# A corpus study of word (root) prominence in Vera'a

*Catalina Torres, Stefan Schnell*

University of Zurich
ARC Centre of Excellence for the Dynamics of Language
catalina.torres@ivs.uzh.ch, stefan.schnell@uzh.ch

## Abstract

This study presents an acoustic investigation of word level prosody in the Oceanic language Vera'a. The analysis is based on a corpus of speech data collected during fieldwork from multiple speakers. A previous description of Vera'a suggests the language has lexical stress but its acoustic realisation was not further investigated. This study provides the first instrumental examination of five acoustic measures and their relation to prominence marking. The evidence indicates that vowels in the last syllable of the root are more prominent. However, these observations are restricted to some acoustic parameters.

**Index Terms**: root, word prominence, acoustic correlates, Oceanic languages, prosodic typology

## 1. Introduction

Despite a growing body of research dealing with the phonetics and phonology of word prosodic systems in the Austronesian language family [1, 2, 3, 4, 5, 6, 7, 8], instrumental quantitative studies in this field remain under-represented. Word prosodic systems in the Austronesian language family are of particular interest since previously reported lexical stress patterns have been challenged in the literature [9] showing a tendency for phrasal edge-marking prosodic typology [10]. As an example, in Indonesian, a language often spoken by multilingual speakers, studies have shown that lexical stress might not be present. Instead the evidence suggests a pattern of phrasal prominence. Additionally, these studies show that perception and production of prominence may vary depending on other languages spoken by individuals [11, 12].

As Gordon and Roettger [13] point out, one particular issue that studies examining word prosodic systems struggle with is a lack of control for prosodic boundary phenomena. The authors suggest to control for target words not to be placed at phrasal boundaries and to include a sufficiently large sample of lexical items (although they do not specify number of items or speakers). Efforts to implement state of the art phonetic analysis on under-documented languages from the Pacific have yielded important results. Recent instrumental work on Austronesian languages from the Oceanic subgroup indicate that impressionistic reports from the literature do not bear out when tested experimentally. These studies include data from multiple speakers and tested speech produced in different contexts while controlling for phrasal boundary effects. Contra previous claims, acoustic phonetic studies on Nafsan and Drehu show that these are edge-marking languages and prominence is realised post-lexically [5, 14, 8]. Interestingly, Nafsan and Drehu are languages that do not display weight sensitivity in their prominence marking systems. Although in both languages there is a phonemic vowel length distinction [15, 16, 17], prominence lending pitch peaks align with boundaries of phrases [5] or prosodic constituents such as accentual phrases [8]. We here report findings from a corpus study of Vera'a, another Oceanic language that is closely related to Nafsan and Drehu within the subgroup of Eastern Oceanic languages and is spoken in North Vanuatu.

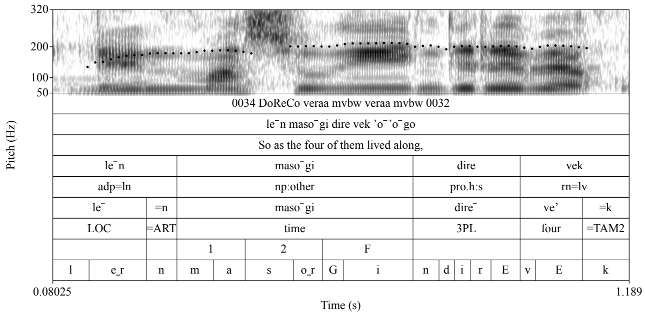


Figure 1: *Waveform, F0 trace, and spectrogram of an utterance from the corpus. The root in the utterance is masogi 'time'.*

