Subject** focus ($p = .0001$). However, inspection of SS ANOVA plots (e.g. Figure 2) revealed that this difference was significant for 2 of 5 sentences.

Finally, the main effect of focus on pre-focus mean f_0 was significant, $X^2(6) = 221.25$, $p < .001$. The Post-hoc test shows that **Corrective-Object focus** had significantly lower pre-focus mean f_0 than corresponding **Object** focus ($p < .0001$). SS ANOVA plots (e.g. Figure 3) show that this difference was observed in all 5 sentences.

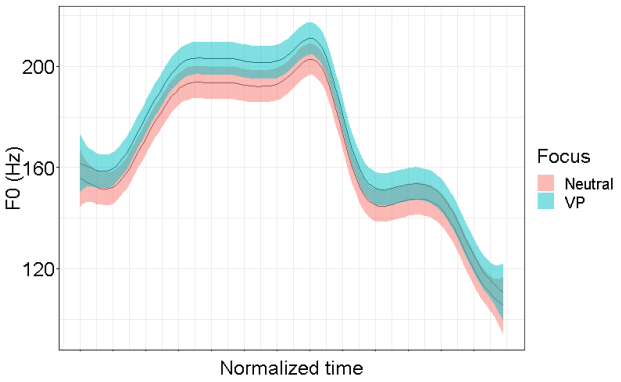

Figure 1: SS ANOVA comparing VP and neutral focus of S2 (see Table 1).

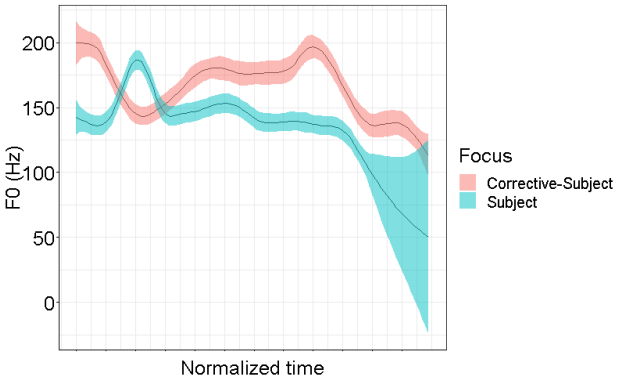

Figure 2: SS ANOVA comparing subject and Corrective-Subjective focus of S5.

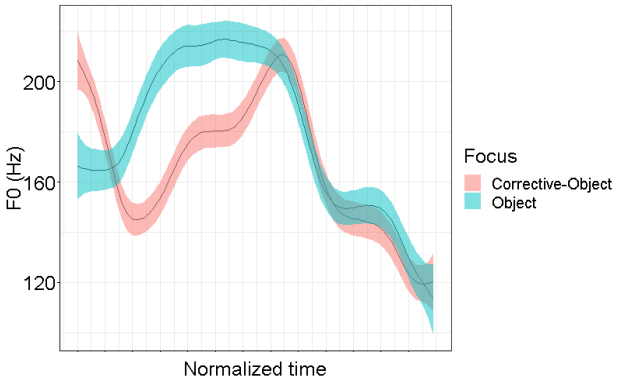

Figure 3: SS ANOVA comparing Object and Corrective-Object focus of S2.

3.2 Intensity

There was a main effect of focus condition on pre-focus ($X^2(6) = 70.24$, $p < .001$), on-focus ($X^2(6) = 115.19$, $p < .001$), and post-focus mean intensity ($X^2(6) = 41.34$, $p < .001$). Compared with neutral focus, **VP focus** had significantly lower on-focus intensity ($p = .0016$). Compared with **Subject focus**, **Corrective-Subject focus** had significantly higher on-focus intensity ($p < .0001$, see Figure 4), and lower pre-focus ($p < .0001$, see Figure 5) and post-focus intensity ($p = .0203$, see Figure 6). Compared with **Object focus**, **Corrective-Object focus** had significantly lower pre-focus ($p = .0241$) and post-focus intensity ($p = .0020$).