### 1.1. Typological features of Vera'a

Vera'a is a canonical Oceanic language in the sense of [18]: Phrase structure is head-initial and word order on clause level is SVO. Adpositional phrases are all prepositional phrases, and on NP level heads precede any modifiers. The only exception is that articles precede NP heads, and so does a small set of quantifiers. The language is mostly analytic, the only affixes being possessive suffixes that attach to nominal roots. The phonology of Vera'a is similar to that of nearby languages: it has 17 consonant phonemes – including a double-articulated labio-velar plosive and a double-articulated labio-velar nasals – and seven vowel phonemes (/ɐ, e, ɛ, i, o, ɔ, u/). Vera'a differs from a range of languages in its neighbourhood though in that vowel length is not phonemic. Syllable structure is (C)V(C), occurring word internally and finally. Complex margins of the type CCV or VCC are only marginally present. In the sketch grammar of the language [19], word stress was impressionistically determined to fall on the final or penultimate syllable of a word (a word equals a bare root in our study). Alternatively, when the last syllable of a word is heavy (e.g., CVC) it is said to bear stress. There is no indication that stress could be distinctive in Vera'a and it is not known that words are distinguished based on stress.

### 1.2. Methodological considerations

In this exploratory study we focus on selected acoustic parameters that have been associated with stress and prominence marking cross-linguistically [20]. Taking into account recent evidence from other Oceanic languages [5, 14, 8] this study seeks to remain agnostic as to whether there is lexical stress in the language, especially considering that there is no evidence of minimal pairs based on this criterion. Instead we set out to examine whether there is acoustic evidence for a prominence pattern at the word/root level. In doing so we leave the question open as to whether Vera'a displays a lexical stress-accent to which post-lexical phrase accents dock on, as typically known from languages such as English [13, 21]. As previous studies have shown, some languages can also display misalignment between stress and post-lexical phrase level prosody meaning that a demarcating and highlighting function need not be aligned [6, 22]. A more detailed explanation for this methodological choice can be found in the discussion. Similar to [4], we draw on narrative discourse data which has the advantage of representing speech production in connected discourse rather than isolated utterances. Our data, however, stems from a larger language documentation corpus, which bears the additional advantage of these data to be representative of the common speech production in everyday language use.

## 2. Research aims

Considering recent findings on word and phrase level prominence patterns in Austronesian and Oceanic languages, it is of interest to examine Vera'a acoustically and determine whether there is phonetic evidence to support the previous description [19]. In particular, it is of interest to evaluate whether the penultimate or final syllable of roots, all other things being equal, represent the most prominent syllable. Additionally, it is our aim to examine whether there is evidence for heavy root-final syllables (CVC) to be acoustically more salient than heavy syllables in other positions. We focus on nominal roots in this paper for practical reasons.

## 3. Materials and Method

### 3.1. Corpus

Our corpus consists of 7 narrative texts, each produced by a single speaker primarily for the purpose of being recorded and added to the ongoing documentation of the Vera'a language (Schnell et al ongoing) as part of the local 'orature'. Speakers came from both genders (2f, 5m) with age ranging between 16 and approximately 65 at time of recording in early 2007.[1] These narratives constitute the Vera'a subcorpus of DoReCo [23, 24] on whose time-aligned segmentation the current study is based, and it is also a major part of the Vera'a corpus within Multi-CAST [25, 26] which features relevant morphosyntactic annotations that enable us to identify nominal roots which are the focus of this study. A detailed description of forced alignment and segmentation process such as location of root and vowel boundaries can be found in [27]. All narratives are in turn part of the larger Vera'a documentation corpus archived with The Language Archive (MPI Nijmegen) (`https://hdl.handle.net/1839/bc035bf8-1d9b-4163-8131-983d5a7b08ab`).

Recordings used in the current study were made as part of the general language documentation project of Vera'a. All recordings were made in bamboo-walled houses that keeps interference from wind and other background noises (community activity, animals, etc.) as minimal as possible in the given conditions. Speakers offered to be recorded as contribution to the documentation project.

[1]Exact age is often unknown by elderly speakers of the language.

### 3.2. Data curation

For the purpose of this study a selection of nominal word roots was carried out, similar to [4]. As NPs in Vera'a can feature both pre- and post-nominal satellites and nominal roots can be expanded by suffixes, the decision was made to base the analysis on forms that represent NP heads and are simple roots. First, the narratives in the corpus were segmented into phrases the size of short utterances. This initial corpus amounts to 1484 utterances. NPs that were not produced directly at the start or the offset of the utterance and that did not co-occur with pauses were identified. For each root an additional segmentation of syllables was included, following [19]. Additionally, the number of syllables per root and the syllable structure of all syllables in the corpus was marked. A hierarchical database was constructed using the EMU Speech Database Management System [28]. It included the following eight tiers: phonemic segments, syllables, morphological glossing, morphologically segmented Vera'a text, syntactic glossing, corresponding Vera'a text, English translation, Vera'a original utterance. Duration and acoustic values were queried using the emuR package in R [29, 30].

### 3.3. Analysis

Following the criteria described above, a set of 412 noun roots and 134 unique word-roots was identified. Table 1 shows an overview of the vowels, syllable structures, and number of syllables per root present in the set. Importantly, the roots selected were never placed at the onset or offset of the utterance. This was done with the aim to control that they were not at a major prosodic boundary. See example in Figure 1. The data set is not balanced containing words of different lengths and with different syllable structures. To allow a comparison across roots and control for potential confounding factors the following analyses are restricted to a set of CV (463 tokens) and another of CVC (280 tokens) syllables. Five acoustic measurements were taken for the vowels of the selected syllables: vowel duration, intensity at mid point, relative intensity, F1, and F2 at mid point. Relative intensity was calculated as the difference between intensity at vowel mid point and intensity at mid point in the last segment preceding the root. Raw values for F1, F2, and vowel duration were Lobanov normalised and these values were used in the statistical analyses.

Table 1: *Summary and counts of vowels, syllable structures, and number of syllables per root found in our corpus.*

| | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| Vowel | **ɐ** | **e** | **ɛ** | **i** | **o** | **ɔ** | **u** | |
| Count | 284 | 94 | 137 | 82 | 75 | 58 | 57 | 787 |
| Syllable structure | | | **cv** | **cvc** | **ccvc** | **v** | **vc** | |
| Count | | | 463 | 280 | 1 | 35 | 8 | 787 |
| Syllables per root | | | **1** | **2** | **3** | **4** | **5** | |
| Count | | | 126 | 201 | 55 | 24 | 3 | 412 |

#### *3.3.1. Statistical analyses*

Data were analysed using linear mixed effects models. Statistical analyses were carried out in R [30] with help of the statistics package lme4 [31]. Values were fitted into a linear mixed effects model to investigate specific factors of interest. Following [32], no random slopes are added to the models as this affects statistical power of small data sets, such as the one used in this study.

## 4. Results

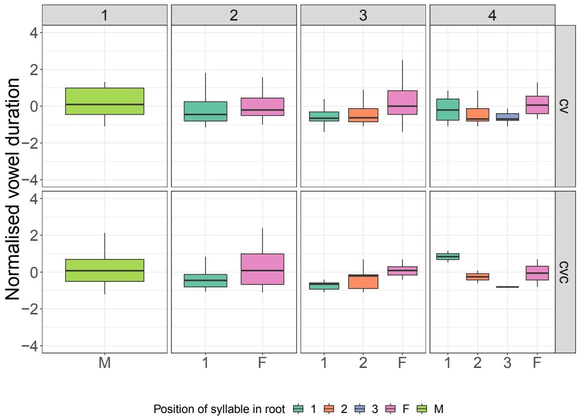


Figure 2: *Normalised vowel duration for CV- and CVC-tokens. Here, 1 stands for first syllable in the root, 2 for second, 3 for third, 4 for fourth, F for final and M for monosyllabic.*

### 4.1. Duration

Figure 2 shows the normalised vowel duration for vowels obtained in words containing one to four syllables in CV and CVC structures. A model that had Position of syllable in the root, Syllables per root, Vowel quality, Sex, and Syllable structure as fixed factors together with speaker and word as random intercepts was performed to identify whether syllable position affects vowel duration. The results show that vowels in root final syllables (Est. 16 $\pm$ 2.6 ms, $p<0.0001$)[2] and in monosyllabic roots (Est. 20 $\pm$ 6.5 ms, $p<0.0001$) are significantly longer than vowels in other positions. The factor Syllable structure was not significant ($p=0.8$), showing that vowels in CV and CVC structure showed comparable duration. Separate models for the CV and CVC sets were additionally performed, but the overall result does not change, apart from estimates being larger in the CVC set, e.g for the final syllable (Est. 25 $\pm$ 5.3 ms, $p<0.0001$).