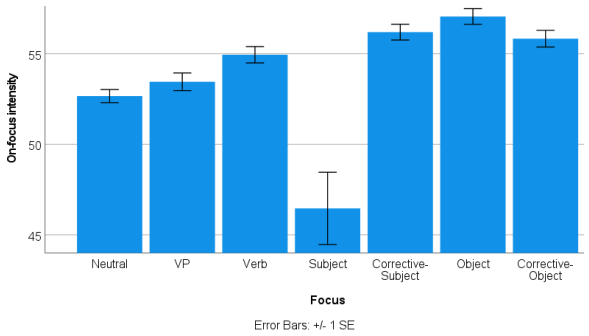

Figure 4: Barplot comparing on-focus mean intensity (dB) in different focus conditions.

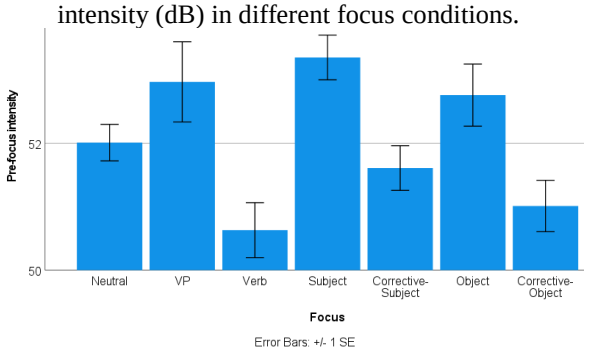

Figure 5: Barplot comparing pre-focus mean intensity (dB) in different focus conditions.

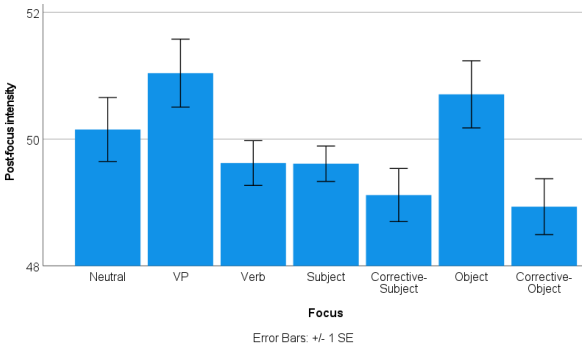

Figure 6: Barplot comparing post-focus mean intensity (dB) in different focus conditions.

3.3 Duration

There was a main effect of focus condition on pre-focus ($X^2(6) = 538.37$, $p < .001$), on-focus ($X^2(6) = 152.29$, $p < .001$), and post-focus mean syllable duration ($X^2(6) = 83.17$, $p < .001$). Compared with neutral focus, **VP focus** had significantly shorter on-focus ($p = .0001$) and pre-focus ($p < .0001$) duration. Compared with **Subject focus**, **Corrective-Subject focus** had significantly shorter on-focus ($p < .0001$, see Figure 7) but longer pre-focus ($p < .0001$, see Figure 8) and post-focus ($p = .0432$) duration. Compared with **Object focus**, **Corrective-Object focus** had significantly longer pre-focus ($p < .0001$) but shorter post-focus duration ($p = .0020$).

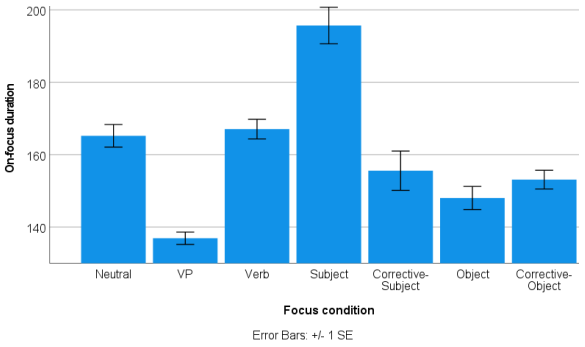

Figure 7: Barplot comparing on-focus mean duration (ms) in different focus conditions.