### 4.2. Intensity

Intensity was taken at vowel mid point to examine vowels as a function of their quality and position in the root. A model with Position of syllable in the root, Syllables per root, Vowel quality, Sex, and Syllable structure as fixed factors together with speaker and word as random intercepts was used to identify which factors affect intensity in the vowel. Similar to results in 4.1, the vowels in final position (Est. 1.3 $\pm$ 0.4 dB, $p<0.0001$) and monosyllabic roots (Est. 1.7 $\pm$ 0.9 dB, $p<0.05$) show significantly greater intensity. Additionally, an effect of syllable structure was found (Est. 0.9 $\pm$ 0.4 dB, $p<0.02$) with monosyllables in CVC showing lower intensity. The differences are, however, fairly modest.

### 4.3. Relative intensity

As noted in other studies examining intensity in relation to prominence [4, 7], it is helpful to evaluate relative intensity as this can provide a more reliable measure than raw values. For this purpose we calculated the difference between the intensity at vowel mid point and that of the last segment preceding the root analysed. Figure 3 shows the relative intensity according to position of the syllable in the root and syllable structure. The median values in the graph suggest that the final syllable displays greater relative intensity. We employed a model with Position of syllable in the root, Syllables per root, Vowel quality, Sex, and Syllable structure as fixed factors together with speaker and word as random intercepts. In line with results in 4.1 and 4.2, the vowels in final position display a significantly greater relative intensity (Est. 1.3 $\pm$ 0.4 relative intensity, $p<0.004$). In this case there was no significant result for monosyllables or syllable structure.

[2]Estimates are provided in ms for a better overview.

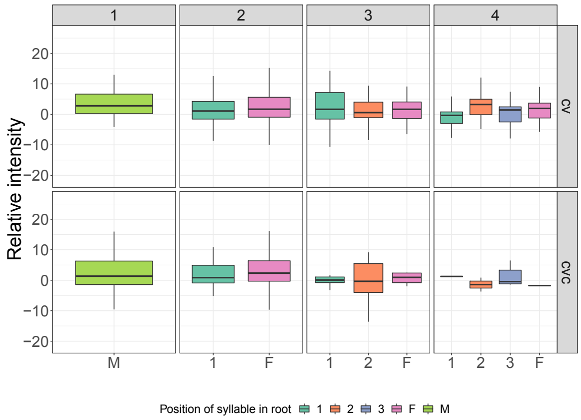


Figure 3: *Relative intensity in mono- and polysyllabic roots in CV and CVC-syllables.*

### 4.4. First and second formant

Figure 4 summarises the Lobanov normalised F1 for all Vera'a vowels in the first and second syllable of disyllabic words with CV. The median F1 of vowels /ɐ, e, o/ suggests these vowels have a raised F1 when in final position. However, this is not the case for the rest of the vowel inventory /ɔ, u ɛ, i/. Figure 5 shows the F1 trajectory of ɐ-vowels as produced by male speakers. The vowel /ɐ/ was selected because it is the vowel with the largest number of tokens (see Table 1). The trajectories indicate that the vowel appears to be realised as more peripheral in its F1 when in the final syllable or in a monosyllabic root. To investigate F1 and F2 at mid point two models with Position of syllable in the root, Syllables per root, Vowel quality, Sex, and Syllable structure as fixed factors together with speaker and word as random intercepts were used. The models included all seven vowel qualities present in our corpus. As can be expected the factor vowel quality was significant for all vowels. However, the factor of main interest to determine prominence patterns in the root, namely Position of syllable in the root did not yield any significant result for F1 or F2. Neither did we find a significant result for Syllable structure and separately run models for the CV and CVC sets do not show other results. Note however that the number of tokens per vowel varies greatly and that there might be too few vowels in all possible positions to determine whether the position in the root truly plays a role. Future investigations of Vera'a prominence patterns should include more vowel tokens.