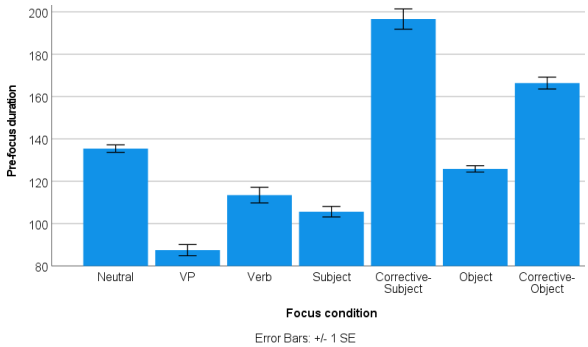

Figure 8: Barplot comparing pre-focus mean duration (ms) in different focus conditions.

4. DISCUSSION

This study set out to identify prosodic focus marking strategies in Fijian. We compared (i) VP (i.e. initial) vs. neutral focus, (ii) subject vs. corrective-subject focus, and (iii) object vs. corrective-object focus. We found significant on-focus mean f_0 raising for VP focus, but no post-focus compression. There was also considerable cross-sentence variability. Significant on-focus *shortening* of syllable duration was also observed. Compared with subject focus, corrective-subject is marked by significant on-focus *increase* in intensity and *shortening* of syllable duration. No significant difference between object and corrective-object focus was observed in the on-focus domain.

Through inferential statistics we have found that focus condition significantly affected prosodic focus markers in different ways. However, note also that there was considerable cross-speaker variability, in part due to the unscripted nature of participants' responses. Cross-sentence variability was also observed, showing that even highly significant effects may not necessarily be applicable to all sentences of the same focus condition.

As the first systematic production study of Fijian focus prosody, we have only been able to cover initial focus (i.e. VP focus), leaving out other positions. To bypass compulsory fronting, one could design production tasks such as nonce sentences or phone digits to elicit in situ prosodic marking strategies (see [11]). In addition, given Fijian focus is marked by multiple prosodic cues alongside syntactic strategies, one could use statistical tests such as linear discriminant analysis [12] to identify their respective predictive power of focus. Much more work is needed to get a better understanding of focus prosody in Fijian, a currently understudied Austronesian language.

6. REFERENCES

[1] Y. Xu, "Speech prosody: A methodological review," *J. Speech Sci.*, vol. 1, no. 1, pp. 85–115, 2011.

[2] S. Calhoun, "The interaction of prosody and syntax in Samoan focus marking," *Lingua*, vol. 165, pp. 205–229, 2015.

[3] F. Kügler and S. Calhoun, "Prosodic encoding of information structure: A typological perspective," in *The Oxford Handbook of Language Prosody*, C. H. M. Gussenhoven and A. Chen, Eds. Oxford: Oxford University Press, 2020, pp. 454–467.

[4] P. A. Geraghty, "Fijian," in *Concise encyclopedia of languages of the world*, K. Brown and S. Ogilvie, Eds. Oxford, 2009, pp. 412–413.

[5] Y. Xu, C. X. Xu, and X. Sun, "On the temporal domain of focus," in *Proceedings of the 2nd International Conference on Speech Prosody (SP2004)*, 2004, pp. 81–84.

[6] P. P. G. Boersma and D. J. M. Weenink, "Praat: Doing phonetics by computer." 2012.

[7] Y. Xu, "ProsodyPro: A tool for large-scale systematic prosody analysis," in *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, 2013, pp. 7–10.

[8] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *J. Stat. Softw.*, vol. 82, no. 13, pp. 1–26, 2017.

[9] R. V. Lenth, "Emmeans: Estimated marginal means, aka least-squares means." 2020, [Online]. Available: https://cran.r-project.org/package=emmeans.

[10] L. S. Davidson, "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance," *J. Acoust. Soc. Am.*, vol. 120, no. 1, pp. 407–415, 2006.

[11] T. C. Mut, C. Simard, Apolonia Tamata, and A. Lee, "Focus prosody in Fijian: In-situ focus marking,", *Proc of ICPhS 2023*, 2023.

[12] M. E. Breen, E. Fedorenko, M. Wagner, and E. A. F. Gibson, "Acoustic correlates of information structure," *Lang. Cogn. Process.*, vol. 25, no. 7–9, pp. 1044–1098, 2010.

6. ACKNOWLEDGEMENTS

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China awarded to AL (Project No. 18600621). We thank Ms. Jueyu Lu for helping with data annotation.