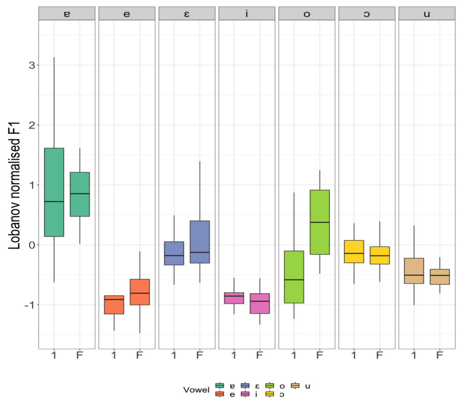


Figure 4: *Normalised F1 of disyllabic roots with CV-syllables.*

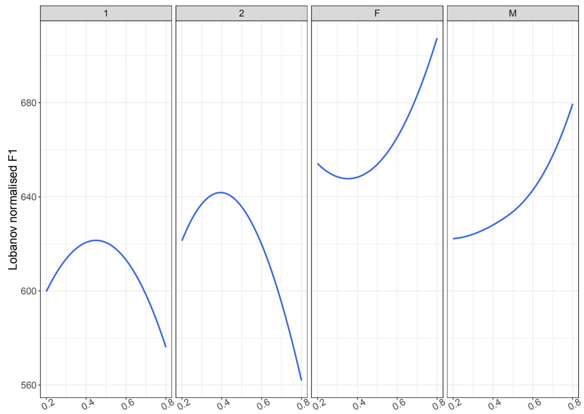


Figure 5: *Smoothed first formant trajectories based on 10 points between 0.20% and 0.80% of ɐ-vowels in CV. Only vowels that were preceded or followed by plosive and fricative consonants are included. Tokens were produced by male speakers.*

## 5. Discussion and Conclusion

Based on a speech corpus of Vera'a narratives this study examined five acoustic parameters in relation to prominence marking. To this end noun roots consisting of mono- and polysyllabic words were compared. Our analysis indicates that the final syllable (CV and CVC) of polysyllabic Vera'a roots is more salient than others through means of duration and relative intensity. More precisely we found that vowels in CV and CVC root final syllables and in monosyllables, display significantly longer duration values than other syllables in the root. Additionally, it was established that relative intensity is greater in root final syllables. Since in monosyllabic roots prominence can only be marked on this syllable it is not surprising that final syllables and monosyllables show similar results regarding duration. In addition, this study also presents results on the acoustic measures raw intensity, normalised F1, and normalised F2. It was found that intensity is greater in vowels in root final syllables. Moreover, vowels in CV-syllables display greater intensity than those in CVC. The differences found in intensity appear to be of small magnitude and although they provide additional evidence in favour of a more prominent root final syllable, they should be interpreted with caution.

Figures 4 and 5 suggest that F1 is raised in the vowels /ɐ, e, o/ when in syllables in root final position. However, the statistical analysis did not confirm this observation. There are different factors that could have influenced this result. Note that a survey on acoustic cues to stress [20] found that in languages in which stress affects vowel quality, the effect is often limited to certain vowels and/or one formant. Additionally, due to the nature of the speech present in our corpus, this study does not contain an even number of vowel tokens in every possible syllable structure (CV and CVC) or position of the syllables (first, second, third, fourth, final, monosyllabic), this means that some vowels are not present in all positions (e.g., the vowel /ɛ/ in the third syllable in CV, or vowel /i/ in the first and third position in CVC). This shortcoming is due to the lower frequency of some vowels in these positions in our corpus but could be mitigated in future research by including a larger set of roots. No statistical evidence was found for F2 as a correlate of prominence in our corpus.

This study presents the first acoustic evidence for word/root prominence being marked on the last syllable in Vera'a, confirming some observations from [19]. In other words, it was found that root final syllables and heavy root final syllables are acoustically more salient than syllables in other positions, although similarly salient compared to monosyllables. Note however that no evidence was found for penultimate stress, contra the previous impressionistic report.

One aspect that was not examined in this study is the effect of fundamental frequency (F0). This was omitted for several reasons. First, to date there is no description of Vera'a phrase level prosody making it difficult to make any well-informed predictions. Second, cross-linguistically, it is known that F0 is often associated with phrase level prosody, marking not only word level but also post-lexical prominence, as well prosodic boundaries [33, 13, 20]. Although this study sought to control for effects relative to prosodic boundary marking of major intonation breaks, it is possible that other factors related to prosodic phrasing could have an effect on the roots investigated. Moreover, there is too little evidence to date to make predictions on Vera'a being a bottom-up language like English, or a top-down language such as Chickasaw [13]. This means that we could not reliably predict whether in Vera'a pitch accents would dock on to a potentially stressed syllable, or whether stress would be orthogonal to phrasal accent assignment, displaying word-level unstressed syllables carrying tonal events like Chickasaw [22] or potentially display a yet unknown alignment of F0. For these reasons we have restricted our analysis to the five acoustic measures presented above.

In conclusion, we find evidence that root final syllables are acoustically salient in Vera'a simple roots. This finding could be indicative of regular demarcative right edge prominence. A more detailed account of Vera'a prosodic typology needs to be completed to be able to determine whether the language conforms to patterns observed in more closely related languages with post-lexical prominence marking [5, 14, 8], whether it could show similarities to languages that display misalignment of F0 in highlighting and demarcating [6, 22] or whether it displays stress-accent [21]. This study is part of a larger research project examining the acoustics of Vera'a and prominence patterns in the language. Future work including a larger set of data and extending to examine the acoustic correlates of phrase level prosodic structure is planned.

## 6. Acknowledgements

We thank the speakers who participated, as well as Sabrina Ryffel and Pelin Teberoğlu for supporting data processing.

## 7. References

[1] R. Maskikit-Essed and C. Gussenhoven, "No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay," *Phonology*, vol. 33, no. 2, pp. 353–389, 2016.

[2] R. Billington, J. Fletcher, N. Thieberger, and B. Volchok, "Acoustic correlates of prominence in Nafsan," in *Proceedings of the 17th Australasian International Speech Science and Technology Conference*, 2018, pp. 137–140.

[3] C. Torres, J. Fletcher, and G. Wigglesworth, "Investigating word prominence in Drehu," in *Proceedings of the 17th Australasian International Speech Science and Technology Conference*, 2018, pp. 141–144.

[4] C. Kaland, "Acoustic correlates of word stress in papuan malay," *Journal of Phonetics*, vol. 74, pp. 55–74, 2019.

[5] J. Fletcher, R. Billington, and N. Thieberger, "Prosodic marking of focus in Nafsan," in *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne*, 2019, pp. 1758–1764.

[6] C. Kaland and S. Baumann, "Demarcating and highlighting in Papuan Malay phrase prosody," *The Journal of the Acoustical Society of America*, vol. 147, no. 4, pp. 2974–2988, 2020.

[7] R. Billington, J. Fletcher, N. Thieberger, and B. Volchok, "Acoustic evidence for right-edge prominence in Nafsan," *The Journal of the Acoustical Society of America*, vol. 147, no. 4, pp. 2829–2844, 2020.

[8] C. Torres and J. Fletcher, "Phrase-level and edge marking in drehu," *Glossa: a journal of general linguistics*, p. 1–32, 2022.

[9] N. Himmelmann and D. Kaufman, "Austronesia," in *The Oxford Handbook of Language Prosody*, 2020.

[10] S.-A. Jun, "Prosodic typology: By prominence type, word prosody, and macro-rhythm," in *Prosodic typology II: The phonology of intonation and phrasing*, jun, sun-ah ed. Oxford University Press, 2014, pp. 520–539.

[11] E. van Zanten and V. J. van Heuven, "Word stress in Indonesian: Its communicative relevance," *Bijdragen tot de Taal-, Land-en Volkenkunde*, vol. 154, pp. 129–149, 1998.

[12] R. Goedemans and E. van Zanten, "Stress and accent in Indonesian," in *LOT Occasional series*. LOT, Netherlands Graduate School of Linguistics, 2007, vol. 9, pp. 35–62.

[13] T. Roettger and M. Gordon, "Methodological issues in the study of word stress correlates," *Linguistics Vanguard*, vol. 3, no. 1, 2017.

[14] C. Torres and J. Fletcher, "The alignment of F0 tonal targets under changes in speech rate in Drehu," *The Journal of the Acoustical Society of America*, vol. 147, no. 4, pp. 2947–2958, 2020.

[15] D. T. Tryon, *Dehu grammar*. Australian National University, 1968.

[16] C. Moyse-Faurie, *Le drehu, langue de Lifou (Iles Loyauté). Phonologie, morphologie, syntaxe*. Langues et Cultures du Pacifique Ivry, 1983.

[17] R. Billington, N. Thieberger, and J. Fletcher, "Nafsan," *Journal of the International Phonetic Association*, pp. 1–21, 2021.

[18] M. D. Ross, "The morphosyntactic typology of Oceanic languages," *Language and Linguistics*, vol. 5, no. 2, pp. 491–541, 2004.

[19] S. Schnell, "A grammar of Vera'a," Ph.D. dissertation, Kiel University, 2011.

[20] M. Gordon and T. Roettger, "Acoustic correlates of word stress: A cross-linguistic survey," *Linguistics Vanguard*, vol. 3, no. 1, 2017.

[21] L. M. Hyman, "Do all languages have word accent," in *Word stress: Theoretical and typological issues*, H. van der Hulst, Ed. Cambridge University Press Cambridge, 2014, pp. 56–82.

[22] M. Gordon, "The phonology of pitch accents in Chickasaw," *Phonology*, vol. 20, no. 2, pp. 173–218, 2003.

[23] S. Schnell, "Vera'a doreco dataset," in *Language Documentation Reference Corpus (DoReCo) 1.1*, F. Seifart, L. Paschen, and M. Stave, Eds. Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), 2022. [Online]. Available: https://doreco.huma-num.fr/languages/vera1241

[24] F. Seifart, L. Paschen, and M. Stave, Eds., *Language Documentation Reference Corpus (DoReCo)*. Leibniz-Zentrum Allgemeine Sprachwissenschaft and Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), 2022. [Online]. Available: https://doreco.huma-num.fr/

[25] S. Schnell, "Multi-CAST Vera'a," in *Multi-CAST*, G. Haig and S. Schnell, Eds., 2015, version 2207.

[26] G. Haig and S. Schnell, Eds., *Multi-CAST*, 2015. [Online]. Available: https://multicast.aspra.uni-bamberg.de/

[27] L. Paschen, S. Fuchs, and F. Seifart, "Final lengthening and vowel length in 25 languages," *Journal of Phonetics*, vol. 94, p. 101179, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0095447022000547

[28] R. Winkelmann, J. Harrington, and K. Jänsch, "Emu-sdms: Advanced speech database management and analysis in r," *Computer Speech & Language*, vol. 45, pp. 392–410, 2017.

[29] R. Winkelmann, K. Jaensch, S. Cassidy, and J. Harrington, *emuR: Main Package of the EMU Speech Database Management System*, 2017, r package version 0.2.3.

[30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: https://www.R-project.org/

[31] D. Bates, M. Mächler, and S. Walker, *Linear Mixed-Effects Models using 'Eigen' and S4*, 2019.

[32] H. Matuschek, R. Kliegl, S. Vasishth, H. Baayen, and D. Bates, "Balancing Type I error and power in linear mixed models," *Journal of memory and language*, vol. 94, pp. 305–315, 2017.

[33] D. R. Ladd, *Intonational phonology*. Cambridge University Press, 2008.